Дискретная математика

и комбинаторика



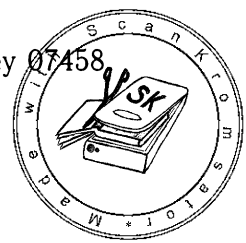
Джеймс Андерсон

DISCRETE MATHEMATICS WITH COMBINATORICS

JAMES A. ANDERSON
University of South Carolina, Spartanburg



Prentice Hall, Upper Saddle River, New Jersey 07458



ДИСКРЕТНАЯ МАТЕМАТИКА И КОМБИНАТОРИКА

ДЖЕЙМС А. АНДЕРСОН

Университет Южной Каролины, Спартанбург

Москва • Санкт-Петербург • Киев
2004

Made with Scankromsator*

ББК 32.973.26-018.2.75
А65
УДК 681.3.07

Издательский дом "Вильямс"

Зав. редакцией *С.Н. Тригуб*

Перевод с английского канд. физ.-мат. наук *М.М. Беловой*

Под редакцией канд. физ.-мат. наук *С.С. Шкильняка* и *М.Р. Саит-Аметова*

Научный консультант докт. физ.-мат. наук, проф. *Ю.В. Казаченко*

По общим вопросам обращайтесь в Издательский дом "Вильямс" по адресу:
info@williamspublishing.com, http://www.williamspublishing.com

Андерсон, Джеймс А.

А65 Дискретная математика и комбинаторика. : Пер. с англ. — М. : Издательский дом "Вильямс", 2004. — 960 с. : ил. — Парал. тит. англ.

ISBN 5-8459-0498-6 (рус.)

Эта книга представляет собой современный учебник по дискретной математике. Кроме таких разделов, как математическая логика, теория множеств, комбинаторика, теория графов, теория алгоритмов и вычислений, традиционно включаемых в основной курс дискретной математики, она содержит обширные сведения по теории вероятностей, алгебре и теории чисел. Особое внимание уделено теории доказательств. Чтение книги требует некоторой математической культуры, хотя для изучения основных глав достаточно знаний по математике в объеме средней школы. Материал сопровождается многочисленными примерами, в конце каждого раздела приводится большое количество упражнений.

Книга адресована в первую очередь преподавателям и студентам технических специальностей. Она будет также полезна тем, кто интересуется дискретной математикой и желает изучить ее самостоятельно.

ББК 32.973.26–018.2.75

Все названия программных продуктов являются зарегистрированными торговыми марками соответствующих фирм.

Никакая часть настоящего издания ни в каких целях не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитный носитель, если на это нет письменного разрешения издательства Prentice Hall, Inc.

Authorized translation from the English language edition published by Prentice-Hall, Inc, Copyright © 2001

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from the Publisher.

Russian language edition published by Williams Publishing House according to the Agreement with R&I Enterprises International, Copyright © 2003

ISBN 5-8459-0498-6 (рус.)
ISBN 0-13-086998-8 (англ.)

© Издательский дом "Вильямс", 2003
© Prentice-Hall, Inc, 2001

Содержание

Предисловие 10

1 Таблицы истинности, логика, доказательства 15

1.1. Высказывания и логические связки 15
1.2. Условные высказывания 23
1.3. Эквивалентные высказывания 27
1.4. Аксиоматические системы: умозаключения и доказательства 34
1.5. Полнота в логике высказываний 45
1.6. Карты Карно 50
1.7. Коммутационные схемы 56

2 Теория множеств 67

2.1. Понятие множества 67
2.2. Операции над множествами 71
2.3. Диаграммы Венна 77
2.4. Булевы алгебры 84
2.5. Отношения 90
2.6. Частично упорядоченные множества 103
2.7. Отношения эквивалентности 106

3 Логика, целые числа и доказательства 113

3.1. Исчисление предикатов 113
3.2. Основные положения теории доказательств и теории целых чисел 123
3.3. Математическая индукция 129
3.4. Делимость 139
3.5. Простые числа 144
3.6. Сравнения 149

4 Функции и матрицы 156

4.1. Функции 156
4.2. Специальные функции 161
4.3. Матрицы 167
4.4. Мощность 177
4.5. Мощность (продолжение) 178

5 Алгоритмы и рекурсия 184

5.1. Циклы и алгоритмы для матриц 184
5.2. Рекурсивные функции и алгоритмы 188
5.3. Сложность алгоритмов 201
5.4. Алгоритмы сортировки 205
5.5. Префиксная и суффиксная записи 214
5.6. Двоичные и шестнадцатеричные числа 219
5.7. Числа со знаком 231
5.8. Дальнейшее изучение матриц 237

6 Графы, ориентированные графы и деревья 244

6.1. Графы 244
6.2. Ориентированные графы 252
6.3. Деревья 259
6.4. Мгновенное безумие 267
6.5. Пути и циклы Эйлера 270
6.6. Матрицы инцидентности и смежности 278
6.7. Гиперкубы и код Грея 290

7 Теория чисел 298

7.1. Решето Эратосфена 298
7.2. Метод выделения множителей Ферма 300
7.3. Алгоритмы деления и алгоритм Евклида 301
7.4. Цепные дроби 306
7.5. Подходящие дроби 310

8 Комбинаторика и вероятность 316

8.1. Основные комбинаторные принципы 316
8.2. Комбинаторный принцип сложения 324
8.3. Перестановки и сочетания 331
8.4. Формирование перестановок и сочетаний 343
8.5. Введение вероятности 347
8.6. Обобщенные перестановки и сочетания 354
8.7. Перестановки и сочетания с повторением 359
8.8. Принцип клеток 363
8.9. Снова о вероятности 369
8.10. Теорема Байеса 384
8.11. Цепи Маркова 386

9 Алгебраические структуры 392

9.1. Вновь о частично упорядоченных множествах 392
9.2. Полугруппы и полурешетки 397
9.3. Решетки 403
9.4. Группы 409
9.5. Группы и гомоморфизмы 415

10 Некоторые специальные вопросы теории чисел 422

10.1. Целочисленные решения линейных уравнений 422
10.2. Решения сравнений 424
10.3. Китайская теорема об остатках 428
10.4. Свойства функции ϕ 433
10.5. Порядок целого числа 439

11 Некоторые специальные вопросы теории рекурсии 448

11.1. Однородные линейные рекуррентные отношения 448
11.2. Неоднородные линейные рекуррентные отношения 460
11.3. Конечные разности 469
11.4. Факториальные многочлены 473
11.5. Суммирование разностей 483

12 Снова о комбинаторных подсчетах 489

12.1. Задачи о размещении 489
12.2. Числа Каталана 495
12.3. Общее включение-исключение и разупорядочения 502
12.4. Ладейные полиномы и запрещенные позиции 509

13 Производящие функции 523

13.1. Определение производящей функции 523
13.2. Производящие функции и рекуррентные отношения 525
13.3. Производящие функции и комбинаторные подсчеты 535
13.4. Разбиения 542
13.5. Экспоненциальные производящие функции 549

14 Некоторые специальные вопросы теории графов 556

14.1. Алгебраические свойства графов 556
14.2. Планарные графы 580
14.3. Раскраска графов 586
14.4. Пути и циклы Гамильтона 600
14.5. Взвешенные графы и алгоритмы поиска кратчайшего пути 611

15 Деревья 624

15.1. Свойства деревьев 624
15.2. Бинарные деревья поиска 631
15.3. Взвешенные деревья 638
15.4. Обход бинарных деревьев 649
15.5. Остовные деревья 658
15.6. Минимальные остовные деревья 682

16 Сети 691

16.1. Сети и потоки 691
16.2. Паросочетание 707
16.3. Сети Петри 716

17 **Теория вычислений** 725

17.1. Регулярные языки 725
17.2. Автоматы 731
17.3. Грамматики 741

18 **Теория кодов** 753

18.1. Введение 753
18.2. Порождающие матрицы 757
18.3. Коды Хемминга 767

19 **Перечисление цветов** 775

19.1. Теорема Бернсайда 775
19.2. Теорема Пойа 781

20 **Кольца, области целостности и поля** 788

20.1. Кольца и области целостности 788
20.2. Области целостности 797
20.3. Полиномы 801
20.4. Алгебры и полиномы 808

21 **Характеры групп и полугрупп** 819

21.1. Комплексные числа 819
21.2. Характеры групп 820
21.3. Характеры полугрупп 825

22 **Приложения теории чисел** 829

22.1. Приложение: поиск по образу 829
22.2. Приложение: функции хеширования 837
22.3. Приложение: криптография 843

Литература 850

Ответы к упражнениям 856

Предметно-именной указатель 942

Список обозначений 954

Памяти Эльвуда В. Стоуна и Орвилла Г. Уиби.
Моей семье: Мэрилин, Энди, Кристин и Филу.
Учителям, наставникам и друзьям: Т.Хэду, Н.Кимура, Л.С.Лоузу и Э.Л.Дубовски.

Предисловие

Несмотря на то, что на рынке учебной литературы на сегодняшний день имеется немало книг по дискретной математике, пожалуй, ни одна из них не может сравниться с этой книгой по размаху и глубине рассматриваемых тем. Отличительной особенностью данной книги является то внимание, которое уделяется в ней теории доказательств. Понятие доказательства вводится уже в первой главе и развивается на протяжении всей книги. Дискретную математику изучают, как правило, на первых курсах математических факультетов и факультетов информатики. Необходимость изучения техники доказательств для будущих математиков очевидна. Она также крайне важна для развития логического мышления у будущих специалистов в области информатики. По сути, правильно составленная компьютерная программа эквивалентна логическому доказательству. По мнению автора, использовать то или иное приложение (или, в крайнем случае, не ошибаться при его использовании) легче, если понимаешь, как оно работает. За редкими исключениями, материал книги является самодостаточным. Использованию любого понятия в тексте предшествует его строгое математическое определение.

Материал данной книги не требует от читателя знания математического анализа. Для понимания основных глав вполне достаточно знакомства со школьным курсом алгебры. Но, несмотря на самодостаточность книги в целом, отдельные ее главы все же требуют от читателя определенной математической подготовки.

Книга планировалась как основа семестрового или двухсеместрового курса по дискретной математике. В первых восьми главах закладываются основы предмета; их можно использовать при чтении начального курса. Эти главы практически независимы друг от друга, так что преподаватель сам может выбрать материал для изложения. Содержание остальных глав соответствует второй части курса по дискретной математике. В них введенные ранее понятия изучаются более детально, и рассматриваются темы повышенного уровня сложности. Темы рассматриваются с различных точек зрения, чтобы показать, как их можно использовать в зависимости от ситуации. В спектр рассматриваемых тем входят:

Логика, включающая таблицы истинности, логику высказываний, коммутационные схемы, исчисление предикатов, индукцию и доказательства;

Теория множеств, включающая понятие мощности множеств, отношения, частично упорядоченные множества, отношения конгруэнтности, графы, ориентированные графы и функции;

Алгоритмы, включающая понятие сложности алгоритмов, алгоритмы поиска и сортировки, а также алгоритмы Евклида, Хаффмана, Прима, Уоршолла, Форда-Фулкерсона, Флойда-Уоршолла и Дейкстры;

Теория графов, включающая ориентированные графы, циклы и пути Эйлера, циклы и пути Гамильтона, плоские и взвешенные графы;

Деревья, включающая бинарные деревья поиска, взвешенные деревья, обход деревьев, коды Хаффмана и остовные деревья;

Комбинаторика, включающая перестановки, сочетания, включение-исключение, разбиения, производящие функции, числа Каталана, числа Стирлинга, ладейные многочлены, разупорядочения и перечисления цветов;

Алгебра, включающая полугруппы, группы, решетки, полурешетки, булевы алгебры, кольца, поля, области целостности, полиномы и матрицы.

Книга содержит обширные сведения по алгебре и теории чисел, что, по мнению автора, усиливает ее содержание, хотя включать эти темы в курс лекций вовсе не обязательно. Материал соответствующих глав полностью независим от остальных частей книги, и в каком объеме его стоит излагать — решать лектору. Данное пособие содержит также элементы теории вероятностей, сведения о конечных разностях и другие темы, обычно не рассматриваемые в книгах по дискретной математике.

СТРУКТУРА КНИГИ

Содержание первых трех глав охватывает логику и теорию множеств. Здесь позиция автора состоит в том, что понимание основ теории доказательств необходимо для логического построения передовых компьютерных технологий. Изложены основные положения теории доказательств и проиллюстрированы многочисленными примерами. В главе 2 студенту предлагается самостоятельно доказать некоторые элементарные утверждения теории множеств. В главе 3 вводится понятие аксиоматической системы для теории чисел. Студенту предоставляется возможность попрактиковаться в доказательстве теорем о хорошо знакомых ему объектах. В этой же главе вводится метод доказательства по индукции. Далее в книге приводится множество доказательств, а также посвященных им задач. Сначала эти задачи довольно просты, но по ходу изложения книги уровень их сложности постепенно возрастает.

Отношения и графы вводятся в главе 2. В главе 4 из понятия отношения естественным образом выводится понятие функции. Однако подробное изучение функций в главе 4 проводится независимо от материала главы 2. Точно так же в главе 6 графы исследуются вне контекста главы 2, где они рассматриваются в связи с отношениями.

Матрицы, перестановки и последовательности вводятся в главе 4 как функции специального вида. Дальнейшее изучение свойств этих функций продолжается в главе 6. Здесь же вводятся алгоритмы операций над матрицами и рассматриваются те свойства матриц, которые используются далее в главах, посвященных алгебре, комбинаторике и теории кодов.

В главе 8 перестановки используются как инструмент комбинаторных подсчетов. Перестановки также используются в последующих главах, посвященных прикладным вопросам алгебры и комбинаторики. Глава 8, хотя она и связана с главой 4, может изучаться независимо.

Содержание главы 5 не связано с предыдущими главами ничем, за исключением того, что касается матриц. Подробно изучены алгоритмы, в частности, алгоритмы сортировки. Понятие сложности алгоритмов также рассмотрено в этой главе. Здесь же даны сведения о префиксной и суффиксной записях. Эта тема опять обсуждается в главе 15 в связи с обходом бинарных деревьев. В главе 5 вводятся также двоичные и шестнадцатеричные числа.

Многие элементарные понятия, связанные с графами, ориентированными графами и деревьями, рассматриваются в главе 6. Более глубоко эти вопросы изучаются в главах 14–16. Содержание главы 6 не зависит от содержания предыдущих глав.

В главах 7 и 10 получает дальнейшее развитие теория чисел. Материал этих глав используется в главе 22 при изучении приложений теории чисел. В остальном они вполне самостоятельны и при желании могут быть опущены.

С главы 8 начинается изучение широкого круга вопросов комбинаторики, которое затем продолжается во многих других частях книги, включая главы 12, 13 и 17. В той же главе 8 приводятся некоторые сведения из теории вероятностей, что само по себе весьма необычно для книг по дискретной математике.

Главы 9 и 20 охватывают основные понятия алгебры, включающие полугруппы, группы, кольца, решетки, полурешетки, области целостности и поля. При построении примеров групп и колец в этих главах используется материал разделов 3.6 и 4.3. Материал главы 9 используется далее в приложениях, рассматриваемых в главах 17–21.

По многим причинам главы 11, 12 и 13 можно выделить в отдельную группу. В главе 11 продолжается изучение рекурсии. Кроме обычных линейных рекуррентных отношений, включаемых, как правило, в учебные курсы по дискретной математике, глава содержит сведения по теории конечных разностей. Если читатель не имеет хотя бы поверхностного представления о рекурсии, то для понимания этой главы ему необходимо ознакомиться с материалом главы 6. В главе 12 продолжается изучение комбинаторики, начатое в главе 8. Здесь рассматриваются такие вопросы, как включение-исключение и задача о размещении. Кроме того, вводится понятие разупорядочения и ладейного полинома. Главы 11 и 12 тесно взаимосвязаны; в них многие темы рассматриваются с различных точек зрения. Одним из объектов такого разностороннего рассмотрения являются числа Стирлинга. Тем не менее, каждая из этих глав вполне самостоятельна.

В главе 13 вводится понятие производящей функции, и на его основе продолжается изучение материала глав 11 и 12. В частности, производящие функции используются для построения эффективного метода решения задач о размещении.

В главах 14–16 продолжается изучение графов и деревьев, начатое в главе 6. Содержание этих глав очевидным образом зависит от материала главы 6, но с материалом большинства предыдущих глав практически не связано. Исключение составляет использование матриц в одном из рассматриваемых алгоритмов. Здесь изучаются стандартные для теории графов и деревьев объекты: плоские графы, циклы Гамильтона, бинарные деревья, остовные деревья, минимальные остовные деревья, алгоритмы нахождения кратчайшего пути и потоки в сетях.

Еще одну группу составляют главы 17–22, посвященные прикладным вопросам теории чисел, алгебры и комбинаторики. Глава 17 посвящена теории вычис-

лений; в ней рассматриваются коды, регулярные языки, автоматы, грамматики и связь между ними. Здесь используются полугруппы, введенные в разделе 9.2. В главе 18 вводятся специальные коды, такие как коды с обнаружением ошибок и коды с исправлением ошибок. Для понимания материала этой главы требуется знание теории групп в объеме раздела 9.4 и теории матриц в объеме глав 4 и 5. Совершенно иное применение теория кодов получает в главе 22, где рассматриваются вопросы криптографии. Эта глава требует знаний по теории чисел на уровне, выходящем за пределы данной книги.

В главе 19 при изложении теорем Бернсайда и Пойя для перечисления цветов используются как алгебра, так и комбинаторика. Здесь, главным образом, требуется знание перестановок в объеме раздела 9.4.

Глава 21 посвящена простейшим приложениям групп и полугрупп, а также их отображениям на комплексную плоскость. Необходимые предварительные сведения содержатся в разделах 9.2 и 9.5.

Глава 22 содержит три важных приложения теории чисел. Изучение функций хеширования и криптографии является для информатики весьма актуальным.

При чтении вводного курса автор, как правило, полностью излагает материал глав 1–5, разделов 8.1–8.3, а также старается изложить первые три раздела главы 6. Как отмечалось ранее, материал первых восьми глав составлен так, чтобы обеспечить максимальную гибкость. В следующей таблице указано, какие предварительные сведения требуются для усвоения материала каждой из глав:

Глава	*Необходимые главы или разделы*
Глава 1	Никакие
Глава 2	Никакие
Глава 3	Разделы 1.1–1.4 и 2.1
Глава 4	Никакие
Глава 5	Разделы 4.1–4.3
Глава 6	Никакие
Глава 7	Глава 3
Глава 8	Никакие
Глава 9	Раздел 3.6
Глава 10	Глава 7
Глава 11	Разделы 5.1–5.3
Глава 12	Глава 8
Глава 13	Главы 11 и 12
Глава 14	Главы 5 и 6
Глава 15	Главы 5 и 6
Глава 17	Глава 9
Глава 18	Главы 5 и 9
Глава 19	Глава 9
Глава 20	Глава 9
Глава 21	Глава 9
Глава 22	Глава 10

ПОДДЕРЖКА

Имеется подробное руководство к решению всех задач, рассмотренных в этой книге, которое можно заказать в издательстве Prentice Hall. Книга имеет свою страницу в Интернете; ее адрес: *www.prenhall.com/janderson*. Здесь вы найдете ссылки на другие интересные сайты, посвященные дискретной математике, в том числе содержащие контрольные задания и формулировки интересных проблем дискретной математики, не вошедшие в данное издание.

БЛАГОДАРНОСТИ

В первую очередь мне хочется поблагодарить Джорджа Лобелла, руководившего разработкой этой книги, и Барбару Мэк, координировавшую наши усилия. Я благодарю Кристиана и Филиппа Мусик за мастерски выполненное художественное оформление. Особую благодарность я приношу Джеймсу Беллу, внесшему огромный вклад в написание данной книги. Очень жаль, что он не смог выступить в роли моего соавтора. Мне очень не хватает его партнерской поддержки. Кроме того, я хочу поблагодарить за помощь моих коллег Дэна Кука, Эда Вайлда, Рика Чоу, М. Б. Ульмера и Джерома Льюиса. Я благодарен своей студентке Соледад Шугаи за усилия, приложенные ею при изучении моего курса. Также хочу поблагодарить студентов Джоди Дина, Джессику Донс, Грейса Эллисона, Винни Чин Фай Ип, Присциллу Лапьер, Эстер Лай, Бадрала Мадани, Джулию Норрис, Трейси Квина и Роберта Вигерта, бывших первыми, кто прослушал данный материал.

Пожалуйста, присылайте по электронной почте свои замечания и предложения по дальнейшему улучшению книги.

Джеймс А. Андерсон
janderson@gw.uscs.edu

ГЛАВА

1

ТАБЛИЦЫ ИСТИННОСТИ, ЛОГИКА, ДОКАЗАТЕЛЬСТВА

1.1. ВЫСКАЗЫВАНИЯ И ЛОГИЧЕСКИЕ СВЯЗКИ

В этом разделе рассматриваются таблицы истинности, знакомство с которыми будет для нас первым шагом в изучении логики. Далее мы увидим, что таблицы истинности являются также основным инструментом для определения других важных понятий дискретной математики. Логика, созданная как наука знаменитым Аристотелем (384–322 до н.э.), на протяжении столетий использовалась для развития многих областей знания, включая теологию, философию, математику. Она — тот фундамент, на котором построено все здание математики. По сути, логика — это наука о рассуждениях, которая позволяет определить истинность или ложность того или иного математического утверждения, исходя из совокупности первичных предположений, называемых аксиомами. Логика применяется также в информатике для построения компьютерных программ и доказательства их корректности. Понятия, методы и средства логики лежат в основе современных информационных технологий. Одна из основных целей этой книги — изложить основы математической логики, показать, как она используется в информатике, и разработать методы анализа и доказательства математических утверждений.

Высказывание — это утверждение или повествовательное предложение, о котором можно сказать, что оно истинно или ложно. Иными словами, утверждение об истинности или ложности высказывания должно иметь смысл. Истинность или ложность, приписываемые некоторому утверждению, называются его ***значением истинности***, или ***истинностным значением***.

Вот примеры предложений, не являющихся высказываниями:

Кто вы? (вопрос),
Прочтите эту главу до следующего занятия (приказ или восклицание),
Это утверждение ложно (внутренне противоречивое утверждение).

Мы будем обозначать высказывания буквами латинского алфавита p, q, r, ... Например, p может обозначать утверждение *Завтра будет дождь*, а q — утверждение *Квадрат целого числа есть число положительное.*

В обыденной речи для образования сложного предложения из простых используются связки — особые части речи, соединяющие отдельные предложения. Наиболее часто употребляются связки *и, или, нет, если ... то, только если*, и *тогда и только тогда*. В отличие от обыденной речи, в логике смысл таких связок должен быть определен однозначно. Истинность сложного высказывания однозначно определяется истинностью или ложностью составляющих его частей. Высказывание, не содержащее связок, называется ***простым***. Высказывание, содержащее связки, называется ***сложным***.

Пусть p и q обозначают высказывания

p: *Джейн водит автомобиль,*
q: *У Боба русые волосы.*

Сложное высказывание

Джейн водит автомобиль и у Боба русые волосы

состоит из двух частей, объединенных связкой *и*. Это высказывание может быть символически записано в виде

$$p \text{ и } q$$

или просто как

$$p \wedge q,$$

где символ $\wedge$ обозначает слово *и* на языке символических выражений. Выражение $p \wedge q$ называется ***конъюнкцией*** высказываний p и q.

Точно так же высказывание

Джейн водит автомобиль или у Боба рыжие волосы.

символически выражается как

$$p \text{ или } q$$

или

$$p \vee q,$$

где $\vee$ обозначает слово *или* в переводе на символический язык. Выражение $p \vee q$ называется ***дизъюнкцией*** высказываний p и q.

Опровержение, или ***отрицание*** высказывания p обозначается через

$$\sim p.$$

Таким образом, если p есть высказывание *Джейн водит автомобиль*, то $\sim p$ — это утверждение *Джейн не водит автомобиль.*

Если r есть высказывание *Джо нравится информатика*, то *Джейн не водит автомобиль и у Боба русые волосы или Джо любит информатику* символически запишется как $((\sim p) \wedge q) \vee r$. И наоборот, выражение $p \wedge (\sim q) \wedge r$ — это

символическая форма записи высказывания *Джейн водит автомобиль, у Боба волосы не русые и Джо нравится информатика.*

Рассмотрим выражение $p \wedge q$. Если некто говорит: "Джейн водит автомобиль и у Боба русые волосы", то мы, естественно, представляем себе Джейн за рулем автомобиля и русоволосого Боба. В любой другой ситуации (например, если Боб не русоволос или Джейн не водит автомобиль) мы скажем, что говорящий не прав.

Возможны четыре случая, которые нам необходимо рассмотреть. Высказывание p может быть истинным (T) или ложным (F) и независимо от того, какое истинностное значение принимает p, высказывание q может также быть истинным (T) или ложным (F). ***Таблица истинности*** перечисляет все возможные комбинации истинности и ложности сложных высказываний.

Случай	p	q	$p \wedge q$
1	T	T	T
2	T	F	F
3	F	T	F
4	F	F	F

Ранее мы убедились, что только в первом случае высказывание $p \wedge q$ истинно. В остальных же мы имеем ложное значение $p \wedge q$. Например, случай 3 описывает значение истинности для $p \wedge q$, когда неверно, что Джейн водит автомобиль и у Боба русые волосы. Если p — высказывание *Джон богат*, а q — высказывание *Джон красив*, то не знакомая с Джоном девушка, которую убедили в том, что высказывание *Джон богат и Джон красив*, или *Джон богат и красив* истинно, будет представлять себе Джона и богатым, и красивым.

Точно так же рассмотрим высказывание

Джейн водит автомобиль или у Боба русые волосы,

которое символически выражается как $p \vee q$. Если некто скажет: "Джейн водит автомобиль или у Боба русые волосы", то он будет не прав только тогда, когда Джейн не сможет управлять автомобилем, а Боб не будет русоволосым. Для того чтобы все высказывание было истинным, достаточно, чтобы одна из двух составляющих его компонент была истинной. Поэтому $p \vee q$ имеет таблицу истинности

Случай	p	q	$p \vee q$
1	T	T	T
2	T	F	T
3	F	T	T
4	F	F	F

Высказывание $p \vee q$ ложно только в случае 4, когда p и q оба ложны.

Если p — высказывание *Джон богат*, а q — высказывание *Джон красив*, и не знакомая с Джоном девушка уверена в истинности высказывания *Джон богат или Джон красив*, или *Джон богат или красив*, то она вправе ожидать, что

истинно одно из высказываний, но не обязательно оба. Девушка почувствует себя введенной в заблуждение, только если обнаружит, что Джон беден и уродлив.

Таблица истинности для отрицания p имеет вид

Случай	p	$\sim p$
1	T	F
2	F	T

Истинностное значение $\sim p$ всегда противоположно истинностному значению p. В таблицах истинности отрицание всегда оценивается первым, если только за знаком отрицания не следует высказывание, заключенное в скобки. Поэтому $\sim p \vee q$ интерпретируется как $(\sim p) \vee q$, так что отрицание применяется только к p. Если мы хотим отрицать все высказывание $p \wedge q$, то это записывается как $\sim(p \vee q)$.

Символы $\wedge$ и $\vee$ называют ***бинарными*** связками, так как они связывают два высказывания как, например, в выражениях $p \wedge q$ и $p \vee q$. Символ $\sim$ является ***унарной*** связкой, так как применяется только к одному высказыванию.

Еще одна бинарная связка — это ***исключающее или***, которое обозначается через $\veebar$. Высказывание $p \veebar q$ истинно, когда истинно p или q, но не оба одновременно. Эта связка имеет таблицу истинности

Случай	p	q	$p \veebar q$
1	T	T	F
2	T	F	T
3	F	T	T
4	F	F	F

Используя слово *или*, мы можем иметь в виду *исключающее или*. Например, когда мы говорим: "Дик сдаст экзамен по логике или он не сдаст этот экзамен", мы, конечно, предполагаем, что Дик сделает что-то одно. Таким образом, когда мы говорим, что p — либо истина, либо ложь, то, естественно, предполагаем, что это не выполняется одновременно. В логике *исключающее или* используется довольно редко, и в дальнейшем мы, как правило, будем обходиться без него.

Рассмотрим высказывание

Сэм уплатит налог за машину или Сэм утратит свою машину и будет ходить на работу пешком.

Пусть p обозначает высказывание *Сэм уплатит налог за машину*, q — высказывание *Сэм останется при своей машине*, а r — высказывание *Сэм будет ходить на работу пешком*. Тогда наше сложное высказывание можно представить в виде

$$p \vee ((\sim q) \wedge r),$$

где скобки использованы, чтобы показать, какие именно высказывания являются компонентами каждой связки.

Таблица истинности дает возможность однозначно указать те ситуации, когда высказывание $p \vee ((\sim q) \wedge r)$ является истинным; при этом мы должны быть

уверены, что учтены все случаи. Поскольку сложное высказывание содержит три основных высказывания p, q и r, то возможны восемь случаев

Случай	p	q	r	$\sim q$	$(\sim q) \wedge r$	$p \vee ((\sim q) \wedge r)$
1	T	T	T	F	F	T
2	T	T	F	F	F	T
3	T	F	T	T	T	T
4	T	F	F	T	F	T
5	F	T	T	F	F	F
6	F	T	F	F	F	F
7	F	F	T	T	T	T
8	F	F	F	T	F	F

При нахождении значений истинности для столбца $(\sim q) \wedge r$ мы используем столбцы для $(\sim q)$ и r, а также таблицу истинности для $\wedge$. Таблица истинности для $\wedge$ показывает, что высказывание $(\sim q) \wedge r$ истинно лишь в том случае, когда истинны оба высказывания $(\sim q)$ и r. Это имеет место лишь в случаях 3 и 7.

Заметим, что при определении значений истинности для столбца $p \vee ((\sim q) \wedge r)$ играет роль только истинность высказываний p и $(\sim q) \wedge r$. Таблица истинности для $\vee$ показывает, что единственный случай, когда высказывание, образованное с помощью связки *или*, ложно, — это случай, когда ложны обе части этого высказывания. Такая ситуация имеет место только в случаях 5, 6 и 8.

Если Сэм не уплатит налог за машину (т.е. p ложно, или имеет значение F), лишится своей машины (q имеет значение F) и будет ходить на работу пешком (r имеет значение T), то будет иметь место случай 7. Тот, кто скажет: "Сэм уплатит налог за машину или Сэм утратит машину и будет ходить на работу пешком", будет абсолютно прав.

Другой, эквивалентный способ построения таблицы истинности состоит в том, чтобы записывать истинностные значения выражения под связкой. Снова рассмотрим выражение $p \vee ((\sim q) \wedge r)$. Сначала мы записываем истинностные значения под переменными p, q и r. Единицы под столбцами истинностных значений указывают на то, что этим столбцам истинностные значения присваиваются в первую очередь. В общем случае число под столбцом будет показывать номер шага, на котором производятся вычисления соответствующих истинностных значений.

Случай	p	q	r	p	$\vee$	$((\sim$	$q)$	$\wedge$	$r)$
1	T	T	T	***T***			***T***		***T***
2	T	T	F	***T***			***T***		***F***
3	T	F	T	***T***			***F***		***T***
4	T	F	F	***T***			***F***		***F***
5	F	T	T	***F***			***T***		***T***
6	F	T	F	***F***			***T***		***F***
7	F	F	T	***F***			***F***		***T***
8	F	F	F	***F***			***F***		***F***
				1			1		1

Затем мы записываем под символом $\sim$ истинностные значения высказывания $\sim q$.

Случай	p	q	r	p	$\vee$	$((\sim$	$q)$	$\wedge$	$r)$
1	T	T	T	T		$\boldsymbol{F}$	T		T
2	T	T	F	T		$\boldsymbol{F}$	T		F
3	T	F	T	T		$\boldsymbol{T}$	F		T
4	T	F	F	T		$\boldsymbol{T}$	F		F
5	F	T	T	F		$\boldsymbol{F}$	T		T
6	F	T	F	F		$\boldsymbol{F}$	T		F
7	F	F	T	F		$\boldsymbol{T}$	F		T
8	F	F	F	F		$\boldsymbol{T}$	F		F
				1		2	1		1

Далее записываем истинностные значения $(\sim q) \wedge r$ под символом $\wedge$.

Случай	p	q	r	p	$\vee$	$((\sim$	$q)$	$\wedge$	$r)$
1	T	T	T	T		F	T	$\boldsymbol{F}$	T
2	T	T	F	T		F	T	$\boldsymbol{F}$	F
3	T	F	T	T		T	F	$\boldsymbol{T}$	T
4	T	F	F	T		T	F	$\boldsymbol{F}$	F
5	F	T	T	F		F	T	$\boldsymbol{F}$	T
6	F	T	F	F		F	T	$\boldsymbol{F}$	F
7	F	F	T	F		T	F	$\boldsymbol{T}$	T
8	F	F	F	F		T	F	$\boldsymbol{F}$	F
				1		2	1	3	1

Наконец, записываем значения высказывания $p \vee ((\sim q) \wedge r)$ под символом $\vee$.

Случай	p	q	r	p	$\vee$	$((\sim$	$q)$	$\wedge$	$r)$
1	T	T	T	T	$\boldsymbol{T}$	F	T	F	T
2	T	T	F	T	$\boldsymbol{T}$	F	T	F	F
3	T	F	T	T	$\boldsymbol{T}$	T	F	T	T
4	T	F	F	T	$\boldsymbol{T}$	T	F	F	F
5	F	T	T	F	$\boldsymbol{F}$	F	T	F	T
6	F	T	F	F	$\boldsymbol{F}$	F	T	F	F
7	F	F	T	F	$\boldsymbol{T}$	T	F	T	T
8	F	F	F	F	$\boldsymbol{F}$	T	F	F	F
				1	4	2	1	3	1

ПРИМЕР 1.1. Пусть p, q и r обозначают, соответственно, высказывания *Фрэд любит футбол*, *Фрэд любит гольф*, *Фрэд любит теннис*. Требуется записать высказывание *Фрэд любит футбол и неверно, что он любит гольф или теннис* в символической форме и указать соответствующую ему таблицу истинности.

Сначала заменим это высказывание эквивалентным — *Фрэду нравится футбол и неверно, что Фрэд любит гольф или теннис*. Высказывание *Фрэд любит гольф или теннис* в символической форме записывается как $q \vee r$. Высказывание *Неверно, что Фрэд любит гольф или теннис*, символически записывается как $\sim(q \vee r)$, поскольку отрицание применяется ко всему высказыванию, которое следует после "что". Итак, исходное высказывание символически изображается $p \wedge (\sim(q \vee r))$. Таблица истинности этого высказывания имеет вид

Случай	p	q	r	p	$\wedge$	$(\sim$	$(q \vee r))$
1	T	T	T	T	$\boldsymbol{F}$	F	T
2	T	T	F	T	$\boldsymbol{F}$	F	T
3	T	F	T	T	$\boldsymbol{F}$	F	T
4	T	F	F	T	$\boldsymbol{T}$	T	F
5	F	T	T	F	$\boldsymbol{F}$	F	T
6	F	T	F	F	$\boldsymbol{F}$	F	T
7	F	F	T	F	$\boldsymbol{F}$	F	T
8	F	F	F	F	$\boldsymbol{F}$	T	F
				1	$*$	3	2

□

■ УПРАЖНЕНИЯ

1. Найдите среди указанных ниже предложений высказывания. Укажите их истинностные значения.

а) Который час?

б) Целое число 1 есть наименьшее положительное целое число.

в) Если $x = 3$, то $x^2 = 6$.

г) Берегись автомобиля!

д) Южная Дакота — южный штат.

2. Найдите среди указанных ниже предложений высказывания. Укажите их истинностные значения.

а) Все четные числа делятся на 2.

б) Загрузите пакеты в машину.

в) Это утверждение не может быть истинным.

г) Юпитер — ближайшая к солнцу планета.

д) Не следует хранить компакт-диски в микроволновой печи.

3. Пусть p, q и r обозначают следующие высказывания:

p : *Путешествие на Марс является дорогостоящим.*

q : *Я совершу путешествие на Марс.*

r : *У меня есть деньги.*

Запишите в символической форме такие высказывания:

а) *У меня нет денег и я не совершу путешествие на Марс.*

б) *У меня нет денег и путешествие на Марс является дорогостоящим или я совершу путешествие на Марс.*

в) *Неверно, что у меня есть деньги и я полечу на Марс.*

г) *Путешествие на Марс не является дорогостоящим и я полечу на Марс или путешествие на Марс является дорогостоящим и я не полечу на Марс.*

4. Пусть p, q и r обозначают следующие высказывания:

p : *Мой компьютер — быстродействующий.*

q : *Я окончу проект вовремя.*

r : *Я сдам экзамен.*

Запишите в символической форме такие высказывания:

а) *У меня не быстродействующий компьютер или я закончу проект вовремя.*

б) *Я не закончу проект вовремя и не сдам экзамен.*

в) *Неверно, что я закончу проект и сдам экзамен.*

г) *У меня быстродействующий компьютер или я не закончу проект вовремя и сдам экзамен.*

5. Постройте таблицы истинности для каждого высказывания в упражнении 3.

6. Постройте таблицы истинности для каждого высказывания в упражнении 4.

7. Пусть p, q и r обозначают следующие высказывания:

p : *Эта игра очень трудна.*

q : *Я играю в шахматы.*

r : *Игра в шахматы требует времени.*

Интерпретируйте следующие выражения как обычные высказывания:

а) $q \wedge r$;

б) $\sim p \vee \sim q$;

в) $(p \vee r) \wedge q$;

г) $p \wedge q \wedge r$.

8. Пусть p, q и r обозначают следующие высказывания:

p : *Доги — большие собаки.*

q : *У меня маленький дом.*

r : *У меня есть дог.*

Представьте следующие символические выражения как обычные высказывания:

а) $p \wedge q \wedge \sim r$;

б) $p \wedge (\sim q \vee \sim r)$;

в) $(p \vee \sim q) \wedge r$;

г) $(p \wedge r) \vee (q \wedge \sim r)$.

9. Постройте таблицы истинности для высказываний в упражнении 7.

10. Постройте таблицы истинности для высказываний в упражнении 8.

11. Постройте таблицы истинности для следующих высказываний:

а) $p \wedge (q \vee \sim r)$;

б) $(q \wedge \sim r) \vee (\sim p \wedge r)$;

в) $\sim(p \wedge r) \vee (\sim q \wedge r)$;

г) $\sim(\sim p \vee (q \wedge \sim r))$;

д) $(p \wedge r) \vee (p \wedge \sim q)$.

12. Постройте таблицы истинности для следующих высказываний:

а) $(p \vee q) \wedge (r \vee q)$;

б) $(\sim q \wedge r) \vee \sim(p \wedge r)$;

в) $\sim((p \wedge r) \vee \sim q)$;

г) $\sim(\sim p \wedge (q \vee \sim r))$;

д) $(p \vee \sim r) \wedge \sim(p \vee \sim q)$.

1.2. УСЛОВНЫЕ ВЫСКАЗЫВАНИЯ

Допустим, некто утверждает, что если случится одно событие, то случится и другое. Это утверждение зачастую принимает форму угрозы, но мы рассмотрим более позитивный пример. Предположим, отец говорит сыну: "Если в этом семестре ты сдашь все экзамены на «отлично», я куплю тебе машину". Заметьте, что высказывание имеет вид: *если p, то q*, где p — высказывание *В этом семестре ты сдашь все экзамены на «отлично»*, а q — высказывание *Я куплю тебе машину*. Сложное высказывание мы обозначим символически через $p \rightarrow q$. Спрашивается, при каких условиях отец говорит правду? Предположим, высказывания p и q истинны. В этом случае счастливый студент получает отличные оценки по всем предметам, и приятно удивленный отец покупает ему машину. Естественно, ни у кого не вызывает сомнения тот факт, что высказывание отца было истинным. Однако существуют еще три других случая, которые необходимо рассмотреть. Допустим, студент действительно добился отличных результатов, а отец не купил ему машину. Самое мягкое, что можно сказать об отце в таком случае, — это то, что он солгал. Следовательно, если p истинно, а q ложно, то $p \rightarrow q$ ложно. Допустим теперь, что студент не получил положительные оценки, но отец тем не менее купил ему машину. В этом случае отец предстает очень щедрым, но его никак нельзя назвать лжецом. Следовательно, если p ложно и q истинно, то высказывание *если p, то q* (т.е. $p \rightarrow q$) истинно. Наконец, предположим, что студент не добился отличных результатов, и отец не купил ему машину. Поскольку студент не выполнил свою часть соглашения, отец тоже свободен от обязательств. Таким образом, если p и q ложны, то $p \rightarrow q$ считается истинным. Итак, единственный случай, когда отец солгал, — это когда он дал обещание и не выполнил его.

Таким образом, таблица истинности для высказывания $p \rightarrow q$ имеет вид

Случай	p	q	$p \rightarrow q$
1	T	T	T
2	T	F	F
3	F	T	T
4	F	F	T

Символ $\rightarrow$ называется ***импликацией***, или ***условной связкой***.

Еще одним примером, который подтверждает таблицу истинности для импликации, является высказывание

Если целое число равно 3, *то его квадрат равен* 9.

Конечно, желательно, чтобы это высказывание было истинным всегда. Пусть p — высказывание *Целое число равно 3*, а q — высказывание *Квадрат целого числа равен 9*. Если p и q истинны, то целое число равно 3 и его квадрат равен 9. Это соответствует первой строке таблицы истинности. Если целое число равно -3, то его квадрат по-прежнему равен 9. В этом случае p ложно и q истинно, что совпадает со случаем 3 и подтверждает правильность третьей строки таблицы. Если целое число равно 4, то ложны и p, и q, что соответствует случаю 4. Этим подтверждается правильность четвертой строки таблицы. Заметим, что случай 2 здесь не имеет места, но во всех других случаях эта таблица истинности обеспечила нам желаемый результат.

Может показаться, что $p \rightarrow q$ носит характер причинно-следственной связи, но это не является необходимым. Чтобы увидеть отсутствие причины и следствия в импликации, вернемся к примеру, в котором p есть высказывание *Джейн управляет автомобилем*, а q — утверждение *У Боба русые волосы*. Тогда высказывание *Если Джейн управляет автомобилем, то у Боба русые волосы* запишется как

если p, то q

или как

$$p \rightarrow q\,.$$

То, что Джейн управляет автомобилем, никак причинно не связано с тем, что Боб русоволосый. Однако нужно помнить, что истинность или ложность бинарного сложного высказывания зависит только от истинности составляющих его частей и не зависит от наличия или отсутствия между ними какой-либо связи.

ПРИМЕР 1.2. Требуется найти таблицу истинности для выражения

$$(p \rightarrow q) \wedge (q \rightarrow r)\,.$$

Используя таблицу истинности для $\rightarrow$, приведенную выше, построим сначала таблицы истинности для $(p \rightarrow q)$ и $(q \rightarrow r)$, учитывая, что импликация ложна только в случае, когда $T \rightarrow F$.

Случай	p	q	r	$(p$	$\rightarrow$	$q)$	$\wedge$	$(q$	$\rightarrow$	$r)$
1	T	T	T	T	**T**	T		T	**T**	T
2	T	T	F	T	**T**	T		T	**F**	F
3	T	F	T	T	**F**	F		F	**T**	T
4	T	F	F	T	**F**	F		F	**T**	F
5	F	T	T	F	**T**	T		T	**T**	T
6	F	T	F	F	**T**	T		T	**F**	F
7	F	F	T	F	**T**	F		F	**T**	T
8	F	F	F	F	**T**	F		F	**T**	F
				1	**2**	1		1	**2**	1

Теперь используем таблицу для $\wedge$, чтобы получить для высказывания

$$(p \to q) \wedge (q \to r)$$

таблицу истинности

Случай	p	q	r	$(p$	$\to$	$q)$	$\wedge$	$(q$	$\to$	$r)$
1	T	T	T	T	T	T	**T**	T	T	T
2	T	T	F	T	T	T	**F**	T	F	F
3	T	F	T	T	F	F	**F**	F	T	T
4	T	F	F	T	F	F	**F**	F	T	F
5	F	T	T	F	T	T	**T**	T	T	T
6	F	T	F	F	T	T	**F**	T	F	F
7	F	F	T	F	T	F	**T**	F	T	T
8	F	F	F	F	T	F	**T**	F	T	F
				1	2	1	*	1	2	1

□

Высказывание вида $(p \to q) \wedge (q \to p)$ обозначается через $p \leftrightarrow q$. Символ $\leftrightarrow$ называется ***эквиваленцией***. Очевидно, таблица истинности для $(p \to q) \wedge (q \to p)$ определяет таблицу истинности для $p \leftrightarrow q$.

Случай	p	q	$(p \to q)$	$\wedge$	$(q \to p)$	$p \leftrightarrow r$
1	T	T	T	T	T	T
2	T	F	F	F	T	F
3	F	T	T	F	F	F
4	F	F	T	T	T	T
				*		*

Непосредственно из определения вытекает, что эквиваленция истинна только в случае, когда p и q имеют одинаковые истинностные значения.

Может возникнуть вопрос о том, как интерпретировать такие выражения, как $\sim p \vee q$, $p \wedge q \vee r$, $p \wedge q \to r$ и $p \wedge q \leftrightarrow q \vee r$, в которых отсутствуют скобки. Во избежание неоднозначности лучше всегда использовать скобки. Однако здесь, как и в алгебре, имеется приоритет выполнения операций. Операции выполняются в следующей последовательности: $\sim$, $\wedge$, $\vee$, $\to$ и $\leftrightarrow$. Поэтому указанные выражения можно интерпретировать как $(\sim p) \vee q$, $(p \wedge q) \vee r$, $(p \wedge q) \to r$ и $(p \wedge q) \leftrightarrow (q \vee r)$.

■ УПРАЖНЕНИЯ

1. Пусть p, q и r обозначают следующие высказывания:

p : *Он купит компьютер.*
q : *Он будет праздновать всю ночь.*
r : *Он выиграет в лотерею.*

Запишите следующие высказывания в виде символических выражений:

а) *Если он выиграет в лотерею, он купит компьютер и будет праздновать всю ночь.*

б) *Если он не купит компьютер, то и праздновать всю ночь не будет.*

в) *Если он выиграет в лотерею, то будет праздновать всю ночь и если он не выиграет в лотерею, то не купит компьютер.*

г) *Если он не выиграет в лотерею или не купит компьютер, то праздновать всю ночь не будет.*

2. Пусть p, q и r обозначают следующие высказывания:

p : *Он читает комиксы.*

q : *Он любит научную фантастику.*

r : *Он — ученый-информатик.*

Запишите следующие высказывания в виде символических выражений:

а) *Если он читает комиксы и любит научную фантастику, то он — ученый-информатик.*

б) *Если он не читает комиксы и не любит научную фантастику, то он не является ученым-информатиком.*

в) *Если он читает комиксы, то он любит научную фантастику и если он не читает комиксы, то он — ученый-информатик.*

г) *Если он — ученый-информатик, то он читает комиксы или он не любит научную фантастику.*

3. Пусть p, q и r обозначают следующие высказывания:

p : *Ему нравятся фиолетовые галстуки.*

q : *Он популярен.*

r : *У него странные друзья.*

Запишите следующие символические выражения в виде высказываний:

а) $(p \wedge q) \rightarrow r$;

б) $q \rightarrow \sim r$;

в) $p \rightarrow (q \vee r)$;

г) $(p \rightarrow \sim q) \wedge (q \rightarrow r)$.

4. Пусть p, q и r обозначают следующие высказывания:

p : *Он удачлив.*

q : *Он популярен.*

r : *Он богат.*

Запишите следующие символические выражения как высказывания:

а) $\sim(p \rightarrow q)$;

б) $(p \vee r) \rightarrow q$;

в) $q \leftrightarrow (p \wedge r)$;

г) $(p \rightarrow q) \wedge (\sim r \rightarrow (\sim p \vee \sim q))$.

5. Постройте таблицы истинности для следующих выражений:

а) $(p \rightarrow q) \rightarrow r$;

б) $p \rightarrow (q \rightarrow r)$;

в) $q \rightarrow (p \wedge r) \leftrightarrow ((q \rightarrow p) \wedge (q \rightarrow r))$;

г) $((p \rightarrow q) \vee r) \rightarrow (\sim p \vee \sim q)$.

6. Постройте таблицы истинности для следующих выражений:

а) $(p \rightarrow q) \rightarrow (q \rightarrow r)$;

б) $(p \to q) \vee \sim(r \wedge q)$;
в) $(p \vee r) \to (p \wedge q)$;
г) $\sim((p \to q) \wedge \sim r) \to (p \vee \sim r)$;
д) $((p \to q) \wedge \sim(r \vee p)) \to (\sim p \vee \sim q)$.

7. Постройте таблицы истинности для следующих выражений:
а) $(p \to q) \leftrightarrow (\sim q \to \sim p)$;
б) $(p \wedge \sim(q \vee \sim r)) \leftrightarrow (p \to q)$;
в) $(p \vee r) \to q$;
г) $((p \to q) \wedge (q \to \sim r)) \to (r \to p)$;
д) $(p \wedge (q \vee r)) \to ((p \wedge q) \vee (p \wedge r))$.

8. Постройте таблицы истинности для следующих выражений:
а) $(\sim(p \wedge \sim r) \vee q) \to (q \vee r)$;
б) $\sim((p \wedge \sim q) \vee r) \leftrightarrow (r \to q)$;
в) $(\sim(p \to q) \to (q \to r))$;
г) $(p \vee q) \wedge (p \to r)$;
д) $((p \vee r) \to q) \to ((p \to q) \vee (p \to r))$.

9. Укажите, какие из следующих высказываний являются истинными:
а) *Если* $2^2 = 4$, *то* $3^2 = 9$;
б) *Если* $2^2 = 5$, *то* $3^2 = 9$;
в) *Если* $2^2 = 5$, *то* $3^2 = 10$;
г) *Если* $2^2 = 4$, *то* $3^2 = 10$.

10. Таблица истинности простого высказывания содержит две строки. Высказывание, состоящее из двух компонент, имеет таблицу истинности из четырех строк; сложное высказывание, составленное из трех простых, имеет таблицу истинности из восьми строк. Сколько строк содержит таблица истинности высказывания, состоящего из четырех компонент? А если высказывание состоит из n компонент?

1.3. ЭКВИВАЛЕНТНЫЕ ВЫСКАЗЫВАНИЯ

Особый интерес представляют сложные высказывания, имеющие различное строение, но являющиеся истинными в одних и тех же случаях. Такие высказывания называются ***логически эквивалентными***. Эквивалентность двух высказываний легко установить посредством сравнения их таблиц истинности.

Например, пусть p и q обозначают высказывания

p: *Сегодня шел дождь.*
q: *Сегодня шел снег.*

Рассмотрим сложные высказывания

Неверно, что сегодня шел дождь или снег,

или символически

$$\sim(p \vee q)\,,$$

и

Сегодня не шел дождь и сегодня не шел снег,

или символически

$$\sim p \wedge \sim q\,.$$

Построим таблицы истинности для обоих высказываний.

Случай	p	q	$\sim$	$(p \vee q)$	$\sim p$	$\wedge$	$\sim q$
1	T	T	F	T	F	F	F
2	T	F	F	T	F	F	T
3	F	T	F	T	T	F	F
4	F	F	T	F	T	T	T
			*			#	

Итак, во всех четырех строках истинностные значения для $\sim(p \vee q)$ (обозначенные *) и для $\sim p \wedge \sim q$ (обозначенные #) совпадают. Это означает, что два рассматриваемых высказывания логически эквивалентны, т.е.

$$\sim(p \vee q) \equiv \sim p \wedge \sim q\,.$$

Эквивалентность — очень полезное свойство; используя его, можно строить отрицание высказываний с "или", осуществляя отрицание каждой из его частей и меняя "или" на "и".

С условным высказыванием — импликацией $p \to q$ — связаны еще три типа высказываний: ***конверсия***, ***инверсия*** и ***контрапозиция*** высказывания $p \to q$. Они определяются следующим образом:

$p \to q$	импликация
$q \to p$	конверсия высказывания $p \to q$
$\sim p \to \sim q$	инверсия высказывания $p \to q$
$\sim q \to \sim p$	контрапозиция высказывания $p \to q$

Пусть дано высказывание-импликация *Если он играет в футбол, то он популярен*. Для этой импликации имеем:

конверсия:	*Если он популярен, то он играет в футбол*
инверсия:	*Если он не играет в футбол, то он не популярен*
контрапозиция:	*Если он не популярен, то он не играет в футбол*

Важно понимать, что высказывания *Если он живет в Детройте, то Боб навестит его* и *Боб навестит его, если он живет в Детройте* по сути являются одним и тем же высказыванием. Однако высказывание *Если Боб навестит его, то он живет в Детройте* не совпадает с предыдущими высказываниями. Не важен порядок, в котором p и q присутствуют в предложении, а важно, какая

часть предложения является частью "если", а какая часть является частью "то". Может показаться, что при замене *если p, то q* на *q, если p* получается конверсия, но с логической точки зрения последнее высказывание совпадает с исходным.

Часть (ж) теоремы, которая будет сформулирована ниже, говорит о том, что импликация и ее контрапозиция логически эквивалентны. Эквивалентность и контрапозиция условных высказываний имеют в математике большое значение. Зачастую гораздо легче доказать теорему от противного, чем дать ее прямое доказательство. Используя эквивалентность импликации и ее контрапозиции, нетрудно показать, что конверсия и инверсия импликации имеют одну и ту же таблицу истинности. В то же время импликация и ее конверсия (или инверсия) имеют различные таблицы истинности. Непонимание этого факта — источник распространенных ошибок в так называемых "логических" рассуждениях.

ТЕОРЕМА 1.3. Используя таблицы истинности, можно доказать следующие логические эквивалентности:

а) ***Законы идемпотентности***

$$p \wedge p \equiv p;$$
$$p \vee p \equiv p.$$

б) ***Закон двойного отрицания***

$$\sim(\sim p) \equiv p.$$

в) ***Законы де Моргана***

$$\sim(p \vee q) \equiv \sim p \wedge \sim q;$$
$$\sim(p \wedge q) \equiv \sim p \vee \sim q.$$

г) ***Свойства коммутативности***

$$p \wedge q \equiv q \wedge p;$$
$$p \vee q \equiv q \vee p.$$

д) ***Свойства ассоциативности***

$$p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r;$$
$$p \vee (q \vee r) \equiv (p \vee q) \vee r.$$

е) ***Свойства дистрибутивности***

$$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r);$$
$$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r).$$

ж) ***Закон контрапозиции***

$$p \to q \equiv \sim q \to \sim p.$$

з) ***Другие полезные свойства***

$$p \to q \equiv \sim p \vee q;$$
$$p \leftrightarrow q \equiv (p \to q) \wedge (q \to p).$$

Отметим, что благодаря свойству ассоциативности высказывания $(p \wedge q) \wedge r$ и $p \wedge (q \wedge r)$ могут быть записаны в виде $p \wedge q \wedge r$. Аналогично, $(p \vee q) \vee r$ и $p \vee (q \vee r)$ можно записать просто как $p \vee q \vee r$.

Высказывание, истинное во всех случаях, называется ***логически истинным***, или ***тавтологией***; высказывание, построенное так, что оно ложно в каждом случае, называется ***логически ложным***, или ***противоречием***. Теоремы в математике являются примерами тавтологий. Было бы неприятно сознавать, что математические теоремы справедливы, к примеру, только в 80 процентах случаев. Высказывание *Он сдаст или не сдаст экзамен* есть пример тавтологии, поскольку либо одно событие, либо другое обязательно должно иметь место. Каждое высказывание вида $p \vee \sim p$ — тавтология. Шекспир должен был бы провозгласить: "Быть или не быть, это и есть тавтология". Высказывание *Если сдаст, то сдаст* — тоже тавтология, хотя оно и не отличается глубиной. Примером тавтологии является высказывание *Если он богат и удачлив, то он удачлив*. Высказывание *Она движется в направлении Техаса и она не движется в направлении Техаса* всегда ложно, т.к. нельзя делать одновременно и то, и другое. Следовательно, это противоречие.

Рассмотрим высказывание вида

$$(p \wedge (p \rightarrow q)) \rightarrow q\,.$$

Ему соответствует таблица истинности

Случай	p	q	$(p$	$\wedge$	$(p \rightarrow q))$	$\rightarrow$	q
1	T	T	T	T	T	T	T
2	T	F	T	F	F	T	F
3	F	T	F	F	T	T	T
4	F	F	F	F	T	T	F
						*	

Столбец, обозначенный звездочкой, дает истинностные значения исходного сложного высказывания. Высказывание истинно во всех четырех возможных случаях, следовательно, является тавтологией.

Имея логически истинное высказывание-тавтологию, легко построить логически ложное высказывание-противоречие. Для этого достаточно взять отрицание логически истинного высказывания. Поэтому высказывание

$$\sim((p \wedge (p \rightarrow q)) \rightarrow q)$$

логически ложно.

Давайте теперь рассмотрим некоторые свойства, связанные с логически истинными и логически ложными высказываниями. Пусть символ $\boldsymbol{T}$ обозначает высказывание, которое есть тавтология, поэтому имеет таблицу истинности, состоящую из одних T. Символом $\boldsymbol{F}$ обозначим противоречие, т.е. высказывание, таблица истинности которого содержит F во всех строках. Используя таблицы

истинности, можно показать, что

$$\begin{array}{lllll} p & \wedge & \boldsymbol{T} & \equiv & p\,, \\ p & \wedge & \boldsymbol{F} & \equiv & \boldsymbol{F}, \\ p & \vee & \boldsymbol{T} & \equiv & \boldsymbol{T}, \\ p & \vee & \boldsymbol{F} & \equiv & p\,, \\ p & \wedge & \sim p & \equiv & \boldsymbol{F}, \\ p & \vee & \sim p & \equiv & \boldsymbol{T}, \\ p & \rightarrow & p & \equiv & \boldsymbol{T}. \end{array}$$

В каждом высказывании можно заменить любую его компоненту на логически эквивалентное ей высказывание. Полученное в результате такой замены высказывание будет логически эквивалентно исходному, поскольку истинностное значение высказывания зависит исключительно от истинностных значений составляющих его компонент (но не от их формы или сложности). Например,

$$\begin{array}{lll} (q \vee r) \vee (p \wedge \sim r) & \equiv & q \vee (r \vee (p \wedge \sim r) \equiv \quad \text{свойство ассоциативности} \\ & \equiv & q \vee ((r \vee p) \wedge (r \vee \sim r)) \equiv \quad \text{свойство дистрибутивности} \\ & \equiv & q \vee ((r \vee p) \wedge \boldsymbol{T}) \equiv \quad \text{эквивалентность} \\ & \equiv & q \vee (r \vee p) \equiv \quad \text{эквивалентность} \\ & \equiv & q \vee (p \vee r) \equiv \quad \text{свойство коммутативности} \\ & \equiv & (q \vee p) \vee r \equiv \quad \text{свойство ассоциативности} \\ & \equiv & (p \vee q) \vee r\,. \quad \text{свойство коммутативности} \end{array}$$

Здесь мы впервые использовали законы логики и фиксированное множество "истинных высказываний" для образования новых "истинных" высказываний.

Условные высказывания могут выражаться в виде различных языковых конструкций, но символически все они записываются как $p \rightarrow q$. Вот несколько примеров таких конструкций:

Если p, *то* q.

p *достаточно для* q.
p *является достаточным условием для* q.

q *необходимо для* p.
q *является необходимым условием для* p.

p, *только если* q *(или: только если* q, *то* p*)*.

Таблица для $p \rightarrow q$ показывает, что если $p \rightarrow q$ истинно и p истинно, тогда q должно быть истинным, т.е. истинность p достаточна для истинности q. Поэтому p *достаточно для* q имеет тот же смысл, что и $p \rightarrow q$. Таким образом, если q ложно и q *необходимо для* p, тогда p должно быть ложно. Поэтому, если $\sim q$ истинно, тогда $\sim p$ должно быть истинно и $\sim q \rightarrow \sim p$. Но последнее выражение — контрапозиция для $p \rightarrow q$; следовательно, q *необходимо для* p имеет то же значение, что и $p \rightarrow q$.

Анализ значения p *только если* q проводится аналогично. Получаем, что p может быть истинным, только если q истинно. Если q не истинно, то p не может

быть истинным. Но это эквивалентно утверждению, что если $\sim q$ истинно, то $\sim p$ должно быть истинно и $\sim q \rightarrow \sim p$. Последнее есть контрапозиция высказывания $p \rightarrow q$, поэтому p *только если* q имеет то же значение, что и $p \rightarrow q$.

Ранее было установлено, что порядок простых высказываний в составе условного высказывания не имеет значения; важно то, какое из простых высказываний следует за *если* (или в данном случае *только если*), а какое следует за *то*. Поэтому утверждения *Только накопив денег, Фрэд сможет поступить в колледж* и *Фрэд сможет поступить в колледж, только если накопит денег* рассматриваются как одно и то же высказывание.

ПРИМЕР 1.4. Приведите следующие высказывания к виду $p \rightarrow q$ или $q \rightarrow p$:

а) *Он добьется успеха, только если будет усердно работать.*

б) *Он счастлив, только если управляет автомобилем.*

в) *Наличия денег достаточно, чтобы быть счастливым.*

г) *Наличие денег необходимо, чтобы быть счастливым.*

д) *Чтобы победить на выборах, нужно набрать достаточно голосов.*

Ответ:

а) Если через p обозначить высказывание *Он добьется успеха*, а через q — *Он будет усердно работать*, то высказывание *если* p, *то* q, или $p \rightarrow q$ примет вид *Если он добивается успеха, то он работает усердно.*

б) Если p обозначает *Он управляет автомобилем*, а q — *Он счастлив*, то высказывание q *только если* p, или $q \rightarrow p$ имеет вид *Если он счастлив, то он управляет автомобилем.*

в) Пусть p обозначает *У него есть деньги*, а q есть *Он счастлив*. Тогда высказывание p достаточно для q, или $p \rightarrow q$ может быть записано как *Если у него есть деньги, он счастлив.*

г) Пусть p обозначает *У него есть деньги*, а q обозначает *Он счастлив*. Тогда высказывание p *необходимо для* q, или $q \rightarrow p$ будет иметь вид *Если он счастлив, то у него есть деньги.*

д) Пусть p обозначает *Он победит на выборах*, а q обозначает *Он получит достаточное количество голосов*, тогда высказывание q *необходимо для* p, или *если* p, *то* q, или $p \rightarrow q$ может быть записано в виде *Если он победит на выборах, то он получит достаточное количество голосов.* □

Вернемся к рассмотрению логической связки $\leftrightarrow$. Поскольку высказывания вида $p \leftrightarrow q$ и $(p \rightarrow q) \wedge (q \rightarrow p)$ логически эквивалентны, то $p \leftrightarrow q$ означает то же, что и p *тогда и только тогда, когда* q, или p *если и только если* q. Пусть, например, p обозначает высказывание *Джим будет играть в футбол*, а q — высказывание *Джейн организует поддержку зрителей*. Тогда $p \leftrightarrow q$ может быть выражено как

Джим будет играть в футбол, если и только если
Джейн организует поддержку зрителей.

Следующие языковые конструкции, выражающие эквиваленцию высказыва-

ний $p \leftrightarrow q$, равносильны:

p если и только если q.
p необходимо и достаточно для q.
p есть необходимое и достаточное условие для q.

Пусть имеется высказывание

Сэм играет в гольф тогда и только тогда, когда тепло,

тогда ожидается, что если будет тепло, то мы увидим Сэма, играющего в гольф, и если Сэм играет в гольф, то погода, безусловно, теплая.

Точно так же, если имеется высказывание

Быть счастливым является необходимым и достаточным условием, чтобы быть удачливым,

то имеется в виду, что если Джон счастливчик, то он и удачлив, а если Джону везет, то он, конечно, счастливчик.

■ УПРАЖНЕНИЯ

1. Используя таблицы истинности, докажите следующие эквивалентности:

а) Закон де Моргана

$$\sim(p \wedge q) \equiv \sim p \vee \sim q\,.$$

б) Свойство ассоциативности связки $\vee$

$$p \vee (q \vee r) \equiv (p \vee q) \vee r\,.$$

в) Свойство дистрибутивности связки *или* относительно *и*

$$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)\,.$$

г) Эквивалентность импликации и высказывания со связкой *или*

$$p \rightarrow q \equiv \sim p \vee q\,.$$

2. Используя пункт (г) предыдущего упражнения, покажите, что отрицание для $p \rightarrow q$ эквивалентно $p \wedge \sim q$.

3. Ранее были рассмотрены связанные между собой условные высказывания, включающие в качестве компонент высказывания p и q. Ими являлись

$p \rightarrow q$	импликация,
$q \rightarrow p$	конверсия (для $p \rightarrow q$),
$\sim q \rightarrow \sim p$	контрапозиция (для $p \rightarrow q$),
$\sim p \rightarrow \sim q$	инверсия (для $p \rightarrow q$).

Используя таблицы истинности, докажите, что

$$p \rightarrow q \not\equiv q \rightarrow p\,.$$

Импликация эквивалентна своей контрапозиции, но не эквивалентна своей конверсии. Часто используется выражение *если p, то q, и наоборот*. На самом деле это означает *если p, то q, и если q, то p*, или

$$(p \rightarrow q) \wedge (q \rightarrow p),$$

что эквивалентно $p \leftrightarrow q$, или *p, если и только если q*. Докажите, что $p \rightarrow q \equiv \sim q \rightarrow \sim p$, не используя непосредственно таблицы истинности. На основании этого результата докажите, что инверсия импликации эквивалентна ее конверсии.

4. Используя логически эквивалентные высказывания и не применяя непосредственно таблицы истинности, покажите, что
   а) $p \equiv \sim(p \wedge s) \rightarrow (\sim s \wedge p)$.
   б) $\sim(p \leftrightarrow q) \equiv (p \wedge \sim q) \vee (q \wedge \sim p)$.
5. Преобразуйте следующие высказывания к виду *если p, то q*:
   а) *Он кентавр, только если он имеет шесть ног.*
   б) *Чтобы быть преуспевающим политиком, нужно быть избранным.*
   в) *Достаточно иметь деньги, чтобы быть популярным.*
6. Преобразуйте следующие высказывания к виду *если p, то q*:
   а) *Необходимо иметь шлем, чтобы играть в американский футбол.*
   б) *Только если я читаю Шекспира, я литературно образован.*
   в) *Для меня сдать этот курс достаточно, чтобы получить диплом.*
7. Дано высказывание *Если я голосую, то я хороший гражданин.*
   а) Сформулируйте конверсию этого выражения.
   б) Сформулируйте инверсию этого выражения.
   в) Сформулируйте контрапозицию этого выражения.
8. Дано высказывание *Если я не буду выплачивать ссуду, у меня отберут участок.*
   а) Сформулируйте конверсию этого высказывания.
   б) Сформулируйте инверсию этого высказывания.
   в) Сформулируйте контрапозицию этого высказывания.

1.4. АКСИОМАТИЧЕСКИЕ СИСТЕМЫ: УМОЗАКЛЮЧЕНИЯ И ДОКАЗАТЕЛЬСТВА

Математики в большинстве своем имеют дело с теоремами и их доказательствами. Теоремы представляют собой "истинные" утверждения относительно рассматриваемых математических систем. Например, утверждение

Гипотенуза прямоугольного треугольника
длиннее любого из катетов

— это известная из геометрии теорема Евклида. Это утверждение считается истинным, поскольку оно "выводимо" из ранее принятых или выведенных истин геометрии Евклида.

Математическая система начинается с неопределяемых понятий и утверждений, точно описывающих фундаментальные характеристики или истинные утверждения относительно этих понятий, которые математики используют для образования системы. Эти фундаментальные характеристики называются аксиомами или постулатами. Утверждения, выведенные (доказанные) только на основе этих фундаментальных свойств (аксиом и постулатов) и ранее доказанных утверждений с помощью логических правил, называются ***теоремами***.

Таким образом, в математических системах вся информация, необходимая для доказательства теоремы, должна содержаться в аксиомах и ранее доказанных теоремах. Развивая конкретный раздел математики, можно не включать в него все аксиомы и доказанные теоремы. Вместо этого можно принять доказанные теоремы в качестве аксиом. Например, аксиомы для целых чисел и аксиомы Пеано для положительных целых чисел неявным образом предполагают выполнение аксиом теории множеств, но, поскольку в теории чисел акцент делается на свойствах целых чисел, было бы излишним стремиться в ней одновременно к полному развитию теории множеств.

Важно, что логические правила, которые используются для вывода новых теорем из аксиом, постулатов и ранее доказанных в данной системе теорем, не порождают в качестве "теорем" ложные высказывания. Эти логические правила называются ***правилами вывода***. ***Умозаключение*** состоит из совокупности утверждений, называемых гипотезами, или посылками, и утверждения, называемого заключением. ***Правильным умозаключением*** называется такое умозаключение, заключение которого истинно всякий раз, когда истинны его гипотезы. Правила вывода выбираются так, чтобы они были правильными умозаключениями.

Умозаключения часто представляют в виде

$$\begin{array}{ll} & H_1 \\ & H_2 \quad \text{гипотезы} \\ & \underline{H_3} \\ \therefore & C \quad \text{заключение} \end{array}$$

Символ $\therefore$ означает "следовательно". Гипотезы представляют собой перечень одного или более высказываний, или ***посылок***. Умозаключение правильно, если всякий раз, когда $H_1, H_2,$ и H_3 истинны, то истинно и C или, что равносильно, всякий раз, когда $H_1 \wedge H_2 \wedge H_3$ истинно, истинно и C.

Правильность умозаключения можно проверить двумя способами. Во-первых, мы можем построить таблицу истинности и показать, что всякий раз, когда гипотезы истинны, истинно и заключение. Во-вторых, мы можем использовать таблицы истинности для обоснования правил вывода, а затем использовать правила вывода для доказательства справедливости заключения. Длинные умозаключения, как правило, проще обосновывать при помощи правил вывода. В частности, используя правила вывода, легко проверить правильность умозаключения

$$\begin{array}{l} p \rightarrow q \\ q \rightarrow r \\ r \rightarrow s \\ \underline{s \rightarrow t} \\ \therefore \; p \rightarrow t \end{array}$$

Проверка того же утверждения с использованием таблиц истинности намного сложнее. Однако, при помощи таблиц истинности мы можем доказать ложность утверждения, чего правила вывода сделать не позволяют.

Рассмотрим умозаключение

$$\begin{array}{l} p \\ p \to q \\ q \to r \\ \hline \therefore\ p \wedge q \wedge r \end{array}$$

Таблицы истинности для посылок и заключения имеют следующий вид.

Случай	p	q	r	p	$p \to q$	$q \to r$	$p \wedge q \wedge r$
1	T	T	T	T	T	T	T
2	T	T	F	T	T	F	F
3	T	F	T	T	F	T	F
4	T	F	F	T	F	T	F
5	F	T	T	F	T	T	T
6	F	T	F	F	T	F	F
7	F	F	T	F	T	T	T
8	F	F	F	F	T	T	T
				1	2	3	*

Заметим, что, когда истинны все посылки (что имеет место в случае 1), истинным также является и заключение, а само умозаключение является правильным.

Рассмотрим умозаключение

$$\begin{array}{l} p \vee q \\ p \to r \\ q \to r \\ \hline \therefore\ r \end{array}$$

Построим для него таблицу истинности.

Случай	p	q	r	$p \wedge q$	$p \to q$	$q \to r$	r
1	T	T	T	T	T	T	T
2	T	T	F	T	F	F	F
3	T	F	T	T	T	T	T
4	T	F	F	T	F	T	F
5	F	T	T	T	T	T	T
6	F	T	F	T	T	F	F
7	F	F	T	F	T	T	T
8	F	F	F	F	T	T	F
				1	2	3	*

Мы снова приходим к выводу, что, когда все посылки истинны (что имеет место в случаях 1, 3 и 5), заключение истинно, и умозаключение правильно.

Однако, если рассмотреть следующее умозаключение

$$\begin{array}{l} p \rightarrow q \\ q \rightarrow r \\ r \\ \hline \therefore\ p \end{array}$$

и посмотреть на его таблицу истинности,

Случай	p	q	r	$p \rightarrow q$	$q \rightarrow r$	r	p
1	T	T	T	T	T	T	T
2	T	T	F	T	F	F	T
3	T	F	T	F	T	T	T
4	T	F	F	F	T	F	T
5	F	T	T	T	T	T	F
6	F	T	F	T	F	F	F
7	F	F	T	T	T	T	F
8	F	F	F	T	T	F	F
				1	2	3	*

мы обнаружим, что хотя в случае 1 истинны и посылки, и заключение, в случаях 5 и 7 посылки истинны, а заключение ложно. Следовательно, умозаключение не является правильным.

Рассмотрим метод проверки правильности умозаключений, альтернативный методу таблиц истинности. Пока предлагается принять метод на веру, вернувшись к его обоснованию несколько позже. Это не прямой метод, поскольку он направлен на доказательство неправильности умозаключения. В случае успеха такого доказательства это будет свидетельством неправильности умозаключения. Если, предполагая неправильность умозаключения, мы приходим к противоречию, то умозаключение является правильным. Например, рассмотрим умозаключение

$$\begin{array}{l} p \vee q \\ p \rightarrow r \\ q \rightarrow r \\ \hline \therefore\ r \end{array}$$

Если умозаключение неправильное, существуют истинностные значения p, q и r, для которых посылки истинны, а заключение ложно. Если заключение ложно, то r ложно. Если $q \rightarrow r$ истинно , а r ложно, то q должно быть ложно. Аналогично, если $p \rightarrow r$ истинно, тогда p должно быть ложно. Но тогда $p \vee q$ ложно, что приводит к противоречию с утверждением, что заключение ложно, а посылки истинны. На основании этого делаем вывод, что умозаключение правильно.

Рассмотрим умозаключение

$$\begin{array}{l} p \to q \\ q \to r \\ r \to s \\ s \to t \\ \hline \therefore\ p \to t \end{array}$$

Снова попытаемся свести рассмотрение к случаю, когда заключение ложно, а посылки истинны. Если $p \to t$ ложно, то p должно быть истинно и t ложно. Поскольку t ложно, то из истинности $s \to t$ следует, что s ложно. Если s ложно и $r \to s$ истинно, то r ложно. Если r ложно и $q \to r$ истинно, то q должно быть ложно. Но поскольку p истинно, а q ложно, то $p \to q$ ложно, что невозможно. Опять приходим к противоречию с тем, что заключение ложно, а все посылки истинны. Таким образом, умозаключение правильно.

Рассмотрим умозаключение

$$\begin{array}{l} p \to q \\ q \to r \\ \hline \therefore\ p \vee r \end{array}$$

Если заключение ложно, то и p и r — оба ложны. Однако, если q ложно, то $p \to q$ и $q \to r$ — оба истинны. Поскольку обе посылки истинны, а заключение ложно, то умозаключение неправильно. Если посмотреть на таблицу истинности данного умозаключения,

Случай	p	q	r	$p \to q$	$q \to r$	$p \vee r$
1	T	T	T	T	T	T
2	T	T	F	T	F	T
3	T	F	T	F	T	T
4	T	F	F	F	T	T
5	F	T	T	T	T	F
6	F	T	F	T	F	F
7	F	F	T	T	T	T
8	F	F	F	T	T	F
				1	2	*

легко заметить, что рассмотренный случай соответствует строке 8 таблицы истинности, откуда следует неправильность умозаключения.

Рассмотрим умозаключение

$$\begin{array}{l} p \to q \\ q \to r \\ r \\ \hline \therefore\ p \end{array}$$

Если заключение ложно, это означает, что p ложно. Но если посылки истинны, то r истинно. В этом случае $p \to q$ и $q \to r$ истинны вне зависимости от того,

истинно q или ложно. Таким образом, имеется два случая, когда можно показать неправильность нашего умозаключения: (1) p ложно, q истинно, r истинно и (2) p ложно, q ложно, r истинно. В этих случаях посылки истинны и заключение ложно. Из таблицы истинности

Случай	p	q	r	$p \to q$	$q \to r$	r	p
1	T	T	T	T	T	T	T
2	T	T	F	T	F	F	T
3	T	F	T	F	T	T	T
4	T	F	F	F	T	F	T
5	F	T	T	T	T	T	F
6	F	T	F	T	F	F	F
7	F	F	T	T	T	T	F
8	F	F	F	T	T	F	F
				1	2	3	*

легко видеть, что это случаи 5 и 7.

Следует обратить внимание, что любое умозаключение с посылками H_1, H_2, H_3, ..., H_n и заключением C является правильным тогда и только тогда, когда высказывание

$$(H_1 \wedge H_2 \wedge H_3 \wedge \ldots \wedge H_n) \to C$$

есть тавтология. Заметим, что порядок следования посылок не является существенным, поскольку

$$H_1 \wedge H_2 \equiv H_2 \wedge H_1.$$

Примем в качестве правила вывода следующее умозаключение, в правильности которого легко убедиться.

$$\begin{array}{l} p \\ \underline{p \to q} \\ \therefore\ q \end{array}$$

Это правильное умозаключение называется ***правилом отделения*** (***modus ponens***).

Рассмотрим пример использования правила отделения. Предположим, b — целое число. Пусть p и q заданы следующим образом

$$\begin{array}{ll} p: & b \text{ четно} \\ q: & b \text{ делится на } 2, \end{array}$$

так что

$$p \to q: \quad \text{если } b \text{ четно, то } b \text{ делится на } 2.$$

Правило отделения дает

$$\begin{array}{ll} p \to q & \text{если } b \text{ четное, то } b \text{ делится на } 2 \\ \underline{p\qquad\quad} & \underline{b \text{ четное}\qquad\qquad\qquad\qquad} \\ \therefore\ q & \therefore\ b \text{ делится на } 2 \end{array}$$

Допустим, что высказывание *если b четно, то b делится на* 2 получено как свойство целых чисел и $b = 12$. Тогда обе посылки истинны, так что нет сомнения в том, что 12 делится на 2. С другой стороны, если $b = 13$, тогда p ложно, и хотя умозаключение правильно, нельзя утверждать что-либо о делимости $b = 13$ на 2. Если одна из посылок ложна, то истинность заключения никоим образом не зависит от правильности умозаключения.

Перечислим некоторые правила вывода, на которые мы будем ссылаться в дальнейшем:

а) *Правило отделения (Modus Ponens)*

$$\begin{array}{l} p \to q \\ \underline{p} \\ \therefore\ q \end{array}$$

б) *Силлогизм*

$$\begin{array}{l} p \to q \\ \underline{q \to r} \\ \therefore\ p \to r \end{array}$$

в) *Modus Tollens*

$$\begin{array}{l} p \to q \\ \underline{\sim q} \\ \therefore\ \sim p \end{array}$$

г) *Расширение*

$$\begin{array}{l} p \\ \hline \therefore\ p \vee q \end{array}$$

д) *Специализация*

$$\begin{array}{l} p \wedge q \\ \hline \therefore\ p \end{array}$$

е) *Конъюнкция*

$$\begin{array}{l} p \\ \underline{q} \\ \therefore\ p \wedge q \end{array}$$

ж) *Выбор*

$$\begin{array}{l} p \\ p \to (r \vee s) \\ r \to q \\ \underline{s \to q} \\ \therefore\ q \end{array}$$

з) *Исключающий выбор*

$$\begin{array}{l} p \vee q \\ \underline{p \to (r \wedge \sim r)} \\ \therefore\ q \end{array}$$

и) *Сведение к абсурду (Reductio ad Absurdum)*

$$\begin{array}{l} \sim p \rightarrow (r \wedge \sim r) \\ \therefore \underline{\quad p \quad} \end{array}$$

Правильность всех перечисленных умозаключений можно показать с помощью таблиц истинности.

Сведение к абсурду используется в методе доказательства, известном как ***доказательство от противного***. Он состоит в следующем: мы допускаем, что истинным является отрицание того высказывания, которое необходимо доказать; затем пытаемся прийти к противоречию. Если это удается, исходное утверждение доказано. Обратите внимание, именно такое рассмотрение было основой ранее упомянутого метода доказательства правильности умозаключения, альтернативного использованию таблиц истинности. Таким образом, этот альтернативный метод рассуждения является правильным.

Рассмотрим такое умозаключение:

Если яблоко красное, то оно спелое
Яблоко спелое
$\therefore$ *Яблоко красное*

В обозначениях

p : *яблоко красное*
q : *яблоко спелое*

умозаключение принимает вид

$$\begin{array}{l} p \rightarrow q \\ \underline{\quad q \quad} \\ \therefore p \end{array}$$

Из таблицы истинности

Случай	p	q	$((p \rightarrow q)$	$\wedge$	$q)$	$\rightarrow$	p
1	T	T	T	T	T	T	T
2	T	F	F	F	F	T	T
3	F	T	T	T	T	F	F
4	F	F	T	F	F	T	F
						*	

очевидным образом следует неправильность умозаключения. В случае 3 обе посылки $p \rightarrow q$ и q истинны; однако, заключение p таковым не является. Таким образом, умозаключение неправильно. Данный вид неправильного умозаключение называется ***ложной конверсией***. Точно так же умозаключение

$$\begin{array}{l} \sim p \rightarrow \sim q \\ \underline{\quad p \quad} \\ \therefore q \end{array}$$

является неправильным и называется ***ложной инверсией***.

В большинстве аксиоматических систем теоремы и аксиомы могут быть весьма сложными, а их переплетение с правилами вывода — весьма запутанным. Сейчас мы попытаемся объяснить, что происходит в процессе доказательства и как доказательства конструируются. Сложные доказательства обычно представляют собой длинную цепочку правильных умозаключений вышеуказанных типов. Поэтому доказательство можно определить как последовательность утверждений, каждое из которых истинно в силу одной из следующих причин:

а) по предположению;
б) по аксиоме или определению;
в) по ранее доказанной теореме или лемме;
г) выведено из предыдущих утверждений;
д) логически эквивалентно предыдущему утверждению.

Используя логическую символику, в качестве примера покажем, что

$$\begin{array}{l} p \rightarrow q \\ \sim r \rightarrow \sim q \\ \sim r \\ \hline \therefore \sim p \end{array}$$

есть правильное умозаключение.

Номер	Утверждение	Как получено
1	$p \rightarrow q$	предположение
2	$\sim r \rightarrow \sim q$	предположение
3	$\sim r$	предположение
4	$\sim q$	2, 3 и правило отделения
5	$\sim q \rightarrow \sim p$	1 и эквивалентность $p \rightarrow q \equiv \sim q \rightarrow \sim p$
6	$\sim p$	4, 5 и правило отделения

Итак, из трех посылок следует заключение $\sim p$, и мы доказали $\sim p$.

В большинстве математических доказательств логика "скрыта" в том смысле, что о ней специально не упоминается. Предполагается, что читатель может отслеживать логику без посторонней помощи, а ее рассмотрение неизбежно усложнило бы сам процесс доказательства. В некоторых доказательствах, которые последуют ниже, мы заглянем за кулисы с тем, чтобы проиллюстрировать использование логики.

■ УПРАЖНЕНИЯ

1. Используя таблицы истинности, докажите правильность следующих умозаключений.

а) Правило силлогизма

$$\begin{array}{l} p \rightarrow q \\ \underline{q \rightarrow r} \\ \therefore p \rightarrow r \end{array}$$

б) Правило выбора

$$\begin{array}{l} p \\ p \rightarrow (r \vee s) \\ r \rightarrow q \\ s \rightarrow q \\ \hline \therefore\ q \end{array}$$

в) Правило сведения к абсурду (reductio ad absurdum)

$$\begin{array}{l} \sim w \rightarrow (r \wedge \sim r) \\ \hline \therefore\ w \end{array}$$

2. Покажите, что следующее умозаключение не является правильными:

$$\begin{array}{l} p \rightarrow q \\ q \rightarrow r \\ \hline \therefore\ r \end{array}$$

3. Определите, какое из следующих умозаключений является правильным:

а)
$$\begin{array}{l} p \rightarrow q \\ p \rightarrow r \\ q \vee r \\ \hline \therefore\ p \end{array}$$

б)
$$\begin{array}{l} \sim p \vee q \\ \sim q \vee r \\ \sim r \\ \hline \therefore\ \sim p \end{array}$$

в)
$$\begin{array}{l} p \rightarrow q \\ p \rightarrow r \\ \sim(p \wedge q) \\ \hline \therefore\ \sim p \end{array}$$

г)
$$\begin{array}{l} \sim p \vee \sim q \\ r \vee \sim q \\ \sim p \\ \hline \therefore\ r \vee \sim p \end{array}$$

4. Определите, какое из следующих умозаключений является правильным:

а)
$$\begin{array}{l} \sim r \\ p \rightarrow r \\ q \rightarrow r \\ \hline \therefore\ \sim(p \wedge q) \end{array}$$

б)
$$\begin{array}{l} p \vee q \\ \sim q \vee r \\ \sim r \\ \hline \therefore\ p \end{array}$$

в)
$$\begin{array}{l} p \rightarrow q \\ q \rightarrow r \\ q \vee r \\ \hline \therefore\ p \end{array}$$

г)
$$\begin{array}{l} p \vee \sim q \\ r \vee \sim q \\ \sim p \\ \hline \therefore\ \sim r \end{array}$$

5. Установите правильность следующих умозаключений, используя метод, альтернативный применению таблиц истинности:

а)
$$\begin{array}{l} s \vee t \\ t \rightarrow r \\ s \rightarrow w \\ \hline \therefore\ r \vee w \end{array}$$

б)
$$\begin{array}{l} p \rightarrow q \\ q \rightarrow r \\ r \\ \hline \therefore\ p \end{array}$$

в)
$$\begin{array}{l} p \rightarrow q \\ \sim q \rightarrow\ \sim s \\ s \rightarrow t \\ t \vee q \\ \hline \therefore\ p \vee s \end{array}$$

г)
$$\begin{array}{l} \sim p \vee q \\ r \vee \sim q \\ p \\ s \\ \hline \therefore\ r \vee s \end{array}$$

6. Установите правильность следующих умозаключений, используя метод, аль-

тернативный методу таблиц истинности:

а)
$$\begin{array}{l} s \vee t \\ t \to r \\ s \to w \\ \hline \therefore\ r \vee w \end{array}$$

б)
$$\begin{array}{l} p \to q \\ q \to r \\ r \\ \hline \therefore\ p \end{array}$$

в)
$$\begin{array}{l} p \to q \\ \sim q \to \sim s \\ s \to t \\ t \vee q \\ \hline \therefore\ p \vee s \end{array}$$

г)
$$\begin{array}{l} \sim p \vee q \\ r \vee \sim q \\ p \\ s \\ \hline \therefore\ r \vee s \end{array}$$

7. Покажите правильность следующих умозаключений, используя правила вывода и эквивалентные высказывания:

а)
$$\begin{array}{l} \sim(s \wedge t) \\ \sim w \to t \\ \hline \therefore\ s \to w \end{array}$$

б)
$$\begin{array}{l} p \to q \\ \sim r \to \sim q \\ \sim r \\ \hline \therefore\ \sim p \end{array}$$

в)
$$\begin{array}{l} \sim(\sim p \vee q) \\ \sim z \to \sim s \\ (p \wedge \sim q) \to s \\ \sim z \vee r \\ \hline \therefore\ r \end{array}$$

г)
$$\begin{array}{l} \sim x \to \sim w \\ (x \vee \sim w) \to z \\ \sim p \to \sim z \\ p \to (\sim r \vee \sim s) \\ \hline \therefore\ \sim r \vee \sim s \end{array}$$

8. Докажите, что логическая эквивалентность двух высказываний означает, что эквиваленция двух высказываний есть тавтология. Иными словами, для высказываний p и q докажите, что

$$p \equiv q$$

означает то же самое, что и

$$p \leftrightarrow q \text{ — } \textit{тавтология.}$$

Вслед за этим докажите, что последнее утверждение эквивалентно утверждению:

$$\textit{как } p \to q\textit{, так и } q \to p \textit{ являются тавтологиями.}$$

В случае, когда импликация $p \to q$ есть тавтология, говорят, что p имеет следствием q, и часто записывают это как $p \Rightarrow q$.

9. Найдите среди указанных ниже выражений тавтологии:

а) $p \to (p \wedge q)$;

б) $p \to (p \vee q)$;

в) $(p \to q) \to (q \to p)$;

г) $((p \to q) \wedge \sim q) \to \sim p)$;

д) $((p \to q) \wedge q) \to p)$.

1.5. ПОЛНОТА В ЛОГИКЕ ВЫСКАЗЫВАНИЙ

Кроме использования в логике, одно из важных применений таблиц истинности состоит в конструировании коммутационных схем. Прежде, чем приступить к изучению коммутационных схем, мы рассмотрим вопрос о минимальном количестве логических связок, необходимых для выражения любого высказывания, образованного с помощью определенных нами выше логических связок. Известно, что $p \leftrightarrow q$ можно выразить как $(p \to q) \wedge (q \to p)$, так что использовать $\leftrightarrow$ удобно, но не необходимо. К тому же $p \veebar q$ эквивалентно

$$(p \wedge \sim q) \vee (\sim p \wedge q)\,.$$

Также $p \to q$ эквивалентно $\sim p \vee q$, поэтому нет необходимости использовать $\to$, если применяется $\sim$ и $\vee$. Кроме того, $p \wedge q$ эквивалентно $\sim(\sim p \vee \sim q)$ и $p \vee q$ эквивалентно $\sim(\sim p \wedge \sim q)$. Следовательно, любое высказывание может быть выражено через пару связок $\sim$ и $\wedge$ или $\sim$ и $\vee$, причем в любом случае необходимы обе связки. Однако существуют две связки, обладающие тем свойством, что любое высказывание может быть выражено с использованием только одной из них. Эти связки: $|$ — так называемый ***штрих Шеффера*** и $\downarrow$ — так называемая ***стрелка Пирса***. Свои названия эти связки получили в честь математиков Г.Шеффера и Ч.Пирса. Этим связкам соответствуют таблицы истинности

Случай	p	q	$p \mid q$
1	T	T	F
2	T	F	T
3	F	T	T
4	F	F	T

Случай	p	q	$p \downarrow q$
1	T	T	F
2	T	F	F
3	F	T	F
4	F	F	T

Для того, чтобы показать, что любую связку можно заменить связкой $|$, достаточно показать это для пар связок $\sim$ и $\wedge$ или $\sim$ и $\vee$, поскольку возможность выразить любую связку одной из этих пар уже показана. Эквивалентность $p \mid p$ и $\sim p$ устанавливается при помощи следующей таблицы истинности:

Случай	p	q	p	$\mid$	p
1	T	T	T	F	T
2	T	F	T	F	T
3	F	T	F	T	F
4	F	F	F	T	F
				*	

Точно так же таблица

Случай	p	q	$(p$	$\mid$	$p)$	$\mid$	$(q$	$\mid$	$q)$
1	T	T	T	F	T	T	T	F	T
2	T	F	T	F	T	T	F	T	F
3	F	T	F	T	F	T	T	F	T
4	F	F	F	T	F	F	F	T	F
						*			

показывает, что $(p \mid p) \mid (q \mid q)$ эквивалентно $p \vee q$. Можно также показать, что $(p \mid q) \mid (p \mid q)$ эквивалентно $p \wedge q$.

Таким образом, если показать, что $\sim$ и $\wedge$ или $\sim$ и $\vee$ можно выразить, используя только $\downarrow$, тогда и любую связку можно выразить, используя лишь $\downarrow$. Предоставляем читателю показать, что $p \downarrow p$ эквивалентно $\sim p$, $(p \downarrow p) \downarrow (q \downarrow q)$ эквивалентно $p \wedge q$, а $(p \downarrow q) \downarrow (p \downarrow q)$ эквивалентно $p \vee q$. Заметим, что $p \mid q$ эквивалентно $\sim(p \wedge q)$, а $p \downarrow q$ эквивалентно $\sim(p \vee q)$. Поэтому в дальнейшем связка $\mid$ будет называться ***не-и***, а связка $\downarrow$ будет называться ***не-или***.

Допустим, что задана произвольная таблица истинности. Существует простой способ найти высказывание, которому она соответствует. Например, предположим, что имеется таблица истинности

Случай	p	q	
1	T	T	T
2	T	F	T
3	F	T	F
4	F	F	T

Известно, что $p \wedge q$ истинно в случае 1 и ложно во всех остальных случаях. Аналогично, $p \wedge \sim q$ истинно только в случае 2, $\sim p \wedge q$ истинно только в случае 3, а $\sim p \wedge \sim q$ истинно только в случае 4. Пусть высказывание должно быть истинным точно в указанных нами случаях. Если для каждого такого случая (строки таблицы) выбрать высказывание, которое истинно только в этом случае, и связать эти высказывания связкой $\vee$, то мы получим высказывание, истинное только в требуемых случаях (строках). В приведенном выше примере рассматриваемая таблица истинности соответствует высказыванию

$$(p \wedge q) \vee (p \wedge \sim q) \vee (\sim p \wedge \sim q)\,.$$

В случае таблиц истинности с тремя переменными имеем аналогичную ситуацию. Для каждой строки следующей таблицы приведено высказывание, истинное только для этой строки.

Случай	p	q	r	
1	T	T	T	$p \wedge q \wedge r$
2	T	T	F	$p \wedge q \wedge \sim r$
3	T	F	T	$p \wedge \sim q \wedge r$
4	T	F	F	$p \wedge \sim q \wedge \sim r$
5	F	T	T	$\sim p \wedge q \wedge r$
6	F	T	F	$\sim p \wedge q \wedge \sim r$
7	F	F	T	$\sim p \wedge \sim q \wedge r$
8	F	F	F	$\sim p \wedge \sim q \wedge \sim r$

Заметьте, что если в какой-либо строке переменная имеет ложное значение, то в соответствующем высказывании она использована с отрицанием. Если требуется построить высказывание, соответствующее конкретной таблице истинности,

необходимо выбрать выражения, соответствующие случаям (строкам), где высказывание истинно, и соединить их связкой $\vee$. Например, построение высказывания с таблицей истинности

Случай	p	q	r	
1	T	T	T	T
2	T	T	F	T
3	T	F	T	F
4	T	F	F	F
5	F	T	T	T
6	F	T	F	F
7	F	F	T	F
8	F	F	F	T

дает высказывание

$$(p \wedge q \wedge r) \vee (p \wedge q \wedge \sim r) \vee (\sim p \wedge q \wedge r) \vee (\sim p \wedge \sim q \wedge \sim r).$$

Построение высказывания с таблицей

Случай	p	q	r	
1	T	T	T	F
2	T	T	F	T
3	T	F	T	T
4	T	F	F	F
5	F	T	T	F
6	F	T	F	T
7	F	F	T	F
8	F	F	F	F

дает высказывание

$$(p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge r) \vee (\sim p \wedge q \wedge \sim r).$$

Такая форма выражения высказывания называется ***дизъюнктивной нормальной формой***. Выражения $p \wedge q \wedge \sim r$, $p \wedge \sim q \wedge r$ и $\sim p \wedge q \wedge \sim r$ называются ***элементарными конъюнкциями***. Дадим более точное определение.

ОПРЕДЕЛЕНИЕ 1.5. Пусть $p_1, p_2, p_3, \ldots, p_n$ — простые высказывания. Назовем выражение $x_1 \wedge x_2 \wedge x_3 \wedge \cdots \wedge x_n$, в котором $x_i = p_i$ или $x_i = \sim p_i$, ***элементарной конъюнкцией***. Выражение, представляющее собой дизъюнкцию элементарных конъюнкций, называется ***дизъюнктивной нормальной формой***, так что если $m_1, m_2, m_3, \ldots, m_n$ есть элементарные конъюнкции, тогда $m_1 \vee m_2 \vee m_3 \vee \cdots \vee m_n$ есть дизъюнктивная нормальная форма.

Хотя любое высказывание может быть выражено в дизъюнктивной нормальной форме, эта форма высказывания не является простейшей. Карты Карно, которым посвящен следующий раздел, как раз позволяют упростить выражение высказывания в дизъюнктивной нормальной форме.

Действуя по той же схеме, мы замечаем, что $p \vee q \vee r$ ложно, только когда p, q и r ложны. Вообще, в таблице

Случай	p	q	r	
1	T	T	T	$\sim p \vee \sim q \vee \sim r$
2	T	T	F	$\sim p \vee \sim q \vee r$
3	T	F	T	$\sim p \vee q \vee \sim r$
4	T	F	F	$\sim p \vee q \vee r$
5	F	T	T	$p \vee \sim q \vee \sim r$
6	F	T	F	$p \vee \sim q \vee r$
7	F	F	T	$p \vee q \vee \sim r$
8	F	F	F	$p \vee q \vee r$

каждое выражение ложно в строке, где оно расположено, и истинно в любой другой строке. Если требуется найти высказывание, обладающее данной таблицей истинности, зная все случаи (строки), где в таблице истинности стоит ложное значение, то используются высказывания, каждое из которых ложно только в соответствующей строке, и все эти высказывания объединяются связкой $\wedge$. Например, таблице истинности вида

Случай	p	q	r	
1	T	T	T	T
2	T	T	F	T
3	T	F	T	F
4	T	F	F	F
5	F	T	T	T
6	F	T	F	F
7	F	F	T	F
8	F	F	F	T

соответствует высказывание

$$(\sim p \vee q \vee \sim r) \wedge (\sim p \vee q \vee r) \wedge (p \vee \sim q \vee r) \wedge (p \vee q \vee \sim r)\,.$$

Такая форма выражения высказывания называется ***конъюнктивной нормальной формой***. Выражения $\sim p \vee q \vee \sim r$, $\sim p \vee q \vee r$, $p \vee \sim q \vee r$ и $p \vee q \vee \sim r$ носят название ***элементарных дизъюнкций***. Теперь дадим более формальное определение.

ОПРЕДЕЛЕНИЕ 1.6. Пусть p_1, p_2, p_3, ..., p_n — простые высказывания. Назовем выражение $x_1 \vee x_2 \vee x_3 \vee \ldots \vee x_n$,в котором $x_i = p_i$ или $\sim p_i$, ***элементарной дизъюнкцией***.

Выражение, представляющее собой конъюнкцию элементарных дизъюнкций, называется ***конъюнктивной нормальной формой***, так что если m_1, m_2, m_3, ..., m_n — элементарные дизъюнкции, то $m_1 \wedge m_2 \wedge m_3 \wedge \ldots \wedge m_n$ есть конъюнктивная нормальная форма.

■ УПРАЖНЕНИЯ

1. Используя таблицы истинности, докажите, что

а) $p \downarrow p$ эквивалентно $\sim p$;

б) $(p \downarrow p) \downarrow (q \downarrow q)$ эквивалентно $p \wedge q$;

в) $(p \downarrow q) \downarrow (p \downarrow q)$ эквивалентно $(p \vee q)$.

2. Найдите высказывания, которым отвечают следующие таблицы истинности:

а)

Случай	p	q	r	
1	T	T	T	T
2	T	T	F	T
3	T	F	T	F
4	T	F	F	F
5	F	T	T	F
6	F	T	F	T
7	F	F	T	F
8	F	F	F	T

б)

Случай	p	q	r	
1	T	T	T	T
2	T	T	F	F
3	T	F	T	T
4	T	F	F	F
5	F	T	T	F
6	F	T	F	T
7	F	F	T	F
8	F	F	F	T

в)

Случай	p	q	r	
1	T	T	T	T
2	T	T	F	F
3	T	F	T	T
4	T	F	F	T
5	F	T	T	F
6	F	T	F	F
7	F	F	T	F
8	F	F	F	T

3. Найдите высказывания, которым отвечают следующие таблицы истинности:

а)

Случай	p	q	r	
1	T	T	T	F
2	T	T	F	T
3	T	F	T	T
4	T	F	F	T
5	F	T	T	F
6	F	T	F	F
7	F	F	T	F
8	F	F	F	T

б)

Случай	p	q	r	
1	T	T	T	F
2	T	T	F	F
3	T	F	T	F
4	T	F	F	F
5	F	T	T	T
6	F	T	F	T
7	F	F	T	F
8	F	F	F	T

в)

Случай	p	q	r	
1	T	T	T	T
2	T	T	F	F
3	T	F	T	T
4	T	F	F	T
5	F	T	T	F
6	F	T	F	F
7	F	F	T	T
8	F	F	F	F

1.6. КАРТЫ КАРНО

Для простых высказываний p_1, p_2, p_3, ... и p_n существует 2^n различных элементарных конъюнкций. (Это будет показано в главе 8.) Например, для высказываний p и q элементарными конъюнкциями будут $p \wedge q$, $p \wedge \sim q$, $\sim p \wedge q$ и $\sim p \wedge \sim q$. ***Карта Карно*** — это таблица, каждый элемент которой является элементарной конъюнкцией. Например, для высказываний p и q карта Карно должна иметь вид, изображенный на рис. 1.1, а на рис. 1.2 внутри прямоугольников представлены соответствующие элементарные конъюнкции.

	q	$\sim q$
p		
$\sim p$		

Рис. 1.1

	q	$\sim q$
p	$p \wedge q$	$p \wedge \sim q$
$\sim p$	$\sim p \wedge q$	$\sim p \wedge \sim q$

Рис. 1.2

Для представления картой Карно высказывания, записанного в дизъюнктивной нормальной форме, необходимо поместить $\times$ в прямоугольниках, соответствующих элементарным конъюнкциям. Например, высказыванию

$$(p \wedge q) \vee (\sim p \wedge \sim q)$$

соответствует карта Карно, изображенная на рис. 1.3.

	q	$\sim q$
p	$\times$	
$\sim p$		$\times$

Рис. 1.3

	q	$\sim q$
p	$\times$	$\times$
$\sim p$		

Рис. 1.4

Заметим, что если высказыванию соответствует карта Карно с двумя соседствующими в строке или в столбце $\times$, тогда выражение можно упростить, сведя

две элементарные конъюнкции к одной, содержащей на одну компоненту меньше (т.е. либо p, либо q не будут присутствовать в выражении). Например, высказывание $(p \wedge q) \vee (p \wedge \sim q)$, которому соответствует карта Карно, изображенная на рис. 1.4, эквивалентно высказыванию p, так как $(p \wedge q) \vee (p \wedge \sim q) \equiv p \wedge (q \vee \sim q) \equiv p \wedge T \equiv p$.

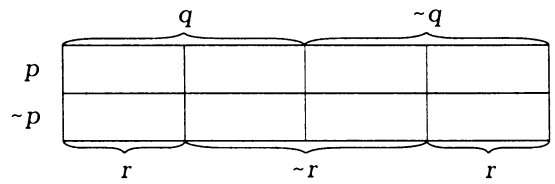


Рис. 1.5

Карта Карно для p, q и r может иметь вид, изображенный на рис. 1.5, а на рис. 1.6 в прямоугольниках внутри представлены элементарные конъюнкции.

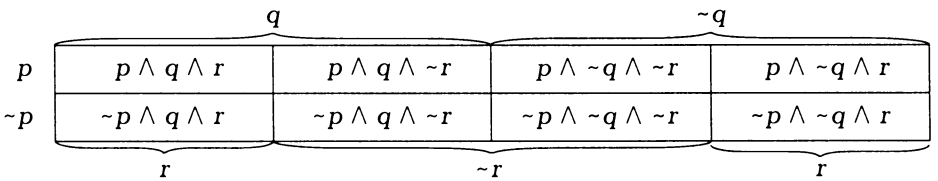


Рис. 1.6

Следовательно, высказыванию

$$(p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge r) \vee (\sim p \wedge q \wedge \sim r)$$

будет соответствовать карта Карно, изображенная на рис. 1.7.

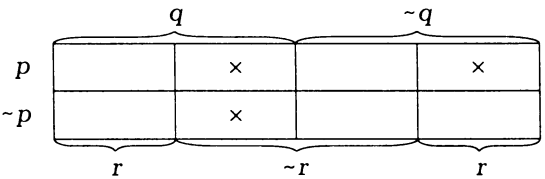


Рис. 1.7

Как и прежде, поскольку два знака соседствуют, две элементарные конъюнкции могут быть сведены к одной, содержащей на одну из компонент p, q и r меньше. В данном случае $(p \wedge q \wedge \sim r) \vee (\sim p \wedge q \wedge \sim r)$ сводится к $(q \wedge \sim r)$, так что выражение принимает вид $(q \wedge \sim r) \vee (p \wedge \sim q \wedge r)$.

В случае, когда четыре значка расположены в прямоугольнике рядом, как показано на рис. 1.8

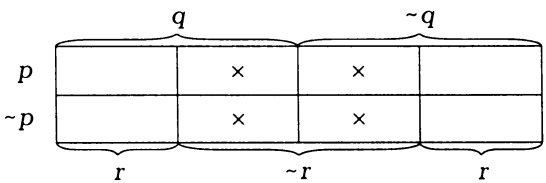


Рис. 1.8

или на рис. 1.9,

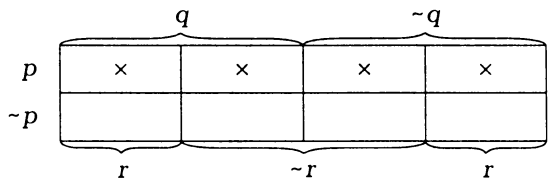


Рис. 1.9

тогда четыре элементарные конъюнкции, отмеченные значками, могут быть сведены к одному члену, содержащему только одну из компонент p, q и r. Например, первая карта Карно представляет

$$(p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge r) \vee (\sim p \wedge q \wedge \sim r) \vee (\sim p \wedge \sim q \wedge \sim r),$$

что сводится к $\sim r$.

Вторая карта Карно представляет выражение

$$(p \wedge q \wedge r) \vee (p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge \sim r) \vee (p \wedge \sim q \wedge r),$$

что может быть сведено к p.

Поскольку r появляется на обоих концах карты Карно, ее можно "скрутить" и считать, что четыре значка на карте Карно образуют прямоугольник из четырех значков (см. рис. 1.10), поэтому выражение

$$(p \wedge q \wedge r) \vee (\sim p \wedge q \wedge r) \vee (p \wedge \sim q \wedge r) \vee (\sim p \wedge \sim q \wedge r)$$

сводится просто к r.

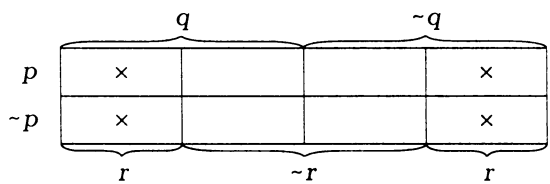


Рис. 1.10

Выражение

$$(p \wedge q \wedge r) \vee (\sim p \wedge q \wedge r) \vee (p \wedge q \wedge \sim r) \vee (\sim p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge \sim r),$$

представленное на рис. 1.11, может быть преобразовано к $q \vee (p \wedge \sim q \wedge \sim r)$ на основе правостороннего блока из четырех значков. Более того, благодаря наличию блока из двух значков в середине первой строки, его можно еще больше упростить, приведя к виду $q \vee (p \wedge \sim r)$.

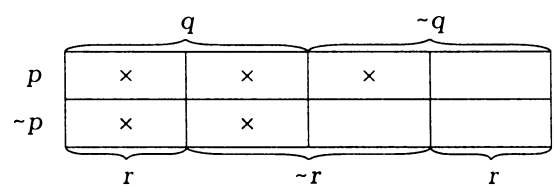


Рис. 1.11

Перечислим четыре последовательных шага при использовании карты Карно:

1. Для каждой элементарной конъюнкции обозначьте на карте соответствующий прямоугольник.
2. Покройте знаки, используя, по возможности, несколько прямоугольных блоков.
3. Используйте блоки, максимальные по величине, не меняя числа блоков.
4. Оцените блоки с соответствующими высказываниями, которые описывают их, и объедините эти высказывания символом $\vee$.

Проиллюстрируем этот процесс путем построения карты Карно для четырех высказываний p, q, r и s, изображенной на рис. 1.12, где во внутренних прямоугольниках представлены элементарные конъюнкции, составленные из тех высказываний, расположенных на краях, которые принадлежат строке и столбцу соответствующих прямоугольников. Например, крайние высказывания, соответствующие второй строке и третьему столбцу, есть p, $\sim q$, $\sim r$ и $\sim s$, что порождает элементарную конъюнкцию $p \wedge \sim q \wedge \sim r \wedge \sim s$.

Рассмотрим выражение

$$(p \wedge q \wedge r \wedge s) \vee (p \wedge q \wedge \sim r \wedge s) \vee (p \wedge q \wedge r \wedge \sim s) \vee$$
$$\vee (p \wedge q \wedge \sim r \wedge \sim s) \vee (\sim p \wedge q \wedge r \wedge \sim s) \vee (\sim p \wedge q \wedge \sim r \wedge \sim s) \vee$$
$$\vee (\sim p \wedge q \wedge r \wedge s) \vee (\sim p \wedge q \wedge \sim r \wedge s) \vee (p \wedge \sim q \wedge \sim r \wedge s) \vee$$
$$\vee (p \wedge \sim q \wedge r \wedge s) \vee (p \wedge \sim q \wedge r \wedge \sim s).$$

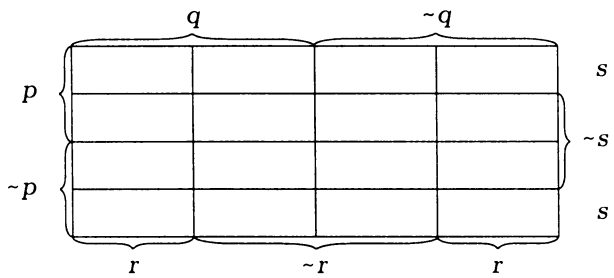


Рис. 1.12

Размещая соответствующие значки в карте Карно, получаем карту, изображенную на рис. 1.13, которую можно покрыть восьмикратным, четырехкратным и двукратным блоками. Восьмикратный блок можно описать, используя только одну из компонент p, q, r или s. В данном случае восьмикратный блок описывает q. Четырехкратный блок можно описать, используя только два простых высказывания. В данном случае этот блок можно описать, используя $p \wedge s$. Двукратный блок можно описать, используя три простых высказывания. В данном случае это $p \wedge \sim q \wedge r$. Следовательно, исходное высказывание можно упростить и привести к виду $q \vee (p \wedge s) \vee (p \wedge \sim q \wedge r)$. Поскольку r находится на обоих концах строки, а s находится на обоих концах столбца, "скручивание" карты Карно объединяет эти

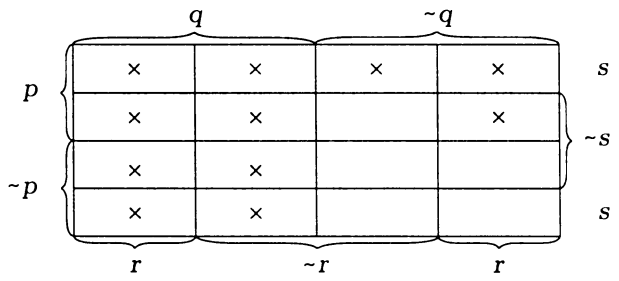


Рис. 1.13

случаи, так что верхний край можно считать прилегающим к нижнему, а левый — прилегающим к правому.

Например, четырехкратный блок на карте Карно на рис. 1.14 можно описать посредством $r \wedge s$.

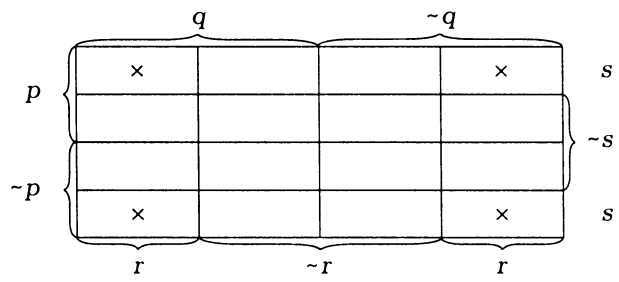
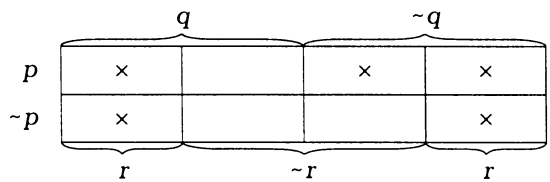


Рис. 1.14

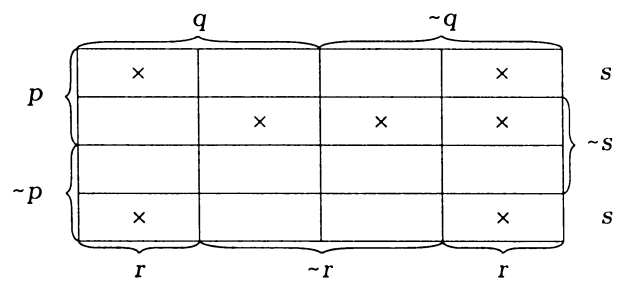
■ УПРАЖНЕНИЯ

1. Упростите высказывания, выраженные следующими картами Карно:

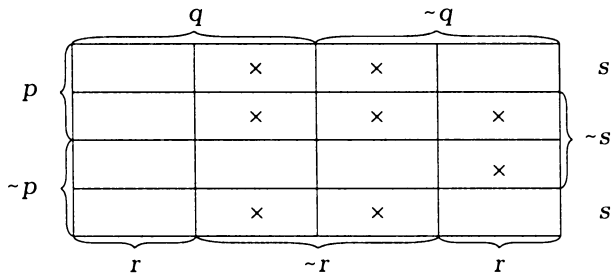
а)



б)

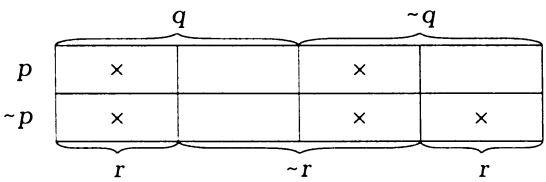


в)

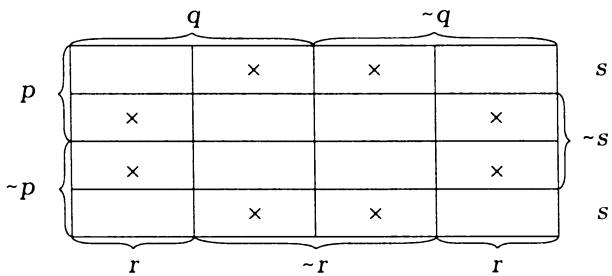


2. Упростите высказывания, выраженные следующими картами Карно:

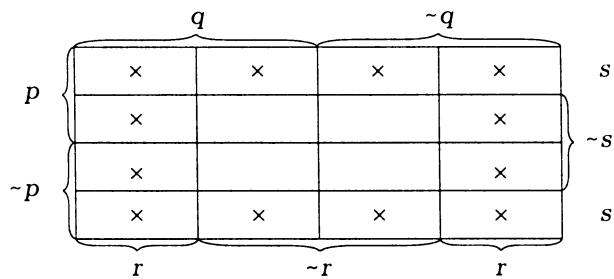
а)



б)



в)



3. Используйте карты Карно для упрощения следующих выражений:

а) $(p \wedge q \wedge r) \vee (\sim p \wedge q \wedge r) \vee (\sim p \wedge q \wedge \sim r) \vee$
$\vee (\sim p \wedge \sim q \wedge \sim r) \vee (p \wedge q \wedge \sim r);$

б) $(p \wedge q \wedge \sim r \wedge s) \vee (p \wedge \sim q \wedge \sim r \wedge s) \vee (p \wedge q \wedge \sim r \wedge \sim s) \vee$
$\vee (p \wedge \sim q \wedge \sim r \wedge \sim s) \vee (\sim p \wedge q \wedge \sim r \wedge \sim s) \vee$
$\vee (\sim p \wedge q \wedge \sim r \wedge s) \vee (\sim p \wedge \sim q \wedge \sim r \wedge s) \vee (\sim p \wedge q \wedge r \wedge \sim s) \vee$
$\vee (\sim p \wedge \sim q \wedge r \wedge \sim s);$

в) $(p \wedge q \wedge r \wedge s) \vee (\sim p \wedge q \wedge r \wedge s) \vee (p \wedge \sim q \wedge \sim r \wedge \sim s) \vee$
$\vee (p \wedge \sim q \wedge r \wedge \sim s) \vee (\sim p \wedge \sim q \wedge \sim r \wedge \sim s) \vee$
$\vee (\sim p \wedge \sim q \wedge r \wedge \sim s).$

4. Используйте карты Карно для упрощения следующих выражений:

а) $(p \wedge q \wedge r) \vee (\sim p \wedge q \wedge \sim r) \vee (\sim p \wedge q \wedge r) \vee$
$\vee (p \wedge \sim q \wedge \sim r) \vee (\sim p \wedge \sim q \wedge r);$

б) $(p \wedge q \wedge \sim r \wedge s) \vee (p \wedge \sim q \wedge \sim r \wedge s) \vee (\sim p \wedge q \wedge r \wedge s) \vee$
$\vee (\sim p \wedge q \wedge \sim r \wedge s) \vee (\sim p \wedge \sim q \wedge \sim r \wedge s) \vee (\sim p \wedge \sim q \wedge r \wedge s);$

в) $(p \wedge q \wedge r \wedge s) \vee (p \wedge q \wedge r \wedge \sim s) \vee (\sim p \wedge q \wedge r \wedge \sim s) \vee$
$\vee (p \wedge \sim q \wedge r \wedge s) \vee (p \wedge \sim q \wedge r \wedge \sim s) \vee (\sim p \wedge \sim q \wedge r \wedge \sim s).$

1.7. КОММУТАЦИОННЫЕ СХЕМЫ

Высказывания, соответствующие коммутационным (релейно-контактным) схемам, принято выражать в системе обозначений булевой алгебры, которая будет введена в разделе 2.4. Поэтому, прежде чем перейти к изучению коммутационных схем, мы перейдем от обозначений, принятых в логике, к булевой записи. Символы $\wedge$, $\vee$ и $\sim$ заменяются, соответственно, на $\cdot$, $+$ и $'$. Таким образом, $(p \wedge q) \vee \sim r$ превращается в $(p \cdot q) + r'$, а

$$(p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge r) \vee (\sim p \wedge q \wedge \sim r)$$

принимает вид

$$(p \cdot q \cdot r') + (p \cdot q' \cdot r) + (p' \cdot q \cdot r').$$

Как и в обычной алгебре, знак произведения, как правило, опускается, и предполагается, что произведение выполняется перед сложением, поэтому приведенное выше выражение можно переписать в виде

$$pqr' + pq'r + p'qr'.$$

В таблицах истинности T заменяется на 1, а F на 0, так что таблица истинности

Случай	p	q	r	p	$\vee$	$((\sim$	$q)$	$\wedge$	$r)$
1	T	T	T	T	T	F	T	F	T
2	T	T	F	T	T	F	T	F	F
3	T	F	T	T	T	T	F	T	T
4	T	F	F	T	T	T	F	F	F
5	F	T	T	F	F	F	T	F	T
6	F	T	F	F	F	F	T	F	F
7	F	F	T	F	T	T	F	T	T
8	F	F	F	F	F	T	F	F	F

преобразуется в таблицу

Случай	p	q	r	p	$+$	$(q'$	$\cdot$	$r)$
1	1	1	1	1	1	0	0	1
2	1	1	0	1	1	0	0	0
3	1	0	1	1	1	1	1	1
4	1	0	0	1	1	1	0	0
5	0	1	1	0	0	0	0	1
6	0	1	0	0	0	0	0	0
7	0	0	1	0	1	1	1	1
8	0	0	0	0	0	1	0	0

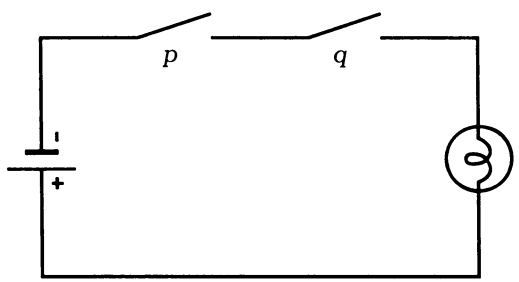


Рис. 1.15

В 1938 г. Клод Шеннон заметил связь между таблицами истинности и электрическими цепями. Рассмотрим схему переключения на рис. 1.15, которая состоит из источника питания (рис. 1.16) и электрической лампочки (рис. 1.17).

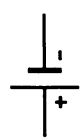


Рис. 1.16



Рис. 1.17

Присвоим значение 1 переключателям p и q, если они замкнуты (т.е. электрический ток проходит через них). В противоположной ситуации присвоим им

значение 0. Присвоим значение 1 схеме, когда лампочка светится (т.е. электрический ток через нее проходит). Заметим, что при последовательном соединении элементов цепи p и q, как это имеет место на приведенной выше схеме, лампочка загорается, и значение схемы становится равным 1 только в случае, когда оба переключателя замкнуты, т.е. и p, и q имеют значение 1. Таким образом, схема соответствует высказыванию $p \cdot q$. Такое расположение переключателей называется логическим элементом ***p и q***, или ***схемой логического умножения***. Этот логический элемент обозначается символом, изображенным на рис. 1.18.

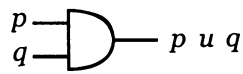


Рис. 1.18

Теперь рассмотрим схему переключения, показанную на рис. 1.19, где переключатели p и q соединены параллельно.

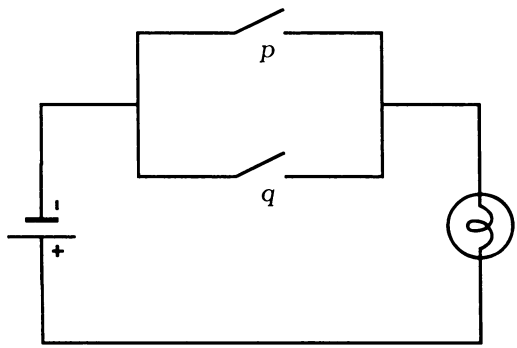


Рис. 1.19

Отметим, что теперь лампочка загорается, и значение схемы становится равным 1, когда один из двух переключателей p или q замкнут, т.е. либо значение $p = 1$, либо $q = 1$ (либо оба они равны 1). Эта схема соответствует высказыванию $p + q$. Такое расположение выключателей называется логическим элементом ***p или q***, или ***схемой логического сложения***. Этот логический элемент обозначается символом, изображенным на рис. 1.20.

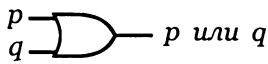


Рис. 1.20

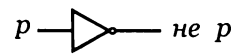


Рис. 1.21

Предположим, имеется схема (мы не будем пытаться ее изобразить), с одним переключателем p, который обладает таким свойством, что лампочка загорается тогда и только тогда, когда p разомкнут. Следовательно, схема имеет значение 1, когда p имеет значение 0, и имеет значение 0, когда p имеет значение 1. Эта схема соответствует p', а соответствующий логический элемент называется логическим элементом ***не***, или ***инвертором***. Логический элемент *не* обозначается символом, изображенным на рис. 1.21.

ПРИМЕР 1.7. Схема на рис. 1.22 содержит логический элемент p *и* q, за которым следует инвертор, так что схема соответствует выражению $(p \cdot q)'$. Заметим, что инвертор отрицает всю предшествующую ему схему.

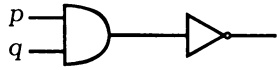


Рис. 1.22 □

ПРИМЕР 1.8. Схема на рис. 1.23 содержит соединение логического элемента p *или* q с логическим элементом *не* r посредством логической схемы умножения. Следовательно, она соответствует выражению $(p + q) \cdot r'$.

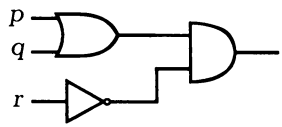


Рис. 1.23 □

ПРИМЕР 1.9. Булево выражение, соответствующее схеме на рис. 1.24, имеет вид $(p' \cdot q) + (p \cdot r')'$.

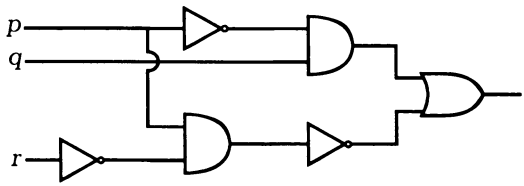


Рис. 1.24 □

ПРИМЕР 1.10. Коммутационная схема, соответствующая выражению $(p' \cdot q) + r$, показана на рис. 1.25.

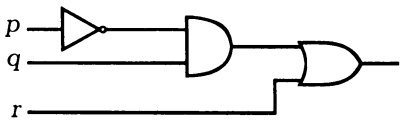


Рис. 1.25 □

ПРИМЕР 1.11. Коммутационная схема, соответствующая выражению $((p + q)' \cdot (p + r)) + r'$, показана на рис. 1.26.

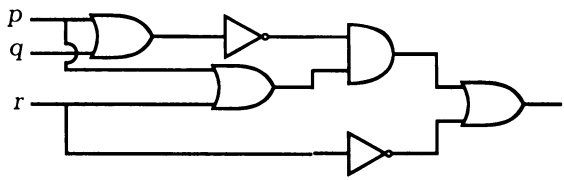


Рис. 1.26 □

ПРИМЕР 1.12. Построим схему трехклавишного переключателя, при помощи которого свет включается тремя различными двупозиционными переключателями. Рассмотрим сначала соответствующее булево выражение. Свет должен включаться, когда все три переключателя замкнуты, т.е. необходимо иметь pqr. Если один из переключателей разомкнут, то свет должен быть выключен. Однако, если разомкнуть другой переключатель, то свет должен включиться. Следовательно, искомое выражение имеет вид $pqr + pq'r' + p'q'r + p'qr'$. Для простоты, вместо схемы, представленной на рис. 1.27, для выражения pqr мы будем использовать схему, изображенную на рис. 1.28.

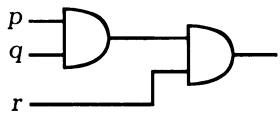


Рис. 1.27



Рис. 1.28

А для выражения $p + q + r$ вместо схемы на рис. 1.29 мы будем использовать схему, показанную на рис. 1.30.

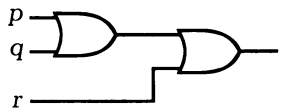


Рис. 1.29

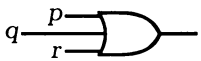


Рис. 1.30

Тогда искомая схема будет такой, как показано на рис. 1.33. □

Ранее было отмечено, что штрих Шеффера, обозначаемый через $|$, имеет ту же таблицу истинности, что и $(pq)'$ (в булевой записи), поэтому мы и упоминали его как логическую связку ***не-и***. В свою очередь, стрелка Пирса, обозначенная $\downarrow$, имеет ту же самую таблицу истинности, что и $(p+q)'$, поэтому она упоминалась как связка ***не-или***. Логические элементы *не-и* и *не-или* обозначаются символами, показанными на рис. 1.31 и 1.32.

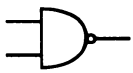


Рис. 1.31



Рис. 1.32

Выражению $(p|q) \downarrow (p|r)$ соответствует схема, изображенная на рис. 1.34.

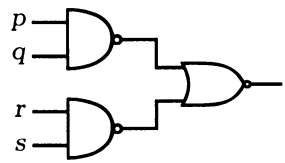


Рис. 1.34

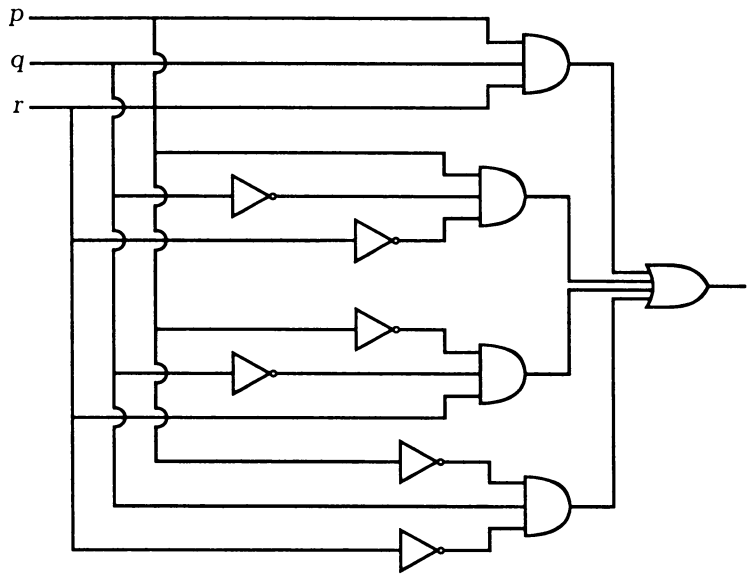


Рис. 1.33

ПРИМЕР 1.13. ***Полусумматор*** находит сумму двух двоичных чисел 1 и 0 согласно таблице сложения:

+	0	1
0	0	1
1	1	10

Свое название полусумматор получил в связи с тем, что при сложении двоичных чисел с более чем одним разрядом (что станет предметом обсуждения в последующей главе) суммируются только низшие разряды, поскольку нет возможности учесть в сумме число, которое "переносится". Для удобства суммирования одноразрядных двоичных чисел таблица, приведенная выше, преобразована к виду

+	0	1
0	00	01
1	01	10

Пусть p и q обозначают числа, которые требуется сложить, а d_1, d_0 — первый и второй разряды суммы, тогда приходим к следующим таблицам истинности:

Случай	p	q	d_0
1	1	1	0
2	1	0	1
3	0	1	1
4	0	0	0

Случай	p	q	d_1
1	1	1	1
2	1	0	0
3	0	1	0
4	0	0	0

Следовательно, $d_0 \equiv pq' + p'q$, что эквивалентно $(p+q)\cdot(pq)'$. Это можно показать, используя таблицы истинности или правило эквивалентности (см. упражнения). Также $d_1 \equiv pq$. Коммутационная схема для полусумматора приведена на рис. 1.35.

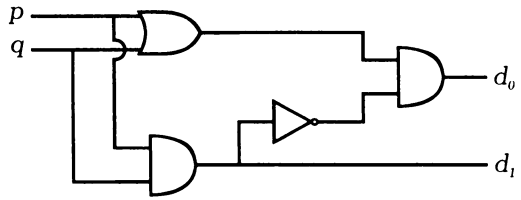


Рис. 1.35

Поскольку полусумматор дает сумму двух чисел, он обозначается символом, изображенным на рис. 1.36.

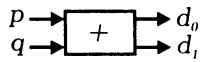


Рис. 1.36 □

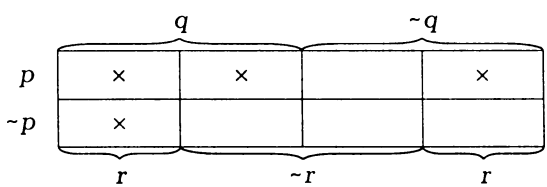
ПРИМЕР 1.14. ***Полный сумматор*** складывает три одноразрядных двоичных числа. Следовательно, он может сложить два двоичных числа с тем числом, которое "переносится". Пока необходимо рассматривать это как сложение трех одноразрядных двоичных чисел. Если предположить, что p, q и r обозначают числа, которые необходимо сложить, а $d_1^{\#}, d_0^{\#}$ — первый и второй разряды их суммы, получаем следующие таблицы истинности:

Случай	p	q	r	$d_1^{\#}$	$d_0^{\#}$
1	1	1	1	1	1
2	1	1	0	1	0
3	1	0	1	1	0
4	1	0	0	0	1
5	0	1	1	1	0
6	0	1	0	0	1
7	0	0	1	0	1
8	0	0	0	0	0

$d_0^{\#}$ действительно есть результат сложения d_0 с r, где d_0 — второй разряд суммы p и q. Следовательно, его схему легко описать. Значение $d_1^{\#}$ проще получить, используя приведенную выше таблицу истинности

$$d_1^{\#} \equiv pqr + pqr' + pq'r + p'qr\,.$$

Воспользовавшись картой Карно для выражения $d_1^{\#}$, получаем следующую карту:



Таким образом, $d_1^{\#} \equiv pq + pr + qr \equiv pq + (p+q)r$, поэтому схема может быть приведена к виду, показанному на рис. 1.37. (Более подробно эта схема представлена на рис. 1.39, приведенном ниже.)

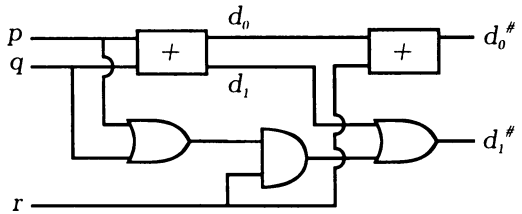


Рис. 1.37

Поскольку полный сумматор складывает три числа, он будет обозначаться символом, показанным на рис. 1.38.

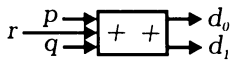


Рис. 1.38

□

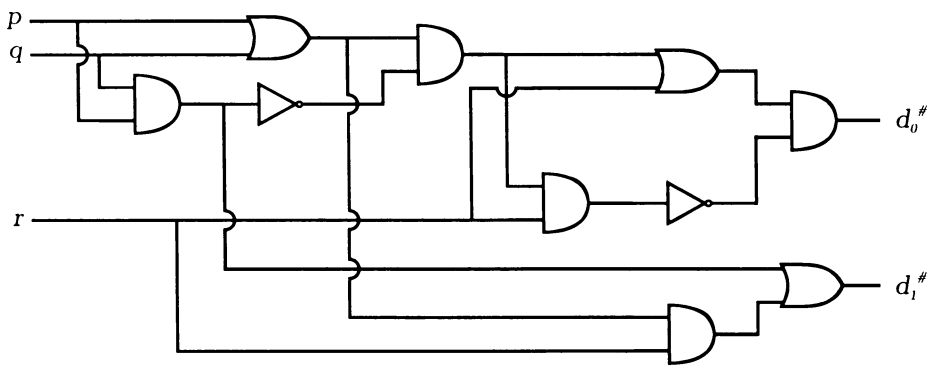


Рис. 1.39

■ УПРАЖНЕНИЯ

1. Перепишите следующие высказывания в обозначениях булевой алгебры:

а) $(p \wedge q) \wedge (q \vee \sim r)$;

б) $(q \vee \sim r) \wedge (\sim p \vee r) \wedge (q \vee \sim r \vee s)$;

в) $\sim(p \wedge \sim q \wedge r) \vee (\sim q \wedge r) \vee (p \wedge \sim q \wedge \sim r)$;

г) $\sim(\sim p \vee (q \wedge \sim r))$;

д) $(p \wedge r) \vee (p \wedge \sim q) \wedge (p \vee r \vee \sim s)$.

2. Перепишите следующие высказывания в обозначениях булевой алгебры:

а) $(p \vee q \vee \sim r) \wedge \sim(r \vee \sim q)$;

б) $(\sim q \wedge r) \vee \sim(p \wedge r)$;

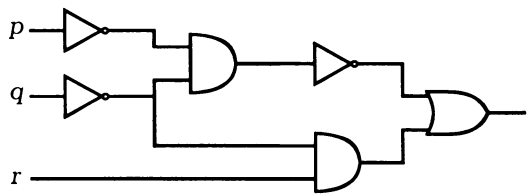
в) $\sim((p \wedge r) \vee \sim q)$;

г) $(p \wedge q \wedge r) \vee (\sim p \wedge q \wedge \sim r) \vee (\sim p \wedge q \wedge r) \vee (\sim p \wedge q \wedge \sim r)$;

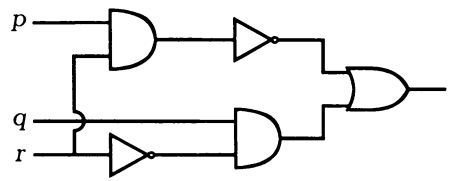
д) $(p \wedge \sim q \wedge \sim r \wedge s) \vee (\sim p \wedge q \wedge r \wedge s) \vee (\sim p \wedge q \wedge \sim r \wedge s)$.

3. Приведите булевы выражения, соответствующие коммутационным схемам:

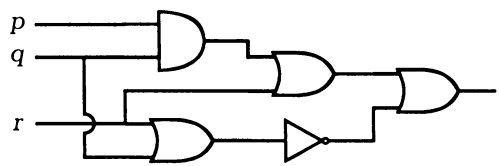
а)



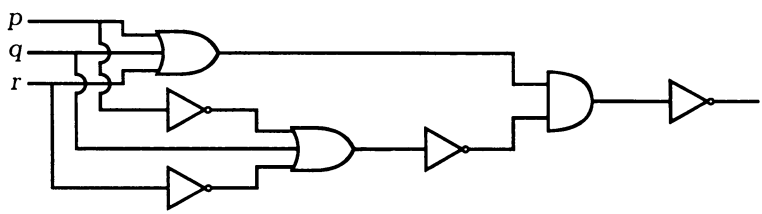
б)



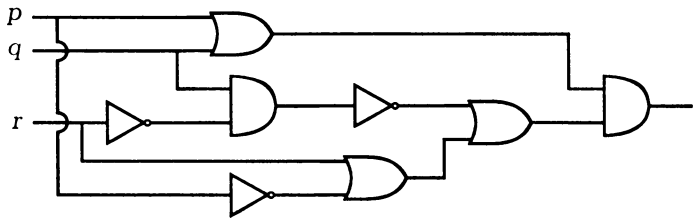
в)



г)

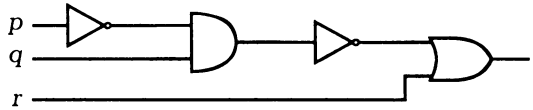


д)

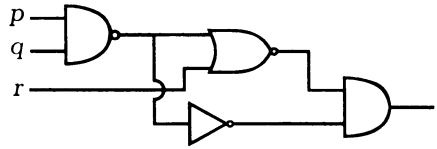


4. Приведите булевы выражения, соответствующие коммутационным схемам:

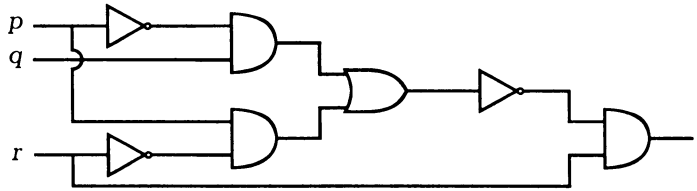
а)



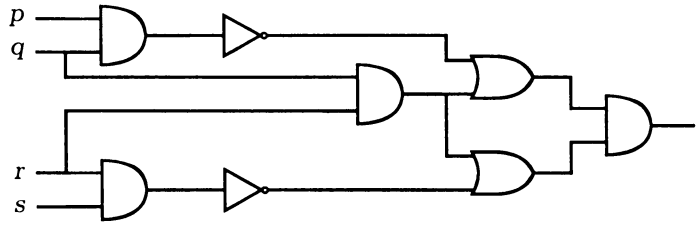
б)



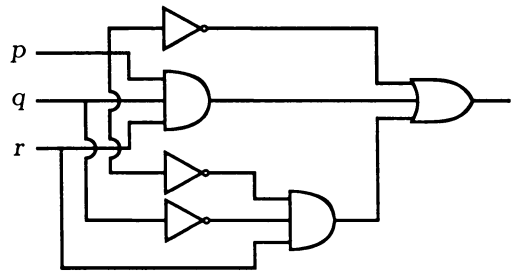
в)



г)



д)



5. Постройте коммутационные схемы, соответствующие булевым выражениям:
- **а)** $(p' + q)(p + (qr))$;
- **б)** $(pq') + ((qr') + (p'r))$;
- **в)** $(q'r') + ((pq)'(q'r))$;
- **г)** $(q(r + s))((p'q') + (qr'))$;
- **д)** $(pq) + ((pq'r') + r)'$.

6. Постройте коммутационные схемы, соответствующие булевым выражениям:
- **а)** $(pq')' + (qr)'$;
- **б)** $((p + q)z)'$;
- **в)** $((p'(q + p'r')) + (pq' + r)$;
- **г)** $(p'q + (q'r))' + ps'$;
- **д)** $((pq')' + (r's))(p + r')$.

7. Муниципальный совет состоит из пяти членов. Каждый член совета имеет для голосования кнопку "за" и кнопку "против". Решение принимается, если за него проголосует большинство. Постройте коммутационную схему устройства, сигнализирующего о том, что решение принято, путем высвечивания индикатора. *Указание*: при построении схем используйте последовательное и параллельное соединение переключателей, как показано в начале раздела.

8. Муниципальный совет состоит из пяти членов, включая председателя совета. Каждый член совета имеет для голосования кнопку "за" и кнопку "против". Решение принимается, если за него проголосует большинство. Председатель совета голосует только в том случае, если голоса "за" и "против" разделились поровну. Постройте коммутационную схему для определения принятия или непринятия решения путем высвечивания индикатора. *Указание*: при построении схем используйте последовательное и параллельное соединение переключателей, как показано в начале раздела.

9. Муниципальный совет состоит из пяти членов, включая председателя совета. Каждый член совета имеет для голосования кнопку "за" и кнопку "против". Решение принимается, если за него проголосует большинство, исключая председателя, который имеет право вето. Постройте коммутационную схему для определения принятия или непринятия решения путем высвечивания индикатора.

10. Электрическая схема содержит три двупозиционных переключателя. Сконструируйте схему для включения и выключения любым переключателем.

ГЛАВА

2

ТЕОРИЯ МНОЖЕСТВ

2.1. ПОНЯТИЕ МНОЖЕСТВА

Под ***множеством*** понимается некоторая, вполне определенная совокупность объектов или элементов. Это утверждение не следует рассматривать как строгое определение. Такое "определение" напоминает данное основоположником теории множеств Георгом Кантором определение множества как "объединения в одно целое объектов, хорошо различимых нашей интуицией или мыслью". Расплывчатость, недостаточность этого определения стала понятной, когда в 1879 году итальянский логик Бурали-Форти, а немного позже выдающийся философ и логик Бертран Рассел открыли парадоксы, указывающие на внутреннюю противоречивость канторовой теории множеств. Для устранения таких противоречий и парадоксов для теории множеств были предложены аксиоматические системы. Наиболее известны системы Цермело-Френкеля-фон Неймана, Гильберта-Бернайса-Геделя и Рассела-Уайтхеда. В силу ограниченности объема книги мы не будем изучать эти системы. Поэтому, по сути, мы оставим понятие множества неопределенным и будем считать множество заданным, если его элементы однозначно определены и это не приводит к каким-либо противоречиям.

Конечные множества можно описывать, перечисляя их элементы. Элементы, принадлежащие конечному множеству, условимся записывать между двумя фигурными скобками и разделять их запятыми. Например, $\{1, 2, 3, 4\}$ есть множество, содержащее натуральные числа 1, 2, 3 и 4. Множество гласных можно представить как {а, о, у, э, и, ы}. Как правило, для обозначения множеств будем использовать прописные буквы. A = {Боб, Джейн, Нэнси} есть множество, состоящее из Боба, Джейн и Нэнси. Множество первых n положительных целых чисел обозначаем $\{1, 2, 3, \ldots, n\}$, где точками показано продолжение перечисления элементов. Это же обозначение можно использовать для некоторых бесконечных множеств. Например, множество положительных целых чисел можно обозначить как $\{1, 2, 3, 4, \ldots\}$. Часто при перечислении множества используется описание характеристического свойства элементов этого множества. Например, $C = \{1, 8, 27, \ldots, k^3, \ldots\}$ описывает множество кубов всех положительных чисел, а $S = \{1, 4, 9, \ldots, n^2\}$ описывает множество квадратов всех положительных чи-

сел, которые меньше или равны n. Очевидно, что перечисление элементов удобно только в том случае, когда множество элементов мало или произвольный элемент характеризуется свойством, которое легко описать. Используя этот метод, не так легко описать множество граждан Соединенных Штатов Америки и совершенно немыслимо описать множество действительных чисел.

В общем случае множество задается путем указания характеристического свойства, т.е. свойства, которому удовлетворяют элементы данного множества, и только они. Для задания обычно используются фигурные скобки, а внутри них приводится характеристическое свойство, описывающее множество. Таким образом, множество $\{x : x$ обладает свойством $P\}$ предполагается содержащим только те объекты, которые имеют свойство P. Например, $\{x : x$ — футболист, играющий за Юго-западный колледж$\}$ — множество, состоящее из всех футбольных игроков, выступающих за Юго-западный колледж. Запись $\{x : x$ — гражданин Англии$\}$ описывает множество всех граждан Англии. Способ задания множества должен быть адекватным, т.е. должен полностью определять множество. Это не представляет труда, если объекты множества перечислены. Рассмотрим, однако, множество $A = \{x : x$ — высокий студент данной группы$\}$ или $B = \{x : x$ — хороший студент данной группы$\}$. Если различным студентам группы предложить определить множества A и B, они могут сделать это неоднозначно, выбирая в качестве элементов как множества A, так и множества B не одних и тех же людей. При рассмотрении множества $C = \{x : x$ — привлекательная (или красивая) студентка группы$\}$ выбрать элементы множества C не только трудно, но не стоит даже пытаться это сделать. Однако, если множество $A = \{x : x$ — студент данной группы, рост которого выше 180см$\}$ и $B = \{x : x$ — студент данной группы, средний балл которого не ниже 4$\}$, то можно сказать определенно, является ли данный студент элементом A или B, так что A и B действительно есть множества.

Таким образом, мы приходим к следующему формальному определению.

ОПРЕДЕЛЕНИЕ 2.1. Если a есть один из объектов множества A, мы говорим, что a есть ***элемент*** A, или a ***принадлежит*** A. Принадлежность элемента a множеству A записывается как $a \in A$. Если a не является элементом A, это записывается как $a \notin A$.

Например, $3 \in \{1, 2, 3, 4\}$, но $5 \notin \{1, 2, 3, 4\}$. Если P есть множество $\{x : x$ был президентом Соединенных Штатов$\}$, то Авраам Линкольн $\in P$, а Патрик Генри $\notin P$.

ОПРЕДЕЛЕНИЕ 2.2. Множество A есть ***подмножество*** множества B (обозначается $A \subseteq B$), если каждый элемент A есть элемент B; т.е. если $x \in A$, то $x \in B$. В частности, каждое множество есть подмножество самого себя. Если A не является подмножеством B, это записывается как $A \not\subseteq B$. Таким образом, $A \not\subseteq B$, если существует элемент A, не принадлежащий B.

Следовательно, $\{1,2,3\} \subseteq \{1,2,3,4\}$, но $\{1,2,5\} \not\subseteq \{1,2,3,4\}$. Если $A = \{x : x$ — футболист колледжа$\}$, $B = \{x : x$ — спортсмен колледжа$\}$, а $C = \{x : x$ — самый сильный математик колледжа$\}$, то $A \subseteq B$, но $C \not\subseteq B$.

Множества равны, если они содержат одни и те же элементы. Если A есть множество $\{2,4,6\}$, а B есть множество $\{x : x$ есть четное положительное целое число, которое меньше $7\}$, тогда A и B — равные множества. Таким образом, мы приходим к следующему определению.

ОПРЕДЕЛЕНИЕ 2.3. Пусть A и B — некоторые множества. Говорят, что A ***равно*** B, и пишут $A = B$, если для любого x имеем: $x \in A$ тогда и только тогда, когда $x \in B$. Иначе говоря, $A = B$ тогда и только тогда, когда $A \subseteq B$ и $B \subseteq A$. Если $A \subseteq B$ и $A \neq B$, то это записывают $A \subset B$ и говорят, что A есть ***собственное подмножество*** B.

Таким образом, доказательство равенства множеств A и B состоит из двух этапов:

1) Доказать, что A есть подмножество B.
2) Доказать, что B есть подмножество A.

Поскольку множество однозначно определяется только элементами, которые оно содержит, порядок их перечисления роли не играет. Например, $\{1,2,4,6\} = \{2,1,6,4\}$. Кроме того, любой элемент либо принадлежит данному множеству, либо нет. Каждый элемент может входить во множество не более одного раза.

С этого момента вводится два новых множества: универсальное множество, или универсум, и пустое множество. В известном смысле они представляют собой противоположности, поскольку пустое множество не содержит элементов, а универсальное множество содержит "все" элементы.

ОПРЕДЕЛЕНИЕ 2.4. ***Пустое множество***, обозначаемое $\varnothing$ или $\{\}$, есть множество, которое не содержит элементов. ***Универсальное множество*** $\boldsymbol{U}$ есть множество, обладающее таким свойством, что все рассматриваемые множества являются его подмножествами.

В теории чисел универсальное множество обычно совпадает со множеством всех целых или натуральных чисел. В математическом анализе универсальное множество может быть множеством всех действительных чисел или множеством всех точек n-мерного пространства. Следует отметить, что универсальное множество $\boldsymbol{U}$, хотя и названо универсальным, однозначно не определено, если точно не указана область рассмотрения (предметная область). Конечно, любое множество, содержащее $\boldsymbol{U}$, может быть использовано как универсальное множество.

По определению, каждое множество есть подмножество универсального множества. Пустое множество есть подмножество любого данного множества A, поскольку каждый элемент пустого множества содержится в A. Можно сказать, что не существует элементов пустого множества, которые не принадлежали бы A.

■ УПРАЖНЕНИЯ

1. Перечислите элементы множества $\{x : x$ — целое и $x^2 < 100\}$.
2. Перечислите элементы множества $\{x : x$ — президент США после 1940 года$\}$.
3. Перечислите элементы множества $\{x : x$ — гласный звук$\}$.
4. Перечислите элементы множества $\{x : x$ — положительное четное целое число, меньшее чем 21$\}$.
5. Опишите множество $\{3, 6, 9, 12, 15, 18, 21, 24\}$ при помощи характеристического свойства.
6. Опишите множество $\{a, b, c, d, e, f, g, h, i, j\}$ при помощи характеристического свойства.
7. Опишите множество {Огайо, Оклахома, Орегон} при помощи характеристического свойства.
8. Опишите множество $\{1, 4, 9, 16, 25, 36, \ldots\}$ при помощи характеристического свойства.
9. Перечислите подмножества множества $\{a\}$.
10. Перечислите подмножества множества $\{a, b\}$.
11. Перечислите подмножества множества $\{a, b, c\}$.
12. Перечислите подмножества множества $\{a, b, c, d\}$.
13. Перечислите подмножества множества $\varnothing$.
14. Используя результаты пяти предыдущих упражнений, определите число подмножеств для множества, состоящего из n элементов.
15. Установите истинность или ложность каждого из следующих утверждений:
    а) $\varnothing \subseteq \varnothing$;
    б) $\varnothing \subset \varnothing$;
    в) $\varnothing \in \varnothing$;
    г) $\varnothing \subseteq A$, где A — произвольное множество;
    д) $\varnothing \in A$, где A — произвольное множество.
16. Установите истинность или ложность каждого из следующих высказываний:
    а) $\{2\} \in \{1, 2, 3, 4, 5\}$;
    б) $\{2\} \subseteq \{1, 2, 3, 4, 5\}$;
    в) $\varnothing = \{\varnothing\}$;
    г) $\{1, 2, 3\} \in \{1, 2, 3, \{1, 2, 3\}\}$;
    д) $\{1, 2, 3\} \subseteq \{1, 2, 3, \{1, 2, 3\}\}$.
17. Определите количество элементов в каждом множестве:
    а) $\{\varnothing, \{\varnothing\}\}$;
    б) $\{\{\varnothing, \{\varnothing\}\}\}$;
    в) $\{1, 2, 3, \{1, 2, 3\}\}$;
    г) $\{\varnothing, \{\varnothing\}, a, b, \{a, b\}, \{a, b, \{a, b\}\}\}$;
    д) $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$.

2.2. ОПЕРАЦИИ НАД МНОЖЕСТВАМИ

Рассмотренные ниже операции над множествами позволяют строить новые множества, используя уже существующие.

ОПРЕДЕЛЕНИЕ 2.5. ***Пересечением*** множеств A и B называется множество, состоящее из всех тех и только тех элементов, которые принадлежат и A, и B. Пересечение множеств A и B обозначается $A \cap B$. Это определение равносильно следующему: $A \cap B = \{x : x \in A \text{ и } x \in B\}$.

Например, если $A = \{1, 2, 3, 4, 5\}$ и $B = \{1, 3, 5, 7, 9\}$, тогда $A \cap B = \{1, 3, 5\}$. Если $C = \{x : x$ имеет рост выше 180см$\}$ и $D = \{x : x$ любит играть в шахматы$\}$, тогда $C \cap D = \{x : x$ имеет рост выше 180см и любит играть в шахматы$\}$. Обратите внимание, что в описании пересечения множеств $B \cap C$ использована связка "и". В дальнейшем мы убедимся, что символы $\cap$ и $\wedge$, введенные в главе 1, связаны между собой и имеют схожие свойства.

Определим пересечение трех и более множеств. Пусть A_1, A_2 и A_3 — множества. Их пересечение можно определить следующим образом:

$$B = A_1 \cap (A_2 \cap A_3).$$

Далее будет показано, что $A_1 \cap (A_2 \cap A_3) = (A_1 \cap A_2) \cap A_3$, поэтому можно использовать запись

$$B = A_1 \cap A_2 \cap A_3.$$

Очевидно, $x \in B$ тогда и только тогда, когда $x \in A_1$, $x \in A_2$ и $x \in A_3$; иными словами, $x \in B$ тогда и только тогда, когда x принадлежит всем трем множествам A_1, A_2 и A_3. Пусть $J = \{1, 2, 3\}$. В таком случае $x \in B$ тогда и только тогда, когда $x \in A_j$ для всех $j \in J$, что равносильно записи

$$B = \{x : x \in A_j \text{ для всех } j \in J\}.$$

Пересечение множеств в общем случае определяется следующим образом.

ОПРЕДЕЛЕНИЕ 2.6. Если $I = \{1, 2, 3, \ldots, k\}$, то

$$\bigcap_{i \in I} A_i = A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_k = \\ = \{x : x \in A_i \text{ для всех } i \in I\}.$$

Объединением множеств A и B называется множество, состоящее из всех тех элементов, которые принадлежат хотя бы одному из множеств A или B. Объединение множеств A и B обозначается $A \cup B$. Сформулированное выше определение можно записать так: $A \cup B = \{x : x \in A$ или $x \in B\}$.

ПРИМЕР 2.7. Например, если $A = \{1, 2, 6, 7\}$, а $B = \{2, 3, 5, 6\}$, тогда $A \cup B = \{1, 2, 3, 5, 6, 7\}$. Объединение $A \cup B$ образовано из A и B путем соединения вместе элементов A и B. □

Если $A = \{x : x$ — политик$\}$, а $B = \{x : x$ — выпускник колледжа$\}$, то $A \cup B = \{x : x$ — политик или выпускник колледжа$\}$. Обратите внимание, что в описании объединения $A \cup B$ использована связка "или" так же, как в описании пересечения множеств использована связка "и".

Объединение множеств A_1, A_2 и A_3 определяется следующим образом:

$$B = A_1 \cup (A_2 \cup A_3).$$

Далее будет показано, что $A_1 \cup (A_2 \cup A_3) = (A_1 \cup A_2) \cup A_3$, поэтому можно использовать запись

$$B = A_1 \cup A_2 \cup A_3.$$

Очевидно, $x \in B$ тогда и только тогда, когда $x \in A_1$ или $x \in A_2$ или $x \in A_3$; иными словами, $x \in B$ тогда и только тогда, когда x принадлежит хотя бы одному из трех множеств A_1, A_2 или A_3. Таким образом, $x \in B$ тогда и только тогда, когда для некоторого $j \in \{1, 2, 3\}$, $x \in A_j$, что равносильно записи

$$B = \{x : x \in A_j \text{ для некоторого } j \in \{1, 2, 3\}\}.$$

Объединение множеств в общем случае определим следующим образом.

ОПРЕДЕЛЕНИЕ 2.8. Пусть $I = \{1, 2, 3, \ldots, k\}$, тогда

$$\bigcup_{i \in I} A_i = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_k =$$
$$= \{x : \text{существует } i \in I \text{ такое, что } x \in A_i\}.$$

ОПРЕДЕЛЕНИЕ 2.9. Пусть A и B множества. ***Разностью множеств*** $A - B$ называется множество всех тех и только тех элементов A, которые не содержатся в B. Или, что то же самое, $A - B = \{x : x \in A$ и $x \notin B\}$. ***Симметрическая разность*** множеств A и B, обозначаемая $A \triangle B$, есть множество $(A - B) \cup (B - A)$.

Например, если $A = \{1, 2, 4, 6, 7\}$, а $B = \{2, 3, 4, 5, 6\}$, то $A - B = \{1, 7\}$, а $A \triangle B$ есть множество $\{1, 3, 5, 7\}$. Симметричная разность множеств A и B состоит из тех элементов, которые принадлежат в точности одному из двух множеств A или B. Если $A = \{x : x$ играет в теннис$\}$, а $B = \{x : x$ играет в гольф$\}$, то $A - B = \{x : x$ играет в теннис, но не играет в гольф$\}$. Множество $A \triangle B = \{x : x$ играет только в теннис или играет только в гольф$\}$. Обратите внимание на сходство со связкой "исключающее или" из главы 1.

ОПРЕДЕЛЕНИЕ 2.10. ***Дополнение*** множества A, обозначаемое A', — это множество элементов универсума, которые не принадлежат A. Следовательно,

$$A' = \boldsymbol{U} - A = \{x : x \in \boldsymbol{U} \text{ и } x \notin A\}.$$

Если $\boldsymbol{U}$ — множество положительных целых чисел, а $A = \{2, 4, 6, 8, \ldots\}$ — множество всех четных положительных чисел, то $A' = \{1, 3, 5, 7, \ldots\}$ — множество всех нечетных положительных чисел. Если $\boldsymbol{U}$ — множество всех букв английского алфавита, а $V = \{a, e, i, o, u\}$, то V' — множество всех букв, обозначающих в английском языке согласный звук. Если $A = \{x : x$ — любитель научной фантастики$\}$, тогда $A' = \{x : x$ не любит научную фантастику$\}$. Обратите внимание, что дополнение множества связано с символом $\sim$ в логике.

В соответствии с определением равенство двух множеств может быть установлено в два этапа.

Докажите, что первое множество есть подмножество второго.
Докажите, что второе множество есть подмножество первого.

Этот способ используется, например, в доказательствах приведенных ниже теорем. Первый случай расписан более подробно.

ТЕОРЕМА 2.11. Для произвольных множеств A и B справедливо равенство

$$A - B = A \cap B'.$$

ДОКАЗАТЕЛЬСТВО. Для доказательства равенства двух множеств нужно показать, что каждое из множеств является подмножеством другого. Это можно осуществить, выбирая произвольный элемент одного множества и доказывая, что он принадлежит другому множеству. Такое доказательство гораздо легче провести, следуя приведенным ниже рассуждениям, поскольку в них каждый шаг обратим.

$$\begin{aligned} a \in A - B &\Leftrightarrow (a \in A) \wedge (a \notin B) \Leftrightarrow && \text{определение разности } A - B \\ &\Leftrightarrow (a \in A) \wedge (a \in B') \Leftrightarrow && \text{определение дополнения} \\ &\Leftrightarrow a \in (A \cap B'). && \text{определение пересечения} \end{aligned}$$

■

Обратите внимание, что при доказательстве следующей теоремы один из законов логики — закон де Моргана — существенно используется для доказательства соответствующего закона де Моргана в теории множеств.

ТЕОРЕМА 2.12. Для произвольных множеств A и B имеет место

а) $(A \cap B)' = A' \cup B'$;
б) $(A \cup B)' = A' \cap B'$.

ДОКАЗАТЕЛЬСТВО. Ниже приведено доказательство части (а). Доказательство части (б) предоставляется читателю.

Как и в предыдущем доказательстве, мы покажем, что каждое из множеств, входящих в равенство, есть подмножество другого.

$$\begin{aligned} a \in (A \cap B)' &\Leftrightarrow a \notin (A \cap B)' \Leftrightarrow && \text{определение дополнения} \\ &\Leftrightarrow \sim(a \in (A \cap B)) \Leftrightarrow && \text{определение } \notin \\ &\Leftrightarrow \sim((a \in A) \wedge (a \in B)) \Leftrightarrow && \text{определение пересечения} \end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow \sim(a \in A) \vee \sim(a \in B) \Leftrightarrow && \text{закон логики де Моргана}\\
&\Leftrightarrow (a \notin A) \vee (a \notin B) \Leftrightarrow && \text{определение } \notin\\
&\Leftrightarrow (a \in A') \vee (a \in B') \Leftrightarrow && \text{определение дополнения}\\
&\Leftrightarrow a \in (A' \cup B'). && \text{определение объединения}
\end{aligned}$$

■

Обратите внимание, что при доказательстве следующей теоремы логический закон дистрибутивности существенно используется для доказательства соответствующего закона дистрибутивности в теории множеств.

ТЕОРЕМА 2.13. Для произвольных множеств A, B и C справедливы равенства

а) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
б) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

ДОКАЗАТЕЛЬСТВО. Ниже приведено доказательство части (а). Доказательство части (б) предоставляется читателю. Опять покажем,что каждое из множеств, входящих в равенство, есть подмножество другого.

$$\begin{aligned}
a \in A \cap (B \cup C) &\Leftrightarrow (a \in A) \wedge (a \in (B \cup C)) \Leftrightarrow && \text{определение пересечения}\\
&\Leftrightarrow (a \in A) \wedge ((a \in B) \vee a \in C) \Leftrightarrow && \text{определение объединения}\\
&\Leftrightarrow ((a \in A) \wedge (a \in B)) \vee && \text{определение пересечения}\\
&\quad \vee ((a \in A) \wedge (a \in C)) \Leftrightarrow && \text{закон логики де Моргана}\\
&\Leftrightarrow (a \in (A \cap B)) \vee (a \in (A \cap C)) \Leftrightarrow && \text{определение пересечения}\\
&\Leftrightarrow a \in ((A \cap B) \cup (A \cap C)) \Leftrightarrow && \text{определение } \notin\\
&\Leftrightarrow (a \in A') \cup (a \in B'). && \text{определение объединения}
\end{aligned}$$

■

Мы убедились, что свойства, доказанные в теории множеств, имеют свой аналог в логике. В разделе 2.4 при рассмотрении булевой алгебры будет дано обобщение как операций теории множеств, так и логических операций.

ОПРЕДЕЛЕНИЕ 2.14. ***Множество всех подмножеств*** множества A, или ***булеан*** множества A, обозначаемый $\mathcal{P}(A)$, есть множество, состоящее из всех подмножеств множества A.

Следовательно, булеан множества $A = \{1, 2, 3\}$ есть множество

$$\mathcal{P}(A) = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{2,3\}, \{1,3\}, \{1,2,3\}\}.$$

Когда A содержит 3 элемента, $\mathcal{P}(A)$ состоит из $2^3 = 8$ элементов или, что то же самое, A включает $2^3 = 8$ подмножеств. Это будет показано в главе 8. В общем случае, если A содержит n элементов, множество $\mathcal{P}(A)$ включает 2^n элементов, т.к. A имеет 2^n подмножеств. По этой причине $\mathcal{P}(A)$ часто обозначают через 2^A.

Еще одной часто используемой операцией над множествами является декартово произведение, которое определяется следующим образом.

ОПРЕДЕЛЕНИЕ 2.15. ***Декартово произведение*** множеств A и B, обозначаемое $A \times B$, есть множество $\{(a,b) : a \in A \text{ и } b \in B\}$. Объект (a,b) называется ***упорядоченной парой*** с первой компонентой a и второй компонентой b.

Множество $A \times B$ состоит из всех упорядоченных пар, имеющих в качестве первой компоненты элемент из A, а в качестве второй компоненты — элемент из B. По существу, это та же упорядоченная пара, которую мы обычно используем в алгебре. Порядок компонент в паре существенен! Например, рисуя график функции, мы знаем, что точка $(1,2)$ не совпадает с точкой $(2,1)$.

Пусть $A = \{1,2,3\}$, а $B = \{r,s\}$. Тогда

$$A \times B = \{(1,r), (1,s), (2,r), (2,s), (3,r), (3,s)\}.$$

Если каждое из множеств A и B представляет собой множество действительных чисел, то $A \times B$ представляет собой *декартову плоскость*, на которой упорядоченные пары чисел используются для графического изображения функций.

Если A содержит n элементов, а B содержит m элементов, тогда $A \times B$ содержит $n \cdot m$ элементов. В частности, если A пусто или B пусто, то, по определению, $A \times B$ пусто.

■ УПРАЖНЕНИЯ

1. Пусть $A = \{1,2,3,4,5,6,7\}$, $B = \{4,5,6,7,8,9,10\}$, $C = \{2,4,6,8,10\}$, а $\boldsymbol{U} = \{1,2,3,4,5,6,7,8,9,10\}$. Определите следующие множества:

а) $A \cup C$; **б)** $A \cap B$;
в) $A \cap (B \cup C)$; **г)** $(A \cap B) \cup C$;
д) $(A \cap B)'$; **е)** $A' \cap B'$;
ж) $A \triangle B$; **з)** $A - B$.

2. Пусть $A = \{1,2,3,4,5,6,7\}$, $B = \{4,5,6,7,8,9,10\}$, $C = \{2,4,6,8,10\}$, а $\boldsymbol{U} = \{1,2,3,4,5,6,7,8,9,10\}$. Определите следующие множества:

а) $A - C$; **б)** $(A - B) \cup (B - A)$;
в) $A \cap (B \cap C')$; **г)** $(A \cup C) - B'$;
д) $(A - \varnothing) \cup (A - A)$; **е)** $B \triangle C$;
ж) $C - A$.

3. Пусть $A = \{1,2,3\}$, а $B = \{a,b\}$. Определите

а) $A \times B$; **б)** $B \times B$; **в)** $A \times \varnothing$.

4. Пусть $A = \{1,2,3\}$, а $B = \{a,b\}$. Определите

а) $A \times A$; **б)** $B \times A$; **в)** $A \triangle B$.

5. Определите $\mathcal{P}(A)$, если $A = \varnothing$.

6. Определите $\mathcal{P}(A)$, если $A = \{\varnothing, \{\varnothing\}\}$.

7. Определите $\mathcal{P}(\mathcal{P}(A))$, если $A = \varnothing$.

8. Определите, какие из следующих утверждений истинны, а какие ложны:

а) $A \cap \varnothing = A$; **б)** $A \triangle A = \varnothing$;
в) если $A \subseteq B$, то $A \cap B = A$; **г)** если $A \cap B = A$, то $B \subseteq A$;
д) $A - A = A$; **е)** $(A \times B)' = A' \times B'$.

9. Определите, какие из приведенных утверждений истинны, а какие ложны:
а) $A \cup \varnothing = A$; **б)** $A \triangle \varnothing = A$;
в) если $A \subseteq B$, то $A \cup B = A$; **г)** если $A \cup B = A$, то $B \subseteq A$;
д) $A - \varnothing = A$.

10. Докажите, что $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

11. Докажите, что $(A \cup B)' = A' \cap B'$.

12. Определим (a, b) как множество $\{a, \{a, b\}\}$. Покажите, что при таком определении мы вправе называть компоненты "первой" и "второй", т.к. $(a, b) = (c, d)$ тогда и только тогда, когда $a = c$ и $b = d$. Таким образом, это определение позволяет ввести понятие порядка через понятие множества, в котором порядок следования элементов не играет никакой роли.

13. Наследником множества A называется множество $A \cup \{A\}$. Определите наследников следующих множеств:
а) $\varnothing$; **б)** $\{\varnothing\}$; **в)** $\{\varnothing, \{\varnothing\}\}$.

14. Пусть I — некоторое множество и пусть для каждого $i \in I$ A_i — множество, обозначенное этим i, тогда

$$\bigcap_{i \in I} A_i = \{x : x \in A_i \text{ для всех } i \in I\}.$$

Множество $\bigcap_{i \in I} A_i$ называется ***обобщенным пересечением*** совокупности множеств $\{A_i : i \in I\}$ и обозначается иногда в сокращенном виде как $\{A_i\}$. Множество I называется ***индексным множеством***, а совокупность $\{A_i\}$ называется ***индексированной*** множеством I.

а) Пусть $N = \{1, 2, 3, \ldots\}$, а $I = \{\square, \triangle, \bigcirc\}$. Пусть

$$A_\square = \{x : x = 2k, \text{ где } k \text{ — некоторое целое число}\};$$
$$A_\triangle = \{x : x = 3k, \text{ где } k \text{ — некоторое целое число}\};$$
$$A_\bigcirc = \{x : x = 5k, \text{ где } k \text{ — некоторое целое число}\}.$$

Найдите такое C, что

$$C = \bigcap_{i \in I} A_i = \{x : x \in A_i \text{ для всех } i \in \{\square, \triangle, \bigcirc\}.$$

б) Пусть $I = N = \{1, 2, 3, \ldots\}$, а

$$A_i = \{x : x \in N \text{ и } x \geq i\} = \{i, i+1, i+2, \ldots\}.$$

Определите $\bigcap_{i \in I} A_i$.

15. Если I — множество, а A_i — множество для каждого $i \in I$, тогда

$$\bigcup_{i \in I} A_i = \{x : \text{существует } i \in I \text{ такое, что } x \in A_i\}.$$

Множество $\bigcup_{i \in I} A_i$ называется ***обобщенным объединением*** совокупности множеств $\{A_i : i \in I\}$. Пусть $I = N = \{1, 2, 3, \ldots\}$, а

$$A_i = \{x : x \in N \text{ и } x \geq i\} = \{i, i+1, i+2, \ldots\}.$$

Определите $\bigcup_{i \in I} A_i$.

2.3. ДИАГРАММЫ ВЕННА

Диаграммы Венна — очень удобный инструмент, позволяющий изображать множества и иллюстрировать операции над ними. Множества в диаграммах Венна изображаются внутренними частями кругов, их пересечениями, объединениями и т.д. Прямоугольник изображает универсальное множество. На рис. 2.1 приведена диаграмма Венна для множества A, которое изображено внутренней частью круга. Внешняя часть круга, находящаяся внутри прямоугольника, изображает A'.

На рис. 2.2 приведена диаграмма Венна для двух множеств, скажем, A и B, каждое множество изображено кругом, и круги пересекаются.

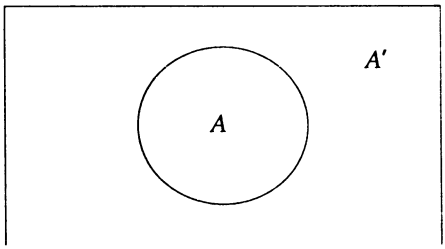


Рис. 2.1

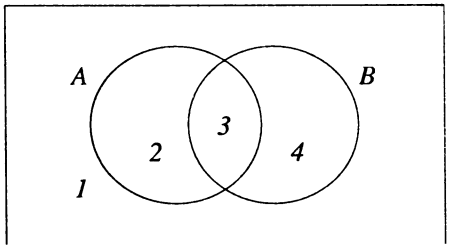


Рис. 2.2

Как показывает диаграмма, внутренняя часть прямоугольника разделена на четыре части. Множеству $A \cap B$ соответствует закрашенная часть диаграммы на рис. 2.3. Закрашенная область на рис. 2.4 представляет $A \cup B$.

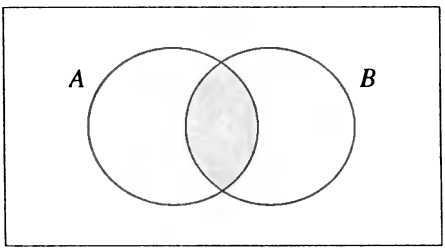


Рис. 2.3

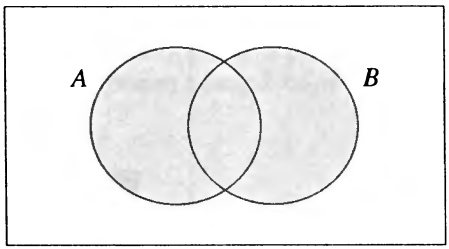


Рис. 2.4

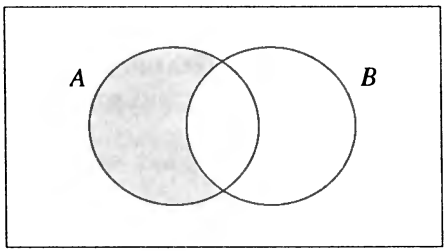


Рис. 2.6

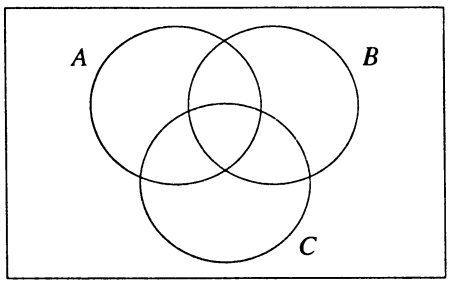
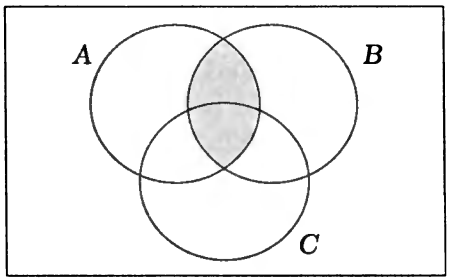
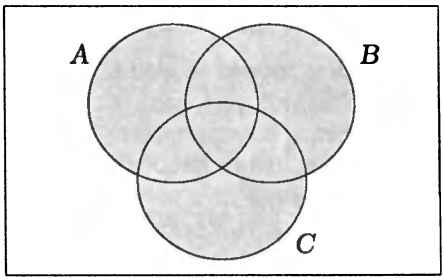


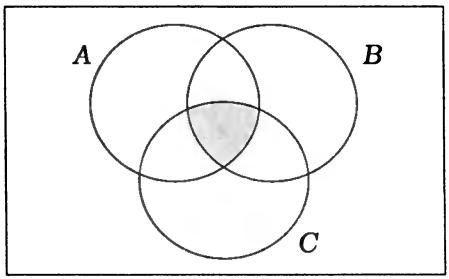
Рис. 2.7



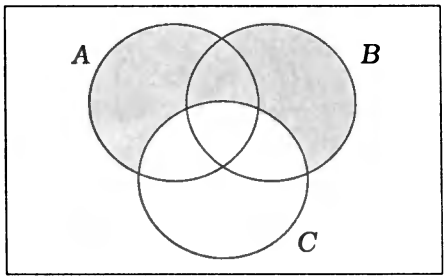
$A \cap B$



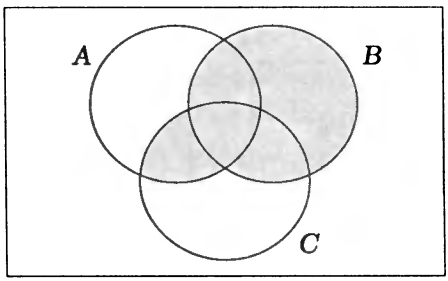
$A \cup B \cup C$



$A \cap B \cap C$



$(A \cup B)$-C



$(A \cap C) \cup B$

Рис. 2.5

Закрашенные области на рис. 2.5 изображают множества $A \cap B$, $(A \cup B \cup C)$, $(A \cap B \cap C)$, $(A \cup B) - C$ и $(A \cap C) \cup B$. Множество $A - B$ представлено закрашенной областью на рис. 2.6. Диаграмма Венна для трех множеств, например, A, B и C, показана на рис. 2.7. Эта диаграмма состоит из восьми частей.

Используя диаграммы Венна, можно показать равенство двух множеств.

ПРИМЕР 2.16. Покажем, что $(A \cup B)' = A' \cap B'$. Множество $(A \cup B)'$ — дополнение множества $A \cup B$, представленного диаграммой Венна на рис. 2.4, поэтому его изображает закрашенная область, показанная на рис. 2.8.

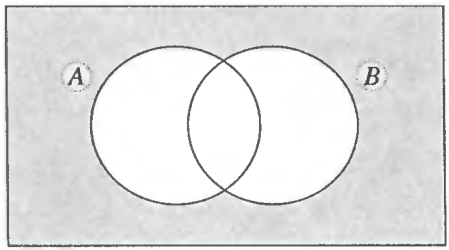


Рис. 2.8

Множеству A' соответствует закрашенная область на рис. 2.9. А множеству B' — закрашенная область на рис. 2.10.

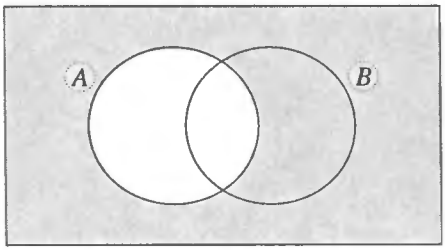


Рис. 2.9

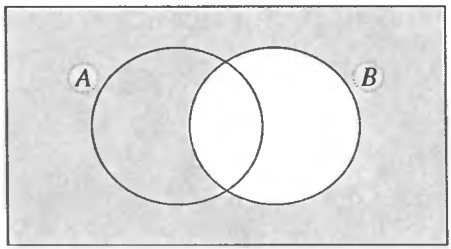


Рис. 2.10

Множеству $A' \cap B'$ соответствуют части, закрашенные на обеих предыдущих диаграммах, поэтому на рис. 2.11 оно изображено более темной областью.

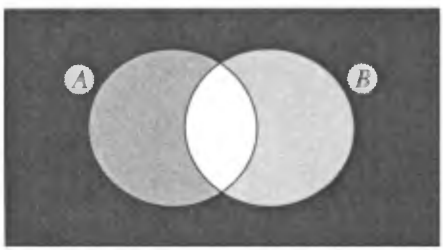


Рис. 2.11

Получили, что и $(A \cup B)'$, и $A' \cap B'$ одинаково изображаются на диаграмме Венна, поэтому $(A \cup B)' = A' \cap B'$. □

ПРИМЕР 2.17. Покажите, что $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. Множество A представлено закрашенной областью на рис. 2.12. Множеству $B \cup C$ соответствует закрашенная область на рис. 2.13.

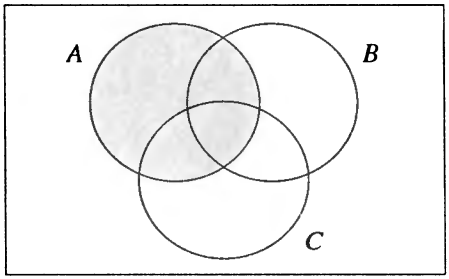


Рис. 2.12

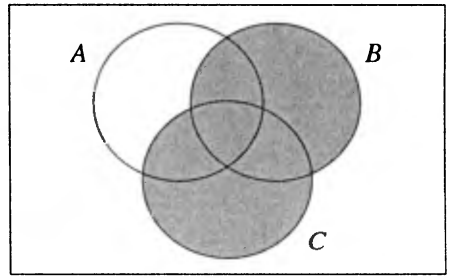


Рис. 2.13

Множество $A \cap (B \cup C)$ изображает область, закрашенная на обеих предыдущих диаграммах, поэтому оно представлено на рис. 2.14 более темной областью. Множество $A \cap B$ представлено на рис. 2.15 закрашенной областью.

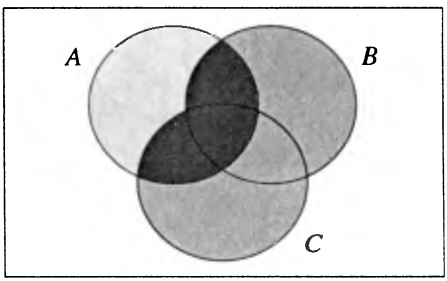


Рис. 2.14

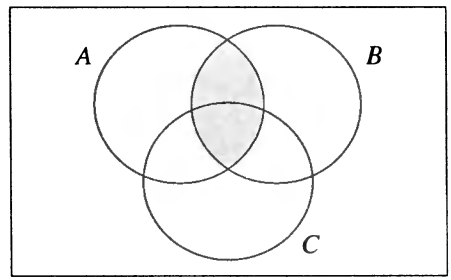


Рис. 2.15

Множество $A \cap C$ изображено на рис. 2.16 более темной областью, а множество $(A \cap B) \cup (A \cap C)$ изображено закрашенной областью на рис. 2.17.

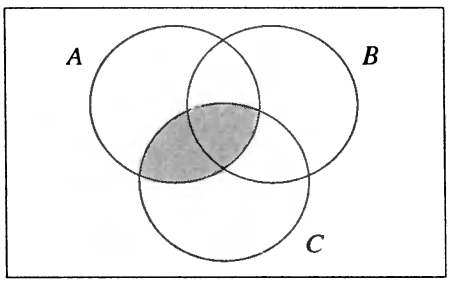


Рис. 2.16

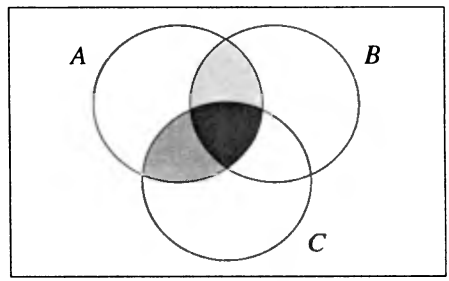


Рис. 2.17

Следовательно, $A \cap (B \cup C)$ и $(A \cap B) \cup (A \cap C)$ изображаются одинаково на диаграммах Венна, поэтому $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. □

Свойства множеств, сформулированные в приведенных ниже теоремах, могут быть проверены путем формальных доказательств или на диаграммах Венна. Обратите внимание, что они дублируют соответствующие свойства в исчислении высказываний (см. теорему 1.3).

ТЕОРЕМА 2.18. Пусть A, B и C — подмножества универсального множества $\boldsymbol{U}$. Тогда справедливы

а) ***Законы идемпотентности***

$$A \cap A = A;$$
$$A \cup A = A.$$

б) ***Двойное дополнение***

$$(A')' = A.$$

в) ***Законы де Моргана***

$$(A \cup B)' = A' \cap B';$$
$$(A \cap B)' = A' \cup B'.$$

г) ***Свойства коммутативности***

$$A \cap B = B \cap A;$$
$$A \cup B = B \cup A.$$

д) ***Свойства ассоциативности***

$$A \cap (B \cap C) = (A \cap B) \cap C;$$
$$A \cup (B \cup C) = (A \cup B) \cup C.$$

е) ***Свойства дистрибутивности***

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C);$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

ж) ***Свойства тождества***

$$A \cup \varnothing = A;$$
$$A \cap \boldsymbol{U} = A.$$

з) ***Свойства дополнения***

$$A \cup A' = \boldsymbol{U};$$
$$A \cap A' = \varnothing.$$

■ УПРАЖНЕНИЯ

1. Для каждого из приведенных ниже множеств используйте диаграммы Венна для двух множеств и заштрихуйте те ее части, которые изображают заданные множества:

а) A'; **б)** $(A \cap B)'$;

в) $(A \cup B) - A \cap B$; **г)** $A - B'$.

2. Для каждого из приведенных ниже множеств используйте диаграммы Венна для двух множеств и заштрихуйте те ее части, которые изображают заданные множества:

а) B'; **б)** $(A \cup B)'$;

в) $A - A \cap B$; **г)** $A \triangle B$.

3. Для каждого из приведенных ниже множеств используйте диаграммы Венна для двух множеств и заштрихуйте те ее части, которые изображают заданные множества:

а) $A - B$; **б)** $(A \cap B)'$;

в) $(A \cup B) - A \cap B$; **г)** $A \cup (B \cap C)$;

д) $(B \cap C) - A$; **е)** $B - (A \cup C)$;

ж) $(A \cap B \cap C)'$.

4. Для каждого из приведенных ниже множеств используйте диаграммы Венна для двух множеств и заштрихуйте те ее части, которые изображают заданные множества:

а) $A \triangle B$; **б)** $(A \cup B)'$;

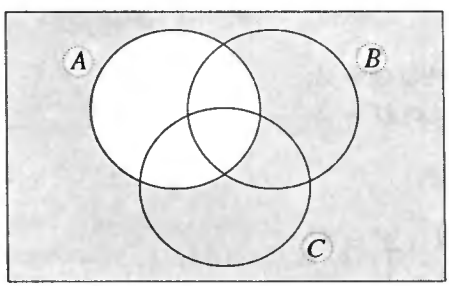
в) $A - A \cap B$; **г)** $(A \cap B) \triangle C$;

д) $(A \cup B \cup C) - (A \cap B \cap C)$; **е)** $(A \cap B) \cup (B \cap C) \cup (A \cap C)$;

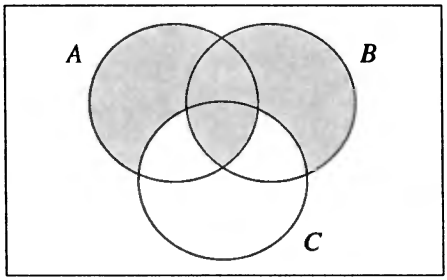
ж) $(A - B) \cup (B - C)$.

5. Опишите множества, соответствующие закрашенной части каждой диаграммы Венна:

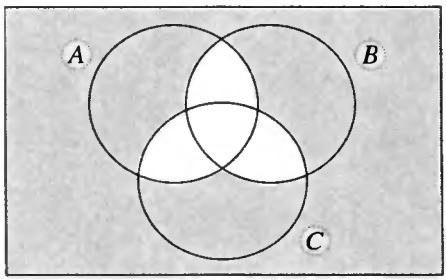
а)



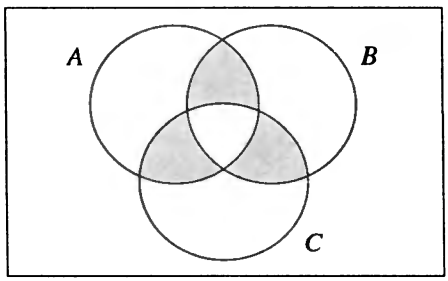
б)



в)

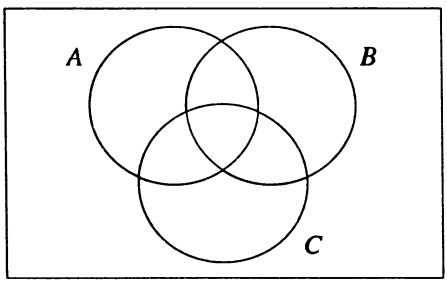


г)

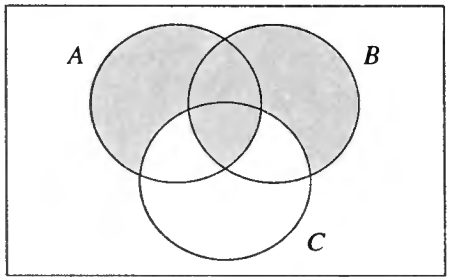


6. Опишите множества, соответствующие закрашенной части каждой диаграммы Венна:

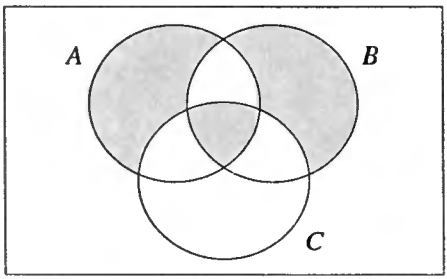
а)



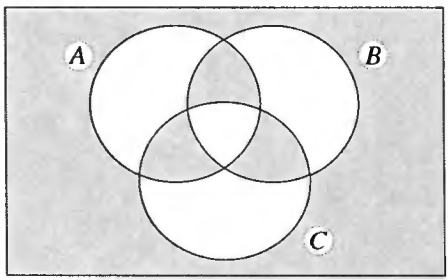
б)



в)



г)



7. С помощью диаграммы Венна покажите, что $(A \cap B)' = A' \cup B'$.
8. С помощью диаграммы Венна покажите, что $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

2.4. БУЛЕВЫ АЛГЕБРЫ

Сравнение основных свойств множеств и логики высказываний показало, что эти свойства имеют много общих черт. Данное обстоятельство нашло свое воплощение в общей теории, известной как ***булева алгебра***. Свое название теория получила в честь Дж. Буля, основоположника математической логики. Главное назначение данной главы состоит в установлении основных аксиом булевой алгебры и их использовании для доказательства теорем, справедливых для любой системы, обладающей подобными свойствами. В задачах и примерах этой главы

предполагаются известными элементарные свойства целых чисел, хотя формально их изучению будет посвящена следующая глава.

ОПРЕДЕЛЕНИЕ 2.19. Операция, заданная на некотором множестве, называется ***бинарной***, если она действует на два элемента этого множества и ее результатом является элемент этого же множества.

ОПРЕДЕЛЕНИЕ 2.20. Операция, заданная на множестве, называется ***унарной***, если она действует на один элемент множества и ее результатом является элемент этого же множества.

Обозначения, использованные ниже, были введены в главе 1.

ОПРЕДЕЛЕНИЕ 2.21. ***Булева алгебра*** есть множество B, содержащее специальные элементы 1 и 0, на котором заданы бинарные операции $+$ и $\cdot$ и унарная операция $'$. Для всех x, y и z из B должны выполняться следующие аксиомы:

а) ***Законы коммутативности***

$$x \cdot y = y \cdot x;$$
$$x + y = y + x.$$

б) ***Законы ассоциативности***

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z;$$
$$x + (y + z) = (x + y) + z.$$

в) ***Законы дистрибутивности***

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z);$$
$$x + (y \cdot z) = (x + y) \cdot (x + z).$$

г) ***Законы тождества***

$$x + 0 = x;$$
$$x \cdot 1 = x.$$

д) ***Законы дополнения***

$$x + x' = 1;$$
$$x \cdot x' = 0.$$

Элемент 1 называется ***единичным элементом***, или ***единицей***; элемент 0 называется ***нулевым элементом***, или ***нулем***; а x' называется ***дополнением*** x. Знак бинарной операции $\cdot$ часто опускается, и $x \cdot y$ записывается просто как xy.

Используя приведенные выше аксиомы, легко доказать ряд теорем о булевых алгебрах, в частности, справедливы приведенные ниже теоремы.

ТЕОРЕМА 2.22. Для всех элементов x и y булевой алгебры выполняются такие соотношения:

а) ***Законы идемпотентности***

$$x + x = x;$$
$$x \cdot x = x.$$

б) ***Свойства констант***

$$x + 1 = 1;$$
$$x \cdot 0 = 0.$$

в) ***Законы поглощения***

$$x + (x \cdot y) = x;$$
$$x \cdot (x + y) = x.$$

ДОКАЗАТЕЛЬСТВО. В каждом случае доказано только первое из утверждений теоремы. Доказательство второго утверждения предоставляется читателю.

а)
$$\begin{aligned} x + x &= (x + x) \cdot 1 = && \text{свойство констант} \\ &= (x + x) \cdot (x + x') = && \text{закон дополнения} \\ &= x + (x \cdot x') = && \text{закон дистрибутивности} \\ &= x + 0 = && \text{закон дополнения} \\ &= x; && \text{закон тождества} \end{aligned}$$

б)
$$\begin{aligned} x + 1 &= (x + 1) \cdot 1 = && \text{закон тождества} \\ &= (x + 1) \cdot (x + x') = && \text{закон дополнения} \\ &= x + (1 \cdot x') = && \text{закон дистрибутивности} \\ &= x + (x' \cdot 1) = && \text{закон коммутативности} \\ &= x + x' = && \text{закон тождества} \\ &= 1; && \text{закон дополнения} \end{aligned}$$

в)
$$\begin{aligned} x + (x \cdot y) &= (x \cdot 1) + (x \cdot y) = && \text{закон тождества} \\ &= x \cdot (1 + y) = && \text{закон дистрибутивности} \\ &= x \cdot (y + 1) = && \text{закон коммутативности} \\ &= x \cdot 1 = && \text{свойство констант} \\ &= x. && \text{закон тождества} \end{aligned}$$

■

ТЕОРЕМА 2.23. (Закон единственности дополнения) Дополнение произвольного элемента x булевой алгебры единственным образом определяется его свойствами: если $x + x' = 1$, $x \cdot x' = 0$, $x + x^* = 1$, а $x \cdot x^* = 0$, то $x' = x^*$.

ДОКАЗАТЕЛЬСТВО. Если $x + x' = 1$ и $x + x^* = 1$, тогда

$$\begin{aligned}
x' &= x' \cdot 1 = && \text{закон тождества} \\
&= x' \cdot (x + x^*) = && \text{задано} \\
&= x' \cdot x + x' \cdot x^* = && \text{закон дистрибутивности} \\
&= x \cdot x' + x' \cdot x^* = && \text{закон дополнения} \\
&= 0 + x' \cdot x^* = && \text{задано} \\
&= x' \cdot x^* + 0 = && \text{закон коммутативности} \\
&= x' \cdot x^* && \text{закон тождества}
\end{aligned}$$

и

$$\begin{aligned}
x^* &= x^* \cdot 1 = && \text{закон тождества} \\
&= x^* \cdot (x + x') = && \text{задано} \\
&= x^* \cdot x + x^* \cdot x' = && \text{закон дистрибутивности} \\
&= x \cdot x^* + x' \cdot x^* = && \text{закон коммутативности} \\
&= 0 + x' \cdot x^* = && \text{задано} \\
&= x' \cdot x^* + 0 = && \text{закон коммутативности} \\
&= x' \cdot x^*, && \text{закон тождества}
\end{aligned}$$

так что $x^* = x'x^* = x'$. ■

ТЕОРЕМА 2.24. Для всех элементов x и y булевой алгебры имеют место такие соотношения:

а) ***Закон инволюции***

$$(x')' = x.$$

б) ***Дополнение законов тождества***

$$0' = 1;$$
$$1' = 0.$$

в) ***Законы де Моргана***

$$(x + y)' = x' \cdot y';$$
$$(x \cdot y)' = x' + y'.$$

ДОКАЗАТЕЛЬСТВО. Для доказательства пункта (а) заметим, что

$$\begin{aligned}
x' + x = x + x' &= && \text{закон коммутативности} \\
&= 1; && \text{закон дополнения} \\
x' \cdot x = x \cdot x' &= && \text{закон коммутативности} \\
&= 0. && \text{закон дополнения}
\end{aligned}$$

Следовательно, x есть дополнение x' и в соответствии с законом единственности дополнения

$$(x')' = x\,.$$

Доказательство пункта (б) оставляем читателю. Доказательство пункта (в) приведено ниже. ■

Обратите внимание, что каждая аксиома булевой алгебры состоит из пары равенств, которые являются двойственными в том смысле, что если в одном равенстве заменить $+$ на $\cdot$, $\cdot$ на $+$, 0 на 1 и 1 на 0, получим второе равенство. В результате каждая теорема обладает двойственностью в том смысле, что если в любой теореме булевой алгебры заменить $+$ на $\cdot$, $\cdot$ на $+$, 0 на 1, а 1 на 0, снова получим теорему (хотя она может и не отличаться от исходной). Это соответствие имеет место, поскольку каждый шаг в доказательстве двойственной теоремы является двойственным соответствующему шагу в доказательстве исходной теоремы. Для примера рассмотрим приведенные ниже доказательства первого из законов де Моргана

$$(x + y)' = x' \cdot y'$$

и двойственное к нему соотношение

$$(x \cdot y)' = x' + y'\,.$$

Сначала докажем, что $(x+y)' = x'\cdot y'$. Для этого покажем, что $(x+y)+x'\cdot y' = 1$ и $(x+y)\cdot(x'\cdot y') = 0$. После этого, используя закон единственности дополнения, приходим к тому, что $(x + y)' = x' \cdot y'$.

$$\begin{aligned}
(x + y) + x' \cdot y' &= ((x + y) + x') \cdot ((x + y) + y') = && \text{закон дистрибутивности}\\
&= ((y + x) + x') \cdot ((x + y) + y') = && \text{закон коммутативности}\\
&= (y + (x + x')) \cdot (x + (y + y')) = && \text{закон ассоциативности}\\
&= (y + 1) \cdot (x + 1) = && \text{закон дополнения}\\
&= 1 \cdot 1 = && \text{свойство констант}\\
&= 1; && \text{закон тождества}
\end{aligned}$$

$$\begin{aligned}
(x + y) \cdot (x' \cdot y') &= (x' \cdot y') \cdot (x + y) = && \text{закон коммутативности}\\
&= ((x' \cdot y') \cdot x) + ((x' \cdot y') \cdot y) = && \text{закон дистрибутивности}\\
&= (x \cdot (x' \cdot y')) + ((x' \cdot y') \cdot y) = && \text{закон коммутативности}\\
&= ((x \cdot x') \cdot y') + (x' \cdot (y' \cdot y)) = && \text{закон ассоциативности}\\
&= ((x \cdot x') \cdot y') + (x' \cdot (y \cdot y')) = && \text{закон коммутативности}\\
&= (0 \cdot y') + (x' \cdot 0) = && \text{закон дополнения}\\
&= (y' \cdot 0) + (x' \cdot 0) = && \text{закон коммутативности}\\
&= 0 + 0 = && \text{свойство констант}\\
&= 0. && \text{закон тождества}
\end{aligned}$$

Теперь докажем, что $(x\cdot y)' = x'+y'$. Сначала покажем, что $(x\cdot y)\cdot(x'+y') = 0$ и $(x \cdot y) + (x' + y') = 1$. После этого, на основании закона единственности дополнения, получаем, что $(x \cdot y)' = x' + y'$.

$$
\begin{aligned}
(x \cdot y) \cdot (x' + y') &= ((x \cdot y) \cdot x') + \cdot((x \cdot y) \cdot y') = && \text{закон дистрибутивности} \\
&= ((y \cdot x) \cdot x') + ((x \cdot y) \cdot y') = && \text{закон коммутативности} \\
&= (y \cdot (x \cdot x')) + (x \cdot (y \cdot y')) = && \text{закон ассоциативности} \\
&= (y \cdot 0) \cdot (x \cdot 0) = && \text{закон дополнения} \\
&= 0 \cdot 0 = && \text{свойство констант} \\
&= 0; && \text{закон тождества}
\end{aligned}
$$

$$
\begin{aligned}
(x \cdot y) + (x' + y') &= (x' + y') + (x \cdot y) = && \text{закон коммутативности} \\
&= ((x' + y') + x) \cdot ((x' + y') + y) = && \text{закон дистрибутивности} \\
&= (x + (x' + y')) \cdot ((x' + y') + y) = && \text{закон коммутативности} \\
&= ((x + x') + y') \cdot (x' + (y' + y)) = && \text{закон ассоциативности} \\
&= ((x + x') + y') \cdot (x' + (y + y')) = && \text{закон коммутативности} \\
&= (1 + y') \cdot (x' + 1) = && \text{закон дополнения} \\
&= (y' + 1) \cdot (x' + 1) = && \text{закон коммутативности} \\
&= 1 \cdot 1 = && \text{свойство констант} \\
&= 1. && \text{закон тождества}
\end{aligned}
$$

Обратите внимание, что каждая строка второго доказательства двойственна соответствующей строке первого доказательства.

Итак, всякий раз, доказывая теорему относительно булевой алгебры, мы фактически доказываем две теоремы: саму теорему и двойственную ей (если они отличаются). Таким образом, доказывая одну теорему, вторую теорему получаем даром.

Используя правила теории множеств (теорема 2.18), легко можно показать, что подмножества произвольного множества A образуют булеву алгебру, где $\cap$ и $\cup$ аналогичны бинарным операциям $\cdot$ и $+$, соответственно, а $A - B$ соответствует B'. Множество A является единичным элементом 1, а пустое множество есть нулевой элемент 0 этой булевой алгебры.

■ УПРАЖНЕНИЯ

1. Докажите непосредственно, что $x \cdot x = x$.
2. Докажите непосредственно, что $x \cdot 0 = 0$.
3. Докажите непосредственно, что $x \cdot (x + y) = x$.
4. Докажите непосредственно, что $0' = 1$.
5. Докажите непосредственно, что $1' = 0$.
6. Найдите выражения, двойственные приведенным ниже:
   а) $x \cdot y' + x \cdot z' + y \cdot x'$;
   б) $x \cdot y \cdot z' + x \cdot y' \cdot z$;
   в) $x \cdot y \cdot (x + 0 + (z \cdot 1))$.
7. Найдите выражения, двойственные приведенным ниже:
   а) $(x + y') \cdot (z' + y)'$;
   б) $(1 + x) \cdot y + x \cdot y' \cdot z$;
   в) $(x \cdot y + 1) \cdot (0 + x) \cdot z$.

8. Докажите, что если $xy = xz$ и $x'y = x'z$, то $y = z$.
9. Сформулируйте задачу, двойственную к предыдущей.
10. Докажите, что $xy' = 0$ тогда и только тогда, когда $xy = x$.
11. Сформулируйте задачу, двойственную к предыдущей.
12. Докажите, что нулевой элемент 0 и единичный элемент 1 определены своими свойствами единственным образом.
13. Множество называется ***коконечным***, если его дополнение конечно. Пусть универсальное множество $\boldsymbol{U}$ есть множество всех конечных и всех коконечных подмножеств множества положительных целых чисел. Докажите, что подмножества $\boldsymbol{U}$ вместе с операциями объединения, пересечения и дополнения образуют булеву алгебру.

2.5. ОТНОШЕНИЯ

Среди рассмотренных операций над множествами было декартово произведение множеств A и B, которое обозначается через $A \times B$. Оно представляет собой множество $\{(a, b) : a \in A \text{ и } b \in B\}$. Таким образом, множество $A \times B$ состоит из всех упорядоченных пар, имеющих в качестве первой компоненты элемент из A, а в качестве второй компоненты — элемент из B.

ОПРЕДЕЛЕНИЕ 2.25. ***Отношением*** R множеств A и B называется произвольное подмножество $A \times B$. Если $(a, b) \in \mathbf{R}$, это записывают как $a\mathbf{R}b$; при этом говорят, что a и b находятся в отношении $\mathbf{R}$, или просто, что a ***относится*** к b. Если $A = B$, то отношение есть подмножество $A \times A$; такое отношение называют ***бинарным отношением*** на A.

В дальнейшем на множестве будем обычно рассматривать бинарные отношения, поэтому вместо термина "бинарное отношение" будем употреблять термин "отношение".

Если $A = \{1, 2, 3\}$, а $B = \{r, s\}$, так что

$$A \times B = \{(1, r),\ (1, s),\ (2, r),\ (2, s),\ (3, r),\ (3, s)\},$$

тогда $R = \{(1, r), (1, s), (3, s)\}$ есть отношение множеств A и B. Можно также записать $3Rs$, поскольку $(3, s) \in R$. Множество $A \times B$ содержит шесть элементов, поэтому существует $2^6 = 64$ подмножества множества $A \times B$. Следовательно, существует 64 различных отношения на $A \times B$.

В примерах и упражнениях, приведенных ниже, предполагаются известными элементарные свойства действительных и целых чисел, а также заданных на этих числах функций.

Вот несколько примеров отношений:

1. Все множество $A \times B$ есть отношение множеств A и B.
2. Если A — множество действительных чисел, то $\{(x, y) \in A \times A : x^2 + y^2 = 4\}$ есть бинарное отношение на A.

3. Пусть A — множество товаров в магазине, а B — множество действительных чисел. Тогда $\{(x, y) \in A \times B : y$ — цена $x\}$ — отношение множеств A и B.
4. Пусть A — множество женщин, а B — множество мужчин, тогда $\{(x, y) : y$ является мужем $x\}$ есть отношение множеств A и B.
5. Если A — множество людей, то $\{(x, y) \in A \times A : y$ является родственником $x\}$ есть бинарное отношение на A.

ОПРЕДЕЛЕНИЕ 2.26. ***Область определения*** отношения R на A и B есть множество всех $x \in A$ таких, что для некоторых $y \in B$ имеем $(x, y) \in R$. Другими словами, область определения R есть множество всех первых координат упорядоченных пар из R. ***Множество значений*** отношения R на A и B есть множество всех $y \in B$ таких, что $(x, y) \in R$ для некоторого $x \in A$. Другими словами, множество значений R есть множество всех вторых координат упорядоченных пар из R.

В примерах отношений, приведенных выше, в частности, в (1), область определения есть все множество A, а множество значений — все множество B. В (2) как область определения, так и множество значений совпадают с множеством $\{t : -2 \leq t \leq 2\}$. В (3) область определения есть множество A, а множество значений есть множество всех действительных чисел, каждое из которых совпадает с ценой некоторого товара в магазине. В (4) область определения есть множество всех замужних женщин, а множество значений — множество всех женатых мужчин. В (5) область определения и множество значений есть множество всех людей, имеющих родственников.

С каждым отношением R на $A \times B$ связано отношение R^{-1} на $B \times A$.

ОПРЕДЕЛЕНИЕ 2.27. Пусть $R \subseteq A \times B$ есть отношение на $A \times B$. Тогда отношение R^{-1} на $B \times A$ определяется следующим образом:

$$R^{-1} = \{(b, a) : (a, b) \in R\}.$$

Другими словами, $(b, a) \in R^{-1}$ тогда и только тогда, когда $(a, b) \in R$ или, что равносильно, $bR^{-1}a$ тогда и только тогда, когда aRb. Отношение R^{-1} называется ***обратным отношением*** к данному отношению R.

Пусть $R = \{(1, r), (1, s), (3, s)\}$, тогда $R^{-1} = \{(r, 1), (s, 1), (s, 3)\}$. Пусть R — отношение $\{(x, y) : y$ является мужем $x\}$, тогда R^{-1} — отношение $\{(x, y) : y$ — жена $x\}$. Пусть R — отношение $\{(x, y) : y$ является родственником $x\}$ либо R — отношение $\{(x, y) : x^2 + y^2 = 4\}$, тогда $R^{-1} = R$.

Имея два заданных отношения, можно образовать новые отношения указанным ниже способом.

ОПРЕДЕЛЕНИЕ 2.28. Пусть $R \subseteq A \times B$ — отношение на $A \times B$, а $S \subseteq B \times C$ — отношение на $B \times C$. ***Композицией*** отношений S и R называется отношение $T \subseteq A \times C$, определенное таким образом:

$$T = \{(a, c) : \text{ существует такой элемент } b \text{ из } B, \text{ что } (a, b) \in R \text{ и } (b, c) \in S\}.$$

Это множество обозначается $T = S \circ R$.

ПРИМЕР 2.29. Пусть $A = \{1, 2, 3\}$, $B = \{x, y\}$, а $C = \{\square, \triangle, \bigcirc, \bigstar\}$,, и пусть отношения R на $A \times B$ и S на $B \times C$ заданы в виде:

$$R = \{(1, x), (1, y), (3, x)\};$$
$$S = \{(x, \square), (x, \triangle), (y, \bigcirc), (y, \bigstar)\}.$$

Тогда

$$S \circ R = \{(1, \square),\ (1, \triangle),\ (1, \bigcirc),\ (1, \bigstar),\ (3, \square),\ (3, \triangle)\},$$

поскольку

$$\text{из } (1, x) \in R \text{ и } (x, \square) \in S \text{ следует } (1, \square) \in S \circ R;$$
$$\text{из } (1, x) \in R \text{ и } (x, \triangle) \in S \text{ следует } (1, \triangle) \in S \circ R;$$
$$\text{из } (1, y) \in R \text{ и } (y, \bigcirc) \in S \text{ следует } (1, \bigcirc) \in S \circ R;$$
$$\vdots$$
$$\text{из } (3, x) \in R \text{ и } (x, \triangle) \in S \text{ следует } (3, \triangle) \in S \circ R.$$

□

ПРИМЕР 2.30. Пусть R и S — бинарные отношения на множестве положительных целых чисел, заданные в виде $S = \{(x, x + 2) : x$ — положительное целое число$\}$ и $R = \{(x, x^2) : x$ — целое положительное число$\}$. Тогда $S \circ R = \{(x, x^2 + 2) : x$ — положительное целое число$\}$ и $R \circ S = \{(x, (x + 2)^2) : x$ — положительное целое число$\}$. □

ТЕОРЕМА 2.31. Композиция отношений ассоциативна; т.е., если A, B и C — множества и если $R \subseteq A \times B$, $S \subseteq B \times C$ и $T \subseteq C \times D$, тогда $T \circ (S \circ R) = (T \circ S) \circ R$.

ДОКАЗАТЕЛЬСТВО. Покажем сначала, что $T \circ (S \circ R) \subseteq (T \circ S) \circ R$. Пусть $(a, d) \in T \circ (S \circ R)$, тогда существует такое $c \in C$, что $(a, c) \in S \circ R$ и $(c, d) \in T$. Поскольку $(a, c) \in S \circ R$, существует такое $b \in B$, что $(a, b) \in R$ и $(b, c) \in S$. Поскольку $(b, c) \in S$ и $(c, d) \in T$, $(b, d) \in T \circ S$. Поскольку $(b, d) \in T \circ S$ и $(a, b) \in R$, $(a, d) \in (T \circ S) \circ R$. Таким образом, $T \circ (S \circ R) \subseteq (T \circ S) \circ R$. Вторая часть доказательства, показывающая, что $(T \circ S) \circ R \subseteq T \circ (S \circ R)$, аналогична и предоставляется читателю как упражнение. ■

Теперь рассмотрим специальные свойства отношений на A.

ОПРЕДЕЛЕНИЕ 2.32. Отношение R на $A \times A$ называется ***рефлексивным***, если (a, a) принадлежит R для всех a из A. Отношение R называется ***антирефлексивным***, если из $(a, b) \in R$ следует $a \neq b$. Отношение R ***симметрично***, если для всех a и b, принадлежащих A, из $(a, b) \in R$ следует, что $(b, a) \in R$. Отношение R ***транзитивно***, если для всех a, b и c из A из того, что $(a, b) \in R$ и $(b, c) \in R$, следует, что $(a, c) \in R$. Отношение R называется ***антисимметричным***, если для всех a и b из A, из принадлежности (a, b) и (b, a) отношению R следует, что $a = b$.

ПРИМЕР 2.33. Пусть $A = \{1, 2, 3, 4, 5, 6\}$ и пусть отношение $R_1 \subseteq A \times A$ есть множество $R_1 = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6), (1,2), (1,4), (2,1), (2,4), (3,5), (5,3), (4,1), (4,2)\}$. Отношение R_1 рефлексивно, т.к. для каждого $a \in A$, $(a, a) \in R_1$.

Рассмотрев все возможные случаи и показав, что в каждом из них из $(a, b) \in R_1$ следует $(b, a) \in R_1$, можно показать, что отношение R_1 является симметричным.

Случай	$(a, b) \in R_1$	(b, a)	$(b, a) \in R_1$?
1	$(1, 2)$	$(2, 1)$	Да
2	$(1, 4)$	$(4, 1)$	Да
3	$(2, 1)$	$(1, 2)$	Да
⋮	⋮	⋮	⋮

Можно также показать, что R_1 транзитивно, используя метод прямого перебора, как показано на примере следующей таблицы.

Случай	$(a, b) \in R_1$	$(b, c) \in R_1$	(b, c)	$(a, c) \in R_1$?
1	$(1, 2)$	$(2, 1)$	$(1, 1)$	Да
2	$(1, 2)$	$(2, 2)$	$(1, 2)$	Да
3	$(1, 2)$	$(2, 4)$	$(1, 4)$	Да
4	$(1, 4)$	$(4, 1)$	$(1, 1)$	Да
5	$(1, 4)$	$(4, 2)$	$(1, 2)$	Да
⋮	⋮	⋮	⋮	⋮

Проанализировав каждый возможный случай, когда $(a, b) \in R_1$ и $(b, c) \in R_1$, получаем, что $(a, c) \in R_1$.

R_1 не является антисимметричным, поскольку $(1, 2) \in R_1$ и $(2, 1) \in R_1$, но $1 \neq 2$. □

ПРИМЕР 2.34. Пусть $A = \{\square, \triangle, \bigcirc, \bigstar\}$ и пусть $R_2 \subseteq A \times A$ определено в виде

$$R_2 = \{(\square, \square), (\square, \triangle), (\square, \bigstar), (\triangle, \square), (\bigstar, \square), (\bigstar, \bigstar), (\bigcirc, \bigstar), (\bigcirc, \bigcirc)\}.$$

R_2 не является рефлексивным, т.к. $\triangle \in A$, но $(\triangle, \triangle) \notin R_2$. R_2 не является симметричным, поскольку $(\bigcirc, \bigstar) \in R_2$, но $(\bigstar, \bigcirc) \notin R_2$. R_2 не является антисимметричным, поскольку $(\triangle, \square) \in R_2$ и $(\square, \triangle) \in R_2$, но $\triangle \neq \square$. R_2 не является транзитивным, т.к. $(\triangle, \square) \in R_2$ и $(\square, \bigstar) \in R_2$, но $(\triangle, \bigstar) \notin R_2$. □

ПРИМЕР 2.35. Пусть A — множество положительных целых чисел. Определим отношение R, задавая $(x, y) \in R$ условием: y кратно x. R рефлексивно, поскольку для каждого положительного целого числа n, $n = 1 \cdot n$ и $(n, n) \in R$. R не является симметричным, т.к $(2, 4) \in R$, но $(4, 2) \notin R$; однако, R антисимметрично, т.к., если $(m, n) \in R$ и $(n, m) \in R$, тогда n кратно m и m кратно n, так что $m = n$. R транзитивно, потому что если $(m, n) \in R$ и $(n, p) \in R$, тогда n кратно m и p кратно n, так что p кратно m и $(m, p) \in R$. □

ОПРЕДЕЛЕНИЕ 2.36. Пусть R — бинарное отношение на множестве A. ***Рефлексивное замыкание*** R есть наименьшее рефлексивное отношение на A, содержащее R как подмножество. ***Симметричное замыкание*** R есть наименьшее симметричное отношение на A, содержащее R как подмножество. ***Транзитивное замыкание*** R есть наименьшее транзитивное отношение на A, содержащее R как подмножество.

ТЕОРЕМА 2.37. Пусть R — отношение на множестве A и $I = \{x : x = (a, a)$ для некоторого $a \in A\}$. Тогда

а) $R \cup I$ есть рефлексивное замыкание R;
б) $R \cup R^{-1}$ есть симметричное замыкание R;
в) если A — конечное множество, содержащее n элементов, то отношение $R \cup R^2 \cup R^3 \cup \cdots \cup R^n$ есть транзитивное замыкание R.

ДОКАЗАТЕЛЬСТВО. Доказательство утверждений (а) и (б) предоставляется читателю. Обозначим транзитивное замыкание R через $\overline{R}$. Для доказательства утверждения (с) сначала покажем, что $R \cup R^2 \cup R^3 \cup \cdots \cup R^n \subseteq \overline{R}$. Проведем индукцию по n. Для $n = 1$ имеем $R \subseteq \overline{R}$, что достоверно истинно. Допустим, что $R \cup R^2 \cup R^3 \cup \cdots \cup R^k \subseteq \overline{R}$. Необходимо показать, что $R \cup R^2 \cup R^3 \cup \cdots \cup R^{k+1} \subseteq \overline{R}$ или, что то же самое, $R^{k+1} \subseteq \overline{R}$. Пусть $(a, c) \in R^{k+1}$. Тогда существует b такое, что $(a, b) \in R^k$ и $(b, c) \in R$. Но, согласно индуктивному предположению, (a, b) и $(b, c) \in \overline{R}$. Поскольку $\overline{R}$ транзитивно, $(a, c) \in \overline{R}$. Поэтому $R \cup R^2 \cup R^3 \cup \cdots \cup R^{k+1} \subseteq \overline{R}$. Для того, чтобы показать, что $\overline{R} \subseteq R \cup R^2 \cup R^3 \cup \cdots \cup R^{k+1}$, просто покажем, что $R \cup R^2 \cup R^3 \cup \cdots \cup R^{k+1}$ транзитивно. Пусть $(a, b) \in R^j$ и $(b, c) \in R^k$. Тогда $(a, c) \in R^{j+k}$. Если $a = c$, утверждение доказано. В противном случае существуют $b_2, b_3, b_4, \ldots, b_{j+k-1} \in A$ такие, что $(a, b_2), (b_2, b_3), (b_3, b_4), \ldots, (b_{j+k-2}, b_{b_{j+k-1}}), (b_{j+k-1}, c) \in R$. Обозначим a через b_1, а c через b_{j+k}. Если некоторые из b_i равны, например, $b_p = b_q$, из указанной выше последовательности упорядоченных пар, находящихся в отношении R, можно удалить $(b_p, b_{p+1}), (b_{p+1}, b_{p+3}), \ldots, (b_{q-1}, b_q)$ и после этого получить последовательность $a, b_2, b_3, b_{p-1}, b_q, \ldots, b_{j+k-1}, c$, в которой каждый предыдущий элемент находится в R-отношении к последующему. Так можно продолжать до тех пор, пока все элементы окажутся отличными, но при этом каждый из них будет нахо-

диться в R-отношении к последующему. Поскольку во множестве A существует только n различных элементов, получаем, что $(a, c) \in R^n$ и $R \cup R^2 \cup R^3 \cup \cdots \cup R^{k+1}$ транзитивно. ■

Одним из наглядных способов представления конечного антирефлексивного симметричного отношения является граф. В дальнейшем мы убедимся, что легко установить связь между конечными антирефлексивными симметричными отношениями и графами. Однако, помимо этого конкретного использования, графы получили самостоятельное развитие, и в настоящее время теория графов является одной из важных областей математики.

ОПРЕДЕЛЕНИЕ 2.38. ***Граф*** есть конечное множество V, называемое ***множеством вершин***, на котором задано симметричное антирефлексивное отношение R и выделено множество E двухэлементных подмножеств V, определяемое как $\{a, b\} \in E$ тогда и только тогда, когда $(a, b) \in R$ и $a \neq b$. Множество E называется ***множеством ребер***. Всякий элемент E называется ***ребром***. Граф обозначается $G(V, E)$. Говорят, что элементы a и b графа V ***соединены*** или ***связаны*** ребром $\{a, b\}$, если $\{a, b\} \in E$.

Конечный граф обычно изображается при помощи диаграммы, в которой вершины представлены точками, а ребра, соединяющие две вершины, — линиями между этими точками.

ПРИМЕР 2.39. Граф, в котором множество вершин $V = \{a, b, c\}$, а $E = \{\{a, b\}, \{b, c\}\}$, может иметь вид, как на рис. 2.18 или рис. 2.19.

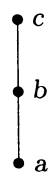


Рис. 2.18

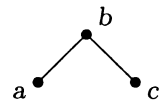


Рис. 2.19

$R = \{(a, b), (b, a), (b, c), (c, b)\}$. □

ПРИМЕР 2.40. Граф, в котором множество вершин $V = \{a, b, c, d, e\}$, а

$$E = \{\{a, b\}, \{a, e\}, \{b, e\}, \{b, d\}, \{b, c\}, \{c, d\}\},$$

имеет диаграмму, изображенную на рис. 2.20.

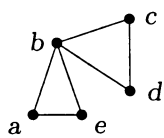


Рис. 2.20

$R = \{(a, b), (b, a), (e, a), (a, e), (e, b), (b, e), (b, d), (d, b), (b, c), (c, b), (d, c), (c, d)\}$. □

Мы видели, что граф описывал симметричное отношение R на множестве V, в котором никакой элемент не находился в отношении с самим собой. Для отношения R более общего вида необходимо представление элемента $(a, b) \in R$, для которого возможно $(b, a) \notin R$. Необходимо также иметь возможность представлять элемент $(a, a) \in R$. Все сказанное можно осуществить с помощью ориентированного графа.

ОПРЕДЕЛЕНИЕ 2.41. ***Ориентированный граф***, или ***орграф*** G, обозначаемый $G(V, E)$, состоит из множества V вершин и отношения E на V, называемого множеством ***ориентированных ребер***, или просто ***ребер***, если понятно, что граф ориентирован. Элемент множества E называется ***ориентированным ребром***. Если $(a, b) \in E$, тогда a называется ***начальной вершиной*** (a, b), а b — его ***конечной вершиной***.

Обратите внимание, что, в отличие от неориентированного графа, в случае ориентированного графа мы допускаем наличие петель. Причин этому две. Во-первых, ориентированные ребра естественным образом включаются в наше определение, поскольку петля при вершине a есть просто ребро (a, a). Этого нельзя было сделать в случае неориентированного графа, поскольку его ребра имеют вид $\{a, b\}$, так что петля имела бы вид $\{a, a\}$. Во-вторых, вообще говоря, элементы могут находиться в отношении к самим себе, что для множеств не имеет смысла, поскольку любой элемент присутствует во множестве только один раз. Тем не менее, дальнейшее рассмотрение покажет, что петли в ориентированных графах встречаются чаще, чем в неориентированных.

Ребро (a, b) орграфа обозначается на диаграмме стрелкой от a к b. Заметим, что в простом графе ребро представляется двухэлементным подмножеством, чтобы подчеркнуть, что ребро как отношение симметрично, тогда как в орграфе ребро представлено упорядоченной парой, чтобы подчеркнуть то, что порядок имеет значение, и то, что (a, b) может быть ребром диаграммы, а (b, a) — нет.

ПРИМЕР 2.42. Орграф с вершинами

$$V = \{a, b, c\}$$

и ребрами

$$E = \{(a, a), (a, b), (b, c), (c, b), (c, a)\}$$

показан на рис. 2.21.

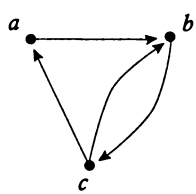


Рис. 2.21

□

ПРИМЕР 2.43. Орграф с вершинами

$$V = \{a, b, c, d\}$$

и ребрами

$$E = \{(a,b),(b,c),(c,c),(b,d),(d,b),(c,d),(d,a)\}$$

показан на рис. 2.22.

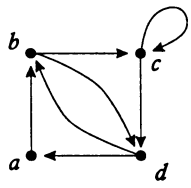


Рис. 2.22 □

■ УПРАЖНЕНИЯ

1. Найдите область определения и множество значений отношений:

а) $\{(a,1),(a,2),(c,1),(c,2),(c,4),(d,5)\}$;

б) $\{(1,2),(2,4),(3,6),(4,8),\ldots\}$;

в) $\{(x,y) : x,y \in R$ и $x = y^2\}$.

2. Найдите область определения и множество значений отношений:

а) $\{(a,1),(a,2),(a,3),(a,4),(a,5),(a,6)\}$;

б) $\{(x,y) : x,y \in I$ и $x^2 + y^2 \leq 16\}$;

в) $\{(x,y) : 0 \leq x,y \leq 10$ и $x > 2y\}$.

3. Пусть $A = \{1,2,3,4,5\}$;

$B = \{6,7,8,9\}$;

$C = \{10,11,12,13\}$;

$D = \{\square, \triangle, \bigcirc, *\}$.

Пусть $R \subseteq A \times B$, $S \subseteq B \times C$ и $T \subseteq C \times D$ определены следующим образом

$$R = \{(1,7),(4,6),(5,6),(2,8)\};$$
$$S = \{(6,10),(6,11),(7,10),(8,13)\};$$
$$T = \{(11,\triangle),(10,\triangle),(13,*),(12,\square),(13,\bigcirc)\}.$$

Определите отношения:

а) R^{-1} и S^{-1};

б) $S \circ R$;

в) $S \circ S^{-1}$ и $S^{-1} \circ S$;

г) $R^{-1} \circ S^{-1}$;

д) $T \circ (S \circ R)$;

е) $T \circ S$;

ж) $(T \circ S) \circ R$.

4. Пусть $A = \{(b,a),(c,e),(d,i),(f,o)(g,u)\}$ и $B = \{(v,a),(w,e),(x,i),(y,o)(z,u)\}$.

а) Опишите отношение A^{-1}.

б) Опишите отношение B^{-1}.

в) Опишите отношение $A^{-1} \circ B$.

г) Опишите отношение $B^{-1} \circ A$.

5. Пусть отношения $U, V \subseteq R \times R$ определены указанным ниже способом $U = \{(x, y) : y = x^2 + 5)\}$ и $V = \{(x, y) : y = 3x\}$.

а) Опишите отношение $U \circ V$.

б) Опишите отношение $V \circ U$.

в) Опишите отношение U^{-1}.

г) Опишите отношение V^{-1}.

д) Найдите область значений U.

6. Пусть $A = \{a, b, c, d, e\}$. Опишите наименьшее рефлексивное отношение на множестве A.

7. Пусть $A = \{a, b, c, d, e\}$ и $S = \{(a, a), (a, b), (b, c), (b, d), (c, e), (e, d), (c, b)\}$.

а) Опишите наименьшее симметричное отношение на A, содержащее S.

б) Опишите наименьшее рефлексивное и симметричное отношение на A, содержащее S.

в) Опишите наибольшее симметричное отношение, содержащееся в S.

г) Опишите наименьшее транзитивное отношение на A, содержащее S.

8. Пусть $A = \{a, b, c, d, e\}$, а S, T, U и V — отношения на A, где

$$S = \{(a, a), (a, b), (b, c), (b, d), (c, e), (e, d), (c, a)\};$$
$$T = \{(a, b), (b, a), (b, c), (b, d), (e, e), (d, e), (c, b)\};$$
$$U = \{(a, b), (a, a), (b, c), (b, b), (e, e), (b, a), (c, b), (c, c), (d, d), (a, c), (c, a)\};$$
$$V = \{(a, b), (b, c), (b, b), (e, e), (b, a), (c, b), (d, d), (a, c), (c, a)\}.$$

а) Какое из отношений является симметричным?

б) Какое из отношений является рефлексивным?

в) Какое из отношений является транзитивным?

г) Какое из отношений является антисимметричным?

9. Пусть $A = \{a, b, c, d, e\}$, а S, T, U и V — отношения на A, где

$$S = \{(a, a), (a, b), (b, c), (b, d), (c, e), (e, d), (c, a)\};$$
$$T = \{(a, b), (b, a), (b, c), (b, d), (e, e), (d, e), (c, b)\};$$
$$U = \{(a, b), (a, a), (b, c), (b, b), (e, e), (b, a), (c, b), (c, c), (d, d), (a, c), (c, a)\};$$
$$V = \{(a, b), (b, c), (b, b), (e, e), (b, a), (c, b), (d, d), (a, c), (c, a)\}.$$

а) Опишите $U \cap V$.

б) Опишите $S \cup T$.

в) Опишите $U - T$.

г) Опишите $U \triangle S$.

10. Докажите, что пересечение рефлексивных отношений рефлексивно.

11. Докажите, что пересечение симметричных отношений симметрично.

12. Пусть $A = \{a, b, c, d, e\}$.

а) Опишите отношение на A, которое рефлексивно, но не является ни симметричным, ни транзитивным.

б) Опишите отношение на A, которое симметрично, но не является ни рефлексивным, ни транзитивным.

в) Опишите отношение на A, которое транзитивно, но не является ни рефлексивным, ни симметричным.

13. Пусть $A = \{a, b, c, d, e\}$.

а) Опишите отношение на A, которое рефлексивно и симметрично, но не является транзитивным.

б) Опишите отношение на A, которое симметрично и транзитивно, но не является рефлексивным.

в) Опишите отношение на A, которое рефлексивно и транзитивно, но не является симметричным.

14. Установите истинность или ложность каждого из приведенных ниже высказываний. Для каждого ложного высказывания приведите контрпример.

а) Если отношения R и S рефлексивны, то отношение $R \cap S$ рефлексивно.

б) Если отношения R и S рефлексивны, то отношение $R \cup S$ рефлексивно.

в) Если отношения R и S рефлексивны, то отношение $R \circ S$ рефлексивно.

г) Если отношения R и S рефлексивны, то отношение $R - S$ рефлексивно.

д) Если отношения R и S рефлексивны, то отношение $R \triangle S$ рефлексивно.

15. Установите истинность или ложность каждого из следующих высказываний. Для ложных высказываний приведите контрпример.

а) Если отношения R и S симметричны, то отношение $R \cap S$ симметрично.

б) Если отношения R и S симметричны, то отношение $R \cup S$ симметрично.

в) Если отношения R S симметричны, то отношение $R \circ S$ симметрично.

г) Если отношения R и S симметричны, то отношение $R - S$ симметрично.

д) Если отношения R и S симметричны, то отношение $R \triangle S$ симметрично.

16. Установите истинность или ложность приведенных ниже высказываний. Для ложных высказываний приведите контрпример.

а) Если отношения R и S антисимметричны, то отношение $R \cap S$ антисимметрично.

б) Если отношения R и S антисимметричны, то отношение $R \cup S$ антисимметрично.

в) Если отношения R и S антисимметричны, то отношение $R \circ S$ антисимметрично.

г) Если отношения R и S антисимметричны, то отношение $R - S$ антисимметрично.

д) Если отношения R и S антисимметричны, то отношение $R \triangle S$ антисимметрично.

17. Установите истинность или ложность приведенных ниже высказываний. Для каждого ложного высказывания приведите контрпример.

а) Если отношения R и S транзитивны, то отношение $R \cap S$ транзитивно.

б) Если отношения R и S транзитивны, то отношение $R \cup S$ транзитивно,

в) Если отношения R и S транзитивны, то отношение $R \circ S$ транзитивно.

г) Если отношения R и S транзитивны, то отношение $R - S$ транзитивно.

д) Если отношения R и S транзитивны, то отношение $R \triangle S$ транзитивно.

18. Постройте граф для каждого из приведенных ниже отношений на A.

а) $A = \{a, b, c, d, e\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, a), (a, c), (d, e), (e, d)\}$;

б) $A = \{a, b, c, d, e\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, d), (d, c), (c, a), (a, c)\}$;

в) $A = \{a, b, c, d, e\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, d), (d, c), (d, e), (e, d),$
$(b, e), (e, b), (b, d), (d, b)\}$;

г) $A = \{a, b, c, d\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, d), (d, c), (d, a), (a, d),$
$(b, d), (d, b), (a, c), (c, a)\}$.

19. Постройте граф для каждого из приведенных ниже отношений на A.

а) $A = \{a, b, c, d, e\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, d), (d, c), (d, a), (a, d)\}$;

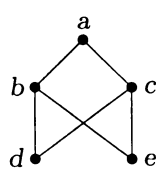
б) $A = \{a, b, c, d, e, f\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, a), (a, c), (d, e), (e, d)\}$;

в) $A = \{a, b, c, d, e\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, d), (d, c), (d, e), (e, d),$
$(a, d), (d, a)\}$;

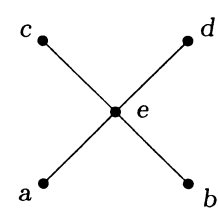
г) $A = \{a, b, c, d, e, f\}$;
$R = \{(a, b), (b, a), (b, c), (c, b), (c, a), (a, c), (a, e), (e, a),$
$(b, e), (e, b), (b, f), (f, b), (c, d), (d, c), (c, f), (f, c),$
$(d, f), (f, d), (e, f), (f, e)\}$.

20. Найдите множество вершин, ребер и соответствующее симметричное отношение для графов, изображенных на нижеследующих рисунках.

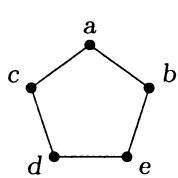
а)



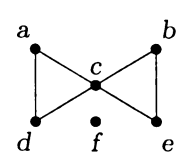
б)



в)

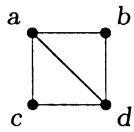


г)

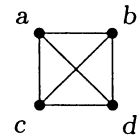


21. Найдите множество вершин, ребер и соответствующее симметричное отношение для графов, изображенных на нижеследующих рисунках.

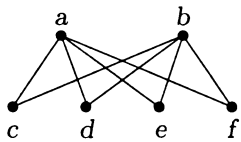
а)



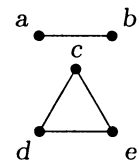
б)



в)



г)



22. Постройте орграфы со следующими свойствами:

а) Множество вершин $\{a, b, c, d, e, f\}$ и отношение R для ребер имеет вид $R = \{(a, b), (b, c), (a, c), (b, e), (c, f), (c, d), (d, f), (f, e)\}$.

б) Множество вершин $\{a, b, c, d\}$ и отношение R для ребер имеет вид $R = \{(b, c), (a, d), (b, a), (d, c), (b, d), (c, a)\}$.

в) Множество вершин $\{a, b, c, d\}$ и отношение R для ребер имеет вид $R = \{(b, a), (a, a), (b, c), (c, d), (d, c), (d, b), (d, a)\}$.

г) Множество вершин $\{a, b, c, d, e\}$ и отношение R для ребер имеет вид $R = \{(a, b), (a, c), (a, d), (c, a), (d, e), (e, d)\}$.

23. Постройте орграф со следующими свойствами:

а) Множество вершин $\{a, b, c, d, e, f\}$ и отношение R для ребер имеет вид $R = \{(a, b), (b, c), (d, c), (d, e), (f, e), (f, a), (b, e)\}$.

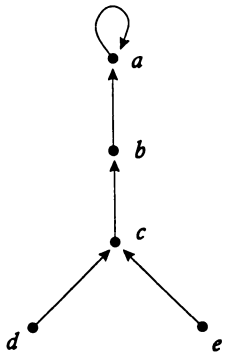
б) Множество вершин $\{a, b, c, d\}$ и отношение R для ребер имеет вид $R = \{(a, c), (a, b), (d, c), (d, b), (a, d), (b, c), (a, a), (c, c)\}$.

в) Множество вершин $\{a, b, c\}$ и отношение R для ребер имеет вид $R = \{(a, b), (a, a), (b, c), (b, b), (c, c), (c, a)\}$.

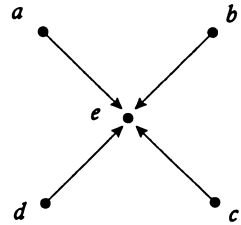
г) Множество вершин $\{a, b, c, d, e\}$ и отношение R для ребер имеет вид $R = \{(a, b), (a, c), (a, d), (c, d), (d, e), (b, e)\}$.

24. Найдите вершины и ориентированные ребра приведенных ниже орграфов.

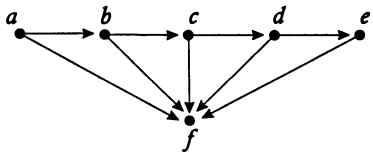
а)



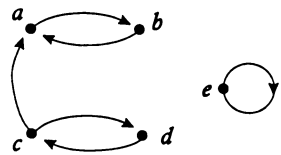
б)



в)

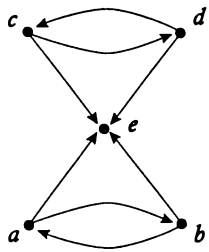


г)

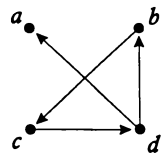


25. Найдите вершины и ориентированные ребра приведенных ниже орграфов.

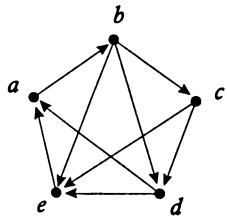
а)



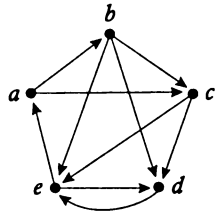
б)



в)



г)



26. Приведенное ниже доказательство предположительно показывает, что если отношение R симметрично и транзитивно, то оно и рефлексивно. Является ли это доказательство правильным? Если нет, то почему?

ДОКАЗАТЕЛЬСТВО. Пусть R — симметричное и транзитивное отношение на множестве A. Пусть $a \in A$ и $(a,b) \in R$. Поскольку $(a,b) \in R$ и R симметрично, $(b,a) \in R$. Поэтому, в силу транзитивности R, имеем $(a,a) \in R$, откуда следует, что R рефлексивно. ■

27. Докажите теорему 2.31.
28. Докажите теорему 2.37.

2.6. ЧАСТИЧНО УПОРЯДОЧЕННЫЕ МНОЖЕСТВА

При изложении оставшейся части этой главы будем предполагать известными элементарные свойства целых чисел, действительных чисел и заданных на них функций, хотя формально до настоящего времени они не обсуждались.

ОПРЕДЕЛЕНИЕ 2.44. Отношение R на A есть отношение ***частичного порядка***, если оно рефлексивно, симметрично и транзитивно. Если отношение R на A является отношением частичного порядка, то (A,R) называют ***частично упорядоченным множеством***, или ***ЧУ-множеством*** с порядком R. Если отношение порядка R предполагается по умолчанию, то (A,R) можно обозначать просто через A.

В примере 2.35 было показано, что если A — множество положительных целых чисел и отношение R определено условием $(x,y) \in R$, если x делит y нацело, то R есть отношение частичного порядка, а (A,R) — ЧУ-множество.

ПРИМЕР 2.45. Пусть $C = \{1,2,3\}$, а X — множество всех подмножеств множества C:

$$X = P(C) = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

Определим отношение R на X посредством $(T,V) \in R$, если $T \subseteq V$. Таким образом, $(\{2\},\{1,2\}) \in R$, поскольку $\{2\} \subseteq \{1,2\}$ и $(\{2,3\},\{3\}) \notin R$, поскольку $\{2,3\} \not\subseteq \{3\}$. Можно легко проверить, что R — отношение частичного порядка, а (A,R) — ЧУ-множество. □

ПРИМЕР 2.46. Пусть S — множество действительных чисел, а R_1 — отношение, определенное условием $(x,y) \in R_1$, если $x \leq y$. Легко показать, что R_1 — отношение частичного порядка, а (S,R_1) — ЧУ-множество. □

Частичное упорядочение принято обозначать через $\leq$, а частично упорядоченное множество — через $(S,\leq)$, где $\leq$ — частичный порядок на множестве S. Если $(a,b) \in \leq$, то, согласно введенным ранее обозначениям для отношений, $a \leq b$.

ОПРЕДЕЛЕНИЕ 2.47. Два элемента a и b частично упорядоченного множества $(S, \leq)$ ***сравнимы***, если $a \leq b$ или $b \leq a$. Если каждые два элемента частично упорядоченного множества $(S, \leq)$ сравнимы, то $(S, \leq)$ называется ***вполне упорядоченным множеством***, или ***цепью***.

ПРИМЕР 2.48. Пусть T — множество положительных делителей числа 30 и $\leq_1$ есть отношение $m \leq_1 n$, если m делит n нацело. Целые числа 5 и 15 сравнимы, поскольку 5 делит 15 нацело, а 5 и 6 — нет. □

ПРИМЕР 2.49. Пусть A — множество целых чисел и $R = \leq_2$ — отношение $x \leq_2 y$, если x меньше или равно y. Упорядоченное множество $(A, \leq_2)$ является цепью. □

ПРИМЕР 2.50. Пусть S — множество всех подмножеств множества $\{a, b, c\}$, а $\leq_3$ есть отношение частичного порядка, определенное в примере 2.45. Множества $\{a, b\}$, $\emptyset$ и $\{a, b, c\}$ сравнимы, но $\{a, b\}$ и $\{b, c\}$ таковыми не являются. Таким образом, $(S, \leq_3)$ — ЧУ-множество, однако цепью не является. □

Для изображения частично упорядоченных множеств имеется графический аппарат, известный как ***диаграммы Гессе***. Для заданного ЧУ-множества $(A, \leq_2)$ диаграмма Гессе состоит из совокупности точек и линий, в которой точки представляют элементы A, и если $a \leq c$ для элементов a и c множества A, тогда a помещено ниже c, и они соединены линией, если не существует такое $b \neq a, c$, что $a \leq b \leq c$. Если рассмотрение отношений ограничено отношениями частичного порядка, для них диаграммы Гессе представляют собой просто ориентированный граф, в котором петли не указаны; и если $a \leq b \leq c$, тогда линия от a к c не указана. Например, диаграмма Гессе, соответствующая множеству $(T, \leq_1)$, описанному выше, показана на рис. 2.23.

Диаграмма Гессе, представляющая $(S, \leq_3)$, показана на рис. 2.24. А на рис. 2.25 представлена диаграмма Гессе для цепи с четырьмя элементами.

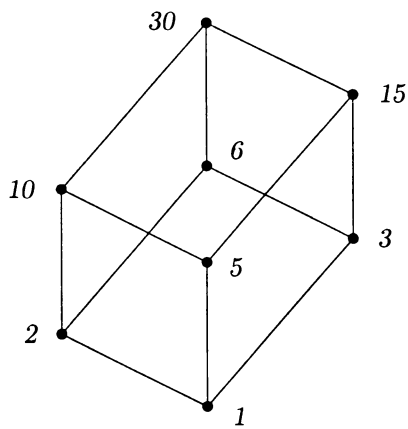


Рис. 2.23

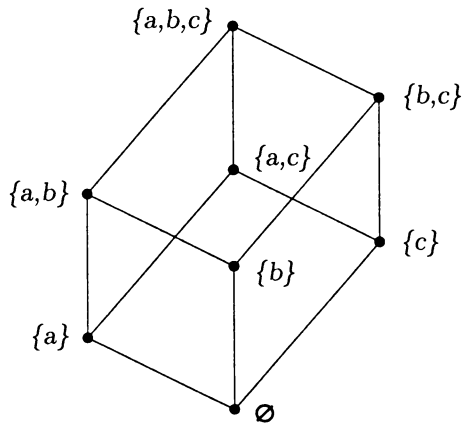


Рис. 2.24

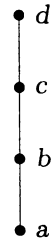


Рис. 2.25

■ УПРАЖНЕНИЯ

1. Какое из приведенных ниже отношений R является отношением частичного порядка на $A = \{a, b, c, d\}$?

а) $R = \{(a,a), (b,b), (c,c), (d,d), (a,c), (b,c), (c,d), (a,d), (b,d)\}$;

б) $R = \{(a,a), (b,b), (c,c), (d,d), (a,b), (b,c), (c,d), (d,a)\}$;

в) $R = \{(b,b), (c,c), (d,d), (a,c), (b,c), (c,d), (a,d), (b,d)\}$;

г) $R = \{(a,a), (b,b), (c,c), (d,d), (a,b), (b,c), (a,c), (a,d), (b,d), (c,d)\}$.

2. Какое из приведенных ниже отношений R является отношением частичного порядка на $A = \{a, b, c, d\}$?

а) A — множество всех людей, а отношение R определено условием: xRy, если x старше y.

б) A — множество всех граждан Соединенных Штатов, и R определено условием: xRy, если x имеет больший номер карточки социального страхования, чем y.

в) A — множество целых чисел, R определено условием: xRy, если $x \geq 2y$.

г) A — множество всех людей, и R определено условием: xRy, если x и y являются братом и сестрой.

д) A — множество всех упорядоченных пар положительных целых чисел, и $(a,b)R(c,d)$, если $a \leq c$, и если $a = c$, то $b \leq d$.

3. Постройте диаграммы Гессе для следующих ЧУ-множеств $(A, \leq)$, где

а) $A = \{a, b, c, d\}$ и
$\leq = \{(a,a), (b,b), (c,c), (d,d), (a,c), (b,c), (c,d), (a,d), (b,d)\}$;

б) $A = \{a, b, c, d\}$ и $\leq = \{(a,a), (b,b), (c,c), (d,d), (a,c), (b,c)\}$;

в) $A = \{a, b, c, d\}$ и $\leq = \{(a,a), (b,b), (c,c), (d,d)\}$;

г) $A = \{a, b, c, d\}$ и
$\leq = \{(a,a), (b,b), (c,c), (d,d), (a,b), (b,c), (a,c), (c,d), (a,d), (b,d)\}$.

4. Постройте диаграммы Гессе для следующих ЧУ-множеств $(A, \leq)$, где

а) $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ и $x \leq y$, если x делит y нацело.

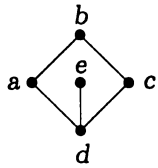
б) $A = B \times B$, где $B = \{1, 2, 3, 6, \}$. Определим $a \leq_1 b$, если a делит b нацело, и $(a, b) \leq (c, d)$, если $a \leq_1 c$, и если $a = c$, то $b \leq_1 d$.

в) A — множество положительных целых чисел, которые делят 27 нацело, и $x \leq y$, если x делит y нацело.

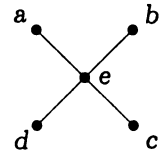
г) A — множество положительных целых чисел, которые делят 54 нацело, и $x \leq y$, если x делит y нацело.

5. Перечислите элементы множества A и выразите $\leq$ как множество упорядоченных пар для каждой из приведенных ниже диаграмм Гессе:

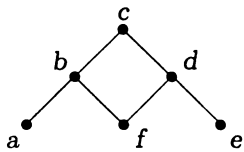
а)



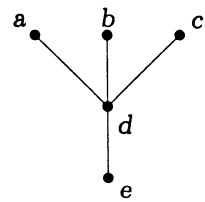
б)



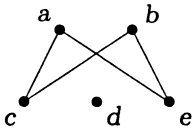
в)



г)



д)



6. Для каждого частичного порядка из предыдущего упражнения перечислите элементы:

а) не сравнимые с a;

б) не сравнимые с b;

в) не сравнимые с c;

г) не сравнимые с d.

7. Опишите симметричные отношения, которые являются отношениями частичного порядка.

2.7. ОТНОШЕНИЯ ЭКВИВАЛЕНТНОСТИ

В разделе 2.5 были определены следующие свойства бинарных отношений. Отношение R на A ***рефлексивно***, если (a, a) принадлежит R для всех a из A. Отношение R ***симметрично***, если для всех a и b из A из того, что (a, b) принадлежит R,

следует, что (b, a) принадлежит R. Отношение R ***транзитивно***, если для всех a, b и c из A таких, что (a, b) и (b, c) принадлежат R, (a, c) также принадлежит R. Эти свойства объединены в приведенном ниже определении.

ОПРЕДЕЛЕНИЕ 2.51. Отношение R на A есть ***отношение эквивалентности***, если оно рефлексивно, симметрично и транзитивно.

Пусть $A = \{1, 2, 3, 4, 5, 6\}$. В примере 2.33 уже было установлено, что отношение R_1 на A, определенное как

$$R_1 = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6), (1,2), \\ (1,4), (2,1), (2,4), (3,5), (5,3), (4,1), (4,2)\},$$

рефлексивно, транзитивно и симметрично, поэтому R_1 есть отношение эквивалентности на множестве A.

ПРИМЕР 2.52. Пусть A — множество целых чисел. Определим отношение $R_3 \subseteq A \times A$ посредством $R_3 = \{(a, b) : a - b = 5 \cdot k$ для некоторого целого числа $k\}$. Например, $(7, 2) \in R_3$, поскольку $7 - 2 = 5 = 5 \cdot 1$, и $(-11, 4) \in R_3$, так как $(-11) - 4 = -15 = 5(-3)$.

Отношение R_3 рефлексивно. Если a — целое число (т.е. $a \in A$), то $a - a = 0 = 5 \cdot 0 = 5 \cdot k$ для $k = 0$, так что $(a, a) \in R_3$.

Отношение R_3 симметрично. Допустим, $(a, b) \in R_3$. Тогда существует такое целое число m, что $a - b = 5 \cdot m$ и

$$\begin{aligned} b - a &= -(a - b) = \\ &= -(5 \cdot m) = \\ &= 5 \cdot (-m) \end{aligned}$$

для целого числа $-m$. Таким образом, $(b, a) \in R_3$.

Отношение R_3 транзитивно. Предположим, что a, b и c — целые числа, $(a, b) \in R_3$ и $(b, c) \in R_3$. По определению, если

$$(a, b) \in R_3, \text{ тогда } a - b = 5 \cdot k \text{ для некоторого целого числа } k,$$

и если

$$(b, c) \in R_3, \text{ тогда } b - c = 5 \cdot m \text{ для некоторого целого числа } m.$$

Суммирование левых и правых частей этих двух равенств дает

$$(a - b) + (b - c) = 5 \cdot k + 5 \cdot m$$

или

$$a - c = 5 \cdot (k + m)$$

для целого числа $k + m$. По определению R_3, $(a, c) \in R_3$, поэтому R_3 транзитивно. Поскольку R_3 рефлексивно, симметрично и транзитивно, оно является отношением эквивалентности. □

Отношение эквивалентности R на множестве A разбивает его на подмножества, элементы которых эквивалентны друг другу и не эквивалентны элементам других подмножеств. В контексте отношений эквивалентности эти подмножества называются ***классами эквивалентности*** по отношению R.

Это разбиение можно представлять себе следующим образом. Пусть множество A — это набор разноцветных шаров, а отношение R задается условием: $(a, b) \in R$ тогда и только тогда, когда a и b имеют одинаковый цвет. Поскольку R — отношение эквивалентности, каждый класс эквивалентности будет состоять из шаров, имеющих одинаковый цвет. Если определить отношение R условием: $(a, b) \in R$ тогда и только тогда, когда шары a и b имеют одинаковый диаметр, то каждый класс эквивалентности будет состоять из шаров одинакового размера.

ОПРЕДЕЛЕНИЕ 2.53. Пусть $a \in A$, и R — отношение эквивалентности на $A \times A$. Пусть $[a]$ обозначает множество $\{x : xRa\} = \{x : (x, a) \in R\}$, называемое ***классом эквивалентности***, содержащим a. Символ $[A]_R$ обозначает множество всех классов эквивалентности множества A по отношению R.

ПРИМЕР 2.54. Ранее было показано, что отношение R_1, определенное выше, есть отношение эквивалентности на множестве $A = \{1, 2, 3, 4, 5, 6\}$. Классы эквивалентности по отношению R_1 были получены путем определения класса эквивалентности каждого элемента множества A:

$$[1] = \{x : (x, 1) \in R_1\} = \{x : xR_1 1\} = \{1, 2, 4\},$$

где $1 \in [1]$, поскольку $(1, 1) \in R_1$, $2 \in [1]$, т.к. $(2, 1) \in R_1$, $4 \in [1]$, поскольку $(4, 1) \in R_1$, и не существует никакого иного x из A такого, что $(x, 1) \in R_1$. Точно так же, получаем

$$[2] = \{x : (x, 2) \in R_1\} = \{2, 1, 4\};$$

$$[3] = \{x : (x, 3) \in R_1\} = \{3, 5\};$$

$$[4] = \{x : (x, 4) \in R_1\} = \{4, 1, 2\};$$

$$[5] = \{x : (x, 5) \in R_1\} = \{5, 3\};$$

$$[6] = \{x : (x, 6) \in R_1\} = \{6\}.$$

Итак, имеется только три различных класса эквивалентности:

$$[1] = [2] = [4] = \{1, 2, 4\},$$
$$[3] = [5] = \{3, 5\}$$

и

$$[6] = \{6\},$$

так что

$$[A]_{R_1} = \{[1], [3], [6]\} = \{\{1, 2, 4\}, \{3, 5\}, \{6\}\}.$$

Этот пример показывает, что любой элемент класса эквивалентности порождает класс эквивалентности; другими словами, если $b \in [a]$, то $[a] = [b]$. На основании этого свойства говорят, что любой элемент класса эквивалентности представляет класс.

Каждый класс эквивалентности содержит по крайней мере один элемент, поэтому, в силу рефлексивности отношения, множество всех элементов, эквивалентных элементу a, должно содержать a. С другой стороны, никакой элемент не может принадлежать двум различным классам эквивалентности. □

ПРИМЕР 2.55. Рассмотрим отношение эквивалентности R_3 из примера 2.52. Для множества A всех целых чисел $R_3 \subseteq A \times A$ было определено посредством $R_3 = \{(a, b) : a - b = 5 \cdot k$ для некоторого целого числа $k\}$. Поскольку

$$\begin{aligned}[a] &= \{x : (x, a) \in R_3\} = \{x : xR_3a\} = \\ &= \{x : x - a = 5 \cdot k \text{ для некоторого целого числа } k\} = \\ &= \{x : x = a + 5 \cdot k \text{ для некоторого целого } k\},\end{aligned}$$

получаем, что классы

$$\begin{aligned}[0] &= \{\ldots, -15, -10, -5, 0, 5, 10, 15, 20, 25, \ldots\} = \\ &= \cdots = [-5] = [0] = [5] = [10] = [15] = \cdots,\end{aligned}$$

$$\begin{aligned}[1] &= \{\ldots, -14, -9, -4, 1, 6, 11, \ldots\} = \\ &= \cdots = [-9] = [-4] = [1] = [6] = \cdots,\end{aligned}$$

$$\begin{aligned}[2] &= \{\ldots, -13, -8, -3, 2, 7, 12, \ldots\} = \\ &= \cdots = [-3] = [-2] = [7] = [12] = \cdots,\end{aligned}$$

$$\begin{aligned}[3] &= \{\ldots, -12, -7, -2, 3, 8, 13, \ldots\} = \\ &= \cdots = [-2] = [3] = [8] = [13] = \cdots,\end{aligned}$$

$$\begin{aligned}[4] &= \{\ldots, -11, -6, -1, 4, 9, 14, \ldots\} = \\ &\cdots = [-6] = [-1] = [4] = [9] = \cdots\end{aligned}$$

представляют собой различные классы эквивалентности по отношению R_3. Таким образом,

$$[A]_{R_3} = \{[0], [1], [2], [3], [4]\}.$$

Элементы [0] "похожи" в том смысле, что каждый из них кратен пяти. Элементы любого другого класса эквивалентности также "похожи" в том смысле, что имеют один и тот же остаток при делении на пять. □

Как отмечалось выше, совокупность классов эквивалентности разделяет все множество A на непустые ***взаимоисключающие***, или ***непересекающиеся*** подмножества, в том смысле, что никакие два из них не имеют общих элементов. Такое разделение множества называется его разбиением.

ОПРЕДЕЛЕНИЕ 2.56. Пусть A и I — множества и пусть $\langle A\rangle = \{A_i : i \in I\}$, где I непусто, есть множество непустых подмножеств множества A. Множество $\langle A\rangle$ называется ***разбиением*** A, если выполнены два условия:

а) $A_i \cap A_j = \varnothing$ для всех $i \neq j$;

б) $A = \bigcup\limits_{i\in I} A_i$, т.е. a принадлежит A тогда и только тогда, когда $a \in A_i$ для некоторого $i \in I$.

ТЕОРЕМА 2.57. Непустое множество подмножеств $\langle A\rangle$ множества A есть разбиение A тогда и только тогда, когда $\langle A\rangle = [A]_R$ по некоторому отношению эквивалентности R.

ДОКАЗАТЕЛЬСТВО. Пусть $\langle A\rangle = \{A_1 : i \in I\}$ есть разбиение A. Определим отношение R на $A \times A$ таким образом: aRb тогда и только тогда, когда a и b принадлежат одному и тому же подмножеству A_i для некоторого i. Несомненно, что для всех a из A имеем aRa, поэтому R рефлексивно. Если a и b принадлежат одному подмножеству A_i, тогда b и a также принадлежат этому подмножеству A_i, поэтому R симметрично. Если элементы a и b принадлежат одному подмножеству и элементы b и c принадлежат одному подмножеству, то a и c тоже находятся в одном подмножестве, в силу условия $A_i \cap A_j = \varnothing$ для $i \neq j$. Следовательно, R транзитивно и представляет собой отношение эквивалентности.

Обратно, предположим, что R — отношение эквивалентности. Необходимо показать, что $[A]_R = \{[a] : a \in A\}$ есть разбиение A. Очевидно, $[a]$ непусто для всех a, т.к. $a \in [a]$. Очевидно также, что A есть объединение $[a]$, так что $a \in A$. Допустим, что пересечение $[a] \cap [b]$ непусто и $x \in [a] \cap [b]$. Тогда xRa и xRb, и, в силу симметричности отношения, aRx. Но поскольку aRx и xRb, то, в силу транзитивности отношения, aRb. Поэтому, $a \in [b]$. Если $y \in [a]$, то yRa, а поскольку aRb, то yRb, в силу транзитивности отношения. Поэтому $[a] \subseteq [b]$. Аналогично можно показать, что $[b] \subseteq [a]$, поэтому $[a] = [b]$ и $[A]_R$ представляет собой разбиение A. ■

ПРИМЕР 2.58. Пусть $A = \{\square, \triangle, \bigcirc, \bigstar\}$. Рассмотрим разбиение

$$A_1 = \{\square\},\ A_2 = \{\triangle, \bigcirc, \bigstar\}.$$

Согласно доказательству предыдущей теоремы, необходимо определить R таким образом: $R = \{(a, b) : a \in A_i \text{ и } b \in A_i \text{ для некоторого } i\}$. Итак,

$$R = \{(\square,\square), (\bigstar,\triangle), (\bigstar,\bigcirc), (\bigstar,\bigstar), (\bigcirc,\triangle),$$
$$(\bigcirc,\bigcirc), (\bigcirc,\bigstar), (\triangle,\triangle), (\triangle,\bigcirc), (\triangle,\bigstar)\}$$

есть отношение, соответствующее заданному разбиению. □

■ УПРАЖНЕНИЯ

1. Установите, является ли каждое из перечисленных ниже отношений на A отношением эквивалентности. Для каждого отношения эквивалентности постройте классы эквивалентности.

а) A — множество целых чисел, и R есть отношение, заданное условием: $(a, b) \in R$, если $a + b = 0$;

б) A — множество целых чисел, и R есть отношение вида $(a, b) \in R$, если $a + b = 5$;

в) A — множество высказываний, и R есть отношение вида $(p, q) \in R$, если p логически эквивалентно q;

г) A — множество упорядоченных пар целых чисел, и $(a, b)R(c, d)$, если $ad = bc$;

д) $A = \{-10, -9, -8, -7, \ldots 0, 1, \ldots, 9, 10\}$, и $(a, b) \in R$, если $a^2 = b^2$;

е) $A = \{-10, -9, -8, -7, \ldots 0, 1, \ldots, 9, 10\}$, и $(a, b) \in R$, если $a^3 = b^3$.

2. Пусть $A = \{1, 2, 3\}$. Установите, является ли каждое из приведенных ниже отношений на A отношением эквивалентности. Для каждого отношения эквивалентности постройте классы эквивалентности.

а) $R_1 = \{(2, 2), (1, 1)\}$;

б) $R_2 = \{(1, 1), (2, 2), (3, 3)\}$;

в) $R_3 = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1), (3, 1), (1, 3)\}$;

г) $R_4 = \{(1, 1), (2, 2), (3, 3), (1, 2), (3, 2), (2, 1)\}$;

д) $R_5 = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1), (2, 3), (3, 2), (1, 3), (3, 1)\}$.

3. Установите, является ли каждое из приведенных ниже отношений на A отношением эквивалентности. Для каждого из отношений эквивалентности постройте классы эквивалентности.

а) A — множество всех подмножеств множества $\{a, b, c, d\}$, отношение R определяется следующим образом: sRt, если s и t содержат одинаковое количество элементов;

б) $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, отношение R определяется следующим образом: aRb, если $a + b$ положительно;

в) $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, отношение R определяется следующим образом: by aRb, если $a + b$ четное;

г) A — множество прямых в плоскости, отношение R определяется следующим образом: nRm, если n и m пересекаются;

д) A — множество прямых в плоскости, отношение R определяется следующим образом: nRm, если n и m параллельны;

е) A — множество компьютерных программ, написанных для NASA, отношение R определяется следующим образом: pRq, если p и q написаны на одном и том же языке программирования.

4. Пусть $A = \{\square, \triangle, \bigcirc, w, t, z, h, \bigstar\}$. Отношение R на A определено, как:
$R = \{(\square, \square), (\triangle, \triangle), (\bigcirc, \bigcirc), (w, w), (t, t), (z, z), (h, h), (\bigstar, \bigstar), (\square, \bigcirc),$
$(\bigcirc, z), (w, h), (t, \bigstar), (\square, z), (z, \square), (z, \bigcirc), (h, w), (\bigstar, t), (\bigcirc, \square)\}$.

а) покажите, что R является отношением эквивалентности на A;

б) постройте все классы эквивалентности по отношению R.

5. Пусть $f(x) = x^2 + 1$, где $x \in [-2, 4] = \{x : x$ действительное число и $-2 \leq x \leq 4\} = A$. Определим отношение R на $A \times A$ следующим образом:

$$(a, b) \in R \text{ тогда и только тогда, когда } f(a) = f(b).$$

а) покажите, что R является классом эквивалентности на $A = [-2, 4]$;

б) определите классы эквивалентности:

$$[1],\ [2],\ [0],\ [3],\ [-1/2],\ \text{и}\ [4].$$

6. Для $X \neq \varnothing$ определим отношение I как $I = \{x : x = (a, a)$ для некоторого $a \in X\} = \{(a, a) \in X \times X : a \in X\}$. I представляет собой отношение эквивалентности на X и называется отношением тождества на X. Пусть R — отношение на X. Докажите, что

а) R рефлексивно тогда и только тогда, когда $I \subseteq R$;

б) R симметрично тогда и только тогда, когда $R = R^{-1}$;

в) R транзитивно тогда и только тогда, когда $R \circ R^{-1} \subseteq R$.

7. Установите, какие из приведенных ниже совокупностей элементов составляют разбиение множества $A = \{1, 2, 3, 4, 5, 6, 7\}$. Если некоторая из совокупностей не входит в разбиение, объясните, почему. Для всех совокупностей, входящих в разбиение, перечислите элементы соответствующего отношения R, и если это перечисление представляется слишком длинным, опишите множество упорядоченных пар, не перечисляя их.

а) $\{\{1, 2\}, \varnothing, \{3, 4, 5\}\{6, 7\}\}$;

б) $\{\{1, 2\}, \{3, 4, 5\}\{6, 7\}\}$;

в) $\{\{1, 7\}, \{3, 4, 6\}\}$;

г) $\{\{1, 5\}, \{3, 4, 5\}\{2, 6, 7\}\}$;

д) $\{\{1, 2, 3, 4, 5, 6, 7\}\}$.

8. Установите, какие из приведенных ниже совокупностей элементов составляют разбиение множества $A = \{1, 2, 3, 4, 5, 6, 7\}$. Если некоторая совокупность не входит в разбиение, объясните, почему. Для совокупностей, входящих в разбиение, перечислите элементы соответствующего отношения R, и если это перечисление представляется слишком длинным, опишите множество упорядоченных пар, не перечисляя их.

а) $\{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}\}$;

б) $\{\{1, 4\}, \{3, 5, 8\}\{6, 2, 7\}\}$;

в) $\{\{1, 6\}, \{3, 4\}, \{5, 2\}\}$;

г) $\{\{1, 7\}, \{3, 5\}, \{2, 4\}, \{6\}\}$;

д) $\{\{1, 2, 3, 4\}, \{4, 5, 6, 7\}\}$.

9. Пусть A — множество неотрицательных целых чисел, B — множество положительных целых чисел. Определим отношение R на $A \times B$ следующим образом: $(a, b)R(c, d)$, если $ad = bc$. Является ли это отношением эквивалентности? Если это так, докажите. Если нет, объясните почему. Какие элементы множества $A \times B$ находятся в отношении с $(1, 2)$?

ГЛАВА

3

ЛОГИКА, ЦЕЛЫЕ ЧИСЛА И ДОКАЗАТЕЛЬСТВА

3.1. ИСЧИСЛЕНИЕ ПРЕДИКАТОВ

Многие утверждения, имеющие форму высказываний, на самом деле таковыми не являются, так как содержат переменные, конкретные значения которых не указаны. Поскольку такое утверждение при одних значениях переменных может быть истинным, а при других — ложным, ему не может быть предписано истинностное значение. Такие утверждения, примерами которых являются

$$\begin{aligned} P(x): &\quad 3 + x = 5; \\ Q(x, y, z): &\quad x^2 + y^2 \geq z^2; \\ R(x, y): &\quad x^2 + y^2 \geq 0; \\ S(x): &\quad -1 \leq \sin(x) \leq 1, \end{aligned}$$

называются ***предикатами***. Предикат с одной переменной называется ***одноместным предикатом***. Предикат, имеющий две переменные, называется ***двуместным предикатом***, а предикат, содержащий n переменных, называется ***n-местным предикатом***. Предикат $P(x)$ — одноместный, поскольку имеет одну переменную. Предикат $Q(x, y, z)$ является трехместным предикатом. Предикаты становятся высказываниями, когда их переменным присваивается значение. Например,

$$P(2): \quad 3 + 2 = 5$$

является высказыванием, и это высказывание истинно.

$$P(7): \quad 3 + 7 = 5$$

тоже является высказыванием, и это высказывание ложно. Говорят, что набор значений переменных ***удовлетворяет предикату***, если на этом наборе предикат принимает значение "истина", т.е. становится истинным высказыванием. Например, 2 удовлетворяет $P(x)$, поскольку $P(2)$ истинно, а 7 не удовлетворяет $P(x)$, поскольку $P(7)$ ложно.

Некоторые предикаты истинны для каждого возможного набора значений переменных, взятых из данного множества. Среди примеров, приведенных выше, $S(x)$ — предикат $-1 \leq \sin(x) \leq 1$, где x принадлежит множеству действительных чисел. Тот факт, что для любого значения x

$$-1 \leq \sin(x) \leq 1,$$

дает основание сказать, что

$$\textit{для любого } x \textit{ истинно } S(x),$$

что символически изображается как

$$\forall x S(x)$$

или

$$\forall x(-1 \leq \sin(x) \leq 1).$$

Символ $\forall x$ называется ***универсальным квантором***, или ***квантором всеобщности***, и читается "для любого x" или "для всех x". Множество значений, которое может принимать x, называется ***универсом***, или ***предметной областью***. В частности, предметной областью может являться множество действительных чисел, множество целых чисел или другие подобные множества. Вообще говоря, истинность утверждения с квантором всеобщности зависит от предметной области. Например, высказывание

$$\forall x(x^2 > 5)$$

не истинно, если универсом (предметной областью) является множество целых чисел; однако, оно было бы истинным, если в качестве предметной области взять множество целых чисел, больших чем 2. Точно так же высказывание

$$\forall x(x^2 \geq 0)$$

истинно, если предметная область — множество действительных чисел, но ложно, если предметная область содержит комплексное число i, т.к. $i^2 = -1$.

Квантифицированный предикат

$$\forall x P(x),$$

где $P(x) : 3+x = 5$, не является истинным, если предметная область — множество целых чисел, потому что имеются значения x из этой области (например, $x = 4$) такие, что $P(x)$ ложно. Для истинности $\forall x P(x)$ $P(x)$ должно быть истинным для каждого значения x из предметной области. Таким образом, $\forall x P(x)$ имеет конкретное истинностное значение, поэтому является высказыванием.

Предикат $\forall x \forall y\, R(x, y)$ читается как "для каждого x, для каждого y имеет место $R(x, y)$" или "для каждого x и для каждого y имеет место $R(x, y)$". Предикат $\forall x \forall y\, R(x, y)$ истинный только тогда, когда для любого набора значений x и y из соответствующей предметной области $x^2 + y^2 \geq 0$ истинно. Если предметной областью для x и для y является множество действительных чисел, то

$$\forall x \forall y\, R(x, y)$$

истинно. Рассматривая приведенный выше предикат $Q(x, y, z)$: $x^2 + y^2 \geq z^2$, мы находим, что

$$(\forall x)\,(\forall y)\,(\forall z)\,Q(x, y, z)$$

не является истинным, когда предметная область для x, y и z — множество целых чисел, поскольку высказывание

$$Q(1, 2, 3)$$

ложно. Для ясности будем иногда заключать квантор в скобки, записывая $(\forall x)P(x)$ вместо $\forall xP(x)$ и $(\forall x)(\forall y)R(x, y)$ вместо $\forall x\forall yR(x, y)$.

Хотя $\forall xP(x)$, вообще говоря, не истинно, $P(x)$ истинно по крайней мере для одного значения x, а именно, $x = 2$. Поэтому существует x такой, что $P(x)$ истинно. Символически это записывается как

$$\exists xP(x)\,.$$

Символ $\exists x$ называется ***квантором существования***. Здесь, как и в случае $\forall x$, выражение $\exists x$ относится к значению x из ***предметной области***.

Высказывание

$$\exists x(x^2 = 4)$$

истинно, если предметной областью есть множество действительных чисел, т.к. $x^2 = 4$ истинно для $x = 2$. Высказывание

$$\exists x(x^2 = 5)$$

также истинно, если предметной областью есть множество действительных чисел; однако, оно не является истинным, если предметная область есть множество целых чисел, так как не существует целого числа, квадрат которого равен 5.

Поскольку для предиката $Q(x, y, z)$ высказывания $Q(1, 2, 0)$ и $Q(-3, 4, 5)$ истинны, то высказывание

$$\exists x\,\exists y\,\exists z\,Q(x, y, z)$$

истинно.

Чтобы предикат был высказыванием, все его переменные должны иметь конкретные значения или быть связаны соответствующим квантором. Например,

$$(\exists x)(\forall z)Q(x, y, z)$$

не является высказыванием, т.к. переменная y не связана никаким квантором.

Если $D(x)$ — предикат, то высказывание $\forall xD(x)$ истинно только тогда, когда высказывание $D(x)$ истинно для любого x. Когда же высказывание $\forall xD(x)$ будет ложным? Отрицание $\forall xD(x)$ может быть записано как

$$\sim\forall xD(x)\,.$$

Чтобы доказать ложность высказывания $\forall xD(x)$, достаточно найти только одно значение x, для которого $D(x)$ ложно (или, что равносильно, $\sim D(x)$ истинно). Таким образом, $\forall xD(x)$ ложно тогда и только тогда, когда

$$\exists x(\sim D(x))$$

истинно, т.е. существует такое значение x (из предметной области), что $\sim D(x)$ истинно (или $D(x)$ ложно). Например, высказывание

$$\forall x(x^2 > 0)$$

не является истинным, если предметная область — множество целых чисел, т.к. x^2 не больше, чем 0, если $x = 0$. Таким образом, $\forall x(x^2 > 0)$ ложно, потому что $\exists x(x^2 \not> 0)$ истинно. Но это означает, что $\forall x(x^2 > 0)$ ложно, потому что $\exists x(\sim(x^2 > 0))$ истинно.

Аналогично, если $G(x)$ — предикат, отрицание существования такого x, что $G(x)$ истинно, записывается в виде

$$\sim(\exists x G(x)) .$$

Очевидно, если не существует значения x, при котором $G(x)$ истинно, то $G(x)$ ложно для любого x. В этом случае $\sim G(x)$ будет истинным для всех значений x или, что равносильно,

$$\forall x(\sim G(x)) .$$

Результат проведенных рассуждений может быть сформулирован в виде тождеств:

$$\sim\forall x(D(x)) \equiv \exists x(\sim D(x)) ;$$
$$\sim\exists x(G(x)) \equiv \forall x(\sim G(x)) .$$

Для образования отрицания с навешенным квантором всеобщности нужно заменить $\forall x$ на $\exists x$ и взять отрицание предиката, следующего за квантором. А для формирования отрицания предиката с навешенным квантором существования нужно заменить $\exists x$ на $\forall x$ и взять отрицание предиката, следующего за квантором.

Отрицание высказывания, содержащего более одного квантора, осуществляется путем последовательного рассмотрения каждого квантора, начиная с первого. Например, для отрицания $(\forall x)(\forall y)R(x, y)$ имеем

$$\begin{aligned}\sim(\forall x)(\forall y)R(x, y) &\equiv \exists x\sim(\forall y)R(x, y) \equiv \\ &\equiv (\exists x)(\exists y)\sim R(x, y) .\end{aligned}$$

Аналогично,

$$\begin{aligned}\sim(\exists x)(\forall y)\,(\exists z)\,Q(x, y, z) &\equiv (\forall x)\sim(\forall y)\,(\exists z)\,Q(x, y, z) \equiv \\ &\equiv (\forall x)(\exists y)\sim(\exists z)\,Q(x, y, z) \equiv \\ &\equiv (\forall x)(\exists y)(\forall z)\sim Q(x, y, z) .\end{aligned}$$

Следовательно, для отрицания высказывания, содержащего кванторы, $\exists$ заменяется на $\forall$ и наоборот, после чего берется отрицание предиката, связанного с этой последовательностью кванторов.

Существует связь между квантифицированными предикатами и предикатами, в которых используются конкретные значения переменных. Рассмотрим, например, предикат вида

$$\forall x S(x)$$

в форме высказывания

$$\textit{для каждого } x \textit{ верно, что } -1 \leq \sin(x) \leq 1\,.$$

Предположим, что это высказывание истинно. Применим это к конкретному значению x, равному 8. Тогда

$$S(8)\,.$$

Если $\forall x S(x)$ истинно и a — конкретное значение x (из предметной области), то

$$S(a)$$

истинно. Таким образом, для случая x равно 8 истинным является

$$-1 \leq \sin(8) \leq 1\,.$$

Обратно, если b — произвольно выбранное значение из универса (т.е. единственным известным и допускаемым свойством b есть принадлежность предметной области), и мы покажем, что $S(b)$ истинно, то

$$\textit{для любого } x \textit{ имеем } S(x)$$

или

$$\forall x S(x)\,.$$

Следующие два правила устанавливают связь между предикатами с навешенным квантором всеобщности и предикатами, примененными к произвольным константам (элементам) из универса.

Из истинности $\forall x S(x)$ следует, что $S(a)$ истинно для произвольного (т.е. произвольно выбранного) элемента a из универса.

Из истинности $S(b)$ для произвольной константы b из универса следует, что $\forall x S(x)$ истинно.

Пусть дано высказывание $\exists x(x^2 + 2x - 3 = 0)$, отсюда заключаем, что существует значение a (а именно, $a = 1$) такое, что $a^2 + 2a - 3 = 0$. Обратно, если для конкретного значения a имеем $a^2 + 2a - 3 = 0$, можно сделать вывод, что $\exists x(x^2 + 2x - 3 = 0)$.

Следующие два правила дают возможность установить связь между предикатами с конкретными значениями переменных и предикатами с навешенным квантором существования.

Из истинности $\exists x S(x)$ следует, что $S(a)$ истинно для некоторого a из универса.

Из истинности $S(b)$ для некоторого b из универса следует, что $\exists x S(x)$.

Суммируя вышесказанное, получаем такие соотношения:

1) ***Универсальная конкретизация***
Из истинности $\forall x P(x)$ следует истинность $P(a)$ для произвольного a из универса.

2) ***Универсальное обобщение***
Если произвольное a из универса обеспечивает истинность $P(a)$, делаем вывод, что $\forall x(P(x)$ истинно.

3) ***Экзистенциональная конкретизация***
Из истинности $\exists x P(x)$ следует, что существует конкретное b такое, что $P(b)$ истинно.

4) ***Экзистенциональное обобщение***
Из существования конкретного c из универса, для которого $P(c)$ истинно, можно сделать вывод, что $\exists x P(x)$.

На основе этих соотношений можно доказать следующие утверждения.

ТЕОРЕМА 3.1. Для произвольных предикатов $P(x)$ и $Q(x)$, имеющих одну предметную область,

а) $\forall x(P(x) \wedge Q(x)) \equiv \forall x P(x) \wedge \forall x Q(x)$;
б) $\exists x(P(x) \vee Q(x)) \equiv \exists x P(x) \vee \exists x Q(x)$;
в) из $\forall x P(x) \vee \forall x Q(x)$ следует, что $\forall x(P(x) \vee Q(x))$;
г) из $\exists x(P(x) \wedge Q(x))$ следует, что $\exists x P(x) \wedge \exists x Q(x)$.

ДОКАЗАТЕЛЬСТВО. Чтобы доказать

$$\forall x(P(x) \wedge Q(x)) \equiv \forall x P(x) \wedge \forall x Q(x)\,,$$

сначала предположим, что $\forall x(P(x) \wedge Q(x))$, откуда следует, что $\forall x P(x) \wedge \forall x Q(x)$. Из $\forall x(P(x) \wedge Q(x))$, используя универсальную конкретизацию, заключаем, что для произвольного a имеет место $P(a) \wedge Q(a)$. Следовательно, согласно правилу специализации (логики высказываний) можно сделать вывод, что для произвольного a имеет место как $P(a)$, так и $Q(a)$. На основе универсального обобщения делаем вывод, что $\forall x P(x)$ и $\forall x Q(x)$. Таким образом, можно заключить, что $\forall x P(x) \wedge \forall x Q(x)$, в силу правила конъюнкции (логика высказываний). Обратно, предположим, что $\forall x P(x) \wedge \forall x Q(x)$. На основе правила специализации можно сделать вывод, что $\forall x P(x)$ и $\forall x Q(x)$. Из правила всеобщей конкретизации следует, что для произвольного a имеют место $P(a)$ и $Q(a)$. Следовательно, используя правило конъюнкции, можно сделать вывод, что $P(a) \wedge Q(a)$ для произвольного a. На основе универсального обобщения делаем вывод, что $\forall x(P(x) \wedge Q(x))$. Оставшуюся часть доказательства предоставляем провести читателю. ■

Если $p(x)$ — утверждение "x — студент", а утверждение $q(x)$ — "x учится усердно", то утверждение $\forall x(p(x) \rightarrow q(x))$ может быть интерпретировано как "Все студенты учатся усердно". Точно так же, если $p(x)$ — утверждение "x — слон", а $q(x)$ — утверждение "x предпочитает арахис в шоколаде", то выражение $\exists x(p(x) \wedge q(x))$ может быть сформулировано как "Некоторые слоны предпочитают арахис в шоколаде". Утверждение "Все люди смертны" логически может быть выражено как $\forall x(p(x) \rightarrow q(x))$, где $p(x)$ — утверждение "x — человек", а $q(x)$ — утверждение "x смертен". Аналогично, утверждение "Некоторые целые числа делятся на 5" может быть записано $\exists x(p(x) \wedge q(x))$, где $p(x)$ — утверждение "x — целое число", а $q(x)$ — утверждение "x делится на 5". Обычно логические выражения, содержащие $\forall x$, можно превратить в утверждения, содержащие "для

всех", а логические выражения, содержащие $\exists x$, можно обратить в утверждения, содержащие "для некоторого".

В силу такого характера перевода с языка логики на естественные языки и обратно, отрицание утверждения типа "для всех" есть утверждение типа "для некоторого", а отрицание утверждения типа "для некоторого" есть утверждение типа "для всех". Например, отрицание утверждения

Все целые числа являются простыми

есть утверждение

Некоторые целые числа не являются простыми,

а отрицание утверждения

Некоторые люди любят есть репу

есть утверждение

Все люди не любят есть репу.

Для неформальной проверки правильности умозаключений, включающих утверждения типа "для всех" и "для некоторого", используются так называемые ***диаграммы Эйлера***, которые состоят из кругов, изображающих множества. Например, утверждению "Все p есть q" соответствует диаграмма, приведенная на рис. 3.1, где круг, изображающий множество p, содержится в круге, изображающем множество q. Утверждение "Некоторые p есть q" представляется диаграммой на рис. 3.2, где мы видим, что пересечение кругов, изображающих множества p и q, непусто. Для проверки умозаключения необходимо попытаться построить диаграмму Эйлера, в которой посылки истинны, а заключение ложно. Если такое построение осуществимо, то умозаключение неправильно, если же нет, то это доказывает правильность нашего умозаключения.

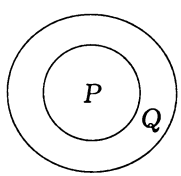


Рис. 3.1

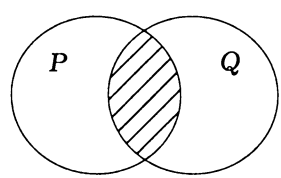
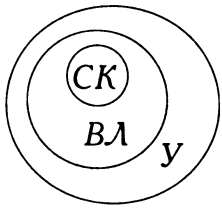


Рис. 3.2

ПРИМЕР 3.2. Рассмотрим умозаключение

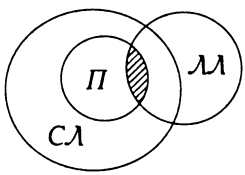
Все студенты колледжа выдающиеся
Все выдающиеся люди — ученые
$\therefore$ Все студенты колледжа — ученые



В соответствии с посылками круг, изображающий студентов колледжа (СК), должен быть внутри круга, изображающего выдающихся людей (ВЛ), который, в свою очередь, должен быть внутри круга (У), изображающего ученых. Следовательно, круг студентов колледжа должен находиться внутри круга ученых, и умозаключение является правильным. □

ПРИМЕР 3.3. Рассмотрим умозаключение

Все поэты счастливы
Некоторые поэты ленивы
∴ Некоторые ленивые люди счастливы



В соответствии с посылками круг, изображающий поэтов (П), должен быть внутри круга, изображающего счастливых людей (СЛ), а пересечение поэтов и ленивых людей (ЛЛ) должно быть непусто. Но это пересечение содержится в круге, изображающем поэтов, так что пересечение ленивых и счастливых людей непусто. Умозаключение правильно. □

ПРИМЕР 3.4. Рассмотрим умозаключение

Некоторые поэты неудачники
Некоторые атлеты неудачники
∴ Некоторые поэты являются атлетами

Соответствующая диаграмма Эйлера изображена на рис. 3.3.

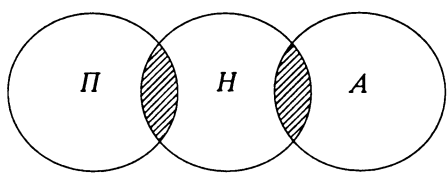


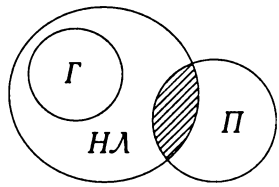
Рис. 3.3

Мы видим, что возможно построить такую диаграмму Эйлера, в которой пересечение кругов поэтов (П) и неудачников (Н) непусто и пересечение кругов атлетов (А) и неудачников непусто, так что посылки истинны, но при этом круги поэтов и атлетов не пересекаются, так что следствие не является верным. Следовательно, умозаключение не является правильным. □

ПРИМЕР 3.5. Рассмотрим умозаключение

Все гении нелогичны
Некоторые политики нелогичны
∴ Некоторые политики гении

На рисунке слева круг для гениев (Г) содержится внутри множества нелогичных людей (НЛ), а пересечение круга политиков (П) и круга нелогичных людей непусто, но круги политиков и гениев не пересекаются. Следовательно, посылки истинны, но заключение ложно. Поэтому умозаключение не является правильным. □



■ УПРАЖНЕНИЯ

1. Заданы предикаты

$$\begin{aligned} P(x,y,z): &\quad x^2+y^2 \geq z^2;\\ Q(x,y): &\quad y=(x-4)/(x+4);\\ R(x,r): &\quad |x-1| \leq r;\\ S(n,y): &\quad y=n!;\\ T(a,b,c): &\quad a \text{ нравится } b \text{ больше, чем } c. \end{aligned}$$

Запишите следующие высказывания:

а) $P(3,4,5)$; **б)** $Q(8,2)$;

в) $R(3,7)$; **г)** $S(4,24)$;

д) T(Джон, Сью, Мэри).

2. Заданы предикаты

$$\begin{aligned} P(x,y,z): &\quad x^2+y^2=z^2;\\ Q(x,y): &\quad \text{если } y^2=x^2, \text{ то } x=y;\\ R(x): &\quad \lfloor\lfloor x\rfloor\rfloor=\lfloor x\rfloor;\\ S(a,b,x): &\quad a \leq x^2 \leq b;\\ T(a,b): &\quad a \text{ играет в теннис лучше, чем } b. \end{aligned}$$

Запишите следующие высказывания:

а) $P(3,4,5)$; **б)** $Q(-2,2)$;

в) $R(3.1416)$; **г)** $S(0,4,-3)$;

д) T(Джон, Фред).

3. Используя P, Q, R, S и T из упражнения 1, запишите высказывания:

а) $(\exists x)(\exists y)P(x,y,25)$; **б)** $(\exists x)Q(x,7)$;

в) $(\forall r)(\exists x)R(x,r)$; **г)** $(\exists n)(\forall y)S(n,y)$;

д) $(\forall c)T$(Джон, Сью, c).

4. Используя P, Q, R, S и T из упражнения 1, запишите высказывания:

а) $(\forall x)(\forall y)(\exists z)P(x,y,z)$; **б)** $(\forall x)(\forall y)Q(x,y)$;

в) $(\forall x)R(x)$; **г)** $(\forall x)(\exists a)(\exists b)S(a,b,x)$;

д) $(\forall a)T(a,$Тед$)$.

5. Запишите приведенные ниже утверждения в символической форме, введя предикаты. В случае необходимости укажите предметную область.

а) На каждой улице будет праздник.

б) Некоторые машины умнее людей.

в) Любой играет в теннис лучше Фрэда.

г) Для каждого действия существует равное и противоположно направленное противодействие.

д) Каждый игрок в гольф, в конце концов, будет обыгран более сильным игроком.

6. Запишите приведенные ниже утверждения в символической форме, введя предикаты. В случае необходимости укажите предметную область.

а) Некоторые композиторы пишут симфонии лучше, чем другие.

б) Для каждого n существует три целых числа таких, что n-я степень одного из них равна сумме n-ых степеней двух других.

в) Существуют лошади, которые могут обогнать некоторых людей.

г) Он лучший атлет в мире.

д) Не существует совершенных героев.

7. Ранее было показано, что если $Q(x, y, z) : x^2 + y^2 \geq z^2$, то утверждение

$$\forall x \forall y \forall z Q(x, y, z)$$

ложно. Если универс — множество положительных целых чисел, то что можно сказать об истинности

а) $\forall x \forall y \exists z Q(x, y, z)$?

б) $\forall x \exists y \exists z Q(x, y, z)$?

в) $\exists x \forall y \exists z Q(x, y, z)$?

г) $\exists x \exists y \exists z Q(x, y, z)$?

д) $\exists x \exists y \forall z Q(x, y, z)$?

8. Пусть известно, что

$$\forall n(n \text{ нечетно } \rightarrow (\exists k)(n = 2k + 1)).$$

а) Какое из этих утверждений истинно?

$$7 \text{ нечетно } \rightarrow (\exists k)(7 = 2k + 1),$$
$$6 \text{ нечетно } \rightarrow (\exists k)(6 = 2k + 1).$$

б) Какое из этих утверждений истинно и почему?

$$(\exists k)(7 = 2k + 1),$$
$$(\exists k)(6 = 2k + 1).$$

9. Докажите, что из $\forall x\ P(x) \vee \forall x\ Q(x)$ следует, что $\forall x\ (P(x) \vee Q(x))$.

10. Докажите, что из $\exists x\ (P(x) \wedge Q(x))$ следует, что $\exists x\ P(x) \wedge \exists x\ Q(x)$.

11. Проверьте правильность умозаключений при помощи диаграмм Эйлера.

Все адвокаты богаты	Некоторые адвокаты богаты
Все богатые едят омаров	Некоторые врачи богаты
$\therefore$ Все адвокаты едят омаров	$\therefore$ Некоторые врачи — адвокаты

12. Проверьте правильность умозаключений при помощи диаграмм Эйлера.

Некоторые марсиане зеленые	Все мужчины любят мясо
Все елки зеленые	Некоторые учителя — мужчины
$\therefore$ Некоторые марсиане — елки	$\therefore$ Некоторые учителя любят мясо

13. Проверьте правильность умозаключений при помощи диаграмм Эйлера.

Все врачи любят музыку	Некоторые врачи умные
Все поэты любят музыку	Все умные люди поэты
$\therefore$ Все врачи — поэты	$\therefore$ Некоторые врачи — поэты

14. Проверьте правильность умозаключений при помощи диаграмм Эйлера.

Все машины дорогие	Все мужчины смотрят телевизор
Велосипед недорогой	Некоторые слесари — мужчины
$\therefore$ Велосипед — не машина	$\therefore$ Некоторые слесари смотрят телевизор

15. Проверьте правильность умозаключений при помощи диаграмм Эйлера.

Все люди смертны	Некоторые гуси — мужчины
Гуси — не люди	Некоторые мужчины играют в гольф
$\therefore$ Гуси бессмертны	$\therefore$ Некоторые гуси играют в гольф

3.2. ОСНОВНЫЕ ПОЛОЖЕНИЯ ТЕОРИИ ДОКАЗАТЕЛЬСТВ И ТЕОРИИ ЦЕЛЫХ ЧИСЕЛ

В данном разделе рассматриваются некоторые свойства целых чисел, которые докажут свою полезность в дальнейшем. Возможно, более важная цель — использовать множество целых чисел как хорошо известную предметную область для развития и применения на практике техники доказательства теорем. Первая проблема — с чего начать. Что можно считать известным и, следовательно, использовать в доказательстве теорем? В общем курсе по теории чисел можно было бы начать с основных аксиом теории чисел и развивать изложение в этом направлении. В нашем случае это не совсем целесообразно. Вторая проблема — решить, какие детали включать в рассмотрение. Нужно ли объяснять, почему в одном месте доказательства имеем $4 + x + 5 = y$, а в другом — $y = x + 9$? Если обращать внимание на все несущественные детали, то основные моменты доказательства затеряются в таком хаосе. Обычно предполагается, что все детали читателю понятны и в случае необходимости он может сам их воспроизвести. Не существует точных правил, которые бы говорили, о чем нужно упомянуть, а что можно опустить. Изящное изложение доказательства теоремы — искусство. В дальнейшем мы будем предполагать, что большинство аксиом, перечисленных в этом разделе, известны и могут использоваться без подробных разъяснений.

Для начала примем следующие аксиомы равенства. Они распространяются на элементы любого множества, в том числе и множества Z целых чисел.

(E1) Для любого a, $a = a$.

(E2) Для любых a и b, если $a = b$, то $b = a$.

(E3) Для любых a, b и c, если $a = b$ и $b = c$, то $a = c$.

Аксиомы E1, E2 и E3 говорят о том, что равенство на любом множестве обладает свойствами отношения эквивалентности. Эти аксиомы распространяются на все математические формы, а не только на теорию чисел.

Примем также следующие свойства сложения, вычитания и умножения на множестве Z целых чисел.

(I1) Если a и b — целые числа, то $a+b$ и $a \cdot b$ — целые числа. Множество целых чисел ***замкнуто*** относительно операций сложения и умножения.

(I2) Если $a = b$ и $c = d$, то $a + c = b + d$ и $a \cdot c = b \cdot d$.

(I3) Для любых целых чисел a и b имеют место равенства $a + b = b + a$ и $a \cdot b = b \cdot a$. Множество целых чисел ***коммутативно*** относительно операций сложения и умножения.

(I4) Для любых целых чисел a, b и c имеют место равенства $(a+b)+c = a+(b+c)$ и $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. Множество целых чисел ***ассоциативно*** относительно операций сложения и умножения.

(I5) Для любых целых чисел a, b и c имеет место равенство $a \cdot (b+c) = (a \cdot b)+(a \cdot c)$. Умножение целых чисел ***дистрибутивно*** относительно сложения.

(I6) Существуют единственные целые числа 0 и 1 такие, что $a + 0 = 0 + a = a$ и $a \cdot 1 = 1 \cdot a = a$ для любого целого числа a. Целое число 0 называется ***нейтральным элементом сложения***, или ***нулем*** множества целых чисел, целое число 1 называется ***нейтральным элементом умножения***.

(I7) Для каждого целого числа a существует единственное целое число $-a$, называемое его ***обратным элементом*** относительно сложения, такое, что $a + (-a) = (-a) + a = 0$.

(I8) Если b и c — целые числа и для некоторого ненулевого числа a имеем $a \cdot b = a \cdot c$, то $b = c$. Это утверждение называется ***мультипликативным свойством сокращения***.

Как уже отмечалось, в математике принято доказывать теоремы без излишнего формализма, поскольку предполагается, что читателю известны многие детали, поэтому нет необходимости излагать их явно. Для иллюстрации такого подхода сначала будет дано формальное доказательство того, что квадрат четного числа есть число четное. Поскольку предметная область — множество целых чисел, формально утверждение теоремы имеет вид

$$\forall n\,(n \text{ — четное} \rightarrow n^2 \text{ — четное}).$$

Формальное доказательство теоремы изложено ниже. Для доказательства истинности утверждения с навешенным квантором всеобщности одновременно используется принцип универсального обобщения.

1.	a — целое число	фиксированный элемент универса
2.	a — четное	задано (предположение)
3.	$\forall n(n$ — четное $\rightarrow (\exists k)(n = 2k))$	определение четного числа
4.	a — четное $\rightarrow (\exists k)(a = 2k)$	1, 3 и универсальная конкретизация
5.	$(\exists k)(a = 2k)$	2, 4 и правило отделения
6.	$a = 2L$ для некоторого L	5 и экзистенциональная конкретизация
7.	$\forall r \forall s(r = s \rightarrow r^2 = s^2)$	следует из аксиомы I2, известно
8.	$(a = 2L) \rightarrow a^2 = (2L)^2$	7 и универсальная конкретизация
9.	$a^2 = (2L)^2$	6, 8 и правило отделения
10.	$(2L)^2 = 2(2L^2)$	9 и свойство ассоциативности целых чисел (шаги опущены для краткости)
11.	$\forall x \forall y \forall z((x = y$ и $y = z) \rightarrow x = z)$	свойство транзитивности равенства
12.	$a^2 = (2L)^2$ и $(2L)^2 = 2(2L^2) \rightarrow a^2 = 2(2L^2)$	9, 11 и универсальная конкретизация
13.	$a^2 = (2L)^2$ и $(2L)^2 = 2(2L^2)$	9, 10 и закон конъюнкции
14.	$a^2 = 2(2L^2)$	12, 13 и правило отделения
15.	$(\exists k)(a^2 = 2k)$	14 и экзистенциональное обобщение для фиксированного значения $(2L^2)$
16.	$\forall x((\exists k)(x = 2k) \rightarrow x$ — четное)	определение четного целого числа
17.	$(\exists k)(a^2 = 2k) \rightarrow a$ — четное	16 и универсальная конкретизация для фиксированного значения a^2
18.	a^2 — четное	15, 17 и правило отделения
19.	a — четное $\rightarrow a^2$ — четное	2, 18 и то, что $((p \wedge q) \rightarrow (p \rightarrow q)$ есть тавтология)
20.	$\forall n(n$ — четное $\rightarrow n^2$ — четное)	1, 19 и универсальное обобщение

Приведенные выше рассуждения выписаны более формально, чем это принято в курсах математической логики. Этот пример иллюстрирует, как в конкретном случае взаимодействуют законы логики, кванторы, правильные умозаключения и вывод. Более типичное для данной книги математическое доказательство приведено ниже.

ТЕОРЕМА 3.6. Если n четно, то n^2 тоже четно.

ДОКАЗАТЕЛЬСТВО. Предположим, что n — (произвольное) четное целое число. По определению четного целого числа существует целое число L такое, что

$$n = 2L\,.$$

Если целые числа равны, то равны квадраты этих чисел, так что

$$n^2 = (2L)^2\,.$$

Но

$$(2L)^2 = (2L)(2L) = 2(2L^2)\,,$$

поэтому $n^2 = 2(2L^2)$ и $n^2 = 2J$ для некоторого целого числа J (а именно $J = 2L^2$). По определению четного целого числа n^2 — четное число. ■

В качестве другого примера рассмотрим утверждение, обратное предыдущему.

ТЕОРЕМА 3.7. Если n^2 четно, то и n четно.

ДОКАЗАТЕЛЬСТВО. Доказательство этой теоремы будет примером ***доказательства методом контрапозиции*** вместо прямого доказательства самой теоремы. Этот метод доказательства возможен, т.к.

$$p \to q \equiv \sim q \to \sim p\,.$$

Это означает, что доказательство $\sim q \to \sim p$ эквивалентно доказательству $p \to q$. Утверждение

$$\sim q \to \sim p$$

означает следующее:

Если n не является четным, то n^2 тоже не является четным.

Пусть n — целое число, которое не является четным (напомним, что четные целые числа имеют вид $2K$, а нечетные целые числа имеют вид $2L + 1$, и целое число является либо четным, либо нечетным, но не может быть и тем, и другим одновременно). Необходимо показать, что число n^2 — нечетное. Поскольку n нечетное, существует целое число L такое, что

$$n = 2L + 1\,.$$

Возводя в квадрат обе части соотношения, получаем

$$\begin{aligned} n^2 &= (2L+1)^2 = \\ &= 4L^2 + 4L + 1 = \\ &= 2(2L^2 + 2L) + 1\,. \end{aligned}$$

Так что, если $J = 2L^2 + 2L$, имеем

$$n^2 = 2J + 1\,.$$

По определению нечетного числа число n^2 нечетно, и теорема доказана. ■

Доказательство следующей теоремы предоставляем читателю.

ТЕОРЕМА 3.8. Пусть a, b и c — целые числа.

а) Если $b + a = c + a$, то $b = c$.
б) Для любого целого числа a имеет место равенство $a \cdot 0 = 0$.
в) Для любого целого числа a имеет место равенство $-(-a) = a$.
г) $a \cdot (-b) = -(ab)$.
д) $(-a) \cdot (-b) = a \cdot b$.
е) $-(a + b) = (-a) + (-b)$.

Множество целых чисел Z содержит подмножество N положительных целых чисел.[1] Примем в качестве аксиом следующие утверждения относительно множества N.

(N1) Целое число 1 есть положительное целое число.

(N2) Множество положительных целых чисел замкнуто относительно сложения и умножения, т.е. если a и b — положительные целые числа, то $a+b$ и $a \cdot b$ — положительные целые числа.

(N3) ***(Аксиома трихотомии)*** Для каждого целого числа a истинным является одно и только одно из перечисленных ниже утверждений:

а) a — положительное целое число;
б) $a = 0$;
в) $-a$ — положительное целое число.

Далее приведено частичное упорядочение целых чисел, в котором целые числа расположены по принципу "большее за меньшим". Обычно целые числа "упорядочены" следующим образом:

$$\ldots, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \ldots,$$

где все положительные числа "идут за" нулем, а нуль "идет за" или "больше,чем" любое отрицательное число. Трихотомия позволяет установить, является ли число положительным, поэтому отношение $>$ на множестве $Z \times Z$ можно определить так: $a > b$ тогда и только тогда, когда $a-b$ положительно. В силу трихотомии, $a-b$ либо положительно, либо нуль, либо отрицательно. Поэтому можно объединить две из этих возможностей в одну и говорить, что $a \geq b$ тогда и только тогда, когда $a > b$ или $a = b$. Сказанное выше формализуем в виде следующего определения.

ОПРЕДЕЛЕНИЕ 3.9. Для целых чисел a и b имеем: $a > b$ тогда и только тогда, когда $a - b$ положительно; $a \geq b$ тогда и только тогда, когда $a > b$ или $a = b$. Кроме того, $b < a$ равносильно $a > b$ и $b \leq a$ равносильно $a \geq b$.

Очевидно, $a > 0$ равносильно утверждению, что a положительное, потому что $a > 0$ тогда и только тогда, когда $a - 0 = a$ положительное. Аналогично, $a < 0$ равносильно утверждению, что a отрицательно.

Хотя приведенные ниже утверждения хорошо известны и могут рассматриваться как предположения, они приведены для практики доказательства теорем.

ТЕОРЕМА 3.10. Для целых чисел a и b:

а) $(a \geq b) \wedge (b \geq a) \rightarrow a = b$;
б) $(a > b) \wedge (b > c) \rightarrow a > c$.

ДОКАЗАТЕЛЬСТВО. а) Здесь мы используем принцип сведения к абсурду. Если, предположив, что утверждение в части (а) ложно, мы получим противоречие, то

[1] В отечественной литературе положительные целые числа принято называть числами натуральными. — *Прим. ред.*

истинность этого утверждения будет доказана. Предположим, что утверждение в части (а) ложно. Тогда, если $a \geq b$, $b \geq a$ и $a \neq b$, то $a > b$ и $b > a$. Но в этом случае $a - b$ положительно и $b - a = -(a - b)$ положительно, что противоречит аксиоме трихотомии N3. Так как мы пришли к противоречию, должен иметь место случай $a = b$. б) Если $a > b$ и $b > c$, то $a-b$ и $b-c$ положительны. Следовательно, согласно N2 число $(a - b) + (b - c) = a - c$ положительно, и $a > c$. ■

ТЕОРЕМА 3.11. Пусть a, b, c и d — целые числа. Тогда

а) $(a > b) \wedge (c > d) \rightarrow a + c > b + d$;
б) $(a > 0) \wedge (c > d) \rightarrow ac > ad$;
в) $(a > b > 0) \wedge (c > d > 0) \rightarrow ac > bd$;
г) $(a \geq b \geq 0) \wedge (c \geq d \geq 0) \rightarrow ac \geq bd$.

ДОКАЗАТЕЛЬСТВО. а) Если $a > b$ и $c > d$, то числа $a-b$ и $c-d$ положительны. Следовательно, $(a-b)+(c-d)$ положительно. Но $(a-b)+(c-d) = (a+c)-(b+d)$, и поэтому $a + c > c + d$. б) Если $c > d$, то разность $c - d$ положительна. Поскольку a положительно, $a(c - d) = ac - ad$ также положительно, и $ac > ad$. в) Если $a > b > 0$ и $c > d > 0$, то $ac > bc$ и $bc > bd$. Поэтому $ac > bd$. г) Доказательство этой части предоставляется читателю. ■

Далее иллюстрируется применение ***доказательства перебором***. Выбор из трех возможностей обычно имеет вид

$$\begin{array}{l} p \\ p \rightarrow (r \vee s \vee t) \\ r \rightarrow q \\ s \rightarrow q \\ t \rightarrow q \\ \hline \therefore\ q \end{array}$$

ТЕОРЕМА 3.12. Пусть n — целое число. Тогда $n^2 \geq 0$.

ДОКАЗАТЕЛЬСТВО. Поскольку n — целое число, согласно аксиоме трихотомии оно либо положительно, либо отрицательно, либо равно 0. Если n положительно, то $n \cdot n$ положительно по аксиоме N2, и $n^2 \geq 0$. Если n отрицательно, то $n = -m$ для некоторого положительного целого числа m. Поэтому $n^2 = (-m)(-m) = m^2$ опять является положительным, и $n^2 \geq 0$. Если $n = 0$, то $n^2 = 0$, так что $n^2 \geq 0$. Следовательно, $n^2 \geq 0$. ■

■ УПРАЖНЕНИЯ

1. Пусть a, b и c — целые числа. Докажите, что
а) если $b + a = c + a$, то $b = c$;
б) $a \cdot 0 = 0$ для любого целого числа a;
в) $-(-a) = a$ для любого целого числа a;
г) $a \cdot (-b) = -(ab)$;
д) $(-a) \cdot (-b) = a \cdot b$;
е) $-(a + b) = (-a) + (-b)$.

2. Докажите, что если $a + b = a$, то $b = 0$.
3. Докажите, что если $a \geq b \geq 0$ и $c \geq d \geq 0$, то $ac \geq bd$.
4. Докажите, что $a^2 \geq a$ для любого целого числа a.
5. Докажите неравенство $a^2 + b^2 \geq 2ab$ для целых чисел a и b.

3.3. МАТЕМАТИЧЕСКАЯ ИНДУКЦИЯ

Одна из наиболее сложных проблем теории чисел состоит в доказательстве истинности утверждений для всех положительных целых чисел. Если утверждение истинно для первых десяти положительных чисел или даже для первых десяти биллионов положительных чисел, то из этого еще не следует, что утверждение истинно для всех положительных целых чисел. Обычно легко доказать, что утверждение не является истинным для всех положительных целых чисел. В самом деле, используя тот факт, что отрицание $(\forall x)P(x)$ есть $(\exists x)\sim P(x)$, необходимо найти только одно положительное целое число, для которого утверждение $P(x)$ не является истинным. Эта процедура известна как нахождение ***контрпримера***. Инструментом для доказательства истинности утверждений относительно всех положительных чисел является принцип индукции, который будет нашей последней аксиомой для множества положительных чисел N. Это вовсе не означает, что доказательство любого утверждения относительно положительных чисел требует индукции. Многие теоремы могут быть доказаны на основе других теорем и аксиом. Тем не менее, в основе доказательств многих рассматриваемых ниже теорем лежит принцип индукции. В дальнейшем принцип индукции будет сформулирован в другой, эквивалентной форме.

(N4) (Принцип математической индукции) Пусть $P(n)$ есть такое утверждение, что

а) $P(1)$ истинно, и
б) для каждого k, если $P(k)$ истинно, то $P(k+1)$ истинно.

Тогда $P(n)$ истинно для любого целого положительного числа n.

В символической записи принцип математической индукции имеет вид

$$(P(1) \wedge ((\forall k)\, P(k) \rightarrow P(k+1)) \rightarrow (\forall n)\, P(n)\,.$$

Математическую индукцию можно сравнить с бесконечным рядом костяшек домино. Свойство выстроенных в ряд домино состоит в том, что если опрокинуть одну костяшку, падает следующая. Пусть утверждение $P(n)$ состоит в том, что падает n-я костяшка. Опрокидывание первой костяшки домино обозначим $P(1)$, т.е. истинность $P(1)$ состоит в том, что первая костяшка падает. Поскольку каждая костяшка опрокидывает следующую, то из истинности $P(k)$ следует истинность $P(k+1)$. Поскольку падают все костяшки, $P(n)$ истинно для каждого положительного целого числа n. В дальнейшем, проводя доказательство при помощи описанного выше принципа математической индукции, мы будем просто говорить, что доказываем утверждение по индукции.

ПРИМЕР 3.13. Допустим, необходимо найти формулу для вычисления суммы первых n положительных целых чисел, $1+2+3+\cdots+n$. Такая формула была бы полезна, если надо было бы, например, найти сумму первых 100 000 положительных целых чисел, не выполняя для этого 99 999 операций сложения. Предположим, у нас имеются основания полагать, что сумма первых n положительных чисел равна $\frac{n(n+1)}{2}$. Докажем это утверждение по индукции.

Пусть S_n есть утверждение

$$1+2+3+\cdots+n=\frac{n(n+1)}{2}. \tag{1}$$

Сначала покажем, что S_1 истинно. При $n=1$ левая часть равенства $1+2+3+\cdots+n$ равна 1. Справа получаем

$$\frac{n(n+1)}{2}=\frac{1(1+1)}{2}=1,$$

так что S_1 истинно. Затем предполагаем, что для любого положительного числа k S_k истинно, т.е. равенство (1) истинно для $n=k$. Это означает, что

$$1+2+3+\cdots+k=\frac{k(k+1)}{2}.$$

Теперь необходимо доказать, что S_{k+1} истинно, т.е. равенство S_n истинно для $n=k+1$. Другими словами, требуется доказать, что

$$1+2+3+\cdots+k+(k+1)=\frac{(k+1)((k+1)+1)}{2}.$$

Используя предположение

$$1+2+3+\cdots+k=\frac{k(k+1)}{2}$$

и сравнивая это равенство с предыдущим, мы замечаем, что, добавив $k+1$ к обеим частям последнего равенства, получим левую часть искомого равенства:

$$\begin{aligned}1+2+3+\cdots+k&=\frac{k(k+1)}{2};\\ 1+2+3+\cdots+k+(k+1)&=\frac{k(k+1)}{2}+(k+1)=\\ &=(k+1)(\frac{k}{2}+1)=\\ &=\frac{(k+1)((k+1)+1)}{2}.\end{aligned}$$

Итак, доказано равенство

$$1+2+3+\cdots+k+(k+1)=\frac{(k+1)((k+1)+1)}{2},$$

которое означает, что соотношение (1) справедливо для $n=k+1$. Этот результат обеспечивает истинность S_{k+1}. Следовательно, в силу аксиомы N4, S_n истинно для любого положительного целого числа n. □

ПРИМЕР 3.14. Требуется доказать, что

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}. \qquad (2)$$

Пусть S_n обозначает утверждение, что равенство в уравнении (2) выполняется. Сначала покажем, что утверждение S_n истинно для $n = 1$. Подставляя 1 в обе части равенства, получаем

$$1^2 = \frac{1(1+1)(2\cdot 1+1)}{6},$$

что дает нам $1 = 1$ и обеспечивает истинность утверждения S_n при $n = 1$.

Затем мы предполагаем, что равенство S_n истинно для $n = k$. То есть,

$$1^2 + 2^2 + 3^2 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}.$$

Теперь требуется доказать, что равенство S_n истинно для $n = k+1$. Иначе говоря, требуется доказать, что

$$1 + 2 + 3 + \cdots + k^2 + (k+1)^2 = \frac{(k+1)((k+1)+1)(2(k+1)+1)}{6} = \\ = \frac{(k+1)(k+2)(2k+3)}{6}.$$

Используя предположение

$$1 + 2 + 3 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6},$$

замечаем, что, прибавив к обеим частям последнего равенства $(k+1)^2$, мы получим левую часть искомого равенства. Таким образом,

$$1 + 2 + 3 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6};$$

$$1 + 2 + 3 + \cdots + k^2 + (k+1)^2 = \frac{k(k+1)(2k+1)}{6} + (k+1)^2 = \\ = \frac{k(k+1)(2k+1)}{6} + \frac{6(k+1)^2}{6} = \\ = \frac{k(k+1)(2k+1) + 6(k+1)^2}{6} = \\ = \frac{(k+1)\left[k(2k+1) + 6(k+1)\right]}{6} = \\ = \frac{(k+1)(2k^2+7k+6)}{6} = \\ = \frac{(k+1)(2k+3)(k+2)}{6}.$$

Итак, доказано равенство

$$1+2+3+\cdots+k^2+(k+1)^2=\frac{(k+1)(k+2)(2k+3)}{6},$$

которое означает, что равенство S_n выполняется для $n=k+1$. Следовательно, S_n истинно для каждого положительного целого числа n. □

ПРИМЕР 3.15. Требуется показать, что для любого положительного целого числа n, число n^3-n делится на 3. (Заметим, что положительное целое число t делится на 3 при условии, что существует положительное целое число m такое, что $t=3m$.) Доказательство будет проведено индукцией. Пусть S_n — утверждение "n^3-n делится на 3". Покажем, что S_1 истинно. Для $n=1$ имеем $n^3-n=1^3-1=1-1=0$, поэтому n^3-n при $n=1$ делится на 3, т.к. $0=3\cdot 0$. Таким образом, S_1 истинно.

Теперь покажем, что для любого положительного числа k, если S_k истинно, то S_{k+1} истинно. Допустим, что S_k истинно. Иначе говоря, k^3-k делится 3, поэтому $k^3-k=3m$ для некоторого положительного целого числа m. Требуется доказать, что S_{k+1} истинно. Последнее означает, что $(k+1)^3-(k+1)$ делится на 3, т.е. существует положительное целое число w такое, что $(k+1)^3-(k+1)=3w$. Обратим внимание на положительное целое число $(k+1)^3-(k+1)$. В данном случае стратегия состоит в том, чтобы преобразовать $(k+1)^3-(k+1)$ таким образом, чтобы была возможность использовать наши знания о k (а именно, что $k^3-k=3m$). Итак,

$$\begin{aligned}(k+1)^3-(k+1)&=(k^3+3k^2+3k+1)-(k+1)=\\&=(k^3-k)+(3k^2+3k)=\\&=3\cdot m+3(k^2+k)=\\&=3\cdot w,\end{aligned}$$

где $w=m+(k^2+k)$. Таким образом, $(k+1)^3-(k+1)$ делится на 3, и, следовательно, S_{k+1} истинно. Поэтому, по индукции, S_n истинно для всех n. □

ПРИМЕР 3.16. Требуется показать, что плоскость, разделенная на области любым числом различных прямых линий, может быть раскрашена черной и белой краской таким образом, что любые две области, имеющие общую границу, будут окрашены в различные цвета, как изображено на рис. 3.4.

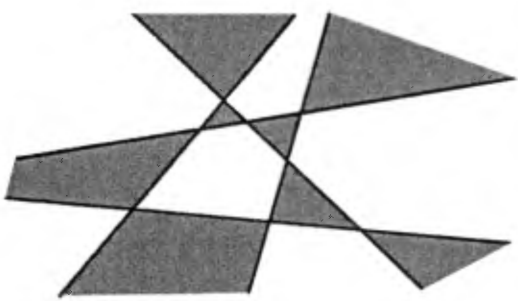


Рис. 3.4. Закрашенные области

Поскольку в формулировке задачи спрашивается о всех возможных случаях расположения областей и линий, это приводит к бесконечно большому числу возможностей. Принцип индукции является наиболее подходящим для решения задач указанного типа. У нас имеется выбор: либо показать, что любые m таких областей может быть раскрашено указанным образом, либо показать, что любое число областей, образованных при помощи m прямых линий, может быть раскрашено нужным образом. В любом из этих случаев покажем, что чертеж может быть раскрашен нужным образом. В первом случае скажем, что доказательство индукцией проводится по числу областей. Во втором случае мы скажем, что доказательство индукцией проводится по числу прямых линий. Далее будем использовать число линий, имеющихся на чертеже.

Пусть $P(m)$ — утверждение, что чертеж, выполненный с использованием m линий, может быть раскрашен требуемым образом (т.е. при помощи черного и белого цветов, и при этом никакие соседние области, имеющие общий участок границы, не будут окрашены в один и тот же цвет).

Выбираем

$$M = \{m : P(m) \text{ истинно}\} =$$
$$= \{m : \text{чертеж, выполненный } m \text{ линиями, может быть раскрашен}\}.$$

Таким образом, если $M = N$, то чертеж, образованный любым числом линий, может быть требуемым образом раскрашен. $M \subseteq N$ по определению.

(а) Покажем, что $1 \in M$. Это равносильно тому, что $P(1)$ истинно, т.е. чертеж, сделанный при помощи только одной линии, может быть раскрашен. Из геометрии известно, что линия, проведенная в плоскости, делит ее на две области. Окрасим одну из них в черный цвет, а другую — в белый, как показано на рис. 3.5. Таким образом, показано, что $1 \in M$.

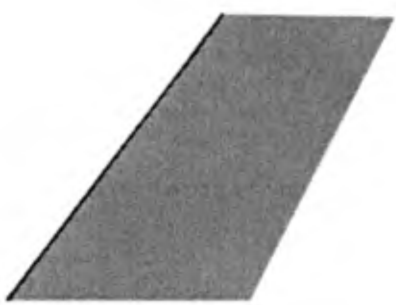


Рис. 3.5. Раскрашено в соответствии с наличием одной линии в плоскости

(б) Покажем, из $m \in M$ следует $m + 1 \in M$. Предположим, что m — положительное целое число из M. Это означает, что $P(m)$ истинно, и любой чертеж, выполненный при помощи m линий, может быть нужным образом раскрашен. Рассмотрим положительное целое число $m + 1$. Необходимо показать, что $P(m + 1)$ истинно, или, что эквивалентно, чертеж, выполненный с использованием $m + 1$ линии, может быть нужным образом раскрашен. Для этого можно использовать "индуктивную гипотезу", а именно: предположение о том, что любой чертеж, выполненный с использованием m линий, может быть нужным образом раскрашен.

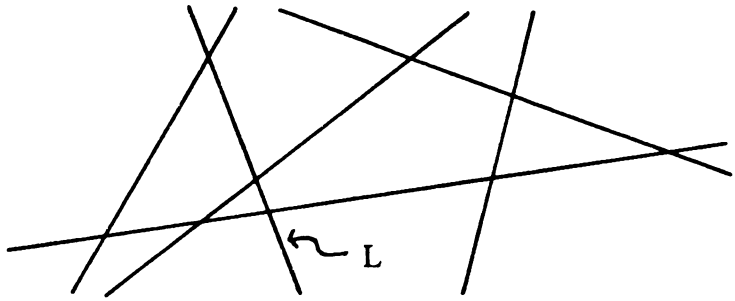


Рис. 3.6. Множество линий

Рассмотрим чертеж, выполненный с использованием $m+1$ линии, как изображено на рис. 3.6.

Теперь нам нужно создать ситуацию, в которой мы сможем использовать тот факт, что $m \in M$, или что любой чертеж, выполненный с использованием m линий, может быть раскрашен нужным образом. Для этого выберем одну из линий, например, линию L, и удалим ее. Полученный в результате чертеж содержит только m линий, поэтому в силу индуктивного предположения, что любой чертеж, состоящий из m линий, может быть требуемым образом раскрашен, можно нужным образом раскрасить чертеж на рис. 3.7.

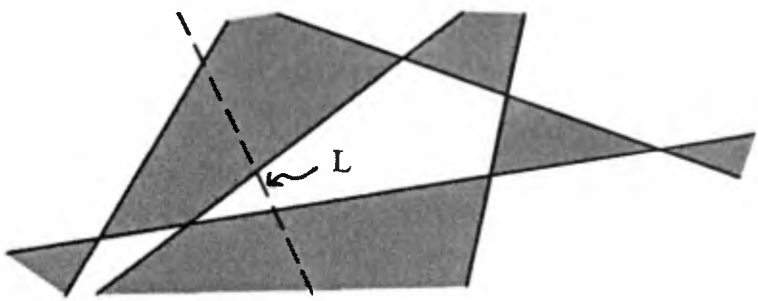


Рис. 3.7. Раскрашено в соответствии с наличием m линий

Теперь вернем линию L на прежнее место и изменим цвет каждой области с одной стороны линии L на противоположный (черные области станут белыми, а белые — черными), как показано на рис. 3.8. Этот чертеж раскрашен требуемым образом, т.к. каждые две соседние области, которые находились по ту сторону от L, где цвета были изменены, имели различные цвета по разные стороны линии, поэтому если цвета были изменены на противоположные, то цвета по разные стороны любой такой границы остались различными. Изменения вносят только те границы, которые находятся вдоль линии L, но если L делила область, то

цвет был изменен по обе стороны от L, делая цвета различными по обе стороны от такой границы. Таким образом, этот чертеж, выполненный с использованием $m+1$ линии, может быть требуемым образом раскрашен. Полученный результат означает, что $P(m+1)$ истинно, т.е. $m+1 \in M$.

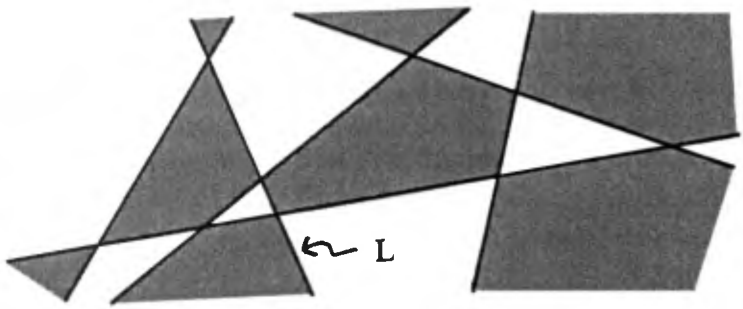


Рис. 3.8. Раскрашено в соответствии с наличием m + 1 линии

Итак, мы доказали части (а) и (б) принципа индукции. Таким образом, $M = N$. В таких случаях, как правило, говорят, что индукцией установлено равенство M и N. Таким образом, чертеж, выполненный с использованием любого (конечного) числа линий, может быть раскрашен требуемым образом. □

Существует два других эквивалентных определения принципа математической индукции. Они эквивалентны исходному определению в том смысле, что любая из трех форм может быть использована для доказательства двух других. Рассмотрим множество $A = \{3, 7, 8\}$. Целое число 3 обладает в отношении множества A уникальным свойством, а именно: число 3 принадлежит A и оно меньше или равно каждому элементу множества A. При этом говорят, что 3 — наименьшее целое число множества A. Не каждое множество целых чисел имеет наименьшее целое число. Пустое множество $\varnothing$ не имеет наименьшего целого числа, поскольку оно пусто и не содержит целых чисел вообще. Множество

$$C = \{n : n \in Z \text{ и } n = 5k \text{ для некоторого целого числа } k\} = \\ = \{\ldots, -15, -10, -5, 0, 5, 10, 15, 20, \ldots\}$$

не имеет наименьшего целого числа, т.к., если m — наименьшее целое число множества C, то $m = 5 \cdot k$ для некоторого целого числа k. Но целое число $s = 5(k-1)$ также принадлежит C и, естественно, $s = 5(k-1) < 5 \cdot k = m$, так что $s < m$, и это противоречит предположению о том, что m — наименьшее целое число множества C.

ОПРЕДЕЛЕНИЕ 3.17. Пусть A — множество целых чисел. Целое число a есть ***наименьшее целое число*** множества A тогда и только тогда, когда
а) a принадлежит A, и
б) если b принадлежит A, то $a \leq b$.

ТЕОРЕМА 3.18. Если множество целых чисел имеет наименьшее целое число, то такое число единственно.

ДОКАЗАТЕЛЬСТВО. Допустим, что числа a и b оба являются наименьшими элементами множества A. По определению наименьшего элемента $a \leq b$ и $b \leq a$. Следовательно, $a = b$. ■

Теперь сформулируем два утверждения, эквивалентные принципу индукции.

(N5) ***(Принцип полного упорядочения)*** Каждое непустое множество положительных целых чисел содержит наименьший элемент.

(N6) ***(Второй принцип индукции)*** Пусть $P(n)$ — утверждение. Если

а) $P(1)$ истинно, и

б) для произвольного k из истинности $P(m)$ для всех $m < k$ следует истинность $P(k)$,

то $P(n)$ истинно для всех положительных целых чисел n.

В символической форме принцип индукции имеет вид

$$(P(1) \wedge ((\forall k)(\forall m < k)P(m)) \rightarrow P(k)) \rightarrow (\forall n)\, P(n)\,.$$

Пример использования второго принципа индукции N6 будет приведен ниже при доказательстве того, что каждое целое число представляет собой произведение простых чисел. Следующая теорема — пример утверждения, которое может показаться слишком очевидным для проведения его доказательства, но поскольку рассматриваемая модель целых чисел определена совокупностью правил, для наших "целых чисел" эти утверждения могут не оказаться столь очевидными или даже истинными. Мы можем установить это, доказывая соответствующую теорему с использованием имеющихся правил. Доказательство является хорошим примером использования принципа полного упорядочения N5. Оно является также примером использования принципа сведения к абсурду.

ТЕОРЕМА 3.19. Не существует такого положительного целого a, что $0 < a < 1$.

ДОКАЗАТЕЛЬСТВО. Предположим, что такое число существует, и пусть S — множество всех положительных целых чисел между 0 и 1. Поскольку S непусто, по аксиоме N5 множество S имеет наименьшее положительное целое число, скажем, a. Умножая неравенство $0 < a < 1$ на a, получаем $0 < a^2 < a$. Теперь, поскольку $a < 1$, имеем $a^2 < 1$. Но это противоречит тому факту, что a — наименьшее положительное целое число, которое меньше 1. Следовательно, S — пустое множество, и не существует положительное целое число между 0 и 1. ■

СЛЕДСТВИЕ 3.20. Если b — целое число, то не существует целое число a, обладающее свойством $b < a < b + 1$.

Аксиомы N4 и N6 сформулированы для целых чисел, начиная с 1, с продолжением условий на целые числа, большие 1. Аналогичные теоремы индукции верны для целых чисел, не меньших некоторого целого j, которое может не совпадать с 1. Следующая теорема служит обоснованием индукции по подмножеству целых, не обязательно положительных, чисел.

ТЕОРЕМА 3.21. (Принцип индукции для целых чисел) Пусть $P(n)$ — высказывание, обладающее свойствами

а) $P(j)$ истинно и

б) для произвольного $k \geq j$, если $P(k)$ истинно, то $P(k+1)$ истинно.

Тогда $P(n)$ истинно для каждого $n \geq j$.

ДОКАЗАТЕЛЬСТВО. Доказательство следует из аксиомы N4. Для этого достаточно положить $n = t - j + 1$. ■

Далее в примере доказывается утверждение, включающее неравенство. Часто в такой ситуации приходится использовать следующее свойство, сформулированное в теореме 3.11:

$$(a > b > 0) \wedge (c > d > 0) \rightarrow ac > bd.$$

ПРИМЕР 3.22. Используя принцип индукции для целых чисел, нужно доказать, что для любого целого числа $n \geq 4$ имеет место неравенство $n! > 2^n$. При использовании индукции по n, в данном случае нельзя начинать с $n = 1$, поскольку утверждение для $n = 1$ неверно. Начальной точкой должно быть $n = 4$, т.к. утверждение неверно также для $n = 2$ и $n = 3$. Согласно обозначениям, введенным для принципа индукции для целых чисел, имеем: $j = 4$, $P(n)$ — утверждение "$n! > 2^n$". Сначала докажем утверждение для $n = 4$. При $n = 4$ имеем $4! = 24$ и $2^4 = 16$, так что $4! > 2^4$.

По индуктивному предположению имеем $k! > 2^k$. Нам нужно доказать, что $(k+1)! > 2^{k+1}$. Заметим, что желаемый результат можно получить, если умножить левую часть неравенства на $(k+1)$, а правую — на 2. Поэтому, если мы покажем, что $(k+1) > 2$, то получим $k! > 2^k$ и $(k+1) > 2$ и сможем сделать вывод, что $(k+1)! > 2^{k+1}$. Поскольку $k \geq 4$, то $k > 2$. Следовательно, $(k+1) \cdot k! > 2 \cdot 2^k$ и $(k+1)! > 2^{k+1}$. Таким образом, $n! > 2^n$ для каждого $n \geq 4$. □

■ УПРАЖНЕНИЯ

1. Докажите, что $1 + 4 + 7 + 10 + \cdots + (3n-2) = \dfrac{n(3n-1)}{2}$.

2. Докажите, что $\dfrac{1}{1 \cdot 3} + \dfrac{1}{3 \cdot 5} + \dfrac{1}{5 \cdot 7} + \dfrac{1}{7 \cdot 9} + \cdots + \dfrac{1}{(2n-1)(2n+1)} = \dfrac{n}{2n+1}$.

3. Докажите, что $1 + 2 + 2^2 + 2^3 + \cdots + 2^{n-1} = 2^n - 1$.

4. Докажите, что $1 + r + r^2 + r^3 + \cdots + r^{n-1} = \dfrac{1 - r^n}{1 - r}$.

5. Докажите, что $1 + 3 + 5 + 7 + \cdots + (2n-1) = n^2$.

6. Докажите, что $\sum_{i=1}^{n} i^3 = \dfrac{n^2(n+1)^2}{4}$.

7. Докажите, что $\sum_{i=1}^{n} i(i+1) = \dfrac{n(n+1)(n+2)}{3}$.

8. Для положительного целого числа n определим a^n как $a^1 = a$ и $a^{k+1} = a^k \cdot a$. Докажите, что $a^{m+n} = a^m a^n$.

9. Используя математическую индукцию, докажите, что $(ab)^k = a^k b^k$.

10. Используя математическую индукцию, докажите, что $n^2 > 2n+1$ для $n \geq 3$.

11. Используя математическую индукцию, докажите, что $2^n > n^2$ для $n \geq 5$.

12. Найдите наибольшее множество положительных целых чисел, для которых $2^n > n!$ истинно. Докажите истинность утверждения.

13. Докажите, что каждое целое число $n \geq 8$ может быть записано как $n = 3k + 5m$ для некоторых неотрицательных целых чисел k и m.

14. Докажите, что для каждого положительного целого числа $n \geq 2$

$$\frac{4^n}{n+1} < \frac{(2n)!}{(n!)^2}.$$

15. Докажите, что для любого целого числа a и положительных целых чисел m и n имеет место равенство $a^{mn} = (a^m)^n$.

16. Используя математическую индукцию, докажите, что $n! > n^3$ для всех $n \geq 6$.

17. Используя математическую индукцию, докажите, что

$$\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{9}\right)\left(1 - \frac{1}{16}\right)\cdots\left(1 - \frac{1}{n^2}\right) = \frac{n+1}{2n} \qquad \text{для} \qquad n \geq 2.$$

18. Докажите, что если W — непустое подмножество множества Z и существует целое число x такое, что $x \leq w$ для всех w из W, то W имеет наименьший элемент.

19. Найдите ошибку в доказательстве того, что все лошади имеют одну и ту же масть.

ДОКАЗАТЕЛЬСТВО. Для доказательства утверждения покажем, что любые n лошадей имеют одну и ту же масть. При $n = 1$ имеется одна лошадь, которая, естественно, имеет такую же масть, как она сама. Для $n = k$ предположим, что любые данные k лошадей имеют одинаковую масть. Предположим, что в загон помещена $k+1$-я лошадь. Выведем из загона одну лошадь. Теперь в загоне k лошадей, поэтому они имеют одну и ту же масть. Вернем выведенную лошадь в загон и выведем другую лошадь. Опять в загоне k лошадей, поэтому все они имеют одну масть. Итак, все лошади имеют ту же масть, что и те лошади, которые были выведены. Поэтому все $k+1$ лошадей имеют одну и ту же масть. Таким образом, все лошади имеют одну и ту же масть. ■

20. Определим $\bigcup_{i=1}^{n} A_i$ следующим образом: $\bigcup_{i=1}^{1} A_i = A_1$ и $\bigcup_{i=1}^{k+1} A_i = \left(\bigcup_{i=1}^{k} A_i\right) \cup A_{k+1}$. Определим $\bigcap_{i=1}^{n} A_i$ как $\bigcap_{i=1}^{1} A_i = A_1$ и $\bigcap_{i=1}^{k+1} A_i = \left(\bigcap_{i=1}^{k} A_i\right) \cap A_{k+1}$. Докажите, что

а) $\left(\bigcup_{i=1}^{n} A_i\right) \cap A = \bigcup_{i=1}^{n} (A_i \cap A)$; **б)** $\left(\bigcup_{i=1}^{n} A_i\right)' = \bigcap_{i=1}^{n} A_i'$.

21. Используя приведенные выше определения, докажите, что

а) $\left(\bigcap_{i=1}^{n} A_i\right) \cap A = \bigcap_{i=1}^{n} (A_i \cap A)$; **б)** $\left(\bigcap_{i=1}^{n} A_i\right)' = \bigcup_{i=1}^{n} A_i'$.

22. Докажите, что из первого принципа индукции вытекает принцип полного упорядочения.

23. Докажите, что из принципа полного упорядочения вытекает второй принцип индукции.

24. Докажите, что из второго принципа индукции вытекает первый принцип индукции.

25. Метод "бесконечного спуска" используется иногда для доказательства того, что никакое положительное целое число не может обладать некоторым заданным свойством. Метод заключается в следующем. Предположим, что существует положительное целое число, скажем, a_1, обладающее каким-то свойством. Теперь образуем второе положительное целое число, скажем, a_2, тоже обладающее этим свойством, но такое, что $a_2 < a_1$. Аналогично, образуем положительное целое число a_3, обладающее нашим свойством, и такое, что $a_3 < a_2$, и так далее. Таким образом, получаем "бесконечно убывающую" последовательность $a_1 > a_2 > a_3 > \cdots$ положительных чисел. Очевидно, что такая последовательность не может существовать. Эта последовательность возникла из предположения о том, что существует положительное целое число, которое меньше произвольного фиксированного числа и обладает заданным свойством. Таким образом, наше предположение ошибочно. Докажите следующее утверждение, которое подтверждает приведенное выше рассуждение. Предположим, что

а) T — подмножество N.

б) Если $n \in T$, то существует $m \in N$ такое, что $m < n$ и $m \in T$.

Тогда $T = \varnothing$.

26. Используя метод бесконечного спуска, докажите, что не существуют положительные целые числа p и q такие, что $p^2 = 3q^2$, путем нахождения положительного числа $p' < p$ такого, что для некоторого положительного числа q' имеем $(p')^2 = 3(q')^2$. Можно считать известным тот факт, что, если a и b — положительные целые числа такие, что $b^2 = 3 \cdot a$, то существует положительное целое число k, обладающее свойством $b = 3 \cdot k$.

3.4. ДЕЛИМОСТЬ

Многие целые числа можно представить как произведение меньших чисел. Важные характеристики и отношения целых чисел могут быть получены на основе анализа их структуры с точки зрения составляющих множителей.

ОПРЕДЕЛЕНИЕ 3.23. Целое число a есть ***кратное*** числа b, если $a = bm$ для некоторого целого числа m. Ненулевое целое число b ***делит*** целое число a, что обозначается как $b \mid a$, если a есть кратное b. Целое число b, которое делит целое число a, называется ***делителем*** числа a.

По определению, $9 \mid 27$, поскольку $27 = 9 \cdot 3$, но 5 не делит 12, т.к. не существует целого числа m такого, что $12 = 5 \cdot m$. Целые числа $1, 2, 3, 4, 6$ и 12, и только они, являются положительными делителями числа 12. Целые числа -1, -2, -3, -4, -6 и -12 также являются делителями числа 12.

ТЕОРЕМА 3.24. Пусть a, b и c — целые числа. Тогда

а) $a \mid a$ для любого a.
б) Для любых a, b и c из $a \mid b$ и $b \mid c$ следует $a \mid c$.
в) Для любых a, b и c имеем: $b \mid a$ и $b \mid c$ тогда и только тогда, когда $b \mid (m \cdot a + n \cdot c)$ для всех целых чисел m и n.

ДОКАЗАТЕЛЬСТВО.

а) $a = a \cdot 1$.
б) Если $a \mid b$ и $b \mid c$, то $b = a \cdot m$ и $c = b \cdot n$. Следовательно, $c = b \cdot n = (a \cdot m) \cdot n = a \cdot (m \cdot n)$ и $a \mid c$.
в) Допустим, что $b \mid a$ и $b \mid c$. Тогда $a = b \cdot p$ и $c = b \cdot q$ для некоторых целых чисел p и q. Следовательно, $m \cdot a + n \cdot c = m \cdot (b \cdot p) + n \cdot (b \cdot q) = b \cdot (m \cdot p + n \cdot q)$ и $b \mid (m \cdot a + n \cdot c)$. Обратно, предположим, что $b \mid (m \cdot a + n \cdot c)$ для всех целых чисел m и n. Тогда, если $m = 1$ $n = 0$, имеем $b \mid a$. Если $m = 0$ и $n = 1$, имеем $b \mid c$.

■

ТЕОРЕМА 3.25. Если a и b — положительные целые числа и $a \mid b$, то $a \leq b$.

ДОКАЗАТЕЛЬСТВО. Доказательство следует из пункта (г) теоремы 3.11. ■

СЛЕДСТВИЕ 3.26. Если для положительных целых чисел a и b имеем $a \mid b$ и $b \mid a$, то $a = b$.

ТЕОРЕМА 3.27. Для целых чисел a и b, если $a \mid b$ и $b \mid a$, то $a = b$ или $a = -b$.

ДОКАЗАТЕЛЬСТВО. Если $a > 0$ и $b > 0$, то в силу следствия 3.26 имеем $a = b$. Если $a < 0$ и $b > 0$, то $-a > 0$ и $(-a) \mid b$, и $b \mid (-a)$, так что $b = -a$ или $a = -b$. Справедливость утверждения теоремы в случае $a > 0$ и $b < 0$ является следствием симметрии. Доказательство случая $a < 0$ и $b < 0$ предоставляется читателю. ■

Из приведенных выше теорем следует, что для положительных целых чисел отношение "делит" рефлексивно, антисимметрично и транзитивно, поэтому оно является частичным порядком на множестве положительных целых чисел. Обратите внимание, что отношение делимости не является отношением частичного порядка на множестве целых чисел, т.к. "делит" не является на множестве целых чисел антисимметричным отношением.

Следующая теорема объясняет действие, которое встречается в элементарной арифметике. Положительное целое число a можно разделить на положительное целое число b и получить остаток r, который меньше, чем b. Например, если 31 делить на 9, получим частное 3 и остаток 4. К примеру, $31 = 3 \cdot 9 + 4$, где $4 < 9$.

Доказательство следующей теоремы также является иллюстрацией принципа полного упорядочения.

ТЕОРЕМА 3.28. (Алгоритм деления) Для положительных целых чисел a и b существуют единственные неотрицательные целые числа q и r, где $0 \leq r < b$ такие, что $a = bq + r$. Такие целые числа r и q называются, соответственно, ***остатком*** и ***частным*** от деления a на b.

ДОКАЗАТЕЛЬСТВО. Пусть a и b — положительные целые числа. Рассмотрим множество S всех неотрицательных целых чисел, имеющих вид $a-bq'$ для неотрицательного целого числа q', т.е.

$$S = \{x : x = a - bq', x \geq 0, \text{ и } q' \geq 0\}.$$

Множество S непусто, поскольку для $q' = 0$ имеем $a-bq' = a-b\cdot 0 = a > 0$, так что $a \in S$. Множество S имеет наименьший элемент, скажем, r', В самом деле, если $0 \in S$, он и является наименьшим элементом. В противном случае S содержит только положительные целые числа и, следовательно, согласно принципу полного упорядочения N5 имеет наименьший элемент. Поскольку r' принадлежит S, пусть q' — неотрицательное целое число такое, что $r' = a - bq'$.

Если $r' \geq b$, то $r'' = r' - b \geq 0$ и

$$\begin{aligned} a - bq' &= r', \\ a - bq' - b &= r' - b, \\ a - b(q' + 1) &= r'', \end{aligned}$$

так что r'' принадлежит S. Кроме того, $r'' < r'$, так как

$$\begin{aligned} b &> 0, \\ -b &< 0, \\ r' - b &< r' + 0 = r'. \end{aligned}$$

Но $r'' < r'$ и то, что $r'' \in S$, противоречит предположению, что r' является наименьшим элементом множества S. Таким образом, $0 \leq r' < b$.

Чтобы показать единственность, предположим, что $a = bq + r = bp + s$, где $0 \leq r \leq s < b$. Поскольку $r \geq 0$, то $-r \leq 0$. Следовательно, $s - r < b + 0 = b$. Но $s - r = b(q - p)$, и по теореме 3.25, если $q - p$ не равно 0, то $s - r \geq b$. Следовательно, $q - p = 0$, $q = p$ и $r = s$. ■

Если $a < b$, то q должно быть равно 0, чтобы выполнялось $a = bq + r$, где $0 \leq r < b$ и $q \geq 0$. Например, для $a = 4$ и $b = 7$ алгоритм деления дает $q = 0$ и $r = a = 4$, так что $4 = 7 \cdot 0 + 4$.

В силу единственности q и r, если можно получить q и r каким-нибудь другим способом, причем $a = bq + r$, $0 \leq r < b$ и $q \geq 0$, то эти q и r должны совпадать с теми, которые существуют согласно теореме.

Поскольку любой положительный делитель положительного целого числа n не может быть больше самого этого числа, все положительные делители числа n можно найти, перебирая все целые числа от 1 до n и проверяя, не делят ли они n. Так, например, мы определили, что положительными делителями числа 12 являются числа 1, 2, 3, 4, 6 и 12. Точно так же можно показать, что положительными делителями числа 90 являются 1, 2, 3, 5, 6, 9,10, 15, 18, 30, 45 и 90.

Проверка показывает, что числа 1, 2, 3 и 6 являются делителями как 12, так и 90. Числа 1, 2, 3 и 6 называются общими делителями чисел 12 и 90. Более того, 6 есть наибольший из этих общих делителей. Обратим внимание, что общие делители 1, 2, 3 и 6 являются также делителями наибольшего общего делителя 6. В контексте наибольших общих делителей рассматриваются только положительные делители.

ОПРЕДЕЛЕНИЕ 3.29. Положительное целое число d называется ***общим делителем*** чисел a и b, если $d \mid a$ и $d \mid b$.

ОПРЕДЕЛЕНИЕ 3.30. Положительное целое число d называют ***наибольшим общим делителем*** целых чисел a и b, если
а) $d \mid a$ и $d \mid b$, и
б) из $c \mid a$ и $c \mid b$ следует $c \mid d$.
Наибольший общий делитель чисел a и b обозначается через НОД(a, b).

Алгоритм нахождения наибольшего общего делителя приведен в главе 7.

ТЕОРЕМА 3.31. Если d и c — наибольшие общие делители целых чисел a и b, то $c = d$; другими словами, существует единственный общий делитель для целых положительных чисел a и b.

ДОКАЗАТЕЛЬСТВО. По определению наибольшего общего делителя $d \mid c$ и $c \mid d$. Следовательно, $c = d$ или $c = -d$. Поскольку c и d — оба положительные, то $c = d$. ■

ТЕОРЕМА 3.32. Наибольший общий делитель положительных целых чисел a и b существует. Такой наибольший общий делитель может быть записан в виде

$$u \cdot a + v \cdot b$$

для некоторых целых чисел u и v. Кроме того, наибольший общий делитель — это наименьшее положительное целое число такого вида.

ДОКАЗАТЕЛЬСТВО. Пусть S — множество всех положительных целых чисел, имеющих форму $na + mb$. Пусть $d = ua + vb$ — наименьший элемент множества S. Тогда $d \leq a$, т.к. $a = 1 \cdot a + 0 \cdot b$ принадлежит S. Кроме того, $a = qd + r$ для некоторых q и r, где $q > 0$ и $0 \leq r < d$. Итак, $a = q(ua + vb) + r$. Решая относительно r, получаем, $r = (1 - qu)a + (-v)qb$, так что r принадлежит S или $r = 0$. Но r меньше, чем d, которое, в свою очередь, является наименьшим элементом множества S, так что $r = 0$. Поэтому $d \mid a$. Аналогично, $d \mid b$. Если c — произвольный делитель чисел a и b, то по теореме 3.24 часть (в) $c \mid d$, поскольку $d = ua + vb$. Следовательно, d — наибольший общий делитель чисел a и b. ■

Алгоритм нахождения u и v приведен в главе 7.

Свойства наибольшего общего делителя, которые сформулированы в следующей теореме, будут использованы в дальнейшем.

ТЕОРЕМА 3.33. Если $a = bq + c$, то НОД(a, b) = НОД(b, c); другими словами, каждый делитель a и b является делителем b и c, и обратно.

ДОКАЗАТЕЛЬСТВО. Пусть e = НОД(a, b) и f = НОД(b, c). Поскольку $e \mid a$ и $e \mid b$, то по теореме 3.24(в) имеем $e \mid c$, т.к. $c = a - bq$. Тогда, по определению наибольшего общего делителя $e \mid f$.

Обратно, поскольку $f \mid b$ и $f \mid c$, то по теореме 3.24(в) $f \mid a$. По определению наибольшего общего делителя $f \mid e$. Следовательно, $e = f$. ■

ТЕОРЕМА 3.34. Если a, b и c — целые числа, НОД$(a,b)=1$ и $a \mid bc$, то $a \mid c$.

ДОКАЗАТЕЛЬСТВО. Поскольку НОД$(a,b)=1$, то существуют целые числа u и v такие, что $au+bv=1$. Умножим каждое слагаемое на c, получая $cau+cbv=c$. Тогда $a \mid cau$ и $a \mid cbv$, т.к. $a \mid bc$. Следовательно, $a \mid c$. ■

ОПРЕДЕЛЕНИЕ 3.35. Если наибольший общий делитель a и b есть 1, то числа a и b называются ***взаимно простыми***.

Непосредственно из определения следует, что если a и b — взаимно простые числа, то существуют целые числа u и v такие, что $au+bv=1$.

Заметим, что если a и b — положительные целые числа, то ab кратно и a, и b. Если рассматривать множество всех чисел, кратных a и b, то согласно принципу полного упорядочения существует наименьшее кратное чисел a и b. Если c — наименьшее кратное чисел a и b, а d — другое кратное чисел a и b, то $c \mid d$. Иначе говоря, существуют q и r такие, что $d=qc+r$, где $r<c$. Поскольку и a, и b делят числа d и c, они также делят r. Но тогда r было бы кратным a и b, которое меньше, чем c, что приводит к противоречию. Таким образом, мы приходим к следующим определениям.

ОПРЕДЕЛЕНИЕ 3.36. Положительное целое число m называется ***общим кратным*** целых чисел a и b, если $a \mid m$ и $b \mid m$.

ОПРЕДЕЛЕНИЕ 3.37. Положительное целое число m называется ***наименьшим общим кратным*** целых чисел a и b, если
а) $a \mid m$ и $b \mid m$, и
б) если $a \mid n$ и $b \mid n$, то $m \mid n$.
Наименьшее общее кратное чисел a и b будем обозначать НОК(a,b).

Использование наибольшего общего делителя оказывается полезным при нахождении решений уравнений вида $ax+by=c$ или для доказательства, что такие решения не существуют. Это вытекает из приведенной ниже теоремы.

Алгоритм для нахождения x и y приведен в главе 7.

ТЕОРЕМА 3.38. Уравнение $ax+by=c$, где a,b и c — целые числа, имеет целочисленное решение (т.е. существуют целые числа x и y такие, что $ax+by=c$) тогда и только тогда, когда c делится на НОД(a,b). Если c делится на НОД(a,b), то решение $ax+by=c$ имеет вид

$$x_0=\frac{u\cdot c}{\text{НОД}(a,b)} \qquad y_0=\frac{v\cdot c}{\text{НОД}(a,b)},$$

где u и v — любые решения уравнения НОД$(a,b)=au+bv$.

ДОКАЗАТЕЛЬСТВО. Известно, что существуют целые числа u и v такие, что $au + bv = \text{НОД}(a, b)$. Если c делится на $\text{НОД}(a, b)$, то $c = e \cdot \text{НОД}(a, b)$ для некоторого целого числа e. Следовательно, $aue + bve = e \cdot \text{НОД}(a, b) = c$, так что $x = u \cdot e$ и $y = v \cdot e$ — решения уравнения. Обратно, если существуют x и y такие, что $ax + by = c$, то, поскольку $\text{НОД}(a, b)$ делит как a, так и b, он делит $ax + by$ и, следовательно, делит c. ■

■ УПРАЖНЕНИЯ

1. Найдите положительные делители каждого из следующих целых чисел:
   а) 54; **б)** 63; **в)** 72; **г)** 73; **д)** 74.
2. Для положительных целых чисел a и b найдите неотрицательные целые числа q и r, где $0 \leq r < b$ таковы, что $a = bq + r$.
   а) $a = 54$, $b = 27$; **б)** $a = 47$, $b = 47$; **в)** $a = 93$, $b = 17$;
   г) $a = 43$, $b = 8$; **д)** $a = 33$, $b = 1$.
3. Для положительных целых чисел a и b найдите $\text{НОД}(a, b)$, $\text{НОК}(a, b)$ и $\text{НОД}(a, b) \cdot \text{НОК}(a, b)$, если они определены.
   а) $a = 54$, $b = 27$; **б)** $a = 17$, $b = 0$; **в)** $a = 6$, $b = 15$;
   г) $a = 12$, $b = 16$; **д)** $a = 33$, $b = 1$.
4. Докажите или опровергните утверждение: для положительных целых чисел a и b $\text{НОД}(a, b) \leq \text{НОК}(a, b)$.
5. Если $a = 0$ или $b = 0$, то что можно было бы принять в качестве $\text{НОК}(a, b)$? Ответ обоснуйте.
6. Докажите, что для взаимно простых чисел a и b и заданного целого числа n существуют целые числа x и y такие, что $ax + by = n$.
7. Докажите, что для положительных целых чисел a и b, если $a > b$, то $\text{НОД}(a, b) = \text{НОД}(a - b, b)$.
8. Рассмотрим теорему: если целые числа a и b таковы, что $a \mid b$ и $b \mid a$, то $a = b$ или $a = -b$. Докажите, что эта теорема также справедлива в случае, когда $a < 0$ и $b < 0$.

3.5. ПРОСТЫЕ ЧИСЛА

Очевидно, что каждое целое число a делится само на себя и на 1, т.к. $a = 1 \cdot a$. Каждое целое число делит 0, но 0 не делит никакое целое число. Для некоторых задач теории чисел необходимо знать, имеет ли некоторое конкретное целое число делители, отличные от него самого и 1. Некоторые целые числа не могут быть разложены в произведение целых чисел, исключая тривиальный способ. Такие целые числа называются простыми.

ОПРЕДЕЛЕНИЕ 3.39. Целое число, большее 1, называется ***простым***, если оно не имеет положительных делителей, кроме 1 и самого себя. Положительное целое число, большее 1, называется ***составным***, если оно не является простым.

Среди первых 10 положительных целых чисел имеется четыре простых числа: 2, 3, 5 и 7. Целые числа $4 = 2 \cdot 2$, $6 = 2 \cdot 3$, $8 = 2 \cdot 4$, $9 = 3 \cdot 3$ и $10 = 2 \cdot 5$ являются составными. Итак, если $n = r \cdot s$, где $1 < r < n$ и $1 < s < n$, то n — составное число. По определению, целое число 1 не является ни простым, ни составным. Число 2 — единственное четное простое число. Определить, является ли небольшое целое число простым, пытаясь разделить его на меньшие простые числа, сравнительно легко, т.к. количество возможных вариантов невелико. Однако, вопрос о том, является ли простым большое целое число, может оказаться достаточно сложным.

Следующая теорема показывает, что существует бесконечно много простых чисел. Доказательство теоремы является классическим примером использования принципа сведения к абсурду.

ТЕОРЕМА 3.40. (Евклид) Существует бесконечно много простых чисел.

ДОКАЗАТЕЛЬСТВО. Допустим, что существует только конечное число простых чисел, например, p_1, p_2, ..., p_k. Рассмотрим целое число $(p_1 p_2 \cdots p_k) + 1$. Предположим, что p_r — некоторое простое число, и $p_r \mid ((p_1 p_2 \cdots p_k) + 1)$. Но тогда $p_r \mid (p_1 p_2 \cdots p_k)$, откуда следует, что $p_r \mid 1$, а это приводит к противоречию, т.к. $p_r > 1$. Следовательно, $(p_1 p_2 \cdots p_k) + 1$ — простое число, что, в свою очередь, также является противоречием, т.к. этого числа нет среди указанной конечной совокупности простых чисел. Таким образом, наше предположение о том, что существует конечное число простых чисел, ложно, поэтому простых чисел должно быть бесконечно много. ■

Поскольку разложение целых чисел на простые множители является важной задачей, необходимо иметь быстрый и простой способ определения, является ли данное положительное целое число простым или составным. Следующая теорема показывает, что для проверки простоты числа необходимо определить только некоторые из его возможных делителей.

ТЕОРЕМА 3.41. Если положительное целое число n является составным, то n имеет простой делитель p такой, что $p^2 \leq n$.

ДОКАЗАТЕЛЬСТВО. Пусть p — наименьший простой делитель числа n. Если n раскладывается на множители r и s, то $p \leq r$ и $p \leq s$. Следовательно, $p^2 \leq rs = n$. ■

Например, чтобы определить, является ли $n = 521$ простым, необходимо рассмотреть только простые числа, которые меньше или равны 22, потому что $22^2 = 484$, а $23^2 = 529$. Простые числа, меньшие или равные 22, — это 2, 3, 5, 7, 11, 13, 17 и 19. Проверяя каждое из них, находим, что ни одно из них не делит 521. Поэтому 521 является простым числом в силу предыдущей теоремы.

Как покажет следующая теорема, простые числа образуют множество своего рода строительных блоков для целых чисел.

ТЕОРЕМА 3.42. Каждое положительное целое число либо равно 1, либо простое, либо может быть записано как произведение простых чисел.

ДОКАЗАТЕЛЬСТВО. В доказательстве этой теоремы использован второй принцип индукции. Несомненно, теорема справедлива для $n = 1$. Допустим, что она справедлива для всех положительных целых чисел m, меньших, чем k. Если k — простое, то теорема справедлива также и для k. Если k не является простым, то оно делится на некоторое простое число p и $k = pq$, где и p, и q не могут быть равными k или 1. Поскольку согласно теореме 3.25 числа p и q меньше, чем k, то по индуктивному предположению они либо простые, либо могут быть записаны как произведение простых чисел. Следовательно, $k = pq$ может быть записано как произведение простых чисел. ■

Целое число 37 — простое. Целое число $1554985071 = 3 \cdot 3 \cdot 4463 \cdot 38713$ — произведение четырех простых чисел, два из которых совпадают.

Следующая теорема показывает, что если некоторое целое число n делится на простое число p, то невозможно разложить n на множители таким образом, чтобы p не был делителем по крайней мере одного из этих множителей.

ТЕОРЕМА 3.43. Если p — простое число и $p \mid ab$, где a и b — положительные целые числа, то $p \mid a$ или $p \mid b$.

ДОКАЗАТЕЛЬСТВО. Если $p \mid a$, то утверждение теоремы имеет место. С другой стороны, предположим, что p не делит a. Поскольку p и a — взаимно простые, то существуют u, v такие, что $pu + av = 1$. Поэтому $pub + avb = b$. В силу того, что $p \mid pub$ и $p \mid abv$ (т.к. $p \mid ab$), имеем $p \mid b$. ■

ЛЕММА 3.44. Если простое число p делит произведение положительных целых чисел $q_1 q_2 \cdots q_n$, то p делит q_i для некоторого i, $1 \leq i \leq n$.

ДОКАЗАТЕЛЬСТВО. В доказательстве леммы использована индукция по n, числу множителей в произведении. Если $n = 1$, правильность леммы очевидна. Допустим, что лемма верна для $n = k$. Это означает, что если p делит произведение любых k целых чисел, то p делит один из k множителей. Предположим, что p делит произведение $k + 1$ целых чисел, скажем, $p \mid q_1 q_2 \cdots q_k q_{k+1}$, так что $p \mid (q_1 q_2 \cdots q_k) q_{k+1}$. Если p делит q_{k+1}, то доказательство завершено. Если p не делит q_{k+1}, то по теореме 3.43 имеем $p \mid (q_1 \cdots q_k)$. Но поскольку $q_1 \cdots q_k$ — произведение k целых чисел, то по индуктивному предположению $p \mid q_i$ для некоторого $1 \leq i \leq k$. Следовательно, $p \mid q_i$ для некоторого i, $1 \leq i \leq k + 1$. ■

ЛЕММА 3.45. Если простое число p делит произведение простых чисел q_1, q_2, ..., q_n, то $p = q_i$ для некоторого i, $1 \leq i \leq n$.

ДОКАЗАТЕЛЬСТВО. Согласно лемме 3.44, p делит q_i для некоторого $1 \leq i \leq n$. Поскольку числа p и q_i — простые, то $p = q_i$. ■

Из совокупности теорем, приведенных выше, получаем главный результат этой главы, который известен как основная теорема арифметики, или теорема о единственности разложения на простые множители.

ТЕОРЕМА 3.46. (Основная теорема арифметики) Любое положительное целое число, большее, чем 1, либо является простым, либо может быть записано в виде произведения простых чисел, причем это произведение единственно с точностью до порядка сомножителей.

ДОКАЗАТЕЛЬСТВО. Допустим, что $q_1q_2\cdots q_n$ и $p_1p_2\cdots p_s$ — два способа записи положительного целого числа m как произведения простых чисел. В доказательстве теоремы будет использована индукция по n, числу простых сомножителей в первом произведении. Если $n = 1$, теорема тривиально верна. Сформулируем индуктивное предположение: пусть теорема верна, когда $q_1q_2\cdots q_k = p_1p_2\cdots p_s$; т.е., если $m = q_1q_2\cdots q_k = p_1p_2\cdots p_s$, то $k = s$ и произведение единственно с точностью до порядка простых сомножителей. Предположим теперь, что $m = q_1q_2\cdots q_{k+1} = p_1p_2\cdots p_{s'}$. Поскольку q_{k+1} делит $p_1p_2\cdots p_{s'}$, отсюда следует, что $q_{k+1} = p_i$ для некоторого $1 \le i \le s'$. Разделим оба произведения на q_{k+1} или используем свойство сокращения (аксиому I8). Тогда $q_1q_2\cdots q_k = p_1p_2\cdots p_{i-1}p_{i+1}\cdots p_{s'}$. Но по индукции $k = s' - 1$, и произведение единственно с точностью до порядка простых сомножителей. Следовательно, $k + 1 = s'$, и разложение на множители числа m единственно с точностью до порядка простых сомножителей. ■

Например,

$$\begin{aligned} n = 39616304 &= \\ &= 2\cdot 13\cdot 7\cdot 2\cdot 23\cdot 13\cdot 2\cdot 13\cdot 2\cdot 7 = \\ &= 2\cdot 2\cdot 2\cdot 2\cdot 7\cdot 7\cdot 13\cdot 13\cdot 13\cdot 23 \end{aligned}$$

представляют собой два разложения на множители числа n; однако, в каждом из произведений одно и то же простое число использовано одинаковое число раз. Различается только порядок записи простых чисел. Фактически имеется 12,600 различных способов разложения на множители числа n с использованием 10 простых сомножителей, однако, каждое такое разложение содержит ровно четыре двойки, две семерки, три множителя, равных 13, и один, равный 23. Обычно простые множители группируют, используя экспоненциальную запись. Например,

$$n = 2^4 7^2 13^3 23^1 .$$

СЛЕДСТВИЕ 3.47. Каждое положительное целое число, большее 1, может быть записано единственным образом с точностью до порядка в виде $q_1^{k(1)}q_2^{k(2)}\cdots q_n^{k(n)}$, где $k(1)$, $k(2)$, ... и $k(n)$ — положительные целые числа.

Теперь понятно, почему 1 не входит во множество простых чисел. В противном случае теорема о единственности разложения на простые множители была бы неверна. При изучении разложения различных целых чисел, используя их представление, данное в предыдущем следствии, зачастую для удобства обозначения допускается нулевая степень простого сомножителя. Обычно такая практика не приводит к путанице, т.к. при $q_i^0 = 1$ всегда $q_i \ge 1$. Если разложение на простые множители известно, то простые числа, формирующие разложение на простые множители любого делителя этого числа, образуют подмножество соответствующего множества для делимого.

Доказательство двух следующих теорем предоставляется читателю.

ТЕОРЕМА 3.48. Если $a = p_1^{a(1)}\cdots p_k^{a(k)}$ и $b \mid a$, то $b = p_1^{b(1)}\cdots p_k^{b(k)}$, где $0 \le b(i) \le a(i)$ для всех i; и обратно.

ТЕОРЕМА 3.49. Пусть $a = p_1^{a(1)} p_2^{a(2)} p_3^{a(3)} \cdots p_k^{a(k)}$ и $b = p_1^{b(1)} p_2^{b(2)} p_3^{b(3)} \cdots p_k^{b(k)}$, где p_i — простые числа, которые делят либо a, либо b, и некоторые показатели степени могут быть равны 0. Пусть $m(i) = \min(a(i), b(i))$ и $M(i) = \max(a(i), b(i))$ для $1 \leq i \leq k$. Тогда

$$\text{НОД}(a,b) = p_1^{m(1)} p_2^{m(2)} p_3^{m(3)} \cdots p_k^{m(k)}$$

и

$$\text{НОК}(a,b) = p_1^{M(1)} p_2^{M(2)} p_3^{M(3)} \cdots p_k^{M(k)}.$$

Применим теорему 3.49 в случае, когда $a = 195000$ и $b = 10435750$. Разложения на простые множители чисел a и b имеют вид

$$a = 2^3 3^1 5^4 13^1 \qquad \text{и} \qquad b = 2^1 5^3 13^3 19^1.$$

Таким образом,

$$\begin{aligned}
\text{НОД}(195000, 10435750) &= 2^{\min(3,1)} 3^{\min(1,0)} 5^{\min(4,3)} 13^{\min(1,3)} 19^{\min(0,1)} = \\
&= 2^1 3^0 5^3 13^1 19^0 = 2^1 5^3 13^1 = 3250, \\
\text{НОК}(195000, 10435750) &= 2^{\max(3,1)} 3^{\max(1,0)} 5^{\max(4,3)} 13^{\max(1,3)} 19^{\max(0,1)} = \\
&= 2^3 3^1 5^4 13^3 19^1 = 626145000.
\end{aligned}$$

Следующая теорема вытекает из теоремы 3.49. Доказательство предоставляется читателю.

ТЕОРЕМА 3.50. Если a и b — положительные целые, то $\text{НОД}(a,b) \cdot \text{НОК}(a,b) = ab$.

■ УПРАЖНЕНИЯ

1. Разложите каждое из следующих целых чисел на простые множители.
   а) 728
   б) 1599 (используйте равенство $1599 = 1600 - 1$)
   в) 4899 **г)** 131 **д)** 523
2. Используйте теоремы данной главы для нахождения НОД и НОК следующих пар чисел:
   а) 162 и 12 **б)** 71 и 23 **в)** 72 и 30
   г) $n!$ и $(n+2)!$ **д)** 75 и 99
3. Два простых числа a и b называются числами-близнецами, если разность между ними равна 2, т.е. если $a + 2 = b$. Например, 3 и 5 являются числами-близнецами. Найдите три другие пары чисел-близнецов.
4. Является ли среднее арифметическое двух простых чисел-близнецов простым числом?
5. Если a и b — простые числа, следует ли отсюда, что $a^2 + b^2$ — простое число?
6. Объясните, почему для любого положительного целого числа n все числа $n! + 1$, $n! + 2$, $n! + 3$, $n! + 4$, $\ldots, n! + n$ являются составными. Что данный факт говорит о расстоянии между простыми числами?

7. Покажите, что числа 479001603 и 479001607 не являются простыми, используя предыдущее упражнение и калькулятор.
8. Если p_1, p_2, p_3, ..., p_n — первые n последовательных простых чисел, то следует ли отсюда, что $p_1 \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_n + 1$ (произведение первых n простых чисел плюс единица) есть простое число. Обоснуйте свой ответ.
9. Покажите, что не существует троек последовательных нечетных чисел, каждое из которых является простым, за исключением тройки $(3, 5, 7)$.
10. Докажите, что если p и q — простые числа, большие или равные 5, то $p+q$ или $p-q$ делится на 3 и, следовательно, $p^2 - q^2$ делится на 24. (*Указание*: $p+2$ или $p-2$ делится на 3 и $q-2$ или $q+2$ делится на 3.)
11. Докажите, что если $a = p_1^{a(1)} \cdots p_k^{a(k)}$ и $b \mid a$, то $b = p_1^{b(1)} \cdots p_k^{b(k)}$, где $0 \leq b(i) \leq a(i)$ для всех i, и обратно.
12. Пусть $a = p_1^{a(1)} p_2^{a(2)} p_3^{a(3)} \cdots p_k^{a(k)}$ и $b = p_1^{b(1)} p_2^{b(2)} p_3^{b(3)} \cdots p_k^{b(k)}$, где p_i — простые числа, которые делят или a, или b, и некоторые показатели степени могут быть равны 0. Пусть $m(i) = \min(a(i), b(i))$ и $M(i) = \max(a(i), b(i))$ для $1 \leq i \leq k$. Затем докажите, что

$$\text{НОД}(a, b) = p_1^{m(1)} p_2^{m(2)} p_3^{m(3)} \cdots p_k^{m(k)}$$

и

$$\text{НОК}(a, b) = p_1^{M(1)} p_2^{M(2)} p_3^{M(3)} \cdots p_k^{M(k)}.$$

3.6. СРАВНЕНИЯ

В примере 2.52 изучалось отношение R_3 на множестве целых чисел, определенное как $R_3 = \{(a, b) : a - b = 5 \cdot k\}$. Отношение эквивалентности, порожденное таким разбиением, называется "сравнением по модулю 5". Фактически такие разбиения можно построить для любого положительного целого числа n, как это было сделано для $n = 5$.

ОПРЕДЕЛЕНИЕ 3.51. Пусть n — положительное целое число. Целое число a ***сравнимо*** с целым числом b по модулю n, что обозначается $a \equiv b \pmod{n}$, если n делит $(a - b)$.

Следующая теорема доказывается по аналогии с примером 2.52.

ТЕОРЕМА 3.52. Отношение $\equiv$ для фиксированного n является отношением эквивалентности на множестве целых чисел. Это означает, что

а) $a \equiv a \pmod{n}$ для каждого целого числа a;

б) если $a \equiv b \pmod{n}$, то $b \equiv a \pmod{n}$ для целых чисел a и b;

в) если $a \equiv b \pmod{n}$ и $b \equiv c \pmod{n}$, то $a \equiv c \pmod{n}$.

ОПРЕДЕЛЕНИЕ 3.53. Пусть n — положительное целое число. Множество всех классов эквивалентности по модулю n обозначается Z_n и называется ***множеством классов вычетов по модулю n***.

Классы вычетов по модулю n представляют собой новые объекты. Они являются классами эквивалентности. Элементы каждого класса эквивалентности сравнимы между собой по модулю n. Например, пусть $n = 3$. Имеется три класса эквивалентности по модулю 3, так что множество

$$Z_3 = \{[0], [1], [2]\}$$

содержит три элемента. Элементы Z_3 — классы эквивалентности и, следовательно, множества. Эти три множества содержат 0, 1 и 2, соответственно, как свои имена. В каждом из этих классов эквивалентности все элементы сравнимы между собой по модулю 3, так что $a \equiv b \pmod 3$ тогда и только тогда, когда a и b принадлежат одному и тому же классу эквивалентности — классу вычетов. Таким образом,

$$[0] = \{\ldots -6, -3, 0, 3, 6, 9 \ldots\};$$
$$[1] = \{\ldots -8, -5, -2, 1, 4, 7 \ldots\};$$
$$[2] = \{\ldots -7, -4, -1, 2, 5, 8 \ldots\}.$$

Теперь нам нужно определить операции между классами вычетов по модулю n. Отношение сравнимости $\equiv$, как подсказывает само обозначение, имеет много общего с отношением равенства. Но с отношением сравнимости по модулю n связано больше ограничений, чем с аналогичным ему отношением равенства.

ТЕОРЕМА 3.54. Отношение сравнимости обладает следующими свойствами:

а) если $a \equiv b \pmod n$ и $c \equiv d \pmod n$, то $a + c \equiv b + d \pmod n$ и $ac \equiv bd \pmod n$;

б) если $a \equiv b \pmod{mn}$, то $a \equiv b \pmod m$ и $a \equiv b \pmod n$.

ДОКАЗАТЕЛЬСТВО. Мы докажем пункт (а); часть (б) предлагаем доказать читателю. По определению, если $a \equiv b \pmod n$ и $c \equiv d \pmod n$, то $a - b = un$ и $c - d = vn$ для некоторых целых чисел u и v. Таким образом,

$$(a + c) - (b + d) = (a - b) + (c - d) = u \cdot n + v \cdot n = (u + v) \cdot n$$

для некоторых целых чисел u и v, и $a + c \equiv b + d \pmod n$. Кроме того, $a = b + un$ и $c = d + vn$, так что

$$\begin{aligned} ac &= (b + (un))(d + (vn)) = \\ &= bd + b(vn) + d(un) + uvn^2 = \\ &= bd + (bv + ud + uvn)\, n\,. \end{aligned}$$

Таким образом, $ac - bd = (bv + ud + vun)n$ и $ac \equiv bd \pmod n$. ■

ПРИМЕР 3.55. Пусть $n = 3$.

$$1 \equiv 7 \pmod 3,$$
$$2 \equiv -4 \pmod 3.$$

Из теоремы 3.54 (а) следует, что

$$1 + 2 \equiv 7 + (-4) \pmod 3,$$

а также, что

$$(1)(2) \equiv (7)(-4) \pmod 3,$$

что верно, поскольку $2 - (-28) = 30 = 3 \cdot 10 = 3 \cdot k$. □

ТЕОРЕМА 3.56. Для произвольного положительного целого числа n,

а) если $r \equiv r' \pmod{n}$, $0 \leq r < n$ и $0 \leq r' < n$, то $r = r'$;

б) если a — любое целое число и n — положительное целое число, то имеется целое число r, где $0 \leq r < n$ такое, что $a \equiv r \pmod{n}$. Целое число r — остаток от деления a на n, поэтому $a = nq + r$.

ДОКАЗАТЕЛЬСТВО. Утверждения теоремы следуют непосредственно из алгоритма деления. ■

Часть (а) утверждения теоремы гарантирует, что классы вычетов по модулю n обладают тем свойством, что классы вычетов, порожденные неотрицательными целыми числами, меньшими, чем n, различны. Таким образом, $[0], [1], [2], \ldots, [n-2], [n-1]$ — различные множества относительно сравнимости по модулю n.

Часть (б) показывает, что каждое целое число a сравнимо по модулю n с одним из целых чисел $0, 1, 2, \ldots, n-1$. Итак, любое целое число a принадлежит одному из n различных классов эквивалентности $[0], [1], \ldots, [n-1]$. Проведенное рассмотрение служит объяснением к следующей теореме.

ТЕОРЕМА 3.57. Пусть n — произвольное положительное целое число. Тогда Z_n, множество классов вычетов по модулю n, состоит точно из n различных классов вычетов

$$[0], [1], [2], \ldots, [n-1],$$

представленных различными остатками, которые могут быть получены при делении на n. Кроме того, для $0 \leq r < n$, класс вычетов $[r]$ состоит точно из тех целых чисел a, для которых $a \equiv r \pmod{n}$.

Следующая теорема является простым следствием предыдущих теорем, которое, тем не менее, часто оказывается полезным.

ТЕОРЕМА 3.58. Если $a = nq + r$ и $b = nq' + r'$ для $0 \leq r < n$ и $0 \leq r' < n$, то $r = r'$ тогда и только тогда, когда $a \equiv b \pmod{n}$.

В целом, при работе с классами вычетов по модулю n желательно для его представления выбрать в классе наименьшее неотрицательное целое число. Это особенно важно для сложения и умножения классов вычетов в Z_n, поскольку необходимо, чтобы сумма и произведение все еще представлялись целыми числами, которые находятся между 0 и $n-1$. Это приводит к следующему определению.

ОПРЕДЕЛЕНИЕ 3.59. Пусть Z_n — множество классов вычетов по модулю n. Для любого заданного целого числа m существует целое число r такое, что $0 \leq r \leq n-1$ и $[m] = [r]$ или $m \equiv r \pmod{n}$. При этом говорят, что $[[m]]_n = r$.

ПРИМЕР 3.60. Для $n = 5$ получаем

$$\begin{aligned} Z_5 &= \{[0], [1], [2], [3], [4]\} = \\ &= \{[r] : 0 \leq r < 5\}. \end{aligned}$$

□

На множестве Z_n можно определить операции сложения и умножения. Если $[a]$ — класс вычетов по модулю n, содержащий a, и $[b]$ — класс вычетов по модулю n, содержащий b, то сложение и умножение определим соотношениями

$$[a] \oplus [b] = [a + b] = [[a + b]]_n ,$$
$$[a] \odot [b] = [a \cdot b] = [[a \cdot b]]_n ,$$

где сложение и умножение в центре и справа осуществляется между целыми числами, а сложение и умножение по левую сторону выполняется между классами эквивалентности.

Использование предыдущей теоремы показывает, что операции сложения и умножения таких классов эквивалентности — классов вычетов — определены правильно; т.е. определение не зависит от представителей классов вычетов. Например, если $[a] = [c]$, то результат сложения не должен зависеть от того, используем мы $[a]$ или $[c]$. Иначе говоря, если $[a] = [c]$ и $[b] = [d]$, то $[a] \oplus [b] = [c] \oplus [d]$ и $[a] \odot [b] = [c] \odot [d]$.

ПРИМЕР 3.61. Для $n = 5$ и $Z_5 = \{[0], [1], [2], [3], [4]\}$ находим, что

$$[2] \oplus [4] = [2 + 4] = [6] = [1], \quad \text{поскольку} \quad 6 \equiv 1 \pmod 5;$$

$$[2] \odot [4] = [2 \cdot 4] = [8] = [3], \quad \text{поскольку} \quad 8 \equiv 3 \pmod 5.$$

Вычисляя все возможные суммы и произведения, можно создать таблицы "сложения" и "умножения" для классов вычетов по модулю 5.

$[a] \oplus [b]$	**[0]**	**[1]**	**[2]**	**[3]**	**[4]**
[0]	[0]	[1]	[2]	[3]	[4]
[1]	[1]	[2]	[3]	[4]	[0]
[2]	[2]	[3]	[4]	[0]	[1]
[3]	[3]	[4]	[0]	[1]	[2]
[4]	[4]	[0]	[1]	[2]	[3]

$[a] \odot [b]$	**[0]**	**[1]**	**[2]**	**[3]**	**[4]**
[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]
[2]	[0]	[2]	[4]	[1]	[3]
[3]	[0]	[3]	[1]	[4]	[2]
[4]	[0]	[4]	[3]	[2]	[1]

□

Теорема 3.57 показывает, что для произвольного положительного целого числа n множество целых чисел может быть разбито на n непересекающихся множеств, или классов эквивалентности. В дальнейшем мы часто будем возвращаться к этим множествам, интересуясь их представителями.

ОПРЕДЕЛЕНИЕ 3.62. Если $b \equiv r \pmod{n}$ для положительного целого числа n, то говорят, что r есть ***вычет*** числа b по модулю n. ***Полная система вычетов*** по модулю n есть множество $S = \{r_1, r_2, \ldots, r_n\}$, где пересечение множества S с каждым классом вычетов по модулю n содержит точно одно целое число, т.е. S содержит одного и только одного представителя из каждого такого класса вычетов. Полная система вычетов $\{0, 1, 2, \ldots, (n-1)\}$ называется ***первичной системой вычетов***. Если b — целое число, $b \equiv r \pmod{n}$ и $0 \leq r \leq n-1$, то этот единственный первичный вычет по модулю n обозначается через $r = [[b]]_n$. ***Приведенная система вычетов*** по модулю n есть подмножество полной системы вычетов, состоящее только из тех целых чисел, которые являются взаимно простыми с n, т.е. $\{r : r \in S \text{ и НОД}(r, n) = 1\}$.

Полная система вычетов получается путем выбора одного целого числа из каждого класса вычетов $[0], [1], \ldots, [n-1]$ множества Z_n. Итак, для $n = 6$, $\{24, 7, -58, 40, 113\}$ — полная система вычетов по модулю 6, так как

$$\begin{aligned}
24 &\equiv 0 \pmod{6}, && \text{поэтому} && 24 \in [0];\\
7 &\equiv 1 \pmod{6}, && \text{поэтому} && 7 \in [1];\\
-58 &\equiv 2 \pmod{6}, && \text{поэтому} && -58 \in [2];\\
15 &\equiv 3 \pmod{6}, && \text{поэтому} && 15 \in [3];\\
40 &\equiv 4 \pmod{6}, && \text{поэтому} && 40 \in [4];\\
113 &\equiv 5 \pmod{6}, && \text{поэтому} && 113 \in [5].
\end{aligned}$$

Очевидно, что $\{0, 1, 2, 3, 4, 5\}$ является полной (и притом первичной) системой вычетов по модулю 6. По определению, $[[24]]_6 = 0$ и $[[-58]]_6 = 2$.

Поскольку все элементы любого класса вычетов по модулю n сравнимы один с другим, но не являются таковыми по отношению к элементам другого класса вычетов, каждое целое число сравнимо ровно с одним элементом полной системы вычетов. Проведенные рассуждения приводят к следующей теореме.

ТЕОРЕМА 3.63. Если n — положительное целое число, $\{r_1, r_2, \ldots, r_n\}$ — полная система вычетов по модулю n и a — некоторое целое число, то $a \equiv r_k \pmod{n}$ для одного и только одного k, $1 \leq k \leq n$.

Элементы полной системы вычетов $\{24, 7, -58, 15, 40, 113\}$ и первичной (полной) системы вычетов $\{0, 1, 2, 3, 4, 5\}$ по модулю 6, которые являются взаимно простыми с $n = 6$, содержат в себе, соответственно, множества $\{7, 113\}$ и $\{1, 5\}$. Поэтому оба последних множества являются приведенными системами вычетов по модулю 6. При этом говорят, что $\{1, 5\}$ является первичной приведенной системой вычетов по модулю 6.

Следующая теорема показывает, что приведенная система вычетов содержит больше существенной информации о целых числах из Z, чем система чисел, взаимно простых с n. Доказательство теоремы предоставляется читателю.

ТЕОРЕМА 3.64. Пусть n — положительное целое число и $\{r_1, r_2, \ldots, r_k\}$ — приведенная система вычетов по модулю n. Если a — целое число, взаимно простое с n, то $a \equiv r_j \pmod{n}$ для одного и только одного j, $1 \leq j \leq k$.

Если дана одна система вычетов, имеется несколько способов воспроизвести другие системы вычетов. Один из способов состоит в прибавлении одного и того же целого числа к каждому вычету в полной системе вычетов. Иной способ демонстрирует следующая теорема. Доказательство предоставляется читателю.

ТЕОРЕМА 3.65. Пусть n — положительное целое число и $\{r_1, r_2, \ldots, r_k\}$ — полная [приведенная] система вычетов по модулю n. Если a — целое число, взаимно простое с n, то $\{ar_1, ar_2, \ldots, ar_k\}$ — также полная [приведенная] система вычетов.

Поскольку НОД$(6, 35) = 1$, имея полную систему вычетов $\{0, 1, 2, 3, 4, 5\}$ и приведенную систему вычетов $\{1, 5\}$ по модулю 6, получаем, что $\{35 \cdot 0, 35 \cdot 1, 35 \cdot 2, 35 \cdot 3, 35 \cdot 4, 35 \cdot 5\}$ и $\{35 \cdot 1, 35 \cdot 5\}$ являются, соответственно, полной и приведенной системами вычетов. Но

$$\begin{array}{lllll} [[0]]_6 = 0 & \text{и} & [[35 \cdot 0]]_6 = [[0]]_6 = 0 \\ [[1]]_6 = 1 & & [[35 \cdot 1]]_6 = [[35]]_6 = 5 \\ [[2]]_6 = 2 & & [[35 \cdot 2]]_6 = [[70]]_6 = 4 \\ [[3]]_6 = 3 & & [[35 \cdot 3]]_6 = [[107]]_6 = 3 \\ [[4]]_6 = 4 & & [[35 \cdot 4]]_6 = [[140]]_6 = 2 \\ [[5]]_6 = 5 & & [[35 \cdot 5]]_6 = [[175]]_6 = 1 \end{array}$$

Таким образом, в результате умножения каждого элемента r_i полной или приведенной системы вычетов по модулю n на целое число a, взаимно простое с n, получаем представителей тех же классов вычетов, возможно, в другом порядке.

■ УПРАЖНЕНИЯ

1. Вычислите

а) $[[37]]_4$; **б)** $[[93]]_5$; **в)** $[[48]]_{23}$; **г)** $[[149]]_{27}$;

2. Вычислите

а) $[[33]]$; **б)** $[[47]]_9$; **в)** $[[49]]_{17}$; **г)** $[[146]]_{21}$;

3. Вычислите

а) $[[8!]]_6$;

б) $[[1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10]]_3$;

в) $[[1^2 + 2^2 + 3^2 + \cdots 25^2]]_2$;

г) $[[(146)^2]]_{21}$.

4. Покажите, что $a^2 \equiv b^2 \pmod{n}$ не имеет следствием $a \equiv b \pmod{n}$.

5. Покажите, что если a — нечетное целое число, то $a^2 \equiv 1 \pmod{8}$.

6. Постройте таблицу сложения для классов вычетов по модулю 4.

7. Постройте таблицу умножения для классов вычетов по модулю 4.

8. Постройте таблицу сложения для классов вычетов по модулю 7.

9. Постройте таблицу умножения для классов вычетов по модулю 7.

10. Используя таблицы сложения и умножения для классов вычетов по модулю 5, приведенные ранее, найдите в Z_5 решения следующих уравнений:

а) $x + [3] = [5]$; **б)** $[3] \cdot x = [2]$;

в) $[4] \cdot x = [3]$; **г)** $[2] \cdot x = [1]$.

11. Используя таблицы сложения и умножения для классов вычетов по модулю 4 из приведенных выше упражнений, найдите в Z_4 решения следующих уравнений:

а) $x + [3] = [2]$; **б)** $[3] \cdot x = [2]$;

в) $[2] \cdot x = [2]$; **г)** $[2] \cdot x = [0]$.

12. Используя таблицы сложения и умножения для классов вычетов по модулю 7 из приведенных выше упражнений, найдите в Z_7 решения следующих уравнений:

а) $x + [5] = [2]$; **б)** $[3] \cdot x = [4]$;

в) $[2] \cdot x = [5]$; **г)** $[4] \cdot x = [6]$.

13. Докажите, что если $a \equiv b \pmod{mn}$, то $a \equiv b \pmod{m}$ и $a \equiv b \pmod{n}$.

14. Докажите, что если $ac \equiv bc \pmod{n}$ и целые числа c и n — взаимно простые, то $a \equiv b \pmod{n}$.

15. Докажите теорему 3.64. Пусть n — положительное целое число, и $\{r_1, r_2, \ldots, r_k\}$ — приведенная система вычетов по модулю n. Если a — целое число, взаимно простое с n, то $a \equiv r_j \pmod{n}$ для одного и только одного j, $1 \leq j \leq k$.

16. Докажите теорему 3.65. Пусть n — положительное целое число, и $\{r_1, r_2, \ldots, r_k\}$ — полная [приведенная] система вычетов по модулю n. Если a — целое число, взаимно простое с n, то $\{ar_1, ar_2, \ldots, ar_k\}$ — тоже полная [приведенная] система вычетов.

ГЛАВА

4

ФУНКЦИИ И МАТРИЦЫ

4.1. ФУНКЦИИ

Функция из множества A в множество B представляет собой специальное отношение на $A \times B$, обладающее следующими свойствами:

1. Областью определения отношения является все множество A. Следовательно, для каждого элемента a из A существует элемент b из B такой, что a и b связаны данным отношением.
2. Если a относится к b и a относится к b', то $b = b'$. В терминах упорядоченных пар это утверждение означает, что если (a, b) и (a, b') принадлежат отношению, то $b = b'$. Дадим более формальное определение функции.

ОПРЕДЕЛЕНИЕ 4.1. Отношение f на $A \times B$ называется ***функцией*** из A в B и обозначается через $f : A \to B$, если для каждого $a \in A$ существует единственный элемент $b \in B$ такой, что $(a, b) \in f$. Если $f : A \to B$ — функция, и $(a, b) \in f$, то говорят, что $b = f(a)$. Множество A называется ***областью определения*** функции f, а множество B называется ***областью потенциальных значений***. Если $E \subseteq A$, то множество $f(E) = \{b : f(a) = b$ для некоторого a из $E\}$ называется ***образом*** множества E. Образ всего множества A называется ***областью значений*** функции f. Если $F \subseteq B$, то множество $f^{-1}(F) = \{a : f(a) \in F\}$ называется ***прообразом*** множества F. Функция $f : A \to B$ называется также ***отображением*** ; при этом говорят, что f отображает A в B. Если $(a, b) \in f$, так что $b = f(a)$, то говорят, что элемент a отображается в элемент b.

ПРИМЕР 4.2. Пусть $A = \{-2, -1, 0, 1, 2\}$, а $B = \{0, 1, 2, 3, 4, 5\}$. Определим отношение $f \subseteq A \times B$ как $f = \{(-2, 5), (-1, 2), (0, 1), (1, 2), (2, 5)\}$. Отношение f — функция из A в B, так как $f \subseteq A \times B$, и каждый из элементов A присутствует в качестве первой компоненты упорядоченной пары из f ровно один раз. □

ПРИМЕР 4.3. Пусть $A = \{-2, -1, 0, 1, 2\}$ и $B = \{0, 1, 2, 3, 4, 5\}$. Функция $f : A \to B$ определена соотношением $f(x) = x^2 + 1$. Если $E = \{1, 2\}$, то

$$\begin{aligned} f(E) &= \{b : (a, b) \in f \text{ для некоторого } a \text{ из } E\} = \\ &= \{b : b = f(a) \text{ для некоторого } a \text{ из } E\} = \\ &= \{2, 5\} \end{aligned}$$

является образом E при отображении f. Если $F = \{0, 2, 3, 4, 5\} \subseteq B$, то

$$\begin{aligned} f^{-1}(F) &= \{b : \text{существует } a \in A \text{ такое, что } f(a) = b\} = \\ &= \{-1, 1, -2, 2\} \end{aligned}$$

является прообразом F, где $-1 \in f^{-1}(F)$, т.к. $f(-1) = 2$, $1 \in f^{-1}(F)$, т.к. $f(1) = 2$, $-2 \in f^{-1}(F)$, т.к. $f(-2) = 5$ и $2 \in f^{-1}(F)$, т.к. $f(2) = 5$. Заметим, что элементы 0, 3 и 4 не вносят никаких элементов в $f^{-1}(F)$, поскольку они не принадлежат области значений функции f. Прообраз может быть пустым. Так, например, в случае $W = \{0, 3\}$ прообраз $f^{-1}(W)$ пуст, поскольку не существует такого $a \in A$, для которого $f(a) = 0$ или $f(a) = 3$.

Область значений функции f имеет вид

$$\begin{aligned} f(A) &= \{b : f(a) = b \text{ для некоторого } a \in A\} = \\ &= \{1, 2, 5\}. \end{aligned}$$

Элементами $f(A)$ являются те и только те элементы области потенциальных значений B, которые "используются" функцией f. □

Мы уже отмечали, что если R — отношение на $A \times B$, а S — отношение на $B \times C$, то можно определить отношение $S \circ R$ на $A \times C$, называемое композицией S и R. Если R и S — функции, то $S \circ R$ — тоже функция, называемая композицией S и R. Следующая теорема описывает некоторые полезные свойства композиций.

ТЕОРЕМА 4.4. Пусть $g : A \to B$ и $f : B \to C$. Тогда

а) композиция $f \circ g$ есть функция из A в C, что может быть записано как

$$f \circ g : A \to C;$$

б) если $a \in A$, то $(f \circ g)(a) = f(g(a))$.

ДОКАЗАТЕЛЬСТВО. а) Если $g : A \to B$ и $f : B \to C$, то $g \subseteq A \times B$ и $f \subseteq B \times C$. По определению композиции отношений $f \circ g \subseteq A \times C$. Пусть $(a, c) \in f \circ g$ и $(a, c') \in f \circ g$. Требуется показать, что $c = c'$. Поскольку $(a, c) \in f \circ g$, существует такой элемент $b \in B$, что $(a, b) \in g$ и $(b, c) \in f$. Точно так же, поскольку $(a, c') \in f \circ g$, то существует такой элемент $b' \in B$, что $(a, b') \in g$ и $(b', c') \in f$. Но g является функцией, так что $(a, b) \in g$ и $(a, b') \in g$, откуда следует равенство $b = b'$. Следовательно, $(b', c') = (b, c') \in f$ и $(b, c) \in f$. Поскольку f также является функцией, то $c = c'$. Поэтому $f \circ g$ — функция из A в C.

б) Если $g : A \to B$ и $a \in A$, тогда существует такое b из B, что $b = g(a)$. Поскольку $b \in B$ $f : B \to C$, имеется некоторое $c \in C$, для которого $(b, c) \in f$, т.е. $c = f(b)$. Но $b = g(a)$, так что $f(g(a)) = f(b) = c = (f \circ g)(a)$, поэтому $(a, b) \in g$ и $(b, c) \in f$, откуда следует, что $(a, c) \in f \circ g$. ■

Поскольку композиция отношений обладает свойством ассоциативности (см. теорему 2.31), то, в частности, имеет место следующая теорема.

ТЕОРЕМА 4.5. Пусть $f : A \to B$, $g : B \to C$ и $h : C \to D$. Тогда $h \circ (g \circ f) = (h \circ g) \circ f$, т.е. композиция функций ассоциативна.

ПРИМЕР 4.6. Пусть $f(x) = \sqrt{x}$ и $g(x) = x + 3$ — функции, заданные на множестве действительных чисел. Функция $f(g(x)) = f(x + 3) = \sqrt{x + 3}$. Функция $g(f(x)) = g(\sqrt{x}) = \sqrt{x} + 3$. □

Функция f из A в B может быть классифицирована в зависимости от того, существуют ли элементы из B, связанные данным отношением с более чем одним элементом из A, и связан ли каждый элемент из области значений B с соответствующим элементом области определения A.

ОПРЕДЕЛЕНИЕ 4.7. Функция $f : A \to B$ называется ***инъективной***, или ***инъекцией***, если из $f(a) = f(a')$ следует $a = a'$. Функция f называется ***отображением "на"***, или ***сюръективной функцией***, или ***сюръекцией***, если для каждого $b \in B$ существует некоторое $a \in A$ такое, что $f(a) = b$. Функция, которая является одновременно и инъективной, и сюръективной, называется ***взаимно однозначным соответствием***, или ***биекцией***. Если $A = B$ и $f : A \to B$ является взаимно однозначным соответствием, то f называется ***перестановкой*** множества A.

ПРИМЕР 4.8. Пусть A и B — множества действительных чисел и $f : A \to B$ определена таким образом: $f(x) = 3x + 5$. Функция f инъективна, т.к. если $f(a) = f(a')$, тогда $3a + 5 = 3a' + 5$ и, следовательно, $a = a'$. Функция f является также сюръекцией. Для любого действительного числа b требуется найти такое a, что $f(a) = b = 3a + 5$. Решая это уравнение относительно a, находим, что если $a = (1/3)(b - 5)$, тогда $f(a) = b$. Поэтому f представляет собой взаимно однозначное соответствие, а в силу $A = B$, f является также перестановкой. □

ПРИМЕР 4.9. Пусть A и B — множество действительных чисел, и функция $f : A \to B$ определена как $f(x) = x^2$. Функция f не является инъективной, т.к. $f(2) = f(-2)$, но $2 \neq -2$. Функция f не является также и сюръективной, поскольку не существует такого действительного числа a, для которого $f(a) = -1$. Заметим, что если A и B — множество неотрицательных действительных чисел, тогда f является как инъективной, так и сюръективной. □

Пусть f — функция из множества A во множество B, т.е. $f : A \to B$. Очевидно, $f \subseteq A \times B$, так как f является отношением на $A \times B$. ***Обратное отношение*** $f^{-1} \subseteq B \times A$ определяется как $f^{-1} = \{(b, a) : (a, b) \in f\}$. При этом отношение f^{-1} может не быть функцией из B в A, даже если f является функцией из A в B. В следующей теореме сформулированы условия, при которых f^{-1} является функцией. Если f^{-1} действительно является функцией, то ее называют ***обращением*** функции f, или ее ***обратной функцией***.

ТЕОРЕМА 4.10. Если функция $f : A \to B$ является биекцией, то обратное отношение f^{-1} является функцией из B в A, причем биекцией. Обратно, для $f : A \to B$, если f^{-1} — функция из B в A, то f является биекцией.

ДОКАЗАТЕЛЬСТВО. Для доказательства того, что f^{-1} — функция из B в A, необходимо убедиться, что областью определения f^{-1} является B, и что если (b,a) и $(b,a') \in f^{-1}$, то $a = a'$. Для доказательства того, что область определения f^{-1} совпадает с B, возьмем произвольное $b \in B$. Поскольку f сюръективна, то существует $a \in A$ такое, что $f(a) = b$ или $(a,b) \in f$. Поэтому $(b,a) \in f^{-1}$ и B является областью определения f^{-1}. Пусть (b,a) и $(b,a') \in f^{-1}$. Тогда (a,b) и $(a',b) \in f$. Поскольку f инъективна, то $a = a'$. Поэтому f^{-1} — функция. При этом она сюръективна, т.к. если $a \in A$, то в силу того, что f — функция, существует некоторое b такое, что $f(a) = b$, т.е. $(a,b) \in f$. Поэтому $(b,a) \in f^{-1}$, и элемент a принадлежит области значений функции f^{-1}. Следовательно, A представляет собой область значений f^{-1}, а сама функция f^{-1} — сюръективна. Для доказательства инъективности функции f^{-1} предположим, что (b,a) и $(b',a) \in f^{-1}$, т.е. $f^{-1}(b) = a$ и $f^{-1}(b') = a$. Тогда $(a,b) \in f$ и $(a,b') \in f$ (или, что эквивалентно, $f(a) = b$ и $f(a) = b'$). Поскольку f — функция, то $b = b'$. Для доказательства обратного утверждения следует в последнем рассуждении f^{-1} заменить на f. ■

ТЕОРЕМА 4.11. Если $f : A \to B$ — биекция, то

а) $f(f^{-1}(b)) = b$ для любого b из B;
б) $f^{-1}(f(a)) = a$ для любого a из A.

ДОКАЗАТЕЛЬСТВО. Пусть $b \in B$ и $a = f^{-1}(b)$. Тогда $f(a) = b$. Но поскольку $a = f^{-1}(b)$, то $f(f^{-1}(b)) = f(a) = b$. Аналогично доказывается, что $f^{-1}(f(a)) = a$ для любого a из A. ■

ПРИМЕР 4.12. Требуется найти обратную функцию для $y = 3x + 6$. Обращая функцию, получаем

$$\{(y,x) : y = 3x + 6\}.$$

Но это то же самое, что

$$\{(x,y) : x = 3y + 6\}.$$

Решая уравнение относительно y, получаем

$$\{(x,y) : y = (x - 6)/3\}.$$ □

ОПРЕДЕЛЕНИЕ 4.13. Пусть $I : A \to A$ определено соотношением $I(a) = a$ для всех $a \in A$. Функция I называется ***тождественной функцией*** на A.

Предлагаем читателю доказать следующую теорему.

ТЕОРЕМА 4.14. Если $f : A \to A$ и I — тождественная функция на A, то $I \circ f = f \circ I = f$. Если для f существует обратная функция, то $f \circ f^{-1} = f^{-1} \circ f = I$.

ТЕОРЕМА 4.15. Пусть $g : A \to B$ и $f : B \to C$. Тогда

а) если g и f — сюръекции A на B и B на C соответственно, то $f \circ g$ есть сюръекция A на C. Другими словами, композиция двух сюръекций — сюръекция;
б) если g и f — инъекции, то $f \circ g$ — также инъекция; иными словами, композиция двух инъекций — инъекция;

в) если g и f — биекции, то $f \circ g$ — также биекция; иными словами, композиция двух биекций — биекция;
г) $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

■ УПРАЖНЕНИЯ

1. Пусть $f \subseteq R \times R$, где R — множество действительных чисел. Найдите область значений и область определения следующих функций:
а) $f(x) = x^2 + 4$;
б) $f(x) = \sqrt{x-2}$;
в) $f(x) = \dfrac{1}{\sqrt{x-2}}$;
г) $f(x) = \dfrac{1}{x^2+4}$;
д) $f(x) = \dfrac{1}{x^2-4}$;
е) $f(x) = |x|$.

2. Какие из приведенных ниже отношений являются функциями, если x и y — действительные числа, x принадлежит области определения, а y — области значений?
а) $y^2 = x^2 + 4$;
б) $y^3 = x^3 + 4$;
в) $y = 5$;
г) $x = 7$;
д) $y = \sqrt{x^2-2}$.

3. Для функций f и g, заданных на множестве действительных чисел, найдите $f(g(x))$ и $g(f(x))$, если:
а) $f(x) = x^2 + 1$ и $g(x) = x + 3$;
б) $f(x) = \sqrt{x^2+2}$ и $g(x) = x^2 + 3$;
в) $f(x) = \dfrac{1}{x}$ и $g(x) = 2x + 3$.

4. Найдите обратную для каждой из следующих функций:
а) $y = \dfrac{x+4}{2}$;
б) $y = x^3$;
в) $y = \dfrac{x-2}{x+3}$.

5. Выясните, какие из приведенных ниже функций, у которых область определения и область значений совпадают с действительной числовой осью, являются инъективными, сюръективными, имеют обратную функцию.
а) $f(x) = |x|$;
б) $f(x) = x^2 + 4$;
в) $f(x) = x^3 + 6$;
г) $f(x) = x + |x|$;
д) $f(x) = x(x-2)(x+2)$.

6. Пусть $f : A \to B$, A_1 и A_2 — подмножества A, а B_1 и B_2 — подмножества B. Докажите, что
а) если $A_1 \subseteq A_2$, то $f(A_1) \subseteq f(A_2)$;
б) если $B_1 \subseteq B_2$, то $f^{-1}(B_1) \subseteq f^{-1}(B_2)$;
в) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$;
г) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$;
д) $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$;
е) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$;

ж) $f^{-1}(B_1') = (f^{-1}(B_1))'$.

з) Приведите пример функции f и множеств A_1 и A_2 таких, что $f(A_1 \cap A_2) \neq f(A_1) \cap f(A_2)$.

7. Пусть $f : A \to B$ — биекция, а $g : B \to A$ — функция, обладающая следующими свойствами:

$$(g \circ f)(a) = a \text{ для всех } a \in A;$$
$$(f \circ g)(b) = b \text{ для всех } b \in B.$$

Докажите, что $g = f^{-1}$. Заметим, что данное утверждение говорит о том, что f^{-1} — функция из B в A, при этом $f \circ f^{-1} = I_B$ и $f^{-1} \circ f = I_A$, где I_A и I_B являются тождественными функциями на A и B соответственно. (*Указание*: используйте свойство ассоциативности композиции.)

8. Пусть $f : A \to B$. Докажите, что f инъективна тогда и только тогда, когда для всех подмножеств X и Y множества A имеем $f(X \cap Y) = f(X) \cap f(Y)$.

9. Для функции $f : A \to B$ докажите, что

а) f инъективна тогда и только тогда, когда $(f^{-1} \circ f)(W) = W$ для любого $W \subseteq A$, где $f^{-1}(K)$ обозначает прообраз множества K;

б) f сюръективна тогда и только тогда, когда $(f \circ f^{-1})(V) = V$ для любого $V \subseteq B$.

10. Пусть $g : A \to B$ и $f : B \to C$. Докажите, что

а) если g и f — сюръекции, то $f \circ g$ — сюръекция;

б) если g и f — инъекции, то $f \circ g$ — инъекция;

в) если g и f — биекции, то $f \circ g$ — биекция;

г) $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

11. Докажите, что если $f : A \to A$ и I — тождественная функция на A, то $I \circ f = f \circ I = f$. Если f имеет обратную функцию, то $f \circ f^{-1} = f^{-1} \circ f = I$.

4.2. СПЕЦИАЛЬНЫЕ ФУНКЦИИ

Перестановка была определена как функция, которая на множестве устанавливает взаимно однозначное соответствие. Если f — перестановка на множестве $\{1, 2, 3, \ldots, n\}$, она может быть представлена в виде

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ f(1) & f(2) & \ldots & f(n) \end{pmatrix}.$$

Напомним, что нам известна еще одна специальная функция — тождественная функция I, определенная соотношением $I(a) = a$ для всех $a \in A$. Она может быть представлена в виде

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}.$$

ПРИМЕР 4.16. Если $A = \{1, 2, 3\}$ и функция $f : A \to A$ определена соотношениями $f(1) = 3$, $f(2) = 2$ и $f(3) = 1$, то f может быть представлена как

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

□

Если $g : A \to A$ определена соотношением $g(1) = 2$, $g(2) = 3$ и $g(3) = 1$, то g можно изобразить в виде

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Композиции этих функций имеют вид

$$f \circ g = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

и

$$g \circ f = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

В предыдущих примерах было показано, что перестановка

$$I = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

также является тождественной функцией в том смысле, что $I \circ f = f \circ I = f$ для любой перестановки f на множестве A. Чтобы построить обратную перестановку, нужно найти число, стоящее над 1, и поместить его под 1; затем найти число, стоящее над 2, и поместить его под 2; наконец, найти число, стоящее над 3, и поместить его под 3. Таким образом, обращение перестановки

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{есть} \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Введем еще три специальные функции.

ОПРЕДЕЛЕНИЕ 4.17. Функция $f : A \to B$, где A — множество действительных чисел, а B — множество целых чисел, называется ***нижним округлением*** и обозначается $f(x) = \lfloor x \rfloor$, если она каждому $a \in A$ ставит в соответствие наибольшее целое число, меньшее или равное a. Функция $f : F \to B$ называется ***верхним округлением*** и обозначается $f(x) = \lceil x \rceil$, если она каждому $a \in A$ ставит в соответствие наименьшее целое число, большее или равное a.

Например, $\lfloor 2.99 \rfloor = 2$, $\lfloor 4 \rfloor = 4$, $\lfloor -4 \rfloor = -4$ и $\lfloor -4.1 \rfloor = -5$. Точно также, $\lceil 2.99 \rceil = 3$, $\lceil 4 \rceil = 4$, $\lceil -4 \rceil = -4$ и $\lceil -4.1 \rceil = -4$.

ОПРЕДЕЛЕНИЕ 4.18. Пусть A и B совпадают со множеством неотрицательных целых чисел. ***Факториалом*** назовем функцию $f : A \to B$, обозначаемую через $f(n) = n!$ и определяемую следующими соотношениями:

$$\begin{array}{lcl@{\qquad}lcl@{\qquad}lcl}
0! & = & 1 & & & & & & \\
1! & = & 1 & & & & & = & 1 \cdot 0! \\
2! & = & 1 \cdot 2 & & = & 2 & & = & 2 \cdot 1! \\
3! & = & 1 \cdot 2 \cdot 3 & & = & 6 & & = & 3 \cdot 2! \\
4! & = & 1 \cdot 2 \cdot 3 \cdot 4 & & = & 24 & & = & 4 \cdot 3! \\
\vdots & & \vdots & & & \vdots & & & \vdots \\
k! & = & 1 \cdot 2 \cdot 3 \cdot 4 \cdot \cdots \cdot k & & & & & = & k \cdot (k-1)!
\end{array}$$

ПРИМЕР 4.19.

$$\frac{12!}{10!} = \frac{12 \cdot 11 \cdot 10!}{10!} = 12 \cdot 11 = 132. \qquad \square$$

ПРИМЕР 4.20.

$$\frac{7!}{4! \cdot 3!} = \frac{7 \cdot 6 \cdot 5 \cdot 4!}{4! \cdot 3 \cdot 2 \cdot 1} = 35. \qquad \square$$

Операции сложения, вычитания, умножения и деления также являются примерами функций. В случае вычитания целых чисел $5 - 3$ означает, что операция вычитания ставит целое число, а именно, 2, в соответствие числам 5 и 3. Чтобы подчеркнуть функциональную природу операции вычитания, можно было бы записать $-((5,3)) = 2$. Вычитание каждой паре целых чисел (r, s) ставит в соответствие целое число. Функция двух переменных такого частного вида называется бинарной операцией и определяется следующим образом.

ОПРЕДЕЛЕНИЕ 4.21. ***Бинарной операцией*** на множестве A называется функция $b : A \times A \to A$. Образ пары (r, s) при отображении b записывается как $b((r, s))$, или как $r\,b\,s$.

Поскольку область значений бинарной операции на A по определению есть подмножество A, то множество A обладает свойством замкнутости относительно бинарной операции (в том смысле, что значение операция на двух элементах r и s множества A также является элементом A).

Последовательности также являются частными видами функций, область определения которых есть начальное подмножество множества положительных или неотрицательных целых чисел. Полагаем, что читатель знаком с такими последовательностями, как 1, 2, 4, 9, 16, ... или 2, 4, 6, 8, 10, ..., хотя, может быть, никогда и не думал о них, как о функциях.

ОПРЕДЕЛЕНИЕ 4.22. ***Конечной последовательностью*** назовем функцию из $\{1, 2, 3, 4, \ldots, n\}$ в некоторое множество S. ***Бесконечной последовательностью*** назовем функцию из $\{1, 2, 3, 4, \ldots\}$ в некоторое множество S. Любая конечная или бесконечная последовательность может быть названа просто ***последовательностью***.

В большинстве случаев S представляет собой множество положительных целых, целых, рациональных или действительных чисел. Последовательность обычно изображают списком или перечислением элементов. Если A — функция, то значение $A(1)$ можно обозначить через A_1, значение $A(2)$ — через A_2 и так далее. Поэтому, если A — конечная последовательность, она может быть представлена как

$$A(1), A(2), A(3), \ldots, A(n)$$

или

$$A_1, A_2, A_3, \ldots, A_n\,.$$

Если же A — бесконечная последовательность, она может быть представлена как

$$A(1), A(2), A(3), \ldots$$

или

$$A_1, A_2, A_3, \ldots .$$

Например, последовательности $1, 4, 9, 16, \ldots$ соответствует запись $A_i = i^2$. Здесь $A(1) = 1$, $A(2) = 4$, $A(3) = 9$, $A(4) = 16$ и т.д. Заметьте, что $1, 4, 9, 16$ и $1, 9, 4, 16$ — это две разные последовательности.

ПРИМЕР 4.23. Пусть $A(n) = n + 5$. Запишем первые пять членов последовательности: $A(1) = 1 + 5 = 6$, $A(2) = 2 + 5 = 7$, $A(3) = 3 + 5 = 8$, $A(4) = 4 + 5 = 9$ и $A(5) = 5 + 5 = 10$. □

ПРИМЕР 4.24. Пусть $A(n) = n^2 - 3$. Запишем первые пять членов последовательности: $A(1) = 1^2 - 3 = -2$, $A(2) = 2^2 - 3 = 1$, $A(3) = 3^2 - 3 = 6$, $A(4) = 4^2 - 3 = 13$ и $A(5) = 5^2 - 3 = 22$. □

ПРИМЕР 4.25. Заданы первые пять членов последовательности: $1, 6, 11, 16, 21$. Требуется описать ее как функцию. Поскольку каждый член последовательности больше предыдущего на 5, то $A(n) = 1 + 5(n - 1)$, где $1 \leq n \leq 5$. □

ПРИМЕР 4.26. Заданы первые пять членов последовательности: $0, 3, 8, 15, 24$, и требуется описать ее как функцию. Поскольку первые пять членов последовательности $B(n) = n^2$ совпадают с $1, 4, 9, 16, 25$, эта последовательность может быть задана формулой $A(n) = n^2 - 1$. □

ПРИМЕР 4.27. Заданы первые пять членов последовательности: $1, -2, 3, -4, 5$. Эта последовательность может быть описана формулой $A(n) = (-1)^{n+1}n$. Если бы знакопеременная последовательность начиналась с отрицательного числа, то вместо $(-1)^{n+1}$ мы бы подставили $(-1)^n$. □

Два вида последовательностей — ***арифметическая прогрессия*** и ***геометрическая прогрессия*** — представляют особый интерес. Каждый член арифметической прогрессии может быть получен из предыдущего прибавлением константы c. Таким образом, арифметическая прогрессия имеет вид $a + (n - 1)c$, где a — первый член прогрессии. Последовательность 3, 5, 7, 9, 11, ... представляет собой арифметическую прогрессию с $a = 3$ и $c = 2$.

В геометрической прогрессии каждый член может быть получен из предыдущего путем умножения на константу r. Таким образом, геометрическая прогрессия имеет вид $A(n) = ar^{(n-1)}$, где a — первый член прогрессии. Последовательность 4, 12, 36, 108, 324, ... является геометрической прогрессией, в которой $a = 4$ и $r = 3$. Последовательность 32, 16, 8, 4, 2, ... — геометрическая прогрессия, где $a = 32$ и $r = \frac{1}{2}$.

ОПРЕДЕЛЕНИЕ 4.28. Сумма $A_r + A_{r+1} + A_{r+2} + A_{r+3} + \cdots + A_{r+k}$ может быть записана с использованием ***знака суммирования*** как $\sum_{i=r}^{r+k} A_i$.

Например,

$$\sum_{i=3}^{7} i^2 = 3^2 + 4^2 + 5^2 + 6^2 + 7^2;$$

$$\sum_{i=4}^{4} i^3 = 4^3;$$

$$\sum_{i=1}^{4} i = 1 + 2 + 3 + 4;$$

$$\sum_{i=1}^{3} (i+3) = (1+3) + (2+3) + (3+3).$$

Часто требуется найти сумму $\sum\limits_{i=1}^{n} A_i$ первых n членов последовательности A.

Выведем формулу суммы первых n членов арифметической прогрессии

$$S = a + (a+c) + (a+2c) + \ldots + (a+(n-2)c) + (a+(n-1)c).$$

Для этого запишем S в обратном порядке и сложим саму с собой, так что

$$\begin{aligned}
S &= a + (a+c) + (a+2c) + \cdots + [a+(n-1)c],\\
S &= [a+(n-1)c] + [a+(n-2)c] + [a+(n-2)c] + \cdots + a,\\
2S &= [a+(a+(n-1)c)] + [a+(a+(n-1)c)] + [a+(a+(n-1)c])+\\
&\quad + \cdots + [a+(a+(n-1)c] =\\
&= [2a+(n-1)c] + [2a+(n-1)c] + [2a+(n-1)c] + \cdots + [2a+(n-1)c].
\end{aligned}$$

Таким образом, имеем

$$2S = n(2a + (n-1)c),$$

так что

$$S = \frac{n}{2}(2a + (n-1)c).$$

Поскольку $A_1 = a$ и $A_n = a + (n-1)c$), то

$$S = \frac{n}{2}(A_1 + A_n).$$

ПРИМЕР 4.29. $S = 3 + 5 + 7 + 9 + 11 + 13 = \dfrac{6}{2}(3 + 13) = 48.$ □

Теперь выведем формулу для вычисления суммы геометрической прогрессии

$$S = a + ar + ar^2 + ar^3 + ar^4 + \ldots + ar^{n-2} + ar^{n-1}.$$

Заметим, что $rS = ar + ar^2 + ar^3 + ar^4 + ar^{54} + \ldots + ar^{n-1} + ar^n$. Вычитая S из rS, получаем $rS - S = ar^n - a$, так что

$$S = \frac{a(r^n - 1)}{r - 1}.$$

ПРИМЕР 4.30. $2+6+18+54+162 = \dfrac{2(3^5-1)}{3-1} = \dfrac{2(486-1)}{3-1} = 243$, так как здесь $a = 2$, $r = 3$ и $n = 5$. □

■ УПРАЖНЕНИЯ

1. В приведенных ниже упражнениях для перестановок f и g найдите перестановки $f \circ g$, $g \circ f$, f^{-1} и g^{-1}.

а) $f = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$;

б) $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$;

в) $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$;

г) $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$.

2. Вычислите значения следующих выражений:

а) $\dfrac{10!}{8!}$; **б)** $\dfrac{12!}{8! \cdot 4!}$; **в)** $\dfrac{n!}{(n-2)!}$;

г) $-\lfloor -2.999 \rfloor$; **д)** $\lfloor 1.001 \rfloor$; **е)** $\lceil 3.5 \rceil$;

ж) $\lceil -4.01 \rceil$.

3. Вычислите значения следующих выражений:

а) $\dfrac{14!}{13! \cdot 0!}$; **б)** $\dfrac{15!}{5! \cdot 4! \cdot 6!}$; **в)** $\lfloor -9 \rfloor$;

г) $-\lfloor 2.999 \rfloor$; **д)** $\lfloor -1.001 \rfloor$; **е)** $\lceil -3.7 \rceil$;

ж) $\lceil 4.01 \rceil$.

4. Найдите первые пять членов следующих последовательностей:

а) $A_n = n^3 + 3$; **б)** $A_n = (-1)^{n+1} \dfrac{n^2}{n!}$;

в) $A_n = \left\lceil \dfrac{n}{3} \right\rceil$; **г)** $A_n = \dfrac{(n+3)!}{(n!)(3!)}$.

5. Найдите первые пять членов следующих последовательностей:

а) $A_n = n^2 + 2n + 3$; **б)** $A_n = \dfrac{n(n+1)}{n!}$;

в) $A_n = \left\lfloor \dfrac{n-5}{n} \right\rfloor$; **г)** $A_n = (-1)^{n+1} \left(\dfrac{1}{n} \right)$.

6. Определите по первым пяти членам последовательности A формулу для A_n.

а) 3, 8, 13, 18, 23; **б)** $\dfrac{1}{2}, -\dfrac{2}{3}, \dfrac{3}{4}, -\dfrac{4}{5}, \dfrac{5}{6}$;

в) 4, 12, 36, 78, 144; **г)** 2, 2, 4, 12, 48.

7. Определите по первым пяти членам последовательности A формулу для A_n.

а) 2, 8, 18, 32, 50; **б)** $\dfrac{1}{2}, -\dfrac{1}{6}, \dfrac{1}{18}, -\dfrac{1}{54}, \dfrac{1}{162}$;

в) 1, −1, 2, −2, 3; **г)** 3, 6, 11, 18, 27.

4.3. МАТРИЦЫ

До сих пор мы рассматривали лишь "функции одной переменной". Они, как правило, обозначаются через $f(x)$, где x — так называемая "переменная". Например, $f(x) = x^2$ с областью определения R есть функция одной переменной. Если область определения f представляет собой декартово произведение двух множеств, скажем, $C \times D$, то функцию $f : C \times D \to B$ называют "функцией двух переменных" и обычно обозначают через $f(x, y)$, где x и y называются переменными, причем значение x берется из области C, а значение y — из области D. Если $(c, d) \in C \times D$, то пишут $f((c, d))$ или просто $f(c, d)$. Например, $f(x, y) = x^2 + y^2$ с областью определения $R \times R$ — это функция двух переменных. Бинарные операции также являются функциями двух переменных. Точно так же, если область определения f — множество $C \times D \times E$, а $f : C \times D \times E \to B$, то f называется "функцией трех переменных". Например, $f(x, y, z) = x^2 + y^2 + 2yz$ с областью определения $R \times R \times R$ есть функция трех переменных.

Дадим описание специальной функции двух переменных, называемой матрицей или массивом. Информация в компьютере обычно хранится в виде матриц.

ОПРЕДЕЛЕНИЕ 4.31. Для положительных целых чисел m и n ***матрицей*** $m \times n$, или ***массивом*** $m \times n$, называется функция

$$A : \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to D,$$

где D — это, как правило, множество действительных, комплексных, рациональных или целых чисел. Элементы D называются ***скалярами***. Таким образом, для каждого i, $1 \leq i \leq m$, и каждого j, $1 \leq j \leq n$, имеется элемент $A(i, j)$ из D, который находится в i-ой строке и в j-ом столбце соответствующей прямоугольной таблицы. Образ $A(i, j)$ элемента (i, j) области определения сокращенно обозначается через A_{ij}. Таким образом, $n \times m$ матрица A изображается прямоугольной таблицей, где образы $(i, j) \in \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$ при отображении A могут быть перечислены следующим образом:

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ \vdots & & & & \vdots \\ A_{m1} & A_{m2} & A_{m3} & \cdots & A_{mn} \end{bmatrix},$$

где первая строка таблицы состоит из A_{1j} для $j = 1$ до n, вторая строка состоит из A_{2j} для $j = 1$ до n и т.д. При этом говорят, что A содержит m строк и n столбцов и является матрицей ***размера*** $m \times n$. Иногда записывают сокращенно $A = [A_{ij}]$ или даже $A = [a_{ij}]$. Значение A_{ij} называется ***компонентой***, или ***элементом*** матрицы A.

Матрица размера $1 \times m$ называется ***матрицей-строкой***, а матрица размера $n \times 1$ называется ***матрицей-столбцом***. Если в матрице число строк и число столбцов совпадает, она называется ***квадратной матрицей***. Если A — матрица-строка, индекс строки обычно опускают и пишут

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \cdots & A_m \end{bmatrix}.$$

Точно так же, если A — матрица-столбец, то мы опускаем индекс столбца:

$$A = \begin{bmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{n1} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}.$$

Например, $A = \begin{bmatrix} 2 & 1 & 7 \\ 4 & 0 & 6 \end{bmatrix}$ — это матрица 2×3, а $B = \begin{bmatrix} -2 & 5 & 7 \\ -3 & 9 & 0 \\ 25 & 2 & 9 \end{bmatrix}$ представляет собой квадратную матрицу 3×3. По определению, $A_{13} = 7$, $A_{21} = 4$, $B_{12} = 5$ и $B_{31} = 25$. Матрица $C = \begin{bmatrix} 7 & 1 \\ 3 & 9 \end{bmatrix}$ есть квадратная матрица 2×2.

ОПРЕДЕЛЕНИЕ 4.32. Говорят, что две матрицы $A = [A_{ij}]$ и $B = [B_{ij}]$ размера $m \times n$ ***равны***, если равны их соответствующие элементы; т.е. $A = B$ тогда и только тогда, когда $A_{ij} = B_{ij}$ для всех i, $1 \leq i \leq m$, и всех j, $1 \leq j \leq n$.

Например,

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}$$

тогда и только тогда, когда $A_{ij} = B_{ij}$ для $1 \leq i, j \leq 3$. Заметим, что это определение равенства матриц является всего лишь переформулировкой условия равенства двух функций A и B: $A((i,j)) = B((i,j))$ для всех (i,j).

В данном разделе некоторые определения и теоремы будут сформулированы только для матриц 2×2 и 3×3; более общая их формулировка будет дана в последующих главах. Для упрощения выкладок, а также для удобства и определенности, большинство теорем будет доказано только для матриц 2×2 и 3×3.

ОПРЕДЕЛЕНИЕ 4.33. Если d — скаляр, а $A = [A_{ij}]$ — матрица $m \times n$, то dA есть матрица $D = [D_{ij}]$ размера $m \times n$, где $D_{ij} = dA_{ij}$, т.е. каждая компонента D есть произведение соответствующей компоненты A на d. Произведение числа d и матрицы A называется ***умножением матрицы на скаляр***.

ПРИМЕР 4.34. Пусть $A = [A_{ij}] = \begin{bmatrix} 1 & -3 & 5 \\ 6 & 0 & -2 \end{bmatrix}$. Тогда

$$7A = 7\begin{bmatrix} 1 & -3 & 5 \\ 6 & 0 & -2 \end{bmatrix} = \begin{bmatrix} (7)(1) & (7)(-3) & (7)(5) \\ (7)(6) & (7)(0) & (7)(-2) \end{bmatrix} = \begin{bmatrix} 7 & -21 & 35 \\ 42 & 0 & -14 \end{bmatrix}.$$

□

ОПРЕДЕЛЕНИЕ 4.35. Если $A = [A_{ij}]$ и $B = [B_{ij}]$ — $m \times n$ матрицы, тогда $A+B$ есть $m \times n$ матрица $C = [C_{ij}]$, где $C_{ij} = A_{ij} + B_{ij}$, другими словами, матрицы складываются покомпонентно. Матрица C называется ***суммой матриц*** A и B. Запись $A - B$ означает $A + (-1)B$.

ПРИМЕР 4.36. Пусть $A = \begin{bmatrix} -1 & 3 \\ 2 & 7 \\ 4 & -5 \end{bmatrix}$ и $B = \begin{bmatrix} 3 & 11 \\ -5 & 4 \\ 8 & 2 \end{bmatrix}$. Тогда

$$A+B = \begin{bmatrix} -1 & 3 \\ 2 & 7 \\ 4 & -5 \end{bmatrix} + \begin{bmatrix} 3 & 11 \\ -5 & 4 \\ 8 & 2 \end{bmatrix} = \begin{bmatrix} (-1)+3 & 3+11 \\ 2+(-5) & 7+4 \\ 4+8 & (-5)+2 \end{bmatrix} = \begin{bmatrix} 2 & 14 \\ -3 & 11 \\ 12 & -3 \end{bmatrix}.$$

□

ОПРЕДЕЛЕНИЕ 4.37. ***Произведение*** двух матриц определяется следующим образом:

а) Пусть V — матрица-строка или матрица-столбец с n элементами, и W — матрица-строка или матрица-столбец с n элементами:

$$V = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} \quad \text{или} \quad V = \begin{bmatrix} V_1 & V_2 & \cdots & V_n \end{bmatrix},$$

$$W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix} \quad \text{или} \quad W = \begin{bmatrix} W_1 & W_2 & \cdots & W_n \end{bmatrix}.$$

Тогда ***скалярным произведением***, или ***внутренним произведением*** V и W называется число $V_1W_1 + V_2W_2 + \cdots + V_nW_n$, которое обозначается через $V \bullet W$.

б) Пусть

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mp} \end{bmatrix}$$

есть матрица $m \times p$, а

$$B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & & & \vdots \\ B_{p1} & B_{p2} & \cdots & B_{pn} \end{bmatrix}$$

есть матрица $p \times n$.

Тогда (матричное) ***произведение*** A и B, обозначаемое через AB, есть матрица $C = [C_{ij}]$ размера $m \times n$, где C_{ij} — это скалярное произведение i-ой строки матрицы A и j-го столбца матрицы B. То есть,

$$C_{ij} = \begin{bmatrix} A_{i1} & A_{i2} & A_{i3} & \cdots & A_{ip} \end{bmatrix} \bullet \begin{bmatrix} B_{1j} \\ B_{2j} \\ B_{3j} \\ \vdots \\ B_{pj} \end{bmatrix} = \sum_{k=1}^{p} A_{ik} B_{kj} .$$

ПРИМЕР 4.38. Пусть

$$A = [A_{ij}] = \begin{bmatrix} 1 & -3 & 5 \\ 6 & 0 & -2 \end{bmatrix} \quad \text{и} \quad B = [B_{ij}] = \begin{bmatrix} -2 & 4 & 0 & 8 \\ 3 & -1 & -2 & 1 \\ 0 & 5 & 7 & 0 \end{bmatrix}.$$

Вычислим матричное произведение $C = [C_{ij}] = AB$ следующим образом. Матрица A имеет размер 2×3, а матрица B — размер 3×4, так что произведение $AB = C$ определено и будет представлять собой матрицу размера 2×4. C_{11} есть скалярное произведение первой строки матрицы A и первого столбца матрицы B, а C_{23} — скалярное произведение второй строки матрицы A и третьего столбца матрицы B. Таким образом,

$$C_{11} = \begin{bmatrix} 1 & -3 & 5 \end{bmatrix} \bullet \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix} = (1)(-2) + (-3)(3) + (5)(0) = -11 ;$$

$$C_{23} = \begin{bmatrix} 6 & 0 & -2 \end{bmatrix} \bullet \begin{bmatrix} 0 \\ -2 \\ 7 \end{bmatrix} = (6)(0) + (0)(-2) + (-2)(7) = -14 .$$

Продолжая вычисления в том же духе, получаем

$$AB = \begin{bmatrix} 1 & -3 & 5 \\ 6 & 0 & -2 \end{bmatrix} \begin{bmatrix} -2 & 4 & 0 & 8 \\ 3 & -1 & -2 & 1 \\ 0 & 5 & 7 & 0 \end{bmatrix} = \begin{bmatrix} -11 & 32 & 41 & 5 \\ -12 & 14 & -14 & 48 \end{bmatrix}. \qquad \square$$

Доказательство следующей теоремы можно найти в любой книге по линейной алгебре.

ТЕОРЕМА 4.39. Для любых матриц A, B и C размера $n \times n$ и действительных чисел r и s справедливы следующие утверждения:

а) $A + B = B + A$ (свойство коммутативности сложения);
б) $A + (B + C) = (A + B) + C$ (свойство ассоциативности сложения);
в) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ (свойство ассоциативности умножения);
г) $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$ (свойство дистрибутивности матриц);
д) $A \cdot (rB + sC) = r(A \cdot B) + s(A \cdot C)$ (свойство линейности матриц).

ОПРЕДЕЛЕНИЕ 4.40. Пусть A — матрица $n \times m$. Ее ***транспонированной матрицей*** называется матрица A^t размера $m \times n$ такая, что ${A^t}_{ij} = A_{ji}$, где A_{ij} — элемент i-ой строки и j-го столбца матрицы A. Если A — матрица $n \times n$ и $A_{ij} = A_{ji}$ для всех $1 \leq i, j \leq n$, то матрица A называется ***симметричной***. Иными словами, матрица A симметрична тогда и только тогда, когда $A = A^t$.

ПРИМЕР 4.41. Пусть $A = \begin{bmatrix} 1 & -3 & 5 \\ 6 & 0 & -2 \end{bmatrix}$. Тогда A^t — матрица размера 3×2, полученная из A заменой строк на столбцы:

$$A^t = \begin{bmatrix} 1 & -3 & 5 \\ 6 & 0 & -2 \end{bmatrix}^t = \begin{bmatrix} 1 & 6 \\ -3 & 0 \\ 5 & -2 \end{bmatrix}.$$

□

ТЕОРЕМА 4.42. Пусть A и B — матрицы, произведение которых определено. Тогда $(AB)^t = B^t A^t$.

Доказательство этой теоремы мы предоставляем читателю.

ОПРЕДЕЛЕНИЕ 4.43. Пусть $A = \{a_1, a_2, \ldots, a_n\}$ и $B = \{b_1, b_2, \ldots, b_m\}$, и пусть R — отношение на $A \times B$. ***Матричным представлением*** R называется матрица $M = [M_{ij}]$ размера $n \times m$, определенная соотношениями

$$M_{ij} = \begin{cases} 1, & \text{если } (a_i, b_j) \in R; \\ 0, & \text{если } (a_i, b_j) \notin R. \end{cases}$$

ПРИМЕР 4.44. Пусть $A = \{a_1, a_2, a_3\}$ и $B = \{b_1, b_2\}$. Пусть R — отношение $\{(a_1, b_1), (a_2, b_1), (a_2, b_2), (a_3, b_2)\}$. Тогда матричное представление имеет вид

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

□

ПРИМЕР 4.45. Наименьшее рефлексивное отношение на $A = \{a_1, a_2, a_3\}$ есть отношение $\{(a_1, a_1), (a_2, a_2), (a_3, a_3)\}$. Матричное представление этого отношения имеет вид

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Наибольшее отношение на A есть $A \times A$. Его матричное представление имеет вид

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

□

ОПРЕДЕЛЕНИЕ 4.46. Определим булевы операции $\vee$ и $\wedge$ на множестве $\{0,1\}$ следующим образом:

$\vee$	0	1
0	0	1
1	1	1

и

$\wedge$	0	1
0	0	0
1	0	1

.

Булевой матрицей называется матрица, каждый элемент которой есть 0 или 1. Пусть A и B — булевы матрицы размера $m \times n$, а C — булева матрица размера $n \times p$. Тогда

а) $U = A \vee B$ определяется соотношением

$$U_{ij} = A_{ij} \vee B_{ij}, \quad \text{где } 1 \le i \le m, 1 \le j \le n;$$

б) $I = A \wedge B$ определяется соотношением

$$I_{ij} = A_{ij} \wedge B_{ij}, \quad \text{где } 1 \le i \le m, 1 \le j \le n;$$

в) $D = A \odot C$ определяется соотношением

$$D_{ij} = (A_{i1} \wedge B_{1j}) \vee (A_{i2} \wedge B_{2j}) \vee \cdots \\ \cdots \vee (A_{in} \wedge B_{nj}), \quad \text{где } 1 \le i \le m,\ 1 \le j \le p.$$

ПРИМЕР 4.47. Пусть $A = \{a_1, a_2, a_3\}$, $R = \{(a_1,a_1),(a_1,a_2),(a_2,a_1),(a_2,a_2),(a_3,a_1)\}$ и $S = \{(a_1,a_1),(a_1,a_3),(a_3,a_1),(a_2,a_2),(a_2,a_3),(a_3,a_2)\}$. Матричными представлениями отношений R и S, соответственно, являются

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Тогда

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \wedge \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vee \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \odot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

□

ТЕОРЕМА 4.48. Пусть R и S — отношения на конечном множестве $A = \{a_1, a_2, \ldots, a_n\}$ и им соответствуют матричные представления M и N.

а) Если R рефлексивно, то $M_{ii} = 1$ для $1 \leq i \leq n$.
б) Если R симметрично, то $M_{ij} = M_{ji}$ для $1 \leq i, j \leq n$; поэтому $M = M^t$, и M — симметричная матрица.
в) Пусть $M^{\odot 2} = M \odot M$. Если R транзитивно, то, как только $M_{ik}^{\odot 2} = 1$, мы имеем $M_{ik} = 1$ для $1 \leq i, k \leq n$.
г) $M \vee N$ является матричным представлением $R \cup S$.
д) $M \wedge N$ является матричным представлением $R \cap S$.

ДОКАЗАТЕЛЬСТВО. а) Поскольку отношение R рефлексивно, то $(a_i, a_i) \in R$ для всех $1 \leq i \leq n$. Поэтому $M_{ii} = 1$ для всех $1 \leq i \leq n$.

б) Если $M_{ij} = 1$, то $(a_i, a_j) \in R$, а поскольку R симметрично, то $(a_j, a_i) \in R$, так что $M_{ji} = 1$. Обратно, если $M_{ji} = 1$, то $(a_j, a_i) \in R$, а поскольку R симметрично, то $(a_i, a_j) \in R$, так что $M_{ij} = 1$. В силу того, что $M_{ij} = 1$ тогда и только тогда, когда $M_{ji} = 1$, $M = M^t$, M — симметричная матрица.

в) Допустим, что $M_{ik}^{\odot 2} = 1$. Поскольку $M_{ik}^{\odot 2} = (A_{i1} \wedge A_{1k}) \vee (A_{i2} \wedge A_{2k}) \vee \cdots \vee (A_{in} \wedge A_{nk})$, то существует такое m, что $(A_{im} \wedge A_{mk}) = 1$. Поэтому $A_{im} = 1$ и $A_{mk} = 1$. По определению матричного представления $(a_i, a_m) \in R$ и $(a_m, a_k) \in R$. Поскольку R транзитивно, то $(a_i, a_k) \in R$ и $M_{ik} = 1$.

Оставшуюся часть доказательства предоставляем читателю. ■

Доказательство следующей теоремы мы также предоставляем читателю.

ТЕОРЕМА 4.49. Пусть R — отношение на $A \times B$ и S — отношение на $B \times C$. Если M и N — матричные представления R и S соответственно, то $M \odot N$ — матричное представление $S \circ R$.

ПРИМЕР 4.50. Пусть $A = \{a_1, a_2\}$, $B = \{b_1, b_2, b_3\}$ и $C = \{c_1, c_2, c_3, c_4\}$. Пусть $R = \{(a_1, b_1), (a_1, b_2), (a_2, b_3)\}$ и $S = \{(b_1, c_1), (b_1, c_2), (b_2, c_3), (b_3, c_1), (b_3, c_2)\}$. Тогда матричными представлениями R и S являются, соответственно,

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Далее, имеем

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \odot \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

и $S = \{(a_1, c_1), (a_1, c_2), (a_1, c_3), (a_2, c_1), (a_2, c_2)\}$. □

Следующее утверждение является следствием теорем 4.48, 4.49 и 2.37.

ТЕОРЕМА 4.51. Пусть A — матричное представление отношения R.

а) Рефлексивное замыкание R имеет матричное представление $A \cup I$.
б) Симметричное замыкание R имеет матричное представление $A \cup A^t$.
в) Если A — конечное множество, содержащее n элементов, тогда транзитивное замыкание R имеет матричное представление $A \cup A^{\odot 2} \cup A^{\odot 3} \cup \cdots \cup A^{\odot n}$.

ОПРЕДЕЛЕНИЕ 4.52. Пусть M — матрица размера $n \times n$, в каждой строке и в каждом столбце которой только один элемент равен 1, а все остальные равны 0. Такая матрица M называется ***матрицей перестановок***.

ТЕОРЕМА 4.53. Пусть M — матрица перестановок, а I — матрица размера $n \times 1$ с различными элементами. Тогда MI — матрица размера $n \times 1$, элементы которой образуют перестановку элементов матрицы I.

ДОКАЗАТЕЛЬСТВО. Пусть $I = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$ и $I' = MI$. Рассмотрим a_j в j-ой строке. Поскольку

$$C_i = \begin{bmatrix} M_{i1} & M_{i2} & M_{i3} & \cdots & M_{in} \end{bmatrix} \bullet \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = \sum_{k=1}^{n} M_{ik} a_k$$

и имеется одна строка p, в которой $M_{pj} = 1$ и $M_{pk} = 0$ для $k \neq j$, то p-я строка матрицы I' есть a_j. Далее, поскольку $M_{kj} = 0$ для всех $k \neq p$, a_j не может появиться ни в какой другой строке матрицы I'. ■

ПРИМЕР 4.54. Пусть $M = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ и $I = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$. Тогда $MI = \begin{bmatrix} a_3 \\ a_1 \\ a_2 \end{bmatrix}$. □

■ УПРАЖНЕНИЯ

1. Вычислите:

а) $\begin{bmatrix} 3 & -1 \\ 2 & 7 \\ -5 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 4 \\ 5 & 1 & 3 & 0 \end{bmatrix}$; **б)** $\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 6 & -1 & 5 \end{bmatrix} \begin{bmatrix} 2 & -2 & 8 \\ 5 & 5 & 1 \\ 0 & 6 & 4 \end{bmatrix}$;

в) $\begin{bmatrix} 1 \\ 3 \\ 4 \\ 5 \end{bmatrix} \begin{bmatrix} -2 & 3 & -5 & 0 \end{bmatrix}$; **г)** $\begin{bmatrix} -2 & 3 & -5 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \\ 5 \end{bmatrix}$;

д) $(-2) \begin{bmatrix} -5 & 1 \\ 6 & -7 \end{bmatrix}$; **е)** $\begin{bmatrix} -1 & 3 \\ 2 & 7 \end{bmatrix} - 8 \begin{bmatrix} 3 & 4 \\ -8 & 1 \end{bmatrix}$.

2. Пусть $A = \begin{bmatrix} -2 & 4 & 7 \\ 6 & 1 & 5 \\ 2 & 3 & -4 \end{bmatrix}$. Вычислите:

а) A^t; **б)** AA^t; **в)** A^tAx.

3. Пусть $A = \begin{bmatrix} -3 & 2 \\ 7 & 5 \end{bmatrix}$ и $B = \begin{bmatrix} 4 & 1 \\ -2 & 9 \end{bmatrix}$. Вычислите:

а) AB; **б)** BA; **в)** AB^t; **г)** A^tB; **д)** $(AB)^t$;

е) B^tA^t; **ж)** $A-B$; **з)** $B-A$; **и)** $5A$; **к)** $2B-3A$.

4. Пусть $A=\begin{bmatrix}1 & 0 & 5\\ 2 & -4 & 6\end{bmatrix}$ и $B=\begin{bmatrix}-1 & 2\\ 3 & 9\\ 10 & 0\end{bmatrix}$. Вычислите:

а) AB; **б)** BA; **в)** A^t;
г) B^t; **д)** A^tB^t; **е)** $(BA)^t$.

5. Докажите справедливость или постройте контрпример для высказывания *Если для матриц A, B и C имеет место $AB=AC$, то $B=C$.*

6. Докажите справедливость или постройте контрпример для высказывания *Если для матриц A и B произведение AB равно нулевой матрице, то либо A, либо B является нулевой матрицей.*

7. Докажите, что если произведение матриц A и B определено, то $(AB)^t=B^tA^t$.

8. Докажите, что для отношений R и S на конечном множестве $A=\{a_1,a_2,\ldots,a_n\}$ с матричными представлениями M и N, соответственно, $M\vee N$ является матричным представлением отношения $R\cup S$.

9. Докажите, что для отношений R и S на конечном множестве $A=\{a_1,a_2,\ldots,a_n\}$ с матричными представлениями M и N, соответственно, $M\wedge N$ является матричным представлением отношения R.

10. Докажите, что если R — отношение на $A\times B$, а S — отношение на $B\times C$ и если M и N — матричные представления R и S, соответственно, тогда $M\odot N$ является матричным представлением отношения $S\circ R$.

11. Пусть $A=\{1,2,3,4,5\}$, $B=\{6,7,8,9\}$, $C=\{10,11,12,13\}$, $D=\{\square,\triangle,\bigcirc,*\}$. Отношения $R\subseteq A\times B$, $S\subseteq B\times C$ и $T\subseteq C\times D$ определены как

$R=\{(1,7),(4,6),(5,6),(2,8)\}$,

$S=\{(6,10),(6,11),(7,10),(8,13)\}$,

$T=\{(11,\triangle),(10,\triangle),(13,*),(12,\square),(13,\bigcirc)\}$.

Найдите матричные представления отношений R, S и T, затем найдите матричные представления отношений

а) R^{-1} и S^{-1}; **б)** $S\circ R$; **в)** $S\circ S^{-1}$ и $S^{-1}\circ S$;
г) $R^{-1}\circ S^{-1}$; **д)** $T\circ(S\circ R)$; **е)** $T\circ S$;
ж) $(T\circ S)\circ R$.

12. Пусть $A=\{(b,a),(c,e),(d,i),(f,o),(g,u)\}$ и $B=\{(v,a),(w,e),(x,i),(y,o),(z,u)\}$. Найдите матричные представления A и B. Затем найдите матричные представления отношений

а) A^{-1}; **б)** B^{-1}; **в)** $A^{-1}\circ B$; **г)** $B^{-1}\circ A$.

13. Пусть $A=\{a,b,c,d,e\}$ и S,T,U и V — отношения на A, где

$S=\{(a,a),(a,b),(b,c),(b,d),(c,e),(e,d),(c,a)\}$,
$T=\{(a,b),(b,a),(b,c),(b,d),(e,e),(d,e),(c,b)\}$,
$U=\{(a,b),(a,a),(b,c),(b,b),(e,e),(b,a),(c,b),(c,c),(d,d),(a,c),(c,a)\}$,
$V=\{(a,b),(b,c),(b,b),(e,e),(b,a),(c,b),(d,d),(a,c),(c,a)\}$.

Найдите матричные представления S, T, U V. Затем, используя эти матри-

цы, определите какие из этих отношений

а) симметричны;

б) рефлексивны;

в) транзитивны;

г) антисимметричны.

14. Пусть $A = \{a, b, c, d, e\}$ и S, T, U и V — отношения на A, где

$$S = \{(a,a),(a,b),(b,c),(b,d),(c,e),(e,d),(c,a)\},$$
$$T = \{(a,b),(b,a),(b,c),(b,d),(e,e),(d,e),(c,b)\},$$
$$U = \{(a,b),(a,a),(b,c),(b,b),(e,e),(b,a),(c,b),(c,c),(d,d),(a,c),(c,a)\},$$
$$V = \{(a,b),(b,c),(b,b),(e,e),(b,a),(c,b),(d,d),(a,c),(c,a)\}.$$

а) Найдите матричное представление S, T, U и V.

б) Найдите матричное представление отношений $S \circ T$ и $U \circ V$.

в) Найдите матричное представление отношений $S \cap T$ и $U \cap V$.

г) Найдите матричное представление $U \triangle S$.

15. Пусть $A = \{1, 2, 3, 4, \}$ $B = \{a, b, c\}$. Найдите отношение на $A \times B$ с матричным представлением

а) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$; **б)** $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$; **в)** $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

16. Пусть $A = \{a, b, c, d, e\}$. Найдите отношение на A с матричным представлением

а) $\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$; **б)** $\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}$;

в) $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$; **г)** $\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$.

17. Пусть $A = \{a, b, c, d, e\}$. Постройте графы со следующими матричными представлениями:

а) $\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$; **б)** $\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}$; **в)** $\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$.

18. Пусть $A = \{a, b, c, d, e\}$. Постройте орграфы со следующими матричными представлениями:

а) $\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}$; **б)** $\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}$; **в)** $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$.

19. Для каждого отношения R, заданного в предыдущем упражнении матричным представлением, найдите $R \circ R$, используя соответствующее матричное представление.

4.4. МОЩНОСТЬ

В этом разделе мы обсудим понятие мощности множества. Если множество конечно, то его мощность есть просто количество содержащихся в нем элементов. Таким образом, для определения мощности множества нам достаточно подсчитать количество содержащихся в нем элементов. Сформулируем сказанное формально в виде следующего определения.

ОПРЕДЕЛЕНИЕ 4.55. Пустое множество есть ***конечное множество мощности 0***. Если существует взаимно однозначное соответствие между множеством A и множеством $\{1, 2, 3, \ldots n\}$, то A есть ***конечное множество мощности n***.

ПРИМЕР 4.56. Множество $\{a, p, r, x, z\}$ имеет мощность 5, т.к. существует взаимно однозначное соответствие: $\{1, 2, 3, 4, 5\} \rightarrow \{a, p, r, x, z\}$, определяемое равенствами $\phi(1) = a$, $\phi(2) = p$, $\phi(3) = r$, $\phi(4) = x$, $\phi(5) = z$. □

ОПРЕДЕЛЕНИЕ 4.57. Множество A называется ***счетно бесконечным***, если существует взаимно однозначное соответствие между множеством A и множеством положительных целых чисел $\{1, 2, 3, \ldots, n, \ldots\}$. Множество называется ***счетным***, если оно конечно или счетно бесконечно.

ТЕОРЕМА 4.58.

а) Пусть A и B — непересекающиеся конечные множества. Тогда множество $A \cup B$ конечно. Если A имеет мощность n и B имеет мощность m, то $A \cup B$ имеет мощность $m + n$.

б) Пусть A и B — непересекающиеся счетно бесконечные множества. Тогда $A \cup B$ — счетно бесконечное множество.

в) Пусть A и B — непересекающиеся счетные множества. Тогда $A \cup B$ — счетное множество.

ДОКАЗАТЕЛЬСТВО. а) Поскольку множества A и B конечны, то для положительных целых чисел n и m существуют взаимно однозначные соответствия $\phi_1 : \{1, 2, 3, \ldots, n\} \rightarrow A$ и $\phi_2 : \{1, 2, 3, \ldots, m\} \rightarrow B$. Определим отображение $\phi : \{1, 2, 3, \ldots, m + n\} \rightarrow A \cup B$ посредством равенств $\phi(k) = \phi_1(k)$ для $1 \leq k \leq n$

и $\phi(n+i) = \phi_2(i)$ для $1 \leq i \leq m$. Функция ϕ устанавливает взаимно однозначное соответствие между множествами $\{1, 2, 3, \ldots, m+n\}$ и $A \cup B$.

б) Поскольку A и B — счетно бесконечные множества, то существуют взаимно однозначные соответствия $\phi_1 : \{1, 2, 3, \ldots n, \ldots\} \to A$ и $\phi_2 : \{1, 2, 3, \ldots m, \ldots\} \to B$. Определим функцию $\phi : \{1, 2, 3, \ldots\} \to A \cup B$ посредством равенств $\phi(n) = \phi_1((n+1)/2)$ в случае, когда n нечетно, и $\phi(n) = \phi_2(n/2)$ в случае, когда n четно. Функция ϕ устанавливает взаимно однозначное соответствие между множествами $\{1, 2, 3, \ldots\}$ и $A \cup B$.

в) Пусть A и B — непересекающиеся счетные множества. Единственный случай, который остался не рассмотренным в пунктах (а) и (б), — это когда одно из множеств A или B конечно, а другое счетно бесконечно. Предположим, что A — конечное множество. Тогда существует взаимно однозначное соответствие $\phi_1 : \{1, 2, 3, \ldots, n\} \to A$. А поскольку B — счетно бесконечное множество, то существует $\phi_2 : \{1, 2, 3, \ldots, m, \ldots\} \to B$. Определим $\phi : \{1, 2, 3, \ldots\} \to A \cup B$ следующим образом: $\phi(k) = \phi_1(k)$ для $1 \leq k \leq n$ и $\phi(k) = \phi_2(k-n)$ для $k > n$. Функция ϕ устанавливает взаимно однозначное соответствие между множествами $\{1, 2, 3, \ldots\}$ и $A \cup B$. ■

Доказательство следующей теоремы предоставляется читателю.

ТЕОРЕМА 4.59. Множество неотрицательных целых чисел и множество целых чисел являются счетно бесконечными.

Дальнейшие свойства мощности будут рассмотрены в следующем разделе. Будет показано, что существуют множества, которые не являются счетными, и что имеется бесконечно много различных бесконечных мощностей.

■ УПРАЖНЕНИЯ

1. Используя определение, найдите мощности следующих множеств:
   а) $\{a, b, c, d, e, f, g\}$;
   б) $\{3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33\}$;
   в) $\{5, 10, 15, 20, 25, 30, \ldots\}$.
2. Покажите, что множество неотрицательных целых чисел счетно бесконечно.
3. Покажите, что множество отрицательных целых чисел счетно бесконечно.
4. Покажите, что множество целых чисел счетно бесконечно.
5. Покажите, что множество $\{-10, -9, -8, \ldots, -2, -1, 0, 1, 2, 3, \ldots\}$ счетно бесконечно.

4.5. МОЩНОСТЬ (ПРОДОЛЖЕНИЕ)

Исследуем теперь понятие мощности множества более основательно. Прежде всего, покажем, что подмножество счетного множества счетно. Отсюда будет следовать, что всякое множество, для которого существует взаимно однозначное соответствие с подмножеством счетного множества, счетно.

ТЕОРЕМА 4.60. Подмножество счетного множества счетно.

ДОКАЗАТЕЛЬСТВО. Покажем, что если B — счетное множество и $A \subseteq B$, то множество A также счетно. Предположим, что B — конечное множество. Покажем, что в таком случае A тоже конечно. Поскольку B конечно, для некоторого положительного целого числа n существует взаимно однозначное соответствие ϕ между $\{1,2,3,\ldots,n\}$ и B. Следовательно, B можно представить в виде $\{b_1, b_2, b_3, \ldots, b_n\}$, где $b_i = \phi(i)$. Теперь построим взаимно однозначное соответствие $\overline{\phi}$ между $\{1,2,3,\ldots,m\}$ и A, где $m \leq n$. Начинаем с b_1 и находим первое b_i, скажем, b_{a_1}, в последовательности b_1, b_2, b_3, ..., b_n такое, что $b_i \in A$. Пусть $\overline{\phi}(1) = b_{a_1}$. Просматривая последовательность, находим в ней следующее b_i, скажем, b_{a_2} такое, что $b_i \in A$. Пусть $\overline{\phi}(2) = b_{a_2}$. Продолжаем процесс, пока не достигнем b_n. Пусть m — наибольшее целое число, которое $\overline{\phi}$ отображает в элемент A. Поскольку $i \leq a_i$ для всех i, то $m \leq n$ и поэтому, несомненно, $\overline{\phi}$ отображает $\{1,2,3,\ldots,m\}$ в A.

Теперь предположим, что B — счетно бесконечное множество. Если A — конечное, то оно счетное, поэтому предположим, что оно не является конечным. Поскольку B — счетно бесконечное, существует взаимно однозначное соответствие ϕ между $\{1,2,3,\ldots\}$ и B. Следовательно, B можно представить в виде $\{b_1, b_2, b_3, \ldots\}$, где $b_i = \phi(i)$. Используя принцип вполне упорядочения, находим наименьшее i такое, что $b_i \in A$. Пусть $\overline{\phi}(1) = b_i$ и $B_1 = B - \{b_i\}$. Очевидно, что $1 \leq i$. Предположим, что мы определили $\overline{\phi}(1), \overline{\phi}(2), \overline{\phi}(3), \ldots, \overline{\phi}(k) \in A$, $B_k = B - \{\overline{\phi}(1), \overline{\phi}(2), \overline{\phi}(3), \ldots, \overline{\phi}(k)\}$, $\overline{\phi}(i) \leq b$ для всех $1 \leq i \leq k$ и всех $b \in B_k$, и если $\overline{\phi}(k) = b_r$, то $k \leq r$. Выберем наименьшее j такое, что $b_j \in A \cap B_k$. Пусть $\overline{\phi}(k) = b_j$ и $B_{k+1} = B_k - \{b_j\}$. Очевидным образом получаем последовательность $\overline{\phi}(1), \overline{\phi}(2), \overline{\phi}(3), \ldots, \overline{\phi}(k), \overline{\phi}(k+1) \in A$, $B_{k+1} = B - \{\overline{\phi}(1), \overline{\phi}(2), \overline{\phi}(3), \ldots, \overline{\phi}(k), \overline{\phi}(k+1)\}$, $\overline{\phi}(i) \leq b$ для всех $1 \leq i \leq k+1$ и всех $b \in B_{k+1}$, и если $\overline{\phi}(k+1) = b_r$, то $k+1 \leq r$. Итак, функция $\overline{\phi}$ определена по индукции, при этом очевидно, что $\overline{\phi}(i) \in A$ для всех i и $\overline{\phi}(i) \neq \overline{\phi}(j)$, если $i \neq j$. Остается только показать, что отображение $\overline{\phi}$ сюръективно. Пусть $a \in A$, тогда $a = b_n$ для некоторого n. Следовательно, для некоторого $m \leq n$ имеем $\overline{\phi}(m) = b_n$. ■

ТЕОРЕМА 4.61. Пусть S — счетно бесконечное множество, тогда множество $S \times S$ также счетно бесконечно.

ДОКАЗАТЕЛЬСТВО. Сначала покажем, что если N — множество положительных целых чисел, то $N \times N$ счетно. Рассмотрим следующий рисунок:

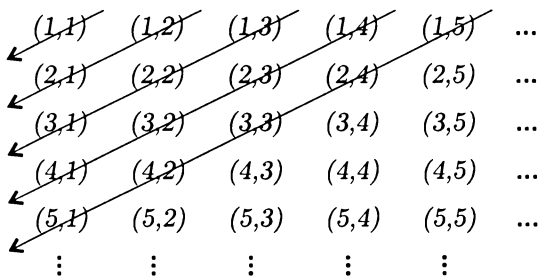


Рис. 4.1

Двигаясь по диагональным стрелкам, начиная с верхней, мы определяем отображение ϕ следующим образом: $\phi(1) = (1,1)$, $\phi(2) = (1,2)$, $\phi(3) = (2,1)$,

$\phi(4) = (1,3,)$, $\phi(5) = (2,3)$ и т.д. Функция ϕ устанавливает искомое взаимно однозначное соответствие. Упорядоченная пара (m,n) лежит на $m+n-1$-ой по порядку диагонали и является m-ым элементом вдоль этой диагональной линии. Следовательно, множество, изображенное на рис. 4.1, счетно. Таким образом, множество $N \times N$ счетно. Поскольку S — счетно бесконечное множество, то существует взаимно однозначное соответствие $\theta : N \to S$. Определим соответствие $\theta \times \theta : N \times N \to S \times S$ следующим образом: $\theta \times \theta(a,b) = (\theta(a), \theta(b))$. Легко показать, что функция $\theta \times \theta$ устанавливает взаимно однозначное соответствие между множествами $N \times N$ и $S \times S$. ■

ТЕОРЕМА 4.62. Множество Q^+ положительных рациональных чисел является счетно бесконечным.

ДОКАЗАТЕЛЬСТВО. Рассмотрим подмножество M множества $N \times N$ вида $\{(a,b) : (a,b) \in N \times N$, при этом целые числа a и b — взаимно простые$\}$. Очевидно, функция $\phi : Q^+ \to M$, определенная соотношением $\phi(a/b) = (a,b)$, — это и есть искомое взаимно однозначное соответствие. Поскольку множество $N \times N$ — счетно бесконечное, множество M также является счетно бесконечным. Следовательно, Q^+ — счетно бесконечное множество. ■

ТЕОРЕМА 4.63. Если A и B — счетные множества, то множество $A \cup B$ также счетно.

ДОКАЗАТЕЛЬСТВО. Множество $A - B$ счетно, как подмножество счетного множества A. Но множества $A - B$ и B не пересекаются, поэтому по теореме 4.58 (в) множество $(A - B) \cup B = A \cup B$ является счетным. ■

Приведенные выше теоремы могут навести на мысль, что все множества счетны. Однако, как показывает следующая теорема, существуют бесконечные, но несчетные множества.

ТЕОРЕМА 4.64. Пусть R — множество действительных чисел. Множество I вида $I = \{x : x \in R$ и $0 < x < 1\}$ не является счетным.

ДОКАЗАТЕЛЬСТВО. Применим метод сведения к абсурду. Предположим, что множество I счетно. Тогда I можно представить в виде $\{a_1, a_2, a_3, a_4, a_5, \ldots\}$, где a_i — образ i при взаимно однозначном соответствии между N и I. Пусть $.a_{i1}a_{i2}a_{i3}a_{i4}a_{i5}\ldots$ — это запись в виде десятичной дроби числа a_i. Тогда мы имеем таблицу

$$\begin{aligned}
a_1 &= .a_{11}a_{12}a_{13}a_{14}a_{15}\ldots \\
a_2 &= .a_{21}a_{22}a_{23}a_{24}a_{25}\ldots \\
a_3 &= .a_{31}a_{32}a_{33}a_{34}a_{35}\ldots \\
a_4 &= .a_{41}a_{42}a_{43}a_{44}a_{45}\ldots \\
a_5 &= .a_{51}a_{52}a_{53}a_{54}a_{55}\ldots \\
&\vdots \quad \vdots \qquad \vdots
\end{aligned}$$

Определим элемент $b \in I$, где $b = .b_1b_2b_3b_4b_5\ldots$, следующим образом: $b_i = 5$, если $a_{ii} \neq 5$, и $b_i = 8$, если $a_{ii} = 5$. Рассмотрим конкретный пример:

$$a_1 = .5135897623\ldots$$

$$
\begin{aligned}
a_2 &= .1425958273\ldots \\
a_3 &= .2837462982\ldots \\
a_4 &= .1985723986\ldots \\
a_5 &= .2309847621\ldots \\
&\vdots\ \ \vdots \qquad \vdots
\end{aligned}
$$

Имея $a_{11} = 5$, $a_{22} = 4$, $a_{33} = 3$, $a_{44} = 5$, $a_{55} = 8$ и т.д., получаем $b_1 = 8$, $b_2 = 5$, $b_3 = 5$, $b_4 = 8$, $b_5 = 5$ и т.д. Таким образом,

$$b = .85585\ldots$$

Число b не принадлежит множеству $\{a_1, a_2, a_3, a_4, a_5, \ldots\}$, т.к. для каждого m число b не может совпадать с a_m. В самом деле, $b_m \neq a_{mm}$, поэтому b и a_m не совпадают в m-ой цифре. Но это противоречит предположению о том, что $I = \{a_1, a_2, a_3, a_4, a_5, \ldots\}$, так как $b \in I$, но $b \notin \{a_1, a_2, a_3, a_4, a_5, \ldots\}$. Следовательно, предположение о счетности I является ложным, и множество I — несчетное. ■

ТЕОРЕМА 4.65. Множество действительных чисел R несчетно.

ДОКАЗАТЕЛЬСТВО. Если бы R было счетным, то множество I, как подмножество R, было бы счетным, однако, I — несчетно. Следовательно, множество R — несчетное. ■

В дальнейшем мы опишем метод построения бесконечного числа бесконечных множеств, никакие два из которых нельзя связать взаимно однозначным соответствием.

ТЕОРЕМА 4.66. Не существует взаимно однозначного соответствия между множеством S и его булеаном $\mathcal{P}(S)$.

ДОКАЗАТЕЛЬСТВО. Очевидно, не существует взаимно однозначного соответствия между пустым множеством $\varnothing$ и его булеаном $\{\varnothing\}$, содержащим один элемент — $\varnothing$. Снова воспользуемся методом сведения к абсурду. Предположим, что взаимно однозначное соответствие ϕ между S и его булеаном существует. Это значит, что функция ϕ биективно отображает элементы множества S на подмножества множества S. Например, пусть S — множество положительных целых чисел, и пусть

$$
\begin{aligned}
\phi(1) &= \{2, 5, 7, 9, 10\}, \\
\phi(2) &= \{2, 4, 6, 8, \ldots\}, \\
\phi(3) &= \varnothing, \\
\phi(4) &= \{1, 2, 3, 4, 5, 6, \ldots\}, \\
\phi(5) &= \{1, 2\}, \\
&\vdots
\end{aligned}
$$

В некоторых случаях $i \in \phi(i)$. В нашем примере $2 \in \phi(2)$ и $4 \in \phi(4)$. Однако, $1 \notin \phi(1)$, $3 \notin \phi(3)$ и $5 \notin \phi(5)$. Понятно, что любой элемент i, который функция ϕ

отображает в пустое множество, будет обладать свойством $i \notin \phi(i)$, а всякий элемент j, который функция ϕ отображает в S, будет обладать свойством $j \in \phi(j)$. Пусть $W = \{x : x \in S \text{ и } x \notin \phi(x)\}$. Поскольку отображение ϕ сюръективно, то существует элемент $a \in S$ такой, что $\phi(a) = W$. Принадлежит ли a множеству W? Если $a \in W$, то a принадлежит множеству тех элементов S, которые ϕ не отображает на множества, их содержащие. Следовательно, $a \notin \phi(a) = W$. Таким образом, мы приходим к противоречию. Если же $a \notin W$, то $a \notin \phi(a) = W$. Поэтому a удовлетворяет условию принадлежности множеству W, т.е. $a \in W$. Снова получаем противоречие. Итак, в любом случае мы приходим к противоречию. Следовательно, утверждение о существовании взаимно однозначного соответствия ϕ между S и его булеаном неверно. ■

ПРИМЕР 4.67. Пусть $S = \{1, 2, 3, 4\}$ и

$$\begin{aligned}\phi(1) &= \{2, 4\},\\ \phi(2) &= \{1, 2, 3, 4\},\\ \phi(3) &= \{1, 3\},\\ \phi(4) &= \varnothing.\end{aligned}$$

Тогда $W = \{1, 4\}$ — множество из доказательства теоремы, и в него не отображается никакой элемент. □

Примером рассуждения, очень похожего на доказательство предыдущей теоремы, является парадокс Рассела. Он показывает нам, что при определении множеств нужно быть очень осторожным. Иначе в своей математической системе мы можем получить внутреннее противоречие. Особое внимание следует обращать на то, чтобы множества не становились "слишком большими".

ПРИМЕР 4.68. (Парадокс Рассела) Пусть S — множество всех множеств. Пусть $W = \{x : x \notin x\}$. Пустое множество принадлежит W, т.к. оно не принадлежит самому себе. На самом деле, большинство известных нам множеств принадлежит W. Однако, S не принадлежит W, т.к. $S \in S$. А вот принадлежит ли множеству W само W? Если $W \in W$, то оно принадлежит множеству всех множеств, которые сами себе не принадлежат. Следовательно, $W \notin W$. Однако, если предположить, что $W \notin W$, то W принадлежит множеству всех множеств, которые не принадлежат сами себе. Но таким множеством является W, так что $W \in W$. В любом случае мы приходим к противоречию. □

■ УПРАЖНЕНИЯ

1. Пусть $S = \{1, 2, 3, 4, 5, 6\}$ и ϕ — инъективная функция из S в его булеан $\mathcal{P}(S)$, определенная следующим образом:

$$\begin{aligned}\phi(1) &= \{1, 2, 3, 4\},\\ \phi(2) &= \{1, 4\},\\ \phi(3) &= \{2, 3, 4\},\\ \phi(4) &= \varnothing,\end{aligned}$$

$$\phi(5) = \{1, 2, 3, 4, 5, 6\},$$
$$\phi(6) = \{1, 3, 6\}.$$

Опишите множество, отсутствие отображения на которое гарантирует нам теорема 4.66.

2. Пусть $S = \{1, 2, 3, 4, 5, 6\}$ и ϕ — инъективная функция из S в его булеан $\mathcal{P}(S)$, определенная следующим образом:

$$\phi(1) = \{2, 3, 4\},$$
$$\phi(2) = \{1, 4\},$$
$$\phi(3) = \{2, 4\},$$
$$\phi(4) = \varnothing,$$
$$\phi(5) = \{1, 2, 3, 4, 6\},$$
$$\phi(6) = \{1, 3\}.$$

Опишите множество, отсутствие отображения на которое гарантирует нам теорема 4.66.

3. Пусть $S = \{1, 2, 3, 4, 5, 6\}$ и ϕ — инъективная функция из S в его булеан $\mathcal{P}(S)$, определенная следующим образом:

$$\phi(1) = \{1, 2, 3, 4\},$$
$$\phi(2) = \{1, 2, 4\},$$
$$\phi(3) = \{2, 3, 4, 5\},$$
$$\phi(4) = \{4\},$$
$$\phi(5) = \{1, 2, 3, 4, 5, 6\},$$
$$\phi(6) = \{1, 3, 6\}.$$

Опишите множество, отсутствие отображения на которое гарантирует нам теорема 4.66.

4. Докажите, что множество неотрицательных рациональных чисел счетно.

5. Докажите, что множество рациональных чисел счетно.

6. Пусть $S = \{a + bi : a, b \in Z\}$, где $i = \sqrt{-1}$. Докажите, что S счетно.

7. Докажите, что множество всех полиномов 5-ой степени с целыми коэффициентами счетно.

8. Докажите, что множество всех полиномов степени не выше n с целыми коэффициентами является счетным.

ГЛАВА

5 АЛГОРИТМЫ И РЕКУРСИЯ

5.1. ЦИКЛЫ И АЛГОРИТМЫ ДЛЯ МАТРИЦ

Алгоритмом называется совокупность инструкций о том, как решить некоторую задачу. Множество зачастую непонятных руководств, прилагаемых ко всевозможным устройствам и обучающих нас, как устанавливать и использовать их, являются примерами алгоритмов. Компьютеру необходимы инструкции, объясняющие, как выполнить задание, поэтому любая компьютерная программа также представляет собой алгоритм. Всегда желательно, чтобы алгоритм носил максимально общий характер. Например, существенно, что алгоритм умножения двух матриц работает в случае любых двух матриц, для которых операция умножения определена. Как мы увидим в дальнейшем, многие алгоритмы содержат периодически повторяющиеся процессы. Например, при умножении двух матриц размера $n \times n$ используется определенная процедура нахождения элементов матрицы-произведения, которая повторяется n^2 раз. Вообще, важность алгоритмов, содержащих повторяющиеся процедуры, для решения задач при помощи компьютеров трудно переоценить.

Введенная нами ранее операция суммирования последовательности, по сути, представляет собой определенный алгоритм. Знак суммы $\sum_{i=1}^{n} p(i)$ говорит о том, что мы должны вычислить элемент $p(i)$ при $i = 1$, затем положить $i = 2$, вычислить $p(i)$ и прибавить его к предыдущей сумме. Далее мы увеличиваем i на единицу и повторяем описанную процедуру до тех пор, пока i не превзойдет n. По аналогии с описанным процессом вводится следующая процедура "цикла перечислением". В соответствии с приведенным ниже набором инструкций, требуется положить $r = 1$ и выполнить шаги с (1)-го по (m)-ый. Затем положить $r = 2$ и повторить процесс.

Цикл по r от 1 до n:
    Шаг (1);
    Шаг (2);
    ⋮
    Шаг (m);
Конец цикла.

Выполнив процесс для $r = k$, нужно положить $r = k + 1$ и повторять процесс до тех пор, пока не выполним его для $r = n$.

Таким образом, для того чтобы вычислить сумму $S = \sum_{i=1}^{n} p(i)$, мы должны:

Положить $S = 0$;
Цикл по i от 1 до n:
    Заменить значение S на $S + p(i)$;
Конец цикла.

В итоге получим значение S, равное $\sum_{i=1}^{n} p(i)$.

ПРИМЕР 5.1. Вычислить $M = r^n$.

Положить $M = 1$;
Цикл по i от 1 до n:
    Заменить значение M на $M * r$;
Конец цикла. □

ПРИМЕР 5.2. Вычислить $p(a)$, где $p(x)$ — полином, используя схему Горнера. Сначала заметим, что

$$a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = x(x(x(a_4x + a_3) + a_2) + a_1) + a_0\,.$$

Используя этот факт, построим алгоритм вычисления $S = p(a)$, где $p(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0$. Алгоритм имеет вид:

Положить $S = a_n$;
Цикл по i от 1 до n:
    Заменить S на $aS + a_{n-i}$;
Конец цикла. □

ОПРЕДЕЛЕНИЕ 5.3. Даны частичные упорядочения, $\leq_1$ и $\leq_2$, упорядочение $\leq_2$ называется ***топологической сортировкой*** упорядочения $\leq_1$, если $\leq_2$ является полным упорядочением, и всякий раз, когда $a \leq_1 b$, тогда $a \leq_2 b$.

Опишем достаточно простой алгоритм нахождения топологической сортировки для упорядочения $\leq_1$ на множестве S. Он состоит в следующем: мы находим произвольный минимальный элемент частичного упорядочения и объявляем его минимальным элементом полного упорядочения; затем выбираем другой минимальный элемент и объявляем его следующим элементом в полном упорядочении и т.д. Это действие повторяется, пока все элементы множества не будут выбраны. Мы всегда выбираем минимальный элемент, поэтому если в исходном упорядочении $a \leq b$, то это соотношение будет иметь место и в новом упорядочении, так как a будет выбрано первым. Теперь опишем алгоритм формально:

Процедура $\mathbf{TC}(S, \leq_1)$:
    Выбрать минимальный элемент s из $(S, \leq_1)$;
    Удалить s из S;
        Выполнять следующие шаги до тех пор, пока S не пусто;
        Выбрать минимальный элемент t из $(S, \leq_1)$ и удалить его;
        Положить $s \leq_1 t$;
        Переименовать t в s;
    Конец процедуры.

Если $A = [A_{ij}]$ — матрица размера $m \times n$, то произведение этой матрицы на скаляр a есть матрица $[B_{ij}] = a[A_{ij}]$ (см. главу 4). Алгоритм нахождения этого произведения можно описать следующим образом:

Процедура **Скаляр**(a, A, m, n):
  Цикл по i от 1 до m:
    Цикл по j от 1 до n:
      $B_{ij} = aA_{ij}$;
    Конец цикла;
  Конец цикла;
Конец процедуры.

Имя этой процедуры состоит из собственного имени "Скаляр" и входных данных, каковыми являются скалярная величина a и матрица A размера $m \times n$.

Если $A = [A_{ij}]$ и $B = [B_{ij}]$ — матрицы размера $m \times n$, то сумма $A + B$, определенная в главе 4, есть матрица $C = [C_{ij}]$ размера $m \times n$, где $C_{ij} = A_{ij} + B_{ij}$. Алгоритм нахождения суммы двух матриц A и B размера $m \times n$ может быть описан следующим образом:

Процедура **Сложение матриц**(A, B, m, n):
  Цикл по i от 1 до m:
    Цикл по j от 1 до n:
      $C_{ij} = A_{ij} + B_{ij}$;
    Конец цикла;
  Конец цикла;
Конец процедуры.

Напомним, что если

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{np} \end{bmatrix}$$

— матрица размера $n \times p$, и

$$B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1m} \\ B_{21} & B_{22} & \cdots & B_{2m} \\ \vdots & & & \vdots \\ B_{p1} & B_{p2} & \cdots & B_{pm} \end{bmatrix}$$

— матрица размера $p \times m$, то произведение матриц A и B, обозначаемое через AB, есть матрица $C = [C_{ij}]$ размера $n \times m$, где C_{ij} — точечное произведение i-ой строки матрицы A и j-ого столбца матрицы B. То есть,

$$C_{ij} = \begin{bmatrix} A_{i1} & A_{i2} & A_{i3} & \cdots & A_{ip} \end{bmatrix} \bullet \begin{bmatrix} B_{1j} \\ B_{2j} \\ B_{3j} \\ \vdots \\ B_{pj} \end{bmatrix} = \sum_{k=1}^{p} A_{ik}B_{kj}\,.$$

Опишем алгоритм умножения матриц A и B размера $n \times p$ и $p \times m$.

Процедура **Умножение матриц**(A, B, m, p, n):
  Цикл по i от 1 до m:
    Цикл по j от 1 до n:
      $C_{ij} = 0$;
      Цикл по k от 1 до p:
        Значение C_{ij} заменить на $C_{ij} + A_{ik}B_{kj}$;
      Конец цикла;
    Конец цикла;
  Конец цикла;
Конец процедуры.

■ УПРАЖНЕНИЯ

1. Опишите алгоритм нахождения транспонированной матрицы.

2. Вычислите:

а) $\begin{bmatrix} 1 & -1 \\ 3 & 4 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} 3 & 0 & -1 & 2 \\ 3 & 1 & 5 & 0 \end{bmatrix}$;

б) $\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 4 & -1 & 5 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 \\ 5 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix}$;

в) $\begin{bmatrix} 1 \\ -2 \\ 3 \\ -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & -5 & 0 \end{bmatrix}$;

г) $\begin{bmatrix} -4 & 3 & -5 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \\ 3 \end{bmatrix}$;

д) $(-4)\begin{bmatrix} -3 & 1 \\ 2 & -4 \end{bmatrix}$;

е) $\begin{bmatrix} -1 & 3 \\ 3 & 2 \end{bmatrix} - 3\begin{bmatrix} 1 & 4 \\ -2 & 1 \end{bmatrix}$.

3. Пусть $A = \begin{bmatrix} -2 & 4 & 3 \\ 1 & 1 & 7 \\ 2 & -4 & 1 \end{bmatrix}$. Вычислите:

а) A^t; **б)** AA^t; **в)** A^tA.

4. Пусть $A = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$ и $B = \begin{bmatrix} 4 & 1 \\ -2 & -3 \end{bmatrix}$. Вычислите:

а) AB; **б)** BA; **в)** AB^t; **г)** A^tB; **д)** $(AB)^t$;
е) B^tA^t; **ж)** $A - B$; **з)** $B - A$; **и)** $5A$; **к)** $2B - 3A$.

5. Пусть $A = \begin{bmatrix} 1 & 0 & -5 \\ 2 & -4 & 0 \end{bmatrix}$ и $B = \begin{bmatrix} -1 & 2 \\ 3 & 4 \\ 6 & 0 \end{bmatrix}$. Вычислите:

а) AB; **б)** BA; **в)** A^t;
г) B^t; **д)** A^tB^t; **е)** $(BA)^t$.

6. Постройте алгоритм нахождения наибольшего элемента последовательности.
7. Постройте алгоритм нахождения наименьшего элемента последовательности.
8. Постройте алгоритм нахождения наименьшего и наибольшего элементов последовательности.
9. Постройте алгоритм для проверки, является ли матрица A симметричной.
10. Постройте алгоритм для проверки рефлексивности отношения, используя его матричное представление.
11. Постройте алгоритм нахождения среднего арифметического для n чисел.
12. Напишите алгоритм сложения двух целых чисел.
13. Напишите алгоритм вычитания двух целых чисел.
14. Напишите алгоритм умножения двух целых чисел.
15. Напишите алгоритм деления одного целого числа на другое целое число.

5.2. РЕКУРСИВНЫЕ ФУНКЦИИ И АЛГОРИТМЫ

Непосредственное задание функции при помощи формулы, как правило, является предпочтительным. Однако, иногда функцию бывает удобно или даже необходимо задавать при помощи так называемого "рекурсивного метода". Из принципа индукции следует, что если

а) функция задана для некоторого данного начального значения a (обычно 0 или 1), и

б) когда функция задана для некоторого значения k, большего a, она задана также для значения $k+1$,

то, тем самым, функция определена для всех целых чисел, больших a.

Вероятно, правильнее говорить не о рекурсивной функции, а о функции, допускающей рекурсивное определение, поскольку многие функции могут быть заданы как рекурсивно, так и непосредственно.

Заметим, что такая форма рекурсивного определения применима лишь к функциям, заданным на множестве неотрицательных целых чисел. В дальнейшем мы рассмотрим возможность рекурсивного определения для более широкого класса функций. Кроме того, заметим, что возможность рекурсивного определения функции не всегда очевидна. Например, рассмотрим функцию, заданную как

$$\begin{cases} f(0) = 2, \\ f(k+1) = \dfrac{f(k)!}{k!}. \end{cases}$$

Может показаться, что такое определение соответствует приведенным выше пунктам (а) и (б) и задает рекурсивную функцию на множестве неотрицательных це-

лых чисел. Мы видим, однако, что $f(1) = 2$, $f(2) = 1$, $f(3) = \frac{1}{6}$, а значение $f(4)$ не определено, т.к. запись $(\frac{1}{6})!$ не имеет смысла.

Аналогично, если алгоритм может быть выполнен для фиксированного значения a и если из выполнимости алгоритма для значения k, большего a, следует выполнимость алгоритма для значения $k+1$, то такой алгоритм выполним для всех значений $n \geq a$.

Для начала рассмотрим несколько примеров рекурсивных функций.

Рассмотрим функцию $f(n) = a^n$, определенную на множестве неотрицательных целых чисел. Ее можно задать как произведение n чисел, каждое из которых равно a. Функция допускает также и рекурсивное определение, а именно:

$$f(0) = 1;$$
$$f(k+1) = a \cdot f(k).$$

ПРИМЕР 5.4. Рассмотрим функцию $f(n) = n!$, называемую "факториалом". Эту функцию легко выразить посредством соотношения

$$n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n\,.$$

Но компьютеру, выполняющему вычисления, невозможно объяснить, что означает троеточие. Однако мы можем составить процедуру, или алгоритм непосредственного вычисления $n!$. Вот один из возможных вариантов такого алгоритма:

Процедура **Факториал(n)**:
Если $n = 0$, то Факториал$(n) = 1$;
Положить Факториал$(n) = 1$;
Цикл по k от 1 до n:
Значение Факториал(n) заменить на $k \cdot$ (Факториал(n));
Конец цикла;
Конец процедуры.

При переходе к рекурсивному определению функция "факториал" принимает вид

$$\text{Факториал}(0) = 1;$$
$$\text{Факториал}\,(k+1) = (k+1) \cdot \text{Факториал}(k).$$

Соответствующая компьютерная программа может быть описана как

Процедура **Факториал(n)**:
Если $n = 0$, то Факториал$(n) = 1$;
Если $n > 0$, то Факториал$(n) = n \cdot$ Факториал$(n-1)$;
Конец процедуры.

Заметим, что в этой программе процедура Факториал(n) вызывает саму себя, что допустимо в большинстве языков программирования. Когда процедура Факториал(n) вызывается при $n \geq 1$, она выполняется до тех пор, пока не встретится Факториал$(n-1)$. Тогда выполнение процедуры приостанавливается (с сохранением всей информации), и вызывается процедура Факториал$(n-1)$. Если $n-1 \geq 0$, то выполнение процедуры Факториал$(n-1)$ продолжается до тех пор, пока в программе не появится Факториал$(n-2)$. Тогда выполнение программы вновь

приостанавливается (с сохранением всей информации), и вызывается процедура Факториал($n-2$). Процесс продолжается до момента, пока не будет достигнуто и вычислено значение Факториал(0). Это дает возможность продолжить выполнение процедуры Факториал(1) и вычислить соответствующее значение. Это, в свою очередь, дает возможность продолжить процедуру Факториал(2) и вычислить соответствующее значение. Процесс продолжается, пока не будет вычислен Факториал($n-1$), после чего, наконец, вычисляется Факториал(n). □

ПРИМЕР 5.5. Еще одним примером функции, которую обычно задают рекурсивно, является последовательность чисел Фибоначчи. Первый и второй элементы последовательности равны 1, а каждый следующий элемент последовательности равен сумме двух предыдущих. Таким образом, первые десять членов последовательности представляют собой числа $1, 1, 2, 3, 5, 8, 13, 21, 34$ и 55.

Приведенная ниже процедура Фиб(n) описывает программу вычисления значения n-го члена последовательности чисел Фибоначчи:

Процедура **Фиб(n)**:
Если $n = 1$, то Фиб(n) = 1;
Если $n = 2$, то Фиб(n) = 1;
Если $n > 2$, то Фиб(n) = Фиб($n-1$) + Фиб($n-2$);
Конец процедуры.

Процедура крайне неэффективна, хотя имеет простой вид, т.к. на каждом шаге k она вычисляет Фиб($k-1$), начинает снова вычислять Фиб($k-2$), игнорируя тот факт, что значение Фиб($k-2$) уже найдено при вычислении Фиб($k-1$). Читателю будет полезно проследить этот процесс и вычислить Фиб(n) для небольших значений n, таких как 5 или 6. □

ПРИМЕР 5.6. Последовательность ***чисел Каталана*** задается формулой

$$\text{Кат}(n) = \frac{(2n)!}{(n+1)!(n!)}.$$

Если последовательность начинается с 0-го члена, то первые 11 элементов последовательности суть числа $1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862$ и 16796. Функция Кат может быть задана рекурсивно как

$$\text{Кат}(0) = 1;$$
$$\text{Кат}(n+1) = \frac{2(2n+1)}{(n+2)}\,\text{Кат}(n).$$

Чтобы доказать это, вычислим сначала $\text{Кат}(0) = \dfrac{0!}{1!0!} = 1$. Затем покажем, что имеет место равенство

$$\frac{2(2n+1)}{(n+2)}\,\text{Кат}(n) = \text{Кат}(n+1).$$

Упрощая левую часть равенства, получаем

$$\begin{aligned}\frac{2(2n+1)}{(n+2)}\,\text{Кат}(n) &= \frac{2(2n+1)}{(n+2)}\cdot\frac{(2n)!}{(n+1)!(n!)} = \\ &= \frac{(n+1)2(2n+1)(2n)!}{(n+1)(n+2)!(n!)} = \\ &= \frac{(2n+2)(2n+1)(2n)!}{(n+2)!(n+1)(n!)} = \\ &= \frac{(2n+2)!}{(n+2)!(n+1)!} = \\ &= \text{Кат}(n+1).\end{aligned}$$

Рассмотрим теперь две задачи, в которых возникают числа Каталана.

1. Пусть задан выпуклый $(n+2)$-угольник. Число способов разбиения такого многоугольника на треугольники, путем соединения вершин многоугольника с использованием $n-1$ отрезка, равно значению Кат(n). Для треугольника $(n=1)$ существует только один способ образовать треугольник с использованием 0 линий, и Кат$(1)=1$. Для четырехугольника $(n=2)$, как показано на рис. 5.1, существует только два способа разбиения с использованием одной линии.

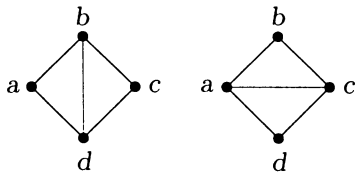


Рис. 5.1

Отметим, что Кат$(2)=2$. Для пятиугольника $(n=3)$, как показано на рис. 5.2, существует пять различных способов разбиения на треугольники с использованием двух линий.

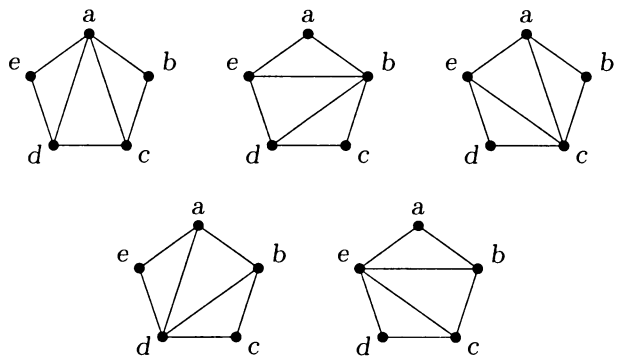


Рис. 5.2

В этом случае мы имеем Кат$(3)=5$.

2. Число способов, которыми можно расставить n пар скобок в выражении

$a_1a_2a_3\cdots a_na_{n+1}$ так, чтобы получить n произведений пар чисел, равно Кат(n). При $n = 1$ имеется один и только один способ, а именно: (a_1a_2). При $n = 2$ имеем $((a_1a_2)a_3)$ и $(a_1(a_2a_3))$, так что существует два способа. Для $n = 3$ имеем $(((a_1a_2)a_3)a_4)$, $(a_1(a_2(a_3a_4)))$, $((a_1a_2)(a_3a_4))$, $((a_1(a_2a_3))a_4)$ и $(a_1((a_2a_3)a_4))$, так что существует пять способов расстановки скобок. □

ПРИМЕР 5.7. Еще одним важным примером является функция Аккермана — Аккер: $N \times N \to N$. Функция определена рекурсивно на $N \times N$, где N — множество целых неотрицательных чисел, следующим образом:

$$\text{Аккер}(0, n) = n + 1;$$
$$\text{Аккер}(m, 0) = \text{Аккер}(m - 1, 1), \text{ если } m > 0;$$
$$\text{Аккер}(m, n) = \text{Аккер}(m - 1,\ \text{Аккер}(m, n - 1)), \text{ если } n, m > 0.$$

Процедуру вычисления функции Аккермана можно описать следующим образом:

Процедура **Аккер(m, n)**:
Если $m = 0$, то Аккер$(m, n) = n + 1$;
Если $m > 0$ и $n = 0$, то Аккер(m, n) = Аккер$(m - 1, 1)$;
Если $m > 0$ и $n > 0$, то Аккер(m, n) = Аккер$(m - 1,$ Аккер$(m, n - 1))$;
Конец процедуры.

Попробуйте вычислить вручную значение Аккер$(3, 2)$. □

ПРИМЕР 5.8. Приведем еще один пример рекурсивного определения функции. Рассмотрим функцию, которая выражает величину ежегодной ренты, когда сумма P инвестируется в течение n периодов с процентной ставкой i за один период, причем подсчет производится в конце каждого периода. Для начала заметим, что если сумма P, которая инвестируется на некоторый период с процентной ставкой i, выражается десятичным числом, то в конце периода объем инвестиции будет равен основной сумме плюс начисленные проценты. Проценты, начисляемые за первый период, составляют iP, так что общий объем инвестиции будет равен $P + iP = (1 + i)P$. Например, если \$1000 инвестировано на год под 5%, то $i = .05$. Тогда в конце года общий объем вклада будет равен \$1000 инвестированных плюс $.05 \times \$1000 = \$50$. Если другие деньги не вкладываются, то в конце второго периода вклад составит $(1 + i)P + i(1 + i)P = (1 + i)^2P$, где $i(1 + i)P$ — проценты от вложенния $(1 + i)P$. К концу второго периода сумма инвестиции составит $(1 + i)^2P$. В нашем примере по прошествии двух лет инвестиция в \$1000 даст сумму \$1050 плюс проценты, которые составят .05(1050). По прошествии n лет начальная инвестиция вырастет до $(1 + i)^nP$.

Теперь предположим, что сумма P инвестируется в конце каждого периода в течение n периодов с процентной ставкой i за каждый период, и подсчет производится в конце каждого периода. На только что вложенную сумму P проценты не начисляются, поэтому она остается равной P. Вклад P, инвестированный в предыдущий период, вместе с начисленными процентами составляет теперь $(1 + i)P$. Вклад, инвестированный в позапрошлый период, вместе с процентами за два периода составляет $(1+i)^2P$. Продолжая подсчет, находим, что общая сумма вклада, накопленная за n периодов, равна

$$P + (1 + i)P + (1 + i)^2P + \cdots + (1 + i)^{n-1}P = P \cdot S(n, i),$$

где $S(n,i) = 1 + (1+i) + (1+i)^2 + \cdots + (1+i)^{n-1}$ и представляет собой геометрическую прогрессию

$$1 + r + r^2 + r^3 + \cdots + r^{n-1} = \frac{r^n - 1}{r - 1}.$$

Следовательно,

$$S(n,i) = \frac{(1+i)^n - 1}{1 + i - 1} = \frac{(1+i)^n - 1}{i}.$$

Очевидно, более эффективно вычислять функцию $S(n,i)$ непосредственно, но она может быть вычислена и рекурсивно. Заметим, что

$$\begin{aligned} S(n+1,i) &= 1 + (1+i) + (1+i)^2 + \cdots + (1+i)^{n-1} + (1+i)^n = \\ &= 1 + (1+i)(1 + (1+i) + (1+i)^2 + \cdots + (1+i)^{n-1}) = \\ &= 1 + (1+i)S(n,i). \end{aligned}$$

Следовательно, рекурсивное определение функции $S(n,i)$ имеет вид:

$$\begin{aligned} &S(1,i) = 1; \\ &S(n,i) = 1 + (1+i)S(n-1,i) \text{ для } n > 1. \end{aligned}$$

□

Теперь рассмотрим рекурсивный алгоритм, который не является определением рекурсивной функции. Согласно старой легенде, в задаче о Ханойской башне фигурируют три стержня и набор дисков разного диаметра. Диски нанизаны на один из стержней в порядке убывания их диаметра, причем диск наибольшего диаметра находится в самом низу, а диск наименьшего диаметра — самый верхний. Задача состоит в том, чтобы переместить все диски на другой стержень с сохранением порядка. При этом диски можно перемещать со стержня на стержень по одному за раз, и диск большего диаметра нельзя помещать над диском меньшего диаметра. Очевидно, что мы имеем дело с рекурсивной задачей, поскольку, если на стержне имеется, например, шесть дисков и известно, как переместить пять из них, то пять верхних дисков можно переместить на средний стержень, затем переместить последний диск на свободный стержень, а затем поместить пять дисков со среднего стержня поверх самого большого диска. Эта ситуация проиллюстрирована на рис. 5.3.

Опишем алгоритм перемещения n дисков с первого стержня на третий. Перенумеруем стержни от 1 до 3 и, кроме того, отождествим их с переменными A, B и C, которые могут принимать значения от 1 до 3.

Процедура **Хан(A, C, n)**:
    Если $n = 1$, переместить единственный диск со стержня A на C;
    Положить $B = 6 - A - C$;
    Если $n > 1$, то
        Сместить верхние $n-1$ дисков на B, используя Хан$(A, B, n-1)$;
        Сместить последний диск с A на C, используя Хан$(A, C, 1)$;
        Сместить верхние $n-1$ дисков на C, используя Хан$(B, C, n-1)$;
Конец процедуры.

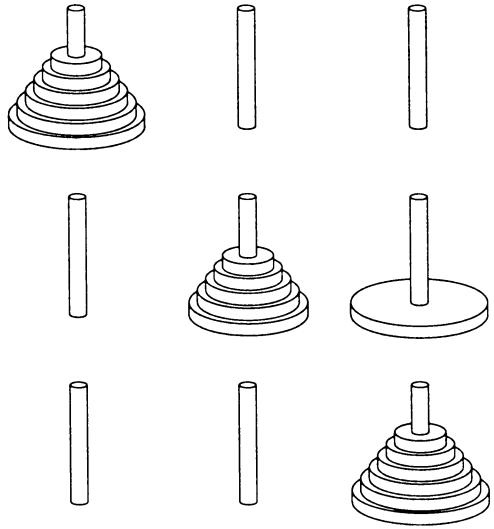


Рис. 5.3

Например, если $n = 3$, начинаем с Хан$(1, 3, 3)$, так что $A = 1$, $B = 2$ и $C = 3$. Находясь внутри процедуры Хан$(1, 3, 3)$, мы вызываем процедуру Хан$(1, 2, 2)$, которая должна переместить два верхних диска на стержень 2. Для Хан$(1, 2, 2)$ имеем: $A = 1$, $B = 3$ и $C = 2$. Из процедуры Хан$(1, 2, 2)$ мы вызываем процедуру Хан$(1, 3, 1)$, которая должна переместить верхний диск со стержня 1 на стержень 3. Далее вызываем Хан$(1, 2, 1)$, чтобы переместить второй диск на стержень 2. Затем вызываем Хан$(1, 3, 1)$, чтобы переместить второй диск со стержня 1 на стержень 3. Затем вызываем Хан$(3, 2, 1)$, перемещая верхний диск со стержня 2 поверх второго диска на стержне 3. Возвращаемся к Хан$(1, 3, 3)$ и вызываем Хан$(1, 3, 1)$, перемещая тем самым самый нижний диск со стержня 3 на стержень 1. Затем вызываем Хан$(2, 3, 2)$ для перемещения двух дисков со стержня 2 на стержень 3. В этот момент $A = 2$, $B = 1$ и $C = 3$. Из процедуры Хан$(2, 3, 2)$ вызываем Хан$(2, 1, 1)$ для перемещения верхнего диска со стержня 2 на стержень 1. Далее вызываем Хан$(2, 3, 1)$ для перемещения второго диска со стержня 2 на стержень 3. Наконец, вызываем Хан$(1, 3, 1)$ для перемещения верхнего диска со стержня 1 поверх всех дисков на стержне 3. Этим мы завершаем выполнение процедур Хан$(2, 3, 2)$ и Хан$(1, 3, 3)$, и игра заканчивается.

Легенда гласит, что на один из трех бриллиантовых стержней нанизаны 64 золотых диска. Некий монах перемещает диски согласно правилу, приведенному выше, со скоростью одно движение за секунду. Согласно легенде, когда монах выполнит эту миссию, наступит конец света. Спрашивается, как долго продлится этот процесс? Построим функцию для вычисления количества движений (и, следовательно, количества секунд), необходимых для переноса n дисков с одного стержня на другой. Обозначим эту функцию H. Известно, что если $n = 1$, то для переноса диска с одного стержня на другой требуется одно движение. Следовательно, $H(1) = 1$. Допустим, что $H(k - 1)$ известно. Тогда $H(k)$, количество движений, необходимых для переноса k дисков со стержня 1 на стержень 3,

складывается из $H(k-1)$ движений, необходимых для переноса $k-1$-го диска со стержня 1 на стержень 2, одного движения по переносу нижнего диска со стержня 1 на стержень 3 и $H(k-1)$ движений, необходимых для переноса $k-1$-го диска со стержня 2 на стержень 3. Таким образом, рекурсивное соотношение $H(k) = 2H(k-1)+1$ задает искомую рекурсивную функцию.

Попробуем задать эту функцию непосредственно. С этой целью вычислим несколько первых значений H и попытаемся найти общее выражение для $H(n)$.

$$H(1) = 1 = 1;$$
$$H(2) = 2 \cdot 1 + 1 = 3;$$
$$H(3) = 2 \cdot 3 + 1 = 7;$$
$$H(4) = 2 \cdot 4 + 1 = 15;$$
$$H(5) = 2 \cdot 15 + 1 = 31.$$

Если к каждому члену последовательности 1, 3, 7, 15, 31 прибавить 1, получим 2, 4, 8, 16, 32. Поэтому можно предположить, что $H(n) = 2^n - 1$. Докажем это по индукции. Очевидно, утверждение справедливо для $n = 1$. Предположим, что оно справедливо для $n = k$, т.е. $H(k) = 2^k - 1$. Нужно показать, что утверждение справедливо для $n = k+1$, т.е., что $H(k+1) = 2^{k+1} - 1$. Но

$$H(k+1) \;=\; 2 \cdot H(k) + 1 \;=\; 2 \cdot (2^k - 1) + 1 \;=\; 2^{k+1} - 1.$$

Следовательно, наше предположение верно. Предлагаем читателю определить, когда, если верить легенде, наступит конец света.

Ранее в этом разделе мы осуществляли переход от рекурсивного описания функции к непосредственному. Для этого мы вычисляли несколько первых значений функции и по ним пытались определить вид выражения, задающего функцию непосредственно. Такая процедура называется ***исключением рекурсии***, или ***решением*** рекурсивной функции. Давайте рассмотрим еще несколько примеров исключения рекурсии, а также примеры перехода от непосредственного задания функции к рекурсивному.

ПРИМЕР 5.9. Исключим рекурсию из следующего рекурсивного определения:

$$f(1) = 1;$$
$$f(k) = f(k-1) + k.$$

Имеем

$$f(1) = 1 = \frac{1 \cdot 2}{2};$$
$$f(2) = 1 + 2 = \frac{2 \cdot 3}{2};$$
$$f(3) = (1 + 2) + 3 = \frac{3 \cdot 4}{2};$$
$$f(4) = (1 + 2 + 3) + 4 = \frac{4 \cdot 5}{2};$$

$$f(5) = (1+2+3+4)+5 = \frac{5\cdot 6}{2}\,.$$

Разумно предположить, что функция имеет вид

$$f(n) = 1+2+3+\cdots+n = \frac{n(n+1)}{2}\,.$$

Чтобы доказать справедливость этой формулы, нужно показать, что данная функция удовлетворяет нашему рекурсивному определению, т.е. показать, что функция $f(n) = \frac{n(n+1)}{2}$ удовлетворяет соотношениям

$$f(1) = 1;$$
$$f(k) = f(k-1) + k\,.$$

Сначала покажем, что $f(1) = 1$:

$$f(1) = \frac{1\cdot 2}{2} = 1.$$

Затем покажем, что $f(k) = f(k-1) + k$:

$$\begin{aligned} f(k-1)+k &= \frac{k(k-1)}{2} + k = \\ &= \frac{k^2-k}{2} + \frac{2k}{2} = \\ &= \frac{k^2+k}{2} = \\ &= \frac{k(k+1)}{2} = \\ &= f(k). \end{aligned}$$

Таким образом, $f(n) = \dfrac{n(n+1)}{2}$ удовлетворяет рекурсивному определению. □

ПРИМЕР 5.10. Исключите рекурсию для функции, заданной соотношением

$$f(1) = 2;$$
$$f(k) = 2\cdot k\cdot f(k-1).$$

Имеем следующие равенства:

$$f(1) = 2 = 2^1\cdot 1!;$$
$$f(2) = 2\cdot 2\cdot 2 = 2^2\cdot 2!;$$
$$f(3) = 2\cdot 2\cdot 2\cdot 2\cdot 3 = 2^3\cdot 3!;$$
$$f(4) = 2\cdot 2\cdot 2\cdot 2\cdot 3\cdot 2\cdot 4 = 2^4\cdot 4!;$$
$$f(5) = 2\cdot 2\cdot 2\cdot 2\cdot 3\cdot 2\cdot 4\cdot 2\cdot 5 = 2^5\cdot 5!\,.$$

Отсюда можно предположить, что искомая функция задается выражением $f(n) = 2^n \cdot n!$. Чтобы это доказать, нам нужно показать, что данная функция удовлетворяет приведенному выше рекурсивному определению. Иными словами, нужно показать, что функция $f(n) = 2^n \cdot n!$ удовлетворяет соотношениям

$$f(1) = 2;$$
$$f(k) = 2kf(k-1).$$

Проверим сначала, что $f(1) = 2$:

$$f(1) = 2^1 \cdot 1! = 2.$$

Далее покажем, что $f(k) = 2kf(k-1)$. Но

$$\begin{aligned} 2kf(k-1) &= 2k(2^{k-1} \cdot (k-1)!) = \\ &= 2(2^{k-1}) \cdot k(k-1)! = \\ &= 2^k \cdot k!. \end{aligned}$$

Итак, функция $f(n) = 2^n \cdot n!$ удовлетворяет рекурсивному определению. □

ПРИМЕР 5.11. Для функции $f(n) = \frac{n(n+1)(2n+1)}{6}$ найдем задающие ее рекурсивные соотношения. Имеем

$$\begin{aligned} f(1) &= 1 = 1^2; \\ f(2) &= 5 = 1^2 + 2^2; \\ f(3) &= 14 = 1^2 + 2^2 + 3^2; \\ f(4) &= 30 = 1^2 + 2^2 + 3^2 + 4^2; \\ f(5) &= 55 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2. \end{aligned}$$

Поэтому можно предположить, что

$$f(1) = 1;$$
$$f(k) = f(k-1) + k^2.$$

Для доказательства применим индукцию. Очевидно, что $f(1) = 1$. Допустим, что $f(k-1) = \frac{(k-1)(k)(2k-1)}{6}$, и покажем, что тогда $f(k) = \frac{k(k+1)(2k+1)}{6}$:

$$\begin{aligned} f(k-1) + k^2 &= \frac{(k-1)(k)(2k-1)}{6} + k^2 = \\ &= \frac{(k-1)(k)(2k-1) + 6k^2}{6} = \\ &= k\frac{(k-1)(2k-1) + 6k}{6} = \\ &= k\frac{2k^2 - 4k + 1 + 6k}{6} = \\ &= k\frac{(k+1)(2k+1)}{6} = \\ &= f(k). \end{aligned}$$

□

■ УПРАЖНЕНИЯ

1. Найдите $f(1)$, $f(2)$, $f(3)$ и $f(4)$ для приведенных ниже рекурсивных функций:

а) $\begin{cases} f(0) = 3, \\ f(k) = 3f(k-1). \end{cases}$

б) $\begin{cases} f(0) = 1, \\ f(k) = 2f(k-1) + 3k. \end{cases}$

в) $\begin{cases} f(0) = 2, \\ f(k) = (f(k-1))^2, \end{cases}$

г) $\begin{cases} f(0) = 1, \\ f(k) = k^2 f(k-1). \end{cases}$

д) $\begin{cases} f(0) = 1, \\ f(k) = 2^{f(k-1)}. \end{cases}$

2. Найдите $f(1)$, $f(2)$, $f(3)$ и $f(4)$ для следующих рекурсивных функций:

а) $\begin{cases} f(0) = 0, \\ f(k) = k + f(k-1). \end{cases}$

б) $\begin{cases} f(0) = 1, \\ f(k) = \left\lfloor \dfrac{f(k-1)}{2} \right\rfloor + 3k. \end{cases}$

в) $\begin{cases} f(0) = 2, \\ f(k) = 8k^2 - (f(k-1))^2. \end{cases}$

г) $\begin{cases} f(0) = 4, \\ f(k) = \dfrac{f(k-1)}{k^2}. \end{cases}$

д) $\begin{cases} f(0) = 2, \\ f(k) = \dfrac{f(k-1)}{k!}. \end{cases}$

3. Найдите $f(2)$, $f(3)$, $f(4)$ и $f(5)$ для следующих рекурсивных функций:

а) $\begin{cases} f(0) = 1, \\ f(1) = 3, \\ f(k) = 2f(k-1) - f(k-2). \end{cases}$

б) $\begin{cases} f(0) = 0, \\ f(1) = 1, \\ f(k) = (f(k-1))^2 - (f(k-2))^2. \end{cases}$

в) $\begin{cases} f(0) = 1, \\ f(1) = 2, \\ f(k) = (f(k-1))^2 - f(k-2) + k^2. \end{cases}$

г) $\begin{cases} f(0) = 1, \\ f(1) = 2, \\ f(k) = (f(k-1) - f(k-2))k!. \end{cases}$

д) $\begin{cases} f(0) = -1, \\ f(1) = 1, \\ f(k) = f(k-1) \div f(k-2). \end{cases}$

4. Найдите $f(2)$, $f(3)$, $f(4)$ и $f(5)$ для следующих рекурсивных функций:

а) $\begin{cases} f(0) = 2, \\ f(1) = 4, \\ f(k) = 3f(k-1) - 2f(k-2). \end{cases}$

б) $\begin{cases} f(0) = 1, \\ f(1) = 2, \\ f(k) = (f(k-1))! - (f(k-2))!. \end{cases}$

в) $\begin{cases} f(0) = 0, \\ f(1) = 2, \\ f(k) = (f(k-1))! \div (f(k-2))!. \end{cases}$

г) $\begin{cases} f(0) = 10, \\ f(1) = 20, \\ f(k) = \left\lfloor \dfrac{f(k-1) + f(k-2)}{k!} \right\rfloor. \end{cases}$

д) $\begin{cases} f(0) = -1, \\ f(1) = 1, \\ f(k) = f(k-1) \div (f(k-2))^2. \end{cases}$

5. Найдите явные выражения для $f(n)$, исключив рекурсию из следующих определений:

а) $\begin{cases} f(0) = 1, \\ f(k) = 2f(k-1). \end{cases}$

б) $\begin{cases} f(0) = 1, \\ f(k) = 1 + f(k-1). \end{cases}$

в) $\begin{cases} f(0) = 1, \\ f(k) = f(k-1) \div k. \end{cases}$

г) $\begin{cases} f(0) = 1, \\ f(k) = f(k-1) + 2. \end{cases}$

д) $\begin{cases} f(0) = 1, \\ f(k) = 5f(k-1). \end{cases}$

6. Найдите явные выражения для $f(n)$, исключив рекурсию из следующих определений:

а) $\begin{cases} f(0) = 2, \\ f(k) = \dfrac{f(k-1)!}{k!}. \end{cases}$

б) $\begin{cases} f(0) = 1, \\ f(k) = 1 + f(k-1). \end{cases}$

в) $\begin{cases} f(0) = -1, \\ f(k) = \dfrac{-1}{f(k-1)}. \end{cases}$

г) $\begin{cases} f(0) = 1, \\ f(k) = -3f(k-1). \end{cases}$

д) $\begin{cases} f(0) = 1, \\ f(k) = 5 + 2f(k-1). \end{cases}$

е) $\begin{cases} f(0) = -1, \\ f(1) = 1, \\ f(k) = \dfrac{f(k-1)}{f(k-2)}. \end{cases}$

7. Докажите, что $a_n = 7 - 2^{n+1}$ удовлетворяет рекурсивному определению

$$\begin{cases} a_0 = 5, \\ a_k = 2a_{k-1} - 7 \text{ при } k > 0. \end{cases}$$

8. Докажите, что $a_n = 3^n - n3^{n+1}$ удовлетворяет рекурсивному определению

$$\begin{cases} a_0 = 1, \\ a_1 = -6, \\ a_k = 6a_{k-1} - 9a_{k-2} \text{ при } k > 1. \end{cases}$$

9. Докажите, что $a_n = -1 - 2^{n+1}$ удовлетворяет рекурсивному определению

$$\begin{cases} a_0 = -3, \\ a_1 = -5, \\ a_k = 6a_{k-1} - 8a_{k-2} - 3 \text{ при } k > 1. \end{cases}$$

10. Докажите, что $a_n = \dfrac{1 - r^{n+1}}{1 - r}$ удовлетворяет рекурсивному определению

$$\begin{cases} a_0 = 1, \\ a_k = a_{k-1} + r^k \text{ при } k > 0. \end{cases}$$

11. Докажите, что $a_n = (n^2 + n + 1)$ удовлетворяет рекурсивному определению

$$\begin{cases} a_1 = 3, \\ a_k = a_{k-1} + 2k \text{ при } k > 1. \end{cases}$$

12. Докажите, что $a_n = n^2(n+1)^2$ удовлетворяет рекурсивному определению

$$\begin{cases} a_1 = 4, \\ a_k = a_{k-1} + 4k^3 \text{ при } k > 1. \end{cases}$$

13. Докажите, что $a_n = -2(-1)^n + 2 \cdot 3^n$ удовлетворяет рекурсивному определению

$$\begin{cases} a_0 = 0, \\ a_1 = 8, \\ a_k = 2a_{k-1} + 3a_{k-2} \text{ при } k > 2. \end{cases}$$

14. Докажите, что $a_n = 2(-3)^n + 5 \cdot 2^n$ удовлетворяет рекурсивному определению

$$\begin{cases} a_0 = 7, \\ a_1 = 4, \\ a_k = -a_{k-1} + 6a_{k-2} \text{ при } k > 1. \end{cases}$$

15. Докажите, что $a_n = 3(-2)^n + 2 \cdot 3^n - 3 \cdot (2)^n$ удовлетворяет рекурсивному определению

$$\begin{cases} a_0 = 2, \\ a_1 = -6, \\ a_k = a_{k-1} + 6a_{k-2} + 3 \cdot (2)^k \text{ при } k > 1. \end{cases}$$

16. Докажите, что

$$a_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right)$$

удовлетворяет рекурсивному определению

$$\begin{cases} a_1 = 1, \\ a_2 = 1, \\ a_k = a_{k-1} + a_{k-2} \text{ при } k > 1. \end{cases}$$

17. Как скоро, согласно легенде о Ханойской башне, наступит конец света после того, как монах начнет перекладывать диски?

5.3. СЛОЖНОСТЬ АЛГОРИТМОВ

В примере с Ханойской башней нам требовалось $2^n - 1$ отдельных перемещений дисков, чтобы переложить n дисков с одного стержня на другой. Зачастую важно иметь возможность оценивать число элементарных операций или время, необходимое компьютеру для выполнения всего алгоритма. Это не всегда одно и то же, поскольку выполнение одних операций длится дольше, чем выполнение других. Необходимо учитывать и другие факторы, такие как требуемый для выполнения алгоритма объем памяти компьютера, точность вычислений и многое другое. Очевидно, что на самом деле это важно только для программ, требующих для выполнения значительного времени. Время выполнения зачастую зависит от такого фактора, как объем входных данных. В вышеупомянутом примере таковым является n — количество дисков. Время или число операций, необходимых для выполнения процедуры, увеличивается по мере увеличения n.

ОПРЕДЕЛЕНИЕ 5.12. Пусть область определения функций f и g — множество целых положительных чисел, а множество их значений — множество действительных чисел. Говорят, что функция g ***мажорирует*** функцию f, если существует действительное число k и целое положительное число m такое, что $|f(n)| \leq k\,|g(n)|$ для всех $n \geq m$. Если g мажорирует f, это обозначается как $f(n) = O(g(n))$. Символ $O(g(n))$ читается как ***"О большое"*** от $g(n)$; при этом говорят, что $f(n)$ имеет порядок О большое от $g(n)$.

Для положительных целых чисел r и s следующую теорему можно доказать методом индукции. Справедливость утверждения теоремы для положительных рациональных чисел r и s можно показать, не используя логарифмы (см. упражнения).

ТЕОРЕМА 5.13. Если r и s — действительные числа, $r \leq s$ и $n > 1$, тогда $n^r \leq n^s$. Следовательно, $n^r = O(n^s)$.

ДОКАЗАТЕЛЬСТВО. Функция $ln(x)$ — возрастающая, поэтому $a \leq b$ тогда и только тогда, когда $\ln(a) \leq \ln(b)$. Отсюда $n^r \leq n^s$ тогда и только тогда, когда $\ln(n^r) \leq \ln(n^s)$, что, в свою очередь, выполняется тогда и только тогда, когда $r\ln(n) \leq s\ln(n)$, т.е. тогда и только тогда, когда $r \leq s$, поскольку $\ln(n)$ для $n > 1$ — величина положительная. ■

Следующая теорема показывает, что свойство функции иметь порядок $O(g(n))$ замкнуто относительно операций сложения и умножения на скаляр. Это значит, что рассматриваемые функции образуют векторное пространство, хотя изучение подобных вопросов выходит за рамки данной книги.

ТЕОРЕМА 5.14. Если $f(n) = O(g(n))$, то $cf(n) = O(g(n))$.

ДОКАЗАТЕЛЬСТВО. По определению, $|f(n)| \leq k\,|g(n)|$ для некоторого постоянного действительного числа k и всех n, больших или равных некоторому целому числу m. Поэтому

$$|cf(n)| \leq k\,|c|\,|g(n)|$$

и $cf(n) = O(g(n))$. ■

ТЕОРЕМА 5.15. Если $f(n) = O(g(n))$ и $h(n) = O(g(n))$, то $(f + h)(n) = O(g(n))$.

ДОКАЗАТЕЛЬСТВО. По определению, для некоторого постоянного k и некоторого целого числа m_1 имеем $|f(n)| \leq k\,|g(n)|$ для всех $n > m_1$. Опять же по определению, для некоторой постоянной l и некоторого целого числа m_2 имеем $|h(n)| \leq l\,|g(n)|$ для всех $n > m_2$. Пусть $m = \max(m_1 m_2)$. Следовательно, для всех $n > m$

$$\begin{aligned} |f(n) + h(n)| &\leq |f(n)| + |h(n)| \leq \\ &\leq k\,|g(n)| + l\,|g(n)| = \\ &= (k + l)\,|g(n)| \end{aligned}$$

и $(f + h)(n) = O(g(n))$. ■

ТЕОРЕМА 5.16. Если $f(n) = O(g(n))$ и $h(n) = O(e(n))$, то $(f \cdot h)(n) = O((g \cdot e)(n))$.

ДОКАЗАТЕЛЬСТВО. По определению, для некоторого постоянного k и некоторого целого числа m_1 имеем $|f(n)| \leq k\,|g(n)|$ для всех $n > m_1$. Также по определению, для некоторого постоянного l и некоторого целого числа m_2 имеем $|h(n)| \leq l\,|e(n)|$ для всех $n > m_2$. Пусть $m = \max(m_1 m_2)$. Поэтому для всех $n > m$

$$\begin{aligned} |f(n) \cdot h(n)| &= |f(n)| \cdot |h(n)| \leq \\ &\leq k\,|g(n)|\,l\,|e(n)| = \\ &= kl\,|g(n) \cdot e(n)| \end{aligned}$$

и $(f \cdot h)(n) = O((g \cdot e)(n))$. ■

Следующая теорема утверждает, что степень полинома единственным образом определяет его мажоранту.

ТЕОРЕМА 5.17. Если $p(n) = a_k n^k + a_{k-1} n^{k-1} + \cdots + a_2 n^2 + a_1 n + a_0$, то $p(n) = O(n^k)$.

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned} |p(n)| &\leq \left|a_k n^k + a_{k-1} n^{k-1} + \cdots + a_2 n^2 + a_1 n + a_0\right| \leq \\ &\leq \left|a_k n^k\right| + \left|a_{k-1} n^{k-1}\right| + \cdots + \left|a_2 n^2\right| + |a_1 n| + |a_0| \leq \\ &\qquad \text{(в силу неравенства треугольника: } |A + B| \leq |A| + |B|\text{)} \\ &\leq |a_k|\,n^k + |a_{k-1}|\,n^{k-1} + \cdots + |a_2|\,n^2 + |a_1|\,n + |a_0| \leq \\ &\leq |a_k|\,n^k + |a_{k-1}|\,n^k + \cdots + |a_2|\,n^k + |a_1|\,n^k + |a_0|\,n^k = \\ &\qquad \text{по теореме 5.13} \\ &= (|a_k| + |a_{k-1}| + \cdots + |a_2| + |a_1| + |a_0|)n^k \end{aligned}$$

и $p(n) = O(n^k)$. ■

ТЕОРЕМА 5.18. Для целых чисел a и b, больших единицы, $\log_a(n) = O(\log_b(n))$.

ДОКАЗАТЕЛЬСТВО. Следует непосредственно из равенства $\log_a(n) = \dfrac{\log_b(n)}{\log_a(b)}$. ■

ТЕОРЕМА 5.19. Пусть n — неотрицательное целое число. Тогда $n < 2^n$ и, следовательно, $n = O(2^n)$.

ДОКАЗАТЕЛЬСТВО. Воспользуемся индукцией, имея для $n = 0$, $0 < 2^0 = 1$. Допустим, что $k < 2^k$, тогда

$$k + 1 \leq k + k \leq 2^k + 2^k = 2 \cdot 2^k = 2^{k+1}$$

и, по индукции, $n < 2^n$. ■

Пусть R — отношение на множестве функций со значениями во множестве целых положительных чисел, определенное как fRg, если $f(n) = O(g(n))$. Очевидно, отношение R рефлексивно. Приведенная ниже теорема утверждает, что это отношение также и транзитивно. Доказательство теоремы предоставляется читателю.

ТЕОРЕМА 5.20. Если $f(n) = O(g(n))$ и $g(n) = O(h(n))$, то $f(n) = O(h(n))$.

Следующая теорема дает ответ на вопрос о том, какие функции могут выступать в роли мажорант для других функций.

ТЕОРЕМА 5.21. Для целых чисел a, больших единицы, $\log_a(n) = O(n)$.

ДОКАЗАТЕЛЬСТВО. Согласно теореме 5.19, имеет место неравенство $n < 2^n$. Поэтому $\log_2(n) < \log_2(2^n) = n$ и $\log_2(n) = O(n)$. Поскольку $\log_a(n) = O(\log_2(n))$, то по теореме 5.18 имеем $\log_a(n) = O(n)$. ■

Доказательство сформулированных ниже теорем предоставляем читателю.

ТЕОРЕМА 5.22. Пусть n — неотрицательное целое число, тогда $n! < n^n$ и, следовательно, $n! = O(n^n)$.

ТЕОРЕМА 5.23. Пусть $a > 1$ и n — неотрицательное целое число, тогда $\log_a(n!) \leq n \log_a(n)$, следовательно, $\log_a(n!) = O(n \log_a(n))$.

Ранее, при решении задачи о Ханойской башне, определялось число шагов, необходимых для перемещения дисков. В примере, приведенном ниже, определяется число арифметических операций, необходимых, соответственно, для сложения двух матриц и умножения двух матриц.

ПРИМЕР 5.24. Рассмотрим алгоритм сложения матриц:

```
Процедура Сложение матриц(A, B, m, n):
    Цикл по i от 1 до m:
        Цикл по j от 1 до n:
            C_ij = A_ij + B_ij;
        Конец цикла;
    Конец цикла;
Конец процедуры.
```

Как видим, сложение выполняется для каждого i и каждого j. Поскольку i принимает m значений, а j принимает n значений, то выполняется mn операций сложения. Пусть $N = \max(m, n)$. Тогда число выполняемых арифметических операций имеет порядок $O(N^2)$. □

ПРИМЕР 5.25. Алгоритм умножения матрицы A размера $m \times p$ и матрицы B размера $p \times n$ имеет вид:

```
Процедура Умножение матриц( A, B, m, p, n):
    Цикл по i от 1 до m:
        Цикл по j от 1 до n:
            C_ij = 0;
            Цикл по k от 1 до p:
                Значение C_ij заменить на C_ij + A_ik B_kj ;
            Конец цикла;
        Конец цикла;
    Конец цикла;
Конец процедуры.
```

Поскольку k принимает значения от 1 до p, то выполняется p операций сложения и p операций умножения. Величина k изменяется от 1 до p для каждого i и каждого j, поэтому k пробегает значения от 1 до p mn раз. Таким образом, выполняется mnp операций сложения и mnp операций умножения. Следовательно, всего выполняется $2mnp$ операций. Пусть $N = \max(m, n, p)$. Тогда число выполняемых арифметических операций имеет порядок $O(N^3)$. □

ПРИМЕР 5.26. Сравним количество операций, которое требуется для непосредственного вычисления значения $p(c)$, где $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0$, с количеством операций, которое потребуется для вычисления того же значения при помощи схемы Горнера.

Если $p(c)$ вычисляется непосредственно, то для подсчета c^k требуется выполнить $k-1$ операций умножения. Еще одна операция нужна для умножения на a_k, так что вычисление $a_k c^k$ требует k операций умножения. Таким образом, нужно выполнить $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$ умножений. Для того чтобы найти сумму $n+1$ слагаемых, требуется выполнить n сложений, так что общее число арифметических операций равно $\frac{n(n+1)}{2} + n$ и имеет порядок $O(n^2)$.

Если мы вычисляем $p(a)$, где $p(x)$ — полином, используя схему Горнера, то прежде всего замечаем, что выражение $a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = x(x(x(a_4x + a_3) + a_2) + a_1) + a_0$ включает четыре операции сложения и четыре операции умножения. Далее мы определяем, что в общем случае $p(x) = x(x(\cdots x((a_n x + a_{n-1}) + a_{n-2}) + \cdots + a_2) + a_1) + a_0$ включает n операций сложения и n операций умножения. Таким образом, общее число арифметических операций равно $2n$ и имеет порядок $O(n)$. □

■ УПРАЖНЕНИЯ

1. Сколько нужно выполнить перемещений в задаче о Ханойской башне, если число дисков равно восьми?
2. Как говорилось ранее, существует древняя легенда (начала девятнадцатого века) о том, что на один из трех бриллиантовых стержней нанизано 64 золотых диска, которые некий монах перемещает со стержня на стержень со скоростью одно движение за секунду. Согласно легенде, по окончании этой миссии наступит конец света. Как скоро это произойдет?

3. Оцените количество арифметических операций, необходимых для вычисления указанных ниже значений полиномов (i) при непосредственном вычислении, (ii) при помощи схемы Горнера.

а) $f(3)$, где $f(x) = 3x^2 + 4x + 5$;

б) $f(2)$, где $f(x) = 2x^4 + 4x^3 + 3x^2 + 2x + 3$;

в) $f(4)$, где $f(x) = x^5 + 4x^4 + 2x^3 + 6x^2 + x + 3$.

4. Оцените количество арифметических операций, необходимых для вычисления следующих произведений:

а) $\begin{bmatrix} 3 & -1 \\ 2 & 7 \\ -5 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 4 \\ 5 & 1 & 3 & 0 \end{bmatrix}$; **б)** $\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 6 & -1 & 5 \end{bmatrix} \begin{bmatrix} 2 & -2 & 8 \\ 5 & 5 & 1 \\ 0 & 6 & 4 \end{bmatrix}$;

в) $\begin{bmatrix} 1 \\ 3 \\ 4 \\ 5 \end{bmatrix} \begin{bmatrix} -2 & 3 & -5 & 0 \end{bmatrix}$.

5. Какие из приведенных ниже функций имеют порядок $O(n^3)$?

а) $n^4 - 3n + 5$; **б)** $n^2 - 6n + 5$; **в)** $(\ln(n))^3$;

г) $n(\ln(n))^2$; **д)** $n^2 \ln(n)$.

6. Какие из приведенных ниже функций имеют порядок $O(n^3)$?

а) $6n^3 + 3n^2 + 2n + 5$; **б)** $n^3 \ln(n)$; **в)** $\ln(n^{n \ln(n)})$;

г) $\log_a(n!)$; **д)** $n!$.

7. Пусть n — некоторое неотрицательное целое число. Докажите, что $n! \leq n^n$ и, следовательно, $n! = O(n^n)$.

8. Докажите, что для $a > 1$ и $n \in N$ справедлива оценка $\log_a(n!) \leq n \log_a(n)$, поэтому $\log_a(n!) = O(n \log_a(n))$.

9. Докажите, что если $f(n) = O(g(n))$ и $g(n) = O(h(n))$, то $f(n) = O(h(n))$.

10. Докажите, что для целого числа a, большего единицы, $\log_a(n) = O(n)$.

11. Используя индукцию, докажите, что если r и s — положительные целые числа, $r \leq s$ и $n > 1$, то $n^r \leq n^s$. Отсюда $n^r = O(n^s)$.

12. Используя результаты предыдущих задач, докажите, что для положительных рациональных чисел r и s, $r \leq s$ и $n > 1$, из неравенства $n^r \leq n^s$ следует, что $n^r = O(n^s)$.

5.4. АЛГОРИТМЫ СОРТИРОВКИ

В этом разделе мы рассмотрим некоторые методы сортировки данных. В дальнейшем мы будем считать выполненными два условия. Во-первых, будем считать, что данные пронумерованы так, что имеется первый элемент данных, второй элемент, и так далее до n-го. Такая организация данных может быть реализована при помощи массива, указателей или еще каким-либо образом. Конкретный метод реализации списка нас не интересует. Второе предположение состоит в том, что на множестве данных введено полное упорядочение, и мы хотим расположить

данные в соответствии с этим порядком. Это может быть лексикографическое упорядочение (т.е. упорядочение по алфавиту) или же упорядочение в соответствии с числовыми значениями, которые приписываются элементам данных. Опять же, нас не интересует, что это за упорядочение, лишь бы оно было задано. В дальнейшем запись $a < b$ будет означать, что в нашем упорядочении a предшествует b.

Для начала рассмотрим ***сортировку выбором*** (СВ). Пусть $a_1, a_2, \ldots, a_n$ — элементы, подлежащие сортировке. Сам процесс сортировки очень прост. Мы фиксируем элемент a_1, находим наименьший элемент и меняем местами этот элемент и a_1. Затем переходим к a_2. Рассматриваем все элементы, начиная с a_2, находим среди них наименьший и меняем местами этот элемент и a_2. Процесс продолжаем до завершения. Более строго: пусть мы находимся на этапе, когда необходимо заменить a_i. Введем переменную min, для того чтобы отслеживать индекс наименьшего элемента среди a_i и всех тех элементов, которые за ним следуют. Сначала полагаем $min = i$ и начинаем сравнивать оставшиеся элементы с a_{min}. Если находим элемент, который меньше a_{min}, скажем a_k, полагаем $min = k$ и продолжаем поиск наименьшего элемента, сравнивая его с a_{min}. Если таковой находится, процесс повторяем. Дойдя до n-го элемента, меняем местами a_i и a_{min} и переходим к a_{i+1}. Теперь мы можем записать алгоритм.

```
Процедура СВ:
    Цикл по i от 1 до n − 1:
        Положить min = i;
        Цикл по j от 1 до n:
            Если a_j < a_min, то пусть min = j;
        Конец цикла;
        Поменять местами a_i и a_min;
    Конец цикла;
Конец процедуры.
```

ПРИМЕР 5.27. Пусть c, b, a, v, h, d — список для сортировки. Начинаем с c на позиции 1 и полагаем $min = 1$. Сравниваем c с b и, поскольку $b < c$ и b находится на второй позиции, полагаем $min = 2$. Далее сравниваем b с a и, поскольку $a < b$, полагаем $min = 3$. Букв, меньших a, не находим, поэтому меняем местами a и b, получая в результате a, b, c, v, h, d. Повторяем процесс, начиная с b, поэтому $min = 2$. Сравнение b с оставшимися буквами показывает, что ни одна из них не меньше b, так что min остается равным 2, и мы меняем местами b с самим собой. Далее переходим к c и полагаем $min = 3$. Опять ни один элемент не меньше c, поэтому c остается на том же месте. Переходим к v и полагаем $min = 4$. Поскольку $h < v$, то $min = 5$. Затем сравниваем h и d. Поскольку $d < v$, полагаем $min = 6$. Меняем местами d и v, получая последовательность a, b, c, d, h, v. Продолжаем процесс для h, а так как $h < v$, то сортировка закончена. □

Рассмотрим число сравнений, которое производится согласно этому методу, для списка из n элементов. Первый элемент сравнивается с остальными $n - 1$ элементами. Второй элемент сравнивается с $n - 2$ элементами и так далее до $(n - 1)$-го элемента, который сравнивается только с одним элементом. Таким

образом, имеем

$$(n-1)+(n-2)+\cdots+2+1=\frac{n(n-1)}{2},$$

так что число сравнений имеет порядок $O(n^2)$.

Теперь рассмотрим весьма популярный алгоритм сортировки, очень напоминающий сортировку выбором. Он называется ***пузырьковой сортировкой*** (ПС), так как в нем меньшие числа поднимаются вверх, как пузырьки в воде. (В данном случае "вверх" означает налево.) Процесс начинают с правого конца списка и сравнивают два последних элемента, сдвигая меньший влево. Затем сравнивают этот меньший с элементом, стоящим слева от него, и опять передвигают меньший из них влево. Процесс продолжают, пока не будет достигнут первый элемент. В результате наименьший элемент становится первым элементом. Затем процесс повторяется до тех пор, пока он не вернется ко второму по величине элементу. Теперь второй по величине элемент располагается на втором месте. Процесс повторяется до тех пор, пока все элементы не будут рассортированы.

На данном этапе новым является то, что при такой сортировке движение происходит с двух сторон. Процесс начинается справа, движется влево вплоть до первого элемента. Затем опять начинается справа, движется влево вплоть до второго элемента. Всякий раз движение влево осуществляется вплоть до j-го элемента данных, при этом значение j возрастает. Для описания такого рода перемещений нам понадобится новая разновидность цикла перечислением. Говоря "цикл по i от n до k с шагом -1", где $n>k$, мы будем иметь в виду, что, начиная с $i=n$, мы будем шаг за шагом уменьшать значение i на 1, вплоть до $i=k$.

```
Процедура ПС:
    Цикл по i от 2 до n:
        Цикл по j от n до i с шагом −1:
            Если a_j < a_{j−1}, то поменять местами a_j и a_{j−1};
        Конец цикла;
    Конец цикла;
Конец процедуры.
```

ПРИМЕР 5.28. Пусть, как и ранее, c,b,a,v,h,d — список для сортировки. Сначала меняем местами h и d, поскольку d меньше, чем h. Затем меняем местами d и v, т.к. d меньше, чем v. На данный момент список выглядит таким образом: c,b,a,d,v,h. Поскольку d не меньше, чем a, мы не меняем местами a и d. Теперь меняем местами b и a, поскольку a меньше, чем b. Наконец, меняем местами a и c, т.к. a меньше, чем c. Состояние списка на данном этапе — a,c,b,d,v,h.

Повторяя процесс, меняем местами h и v, поскольку h меньше v. Оставляем d на своем месте, т.к. h не меньше d. Теперь список имеет вид a,c,b,d,h,v. Также на своих местах остаются d и b, поскольку d не меньше b. Элемент b меньше c, поэтому b и c меняются местами, и мы получаем a,b,c,d,h,v. Несмотря на то, что процесс повторялся трижды, лишние перестановки не потребовались, сортировка закончилась. □

Если проанализировать процесс сравнения в этом просмотре n элементов, то в первом цикле он включает $n-1$ операцию, во втором — $n-2$ операции.

Процесс завершается выполнением одного сравнения в последнем цикле. Отсюда получаем

$$(n-1)+(n-2)+\cdots+2+1=\frac{n(n-1)}{2},$$

так что общее число сравнений имеет порядок $O(n^2)$. Таким образом, полученное значение совпадает с соответствующей величиной для сортировки выбором. Несмотря на то, что число операций сравнения этих двух методов имеет один и тот же порядок, их скорости в реальных случаях могут различаться.

Теперь рассмотрим алгоритм сортировки, который называется ***быстрой сортировкой*** (БС). Это первый метод сортировки, который будет определен рекурсивно. Процесс быстрой сортировки состоит с следующем: мы берем a_1, первый элемент списка данных, затем каждый элемент (исключая a_1) помещаем слева от a_1, если он меньше или равен a_1. Назовем этот список S_1. Затем поместим справа от a_1 каждый элемент, который больше a_1. Назовем этот список S_2. Теперь применим этот же процесс к S_1 и S_2. Процесс продолжаем до тех пор, пока в каждом списке не останется по одному элементу. Затем собираем новый список, каждый раз производя конкатенацию S_i, a_i, S_{i+1} в список $S_i a_i S_{i+1}$, где a_i — элемент, который послужил основой для создания S_i и S_{i+1}. Более формально алгоритм описывается следующим образом:

Процедура **БС(1, n)**:
Положить $l=1$ и $r=n$;
Если $l \neq r$, то
  Положить $m=l$ и разбить список $(a_l, a_{l+1}, a_{l+2}, \ldots, a_r)$ на S_1,
    полученный как $\{a_i : a_i \leq a_l,\ i \neq l\}$,
    и S_2, полученный как $\{a_i : a_i > a_l,\ \}$, так что
    $S_1=(a_l, a_{l+1}, a_{l+2}, \ldots, a_j)$ и $S_2=(a_{j+1}, a_{j+2}, a_{j+3}, \ldots, a_r)$
    после переобозначения;
  Произвести конкатенацию БС$(1, j)$, a_m и БС$(j+1, r)$;
Конец процедуры.

ПРИМЕР 5.29. Пусть, как и ранее, d, b, c, r, v, h, a — список для сортировки. Поскольку d — первый элемент списка, разбиваем список на два подсписка: b, c, a — элементов, меньших d, и r, v, h — список элементов, больших d. Далее, согласно процедуре БС первым рассматриваем левый из вновь созданных подсписков, там мы используем b для разделения c и a на списки, где a будет слева, а c — справа. После конкатенации получаем a, b, c и возвращаемся к первому шагу, где процедура БС вызывается для списка справа от d, а именно, r, v, h. Буква r используется для разделения v и h в списки h и v. После конкатенации имеем h, r, v. И, наконец, последняя конкатенация дает список a, b, c, d, h, r, v. □

Теперь оценим количество операций сравнения, используемых в этом процессе. Быстрая сортировка реализуется наиболее оптимально, когда $n=2^m$ для некоторого m, и каждое разбиение делит список элементов ровно пополам. Предположим, что это так. Если Q_n — число сравнений в быстрой сортировке для списка из n элементов, тогда имеется n операций сравнения для определения, в какой из списков попадет тот или иной элемент. Далее будет уже два списка

оставшихся $n-2$ элементов, каждый из которых потребует $Q_{\frac{n}{2}}$ операций сравнения, поэтому $Q_n = 2Q_{\frac{n}{2}} + n$. Но $n = 2^m$, так что

$$Q_n = Q_{2^m} = 2Q_{2^{m-1}} + 2^m.$$

Поэтому

$$\begin{aligned}
\frac{Q_{2^m}}{2^m} &= \frac{Q_{2^{m-1}}}{2^{m-1}} + 1 = \\
&= \frac{Q_{2^{m-2}}}{2^{m-2}} + 1 + 1 = \\
&= \frac{Q_{2^{m-2}}}{2^{m-2}} + 2 = \\
&= \frac{Q_{2^{m-3}}}{2^{m-3}} + 3 = \\
&\vdots \\
&= \frac{Q_{2^{m-m}}}{2^{m-m}} + m = \\
&= Q_1 + m\,.
\end{aligned}$$

Поскольку $n = 2^m$, то $m = \log_2(n)$, поэтому $Q_n = n(Q_1 + \log_2(n))$, что приблизительно равно $n\ln(n)$. Таким образом, число сравнений имеет порядок $O(n\log_2(n))$. Самый неблагоприятный случай и наибольшее число операций сравнения реализуются, когда элемент, который обычно разделяет множество элементов, всегда остается в конце списка. В этом случае число операций сравнения совпадает с тем, которое требуется в рассмотренных ранее алгоритмах сортировки, поэтому в самом неблагоприятном случае число сравнений имеет порядок $O(n^2)$.

Аналогичной, но более простой для понимания процедурой, является ***сортировка слиянием*** (CC). Процедура состоит в разделении элементов, подлежащих сортировке, пополам или почти пополам. Затем каждый из списков опять делится пополам. Процесс продолжается до тех пор, пока каждое множество не будет содержать только один элемент. После этого процесс обращается, списки собираются вместе таким же способом, как они разделялись, но при этом происходит их слияние. Сначала мы опишем собственно процедуру сортировки слиянием, а затем приведем пример, иллюстрирующий, как ее применять. В большинстве случаев, использующих рекурсию, процедура вызывает себя через свое имя в теле процедуры. Чтобы явно указать тот факт, что процедура вызывает сама себя, мы будем использовать команду "вызов". В приведенной процедуре переменные l, m и r представляют левый, центральный и правый элементы в текущем списке.

Процедура **CC(l, r)**:
  Если $l < r$, то
    Положить $m = \left\lfloor \frac{l+r}{2} \right\rfloor$;
  Вызов CC(l, m);
  Вызов CC($m+1, r$);
  Слияние(l, m, r);
Конец процедуры.

Прежде чем перейти к объяснению процедуры CC, нам нужно описать процедуру "Слияние". Процедура "Слияние" просто берет два уже упорядоченных списка и сравнивает их первые элементы. Меньший из этих двух элементов удаляется из своего списка и помещается в некий новый список в качестве первого элемента. Затем процедура повторяется, и всякий раз удаляемый элемент дописывается в конец нового списка. Это продолжается до тех пор, пока один из списков не опустеет. Затем остаток другого списка дописывается в конец нового списка. Полученный таким образом список и есть результат сортировки. Мы не будем формализовать процедуру "Слияние", т.к. метод ее практической реализации будет зависеть от используемой вычислительной техники.

Процедура **Слияние**(l, m, r):
  Пусть $A = (a_l, a_{l+1}, a_{l+2}, \ldots, a_m)$, $B = (a_{m+1}, a_{m+2}, a_{m+i}, \ldots, a_r)$,
    и C — новый список, результат слияния;
  Если A и B содержат в списках элементы данных, то
    Сравнить первые элементы A и B и удалить меньший
    элемент из его множества;
    Добавить удаленный элемент в конец списка C;
  Если A или B не содержат элементов, то
    добавить остаток другого списка в конец списка C;
  Переобозначить список C как $(a_l, a_{l+1}, a_{l+2}, \ldots, a_r)$;
Конец процедуры.

Пусть c, b, a, v, h, d, x, e — список для сортировки, и символ | обозначает разбиение. На первом шаге список разбивается так: $c, b, a, v|h, d, x, e$. Затем, поскольку в начале каждой записи стоит команда "вызов CC", мы снова вызываем процедуру CC для левой части, получая разбиение $c, b|a, v|h, d, x, e$. Следующий вызов процедуры CC для левой части дает разбиение $c|b|a, v|h, d, x, e$. Левую часть мы больше разбивать не можем, поэтому пытаемся разбить правую часть, чего опять-таки сделать не можем. Тогда мы производим слияние c и b, получая $b, c|a, v|h, d, x, e$. Далее возвращаемся к вызову процедуры CC, в котором разделяются b, c и a, v. Вызов CC(l, m) для левой части полностью завершен, поэтому переходим к вызову процедуры CC(m, r) для правой части, получая разбиение $b, c|a|v|h, d, x, e$. Не имея более возможности вызвать процедуру CC, производим слияние a и v. В результате получаем разбиение $bc|a, v|h, d, x, e$. Затем возвращаемся на тот уровень CC, где разделяются b, c и a, v. Итак, мы завершили вызовы CC(l, m) и CC(m, r) и можем произвести слияние b, c и a, v, получая разбиение $a, b, c, v, |h, d, x, e$. Теперь возвращаемся на тот уровень CC, где было получено разбиение на c, b, a, v и h, d, x, e. Завершив, таким образом, вызов CC(l, m), активизируем вызов CC(m, r), получая разбиение $a, b, c, v, |h, d|x, e$. Снова вызываем CC для разделения h и d; получаем разбиение $a, b, c, v, |h|d|x, e$. Поскольку мы больше не можем использовать процедуру CC, то производим слияние h и d, получая $a, b, c, v, |d, h|x, e$. Затем возвращаемся на уровень CC, где разделяются h, d и x, e. Дальнейшее использование процедуры CC(l, m) невозможно, поэтому мы переходим к вызову CC(m, r), получая разбиение $a, b, c, v, |d, h|x|e$. Теперь переходим к процедуре "Слияние" для осуществления слияния x и e, после чего получаем $a, b, c, v, |d, h|e, x$. Опять возвращаемся на тот уровень CC, где разделя-

ются h, d и x, e. Обе команды CC выполнены, поэтому производим слияние d, h и e, x, получая $a, b, c, v|d, e, h, x$. В конечном счете мы возвращаемся на уровень CC, где разбиение проходит между c, b, a, v и h, d, x, e. На этом уровне выполнены обе команды CC, поэтому мы производим слияние a, b, c, v и d, e, h, x, получая a, b, c, d, e, h, v, s. На этом процедура завершена.

По количеству операций сравнения процедура CC относится к числу наиболее эффективных. Для доказательства этого факта предположим, что $n = 2^m$ для некоторого значения m. Отсюда $m = \log_2 n$. Для фиксированного m это самый неблагоприятный случай. Если рассмотреть слияние, выполняемое начальным вызовом процедуры сортировки, то проводится слияние 2 $(= 2^1)$ списков, каждый из которых содержит 2^{m-1} элементов. Эта ситуация соответствует первому уровню. Вообще, на k-ом уровне имеется 2^k списков, в каждом по 2^{m-k} элементов. Необходимо провести 2^{k-1} сравнений пар списков, по 2^{m-k} элементов в каждом. Каждое сравнение пар списков требует не более $2^{m-k+1} - 1$ сравнений, поскольку один оставшийся элемент не должен сравниваться с каким-либо. Следовательно, на k-ом уровне будет не более

$$2^{k-1} \cdot (2^{m-k+1} - 1) = 2^{k-1} \cdot (2^{m-(k+1)} - 1)$$

операций сравнения. Поэтому общее число операций сравнения не превышает

$$\sum_{k=1}^{m} 2^{k-1} \cdot (2^{m-(k+1)} - 1) = \sum_{i=0}^{m-1} 2^i \cdot (2^{m-i} - 1) =$$

$$= \sum_{i=0}^{m-1} 2^m - 2^i = m2^m - (2^m - 1) = n \log_2 n - (n - 1),$$

и поэтому в самом неудачном случае число сравнений имеет порядок $O(n \ln n)$.

Последней рассмотрим сортировку, которая называется ***сортировка вставками***. Процедура для этой сортировки действительно очень простая. При этом методе сортировки элементы данных выбираются по одному и помещаются во вновь создаваемом списке на соответствующее место. При этом, если место элементу найдено, то все другие вставляемые элементы располагают левее, если они меньше, и правее, если они больше. Опять рассмотрим тот же список для сортировки c, b, a, v, h, d, x, e. Первым выбираем c и помещаем его в новом списке. Далее, поскольку b меньше c, помещаем его перед c, формируя b, c. Следующим выбираем a. Поскольку оно меньше b, помещаем его перед b, формируя a, b, c. Далее выбираем v. Оно больше c, поэтому помещаем его после c, формируя a, b, c, v. Поскольку h меньше, чем v, и больше, чем c, помещаем его между c и v, формируя список a, b, c, h, v. Теперь рассматриваем d. Оно больше, чем c, и меньше, чем h, поэтому помещаем его между c и h, что дает a, b, c, d, h, v. Поскольку x больше v, помещаем его после v, получая a, b, c, d, h, v, x. Наконец, поскольку e больше, чем d, и меньше, чем h, помещаем его между d и h, формируя рассортированный список a, b, c, d, e, h, v, x.

В самом плохом варианте число операций сравнения на каждом этапе такое же, как в пузырьковой сортировке, т.е. имеет порядок $O(n^2)$.

И быстрая сортировка, и сортировка слиянием рекурсивно определяют разбиение множества на меньшие множества, сортировку этих меньших множеств и последующую сборку этих рассортированных множеств. Такой метод разбиения множества на подмножества, выполнения операций над ними и обратной сборки множеств вместе называется методом "***разделяй и властвуй***". Как правило, это весьма эффективный подход к решению задач.

Очень часто процедуры типа "разделяй и властвуй", такие как приведенная выше, удовлетворяют рекурсивному соотношению вида $Q_n = CQ_{\frac{n}{2}} + f(n)$, где Q_n может характеризовать число операций, число сравнений, затрачиваемое время или другие параметры. В данном разделе в качестве такой характеристики будем рассматривать число операций. Функция $f(n)$ описывает число операций, затрачиваемых на разделение множества на два множества и последующую сборку этих двух подмножеств в целое. Когда используется запись $\frac{n}{2}$ как индекс в формуле, это предполагает, что множество данных, над которым производится действие, разбивается на две равные части. Это достаточно распространенная ситуация, т.к. наиболее часто используются бинарные деревья или другие понятия, содержащие бинарные операции. Если бы множество делилось на k равных частей, была бы использована запись $\frac{n}{k}$. Предполагается также, что n является некоторой степенью числа 2. Нас, в основном, будут интересовать случаи, когда $f(n)$ равно константе и когда $f(n) = Dn$ для некоторого постоянного D. Использование записи $f(n) = Dn$ равносильно предположению, что число операций для разбиения и последующего соединения n объектов прямо пропорционально количеству элементов множества. При изучении эффективности быстрой сортировки было показано, что решение рекурсивного соотношения $Q_n = 2Q_{\frac{n}{2}} + n$, где n — степень числа 2, имеет вид $n(Q_1 + \log_2(n))$. Это частный случай решения для соотношения $Q_n = CQ_{\frac{n}{2}} + Dn$, когда $C = 2$ и $D = 1$. Для $C > 2$ решением рекурсивного соотношения будет $Q_n = An^{\log_2 C} + \left(\frac{2D}{2-C}\right)n$, где A выбрано таким образом, что $Q_1 = A + \frac{2D}{2-C}$.

Если $f(n) = D$, где D — некоторая константа, то решением рекурсивного соотношения $Q_n = CQ_{\frac{n}{2}} + D$ для $C \neq 1$ будет $Q_n = \frac{D(Cn^{\log_2 C} - 1)}{C-1}$. В частном случае, когда $C = 2$, решение упрощается до $Q_n = D(2n - 1)$. Изучение такого рода рекурсивных соотношений будет продолжено при решении конкретных задач. К сожалению, целое число n в общем случае не является степенью числа 2. Однако, если предположить, что Q_n — возрастающая функция и n меньше некоторого a, где a — степень числа 2, то $Q_n < Q_a$, т.е. функция Q_n ограничена сверху.

■ УПРАЖНЕНИЯ

1. Рассортируйте последовательность 7, 11, 4, 0, 3, 1, 9, 4, 2, 8, используя:
   а) сортировку выбором;
   б) пузырьковую сортировку;
   в) быструю сортировку;
   г) сортировку слиянием;
   д) сортировку вставками.

2. Рассортируйте последовательность $12, 50, -1, -10, 10, 11, 52, 30, 2, 8, -12$, используя:

а) сортировку выбором;

б) пузырьковую сортировку;

в) быструю сортировку;

г) сортировку слиянием;

д) сортировку вставками.

3. Рассортируйте последовательность $x, a, c, y, p, z, f, t, m, y, u, b, d, n, s, r$, используя:

а) сортировку выбором;

б) пузырьковую сортировку;

в) быструю сортировку;

г) сортировку слиянием;

д) сортировку вставками.

4. Рассортируйте последовательность *Джонсон, Браун, Блэк, Джексон, Мэрфи, Смит, Джонс*, используя:

а) сортировку выбором;

б) пузырьковую сортировку;

в) быструю сортировку;

г) сортировку слиянием;

д) сортировку вставками.

5. Постройте самый неблагоприятный вариант последовательности для осуществления каждой из перечисленных ниже сортировок:

а) сортировка выбором;

б) пузырьковая сортировка;

в) быстрая сортировка;

г) сортировка слиянием;

д) сортировка вставками.

6. Покажите, что для слияния двух уже рассортированных списков длины n требуется не более $n - 1$ операций сравнения.

7. Напишите компьютерную реализацию алгоритма сортировки выбором.

8. Напишите компьютерную реализацию алгоритма пузырьковой сортировки.

9. Напишите компьютерную реализацию алгоритма быстрой сортировки.

10. Напишите компьютерную реализацию алгоритма сортировки слиянием.

11. Напишите компьютерную реализацию алгоритма сортировки вставками.

12. Покажите, что $Q_n = An - D$ удовлетворяет рекурсивному соотношению $Q_n = 2Q_{\frac{n}{2}} + D$.

13. Покажите, что $Q_n = Dn(\log_2 n + A)$ удовлетворяет рекурсивному соотношению $Q_n = 2Q_{\frac{n}{2}} + Dn$.

14. Покажите, что $Q_n = An^{\log_2 C} + \left(\frac{2D}{2-C}\right) n$ для $C > 2$ удовлетворяет рекурсивному соотношению $Q_n = CQ_{\frac{n}{2}} + Dn$.

15. Восстановив пропущенные выкладки, докажите, что

$$Q_n = An^{\log_2 C} + \left(\frac{2D}{2-C}\right) n$$

является решением $Q_n = CQ_{\frac{n}{2}} + Dn$ для $C > 2$, если A выбрано так, что $Q_1 = A + \frac{2D}{2-C}$. Мы предполагаем, что $n = 2^m$.

$$\begin{aligned}
\frac{Q_{2^m}}{2^m} &= \frac{C}{2}\left(\frac{Q_{2^{m-1}}}{2^{m-1}}\right) + D = \\
&= \left(\frac{C}{2}\right)^2 \frac{Q_{2^{m-2}}}{2^{m-2}} + \left(\frac{C}{2}\right) D + D = \\
&= \left(\frac{C}{2}\right)^3 \frac{Q_{2^{m-3}}}{2^{m-3}} + \left(\left(\frac{C}{2}\right)^2 + \frac{C}{2} + 1\right) D = \\
&\vdots \\
&= \left(\frac{C}{2}\right)^m Q_1 + \left(1 + \frac{C}{2} + \left(\frac{C}{2}\right)^2 + \cdots + \left(\frac{C}{2}\right)^{m-1}\right) D = \\
&= \frac{C^{\log_2 n}}{n} Q_1 + \frac{1 - \left(\frac{C}{2}\right)^m}{1 - \left(\frac{C}{2}\right)} D = \\
&= \frac{n^{\log_2 C}}{n} Q_1 + \frac{1 - \left(\frac{C}{2}\right)^m}{1 - \left(\frac{C}{2}\right)} D = \\
&\vdots
\end{aligned}$$

16. Получите решение вида $Q_n = \dfrac{D(Cn^{\log_2 C} - 1)}{C - 1}$ для рекурсивного соотношения $Q_n = CQ_{\frac{n}{2}} + D$, когда $C \neq 1$.

5.5. ПРЕФИКСНАЯ И СУФФИКСНАЯ ЗАПИСИ

В арифметике, как и в логике, существуют определенные правила, касающиеся приоритета выполнения операций над целыми числами, что дает возможность, не приводя к путанице, сокращать число используемых в выражении скобок. Некоторые свойства целых чисел также позволяют уменьшить потребность в скобках. Например, известно, что благодаря закону ассоциативности сложения выражение $3 + 5 + 7$ не требует скобок, т.к. $(3 + 5) + 7 = 3 + (5 + 7)$. Понятно также, что операции возведения в степень должны быть выполнены первыми, за ними следуют умножение и деление и, наконец, сложение и вычитание. Но даже при таком соглашении необходимо использовать скобки. Например, нельзя выразить $(a+b)^2(4+5b)$, не используя скобок. Исключить скобки можно, применив префиксную или суффиксную форму записи. Чтобы подчеркнуть их отличие от обычной

для арифметики формы записи, в которой знак бинарной операции записывается между операндами, назовем обычную форму записи инфиксной. Например, выражения $3 + 4$ и $6 \div 3$ записаны в инфиксной форме.

ОПРЕДЕЛЕНИЕ 5.30. Выражение находится в ***префиксной*** (***польской***) ***записи***, если в нем знак операции непосредственно предшествует операндам, на которые она действует. Выражение находится в ***суффиксной*** (***постфиксной***, или ***обратной польской записи***), если в нем знак операции следует непосредственно за операндами, на которые она действует. Выражение находится в ***инфиксной записи***, если в нем знак операции находится между операндами, на которые она действует.

Например, инфиксное выражение $a + b$ в префиксной записи имело бы вид $+ab$, а в постфиксной записи, соответственно, вид $ab+$. Инфиксное выражение $a \times (b+c)$ имело бы в префиксной записи вид $\times a+bc$, а в постфиксной, соответственно, вид $abc+\times$. Обратите внимание, что постфиксную запись следует "читать" слева направо, а "читать" префиксную запись необходимо справа налево. Операцию возведения в степень для удобства будем обозначать символом ^. Таким образом, a^5 будет записываться как $a\hat{}5$. Выражение $(a+b) \times (c+d)\hat{}3$ в префиксной записи имеет вид $\times + ab\hat{} + cd3$, а в постфиксной записи, соответственно, вид $ab+cd+3\hat{}\times$.

В данном разделе рассмотрены алгоритмы перехода от инфиксной записи к постфиксной и от постфиксной — к инфиксной. Подобные алгоритмы для перехода от инфиксной записи к префиксной и наоборот могут быть построены аналогичным образом.

ОПРЕДЕЛЕНИЕ 5.31. Арифметическое выражение находится в ***полной скобочной записи***, если каждому оператору в нем соответствует пара скобок, в которые заключены оператор и те операнды, на которые он действует.

Упрощенно можно сказать, что арифметическое выражение — полноскобочное, если скобки в нем присутствуют всюду, где это возможно по смыслу. Например, $(a \times (b+c))$ и $((a+b) \times ((c+d)\hat{}x))$ являются полноскобочными выражениями, а выражения $a + (b \times c)$ и $(a \div b)\hat{}c$ полноскобочными не являются.

Переход от инфиксного выражения к постфиксному может быть осуществлен посредством следующей процедуры:

1. Добавлять скобки в выражение, пока оно не станет полноскобочным.
2. Начиная с самых внутренних скобок, удалить пару скобок и поместить находящийся внутри скобок оператор на место соответствовавшей ему правой скобки. При наличии более одной самой внутренней пары скобок начинать нужно с левой пары.
3. Перейти к следующей паре скобок, удалить ее и поместить оператор, находившийся внутри скобок, на место соответствовавшей ему правой скобки.
4. Продолжать шаг (3), пока все скобки не будут удалены.

ПРИМЕР 5.32. Пусть дано инфиксное выражение $(a+b)\times(c+d)\hat{}x$. Добавляем скобки, пока выражение не станет полноскобочным. Получаем выражение $((a+b)\times((c+d)\hat{}x))$. Выполнив шаг (2), получим $(ab+\times(cd+\hat{}x))$. Выполнив шаг (3), получим $(ab+\times cd+x\hat{})$. Повторяя шаг (3), приходим к записи $ab+cd+x\hat{}\times$. □

ПРИМЕР 5.33. Пусть дано инфиксное выражение $a\times b+c\times d$. Добавляем скобки, пока не получим полноскобочное выражение, т.е. выражение $((a\times b)+(c\times d))$. Выполнив шаг (2), получим $(ab\times+cd\times)$. Повторяя шаг (3), получаем $ab\times cd\times+$. □

Неформально ***стек*** определяется как область памяти, из которой данные извлекаются в порядке, обратном порядку их ввода. Такой способ организации памяти часто называют LIFO.[1] Добавление в стек и удаление из него элемента обычно называют операциями ***вталкивания*** и ***выталкивания***, соответственно.

Постфиксное выражение может быть заменено инфиксным посредством следующей процедуры:

1. Начинать считывание выражения слева направо.
2. Если считанный символ не является символом операции, то поместить его в стек и продолжать чтение.
3. Если считанный символ является символом операции, то:
   - **а)** вытолкнуть из стека два верхних элемента;
   - **б)** поместить символ операции между вторым и первым элементом из стека и поставить скобки;
   - **в)** поместить выражение, полученное на шаге (б), обратно в стек.
4. Продолжать чтение слева направо и выполнять шаги (2) и (3) до завершения.

ПРИМЕР 5.34. Пусть дано выражение $abc\times+$. Вначале считываем a и, поскольку это не символ операции, помещаем этот элемент в стек, как показано на рис. 5.4. Затем считываем b и, т.к. это не символ операции, помещаем его в стек, как показано на рис. 5.5.

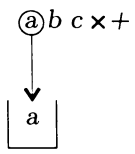


Рис. 5.4

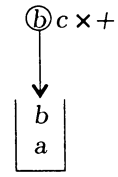


Рис. 5.5

Далее считываем c и, т.к. это не символ операции, помещаем его в стек, как показано на рис. 5.6. Затем считываем $\times$ и, поскольку это символ операции, выталкиваем из стека (удаляем) c (верхний элемент) и b (следующий элемент), помещаем $\times$ между b и c, как показано на рис. 5.7.

[1] От англ. last-in-first-out — "последним вошел, первым вышел". — *Прим. перев.*

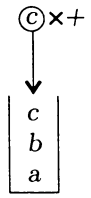


Рис. 5.6

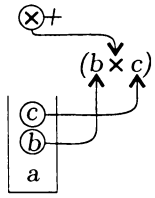


Рис. 5.7

Помещаем $(b \times c)$ в стек, как показано на рис. 5.8. Далее считываем $+$ и, поскольку это символ операции, выталкиваем из стека (удаляем) $(b \times c)$ и a (следующий элемент). Помещаем $+$ между a и $(b \times c)$, как показано на рис. 5.9.

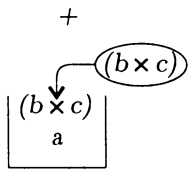


Рис. 5.8

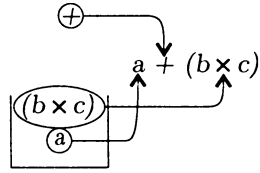


Рис. 5.9

Помещаем $(a + (b \times c))$ в стек, как показано на рис. 5.10. Получаем искомое выражение. Очевидно, лишние скобки можно удалить.

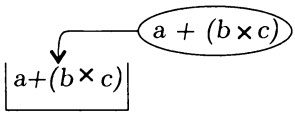


Рис. 5.10

□

ПРИМЕР 5.35. Пусть дано выражение $ab + cd + 3\hat{} \times$. Сначала считываем a и, т.к. это не символ операции, помещаем его в стек. Следующим считываем b и, т.к. это не символ операции, помещаем его в стек. Далее считываем $+$ и, поскольку это символ операции, выталкиваем из стека (удаляем) b (верхний элемент) и a (следующий элемент) и помещаем $+$ между a и b. Переносим $(a + b)$ в стек. Далее считываем c. Поскольку c не является символом операции, то помещаем его в стек. Далее считываем d и, поскольку это не символ операции, переносим его в стек. Затем считываем $+$ и, поскольку это символ операции, выталкиваем из стека (удаляем) d (верхний элемент) и c (следующий элемент), помещаем $+$ между c и d. Переносим $(c + d)$ в стек. Далее считываем 3 и, поскольку это не символ операции, помещаем 3 в стек. Затем считываем символ $\hat{}$ и, поскольку это символ операции, выталкиваем из стека 3 (верхний элемент) и $(c+d)$ (следующий элемент), после чего помещаем символ $\hat{}$ между $(c+d)$ и 3. Переносим $((c+d)\hat{}3)$ в стек. Считываем $\times$ и, т.к. это символ операции, выталкиваем $((c+d)\hat{}3)$ и $(a+b)$ из стека. Помещаем $\times$ между $(a + b)$ и $((c + d)\hat{}3)$, образуя $(a + b) \times ((c + d)\hat{}3)$. Переносим $((a+b) \times ((c+d)\hat{}3))$ в стек. Это и есть искомый результат. Очевидно, лишние скобки можно удалить. □

■ УПРАЖНЕНИЯ

1. Замените следующие инфиксные выражения постфиксными:

а) $a\hat{}2 + b\hat{}2$;

б) $a\hat{}3 + 3\ a\hat{}2 + 7$;

в) $(a\hat{}2 + b)(c + d\hat{}2)$;

г) $(a\hat{}2 + b\hat{}3)\hat{}4$;

д) $a\hat{}2(b + c)$.

2. Замените следующие инфиксные выражения постфиксными:

а) $(c + d) - (a + b)\hat{}2$;

б) $((a + b) - c)\hat{}3$;

в) $(a + b)\hat{}2 - (c + d)\hat{}2$;

г) $(a - b)(a + b)$;

д) $a - b\hat{}(c + d)$.

3. Замените следующие постфиксные выражения инфиксными:

а) $ab + cd - \div$;

б) $ab2\hat{} + cd2\hat{} + \times$;

в) $3ab \times \times 4cd2\hat{} \times \times +$;

г) $ab + \hat{}2cd\hat{} + \div$;

д) $a2\hat{}b + cd2\hat{} + \times$.

4. Замените следующие постфиксные выражения инфиксными:

а) $abc + +3\hat{}$;

б) $ab + 2\hat{}c + cd2\hat{}+$;

в) $3ab \times \times 4cd2\hat{} \times \times +$;

г) $ab + \hat{}2cd\hat{} + \div$;

д) $a2\hat{}b + c + cd2\hat{} + \div$.

5. Замените следующие инфиксные выражения префиксными:

а) $a\hat{}2 + b\hat{}2$;

б) $a\hat{}3 + 3\ a\hat{}2 + 7$;

в) $(a\hat{}2 + b)(c + d\hat{}2)$;

г) $(a\hat{}2 + b\hat{}3)\hat{}4$;

д) $a\hat{}2(b + c)$.

6. Замените следующие инфиксные выражения префиксными:

а) $(c + d) - (a + b)\hat{}2$;

б) $((a + b) - c)\hat{}3$;

в) $(a + b)\hat{}2 - (c + d)\hat{}2$;

г) $(a - b)(a + b)$;

д) $a - b\hat{}(c + d)$.

7. Замените следующие префиксные выражения инфиксными:

а) $\div + ab - cd$;

б) $\div + a2\hat{}d2 + \hat{}c2\hat{}b2$;

в) $+ \times \times 3ab + \times \times 4cd \times \times 5ad$;

г) $\hat{} + \hat{}a2 + \hat{}b2\hat{}c23$;

д) $\times + a\hat{}2b + cd\hat{}2$.

8. Замените следующие префиксные выражения инфиксными:

а) $\hat{} + \times abc3$;

б) $\times + +abc + +cde$;

в) $\hat{} + \times \times 3ab \times \times 4cd2$;

г) $\hat{} \div +a\hat{}b2 + c\hat{}d22$;

д) $\times + -abc - -abc$.

9. Напишите процедуру для перехода от инфиксной формы записи выражения к префиксной.

10. Напишите процедуру для перехода от постфиксной формы записи выражения к инфиксной.

5.6. ДВОИЧНЫЕ И ШЕСТНАДЦАТЕРИЧНЫЕ ЧИСЛА

Использование цифры 0 для обозначения разрядов в записи числа привело к большому прорыву в развитии математики. Оно дало возможность различать такие числа, как 47, 4700, 4007 и 4070. В данном примере значение четверки и семерки определяется позицией, которую они занимают в записи числа. Например, в числе 47 цифры 4 и 7 находятся в таких позициях, что цифра 4 обозначает 4 десятка, а цифра 7 обозначает 7 единиц. В числе 4700 цифра 4 обозначает 4 тысячи, а цифра 7 обозначает 7 сотен. Таким образом, число 47 можно записать как $4 \cdot 10 + 7$, а число 4700 — как $4 \cdot 1000 + 7 \cdot 100$ или $4 \cdot 10^3 + 7 \cdot 10^2$. Точно так же, число 4007 можно записать как $4 \cdot 10^3 + 7 \cdot 1$, а число 4070 — как $4 \cdot 10^3 + 7 \cdot 10$. В общем случае, если число n представляется в ***десятичной системе счисления***, в основе которой лежит число 10, как $a_m a_{m-1} a_{m-2} \cdots a_2 a_1 a_0$, то

$$n = a_m(10)^m + a_{m-1}(10)^{m-1} + \cdots + a_2(10)^2 + a_1(10) + a_0,$$

где a_i — целые числа между 0 и 9 для всех $0 \leq i \leq m$. Например, $3124 = 3(10)^3 + 1(10)^2 + 2(10) + 4$. Аналогично, если число n представлено в ***двоичной системе счисления***, в основе которой лежит число 2, то каждая цифра должна быть либо 1, либо 0, и ее значение определяется занимаемой позицией. Таким образом, $1101 = 1(2)^3 + 1(2)^2 + 0(2) + 1 = 13$, где число 13 представлено в десятичной системе. Первые 20 положительных целых чисел, записанные в двоичной системе, имеют вид: 1, 10, 11, 100, 101, 110, 111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111, 10000, 10001, 10010, 10011 и 10100. В общем случае, двоичному представлению $n = a_m a_{m-1} a_{m-2} \cdots a_2 a_1 a_0$ соответствует десятичное представление

$$n = a_m(2)^m + a_{m-1}(2)^{m-1} + \cdots + a_2(2)^2 + a_1(2) + a_0,$$

где a_i равно либо 1, либо 0 для всех $0 \leq i \leq m$. Для исключения возможной путаницы двоичное число, например, 1101101, будем обозначать как 1101101_2, а десятичное число, например, 312, будем обозначать как 312_{10}.

Когда на одометре машины достигается число 9999 и добавляется еще одна миля, то все девятки сбрасываются и получается 10,000. Аналогично, если к числу 11111_2 прибавить 1, то все единицы сбрасываются и получается число 100000_2. Таким образом, если прибавить 1 к 10111_2, получим 11000_2. Для того чтобы перевести двоичное число $a_m a_{m-1} a_{m-2} \cdots a_2 a_1 a_0$ в десятичное, нам просто нужно вычислить значение этого числа, равное $a_m(2)^m + a_{m-1}(2)^{m-1} + \cdots + a_2(2)^2 + a_1(2) + a_0$, в обычной десятичной системе счисления. Например, 1101011_2 записывается в виде

$$\begin{array}{cccccccc} 1 & 1 & 0 & 1 & 0 & 1 & 1 & \\ 2^6 & 2^5 & 2^4 & 2^3 & 2^2 & 2^1 & 1 & \end{array} \begin{array}{rl} = & 1(2)^6 + 1(2)^5 + 0(2)^4 + 1(2)^3 + \\ & + 0(2)^2 + 1(2)^1 + 1 = \\ = & 64 + 32 + 8 + 4 + 1 = \\ = & 109. \end{array}$$

Точно так же находим

$$10101 = 1(2)^4 + 0(2)^3 + 1(2)^2 + 0(2)^1 + 1 = 16 + 4 + 1 = 21.$$

Этот метод нахождения значения двоичного числа часто называют "подсчет обмена", поскольку он аналогичен подсчету наличных денег, если имеется монета в 25 центов, монета в 10 центов и пенни. Сначала мы опишем метод перехода от десятичного числа к двоичному, который часто называют "разменом". Если необходимо определить, какие монеты отдать в обмен за 42 цента, первой будет выбрана монета в 25 центов, поскольку это наибольшая монета, которая меньше 42 центов. Если вычесть 25 центов, то необходимо еще 17 центов. Теперь выбираем монету в 10 центов, поскольку это наибольшая монета, которая меньше, чем 17 центов. Вычитаем 10 центов из 17 центов и находим, что необходимо еще 7 центов. Следующей выбираем монету в 5 центов и 5 центов вычитаем, остается еще 2 цента, которые необходимо добавить. Подобным образом, если имеется, например, число 83_{10}, которое нужно перевести в двоичное, то сначала выбирается наибольшая степень 2, которая не превосходит 83_{10}. В данном случае это $2^6 = 64_{10}$. Помещаем 1 в колонку, соответствующую 2^6. Теперь вычитаем 64_{10} из 83_{10} и получаем 19_{10}:

$$\begin{array}{ccccccc} 1 & & & & & & \\ 2^6 & 2^5 & 2^4 & 2^3 & 2^2 & 2^1 & 1 \end{array}$$

Нам не нужно использовать $2^5 = 32_{10}$, т.к. это число слишком велико, поэтому в колонке, соответствующей 2^5, помещаем 0:

$$\begin{array}{ccccccc} 1 & 0 & & & & & \\ 2^6 & 2^5 & 2^4 & 2^3 & 2^2 & 2^1 & 1 \end{array}$$

Но $2^4 = 16_{10}$ нам нужно, поэтому помещаем 1 в колонку, соответствующую 2^4:

$$\begin{array}{ccccccc} 1 & 0 & 1 & & & & \\ 2^6 & 2^5 & 2^4 & 2^3 & 2^2 & 2^1 & 1 \end{array}$$

Вычитаем 16_{10} из 19_{10}и получаем 3_{10}. Нам не нужно использовать $2^3 = 8_{10}$ или $2^2 = 4_{10}$, т.к. они слишком велики, поэтому помещаем 0 в соответствующие колонки. Но 2 и 1 нам нужны, поэтому мы помещаем по единице в соответствующие колонки:

1	0	1	0	0	1	1
2^6	2^5	2^4	2^3	2^2	2^1	1

Следовательно, искомым числом является 1010011_2.

Аналогично, $199_{10} = 11000111$, т.к. для этого необходимо одно $2^7 = 128_{10}$, одно $2^6 = 64_{10}$, одно $2^2 = 4_{10}$, одно 2_{10} и одно 1_{10}.

Можно получить другой способ преобразования десятичного числа в двоичное, если проанализировать такой процесс: при делении 234 на 10 обычным арифметическим способом получаем частное 23 и остаток 4. Если теперь делить 23 на 10, частным будет 2, а остатком — 3. Наконец, если разделить 2 на 10, получаем частное 0 и остаток 2. Обратите внимание, что расположение остатков в обратном порядке дает исходное число. Используя приведенный ниже алгоритм, целое десятичное число n можно преобразовать в то же самое по значению, но двоичное число n.

1. Разделить n на 10, получив частное q_1 и остаток r_1.
2. Разделить q_1 на 10, получив частное q_2 и остаток r_2.
3. До тех пор, пока q_k не равно 0, делить q_k на 10, получая частное q_{k+1} и остаток r_{k+1}.
4. Если $q_n = 0$, тогда $r_n r_{n-1} \cdots r_3 r_2 r_1 = n$.

Использование приведенного выше алгоритма в двоичной системе счисления, где 10 соответствует десятичному числу 2, обеспечит такой же результат. Но, поскольку 10_2 есть десятичное число 2, деление на 10 в двоичной системе должно давать такие же остатки, как и деление на 2 в десятичной арифметике. Рассмотрим приведенные ниже соответствующие операции для преобразования десятичного числа 57 в двоичное число, в которых все числа одни и те же, за исключением того, что в первом столбце они выражены в двоичной форме, а во втором — в десятичной форме.

двоичная запись	десятичная запись
Разделить 111001 на 10; $q_1 = 11100$, $r_1 = 1$.	Разделить 57 на 2; $q_1 = 28$, $r_1 = 1$.
Разделить 11100 на 10; $q_2 = 1110$, $r_2 = 0$.	Разделить 28 на 2; $q_2 = 14$, $r_2 = 0$.
Разделить 1110 на 10; $q_3 = 111$, $r_3 = 0$.	Разделить 14 на 2; $q_3 = 7$, $r_3 = 0$.
Разделить 111 на 10; $q_4 = 11$, $r_4 = 1$.	Разделить 7 на 2; $q_4 = 3$, $r_4 = 1$.
Разделить 11 на 10; $q_5 = 1$, $r_5 = 1$.	Разделить 3 на 2; $q_5 = 1$, $r_5 = 1$.
Разделить 1 на 10; $q_6 = 0$, $r_6 = 1$.	Разделить 1 на 2; $q_6 = 0$, $r_6 = 1$.
$111001 = 111001$.	$57 = 111001$.

Следующий алгоритм должен переводить число n из десятичной системы счисления в двоичную:

1. Разделить n на 2, получив частное q_1 и остаток r_1.
2. Разделить q_1 на 2, получив частное q_2 и остаток r_2.

3. До тех пор, пока q_k не равно 0, делить q_k на 2, получая частное q_{k+1} и остаток r_{k+1}.
4. Если $q_n = 0$, тогда $(r_n r_{n-1} \cdots r_3 r_2 r_1)_2 = n_{10}$.

Для того чтобы записать целое положительное число в ***шестнадцатеричной системе счисления***, в основе которой лежит число 16, нам понадобятся новые символы, поскольку все неотрицательные целые числа, меньшие 16, должны представляться одной цифрой. Иными словами, 10_{16} эквивалентно 16 в десятичной форме. Вводим "цифры" A, B, C, D, E и F для обозначения чисел 10, 11, 12, 13, 14 и 15. Таким образом, первые 20 целых положительных чисел в шестнадцатеричной форме имеют вид 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, 10, 11, 12, 13 и 14. Если число n в шестнадцатеричной форме представлено как $a_m a_{m-1} a_{m-2} \cdots a_2 a_1 a_0$, то ему соответствует представление в десятичной форме,

$$n = a_m(16)^m + a_{m-1}(16)^{m-1} + \cdots + a_2(16)^2 + a_1(16) + a_0,$$

где для всех $0 \le i \le m$, $0 \le a_i < 16$. Поэтому числу $n = 2A3B4_{16}$ соответствовало бы десятичное число $n = 2(16)^4 + A(16)^3 + 3(16)^2 + B(16) + 4$. Используем следующую таблицу:

$$\begin{aligned} 2(16)^4 &= 2 \times 65536 = 130072 \\ A(16)^3 &= 10 \times 4096 = 40960 \\ 3(16)^2 &= 3 \times 256 \quad = 768 \\ B(16) &= 11 \times 16 \quad = 176 \\ 4 &= 4 \times 1 \quad = 4 \\ &\qquad \text{Сумма} \quad = 171980 \end{aligned}$$

и $2A3B4_{16} = 171980_{10}$.

Аналогично, $3C9_{16} = 3(16)^2 + 12(16) + 9 = 3 \times 256 + 12 \times 16 + 9 = 969_{10}$. Способом, аналогичным "сбрасыванию" для двоичных чисел, после прибавления 1 к $FFFFF$ получаем 100000. Поэтому, если прибавить 1 к $23CFFFF$, получим $23D0000$.

Первый способ перевода десятичного числа в шестнадцатеричное аналогичен способу перевода десятичного числа в двоичное. Если дано десятичное число, например, 530328, наибольшей степенью числа 16, которая разделит 530328 с ненулевым частным, будет $16^4 = 65536$. Поскольку частное от деления 530328 на 66536 равно 8, запишем 8 в первом столбце:

8				
65536	4096	256	16	1
16^4	16^3	16^2	16^1	16^0

Разделим разность $530328 - 8(65536) = 6040$ на 4096. Частное от деления равно 1, остаток равен 1944, поэтому помещаем 1 во второй столбец:

8	1			
66536	4096	256	16	1
16^4	16^3	16^2	16^1	16^0

и делим 1944 на 256. Поскольку частное от деления равно 7, а остаток равен 152, помещаем 7 в третий столбец:

$$\begin{array}{ccccc} 8 & 1 & 7 & & \\ 66536 & 4096 & 256 & 16 & 1 \\ 16^4 & 16^3 & 16^2 & 16^1 & 16^0 \end{array}$$

и делим 152 на 16.

Частное от деления равно 9, остаток равен 8, поэтому помещаем 9 в четвертом участке, а 8 — в последнем столбце:

$$\begin{array}{ccccc} 8 & 1 & 7 & 9 & 8 \\ 66536 & 4096 & 256 & 16 & 1 \\ 16^4 & 16^3 & 16^2 & 16^1 & 16^0 \end{array}$$

Следовательно, $530328_{10} = 81798_{16}$.

Для использования альтернативного метода перевода шестнадцатеричного числа в десятичное необходимо делить на $16_{10} = 10_{16}$ вместо деления на $2_{10} = 10_2$ аналогично предыдущему алгоритму перевода десятичного числа в двоичное. Таким образом, имеем следующий алгоритм:

1. Разделить n на 16, получив частное q_1 и остаток r_1.
2. Разделить q_1 на 16, получив частное q_2 и остаток r_2.
3. До тех пор, пока q_k не равно 0, делить q_k на 16, получая частное q_{k+1} и остаток r_{k+1}.
4. Если $q_n = 0$, то $(r_n r_{n-1} \cdots r_3 r_2 r_1)_2 = n_{10}$.

ПРИМЕР 5.36. Снова рассмотрим преобразование 530328 в шестнадцатеричное число. Поскольку $530328 = 16(33145) + 8$, то $q_1 = 33145$ и $r_1 = 8$. Поскольку $33145 = 16(2071) + 9$, то $q_2 = 2071$ и $r_2 = 9$. Поскольку $2071 = 16(129) + 7$, то $q_3 = 129$ и $r_3 = 7$. Поскольку $129 = 16(8) + 1$, то $q_4 = 8$ и $r_4 = 1$. Следовательно, $r_5 = 8$ и $530328_{10} = 81798_{16}$. □

Рациональное число в десятичной системе счисления, например, 326.149, представляется в виде

$$3(10)^2 + 2(10)^1 + 6(10)^0 + 1(10)^{-1} + 4(10)^{-2} + 3(10)^{-3},$$

или

$$\begin{array}{cccccc} 3 & 2 & 6 & 1 & 4 & 3 \\ (10)^2 & (10)^1 & (10)^0 & (10)^{-1} & (10)^{-2} & (10)^{-3}. \end{array}$$

Аналогичным образом рациональное число в двоичной системе счисления, например, 1011.0111, представляется в виде

$$1(2)^3 + 0(2)^2 + 1(2)^1 + 1(2)^0 + 0(2)^{-1} + 1(2)^{-2} + 1(2)^{-3} + 1(2)^{-4},$$

или

$$\begin{array}{cccccccc} 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ (2)^3 & (2)^2 & (2)^1 & (2)^0 & (2)^{-1} & (2)^{-2} & (2)^{-3} & (2)^{-4}. \end{array}$$

Поэтому для преобразования двоичной записи рационального числа в десятичную необходимо каждый разряд числа преобразовать в десятичную систему счисления и результаты сложить.

ПРИМЕР 5.37. Преобразовать 1011.1111_2 в десятичную форму.

$$\begin{aligned}1011.1111_2 &= 1(2)^3 + 0(2)^2 + 1(2)^1 + 1(2)^0 + 1(2)^{-1} + \\ &\quad + 1(2)^{-2} + 1(2)^{-3} + 1(2)^{-4} = \\ &= 8 + 0 + 2 + 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \\ &= 11 + .5 + .25 + .125 + .0625 = \\ &= 11.9375.\end{aligned}$$

□

ПРИМЕР 5.38. Преобразовать 1010.0101_2 в десятичную форму.

$$\begin{aligned}1010.0101_2 &= 1(2)^3 + 0(2)^2 + 1(2)^1 + 0(2)^0 + 0(2)^{-1} + \\ &\quad + 1(2)^{-2} + 0(2)^{-3} + 1(2)^{-4} = \\ &= 8 + 0 + 2 + 0 + 0 + \frac{1}{4} + 0 + \frac{1}{16} = \\ &= 10 + .25 + .0625 = \\ &= 10.3125.\end{aligned}$$

□

Заметим, что .3729, умноженное на 10, дает 3.729, где первая цифра справа от десятичной точки передвигается на позицию слева от десятичной точки. Если после этого число .729 умножить на 10, то 7 передвинется на позицию слева от десятичной точки. Умножая теперь .29 на 10, мы передвинем 2 в позицию слева от десятичной точки. Наконец, если умножить .9 на 10, то 9 переместится в позицию слева от десятичной точки. Удалив эту цифру, мы получим 0. Цифры, последовательно перенесенные слева направо относительно десятичной точки, вновь образуют исходное число .3729. Таким образом, если для заданного десятичного числа, меньшего единицы, последовательно повторять процесс, состоящий в умножении числа на десять, выборе цифры, стоящей слева от десятичной точки, и последующем ее удалении, до тех пор, пока не останется ни одной цифры, то мы выберем все цифры, входящие в десятичную запись исходного числа.

Тот же самый процесс применим и в случае, если задано двоичное число, меньшее единицы. Пусть, например, задано двоичное число .1101. Умножая его на 10_2, т.е. на два, получаем 1.101. Запомнив и удалив 1, мы умножаем .101 на 10, получая 1.01. Вновь запомнив и удалив 1, умножаем .01 на 10, получая 0.1. Наконец, запомнив и удалив 0, умножаем .1 на 10; получаем и запоминаем 1. Перемещенные и сохраненные в памяти цифры — это цифры, образующие исходную двоичную запись. Если десятичное число, меньшее 1, умножить на 2, то слева от десятичной точки должна появляться та же цифра, что и при умножении этого числа в двоичной записи на 10_2, и которая является первой цифрой числа в двоичной форме записи. Таким образом, если умножать числа в десятичной системе счисления на 2, запоминать и удалять цифры слева от десятичной точки, то, продолжая этот процесс, мы должны получить цифры, которые образуют

число в двоичной записи. Ниже приведены соответствующие операции, преобразующие .125 в двоичное число. В них фигурируют одни и те же числа, но в первом столбце они выражены в двоичной, а во втором — в десятичной форме:

двоичная запись	десятичная запись
Умножим .001 на 10 и получим 0.01.	Умножим .125 на 2 и получим 0.250.
Удалим 0 и запомним как 1-ю цифру.	Удалим 0 и запомним как 1-ю цифру.
Умножим .01 на 10 и получим 0.1.	Умножим .250 на 2 и получим 0.500.
Удалим 0 и запомним как 2-ю цифру.	Удалим 0 и запомним как 2-ю цифру.
Умножим .1 на 10 и получим 1.0.	Умножим .500 на 2 и получим 1.000.
Удалим 1 и запомним как 3-ю цифру.	Удалим 1 и запомним как 3-ю цифру.
$.001_2 = .001_2$.	$.125_{10} = .001_2$.

ПРИМЕР 5.39. Заменить десятичную запись числа .875 двоичной. Сначала умножаем .875 на 2, получаем 1.750. Удаляем 1, сохраняя 1 как первую цифру двоичного числа. Умножаем .750 на 2, получаем 1.500. Удаляем 1, сохраняя 1 как вторую цифру двоичного числа. Умножаем .500 на 2, получаем 1.000. Удаляем 1, сохраняя 1 как третью цифру двоичного числа. Следовательно, наше число в двоичной форме записи имеет вид .111.

В приведенной ниже таблице первый столбец содержит этапы умножения на 2, второй — результаты умножения, а третий — те цифры, которые нужно удалить и запомнить для формирования двоичной записи исходного десятичного числа.

$$\begin{array}{lll} .875 \times 2 & 1.750 & 1 \\ .750 \times 2 & 1.500 & 1 \\ .500 \times 2 & 1.000 & 1 \end{array}$$

Таким образом, $.875_{10} = .111_2$. □

ПРИМЕР 5.40. Десятичную форму записи числа .4 заменить двоичной формой записи. Используя приведенную выше процедуру, получаем

$$\begin{array}{lll} .4 \times 2 & 0.8 & 0 \\ .8 \times 2 & 1.6 & 1 \\ .6 \times 2 & 1.2 & 1 \\ .2 \times 2 & 0.4 & 0 \\ .4 \times 2 & 0.8 & 0 \\ .8 \times 2 & 1.6 & 1 \\ .6 \times 2 & 1.2 & 1 \\ .2 \times 2 & 0.4 & 0 \end{array}$$

Заметим, что когда 0.8 повторяется во втором столбце, повторяется и весь процесс, так что представление числа 0.8 в двоичной форме состоит из повторяющихся цифр 0110. Следовательно, $.4_{10} = .0110011001100110\cdots = .\overline{0110}$ □

Действуя способом, подобным преобразованию двоичного числа в десятичное, замечаем, что $32F.12$ в шестнадцатеричной системе счисления записывается в виде

$$3(16)^2 + 2(16)^1 + F(16)^0 + 1(16)^{-1} + 2(16)^{-2},$$

что эквивалентно

$$3(16)^2 + 2(16)^1 + 15(16)^0 + 1(16)^{-1} + 2(16)^{-2},$$

или

$$3 \times 256 + 2 \times 16 + 15 + \frac{1}{16} + \frac{2}{256} = 815 + .0625 + .0078125 = \\ = 815.0703125_{10}.$$

Поэтому $32F.12_{16} = 815.0703125_{10}$. Как видим, для чисел, меньших единицы, в шестнадцатеричной системе счисления умножение на 16 — это то же самое, что умножение на 10 в десятичной. Поэтому, по аналогии с переходом от десятичного числа к двоичному, где производилось умножение на $2(= 10_2)$, мы теперь умножаем на $16(= 10_{16})$, запоминаем и удаляем цифру слева от десятичной точки и используем ее как первую цифру. Далее, умножая на 16, используем цифры слева от десятичной точки как очередные цифры записи числа в шестнадцатеричной форме. Например, рассмотрим число $.0703125_{10}$. Умножая на 16, получаем 1.1250000, так что первой цифрой шестнадцатеричной записи будет 1. Умножая далее $.125 \times 16$, получаем 2.0. Поэтому $.0703125_{10} = .12_{16}$:

$$\begin{array}{rcc} .0703125 \times 16 & 1.125 & 1 \\ .125 \times 16 & 2.0 & 2 \end{array}$$

ПРИМЕР 5.41. Требуется заменить десятичную запись числа .875 шестнадцатеричной. Умножая .875 на 16, получаем $14 = E_{16}$. Поэтому $.875_{10} = .E_{16}$. □

ПРИМЕР 5.42. Требуется заменить десятичную запись числа .0078125 шестнадцатеричной. Имеем

$$\begin{array}{rcc} .0078125 \times 16 & 0.125 & 0 \\ .125 \times 16 & 2.0 & 2 \end{array}$$

и $.0078125_{10} = .02_{16}$ □

ПРИМЕР 5.43. Преобразуем десятичное число .8 в шестнадцатеричное. Имеем

$$\begin{array}{rcc} .8 \times 16 & 12.8 & C \\ .8 \times 16 & 12.8 & C \\ .8 \times 16 & 12.8 & C \end{array}$$

и, таким образом, $.8_{10} = .CCCC\cdots = .\overline{C}$. □

Приведенные выше примеры показывают, что периодичность или конечность представления рационального числа является функцией основания соответствующей системы счисления, а не самого рационального числа.

В информатике зачастую важно иметь возможность переходить от шестнадцатеричной формы записи числа к двоичной и обратно. Заметим, что для представления в двоичной форме записи $\cdots a_{11}a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1a_0$ имеем

$$\begin{aligned}\cdots a_{11}a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1a_0 &= \cdots a_{11}2^{11} + a_{10}2^{10} + a_9 2^9 a_8 2^8 + \\ &+ a_7 2^7 + a_6 2^6 + a_5 2^5 + a_4 2^4 + \\ &+ a_3 2^3 + a_2 2^2 + a_1 2^1 + a_0 = \\ &= \cdots (a_{11}2^3 + a_{10}2^2 + a_9 2^1 a_8)2^8 + \\ &+ (a_7 2^3 + a_6 2^2 + a_5 2^1 + a_4)2^4 + \\ &+ a_3 2^3 + a_2 2^2 + a_1 2^1 + a_0 = \\ &= \cdots (a_{11}2^3 + a_{10}2^2 + a_9 2^1 a_8)16^2 + \\ &+ (a_7 2^3 + a_6 2^2 + a_5 2^1 + a_4)16 + \\ &+ a_3 2^3 + a_2 2^2 + a_1 2^1 + a_0.\end{aligned}$$

Тогда, если $\cdots a_{11}2^3 + a_{10}2^2 + a_9 2^1 + a_8$, $a_7 2^3 + a_6 2^2 + a_5 a^2 + a_4$ и $a_3 2^3 + a_2 2^2 + a_1 2^1 + a_0$ преобразованы, соответственно, в шестнадцатеричные числа $\cdots h_2$, h_1 и h_0, получаем число $\cdots + h_2(16)^2 + h_1(16) + h_0$, которое представляет собой шестнадцатеричное число $\cdots h_2 h_1 h_0$. Поэтому объединение двоичных цифр в группы по четыре и преобразование их в шестнадцатеричные числа дает преобразование всего числа из двоичной формы в шестнадцатеричную. Например, рассмотрим двоичное число 1010011001010111. Первые четыре цифры 1010 формируют шестнадцатеричную цифру A, следующие четыре цифры 0110 формируют шестнадцатеричную цифру 6, следующие четыре цифры 0101 дают шестнадцатеричную цифру 5, последние четыре цифры 0111 дают шестнадцатеричную цифру 7. Поэтому шестнадцатеричное представление рассматриваемого числа $A657$ имеет вид:

$$\underbrace{1010}_{A}\ \underbrace{0110}_{6}\ \underbrace{0101}_{5}\ \underbrace{0111}_{7}$$

Если количество цифр в двоичном представлении не кратно четырем, необходимо впереди дописать нули так, чтобы количество цифр в двоичном представлении стало кратно четырем. Например, пусть дано двоичное число 1010011110. Дописываем впереди 00 и получаем 001010011110, после чего делаем следующее преобразование:

$$\underbrace{0010}_{2}\ \underbrace{0110}_{6}\ \underbrace{1001}_{9}\ \underbrace{1110}_{E}$$

Пусть задано двоичное представление рационального числа, такое, например, как 101100.0111101. Начиная с десятичной точки и двигаясь от нее в каждом направлении, формируем группы по четыре двоичных цифры, дописывая нули в конце каждой группы, если это необходимо. Затем преобразовываем каждую группу из четырех двоичных цифр (двоичных разрядов) в шестнадцатеричную цифру (шестнадцатеричный разряд). Полученное шестнадцатеричное число является представлением исходного числа. Отсюда имеем:

$$\underbrace{0010}_{2}\ \underbrace{1100}_{C}\ .\underbrace{0111}_{.7}\ \underbrace{1010}_{A}$$

Таким образом, $101100.0111101_2 = 2C.7A_{16}$. Точно так же получаем $011100.11_2 = 00011100.1100 = 1C.C_{16}$.

Строим обратный процесс, преобразуя шестнадцатеричные числа в двоичные. Пусть дано шестнадцатеричное число $72F5$. Поскольку $7_{16} = 0111_2$, $2_{16} = 0010_2$, $F_{16} = 1111_2$ и $5_{16} = 0101_2$, то $72F5_{16} = 0111001011110101_2$. Аналогично, $2A.3C = 00101010.00111100 = 101010.001111$.

Для чисел, представленных в двоичной форме, справедливы те же правила сложения, вычитания, умножения и деления, как и для чисел, представленных в десятичной форме, исключая тот факт, что сложение и вычитание производится согласно приведенным ниже таблицам:

+	0	1
0	0	1
1	1	10

×	0	1
0	0	0
1	0	1

При вычитании важно помнить, что если 1 "занято" из 10, то остается только 1, т.к. 10_2 в действительности есть число 2.

ПРИМЕР 5.44. Требуется выполнить следующие операции в двоичной системе счисления:

а)
$$\begin{array}{r} 1011 \\ 0110 \\ +0101 \\ \hline 10110 \end{array}$$

б)
$$\begin{array}{r} 1100 \\ -1001 \\ \hline 011 \end{array}$$

в)
$$\begin{array}{r} 1101 \\ \times 101 \\ \hline 1101 \\ 0000 \\ 1101 \\ \hline 1000001 \end{array}$$

□

Аналогично, для шестнадцатеричных чисел правила сложения, вычитания, умножения и деления остаются теми же, за исключением того, что для выполнения сложения и умножения должны быть использованы таблицы, представленные на рис. 5.11 и 5.12.

ПРИМЕР 5.45. Требуется выполнить следующие операции в шестнадцатеричной системе счисления:

а)
$$\begin{array}{r} F051 \\ 315A \\ +4163 \\ \hline 1630E \end{array}$$

б)
$$\begin{array}{r} FD32 \\ -475A \\ \hline B5D8 \end{array}$$

в)
$$\begin{array}{r} 28 \\ \times 34 \\ \hline A0 \\ 78 \\ \hline 820 \end{array}$$

□

■ УПРАЖНЕНИЯ

1. Запишите числа от 1 до 25 в двоичной системе счисления.
2. Запишите числа от 1 до 25 в шестнадцатеричной системе счисления.
3. Преобразуйте следующие десятичные числа в двоичные:
   а) 45; **б)** 86;
   в) 243; **г)** 194;
   д) 312.

+	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	10
2	3	4	5	6	7	8	9	A	B	C	D	E	F	10	11
3	4	5	6	7	8	9	A	B	C	D	E	F	10	11	12
4	5	6	7	8	9	A	B	C	D	E	F	10	11	12	13
5	6	7	8	9	A	B	C	D	E	F	10	11	12	13	14
6	7	8	9	A	B	C	D	E	F	10	11	12	13	14	15
7	8	9	A	B	C	D	E	F	10	11	12	13	14	15	16
8	9	A	B	C	D	E	F	10	11	12	13	14	15	16	17
9	A	B	C	D	E	F	10	11	12	13	14	15	16	17	18
A	B	C	D	E	F	10	11	12	13	14	15	16	17	18	19
B	C	D	E	F	10	11	12	13	14	15	16	17	18	19	1A
C	D	E	F	10	11	12	13	14	15	16	17	18	19	1A	1B
D	E	F	10	11	12	13	14	15	16	17	18	19	1A	1B	1C
E	F	10	11	12	13	14	15	16	17	18	19	1A	1B	1C	1D
F	10	11	12	13	14	15	16	17	18	19	1A	1B	1C	1D	1E

Рис. 5.11

×	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
2	2	4	6	8	A	C	E	10	12	14	16	18	1A	1C	1E
3	3	6	9	C	F	12	15	18	1B	1E	21	24	27	2A	2D
4	4	8	C	10	14	18	1C	20	24	28	2C	30	34	38	3C
5	5	A	F	14	19	1E	23	28	2D	32	37	3C	41	46	4B
6	6	C	12	18	1E	24	2A	30	36	3C	42	48	4E	54	5A
7	7	E	15	1C	23	2A	31	38	3F	46	4D	54	5B	62	69
8	8	10	18	20	28	30	38	40	48	50	58	60	68	70	78
9	9	12	1B	24	2D	36	3F	48	51	5A	63	6C	75	7E	87
A	A	14	1E	28	32	3C	46	50	5A	64	6E	78	82	8C	96
B	B	16	21	2C	37	42	4D	58	63	6E	79	84	8F	9A	A5
C	C	18	24	30	3C	48	54	60	6C	78	84	90	9C	A8	B4
D	D	1A	27	34	41	4E	5B	68	75	82	8F	9C	A9	B6	C3
E	E	1C	2A	38	46	54	62	70	7E	8C	9A	A8	B6	C4	D2
F	F	2E	2D	3C	4B	5A	69	78	87	96	A5	B4	C3	D2	E1

Рис. 5.12

4. Преобразуйте следующие десятичные числа в двоичные:

а) 53; **б)** 99;

в) 287; **г)** 186;

д) 334.

5. Преобразуйте следующие двоичные числа в десятичные:

а) 1011; **б)** 11010;

в) 100100; **г)** 111011;

д) 110101.

6. Преобразуйте следующие двоичные числа в десятичные:
а) 1101; **б)** 10110;
в) 101101; **г)** 110101;
д) 1010101.

7. Преобразуйте следующие десятичные числа в шестнадцатеричные:
а) 63; **б)** 116;
в) 376; **г)** 734;
д) 5246.

8. Преобразуйте следующие десятичные числа в шестнадцатеричные:
а) 183; **б)** 297;
в) 458; **г)** 861;
д) 36304.

9. Преобразуйте следующие шестнадцатеричные числа в десятичные:
а) $4E$; **б)** $2F3$;
в) $4FA$; **г)** $41FD$;
д) $D1E3$.

10. Преобразуйте следующие шестнадцатеричные числа в десятичные:
а) AF; **б)** $3A4$;
в) $AB2$; **г)** $D13A$;
д) $21F2D$.

11. Преобразуйте следующие десятичные числа в двоичные:
а) 23.875; **б)** 43.75;
в) 86.125; **г)** 94.4;
д) 52.3.

12. Преобразуйте следующие десятичные числа в двоичные:
а) 3.1875; **б)** 22.625;
в) 87.71875; **г)** 16.53125;
д) 34.28125.

13. Преобразуйте следующие двоичные числа в десятичные:
а) 1110.11; **б)** 1010.101;
в) 100.1001; **г)** 11.0111;
д) 1101.0101.

14. Преобразуйте следующие двоичные числа в десятичные:
а) 101.01011; **б)** 11100.1010;
в) 1011.01110; **г)** 1101.1011;
д) 10101.01111.

15. Преобразуйте следующие десятичные числа в шестнадцатеричные:
а) 36.625; **б)** 161.78125;
в) 592.15625; **г)** 1876.54625;
д) 5246.34765625.

16. Преобразуйте следующие десятичные числа в шестнадцатеричные:
а) 273.9375; **б)** 349.796875;

в) 576.640625; **г)** 1997.35;
д) 36304.140625.

17. Преобразуйте следующие шестнадцатеричные числа в десятичные:
а) $5C.3A$; **б)** $F12.AE$;
в) $4AF.C48$; **г)** $32AB.C11$;
д) $13CE.D24$.

18. Преобразуйте следующие шестнадцатеричные числа в десятичные:
а) $14.D2$; **б)** $11F.E33$;
в) $47F.344$; **г)** $237F.755$;
д) $233B.FDA$.

19. Преобразуйте следующие двоичные числа в шестнадцатеричные:
а) 101.1010; **б)** 110111.010101;
в) 110110110.011011111; **г)** 1110101110.0011011011;
д) 1111010100.001000101.

20. Преобразуйте следующие двоичные числа в шестнадцатеричные:
а) 110.11001; **б)** 100101.011110;
в) 101110101.010101011; **г)** 11010101010.11001101101;
д) 10001010111.1101001010101.

21. Преобразуйте следующие шестнадцатеричные числа в двоичные:
а) $2C.4B$; **б)** $1F2.A3C$;
в) $AF8.48C$; **г)** $A12E.11C$;
д) $C14A.2D4$.

22. Преобразуйте следующие шестнадцатеричные числа в двоичные:
а) $CE.2D5$; **б)** $F1A.E3D$;
в) $B7E.3C4$; **г)** $FA37.2E5$;
д) $C46B.AC4$.

5.7. ЧИСЛА СО ЗНАКОМ

В предыдущем разделе мы рассматривали представления в двоичной и шестнадцатеричной форме только для положительных рациональных чисел. В данном разделе мы покажем, как строить двоичные и шестнадцатеричные представления положительных и отрицательных целых чисел при помощи ***метода двоичного дополнения***. Этот метод требует, чтобы целые числа были записаны с использованием фиксированного количества битов (двоичных разрядов). Поэтому число в двоичной системе будет всегда записываться с использованием одного и того же количества разрядов. Если используется 8 двоичных разрядов, то двоичное представление числа 3 будет иметь вид 00000011. Для представления целого положительного числа с помощью 8 бит перед ним при необходимости дописывается нужное количество нулей. До конца этого раздела будем использовать 8 бит для чисел в двоичной форме и 16 бит для чисел в шестнадцатеричной форме. Сначала рассмотрим случай использования 8 бит.

Целое положительное число должно начинаться с 0, а целое отрицательное число будет начинаться с 1. Поэтому 10110011 будет представлять целое отрицательное число, а 01110101 будет представлять целое положительное число. Поскольку при таком способе записи наибольшее возможное целое положительное число есть 01111111, то всего имеется 127 целых положительных чисел. Пусть дано положительное целое число, представленное с использованием n бит. Тогда соответствующее ему отрицательное число находят путем вычитания этого числа из 2^n, выраженного в двоичной форме. Это осуществляется вычитанием числа из $2^n - 1$ (что называется нахождением 1-го дополнения) и последующим прибавлением 1. Поэтому в случае, когда используется 8 бит, положительное целое число вычитается из

$$2^n - 1 = 11111111,$$

а затем прибавляется 1. Например, чтобы найти отрицание числа 00101011, мы вычитаем его из 11111111:

$$\begin{array}{r} 11111111 \\ -\ 00101011 \\ \hline 11010100 \end{array},$$

а затем к полученному числу 11010100 прибавляем 1. Таким образом, отрицанием числа 00101011 будет число 11010101.

Заметим, что вычитание 00101011 из 11111111 просто меняет нули на единицы, а единицы — на нули. Всегда ли это справедливо, когда 8-битовая двоичная строка вычитается из 11111111?

ПРИМЕР 5.46. Найти представление числа -56 в 8-битовом двоичном дополнении. Сначала число 56 представляется как 00111000. Вычитая его из 11111111, получаем 11000111. Прибавление 1 дает 11001000, так что -56 имеет представление 11001000. □

Если задано целое отрицательное число, можно найти целое положительное число, для которого данное число является отрицанием, просто выполняя те же самые операции с отрицательным числом. Например, чтобы найти то положительное число, для которого 10110111 представляет отрицание, просто вычтем его из 11111111, получая 01001000. После прибавления 1 получаем 01001001, что является двоичным представлением числа 72. Отсюда 10110111 изображает -72. Отрицательные двоичные дополнения изменяются от 11111111 до 10000000, так что их имеется 128. Таким образом, если использовать 8 двоичных разрядов, можно представить все числа от -128 до 127. В общем случае, если имеется n двоичных разрядов, можно представить все целые числа от -2^{n-1} до $2^{n-1} - 1$.

Если для представления целых чисел используется 16 (или больше) бит, то удобнее всего пользоваться шестнадцатеричной системой счисления. Поскольку каждый шестнадцатеричный разряд состоит из четырех двоичных, то для 16 бит необходимо 4 шестнадцатеричных разряда. Таким образом, $21A5$, $F9AB$ и $004F$ являются примерами шестнадцатиразрядных целых чисел со знаком, представленных в шестнадцатеричной системе счисления. Поскольку двоичное представление целого положительного числа должно начинаться с 0, то наибольшее число, которое может стоять в первом разряде его шестнадцатеричного представления, это 7

($= 0111_2$). Поэтому $691F$, $7FFF$ и 0031 изображают целые положительные числа, а 8124, $A105$ и $FFF1$ изображают целые отрицательные числа. Как и ранее, если целое положительное число представлено с использованием n битов, отрицание числа находят путем его вычитания из 2^n. В случае использования 16 битов это осуществляется путем вычитания целого положительного числа в шестнадцатеричной форме из $FFFF$ и последующего прибавления 1. Как и раньше, для преобразования отрицательного числа обратно в его положительного двойника выполняем операцию дополнения, вычитая его из

$$2^{16} - 1 = FFFF,$$

а затем прибавляя 1.

ПРИМЕР 5.47. Найти представление в шестнадцатеричной системе счисления отрицания шестнадцатеричного числа $78F$ с использованием 16 бит. Сначала заменим запись числа на 16-битовое шестнадцатеричное число $078F$. Затем вычтем $078F$ из $FFFF$, получив $F870$, и далее прибавляем 1. В итоге получаем представление $F871$. □

ПРИМЕР 5.48. Найти 16-битовое представление в шестнадцатеричной форме числа -1158_{10}. Поскольку $1158_{10} = 486_{16} = 0486$, представленное в 16-битовой шестнадцатеричной форме, вычитаем 0486 из $FFFF$, получая $FB79$, после чего, прибавляя 1, получаем $FB7A$. Следовательно, 16-битовое шестнадцатеричное представление числа -1158_{10} есть $FB7A$. □

Сложение чисел фиксированной длины со знаком на основе двоичного или шестнадцатеричного представления ничем не отличается от обычного сложения двоичных и шестнадцатеричных чисел, за исключением того, что если сумма порождает дополнительный разряд, тогда первый разряд удаляется. Например, если 0111 и 1011 складываются как двоичные числа:

$$\begin{array}{r} 0111 \\ \underline{1011}\ , \\ [1]\,0010 \end{array}$$

то цифра 1 слева удаляется. Если шестнадцатеричные числа со знаком $FF14$ и $0F1A$ складываются как шестнадцатеричные:

$$\begin{array}{r} FF14 \\ \underline{0F1A}\ , \\ [1]\,0E2E \end{array}$$

то цифра 1 слева также удаляется.

Если попытаться сложить $7F14$ и $6E23$, то получаем

$$\begin{array}{r} 7F14 \\ \underline{6E23}\ , \\ ED37 \end{array}$$

т.е. сложив два положительных числа, получили отрицательное число. В таком случае говорят, что имеет место переполнение. Значение суммы слишком велико

для представления с использованием 16 бит. Числа со знаком, выражаемые с использованием 16 двоичных разрядов или 4 шестнадцатеричных разрядов, должны находиться между -32768 и 32767.

В десятичной арифметике, чтобы вычесть 25 из 75, к 75 прибавляют -25. Во многих учебниках пишут: "измените знак вычитаемого [того числа, которое вычитается] и выполните сложение". При сложении чисел со знаком необходимо действовать таким же образом. При вычитании одного целого числа из другого находим двоичное дополнение числа, которое нужно вычесть, и производим сложение.

Например, для нахождения разности

$$\begin{array}{r} 01010110 \\ -\ \underline{01110001} \end{array}$$

мы сначала находим двоичное дополнение числа 01110001, которое равно 10001111, после чего выполняем сложение:

$$\begin{array}{r} 01010110 \\ +\ \underline{10001111} \\ 11100101 \end{array}.$$

Как и следовало ожидать, результат — отрицательное число.

Используя тот же самый подход при решении следующей задачи:

$$\begin{array}{r} 01010001 \\ -\ \underline{11101001} \end{array},$$

мы сначала находим двоичное дополнение числа 11101001, которое равно 00010111, после чего выполняем сложение:

$$\begin{array}{r} 01010001 \\ +\ \underline{00010111} \\ 01101000 \end{array}.$$

В данном случае отрицательное число вычиталось из положительного, поэтому отрицательное число мы заменили положительным и произвели сложение.

ПРИМЕР 5.49. Вычесть 56 из 23, используя 8-битовые числа со знаком. Поскольку $56 = 00111000$ и $23 = 00010111$, имеем

$$\begin{array}{r} 00010111 \\ -\ \underline{00111000} \end{array}.$$

Сначала находим двоичное дополнение числа 00111000, которое равно 11001000, а затем выполняем сложение:

$$\begin{array}{r} 00010111 \\ +\ \underline{11001000} \\ 11011111 \end{array}.$$

Так как двоичным дополнением числа 11011111 является 00100001, которое равно 33, число 11011111 равно -33. □

ПРИМЕР 5.50. Вычесть 7328 из 3614, воспользовавшись 16-битовыми шестнадцатеричными числами со знаком. Поскольку $7328_{10} = 1CA0_{16}$ и $3614 = E1E_{16}$, имеем

$$\begin{array}{r} 0E1E \\ -\ \underline{1CA0}\ . \end{array}$$

Сначала находим двоичное дополнение числа $1CA0$, которое равно $E35F$, а затем выполняем сложение:

$$\begin{array}{r} 0E1E \\ +\ \underline{E360}\ . \\ F17E \end{array}$$

Поскольку двоичным дополнением числа $F1E$ является $0E82$ и $0E82 = 3714_{10}$, то $3614 - 7328 = -3714$. □

■ УПРАЖНЕНИЯ

1. Найдите двоичное дополнение приведенных ниже 8-битовых чисел в двоичной форме:

а) 01110110; **б)** 10110101;

в) 00110111; **г)** 11101011.

2. Найдите двоичное дополнение приведенных ниже 8-битовых чисел в двоичной форме:

а) 01100100; **б)** 11011001;

в) 00111011; **г)** 10010000.

3. Выразите следующие десятичные числа как 8-битовые двоичные числа со знаком:

а) 73; **б)** -101;

в) -37; **г)** -14.

4. Выразите следующие десятичные числа как 8-битовые двоичные числа со знаком:

а) 48; **б)** -33;

в) -58; **г)** -114.

5. Выполните следующие операции, используя 8-битовые двоичные числа со знаком:

а) $\begin{array}{r} 10110100 \\ +\ \underline{01110010}\ ; \end{array}$ **б)** $\begin{array}{r} 00110100 \\ +\ \underline{01001010}\ ; \end{array}$

в) $\begin{array}{r} 00110101 \\ -\ \underline{01110110}\ ; \end{array}$ **г)** $\begin{array}{r} 00110101 \\ -\ \underline{01110110}\ . \end{array}$

6. Выполните следующие операции, используя 8-битовые двоичные числа со знаком:

а) $\begin{array}{r} 11001001 \\ +\ \underline{01011110}\ ; \end{array}$ **б)** $\begin{array}{r} 00101001 \\ +\ \underline{01010011}\ ; \end{array}$

в) $\begin{array}{r} 01011100 \\ -\ \underline{01110111}\ ; \end{array}$ **г)** $\begin{array}{r} 01101101 \\ -\ \underline{01111000}\ . \end{array}$

7. Превратите следующие числа в 8-битовые двоичные числа со знаком и выполните указанные операции:

а) $\begin{array}{r} 43 \\ +\,\underline{28} \end{array}$; **б)** $\begin{array}{r} 12 \\ +\,\underline{81} \end{array}$;

в) $\begin{array}{r} 43 \\ -\,\underline{67} \end{array}$; **г)** $\begin{array}{r} 86 \\ -\,\underline{73} \end{array}$.

8. Превратите следующие числа в 8-битовые двоичные числа со знаком и выполните указанные операции:

а) $\begin{array}{r} 29 \\ +\,\underline{61} \end{array}$; **б)** $\begin{array}{r} 34 \\ -\,\underline{56} \end{array}$;

в) $\begin{array}{r} 13 \\ -\,\underline{87} \end{array}$; **г)** $\begin{array}{r} 22 \\ -\,\underline{98} \end{array}$.

9. Найдите двоичное дополнение следующих чисел со знаком, выраженных в шестнадцатеричной форме:

а) $723A$; **б)** $AB20$;

в) $FA12$; **г)** $23BC$.

10. Найдите двоичное дополнение приведенных ниже чисел со знаком, выраженных в шестнадцатеричной форме:

а) $23CF$; **б)** $1B5C$;

в) $FF13$; **г)** 1020.

11. Выразите следующие десятичные числа как 16-битовые числа со знаком в шестнадцатеричной форме:

а) 270; **б)** -893;

в) -17; **г)** -9872.

12. Выразите следующие десятичные числа как 8-битовые двоичные числа со знаком:

а) 48; **б)** -237;

в) -8858; **г)** -12499.

13. Выполните следующие операции, используя 16-битовые числа со знаком в шестнадцатеричной форме:

а) $\begin{array}{r} 3C14 \\ +\,\underline{89AC} \end{array}$; **б)** $\begin{array}{r} C12F \\ +\,\underline{359A} \end{array}$;

в) $\begin{array}{r} B127 \\ -\,\underline{C1AB} \end{array}$; **г)** $\begin{array}{r} 34C5 \\ -\,\underline{FF12} \end{array}$.

14. Выполните следующие операции, используя 16-битовые числа со знаком в шестнадцатеричной форме:

а) $\begin{array}{r} 4713 \\ +\,\underline{0F23} \end{array}$; **б)** $\begin{array}{r} 5379 \\ +\,\underline{FA12} \end{array}$;

в) $\begin{array}{r} 37B4 \\ -\ \underline{AA44} \end{array}$; **г)** $\begin{array}{r} C129 \\ -\ \underline{D48A} \end{array}$.

15. Переведите следующие числа в 16-битовые шестнадцатеричные числа со знаком и выполните указанные операции:

а) $\begin{array}{r} 936 \\ +\ \underline{258} \end{array}$; **б)** $\begin{array}{r} 1639 \\ +\ \underline{3926} \end{array}$;

в) $\begin{array}{r} 8836 \\ -\ \underline{19923} \end{array}$; **г)** $\begin{array}{r} 994 \\ -\ \underline{1713} \end{array}$.

16. Переведите следующие числа в 16-битовые шестнадцатеричные числа со знаком и выполните указанные операции:

а) $\begin{array}{r} 829 \\ +\ \underline{1499} \end{array}$; **б)** $\begin{array}{r} 778 \\ -\ \underline{569} \end{array}$;

в) $\begin{array}{r} 13 \\ -\ \underline{87} \end{array}$; **г)** $\begin{array}{r} 19237 \\ -\ \underline{\quad 18} \end{array}$.

17. Найдите диапазон целых чисел, который может быть выражен числами со знаком с использованием 32 бит.

18. Найдите диапазон целых чисел, который может быть выражен числами со знаком с использованием 64 бит.

5.8. ДАЛЬНЕЙШЕЕ ИЗУЧЕНИЕ МАТРИЦ

Детерминант является важной функцией, отображающей множество матриц с элементами из множества действительных чисел во множество действительных чисел. Детерминант используется при решении систем уравнений (метод Крамера), а также при определении обратной матрицы (если она существует).

Пусть дана $n \times n$ матрица

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}.$$

Символом $\overline{A_{ij}}$ обозначим матрицу размера $(n-1)\times(n-1)$, полученную удалением в исходной матрице i-ой строки и j-го столбца.

ПРИМЕР 5.51. Пусть

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Тогда $\overline{A_{11}} = [A_{22}]$, $\overline{A_{12}} = [A_{21}]$, $\overline{A_{21}} = [A_{12}]$ и $\overline{A_{22}} = [A_{11}]$. □

ПРИМЕР 5.52. Пусть

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}.$$

Тогда

$$\overline{A_{11}} = \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix}, \quad \overline{A_{22}} = \begin{bmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{bmatrix}, \text{ и } \overline{A_{31}} = \begin{bmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{bmatrix}.$$

□

ОПРЕДЕЛЕНИЕ 5.53. ***Детерминант*** матрицы размера $n \times n$

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$

обозначается через $\det(A)$, $|A|$ или

$$\begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{vmatrix}$$

и определяется следующим образом:

а) если $n = 1$, то $\det(A) = |A_{11}| = A_{11}$;

б) если $n \geq 2$, то $\det(A) = \sum_{j=1}^{n} A_{1j}(-1)^{1+j} \det(\overline{A_{1j}})$.

Пусть $M_{ij} = \det(\overline{A_{ij}})$. Число M_{ij} называется ***минором*** элемента A_{ij}, а число $C_{ij} = (-1)^{i+j} \det(\overline{A_{ij}})$ — ***алгебраическим дополнением*** элемента A_{ij}. Следовательно, $\det(A) = \sum_{j=1}^{n} A_{1j}(-1)^{1+j} \det(\overline{A_{1j}}) = \sum_{j=1}^{n} A_{1j}C_{1j}$.

Обратите внимание, что в приведенном выше определении $|A_{11}|$ не является абсолютным значение A_{11}, хотя обозначение кажется таким же.

Давайте теперь составим алгоритм для вычисления детерминанта. Прежде всего отметим, что детерминант порядка $n \times n$ определен через детерминанты порядка $(n-1) \times (n-1)$. Здесь мы сталкиваемся с примером рекурсивного алгоритма. В дальнейшем мы рассмотрим алгоритмы такого типа более подробно. Процедура вычисления детерминанта имеет следующий вид:

Процедура **Детерминант**(A, n):
  Цикл по i от 1 до n:
    Если $n = 1$, то $\det(A) = |A_{11}| = A_{11}$;
    Если $n \geq 2$, то $\det(A) = 0$;
    Цикл по j от 1 до n:
      Значение $\det(A)$ заменить на
      $\det(A) + A_{1j}(-1)^{1+j}$Детерминант$(\overline{A_{1j}}, n-1)$;
    Конец цикла;
  Конец цикла;
Конец процедуры.

Поскольку детерминант (A, n) определен через детерминанты $(\overline{A_{1j}}, n-1)$, для вычисления детерминанта (A, n) прежде необходимо определить детерминанты $(\overline{A_{1j}}, n-1)$. Многие компьютерные программы обладают свойством "вызывать" саму себя. Это дает возможность написать компьютерную программу, очень похожую на приведенную выше процедуру. В более полном объеме этот вопрос будет рассмотрен в последующих разделах.

Следовательно, если

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

то $\det(A) = A_{11} \cdot |A_{22}| - A_{12} \cdot |A_{21}| = A_{11} \cdot A_{22} - A_{12} \cdot A_{21}$.

Если

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix},$$

то

$$\det(A) = A_{11} \cdot \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{12} \cdot \begin{vmatrix} A_{22} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} + A_{13} \cdot \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix}.$$

ПРИМЕР 5.54. Пусть $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$. Тогда

$$\det(A) = \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 2\,|2| + 3(-\,|1|) = 2 \times 2 - 3 \times 1 = 1.$$

В общем случае, если $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, то $\det(A) = ad - bc$. □

ПРИМЕР 5.55. Пусть $A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix}$. Тогда

$$\det(A) = \begin{vmatrix} 1 & 2 & 0 \\ 1 & 1 & 2 \\ 3 & 0 & 1 \end{vmatrix} =$$

$$= 1\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} + 0\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} =$$
$$= 1(1 \times 1 - 2 \times 0) - 2(1 \times 1 - 2 \times 3) + 0 =$$
$$= 1 - 2(-5) = 11.$$

□

Определим δ_{ij} как

$$\delta_{ij} = \begin{cases} 1, & \text{если } i = j; \\ 0, & \text{если } i \neq j. \end{cases}$$

ОПРЕДЕЛЕНИЕ 5.56. Пусть I_n есть матрица $[I_{ij}]$ размера $n \times n$, где $I_{ij} = \delta_{ij}$, т.е. на диагонали матрицы I_n стоят единицы, а все остальные ее элементы равны нулю. Матрица I_n называется ***единичной матрицей*** размера $n \times n$, или просто ***единичной матрицей***. Единичная матрица обладает тем свойством, что для любой матрицы A размера $n \times n$ имеет место равенство $AI_n = I_nA = A$.

ПРИМЕР 5.57. Единичными матрицами размера 2×2 и 3×3 являются

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{и} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

□

ОПРЕДЕЛЕНИЕ 5.58. Пусть A — матрица размера $n \times n$. ***Мультипликативно обратной***, или просто ***обратной*** к матрице A называется матрица A^{-1}, обладающая (если она существует) тем свойством, что $AA^{-1} = A^{-1}A = I_n$.

ПРИМЕР 5.59. Матрица $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ является обратной к $\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$, т.к.

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

□

ПРИМЕР 5.60. Матрица $\begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix}$ является обратной к $\begin{bmatrix} 0 & -1 & 2 \\ 1 & 2 & -4 \\ -1 & -2 & 5 \end{bmatrix}$, поскольку

$$\begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & -1 & 2 \\ 1 & 2 & -4 \\ -1 & -2 & 5 \end{bmatrix} =$$
$$= \begin{bmatrix} 0 & -1 & 2 \\ 1 & 2 & -4 \\ -1 & -2 & 5 \end{bmatrix}\begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

□

Ниже приводится алгоритм нахождения матрицы, обратной к матрице A размера $n \times n$:

1. Найти $\det(A)$. Если он равен нулю, обратная матрица не существует.
2. Для каждого A_{ij} найти алгебраическое дополнение C_{ij} и сформировать матрицу $C = [C_{ij}]$, называемую ***матрицей алгебраических дополнений***.
3. Найти матрицу C^t, транспонированную к матрице C, так что $C^t_{ij} = C_{ji}$. Матрица C^t называется ***сопряженной*** к матрице A и обозначается $\text{adj}(A)$.
4. Разделить каждый элемент сопряженной матрицы $\text{adj}(A)$ на $\det(A)$.
5. $A^{-1} = \dfrac{1}{\det(A)} \text{adj}(A)$.

Следовательно, если $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, то

$$B^{-1} = \begin{bmatrix} \dfrac{d}{\det(B)} & \dfrac{-b}{\det(B)} \\ \dfrac{-c}{\det(B)} & \dfrac{a}{\det(B)} \end{bmatrix},$$

где $\det(B) = ad - bc$. Путем прямых вычислений можно показать, что $BB^{-1} = B^{-1}B = I_2$. Аналогично, для нахождения матрицы, обратной к матрице A, где

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix},$$

определяем матрицу алгебраических дополнений $C = [C_{ij}]$, где C_{ij} — детерминанты матриц, полученных из $A = [A_{ij}]$ удалением строки и столбца, содержащих A_{ij}, умноженные на соответствующий знак, а именно, $(-1)^{i+j}$. Наконец, для получения A^{-1} матрица C транспонируется и делится на $\det(A)$, после чего

$$A^{-1} = \frac{1}{\det(A)} \cdot \begin{bmatrix} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} & \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} \\ -\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} & \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix} \\ \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} & \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \end{bmatrix}.$$

ПРИМЕР 5.61. Пусть $A = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$. Тогда $\det(A) = 1$. Матрицей алгебраических дополнений для матрицы A будет

$$C = \begin{bmatrix} 2 & -(-1) \\ -(-5) & 3 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}.$$

Матрицей, сопряженной к матрице A, будет

$$\mathrm{adj}(A) = C^t = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}.$$

Обратной к матрице A будет матрица

$$A^{-1} = \frac{1}{1}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}.$$

□

ПРИМЕР 5.62. Пусть $A = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 2 & -4 \\ -1 & -2 & 5 \end{bmatrix}$. Тогда

$$\det(A) = 0 \cdot \begin{vmatrix} 2 & -4 \\ -2 & 5 \end{vmatrix} - (-1) \cdot \begin{vmatrix} 1 & -4 \\ -1 & 5 \end{vmatrix} + 2 \cdot \begin{vmatrix} 1 & 2 \\ -1 & -2 \end{vmatrix} =$$
$$= 0 + 1 + 0 = 1.$$

Матрицей алгебраических дополнений для матрицы A будет

$$C = \begin{bmatrix} \begin{vmatrix} 2 & -4 \\ -2 & 5 \end{vmatrix} & -\begin{vmatrix} 1 & -4 \\ -1 & 5 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ -1 & -2 \end{vmatrix} \\ -\begin{vmatrix} -1 & 2 \\ -2 & 5 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ -1 & 5 \end{vmatrix} & -\begin{vmatrix} 0 & -1 \\ -1 & -2 \end{vmatrix} \\ \begin{vmatrix} -1 & 2 \\ 2 & -4 \end{vmatrix} & -\begin{vmatrix} 0 & 2 \\ 1 & -4 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 2 \end{vmatrix} \end{bmatrix} =$$

$$= \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}.$$

Таким образом, матрицей, сопряженной к A, будет

$$\mathrm{adj}(A) = C^t = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix},$$

а обратной

$$A^{-1} = \frac{1}{1}\begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix}.$$

□

■ УПРАЖНЕНИЯ

1. Вычислите детерминанты матриц:

а) $\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}$;

б) $\begin{bmatrix} 3 & 1 \\ -4 & 6 \end{bmatrix}$;

в) $\begin{bmatrix} 3 & -2 & -3 \\ -6 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix}$;

г) $\begin{bmatrix} 6 & -4 & -5 \\ 1 & -11 & 2 \\ 3 & -1 & 1 \end{bmatrix}$.

2. Вычислите детерминанты матриц:

а) $\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}$;

б) $\begin{bmatrix} 3 & 1 \\ -4 & 6 \end{bmatrix}$;

в) $\begin{bmatrix} 3 & -2 & -1 \\ -6 & 1 & 2 \\ 9 & 0 & -3 \end{bmatrix}$;

г) $\begin{bmatrix} 1 & -6 & 3 \\ -6 & -1 & 2 \\ 3 & 2 & 1 \end{bmatrix}$.

3. Вычислите детерминант матрицы

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

4. Найдите A^{-1}, если $A =$

а) $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}$;

б) $\begin{bmatrix} 3 & -5 \\ -4 & 7 \end{bmatrix}$;

в) $\begin{bmatrix} 4 & -1 & 3 \\ -3 & 1 & -3 \\ 2 & 0 & 1 \end{bmatrix}$;

г) $\begin{bmatrix} 7 & -7 & 3 \\ 4 & -3 & 2 \\ 2 & -2 & 1 \end{bmatrix}$.

5. Понятие матрицы перестановок было дано в определении 4.52. Докажите, что если A — матрица перестановок, то $A^{-1} = A^t$.

6. В предположении, что если A и B — матрицы, то $\det(AB) = \det(A)\det(B)$,

а) докажите, что детерминант единичной матрицы равен 1;

б) докажите, что если матрица A имеет обратную матрицу, то

$$\det(A^{-1}) = \frac{1}{\det(A)};$$

в) докажите, что матрица A имеет обратную тогда и только тогда, когда $\det(A) \neq 0$.

ГЛАВА

6

ГРАФЫ, ОРИЕНТИРОВАННЫЕ ГРАФЫ И ДЕРЕВЬЯ

6.1. ГРАФЫ

В главе 2 графы использовались для представления отношений. Однако теория графов имеет свою собственную проблематику и является важным разделом математики. В данной главе мы определим понятие графа непосредственно, вне связи с отношениями.

ОПРЕДЕЛЕНИЕ 6.1. ***Граф*** есть конечное множество V, называемое ***множеством вершин***, и множество E двухэлементных подмножеств множества V. Множество E называется ***множеством ребер***. Элемент множества E называется ***ребром***. Граф обозначается $G(V, E)$. Элементы a и b множества V называются ***соединенными***, или ***связанными*** ребром $\{a, b\}$, если $\{a, b\} \in E$.

Обычно конечный граф изображают в виде диаграммы, на которой вершины обозначаются точками, а ребра, соединяющие две вершины,— линиями между этими точками.

ОПРЕДЕЛЕНИЕ 6.2. Если $\{a, b\}$ — ребро, тогда вершины a и b называются ***концами*** ребра $\{a, b\}$. Ребро $\{a, b\}$ называют также ***инцидентным*** к вершинам a и b. Обратно, говорят, что вершины a и b инцидентны к ребру $\{a, b\}$. Две вершины называются ***смежными***, если они являются концами ребра, или, что то же самое, если они инцидентны к одному ребру. Два ребра называются ***смежными***, если они инцидентны к общей вершине.

ПРИМЕР 6.3. Граф с множеством вершин $V = \{a, b, c\}$ и множеством ребер $E = \{\{a, b\}, \{b, c\}\}$ может быть изображен, как показано на рис. 6.1 или рис. 6.2.

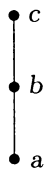


Рис. 6.1

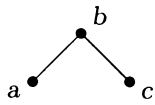


Рис. 6.2

□

ПРИМЕР 6.4. Граф, у которого $V = \{a, b, c, d, e\}$ и

$$E = \{\{a, b\}, \{a, e\}, \{b, e\}, \{b, d\}, \{b, c\}, \{c, d\}\},$$

может быть изображен диаграммой, показанной на рис. 6.3.

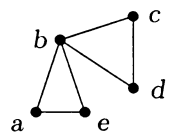


Рис. 6.3

□

Во многих книгах по математике графы, определенные выше, называют простыми графами. Тем не менее, в теории графов чаще всего используются структуры именно такого типа. Ограничение на существование только одного ребра между двумя вершинами дает возможность представлять любое ребро как множество из двух элементов, а именно, вершин ребра. Перечислим теперь основные типы графов, которые будут использоваться в других главах. Заметим, что многие теоремы этой главы будут справедливы для графов более общего вида. Ребро, которое соединяет вершину саму с собой, называется ***петлей***. Если в графе допускается наличие петель, то он называется ***графом с петлями***. Если в графе допускается наличие более одного ребра между двумя вершинами, то он называется ***мультиграфом***. Если каждая вершина графа отмечена, то граф называется ***размеченным***. Если же допускается как наличие петель, так и существование более одного ребра между двумя вершинами, то мы будем называть такой объект ***псевдографом***.

Теория графов берет начало с решения знаменитым математиком Эйлером задачи о кенигсбергских мостах в 1736 году. Существуют две внешне различные версии этой задачи. Задача возникла в прусском городке Кенигсберг на реке Прегал. Жителям Кенигсберга нравилось гулять по дорожке, которая включала семь мостов через реку Прегал. Люди интересовались, могут ли они, начав путь с одного участка суши, обойти все мосты, посетив каждый лишь однажды, и вернуться в точку начала пути, не переплывая реки. Различие между двумя версиями задачи — чисто географическое. Согласно первой версии на реке Прегал был остров Кнайпхоф и участок суши, возникший в результате разветвления реки Прегал на рукава. Этот район и семь мостов показаны на рис. 6.4.

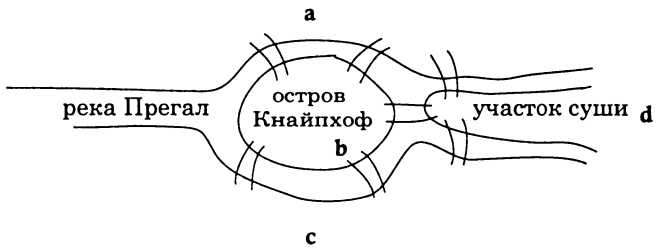


Рис. 6.4

Согласно второй версии на реке Прегал было два острова, один из которых — остров Кнайпхоф. Эта ситуация изображена на рис. 6.5.

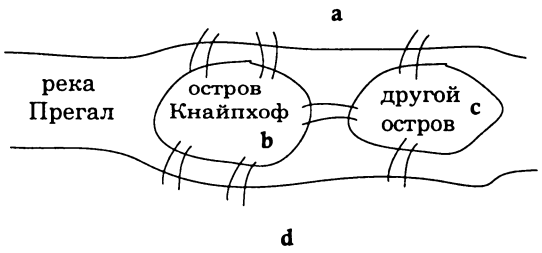


Рис. 6.5

Точная версия авторам не известна. К тому же это не имеет значения для математической формулировки задачи. Эйлер построил мультиграф, показанный на рис. 6.6, который пригоден для каждого из этих случаев. В этом мультиграфе участки суши Эйлер изобразил как вершины, а дорожки через мосты — как ребра. В таком случае задача приобретает следующую формулировку: начиная с любой вершины, проходя по каждому ребру только один раз, вернуться в исходную вершину. Далее будет показано, как на основе понятия мультиграфа Эйлер решил эту задачу. Приведенный пример иллюстрирует полезность математических абстракций, по крайней мере, в двух аспектах. Во-первых, абстракция устраняет в содержании проблемы несущественные детали и концентрирует внимание на тех понятиях, которые действительно важны. Во-вторых, абстракция часто выявляет кажущееся различие задач, совпадающих по сути или, по крайней мере, имеющих сходные решения. Сказанное особо справедливо для теории графов.

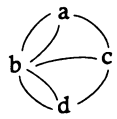


Рис. 6.6

Решение задачи о кенигсбергских мостах будет рассмотрено в разделе 6.2.

Примером другой проблемы, которую можно промоделировать на основе теории графов, является задача о снабжении трех домов тремя видами коммунальных услуг. Согласно условию задачи к каждому из трех домов необходимо подключить три вида коммунальных услуг, например, воду, газ и электричество, посредством подземных линий труб и кабелей. Вопрос состоит в том, можно ли обеспечить эти три дома коммунальными услугами без пересечения линий снабжения. Граф, моделирующий данную задачу, показан на рис. 6.7

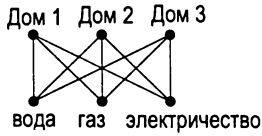


Рис. 6.7

Аналогичная задача более практического характера возникает при создании микросхем. В них пересечение проводов на каждом уровне является нежелательным. Такого рода задачи будут рассмотрены в одной из последующих глав при изучении планарных графов.

ОПРЕДЕЛЕНИЕ 6.5. ***Степенью*** вершины v, обозначается $\deg(v)$, называется количество ребер, инцидентных этой вершине. Вершина степени 0 называется ***изолированной***.

ПРИМЕР 6.6. В графе, показанном на рис. 6.8, вершины a и c — смежные, а e_1, e_2 и e_3 — смежные ребра.

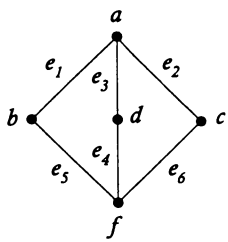


Рис. 6.8

Однако, вершины a и f смежными не являются, а e_2 и e_5 не являются смежными ребрами. Вершины b, c и d имеют степень 2, в то время как вершины a и f имеют степень 3. □

ТЕОРЕМА 6.7. Сумма степеней вершин графа всегда четная.

ДОКАЗАТЕЛЬСТВО. Поскольку каждое ребро графа имеет два конца, степень каждого конца увеличивается на 1 за счет одного ребра. Таким образом, в сумму степеней всех вершин каждое ребро вносит 2 единицы, поэтому сумма должна в два раза превышать число ребер. Следовательно, сумма является четным числом. ■

Теперь воспользуемся этой теоремой для доказательства следующего утверждения.

ТЕОРЕМА 6.8. В любом графе количество вершин нечетной степени четно.

ДОКАЗАТЕЛЬСТВО. Доказательство проведем методом от противного: предположив, что теорема не верна, найдем противоречие, из которого будет следовать, что теорема справедлива. Если теорема не верна, то имеется нечетное количество вершин, степени которых нечетны. Но сумма степеней вершин с четными степенями четна. Сумма степеней всех вершин есть сумма степеней вершин с нечетными

степенями плюс сумма степеней вершин с четными степенями. Поскольку сумма нечетного числа и четного числа есть число нечетное, сумма степеней всех вершин нечетная. Но это противоречит теореме 6.7, поэтому мы пришли к противоречию. Следовательно, делаем вывод, что теорема справедлива. ■

ОПРЕДЕЛЕНИЕ 6.9. Граф $G'(V', E')$ называется ***подграфом*** графа $G(V, E)$, обозначается $G'(V', E') \preceq G(V, E)$, если $V' \subseteq V$ и $E' \subseteq E$. Таким образом, каждая вершина в G' является вершиной в G, и каждое ребро в G' является ребром в G.

ПРИМЕР 6.10. Графы, изображенные на рис. 6.9, 6.10 и 6.11,

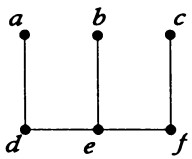


Рис. 6.9

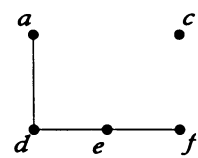


Рис. 6.10

являются подграфами графа на рис. 6.12.

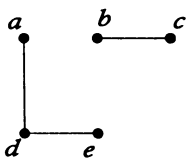


Рис. 6.11

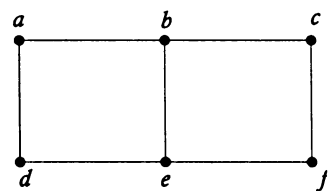


Рис. 6.12 □

Путь (маршрут) в графе — это просто совокупность ребер, которые объединены вместе вершинами так, что вдоль них можно двигаться по графу. В определении, приведенном ниже, свойства пути очерчены более формально. Для удобства договоримся в оставшейся части этого раздела использовать символы $v_0, v_1, v_2, v_3, \ldots, v_k, u$ и v для обозначения вершин.

ОПРЕДЕЛЕНИЕ 6.11. Пусть $G = G(V, E)$ — граф с вершинами v_0, v_1, v_2, $v_3, \ldots, v_k \in V$ и ребрами $e_1, e_2, e_3, \ldots, e_k \in E$. ***Путем длины k*** из v_0 в v_k (или между v_0 и v_k) называется последовательность $v_0 e_1 v_1 e_2 v_2 e_3 v_3 \ldots v_k e_k v_k$ такая, что $e_i = \{v_{i-1}, v_i\}$. Таким образом, путь длины k имеет k ребер. По причине избыточности обозначений в этом определении для графа в общем случае путь будет обозначаться через $v_0 v_1 v_2 v_3 \ldots v_k$. Каждые два последовательных ребра пути имеют общую вершину, поэтому являются смежными. ***Простым путем*** из v_0 в v_k называется путь, в котором нет повторяющихся вершин.

ПРИМЕР 6.12. В графе, изображенном на рис. 6.13,

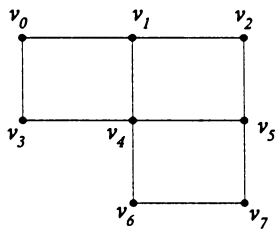


Рис. 6.13

из v_0 в v_7 ведут пути $v_0v_1v_2v_5v_7$, $v_0v_1v_2v_5v_4v_1v_2v_5v_7$, $v_0v_1v_4v_5v_4v_5v_7$ и $v_0v_3v_4v_6v_7$ длины 4, 8, 6 и 4 соответственно. Пути $v_0v_1v_2v_5v_7$ и $v_0v_3v_4v_6v_7$ являются простыми. □

ОПРЕДЕЛЕНИЕ 6.13. Граф G называется ***связным***, если имеется путь между любыми двумя его различными вершинами.

ПРИМЕР 6.14. Граф на рис. 6.14 не связный. Например, нет пути из v_0 в v_3 и между v_2 и v_4.

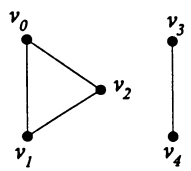


Рис. 6.14 □

Возвращаясь к рассмотрению графа на рис. 6.13, следует отметить, что путь $v_0v_1v_2v_5v_4v_1v_2v_5v_7$ можно сократить до $v_0v_1v_2v_5v_7$. Поскольку вершина v_1 повторялась, необходимо удалить часть пути между двумя появлениями вершины v_1 и после первого появления вершины v_1 переходить сразу к v_2. Аналогично, путь $v_0v_1v_4v_5v_4v_5v_7$ можно сократить до $v_0v_1v_4v_5v_7$. Таким образом, если путь включает какую-либо вершину v_i более чем один раз, его можно сократить, удалив v_i и вершины, лежащие на пути между двумя появлениями вершины v_i. Можно продолжать действовать таким образом, пока не будет исключено повторение любой из вершин. Проведенные рассуждения дают возможность сформулировать следующую теорему.

ТЕОРЕМА 6.15. Пусть $G = G(V, E)$ — граф. Если существует путь из вершины v_i в вершину v_j, тогда существует простой путь из вершины v_i в вершину v_j.

Комбинируя определение 6.13 и теорему 6.15, приходим непосредственно к такому следствию (следствием называется теорема, которая непосредственно вытекает из другой теоремы).

СЛЕДСТВИЕ 6.16. Граф G является связным тогда и только тогда, когда между любыми двумя его вершинами существует простой путь.

ОПРЕДЕЛЕНИЕ 6.17. Пусть $G = G(V, E)$ — граф. Подграф G' графа G называется ***компонентой*** графа G, если выполнены следующие два условия:

1. G' — непустой связный граф.
2. G'' — связный подграф графа G и $G' \preceq G''$, тогда $G' = G''$. Отсюда G' — максимальный связный подграф графа G.

ПРИМЕР 6.18. Графы, изображенные на рис. 6.15 и 6.16, являются компонентами графа, представленного на рис. 6.14.

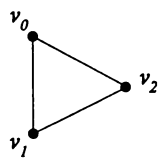


Рис. 6.15



Рис. 6.16

□

ОПРЕДЕЛЕНИЕ 6.19. Пусть $G = (V, E)$ — граф. ***Циклом*** называется путь ненулевой длины, соединяющий вершину саму с собой и не содержащий повторяющихся ребер. ***Простым циклом*** называется цикл, соединяющий вершину v саму с собой и не содержащий повторяющихся вершин, кроме v. Цикл называется ***n-циклом***, если он содержит n ребер и n различных вершин.

ПРИМЕР 6.20. Если опять рассмотреть граф на рис. 6.13, то в нем пути $v_0v_1v_4v_3v_0$, $v_0v_1v_4v_5v_7v_6v_4v_3v_0$, $v_1v_2v_5v_7v_6v_4v_1$ и $v_0v_1v_2v_5v_7v_6v_4v_3v_0$ являются циклами. При этом все циклы , кроме $v_0v_1v_4v_5v_7v_6v_4v_3v_0$, простые. □

Давайте теперь определим графы специального вида, которые пригодятся нам в следующих разделах.

ОПРЕДЕЛЕНИЕ 6.21. Граф называется ***полным***, если любые две его вершины соединены ребром. Полный граф с n вершинами обозначается через K_n.

ПРИМЕР 6.22. На рис. 6.17 показаны, соответственно, графы K_2, K_3, K_4 и K_5.

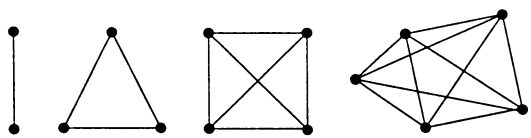


Рис. 6.17

□

ОПРЕДЕЛЕНИЕ 6.23. Граф $G = (V, E)$ называется ***двудольным***, если V можно представить как объединение непересекающихся множеств, скажем, $V = A \cup B$, так что каждое ребро имеет вид $\{a, b\}$, где $a \in A$ и $b \in B$. Таким образом, каждое ребро связывает вершину из A с вершиной из B, но никакие две вер-

шины из A или две вершины из B не являются связанными. Двудольный граф называется ***полным двудольным*** графом $K_{m,n}$, если A содержит m вершин, B содержит n вершин и для каждого $a \in A$, $b \in B$ имеем $\{a, b\} \in E$. Таким образом, для каждого $a \in A$ и $b \in B$ имеется связывающее их ребро.

ПРИМЕР 6.24. Графы $K_{1,2}$, $K_{2,3}$, $K_{2,2}$ и $K_{3,3}$ приведены по порядку на рис. 6.18.

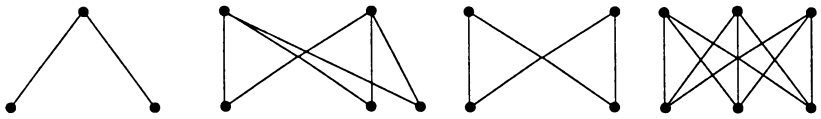


Рис. 6.18

□

■ УПРАЖНЕНИЯ

1. Что из приведенного ниже является путем в графе на рис. 6.19? Которые из них являются простыми путями? Приведите длину каждого из путей.
   а) $aebfcd$; **б)** $aecdaec$;
   в) $aebecfbd$; **г)** $aecfbdafc$.
2. Что из приведенного ниже является путем в графе на рис. 6.20? Которые из них являются простыми путями? Приведите длину каждого из путей.
   а) $abcabcd$; **б)** $bcdeca$;
   в) $debace$; **г)** $decab$.

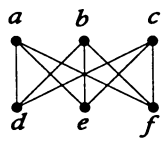


Рис. 6.19

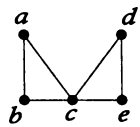


Рис. 6.20

3. Что из приведенного ниже является циклом в графе на рис. 6.21? Которые из них являются простыми циклами? Для каждого n-цикла приведите значение n.
   а) $dabcfbed$; **б)** $bfcedbfcb$;
   в) $abcfebfca$; **г)** $aecfbda$.
4. Что из приведенного ниже является циклом в графе на рис. 6.22? Которые из них являются простыми циклами? Для каждого n-цикла приведите значение n.
   а) $abcdbaea$; **б)** $ebcdbcdae$;
   в) $adcbea$; **г)** $adbcdea$.

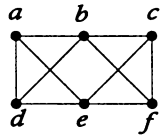


Рис. 6.21

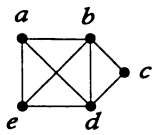


Рис. 6.22

5. Нарисуйте следующие графы:
   а) K_6; б) $K_{1,3}$; в) $K_{1,4}$; г) $K_{3,4}$.
6. Докажите, что если граф содержит цикл от вершины v к самой себе, то он содержит простой цикл от вершины v к самой себе.
7. В игре "вытягивание палочек" некоторое множество их подбрасывается над столом. Содержание игры — вытащить из образовавшейся груды палочки по одной, не нарушая положение других. Представьте себе, что "палочки" — вершины графа. Между двумя вершинами имеется ребро, если одна палочка касается другой. Что представляют собой компоненты этого графа?
8. Покажите, что каждому рефлексивному отношению соответствует граф с петлями и обратно, граф с петлями *в каждой вершине* описывает рефлексивное отношение. Верно ли, что каждому графу с петлями соответствует рефлексивное отношение?

6.2. ОРИЕНТИРОВАННЫЕ ГРАФЫ

Во многих случаях необходимы графы, у которых ребра, по существу, представляют собой улицу с односторонним движением. Это означает, что если рассматривается ребро, выходящее из вершины a в вершину b, то его нельзя рассматривать выходящим из вершины b в вершину a. Например, если граф моделирует поток нефти в трубопроводе, и если нефть течет из пункта a в пункт b, нам не хотелось бы, чтобы поток был и в обратном направлении, из пункта b в пункт a. В последующих главах мы встретим применение ориентированных графов в сетях и теории автоматов.

ОПРЕДЕЛЕНИЕ 6.25. ***Ориентированный граф*** или ***орграф*** G, который обозначается через $G(V, E)$, состоит из множества V вершин и множества E упорядоченных пар элементов из V, называемого множеством ***ориентированных ребер*** или просто ***ребер***, если понятно, что граф ориентирован. Элемент множества E называется ***ориентированным ребром***. Если $(a, b) \in E$, то a называется ***начальной вершиной*** ребра (a, b), а b называется ***конечной вершиной***.

Заметим, что понятие ориентированного графа допускает наличие петель, чего не было в случае простых графов. Тому есть, в основном, две причины. Во-первых, определение ориентированного ребра естественным образом включает понятие петли, т.к. петля у вершины a есть просто ребро (a, a). В случае

неориентированных графов это нельзя было сделать, так как ребро неориентированного графа имеет вид $\{a, b\}$, поэтому петля должна была бы иметь вид $\{a, a\}$. Во-вторых, когда речь идет об отношениях, элемент может находиться в отношении с самим собой, что не имеет смысла для множеств, так как любой элемент присутствует во множестве только один раз. И еще, понятие петли в большей степени присуще ориентированным графам.

Ребро (a, b) ориентированного графа обозначается на диаграмме стрелкой из a в b. Отметим, что в простом графе ребро представляется двухэлементным подмножеством, чтобы подчеркнуть, что отношение симметрично, в то время как в ориентированном графе ребро представлено упорядоченной парой, чтобы акцентировать важность порядка и то, что (a, b) может быть ребром в орграфе, а (b, a) — нет.

ПРИМЕР 6.26. Орграф, у которого $V = \{a, b, c\}$ и $E = \{(a, b), (b, c), (c, b), (c, a)\}$, изображен на рис. 6.23. □

ПРИМЕР 6.27. Орграф, у которого $V = \{a, b, c, d\}$ и $E = \{(a, b), (b, c), (c, c), (b, d), (d, b), (c, d), (d, a)\}$, изображен на рис. 6.24. □

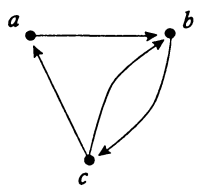


Рис. 6.23

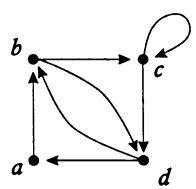


Рис. 6.24

ОПРЕДЕЛЕНИЕ 6.28. Если (a, b) — ребро ориентированного графа $G(V, E)$, так что a — начальная вершина, а b — конечная вершина ребра (a, b), тогда вершины a и b ***инцидентны*** ребру (a, b). Вершина a называется ***смежной*** к вершине b. Вершина b называется также ***смежной*** от вершины a.

ОПРЕДЕЛЕНИЕ 6.29. ***Степенью выхода*** вершины v называется количество ребер, для которых v является начальной вершиной, обозначается $\mathrm{outdeg}(v)$. ***Степенью входа*** вершины v называется количество ребер, для которых v является конечной вершиной, обозначается $\mathrm{indeg}(v)$. Если $\mathrm{indeg}(v) = 0$, то вершина v называется ***источником***. Если $\mathrm{outdeg}(v) = 0$, то вершина v называется ***стоком***.

ПРИМЕР 6.30. В ориентированном графе на рис. 6.25 $\mathrm{indeg}(v_0) = 0$, $\mathrm{indeg}(v_1) = 1$, $\mathrm{indeg}(v_3) = 2$, $\mathrm{indeg}(v_2) = 2$ и $\mathrm{indeg}(v_4) = 3$. Также $\mathrm{outdeg}(v_0) = 3$, $\mathrm{outdeg}(v_1) = 2$, $\mathrm{outdeg}(v_2) = 2$, $\mathrm{outdeg}(v_3) = 1$ и $\mathrm{outdeg}(v_4) = 0$. Таким образом, вершина v_0 — источник, а вершина v_4 — сток.

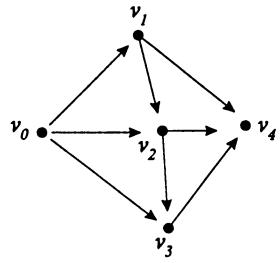


Рис. 6.25

□

Обратим внимание, что, хотя согласно определению ориентированного графа он может иметь петли, до настоящего момента не рассматривалась ситуация, когда из вершины a в вершину b может выходить более чем одно ребро. Ориентированный граф с более чем одним ребром из одной вершины в другую называется ***мультиграфом***, или, более точно, ***ориентированным мультиграфом***. Если каждое ребро помечено, будем говорить, что это ***размеченный ориентированный***, или просто ***размеченный граф***, с пониманием того, что это ориентированный граф. Поскольку в действительности не имеет смысла иметь два ребра из вершины a в вершину b, если нет возможности их отличать, обычно их каким-то образом нумеруют. Формально размеченный граф можно определить следующим образом.

ОПРЕДЕЛЕНИЕ 6.31. ***Размеченный граф*** $G = G(V, L, E)$ представляет собой множество вершин V, множество меток L и множество E, которое является подмножеством $V \times L \times V$. Таким образом, ребро e графа G имеет вид (a, l, b), где l — метка, а a, b — вершины.

Графически ребро $e = (a, l, b)$ размеченного графа обозначается, как на рис. 6.26, или, как на рис. 6.27, если ребро — петля.

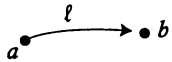


Рис. 6.26



Рис. 6.27

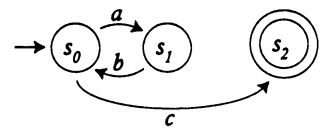


Рис. 6.28

Графы, представленные на рис. 6.28 и 6.29, являются примерами типичных размеченных графов, называемых автоматами.

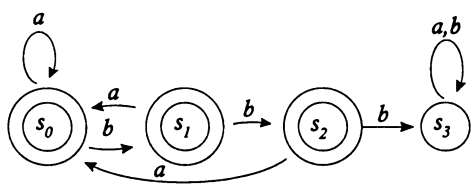


Рис. 6.29

К этим вопросам мы вернемся в главе 17.

ОПРЕДЕЛЕНИЕ 6.32. Ориентированный граф $G'(V', E')$ называется ***ориентированным подграфом*** ориентированного графа $G(V, E)$, обозначается $G'(V', E') \preceq G(V, E)$, если $V' \subseteq V$ и $E' \subseteq E$. Таким образом, каждая вершина в G' является вершиной в G и каждое ориентированное ребро в G' является ориентированным ребром в G.

Понятие ориентированного пути в ориентированном графе вводится по аналогии с понятием пути в графе. Разница лишь в том, что, перемещаясь вдоль ориентированного пути, двигаться нужно в направлении, задаваемом ориентацией ребер.

ОПРЕДЕЛЕНИЕ 6.33. ***Ориентированный путь*** из a в b задается последовательностью вершин $v_0 v_1 v_2 v_3 \cdots v_n$, где $a = v_0$, $b = v_n$ и (v_{i-1}, v_i) для $1 \leq i \leq n$ — ориентированное ребро. ***Длиной*** ориентированного пути называется количество ориентированных ребер, входящих в путь.

ПРИМЕР 6.34. Для графа G, изображенного на рис. 6.30,

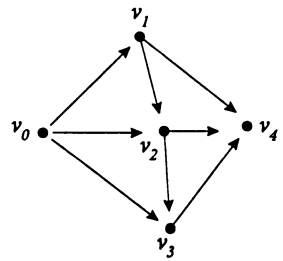


Рис. 6.30

графы на рис. 6.31 и 6.32 являются подграфами. Ориентированные пути в G включают $v_0 v_1 v_2 v_4$, $v_1 v_2 v_4$ и $v_0 v_3 v_4$.

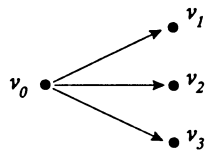


Рис. 6.31

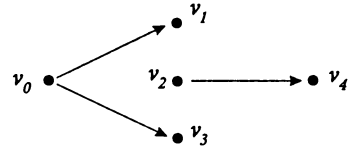


Рис. 6.32

□

Теперь для заданного ориентированного графа G будет построен неориентированный граф G^s такой, что каждое ориентированное ребро G (исключая петли) станет неориентированным ребром графа G^s.

Пусть для каждого ориентированного графа $G(V, E)$ $E' = E - \{(v, v) : v \in V\}$, так что $G'(V, E')$ — ориентированный подграф графа $G(V, E)$, в котором удалены петли. Пусть R — симметричное замыкание множества E', так что если $(a, b) \in E'$, то $(a, b), (b, a) \in R$, а E^s — множество ребер, представляющих отношение R. В таком случае граф $G^s(V, E^s)$ называется ***соотнесенным графом***

ориентированного графа $G(V, E)$. Упрощая формулировку, можно сказать, что множество ребер E^s неориентированного графа $G^s(V, E^s)$ можно определить таким образом: $\{a, b\} \in E^s$ тогда и только тогда, когда для различных вершин a и b ребро $(a, b) \in G$ или $(b, a) \in G$.

ОПРЕДЕЛЕНИЕ 6.35. Ориентированный граф $G(V, E)$ называется ***связным***, если его соотнесенный граф является связным. Ориентированный граф называется ***сильно связным***, если для любой пары вершин $a, b \in V$ существует ориентированный путь из a в b.

ПРИМЕР 6.36. Для ориентированного графа на рис. 6.30 соотнесенным будет граф, показанный на рис. 6.33.

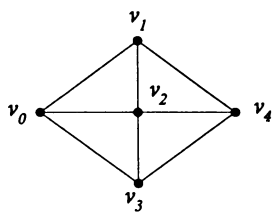


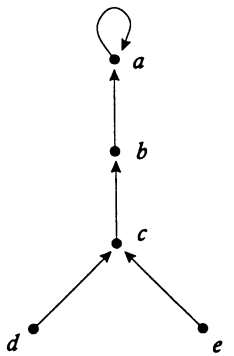
Рис. 6.33

Ориентированный граф является связным, поскольку его соотнесенный граф связный. Рассматриваемый граф однако не является сильно связным, поскольку из v_1 в v_3 не существует ориентированного пути. □

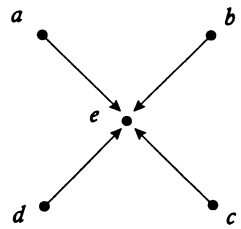
■ УПРАЖНЕНИЯ

1. Найдите вершины и ориентированные ребра для приведенных ниже орграфов. Для каждой вершины определите степень входа и степень выхода. Имеются ли здесь источники и стоки?

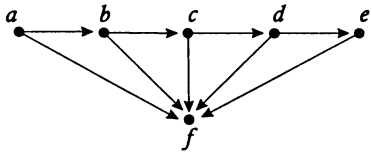
а)



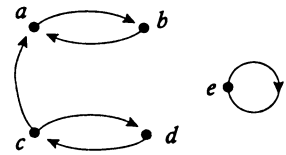
б)



в)

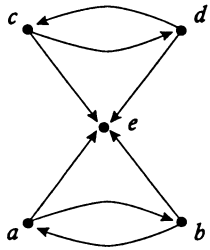


г)

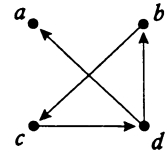


2. Найдите вершины и ориентированные ребра для приведенных ниже орграфов. Для каждой вершины определите степень входа и степень выхода. Имеются ли здесь источники и стоки?

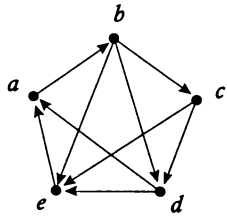
а)



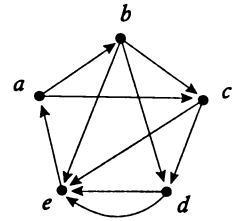
б)



в)



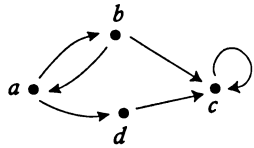
г)



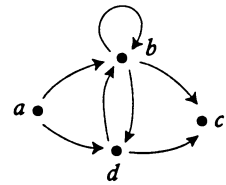
3. Для каждого графа из предыдущего упражнения постройте четыре подграфа.

4. Найдите вершины и ориентированные ребра для приведенных ниже орграфов. Для каждой вершины определите степень входа и степень выхода. Имеются ли здесь источники и стоки?

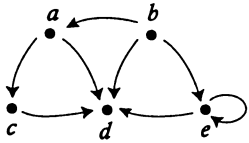
а)



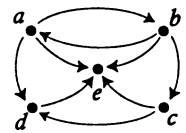
б)



в)

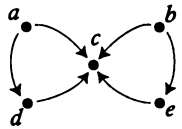


г)

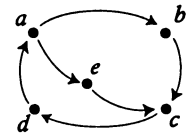


5. Для каждого графа из предыдущего упражнения постройте четыре подграфа.
6. Определите простой ориентированный путь.
7. Определите ориентированный цикл.
8. Определите простой ориентированный цикл.
9. Который из приведенных ниже орграфов является связным? Какой из них является сильно связным?

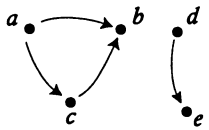
а)



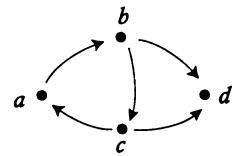
б)



в)

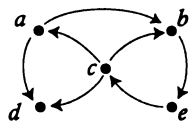


г)

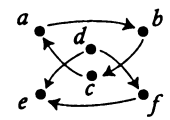


10. Для каждого графа из предыдущего упражнения найдите, если возможно, ориентированный путь длины 2, ориентированный путь длины 3, ориентированный путь длины 4 и ориентированный путь длины 5. Приведите пример пути максимальной длины. Какой самый длинный простой цикл (если таковой существует) может быть построен?
11. Который из приведенных ниже орграфов является связным? Какой из них является сильно связным?

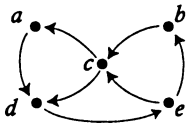
а)



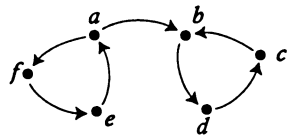
б)



в)



г)



12. Для каждого графа из предыдущего упражнения найдите, если возможно, ориентированный путь длины 2, ориентированный путь длины 3, ориентированный путь длины 4 и ориентированный путь длины 5. Приведите пример пути максимальной длины. Какой самый длинный простой цикл (если таковой существует) может быть построен?

6.3. ДЕРЕВЬЯ

Когда Джойс Килмер (Joyce Kilmer, 1886-1918) писал: "Нет, не увижу, это ясно, поэмы дерева прекрасней", — вряд ли он имел в виду деревья, которые мы будем рассматривать. Разрешим, однако, наши сомнения в пользу автора этих строк. ***Дерево*** — это граф без циклов. ***Лес*** — это граф, компоненты которого являются деревьями (почему бы и нет!). Понятие дерева широко используется во многих областях математики и информатики. Например, они используются как инструмент при вычислениях, как удобный способ хранения данных, способ сортировки или поиска данных. Некоторые из этих приложений будут рассмотрены в последующих разделах. Дерево и названо деревом, поскольку, будучи нарисованным, выглядит как дерево, только перевернутое "вверх ногами". Граф, изображенный на рис. 6.34, является примером дерева.

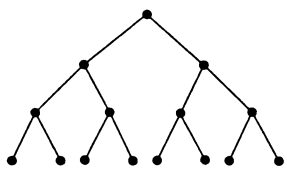


Рис. 6.34

В одной из последующих глав будет показано, что это дерево можно использовать для демонстрации результатов трехкратного подбрасывания монеты. Граф на рис. 6.35 не является деревом, поскольку содержит цикл. Граф на рис. 6.36 — это лес.

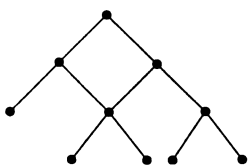


Рис. 6.35

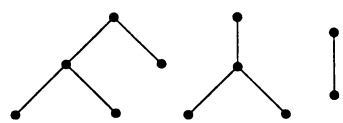


Рис. 6.36

Достаточно развитое генеалогическое дерево образует дерево. Если начать с конкретной (хорошо известной) личности и провести ребра между каждым из родителей и каждым сыном или дочерью, это сформирует дерево. При построении генеалогического дерева, однако, необходимо быть очень осторожным, чтобы браки между дальними родственниками не образовывали циклов. Другим примером дерева может служить организационный устав. Дерево на рис. 6.37 — типичное частичное организационное дерево для университета. В этом разделе, тем не менее, деревья будем рассматривать как графы.

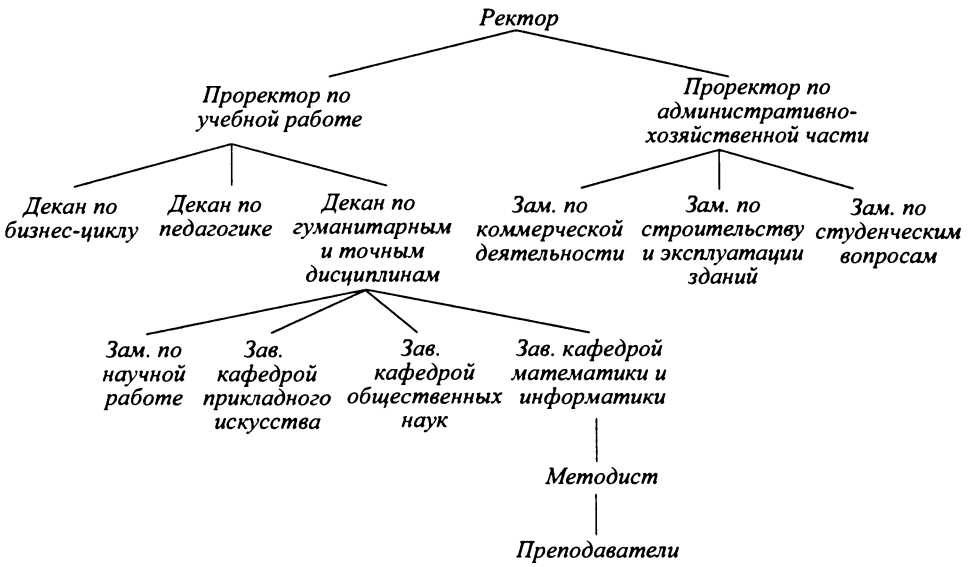


Рис. 6.37

Ориентированное дерево T представляет собой свободный от петель ориентированный граф, соотнесенный граф которого является деревом; так что если существует путь от вершины a к вершине b, то он единственный. Сначала заметим, что если в ориентированном дереве имеется ребро (a, b), тогда не существует ребро (b, a), в противном случае путь aba был бы циклом, и путь из a в b не был бы единственным. Таким образом, множество E, которое для дерева представляет собой как множество ребер, так и отношение, обладает таким свойством, что если $(a, b) \in E$, то $(b, a) \notin E$. Такое отношение называется ***асимметричным***.

Возвращаясь к (неориентированным) деревьям, сначала отметим, что если дерево имеет хотя бы одно ребро, оно имеет хотя бы две вершины со степенью 1. Чтобы убедиться в этом, рассмотрим для заданного дерева T возможные пути. Если у дерева n вершин, тогда длина пути не может превысить $n - 1$, так как каждую вершину можно использовать только один раз. В противном случае получились бы циклы, а у дерева нет циклов. Поэтому существует максимальный путь, который нельзя продлить, чтобы получить более длинный путь. Предположим, что путь начинается в вершине a и заканчивается в вершине b. Как a, так и b

должны иметь степень 1, поскольку в противном случае путь можно было бы продолжить. Но это противоречит тому, что путь максимальный. Вершины степени 1 называются ***листьями***. Другие вершины называются ***внутренними вершинами***. Максимальный путь не обязательно совпадает с самым длинным путем в дереве. Например, в дереве на рис. 6.38 путь $v_0v_2v_5$ является максимальным путем. Вершины v_3, v_4, v_5, v_8 и v_7 представляют собой листья.

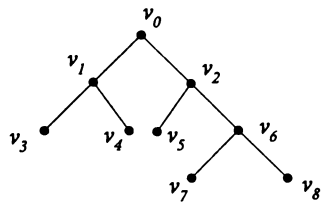


Рис. 6.38

Деревья обладают одним весьма интересным свойством, которое мы сейчас докажем.

ТЕОРЕМА 6.37. Для любых двух вершин a и b дерева T существует единственный путь из a в b.

ДОКАЗАТЕЛЬСТВО. Для доказательства используется тот факт, что контрапозиция утверждения логически эквивалентна самому утверждению. Поэтому предположим, что для некоторых вершин a и b дерева T путь из a в b не является единственным, и покажем, что в таком случае T не будет деревом. Допустим, что существуют два различных пути: $v_0v_1v_2\cdots v_n$ длины n и $v_0v_1'v_2'\cdots v_m'$ длины m, где $a = v_0$ и $b = v_n = v_m'$. В каждом пути должна существовать первая вершина, начиная с которой соответствующие вершины не совпадают, скажем $v_i \neq v_i'$, и в каждом из путей должна существовать точка, начиная с которой вершины опять одни и те же, скажем, $v_j = v_k'$. Тогда $v_{i-1}v_iv_{i+1}v_{i+2}\cdots v_jv_k'v_{k-1}'v_{k-2}'v_i'v_{i-1}$ является циклом, поэтому граф T не является деревом. ■

На основании следующей теоремы можно убедиться в том, что верна также и обратная теорема. Поэтому дерево можно определить как граф, в котором между двумя его вершинами имеется только один путь.

ТЕОРЕМА 6.38. Если для любых двух вершин графа G существует единственный путь из вершины a в вершину b, тогда G — дерево.

ДОКАЗАТЕЛЬСТВО. Опять покажем справедливость теоремы, доказывая контрапозицию исходного утверждения. Предположим, что G не является деревом. Тогда либо G не является связным, либо содержит цикл. Если граф G не связный, то существуют вершины $a, b \in G$, для которых не существует пути из a в b. В таком случае, очевидно, не существует и единственного пути из a в b. Если G содержит цикл $v_0v_1v_2v_3v_4\cdots v_{k-1}v_k$, то как $v_2v_3\cdots v_{k-1}v_kv_0$, так и $v_2v_1v_0$ являются путями из v_2 в v_0. Положив $a = v_2$ и $b = v_0$, видим, что путь между вершинами a и b не является единственным. ■

Предположим, что дерево представляет собой физический объект, подвижный в вершинах, и подвесим дерево за одну из его вершин так, что остальная его часть повиснет ниже этой вершины. Например, пусть задано дерево на рис. 6.39.

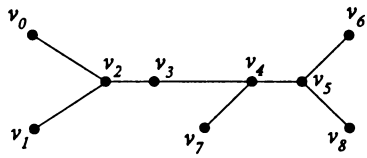


Рис. 6.39

Если подвесить его за вершину v_3, получим дерево, представленное на рис. 6.40. Если подвесим дерево за вершину v_4, оно будет выглядеть так, как показано на рис. 6.41.

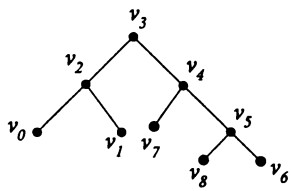


Рис. 6.40

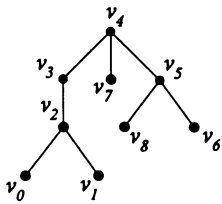


Рис. 6.41

Вершина в самой верхней части каждого из изображений называется ***корнем*** дерева. Если корень дерева определен, дерево называется ***корневым деревом***. При необходимости можно заменить корневое дерево T на ориентированное T', при этом дерево на рис. 6.42 будет заменено деревом на рис. 6.43.

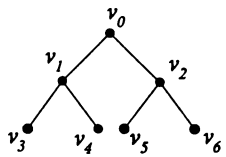


Рис. 6.42

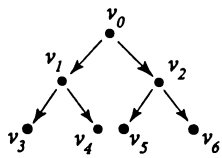


Рис. 6.43

Такое дерево называется ***корневым ориентированным деревом***, T' ***порожденным*** корневым деревом T. При этом следует помнить, что это дерево отличается от неориентированного дерева и что вид ориентированного дерева зависит от выбора корня.

Если корень выбран, ***уровень*** вершины v определяется длиной единственного пути из корня в вершину v. ***Высотой*** дерева называется длина самого длинного пути от корня дерева до листа. Если рассматривается корневое ориентированное дерево T', порожденное данным корневым деревом T, тогда вершина u называется ***родителем*** вершины v, а v называется ***сыном*** вершины u, если существует

ориентированное ребро из u в v. Если u — родитель v и v', тогда v и v' называются ***братьями***. Если существует ориентированный путь из вершины u в вершину v, тогда u называется ***предком*** вершины v, а v называется ***потомком*** вершины u. Если наибольшая из степеней выхода для вершин дерева равна m, тогда дерево называется ***m-арным деревом***. В частном случае, когда $m = 2$, дерево называется ***бинарным деревом***. В каждом бинарном дереве каждый сын родителя обозначается либо как ***левый сын***, либо как ***правый сын*** (но не то и другое одновременно).

ПРИМЕР 6.39. Граф на рис. 6.44 — бинарное дерево. Уровень вершины v_6 равен 2, уровень вершины v_8 равен 3. Высота дерева — 3, поскольку длина пути $v_0v_1v_3v_8$ равна 3 и не существует более длинного пути от корня к листу. Вершина v_1 является родителем для v_3 и v_4.

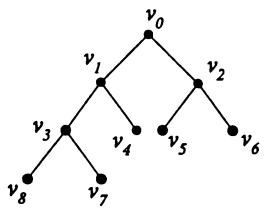


Рис. 6.44

Вершины v_3 и v_4 — братья. Таковыми же являются вершины v_1 и v_2, v_5 и v_6 и v_7 и v_8. Вершина v_1 — предок вершин v_3, v_7, v_8, а v_3, v_7 и v_8 — потомки вершины v_1. Вершина v_8 — левый сын вершины v_3, а v_4 — правый сын вершины v_1. □

Теперь докажем, что в каждом дереве число вершин на единицу больше числа ребер. Предположим, что имеется дерево T. Ранее уже показано, что любое дерево можно представить как корневое дерево, и это никоим образом не меняет ни числа ребер, ни числа вершин. Рассмотрим теперь ориентированное дерево T', порожденное деревом T. У каждого ребра T одна и только одна конечная вершина. Следовательно, число ребер и вершин одно и то же, исключая корневую вершину. Если учесть корневую вершину, получим, что вершин на одну больше, чем ребер. Таким образом, нами доказана следующая теорема.

ТЕОРЕМА 6.40. Если у дерева T имеется e ребер и v вершин, тогда $v = e + 1$.

Справедлива также и обратная теорема.

ТЕОРЕМА 6.41. Если в связном графе G, содержащем e ребер и v вершин, имеем $v = e + 1$, тогда G — дерево.

ДОКАЗАТЕЛЬСТВО. Если G содержит цикл, то, как было установлено при доказательстве теоремы 6.38, если ребро $\{v_i, v_j\}$ входит в цикл, существуют два пути из v_i в v_j. Таким образом, ребро $\{v_i, v_j\}$ можно из цикла удалить, а путь из вершины v_i в вершину v_j будет существовать. Пусть a и b — любые точки в G. Поскольку граф G — связный, существует путь из a в b. Если ребро $\{v_i, v_j\}$ удалено, все равно существует путь из a в b, поскольку ребро $\{v_i, v_j\}$, входящее в путь, можно заменить альтернативным путем из v_i в v_j. Удалим ребро $\{v_i, v_j\}$ из G и, если оставшийся граф все еще содержит цикл, удалим другое ребро, используя ту же процедуру. Будем продолжать, пока все циклы не будут удалены. В

результате получим связный граф, скажем, G', без циклов. Поэтому G' является деревом, и по теореме 6.40 число вершин $v = e' + 1$, где e' — число ребер графа G'. Поскольку ни одна из вершин не была удалена, число вершин остается таким, которое было раньше. Если было удалено n ребер, тогда $e = e' + n$. Но поскольку $v = e + 1$ и $v = e' + 1$, то $e = e'$ и $n = 0$. Следовательно, ни одно ребро не было удалено, поэтому G — дерево. ■

Дерево G', построенное из G в процессе приведенного выше доказательства, называется ***остовным*** (или ***каркасным***) деревом графа G. Дадим более формальное определение.

ОПРЕДЕЛЕНИЕ 6.42. Дерево T называется ***остовным деревом*** графа G, если T — подграф графа G и каждая вершина в G является вершиной в T.

Итак, доказана следующая теорема.

ТЕОРЕМА 6.43. У каждого связного графа существует подграф, который является остовным деревом.

До сих пор рассматривались ориентированные деревья, образованные из корневых (неориентированных) деревьев. Ориентированное дерево называется ***корневым ориентированным деревом***, если существует единственная вершина v_0 такая, что $\text{indeg}(v_0) = 0$, и существует путь из v_0 в каждую другую вершину дерева. Заметим, что ориентированные деревья, которые были рассмотрены до сих пор, в действительности соответствуют этому определению. Рассмотрим, однако, ориентированное дерево на рис. 6.45.

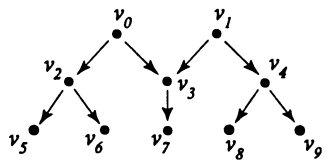


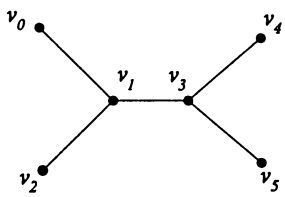
Рис. 6.45

Это ориентированное дерево, но оно не является корневым ориентированным деревом. Большинство ориентированных деревьев, которые мы рассмотрим в дальнейшем, будут, тем не менее, корневыми ориентированными деревьями.

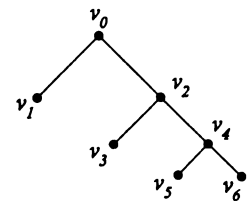
■ УПРАЖНЕНИЯ

1. Которые из приведенных ниже графов являются деревьями?

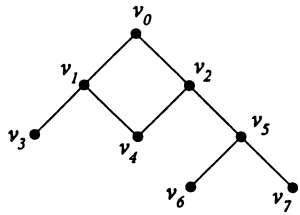
а)



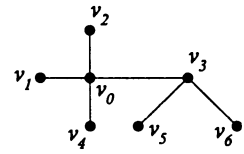
б)



в)



г)



2. Для каждого дерева из предыдущего упражнения

а) используйте в качестве корня вершину v_2 и нарисуйте корневое дерево;

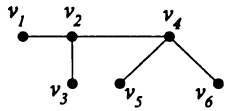
б) нарисуйте порожденное корневое ориентированное дерево;

в) используйте в качестве корня вершину v_3 и нарисуйте корневое дерево;

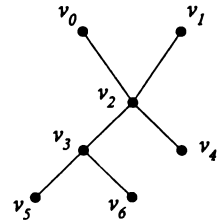
г) нарисуйте порожденное корневое ориентированное дерево.

3. Которые из приведенных ниже графов являются деревьями?

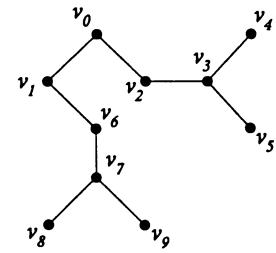
а)



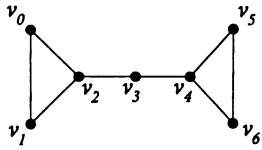
б)



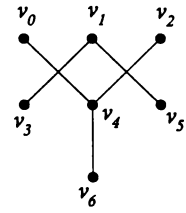
в)



г)



д)



4. Для каждого дерева из предыдущего упражнения

а) используйте в качестве корня вершину v_2 и нарисуйте корневое дерево;

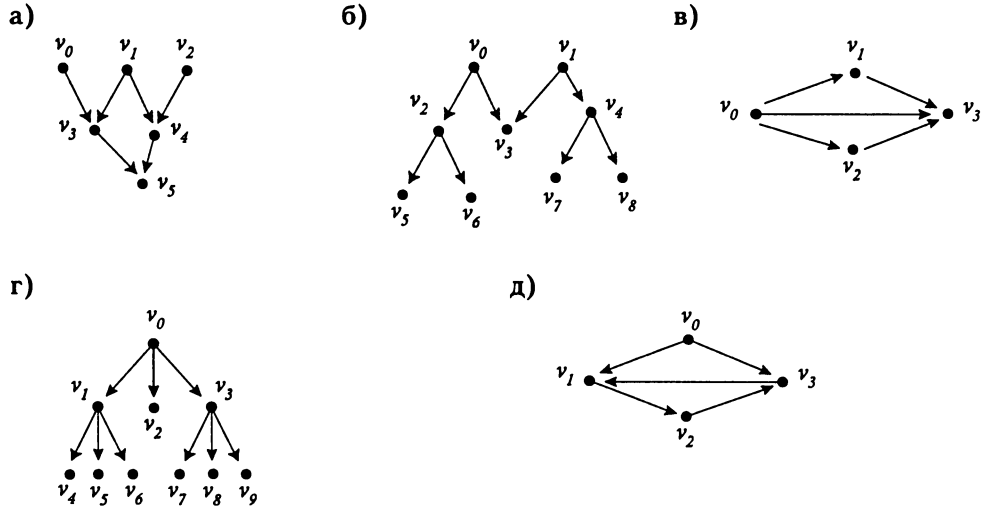
б) нарисуйте порожденное корневое ориентированное дерево;

в) используйте в качестве корня вершину v_3 и нарисуйте корневое дерево;

г) нарисуйте порожденное корневое ориентированное дерево.

5. Покажите, что если лес содержит m компонент, то $v = e + m$.

6. Которые из приведенных ниже графов являются корневыми ориентированными деревьями?



7. Для корневого ориентированного дерева, показанного на рис. 6.46,

а) найдите потомков вершины v_3;

б) найдите предков вершины v_8;

в) найдите родителя вершины v_5;

г) определите уровень вершины v_6;

д) найдите сыновей вершины v_3;

е) найдите высоту дерева;

ж) найдите листья дерева;

з) определите, является ли это дерево бинарным?

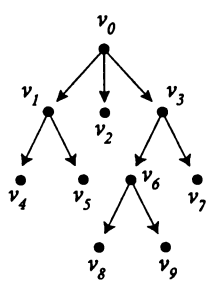


Рис. 6.46

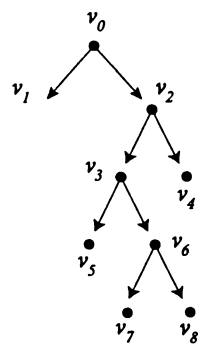


Рис. 6.47

8. Для корневого ориентированного дерева, показанного на рис. 6.47,

а) найдите потомков вершины v_2;

б) найдите предков вершины v_5;

в) найдите родителя вершины v_1;

г) определите уровень вершины v_5;
д) найдите сыновей вершины v_2;
е) найдите высоту дерева;
ж) найдите листья дерева.

9. Нарисуйте генеалогическое дерево, начиная с одного из своих прадедушек.
10. Нарисуйте генеалогическое дерево, начиная с одной из своих прабабушек (но не с жены прадедушки из предыдущей задачи).
11. Используя индукцию, докажите теорему 6.40, которая утверждает, что если у дерева T имеется e ребер и v вершин, то $v = e + 1$.

6.4. МГНОВЕННОЕ БЕЗУМИЕ

В игру ***мгновенное безумие*** играют с помощью четырех кубиков, каждая из шести граней которых окрашена в один из четырех цветов: красный, синий, зеленый или желтый (набор четырех цветов может быть произвольным). Цель игры — составить из кубиков столбик так, что каждая сторона столбика будет окрашена в один из четырех цветов. Решения (если они существуют) зависят от выбора цветов для граней кубика. Решение может быть одно, их может быть несколько, или решение может вообще не существовать. Обозначим грани кубика, показанные на "развертке" на рис. 6.48, как фронтальную (ф), заднюю (з), верхнюю (в), нижнюю (н), левую (л) и правую (п).

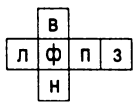


Рис. 6.48

Части кубика, которые должны образовывать сторону столбика, назовем Левой, Правой, Фронтальной и Задней (чтобы отличать их от граней кубика, первоначально фигурирующих в головоломке). Таким образом, Фронтальная и Задняя, а также Левая и Правая части кубика — противоположны друг другу. Предположим, что для заданной игры имеются кубики, показанные на рис. 6.49.

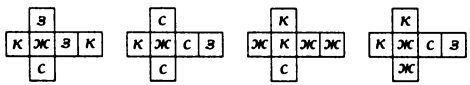


Рис. 6.49

Теперь покажем относительно простой способ нахождения решения, основанный на использовании графов. Перед началом, однако, важно уяснить, что выбор противоположных граней кубика как Фронтальной и Задней никоим образом не определяет, какая другая пара граней может быть Левой и Правой, или какая сторона выбранной пары будет Левой, а какая — Правой. Возможно, в этом легче убедиться, если захватить кубик за противоположные стороны (Фронтальную и Заднюю) и вращать его. Как легко убедиться, четыре вращения реализуют все возможные, упомянутые выше ситуации для других двух сторон (Левой и

Правой). Таким образом, если Фронтальная и Задняя грани для каждого кубика определены, имеется полная свобода выбора пары противоположных граней как Левой и Правой, соответственно. Создаваемый нами граф имеет в качестве вершин четыре цвета. Ребра графа связывают цвета на противоположных гранях кубика. Для первого блока необходимо иметь по одному ребру

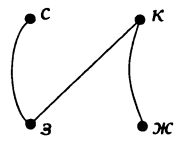


Рис. 6.50

от зеленого к синему, от зеленого к красному и от красного к желтому, что изображено на рис. 6.50. Граф для совокупности четырех блоков показан на рис. 6.51.

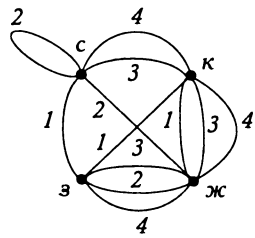


Рис. 6.51

Переупорядочение вершин на диаграмме зачастую позволяет привести граф к более аккуратному виду.

Теперь построим подграф, чтобы иметь возможность выбрать фронтальную и заднюю грань для каждого кубика. Нам нужен такой подграф, что

(1) каждая вершина графа является вершиной подграфа;

(2) грань каждого кубика присутствует в подграфе;

(3) порядок каждой вершины равен 2.

Каждая вершина должна присутствовать, потому что все четыре цвета появляются на фронтальной и задней сторонах. Ребро от каждого кубика должно присутствовать, чтобы мы получили фронтальную и заднюю сторону каждого кубика. Порядок каждой вершины равен двум, потому что каждый цвет появляется один раз на фронтальной поверхности и один раз — на Задней. Примеры возможных подграфов показаны на рис. 6.52.

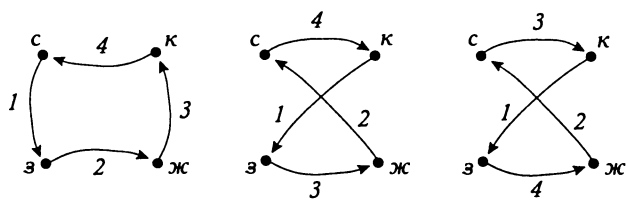


Рис. 6.52

(Попытайтесь найти и другие такие подграфы.) Итак, начинаем с испытания первого из показанных подграфов. Изображая его как ориентированный граф, образующий цикл, можно рассматривать начальную вершину как фронтальную грань, а конечную грань — как заднюю. Таким образом, получаем таблицу

Фронтальная	Задняя
с	*з*
з	*ж*
ж	*к*
к	*с*

Теперь нам нужен другой подграф с такими же свойствами. Единственное ограничение для второго подграфа — он не должен содержать ребра первого подграфа. Эти ребра уже были использованы как Фронтальные и Задние грани для блоков, поэтому не могут фигурировать как боковые. Для упрощения графа удалим использованные ребра, после чего получаем рис. 6.54. По этому графу строим подграф, показанный на рис. 6.53, который можно использовать.

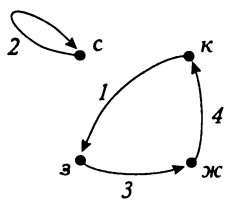


Рис. 6.53

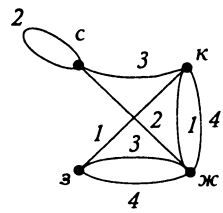


Рис. 6.54

В результате получаем цвета для Левой и Правой частей, показанные в таблице, что и приводит нас к решению.

Левая	Правая
к	*з*
с	*с*
з	*ж*
ж	*к*

Если граф, выбранный для Фронтальной и Задней частей, не дает результата, или желательно найти другое решение, то необходимо испытать другие варианты выбора Фронтальных и Задних частей. Для поиска иных вариантов полезным будет замечание о том, что такие подграфы должны иметь одну из базисных форм, показанных на рис. 6.55.

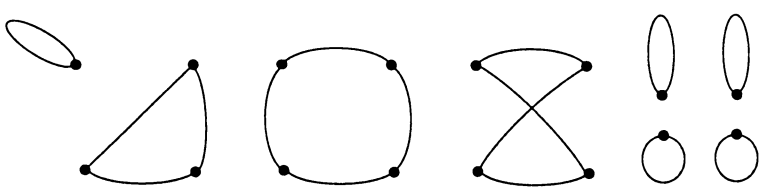
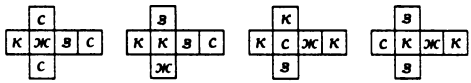


Рис. 6.55

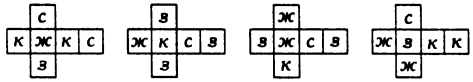
■ УПРАЖНЕНИЯ

1. Найдите решение (если оно существует) для следующих наборов кубиков:

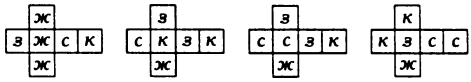
а)



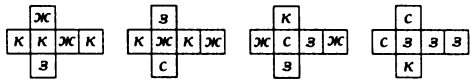
б)



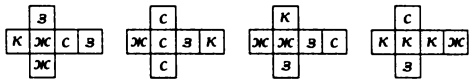
в)



г)



д)



6.5. ПУТИ И ЦИКЛЫ ЭЙЛЕРА

ОПРЕДЕЛЕНИЕ 6.44. Пусть $G = (V, E)$ — граф. Цикл, который включает все ребра и вершины графа G, называется ***эйлеровым циклом***. Если это условие выполняется, говорят, что граф G имеет эйлеров цикл.

Если теперь вернуться к задаче о кенигсбергских мостах, легко убедиться, что она сводится к попытке определить, содержит ли граф, изображающий задачу, эйлеров цикл. Для этой цели нам потребуется приведенная ниже теорема. Эта теорема справедлива также для мультиграфов и псевдографов, исключая тот случай, когда псевдограф имеет только одну вершину. Более того, доказательство теоремы остается тем же. Для ясности будем использовать термин граф, понимая, что каждое утверждение справедливо для мультиграфов и псевдографов.

ТЕОРЕМА 6.45. Граф с более чем одной вершиной имеет эйлеров цикл тогда и только тогда, когда он связный и каждая его вершина имеет четную степень.

ДОКАЗАТЕЛЬСТВО. Предположим, что граф G имеет эйлеров цикл. Граф является связным, т.к. каждая вершина принадлежит циклу. Для всякой вершины v графа G каждый раз, когда эйлеров цикл проходит через v, он вносит 2 в степень v. Поэтому степень v четная.

Обратно, нужно показать, что каждый связный граф, у которого степени вершин четные, имеет эйлеров цикл. Докажем эту теорему, используя индукцию по числу вершин. Поскольку теорема тривиально справедлива при $n \leq 3$, начнем индукцию с $n = 3$. Предположим, что каждый связный граф, имеющий менее k вершин, и все вершины которого обладают четной степенью, содержит эйлеров цикл. Пусть G — связный граф, содержащий k вершин, степени которых четные. Допустим, что v_1 и v_2 — вершины графа G. Поскольку граф G — связный, существует путь из v_1 в v_2. Поскольку степень v_2 — четная, существует неиспользованное ребро, по которому можно продолжить путь. Поскольку граф конечный, то путь, в конце концов, должен вернуться в v_1, и эйлеров цикл C_1 можно считать построенным. Если C_1 является эйлеровым циклом для G, тогда доказательство закончено. Если нет, то пусть G' — подграф графа G, полученный удалением всех ребер, принадлежащих C_1. Поскольку C_1 содержит четное число ребер, инцидентных каждой вершине, каждая вершина подграфа G' имеет четную степень.

Пусть e — ребро графа G', пусть G_e — компонента графа G', содержащая e. Поскольку G' имеет менее, чем k, вершин, и у каждой вершины графа G' четная степень, граф G' имеет эйлеров цикл. Пусть это C_2. Далее, у C_1 и C_2 имеется общая вершина, скажем, a. (Почему?) Теперь можно продолжить эйлеров цикл, начиная его в a, пройти C_1, вернуться в a, затем пройти C_2 и вернуться в a. Если новый эйлеров цикл не является эйлеровым циклом для G, продолжаем использовать этот процесс, расширяя наш эйлеров цикл, пока, в конце концов, не получим эйлеров цикл для G. ■

ПРИМЕР 6.46. Граф на рис. 6.56 имеет эйлеров цикл, поскольку степень каждой его вершины четная. □

ПРИМЕР 6.47. Граф на рис. 6.57 не имеет эйлерова цикла, поскольку степени вершин v_2 и v_4 — нечетные. □

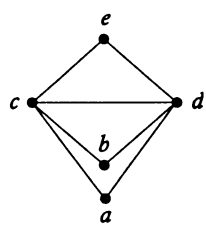


Рис. 6.56

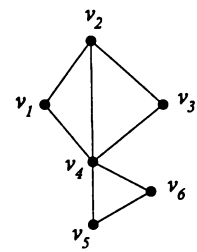


Рис. 6.57

Возвращаясь к задаче о кенигсбергских мостах, обнаруживаем, что мультиграф на рис. 6.58, построенный Эйлером для описания кенигсбергских мостов,

имеет нечетные степени всех его вершин. Итак, этот мультиграф не имеет эйлерова цикла, поэтому невозможно пройти каждый мост по одному разу и вернуться в исходную точку пути. Необходимо также отметить, что хотя задача была решена с использованием мультиграфа, ее можно решить, используя простой граф. Когда имеется более одного

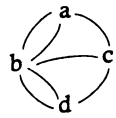


Рис. 6.58

ребра между вершинами, нужно к исходному мультиграфу просто добавить вершину для представления середины каждого моста. Тогда получится простой граф, показанный на рис. 6.59, который также описывает задачу о кенигсбергских мостах.

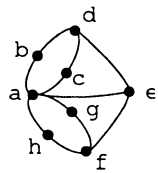


Рис. 6.59

Что касается кенигсбергских мостов, можно также задать вопрос: "Возможно ли не проходить каждый мост по одному разу и не возвращаться обязательно в исходную точку маршрута?". Это предположение приводит к следующему определению и теореме.

ОПРЕДЕЛЕНИЕ 6.48. Пусть $G = (V, E)$ — граф. Путь, который включает каждое ребро графа G только один раз называется ***эйлеровым путем***. В этом случае говорят, что граф G имеет эйлеров путь.

Теперь нас будет интересовать случай, когда эйлеров путь не является эйлеровым циклом. Такой путь будем называть ***собственным эйлеровым путем***. Предоставляем читателю доказать следующую теорему.

ТЕОРЕМА 6.49. Граф (мультиграф или псевдограф) имеет собственный эйлеров путь тогда и только тогда, когда он связный и ровно две его вершины имеют нечетную степень.

Поскольку граф для кенигсбергских мостов имеет четыре вершины с нечетными степенями, можно сделать вывод о невозможности пройти каждый мост по одному разу, даже если не нужно возвращаться в исходную точку маршрута.

ПРИМЕР 6.50. Граф на рис. 6.60 имеет собственный эйлеров путь, т.к. ровно две его вершины имеют нечетную степень. □

ПРИМЕР 6.51. Граф на рис. 6.61 не имеет собственного эйлерова пути, т.к. четыре его вершины имеют нечетную степень. □

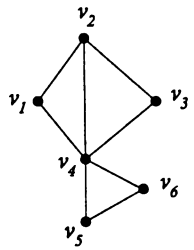


Рис. 6.60

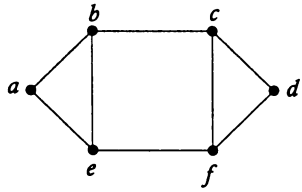


Рис. 6.61

ОПРЕДЕЛЕНИЕ 6.52. Пусть $G = (V, E)$ — ориентированный граф. ***Ориентированным циклом*** называется ориентированный путь ненулевой длины из вершины в ту же вершину без повторения ребер.

ОПРЕДЕЛЕНИЕ 6.53. Пусть $G = (V, E)$ — ориентированный граф. Ориентированный цикл, который включает все ребра и вершины графа G, называется ***эйлеровым циклом***. В этом случае говорят, что ориентированный граф G имеет эйлеров цикл.

Доказательство следующей теоремы предоставляется читателю.

ТЕОРЕМА 6.54. Ориентированный граф имеет эйлеров цикл тогда и только тогда, когда он связный и степень входа каждой вершины равна ее степени выхода.

ПРИМЕР 6.55. Ориентированный граф на рис. 6.62 имеет эйлеров цикл, так как степень входа каждой вершины равна степени выхода. □

ПРИМЕР 6.56. Ориентированный граф на рис. 6.63 не имеет эйлерова цикла, так как степень входа вершины v_1 не равна ее степени выхода. □

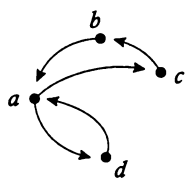


Рис. 6.62

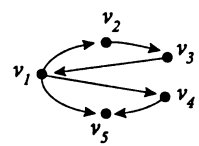
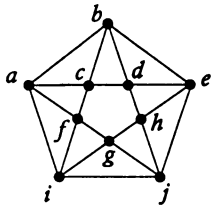


Рис. 6.63

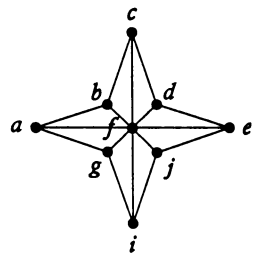
■ УПРАЖНЕНИЯ

1. Среди приведенных ниже графов найдите те, которые имеют эйлеров цикл.

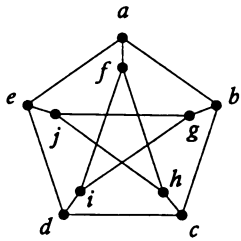
а)



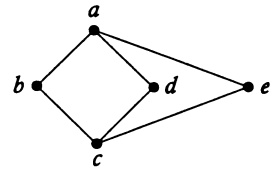
б)



в)

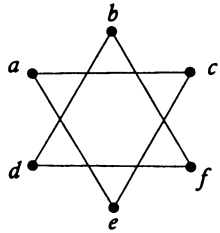


г)

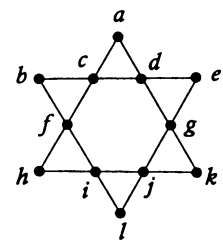


2. Среди приведенных ниже графов найдите те, которые имеют эйлеров цикл.

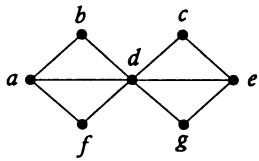
а)



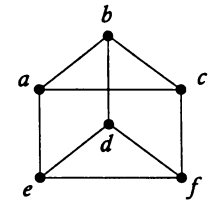
б)



в)

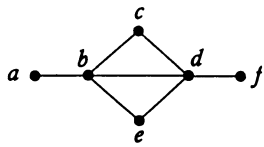


г)

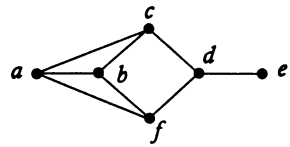


3. Среди приведенных ниже графов найдите те, которые имеют собственный эйлеров цикл.

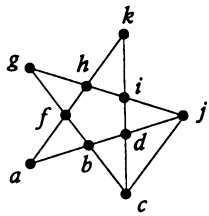
а)



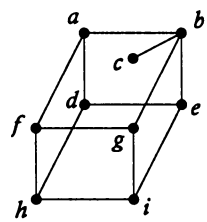
б)



в)

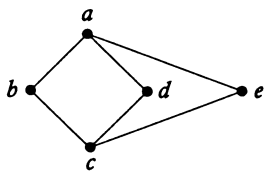


г)

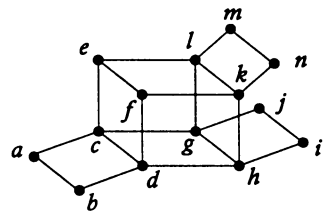


4. Среди приведенных ниже графов найдите те, которые имеют собственный эйлеров цикл.

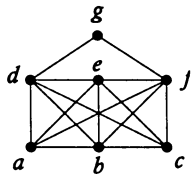
а)



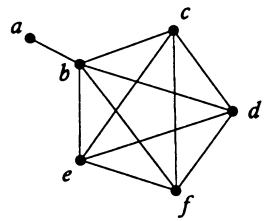
б)



в)

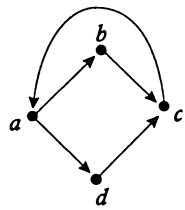


г)

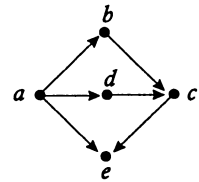


5. Какие из следующих графов являются сильно связными?

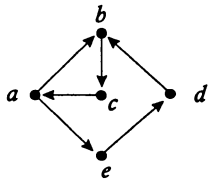
а)



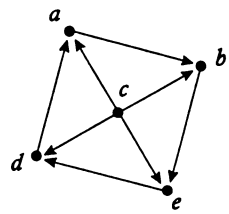
б)



в)

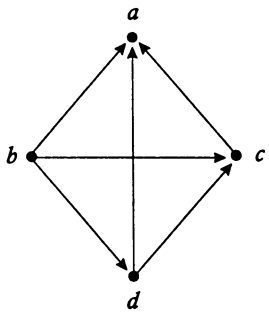


г)

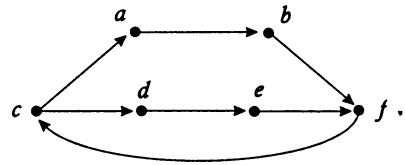


6. Какие из следующих графов являются сильно связными?

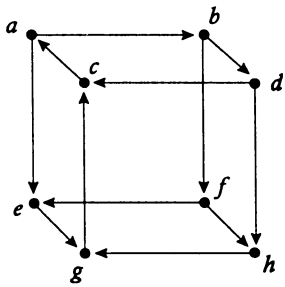
а)



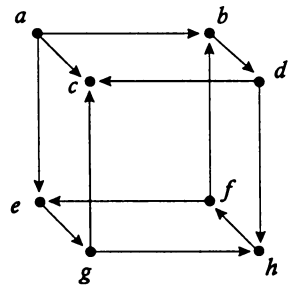
б)



в)

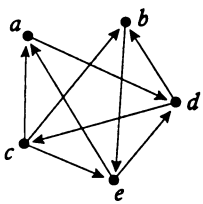


г)

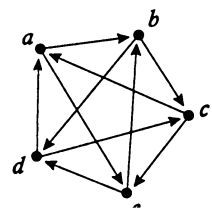


7. Какие из следующих ориентированных графов имеют эйлеровы циклы?

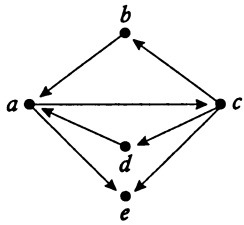
а)



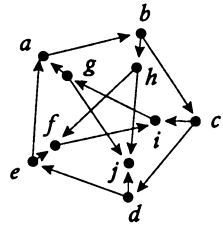
б)



в)

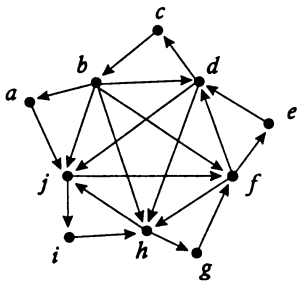


г)

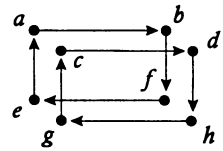


8. Какие из следующих ориентированных графов имеют эйлеровы циклы?

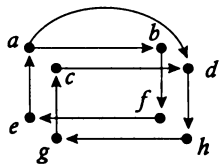
а)



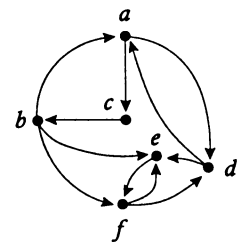
б)



в)



г)



9. Докажите теорему 6.49. Граф (мультиграф или псевдограф) имеет собственный эйлеров путь тогда и только тогда, когда он связный и ровно две его вершины имеют нечетную степень.
10. Докажите, что если граф содержит цикл от вершины v к ней самой, то он содержит простой цикл от вершины v к ней самой.
11. Докажите, что ориентированный граф является сильно связным, если в графе существует вершина v такая, что каждая другая вершина графа достижима из v и вершина v достижима из любой другой вершины графа.
12. Докажите теорему 6.54. Ориентированный граф имеет эйлеров путь тогда и только тогда, когда он сильно связный и степень входа каждой вершины равна ее степени выхода.

6.6. МАТРИЦЫ ИНЦИДЕНТНОСТИ И СМЕЖНОСТИ

В этом разделе будут определены две матрицы, связанные с графами: матрицы инцидентности и матрицы смежности. Будет показано, что задание любой из этих матриц дает возможность восстановить граф. Фактически будет показано, что матрица смежности совпадает с матрицей отношения, определенного этим графом. Как и в случае матрицы отношения для графа, все эти матрицы содержат в качестве элементов 1 или 0, поэтому такие матрицы удобно хранить в памяти компьютера.

ОПРЕДЕЛЕНИЕ 6.57. Пусть G — граф. Пусть B — матрица, строки которой обозначены вершинами графа, а столбцы обозначены ребрами графа. Будем считать, что вершины и ребра графа пронумерованы. Элемент i-ой строки и j-го столбца матрицы B, обозначаемый B_{ij}, равен 1, если i-ая вершина инцидентна j-ому ребру, и равен 0 в противном случае. Матрица B называется ***матрицей инцидентности*** графа G.

ПРИМЕР 6.58. Пусть G — граф, изображенный на рис. 6.64, тогда его матрица инцидентности имеет вид, изображенный на рис. 6.65.

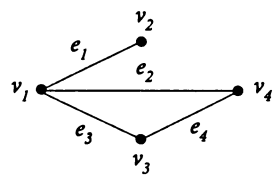


Рис. 6.64

	e_1	e_2	e_3	e_4
v_1	1	1	1	0
v_2	1	0	0	0
v_3	0	0	1	1
v_4	0	1	0	1

Рис. 6.65

□

Легко видеть, что степень вершины равна сумме элементов строки, обозначенной этой вершиной, так как каждая единица в этой строке представляет инцидентность этой вершины ребру. В каждом столбце также будут две единицы, так как каждое ребро инцидентно двум вершинам.

Можно также включить в рассмотрение матрицы инцидентности для графов с петлями. Вид матрицы инцидентности непосредственно показывает, является ли данное ребро петлей, так как ребро представляет собой петлю тогда и только тогда, когда соответствующий столбец содержит только одну единицу. Заметим, что в матрице инцидентности для графа с петлями сумма элементов строки, соответствующей данной вершине, не представляет собой степень вершины, если в ней имеются петли.

ПРИМЕР 6.59. Пусть G — граф, изображенный на рис. 6.66, его матрица инцидентности изображена на рис. 6.67.

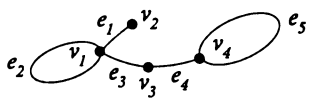


Рис. 6.66

	e_1	e_2	e_3	e_4	e_5
v_1	1	1	1	0	0
v_2	1	0	0	0	0
v_3	0	0	1	1	0
v_4	0	0	0	1	1

Рис. 6.67

Обратите внимание, что наличие петель e_2 и e_5 приводит к тому, что в столбцах, обозначенных этими ребрами, содержится только по одной единице. □

Матрицы инцидентности не имеют большого значения при рассмотрении ориентированных графов, т.к. они не содержат информации о том, как ребро ориентировано. Поэтому, используя матрицу инцидентности, нельзя восстановить ориентированный граф. Такую возможность обеспечивает матрица смежности, определение которой будет сейчас дано.

ОПРЕДЕЛЕНИЕ 6.60. Пусть G — граф (ориентированный граф). Пусть B — матрица, строки которой обозначены вершинами графа и столбцы обозначены теми же вершинами в том же самом порядке. Элемент i-ой строки и j-го столбца матрицы B, обозначаемый B_{ij}, равен 1, если имеется ребро (ориентированное ребро) из i-ой вершины в j-ую вершину, и равен 0 в противном случае. Матрица B называется ***матрицей смежности*** графа G.

Заметим, что свойства матриц смежности в точности совпадают со свойствами матриц отношения из раздела 4.3. Единственное отличие состоит в том, что используется терминология теории графов, и смысл матрицы смежности интерпретируется для графов. Для матриц смежности включен в рассмотрение также алгоритм Уоршолла.

ПРИМЕР 6.61. Пусть G — граф, изображенный на рис. 6.68. Его матрица смежности приведена на рис. 6.69.

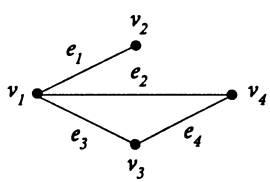


Рис. 6.68

	v_1	v_2	v_3	v_4
v_1	0	1	1	1
v_2	1	0	0	0
v_3	1	0	0	1
v_4	1	0	1	0

Рис. 6.69 □

Поскольку петли отсутствуют, все элементы диагонали матрицы равны 0. Матрица смежности симметрична, так как граф представляет симметричное отношение.

ПРИМЕР 6.62. Пусть G — ориентированный граф, изображенный на рис. 6.70. Его матрица смежности показана на рис. 6.71.

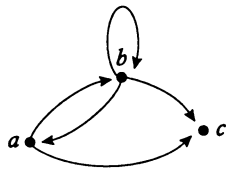


Рис. 6.70

$$\begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

Рис. 6.71 □

Очень часто обозначения вершин несущественны. В таких случаях матрицы будут приводиться без обозначений. Например, матрица

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

является матрицей смежности для ориентированного графа, у которого четыре вершины и восемь ребер.

Важным применением матрицы смежности является возможность находить путь фиксированной длины k. Следующий пример дает ключ к пониманию, как это можно сделать.

Пусть матрица

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

представляет собой матрицу смежности ориентированного графа G с вершинами v_1, v_2, v_3 и v_4. Рассмотрим матрицу

$$A^{\odot 2} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

и вычислим, например, элемент

$$\begin{aligned} A_{12}^{\odot 2} &= (A_{11} \wedge A_{12}) \vee (A_{12} \wedge A_{22}) \vee (A_{13} \wedge A_{32}) \vee (A_{14} \wedge A_{42}) = \\ &= (1 \wedge 0) \vee (0 \wedge 0) \vee (0 \wedge 0) \vee (1 \wedge 1) = \\ &= 0 \vee 0 \vee 0 \vee 1 = \\ &= 1. \end{aligned}$$

Обратите внимание, что значение $A_{12}^{\odot 2}$ равно 1, так как A_{14} и A_{42} оба равны 1, что означает наличие ребра из вершины v_1 в вершину v_4 и из v_4 в вершину v_2. Поэтому из вершины v_1 в вершину v_2 имеется путь длины 2. В общем случае получаем, что $A_{ij}^{\odot 2} = 1$ тогда и только тогда, когда имеется такое число k, что $A_{ik} \wedge A_{kj} = 1$, или, другими словами, имеется ребро из вершины v_i в вершину v_k

и из v_k в вершину v_j. Поэтому существует путь длины 2 или 2-путь из вершины v_i в вершину v_j. Но для $A_{34}^{\odot 2}$ имеем

$$\begin{aligned} A_{34}^{\odot 2} &= (A_{31} \wedge A_{14}) \vee (A_{32} \wedge A_{24}) \vee (A_{33} \wedge A_{34}) \vee (A_{34} \wedge A_{44}) = \\ &= (0 \wedge 1) \vee (0 \wedge 0) \vee (0 \wedge 1) \vee (1 \wedge 0) = \\ &= 0 \vee 0 \vee 0 \vee 0 = \\ &= 0, \end{aligned}$$

так что $A_{34}^{\odot 2} = 0$, потому что не существуют ребра из v_3 в v_i и от v_i к v_4 ни при каком фиксированном i. Другими словами, не существует 2-путь из вершины v_1 в вершину v_2. Отсюда следует, что $A_{ij}^{\odot 2} = 1$, если имеется 2-путь из вершины v_i в вершину v_j, и $A_{ij}^{\odot 2} = 0$, если такой 2-путь не существует. Аналогичным образом, можно убедиться в том, что если использовать обычное умножение матриц, то A_{ij}^2 равно числу значений k таких, что A_{ik} и A_{kj} оба равны 1, поэтому оно равно числу путей из вершины v_i в вершину v_j длины 2.

Просим читателя, используя индукцию, доказать следующие теоремы.

ТЕОРЕМА 6.63. Пусть G — (ориентированный) граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Путь длины k, или k-путь из v_i в v_j для $1 \leq k \leq n$ существует тогда и только тогда, когда $A_{ij}^{\odot k} = 1$.

ТЕОРЕМА 6.64. Пусть G — (ориентированный) граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Из вершины v_i в вершину v_j имеется m путей длины k, где $1 \leq k \leq n$ тогда и только тогда, когда $A_{ij}^k = m$.

Предлагаем читателю, воспользовавшись теоремой 6.63, доказать следующее утверждение.

ТЕОРЕМА 6.65. Пусть G — (ориентированный) граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Пусть $\hat{A} = A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$. Тогда $\hat{A}_{ij} = 1$ тогда и только тогда, когда существует путь из v_i в v_j.

Справедливость теоремы, приведенной ниже, следует непосредственно из определения связности графов или сильной связности ориентированных графов. По сути это та же теорема 6.65.

ТЕОРЕМА 6.66. Пусть G — (ориентированный) граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Пусть $\hat{A}^I = I \vee A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$, где I — мультипликативная единичная матрица. Тогда $\hat{A}_{ij}^I = 1$ для всех i и j тогда и только тогда, когда граф G (сильно) связный.

Из теоремы 6.65 следует, что если R — отношение, задаваемое (ориентированным) графом G, и A — матрица смежности графа G, тогда $\hat{A}$ представляет собой матрицу смежности графа, который описывает транзитивное замыкание R.

В более общем случае, когда G — граф, который не обязательно связный, для матрицы $\hat{A}^I$ справедлив следующий результат.

ТЕОРЕМА 6.67. Пусть G — граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Как и ранее, пусть $\hat{A}^I = I \vee A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$. Тогда (если необходимо) вершины могут быть упорядочены таким образом, что $\hat{A}^I$ будет иметь вид

$$\begin{bmatrix} \hat{A}_1 & & & & \\ & \hat{A}_2 & & 0 & \\ & & \hat{A}_3 & & \\ & 0 & & \ddots & \\ & & & & \hat{A}_m \end{bmatrix},$$

где каждое $\hat{A}_i$ представляет собой квадратную матрицу, главная диагональ которой направлена вдоль главной диагонали $\hat{A}^I$ и содержит все ее элементы, равные 1. Как указано, любой элемент $\hat{A}^I$, который находится вне всех матриц $\hat{A}_i$, равен 0. Матрица A также может быть разбита на блоки такого же размера, что и $\hat{A}^I$, при этом A будет иметь вид

$$\begin{bmatrix} A_1 & & & & \\ & A_2 & & 0 & \\ & & A_3 & & \\ & 0 & & \ddots & \\ & & & & A_m \end{bmatrix}.$$

Здесь каждое A_i имеет такую же форму, как и $\hat{A}_i$, и представляет собой матрицу инцидентности компоненты графа G, а любой элемент матрицы A, находящийся вне всех матриц $\hat{A}_i$, равен 0.

ДОКАЗАТЕЛЬСТВО. Если все вершины G, которые принадлежат одной и той же компоненте, помещены вместе, тогда, поскольку между двумя любыми такими вершинами существует путь, блок матрицы $\hat{A}^I$, состоящий только из строк и столбцов, обозначенных этими вершинами, содержит в качестве элементов только единицы. Далее, любой элемент A в этой же строке или столбце, обозначенных указанными вершинами, должен быть равен нулю, поскольку не существует пути из любой другой вершины в вершину, принадлежащую компоненте.

Поскольку для блоков $\hat{A}_i$ вершины, обозначающие его строки и столбцы, принадлежат одной и той же компоненте, соответствующий блок A_i в матрице A представляет граф — компоненту графа G. Как и ранее, и по той же причине, все другие элементы той же строки или столбца матрицы A, обозначенных этими вершинами, должны быть равны нулю. ■

Матрица $\hat{A}$ может быть вычислена как $\hat{A} = A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$, но такой способ не очень эффективен. Намного лучший метод обеспечивает ***алгоритм Уоршолла***, известный также как алгоритм Роя-Уоршолла. Чтобы понять, как он работает, рассмотрим матрицу смежности, изображенную на рис. 6.72.

$$A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Рис. 6.72

Матрица A представляет множество всех 1-путей. Теперь необходимо найти все 2-пути, где средней вершиной является v_1, чтобы скомбинировать 1-пути, которые уже имеются. Начинаем с первого столбца. Игнорируем 1 в первой строке, если имеется 1 в i-ой строке первого столбца, так как в этом случае существует 1-путь из v_i в v_1. Поскольку 1 имеется в строке 3, существует 1-путь из v_3 в v_1. Теперь посмотрим на первую строку. Игнорируем 1 в первом столбце, если имеется 1 в j-ом столбце, так как в этом случае существует ребро, или 1-путь из v_1 в v_j. Поскольку 1 имеется в четвертом столбце, существует ребро из v_1 в v_4. Таким образом, имеется 2-путь из v_3 через v_1 в v_4. Обозначим это, поместив 1 в третьей строке четвертого столбца, и получим матрицу, показанную на рис. 6.73.

$$\begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{array} \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array}$$

Рис. 6.73

Поскольку в первом столбце и в первой строке нет других единиц, этот шаг закончен. Если бы в любой другой строке первого столбца имелась 1, скажем, в i-ой строке, или в любом другом столбце первой строки имелась 1, скажем, в j-ом столбце, тогда нам нужно было бы поместить 1 в i-ой строке j-ого столбца. В нашем примере, использовав $\vee$ как сложение, мы фактически просуммировали первую строку с третьей. В общем случае, если имеется 1 в i-ой строке первого столбца, то первая строка прибавляется к i-ой строке.

Теперь необходимо найти все пути длины не больше 3, проходящие через v_1 и/или v_2 (если такие имеются). Рассмотрим второй столбец. Игнорируя 1 во второй строке, ищем 1 в любой другой строке столбца 2. Поскольку имеется 1 в строке 4, то существует 1-путь из v_4 в v_2 или 2-путь из v_4 через v_1 в v_2, так как это единственный способ, которым мы можем получить единицу.

Игнорируя 1 во втором столбце, ищем другие 1 в любом другом столбце строки 2. Поскольку имеется 1 в столбце 3, существует 1-путь из v_2 в v_3, или 2-путь из v_2 через v_1 в v_3. В любом случае существует путь из v_4 в v_3 такой, что он проходит только через вершины v_1 и v_2. Опять обозначаем это, помещая 1 в четвертой строке третьего столбца, и получаем матрицу, показанную на рис. 6.74.

$$\begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array} \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array}$$

Рис. 6.74

То же самое можно выполнить, прибавляя вторую строку к четвертой строке. Если бы 1 находилась в i-ой строке столбца 2 и 1 находилась бы в j-ом столбце второй строки, нам надо было бы прибавить строку 2 к строке i.

Аналогично, теперь нам необходимы все пути длины 4 или менее, проходящие через v_1 и/или v_2, и/или v_3 (если такие имеются). Рассмотрим третий столбец и третью строку. Если имеется 1 в i-ой строке столбца 3 и 1 имеется в j-ом столбце строки 3, то существует путь из v_i в v_3, который проходит только через вершины v_1 и/или v_2 (если таковой имеется) и путь из v_3 в v_j, проходящий только через вершины v_1 и/или v_2 (если таковой имеется). Поэтому существует путь из v_i в v_j, и 1 нужно поместить на (i, j)-ой позиции. Это равносильно сложению строки 3 с любой строкой, содержащей 1 в столбце 3. Поскольку 1 имеется в третьем столбце первой и четвертой строк, строка 3 прибавляется к каждой из этих строк. Это дает матрицу, показанную на рис. 6.75.

	v_1	v_2	v_3	v_4
v_1	1	0	0	1
v_2	1	1	1	1
v_3	1	0	1	1
v_4	1	1	1	1

Рис. 6.75

Окончательно, нам необходимы все пути длины 4 или менее, проходящие через v_1 и/или v_2, и/или v_3, и/или v_4 (если таковые существуют). Прибавляем четвертую строку к любой другой строке, в которой 1 появляется в столбце 4.

Ниже приведены два алгоритма вычисления $\hat{A}$. Оба алгоритма следуют из способа, которым проводились вычисления в нашем примере. Первый способ — самый удобный для счета вручную. Следует помнить, что в алгоритме сложение означает булево сложение.

АЛГОРИТМ УОРШОЛЛА 1

1) Просмотреть столбец 1 матрицы A. Найти строку этого столбца, в которой имеется 1, и прибавить строку 1 к этой строке.
2) Просмотреть столбец 2 матрицы, построенной в пункте (1). Найти строку этого столбца, в которой имеется 1, и прибавить строку 2 к этой строке.
3) Просмотреть столбец 3 матрицы, построенной в пункте (2). Найти строку этого столбца, в которой имеется 1, и прибавить строку 3 к этой строке.
4) Продолжать процесс просмотра следующего по счету столбца матрицы, построенной на предыдущем шаге; искать строку, содержащую 1; прибавлять строку, соответствующую исследуемому столбцу, к строке, в которой имеется 1.
5) Продолжать, пока не будут рассмотрены все столбцы.

Второй метод использует тот факт, что процесс начинается с первой строки и первого столбца, и если 1-цы имеются в i-ой строке первого столбца и в j-ом столбце первой строки, тогда 1-ца помещается в i-ой строке j-го столбца матрицы. Другими словами, если $A_{i1} = 1$ и $A_{1j} = 1$, тогда полагаем $A_{ij} = 1$. Если A_{ij} уже было равно 1, тогда оставляем его равным 1. Это эквивалентно следующему: "для всех i $A_{ij} = A_{ij} \vee (A_{i1} \wedge A_{1j})$, поскольку $A_{i1} \wedge A_{1j}$ равно 1 тогда и только тогда, когда $A_{i1} = 1$ и $A_{1j} = 1$, и равно 0 в противоположном случае". Используя вторую строку и второй столбец, можно было бы получить

новые значения для A_{ij}, положив $A_{ij} = A_{ij} \vee (A_{i2} \wedge A_{2j})$. Продолжая процесс, приходим к следующему алгоритму в псевдокодах.

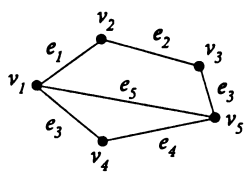
АЛГОРИТМ УОРШОЛЛА 2

```
Цикл по k от 1 до n:
    Цикл по i от 1 до n:
        Цикл по j от 1 до n:
            A_ij = A_ij ∨ (A_ik ∧ A_kj);
        Конец цикла;
    Конец цикла;
Конец цикла.
```

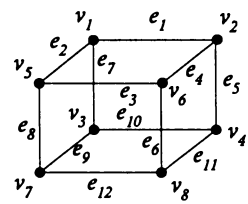
■ УПРАЖНЕНИЯ

1. Найдите матрицы инцидентности следующих графов:

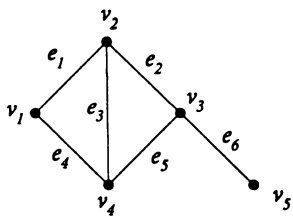
а)



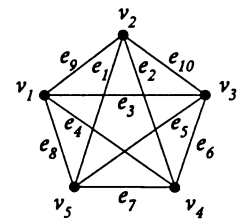
б)



в)

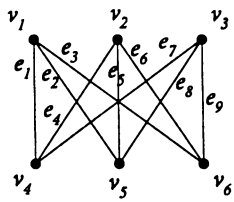


г)

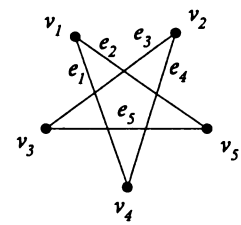


2. Найдите матрицы инцидентности следующих графов:

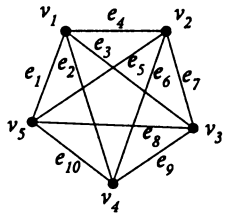
а)



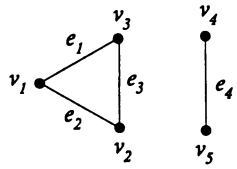
б)



в)



г)



3. Найдите матрицы смежности для графов из упражнения 1.

4. Найдите матрицы смежности для графов из упражнения 2.

5. Для заданной матрицы инцидентности

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

найдите соответствующий граф.

6. Для заданной матрицы инцидентности

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

найдите соответствующий граф.

7. Для заданной матрицы смежности

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

найдите соответствующий граф.

8. Для заданной матрицы смежности

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

найдите соответствующий граф.

9. Для графа, показанного на рис. 6.76,

а) найдите матрицу смежности;

б) используя матрицу смежности, найдите все пути длины 2;

в) используя матрицу смежности, найдите все пути длины 3.

10. Для графа, показанного на рис. 6.77,

а) найдите матрицу смежности;

б) используя матрицу смежности, найдите все пути длины 2;

в) используя матрицу смежности, найдите все пути длины 3.

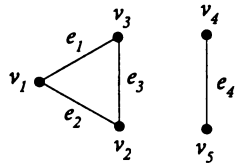


Рис. 6.76

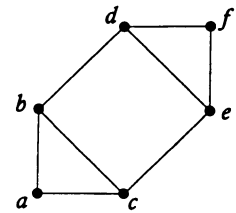


Рис. 6.77

11. Для графа, показанного на рис. 6.78,

а) найдите матрицу смежности;

б) используя матрицу смежности, найдите все пути длины 2;

в) используя матрицу смежности, найдите все пути длины 3.

12. Для графа, заданного на рис. 6.79,

а) найдите матрицу смежности;

б) используя матрицу смежности, найдите все пути длины 2;

в) используя матрицу смежности, найдите все пути длины 3.

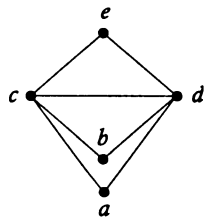


Рис. 6.78

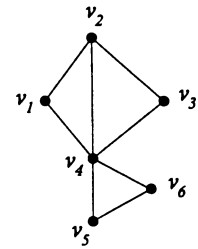


Рис. 6.79

13. Используя тот факт, что $\hat{A} = A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$, определите транзитивное замыкание отношения, представленного графом на рис. 6.80.

14. Используя тот факт, что $\hat{A} = A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$, определите транзитивное замыкание отношения, представленного графом на рис. 6.81.

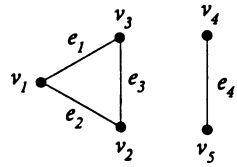


Рис. 6.80

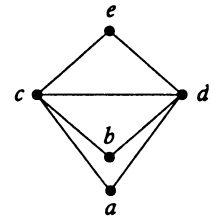


Рис. 6.81

15. Используя тот факт, что $\hat{A} = A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$, определите транзитивное замыкание отношения, представленного ориентированным графом на рис. 6.82.

16. Используя тот факт, что $\hat{A} = A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$, определите транзитивное замыкание отношения, представленного ориентированным графом на рис. 6.83.

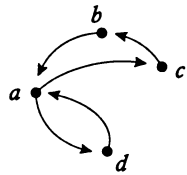


Рис. 6.82

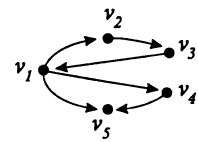


Рис. 6.83

17. Используя алгоритм Уоршолла, определите транзитивное замыкание отношения, представленного графом на рис. 6.84.

18. Используя алгоритм Уоршолла, определите транзитивное замыкание отношения, представленного графом на рис. 6.85.

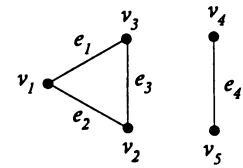


Рис. 6.84

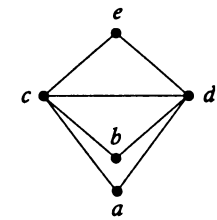


Рис. 6.85

19. Используя алгоритм Уоршолла, определите транзитивное замыкание отношения, представленного графом на рис. 6.86.

20. Используя алгоритм Уоршолла, определите транзитивное замыкание отношения, представленного ориентированным графом на рис. 6.87.

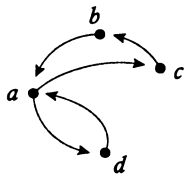


Рис. 6.86

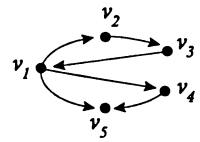


Рис. 6.87

21. Используя алгоритм Уоршолла, определите транзитивное замыкание отношения, представленного ориентированным графом на рис. 6.88.

22. Используя алгоритм Уоршолла, определите транзитивное замыкание отношения, представленного ориентированным графом на рис. 6.89.

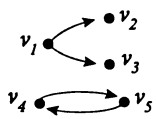


Рис. 6.88

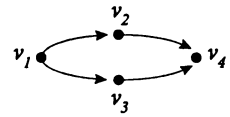


Рис. 6.89

23. Матрицы инцидентности для мультиграфов можно определить таким же способом, как и для обыкновенных графов. Как по матрице инцидентности можно судить о том, имеют ли два ребра мультиграфа общие вершины? (Такие ребра называются кратными.)

24. Если рассматривается матрица инцидентности для мультиграфа, дает ли сумма элементов строки, обозначенной данной вершиной, степень данной вершины? Что означает сумма элементов строки?

25. Используя индукцию, докажите теорему 6.63. Пусть G — (ориентированный) граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Из вершины v_i в вершину v_j тогда и только тогда существует путь длины k, где $1 \leq k \leq n$, когда $A_{ij}^{\odot k} = 1$.

26. Используя индукцию, докажите теорему 6.64. Пусть G — (ориентированный) граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Из вершины v_i в вершину v_j тогда и только тогда существует m путей длины k, где $1 \leq k \leq n$, когда $A_{ij}^{k} = m$.

27. Используя теорему 6.63, докажите теорему 6.65. Пусть G — (ориентированный) граф с вершинами v_1, v_2, v_3, ..., v_n и матрицей смежности A. Если $\hat{A} = A \vee A^{\odot 2} \vee A^{\odot 3} \vee A^{\odot 4} \vee \cdots \vee A^{\odot n}$, то $\hat{A}_{ij} = 1$ тогда и только тогда, когда имеется путь из вершины v_i в вершину v_j.

28. Опишите матрицу смежности для дерева.

29. Опишите транзитивное замыкание для дерева.

6.7. ГИПЕРКУБЫ И КОД ГРЕЯ

ОПРЕДЕЛЕНИЕ 6.68. ***Расстоянием*** между двумя вершинами графа называется длина самого короткого пути между этими двумя вершинами.

ОПРЕДЕЛЕНИЕ 6.69. ***Диаметром*** графа называется наибольшее расстояние между двумя любыми его вершинами.

Во многих случаях вместо использования одиночных процессоров, способных выполнять только одну программу, несколько процессоров соединяется для выполнения программ в параллельном режиме. При этом обеспечивается обмен информацией между процессорами, и различные программы выполняются одновременно. Одним из способов связи компьютеров есть соединение их сериями в кольцо. Указанный метод обладает тем недостатком, что для передачи информации от одного процессора к другому потребуется прохождение ее через значительное число процессоров. В самой неблагоприятной ситуации информация должна будет пройти через половину процессоров. Незначительное улучшение дает использование сетки, или прямоугольного массива процессоров, когда они помещены в каждой точке сетки. Пример такой сетки показан на рис. 6.90.

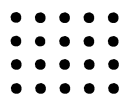


Рис. 6.90

Таким способом достигается лишь некоторое улучшение, так как прохождение информации через значительное число процессоров остается необходимым. В случае сетки 4×5, показанной выше, может возникнуть необходимость прохождения информации через восемь процессоров, включая конечные процессоры. В общем случае сетки $m \times n$ возможна ситуация прохождения информации через $m + n - 1$ процессор.

Намного лучшей конфигурацией является гиперкуб. В частности, ***n-гиперкуб*** может быть использован для соединения вплоть до 2^n компьютеров. Граф n-гиперкуба строится рекурсивно следующим образом. Для $n = 1$ одну вершину представляем посредством 1, а другую — посредством 0, так что получаем граф, изображенный на рис. 6.91.



Рис. 6.91

Таким образом, вершины графа состоят из всех строк длины 1, содержащих 0 и 1. Для $n = 2$ вершины представляем посредством 11, 10, 01 и 00, так что получаем граф, изображенный на рис. 6.92.

Таким образом, вершины представлены строками длины 2, содержащими 0 и 1. Для $n = 3$ вершины представлены посредством 111, 110, 101, 100, 011, 010, 001, 000, так что получаем граф, изображенный на рис. 6.93.

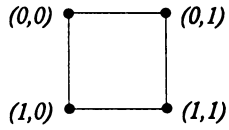


Рис. 6.92

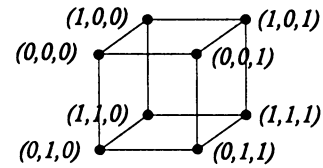


Рис. 6.93

Таким образом, вершины этого графа описывают строки длины 3, состоящие из 0 и 1. Заметим, что две вершины являются смежными, если одна переходит в другую при замене одного символа в строке.

В главе 1 при построении таблиц истинности был составлен список всех возможных комбинаций утверждений.

Если использовать систему обозначений булевой алгебры и заменить T на 1 и F на 0, то обнаружим, что образованы вершины гиперкубов порядка 1, 2, 3 и 4, которые в сокращенном виде можно записать следующим образом:

порядок 1	порядок 2	порядок 3	порядок 4
1	11	111	1111
0	10	110	1110
	01	101	1101
	00	100	1100
		011	1011
		010	1010
		001	1001
		000	1000
			0111
			0110
			0101
			0100
			0011
			0010
			0001
			0000

Что означает для этих точек быть смежными, уже было определено для $n = 2$, 3 и 4. Рассмотрим следующую карту Карно:

	q	$\sim q$
p	11	10
$\sim p$	01	00

Нас интересует, какие блоки являются смежными. Если опять заменить T на 1, F на 0 и придать первой переменной значение p, а второй переменной — значение q, тогда внутренние значения в таблице покажут, где находятся соседние

блоки. Следовательно, каждая вершина является смежной к другой вершине, если они смежные как вершины куба порядка 2.

Аналогично, для трех переменных, где значения p, q и r даны по порядку, получаем таблицу

	q	q	$\sim q$	$\sim q$
p	111	110	100	101
$\sim p$	011	010	000	001
	r	$\sim r$	$\sim r$	r

Вспоминая теперь, что в случае карты Карно таблица скручена так, что два края рассматриваются как смежные, приходим к выводу, что две вершины являются смежными, если они смежные как вершины куба порядка 3.

Рассматривая карту Карно для четырех переменных и присваивая переменным значения p, q, r и s по порядку, аналогичным образом получаем

	q	q	$\sim q$	$\sim q$	
p	1111	1101	1001	1011	s
p	1110	1100	1000	1010	$\sim s$
$\sim p$	0110	0100	0000	0010	$\sim s$
$\sim p$	0111	0101	0001	0011	s
	r	$\sim r$	$\sim r$	r	

Вспоминая, что в случае карты Карно таблица скручена так, что два края рассматриваются как смежные, и смежными являются также верх и низ таблицы, снова приходим к выводу, что две вершины являются смежными в таблице только тогда, когда они смежные как вершины куба порядка 4.

Используя приведенный выше метод, построим последовательность вершин для $k+1$-мерного куба, используя последовательность вершин k-мерного куба следующим образом:

1) Поместить 1 перед каждой вершиной в последовательности вершин k-мерного куба. Вершины, которые были смежными в k-мерном кубе, с приставленной единицей остаются смежными в $k+1$-мерном кубе.

2) Поместить 0 перед каждой вершиной в последовательности вершин k-мерного куба. Вершины, которые были смежными в k-мерном кубе, с приставленным нулем остаются смежными в $k+1$-мерном кубе.

3) Поместить последовательность, построенную в (2), вслед за последовательностью, построенной в (1).

Нами предложен метод для построения вершин гиперкубов, причем для $n = 1, 2, 3$ и 4 можно использовать карты Карно, чтобы показать, когда вершины являются смежными. При $n > 4$ для образования вершин можно использовать приведенные выше шаги (1)–(3).

Предположим, что имеется вращающийся диск, разделенный на секторы, и серия щеток или лазерных лучей, возвращающих информацию о том, на какой

угол диск успел повернуться на текущий момент. Если двоичные строки, которыми пронумерованы соседние секторы, существенно различны в том смысле, что при переходе от одного сектора к другому многие цифры в них изменяются, то устройство, считывающее эту информацию непосредственно в процессе вращения, может выработать число, совершенно не похожее ни на одно из чисел. В данном случае желательно пронумеровать секторы так, чтобы двоичные строки, нумерующие соседние секторы, различались только одной цифрой.

В общем случае, однако, вершины в приведенном выше списке не являются смежными. Можно, тем не менее, изменить ситуацию, внеся незначительные исправления в приведенной выше части (2). Для формирования списка вершин $k+1$-мерного куба, до того как помещать 0 перед элементами списка для k-мерного куба, *реверсируем* список вершин для k-мерного куба, т.е. порядок вершин в списке изменим на обратный. Например, при формировании списка вершин 2-мерного куба помещаем 1 перед столбцом для 1-мерного куба:

$$\begin{matrix}1\\0\end{matrix}$$

Затем меняем порядок элементов в столбце:

$$\begin{matrix}0\\1\end{matrix}$$

Наконец, помещаем 0 перед каждым элементом, имея в итоге

$$\begin{matrix}1&1\\1&0\\0&0\\0&1\end{matrix}$$

Чтобы получить вершины 3-мерного куба, помещаем 1 перед приведенным выше списком для 2-мерного куба и помещаем 0 перед реверсированным списком для 2-мерного куба. В итоге для 3-мерного куба получаем

$$\begin{matrix}1&1&1\\1&1&0\\1&0&0\\1&0&1\\0&0&1\\0&0&0\\0&1&0\\0&1&1\end{matrix}$$

Легко доказать, что приведенная процедура всегда будет давать последовательность вершин для k-мерного куба, которую назовем ***k*-списком**. В таком k-списке (1) каждая вершина последовательности является смежной для следующей вершины и (2) первая вершина последовательности является смежной к последней вершине последовательности. Используя индукцию, начнем с констатации

того, что последовательность вершин для 1-мерного куба очевидным образом обладает перечисленными свойствами. Предположим, что k-список для k-мерного куба обладает этими свойствами. Если поместить 1 перед каждым элементом k-списка вершин k-мерного куба, то каждая вершина последовательности будет очевидным образом смежной к последующей. Также при реверсировании, т.е. при изменении на обратный порядка вершин в k-списке для k-мерного куба, каждая вершина остается смежной к последующей вершине. И после помещения 0 перед каждым элементом реверсированного списка вершин каждая вершина в новом списке остается смежной к последующей. Первый элемент реверсированного списка для k-мерного куба совпадает с последним элементом исходного k-списка для k-мерного куба, поэтому когда первый элемент реверсированного списка с 0 впереди следует за последним элементом исходного k-списка с 1 впереди, они отличаются только первой цифрой и, следовательно, являются смежными. Аналогичным образом, последний элемент в реверсированном k-списке равен первому элементу в k-списке, так что при помещении 0 перед последним элементом в реверсированном k-списке и помещении 1 перед первым элементом k-списка эти две вершины получаются смежными. Таким образом, рассматриваемый $k+1$-список обладает требуемыми свойствами. Эта последовательность, построенная для n-мерного куба, называется ***кодом Грея*** для n.

Итак, правила построения кода Грея для $k+1$ состоят в следующем:

1) Поместите 1 перед каждой вершиной в k-списке k-мерного куба. Вершины, смежные в k-мерном кубе, с приставленной впереди 1 остаются смежными в $k+1$-мерном кубе.

2) Поместите 0 перед каждой вершиной в реверсированном k-списке k-мерного куба. Вершины, смежные в k-мерном кубе, с приставленным впереди 0 остаются смежными в $k+1$-мерном кубе.

3) Разместите последовательность, сформированную в (2), после последовательности, сформированной в (1).

4) Каждая последовательная пара вершин в $k+1$-мерном списке $k+1$-мерного куба является смежной. Первая вершина $k+1$-списка также является смежной к последней вершине списка.

Например, 3-список и реверсированный 3-список представляют собой, соответственно,

1	1	1
1	1	0
1	0	0
1	0	1
0	0	1
0	0	0
0	1	0
0	1	1

и

0	1	1
0	1	0
0	0	0
0	0	1
1	0	1
1	0	0
1	1	0
1	1	1

Заметим, что первый элемент каждого списка равен последнему элементу другого списка. Дописывая 1 перед каждой вершиной в первом списке и 0 перед каждой вершиной во втором списке и помещая второй список в конце первого списка,

получаем

1 1 1 1
1 1 1 0
1 1 0 0
1 1 0 1
1 0 0 1
1 0 0 0
1 0 1 0
1 0 1 1
0 0 1 1
0 0 1 0
0 0 0 0
0 0 0 1
0 1 0 1
0 1 0 0
0 1 1 0
0 1 1 1

Таким образом, мы построили код Грея для $n = 4$.

Ранее упоминалось соединение компьютеров в конфигурацию сетки или решета. Под сеткой понимают граф, вершины которого заданы массивом размерности $m \times n$, и для которого две вершины, соседствующие в одной и той же строке или столбце массива, являются смежными как вершины графа. Возможно ли для $m \leq 2^k$ и $n \leq 2^l$ построить подграф $k + l$-мерного куба, т.е. сетку размера $m \times n$? Это уже делалось при построении карт Карно. Указанное легко сделать, обозначая строки согласно первым m элементам кода Грея для k и обозначая столбцы согласно первым n элементам кода Грея для l. Элемент на позиции (i, j) сетки есть метка i-ой строки, за которой следует метка j-го столбца. Таким образом, если необходимо построить сетку 3×7, то она будет иметь вид

	111	**110**	**100**	**101**	**001**	**000**	**010**
11	11111	11110	11100	11101	11001	11000	11010
10	10111	10110	10100	10101	10001	10000	10010
00	00111	00110	00100	00101	00001	00000	00010

Здесь (i, j)-ый элемент сетки является (i, j)-ым элементом таблицы. Проведенные рассмотрения доказывают следующую теорему.

ТЕОРЕМА 6.70. Каждая сетка размера $m \times n$ представляет собой подграф $i + j$-мерного куба, где $m \leq 2^i$ и $n \leq 2^j$.

Предоставляем читателю доказать следующую теорему.

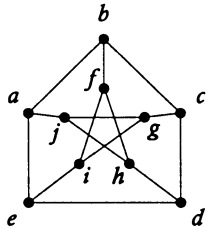
ТЕОРЕМА 6.71. Каждый гиперкуб для $n \geq 1$ является двудольным графом, в котором непересекающиеся множества вершин состоят из множества тех вершин, которые изображаются строками, содержащими четное число единиц, и множества вершин, которые изображаются строками, содержащими нечетное число единиц.

■ УПРАЖНЕНИЯ

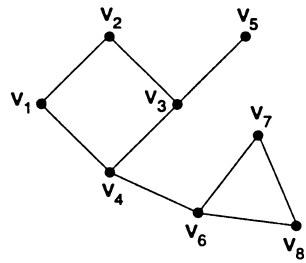
1. Найдите диаметр следующих графов:

а) K_n **б)** $K_{m,n}$

в)

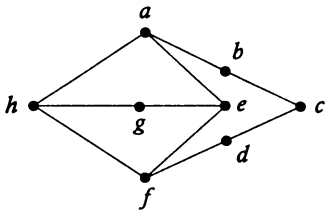


г)

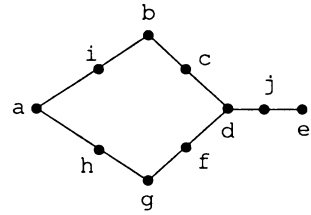


2. Найдите диаметр следующих графов:

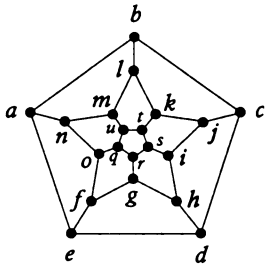
а)



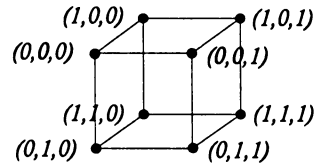
б)



в)



г)



д) 4-гиперкуб.

3. Постройте код Грея для 5.

4. Постройте код Грея для 6.

5. Постройте сетку 4×5.

6. Постройте сетку 3×8.

7. Постройте сетку 5×6.

8. Докажите теорему 6.71. Каждый гиперкуб для $n \geq 1$ является двудольным графом, в котором непересекающиеся множества вершин состоят из мно-

жества тех вершин, которые изображаются строками, содержащими четное число единиц, и множества вершин, которые изображаются строками, содержащими нечетное число единиц.

9. Докажите, что если $m = 2^i$ и $n = 2^j$, то построенная сетка $m \times n$ обладает свойством, что соответствующие вершины в первой и последней строке сетки являются смежными, и соответствующие вершины в первом и последнем столбце являются смежными.

ГЛАВА

7

ТЕОРИЯ ЧИСЕЛ

7.1. РЕШЕТО ЭРАТОСФЕНА

В предыдущей главе отмечалось, что ***алгоритм*** представляет собой конечное множество правил для чисто механического решения задач. Первым примером алгоритма в теории чисел будет древний метод нахождения простых чисел, называемый "решетом Эратосфена", который представляет собой алгоритм определения простых чисел, меньших заданного целого числа. Проиллюстрируем этот метод, определяя простые числа в интервале от 1 до 100. Прежде всего перечислим целые числа от 1 до 100:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20,
21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40,
41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60,
61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80,
81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100.

Начиная с 2, которая является первым простым числом, выделим жирным шрифтом все числа, кратные 2. Эта процедура очень проста, так как отмеченным должно быть каждое второе целое число, больше 2.

1, ***2***, 3, ***4***, 5, ***6***, 7, ***8***, 9, ***10***, 11, ***12***, 13, ***14***, 15, ***16***, 17, ***18***, 19, ***20***,
21, ***22***, 23, ***24***, 25, ***26***, 27, ***28***, 29, ***30***, 31, ***32***, 33, ***34***, 35, ***36***, 37, ***38***, 39, ***40***,
41, ***42***, 43, ***44***, 45, ***46***, 47, ***48***, 49, ***50***, 51, ***52***, 53, ***54***, 55, ***56***, 57, ***58***, 59, ***60***,
61, ***62***, 63, ***64***, 65, ***66***, 67, ***68***, 69, ***70***, 71, ***72***, 73, ***74***, 75, ***76***, 77, ***78***, 79, ***80***,
81, ***82***, 83, ***84***, 85, ***86***, 87, ***88***, 89, ***90***, 91, ***92***, 93, ***94***, 95, ***96***, 97, ***98***, 99, ***100***

Следующее простое число равно 3. Продолжая процедуру, отметим все числа, кратные 3. Эта процедура также является простой, так как должно быть отмечено каждое третье целое число, больше 3.

1, **2**, 3, **4**, 5, **6**, 7, **8**, **9**, **10**, 11, **12**, 13, **14**, **15**, **16**, 17, **18**, 19, **20**, **21**, **22**, 23, **24**, 25, **26**, **27**, **28**, 29, **30**, 31, **32**, **33**, **34**, 35, **36**, 37, **38**, **39**, **40**, 41, **42**, 43, **44**, **45**, **46**, 47, **48**, 49, **50**, **51**, **52**, 53, **54**, 55, **56**, **57**, **58**, 59, **60**, 61, **62**, **63**, **64**, 65, **66**, 67, **68**, **69**, **70**, 71, **72**, 73, **74**, **75**, **76**, 77, **78**, 79, **80**, **81**, **82**, 83, **84**, 85, **86**, **87**, **88**, 89, **90**, 91, **92**, **93**, **94**, 95, **96**, 97, **98**, **99**, **100**

Теперь отмечаем все числа, кратные простому числу 5.

1, **2**, 3, **4**, **5**, **6**, 7, **8**, **9**, **10**, 11, **12**, 13, **14**, **15**, **16**, 17, **18**, 19, **20**, **21**, **22**, 23, **24**, **25**, **26**, **27**, **28**, 29, **30**, 31, **32**, **33**, **34**, **35**, **36**, 37, **38**, **39**, **40**, 41, **42**, 43, **44**, **45**, **46**, 47, **48**, 49, **50**, **51**, **52**, 53, **54**, **55**, **56**, **57**, **58**, 59, **60**, 61, **62**, **63**, **64**, **65**, **66**, 67, **68**, **69**, **70**, 71, **72**, 73, **74**, **75**, **76**, 77, **78**, 79, **80**, **81**, **82**, 83, **84**, **85**, **86**, **87**, **88**, 89, **90**, 91, **92**, **93**, **94**, **95**, **96**, 97, **98**, **99**, **100**

Далее отмечаем все числа, кратные простому числу 7.

1, **2**, 3, **4**, **5**, **6**, 7, **8**, **9**, **10**, 11, **12**, 13, **14**, **15**, **16**, 17, **18**, 19, **20**, **21**, **22**, 23, **24**, **25**, **26**, **27**, **28**, 29, **30**, 31, **32**, **33**, **34**, **35**, **36**, 37, **38**, **39**, **40**, 41, **42**, 43, **44**, **45**, **46**, 47, **48**, **49**, **50**, **51**, **52**, 53, **54**, **55**, **56**, **57**, **58**, 59, **60**, 61, **62**, **63**, **64**, **65**, **66**, 67, **68**, **69**, **70**, 71, **72**, 73, **74**, **75**, **76**, **77**, **78**, 79, **80**, **81**, **82**, 83, **84**, **85**, **86**, **87**, **88**, 89, **90**, **91**, **92**, **93**, **94**, **95**, **96**, 97, **98**, **99**, **100**

Поскольку 7 — наибольшее простое число, квадрат которого меньше или равен 100, по теореме 3.41 продолжать процесс нет необходимости. Числа, большие 1 и не отмеченные жирным шрифтом, представляют собой простые числа, которые не превышают 100.

■ УПРАЖНЕНИЯ

1. Постройте решето Эратосфена для положительных целых чисел, меньших 200.
2. Используйте решето Эратосфена и калькулятор для разложения на простые множители числа 1726.
3. Используйте решето Эратосфена и калькулятор для разложения на простые множители числа 481.
4. Используйте решето Эратосфена и калькулятор для разложения на простые множители числа 2502.
5. Используйте решето Эратосфена и калькулятор для разложения на простые множители числа 1739.
6. Используйте решето Эратосфена и калькулятор для разложения на простые множители числа 391.
7. Используйте решето Эратосфена и калькулятор для разложения на простые множители числа 3901.

7.2. МЕТОД ВЫДЕЛЕНИЯ МНОЖИТЕЛЕЙ ФЕРМА

Следущая теорема является основой алгоритма разложения на простые множители, названного ***методом выделения множителей Ферма***. Метод используется для определения того, является ли число простым.

ТЕОРЕМА 7.1. Целое нечетное число $n > 1$ не является простым тогда и только тогда, когда существуют неотрицательные целые числа p и q такие, что $n = p^2 - q^2$, и при этом $p - q > 1$.

Очевидно, если n можно представить как разность квадратов двух неотрицательных целых чисел, скажем, $n = p^2 - q^2$, тогда $n = (p - q)(p + q)$. Поскольку $p - q > 1$, то также $p + q > 1$; и n не является простым числом.

Обратно, если $n = rs$, где $r \geq s > 1$, тогда n можно представить как $((r+s)/2)^2 - ((r-s)/2)^2$. Поскольку n нечетное, r и s также являются нечетными и, следовательно, $r+s$ и $r-s$ — четные числа. Положив $p = (r+s)/2$ и $q = (r-s)/2$, находим, что p и q — целые неотрицательные числа и $p - q = s > 1$. При $n = 1$ положим $p = 1$, а $q = 0$. ■

Применение этого метода состоит в попытке найти целые числа p и q такие, что $n = p^2 - q^2$ или, что эквивалентно, $p^2 = n + q^2$, или $q^2 = p^2 - n$. Если используется первое уравнение, полагаем $q = 1, 2, \ldots$ до тех пор, пока $n + q^2$ не станет полным квадратом. Если до значения $q = (n - 1)/2$ полный квадрат не достигнут, рассмотрим ситуацию, когда $q = (n-1)/2$, что дает $n+q^2 = ((n+1)/2)^2$ и приводит к разложению n на множители $n \cdot 1$. Поскольку q имеет вид $(r - s)/2$, где r и s — делители n, то очевидно, что q не может превысить $(n - 1)/2$. Следовательно, если до значения $q = (n - 1)/2$ полный квадрат не достигнут, число n является простым.

При использовании второго уравнения, т.е. $q^2 = p^2 - n$, берем в качестве m наименьшее целое число такое, что $m^2 \geq n$, и последовательно полагаем, $p = m, m+1, \ldots$ до тех пор, пока $p^2 - n$ не станет полным квадратом. Как и прежде, q не может превысить $(n - 1)/2$, так что если до значения $p = (n + 1)/2$ полный квадрат не получен, число n является простым. Преимущество использования второго соотношения состоит в том, что проверке на полный квадрат подлежит меньшее количество чисел.

Например, рассмотрим применение записи $p^2 = n+q^2$ для проверки, является ли простым число $n = 527$. Рассмотрим $q = 1, 2, \ldots, (n - 1)/2$.

q	$n + q^2$
1	$527 + 1 = 528$
2	$527 + 4 = 531$
3	$527 + 9 = 536$
4	$527 + 16 = 543$
5	$527 + 25 = 552$
6	$527 + 36 = 563$
7	$527 + 49 = 576 = (24)^2$

Итак, $n = 527$ является составным, и его делители легко вычисляются:

$$\begin{aligned} 527 &= (24)^2 - 7^2 = \\ &= (24-7)(24+7) = \\ &= 17 \cdot 31. \end{aligned}$$

Решение вопроса о том, является ли число n простым, не всегда требует проверки всех значений q от 1 до $(n-1)/2$.

■ УПРАЖНЕНИЯ

Воспользовавшись методом выделения множителей Ферма, определите, являются ли следующие числа простыми.

1. 1001
2. 1349
3. 4851
4. 1079
5. 8051
6. 567
7. 7931

7.3. АЛГОРИТМЫ ДЕЛЕНИЯ И АЛГОРИТМ ЕВКЛИДА

Ранее была доказана теорема, названная ***алгоритмом деления***. Теорема гласила, что для положительных целых чисел a и b существуют единственные неотрицательные числа q и r такие, что $0 \leq r < b$ и $a = bq + r$. Это утверждение не является алгоритмом, но алгоритмы для нахождения q и r существуют. Если числа a и b велики, число q можно "угадать". Если умножить b на q и результат слишком велик, следует q заменить на $q-1$ и повторять процесс, пока не получим $bq \leq a$, после чего положить $r = a - bq$. Однако, если $bq < a$, но $a - bq \geq b$, то следует заменить q на $q+1$ и повторять процесс, пока не получим $a - bq < b$, после чего положить $r = a - bq$. Более системно метод описывается следующим образом:

1. Положить $q = 1$ и $r = a - bq$.
2. Если $r \geq b$, положить $q = q + 1$ и $r = a - bq$.
3. Продолжать процесс, пока $r < b$.

Изложенное представляет собой алгоритм нахождения наибольшего общего делителя больших чисел, названный алгоритмом Евклида. Нахождение наибольшего общего делителя необходимо при сложении дробей, а также при решении уравнений в целых числах.

ТЕОРЕМА 7.2. (Алгоритм Евклида) Допустим, что a и b — положительные целые числа, и последовательное применение алгоритма деления приводит к последовательности следующего вида:

$$\begin{aligned}
a &= bq_0 + r_0 & 0 \le r_0 < b \\
b &= r_0 q_1 + r_1 & 0 \le r_1 < r_0 \\
r_0 &= r_1 q_2 + r_2 & 0 \le r_2 < r_1 \\
r_1 &= r_2 q_3 + r_3 & 0 \le r_3 < r_2 \\
r_2 &= r_3 q_4 + r_4 & 0 \le r_4 < r_3 \\
&\vdots & \vdots \\
r_k &= r_{k+1} q_{k+2} + r_{k+2} & 0 \le r_{k+2} < r_{k+1} \\
&\vdots & \vdots
\end{aligned}$$

Существует $r_k = 0$. Пусть s — первое целое число такое, что $r_s = 0$. Тогда $r_{s-1} = \text{НОД}(a, b)$, если $s > 0$, и $b = \text{НОД}(a, b)$, если $s = 0$.

ДОКАЗАТЕЛЬСТВО. Пусть $S = \{r_0, r_1, \ldots\}$. Если $r_0 = 0$, тогда результат очевиден, поэтому предположим, что $r_0 > 0$. Согласно принципу полного упорядочения множество S содержит наименьшее положительное целое число, скажем, r_i. Но $0 \le r_{i+1} < r_i$, и поэтому $r_{i+1} = 0$. Положим $s = i + 1$.

По теореме 3.33,

$$\text{НОД}(a, b) = \text{НОД}(b, r_0) = \text{НОД}(r_0, r_1) = \cdots = \text{НОД}(r_{s-1}, r_s),$$

но поскольку $r_s = 0$, $\text{НОД}(r_{s-1}, r_s) = r_{s-1}$, так что $\text{НОД}(a, b) = r_{s-1}$. ■

ПРИМЕР 7.3. Используя алгоритм Евклида для нахождения НОД(203, 91), действуем следующим образом. Сначала делим 203 на 91 нацело и получаем

$$\begin{aligned}
203 &= 91 \cdot 2 + 21; \\
a &= b \cdot q_0 + r_0.
\end{aligned}$$

Теперь делим 91 на остаток 21 нацело:

$$\begin{aligned}
91 &= 21 \cdot 4 + 7; \\
b &= r_0 \cdot q_1 + r_1.
\end{aligned}$$

Разделяя 21 на 7 нацело, получаем

$$\begin{aligned}
21 &= 7 \cdot 3 + 0; \\
r_0 &= r_1 \cdot q_2 + r_2.
\end{aligned}$$

Итак, $s = 2$ и $\text{НОД}(203, 91) = \text{НОД}(91, 21) = \text{НОД}(21, 7) = 7 = r_{s-1} = r_1$. □

ПРИМЕР 7.4. Найти НОД(99, 205). Действуя вышеуказанным образом, делим 205 на 99 нацело, получая

$$205 = 99 \cdot 2 + 7.$$

Теперь делим 99 нацело на остаток 7, получая

$$99 = 7 \cdot 14 + 1.$$

Далее делим 7 на 1 и получаем

$$7 = 1 \cdot 7 + 0.$$

Таким образом, НОД(99, 205) = 1. □

Алгоритм Евклида позволяет находить такие значения u и v, что НОД$(a, b) = r_t$, $au + vb =$ НОД(a, b). Используя обозначения теоремы 7.2, запишем $r_{t-2} = r_{t-1} \cdot q_t + r_t$. Отсюда $r_t = r_{t-2} - r_{t-1} \cdot q_t$. Аналогично, $r_{t-1} = r_{t-3} - r_{t-2} \cdot q_{t-1}$. Подстановка в предыдущее уравнение дает

$$r_t = r_{t-2} - (r_{t-3} - r_{t-2} \cdot q_{t-1}) \cdot q_t.$$

Аналогичным образом можно выразить r_{t-2} через r_{t-3} и r_{t-4} и, подставив в предыдущее уравнение, исключить r_{t-2}. Продолжая в том же духе, можно исключить все r_i и вернуться обратно к a и b, получив НОД(a, b) в виде $ua + vb$.

ПРИМЕР 7.5. Выразить НОД(85, 34) в виде $85u + 34v$. Для использования алгоритма Евклида сначала разделим 85 на 34 нацело:

$$85 = 34 \cdot 2 + 17.$$

Разделяя 34 на 17 нацело, получаем

$$34 = 17 \cdot 1 + 0.$$

Таким образом, НОД(85, 34) = 17 и НОД(85, 34) = 17 = (85)(1) + (34)(−2). □

ПРИМЕР 7.6. Выразить НОД(252, 580) в виде $252u + 580v$. Делим 580 на 252 нацело, получая

$$580 = 252 \cdot 2 + 76;$$
$$a = b \cdot q_0 + r_0.$$

Теперь делим 252 нацело на 76:

$$252 = 76 \cdot 3 + 24;$$
$$b = r_0 \cdot q_1 + r_1.$$

Продолжая процесс, получаем

$$76 = 24 \cdot 3 + 4;$$
$$r_0 = r_1 \cdot q_2 + r_2$$

и

$$24 = 4 \cdot 6 + 0;$$
$$r_1 = r_2 \cdot q_3 + r_3.$$

Обратная подстановка дает

$$\begin{aligned} 4 &= \underline{76} - \underline{24} \cdot 3 = \\ &= \underline{76} - [\underline{252} - \underline{76} \cdot 3] \cdot 3 = \\ &= \underline{76} \cdot 10 + \underline{252} \cdot (-3) = \\ &= [\underline{580} - \underline{252} \cdot 2] \cdot 10 + \underline{252} \cdot (-3) = \\ &= \underline{580} \cdot 10 + \underline{252} \cdot (-23), \end{aligned}$$

где r_i выделены подчеркиванием. □

ПРИМЕР 7.7. Выразить НОД$(252, 576)$ в виде $252u + 576v$. Сначала разделим 576 на 252 нацело:

$$576 = 252 \cdot 2 + 72.$$

Теперь разделим 252 нацело на 72:

$$252 = 72 \cdot 3 + 36.$$

После обратной подстановки получаем

$$\begin{aligned} \underline{36} &= \underline{252} - \underline{72} \cdot 3 = \\ &= \underline{252} - [\underline{576} - \underline{252} \cdot 2] \cdot 3 = \\ &= (7)(\underline{252}) + (-3)(\underline{576}). \end{aligned}$$ □

В примере 7.3 показано, что НОД$(91, 203) = 7$. Если разделить числа 91 и 203 на их наибольший общий делитель 7, получим $\frac{91}{7} = 13$ и $\frac{203}{7} = 29$. Легко видеть, что целые числа 13 и 29 — взаимно простые; т.е. НОД$(13, 29) = 1$. Таким образом, оказалось, что результаты деления двух целых чисел на их наибольший общий делитель не имеют общих делителей, за исключением 1. Доказательство следующей теоремы предоставляем читателю.

ТЕОРЕМА 7.8. Пусть заданы целые числа a и b, не равные нулю, тогда числа $\frac{a}{\text{НОД}(a,b)}$ и $\frac{b}{\text{НОД}(a,b)}$ являются взаимно простыми, т.е. НОД $\left(\frac{a}{\text{НОД}(a,b)}, \frac{b}{\text{НОД}(a,b)}\right) = 1$.

Было определено ранее, что положительное целое число m представляет собой ***наименьшее общее кратное*** целых чисел a и b при условии, что (а) $a \mid m$ и $b \mid m$, и (б) если n — любое общее кратное чисел a и b, то $m \mid n$. Наименьшее общее кратное чисел a и b обозначим НОК(a, b).

Необходимо построить алгоритм нахождения наименьшего общего кратного двух целых чисел. Цель достигается определением наибольшего общего делителя этих чисел (алгоритм нахождения которого уже имеется) и использованием теоремы 3.50. Указанная теорема утверждает, что для положительных целых чисел a и b НОД$(a, b) \cdot$ НОК$(a, b) = ab$.

Для нахождения НОК$(91, 203)$ сначала определим НОД$(91, 203)$, воспользовавшись алгоритмом Евклида, как это было сделано ранее, а затем разделим на него произведение чисел 91 и 203. Поскольку НОД$(91, 203) = 7$, находим

$$\text{НОК}(91, 203) = \frac{(91)(203)}{7} = 2639.$$

■ УПРАЖНЕНИЯ

1. Задан алгоритм деления $a = bq + r$. Найдите q и r для указанных ниже значений a и b:
   а) $a = 75$, $b = 8$; **б)** $a = 102$, $b = 5$;
   в) $a = 81$, $b = 9$; **г)** $a = 16$, $b = 25$.
2. Найдите наибольший общий делитель для следующих пар чисел:
   а) 75, 25; **б)** 27, 18;
   в) 621, 437; **г)** 289, 377;
   д) 822, 436.
3. Найдите наименьшее общее кратное для пар чисел из упражнения 2.
4. Для приведенных ниже пар чисел a и b найдите u, v и d такие, что $au + bv = d$, где d — наибольший общий делитель чисел a и b:
   а) $83, 17$; **б)** $361, 418$;
   в) $25, 15$; **г)** $81, 9$;
   д) $216, 324$.
5. Докажите, что если НОД(a, b) выражен как $ax + by$, то x и y — взаимно простые числа.
6. Докажите, что для целых чисел a и b, не равных нулю, $\frac{a}{\text{НОД}(a,b)}$ и $\frac{b}{\text{НОД}(a,b)}$ взаимно просты, т.е. НОД $\left(\frac{a}{\text{НОД}(a,b)}, \frac{b}{\text{НОД}(a,b)}\right) = 1$ (теорема 7.8).
7. Запишите в псевдокодах процедуру нахождения согласно алгоритму деления величин q и r, когда произвольное положительное целое число a делится на произвольное положительное целое число b.
8. Запишите в псевдокодах процедуру для алгоритма Евклида.
9. Задано целое число a и положительное целое число b. Докажите, что целые числа, произведение которых равно a и наибольший общий делитель которых равен b, существуют тогда и только тогда, когда $b^2 \mid a$.
10. Предполагая, что
$$F(k) = 2^{2^k} + 1 \text{ при } k \geq 0,$$
   а) докажите, что $F(m+1) = F(0)F(1)F(3)\cdots F(m)+2$ для каждого $m \geq 0$;
   б) докажите, что $F(m)$ и $F(n)$ — взаимно простые числа, если $m \neq n$.
11. Покажите, что если НОД(a, b) выражен как $ax+by$, то x и y не определяются единственным образом.
12. Докажите, что если НОД$(a, b) = 1$, то НОД$(a + nb, b) = 1$.
13. Докажите, что $n \cdot$ НОД$(a, b) =$ НОД(na, nb).
14. Пусть S — множество всех целых чисел вида $ax+by$. Покажите, что НОД(a, b) делит все элементы множества S.

7.4. ЦЕПНЫЕ ДРОБИ

В данной главе алгоритм деления и алгоритм Евклида будут применяться достаточно часто. Рассмотрим сначала рациональные числа a/b, где a и b — целые числа и $b > 0$. Если алгоритм деления применен к целым числам a и b, где $b > 0$, то существуют единственные целые числа t (частное) и r (остаток) такие, что

$$a = b \cdot t + r \qquad \text{при} \qquad 0 \leq r < b.$$

Для $a = -124$ и $b = 35$ имеем

$$-124 = (35)(-4) + 16.$$

Переписав последние два равенства и перейдя к рациональным числам, получаем

$$\frac{a}{b} = t + \frac{r}{b} \qquad \text{и} \qquad \frac{-124}{35} = -4 + \frac{16}{35}\ .$$

Воспользовавшись функцией целой части числа, отношения можем переписать в таком виде

$$\left\lfloor \frac{a}{b} \right\rfloor = t \qquad \text{и} \qquad \left\lfloor \frac{-124}{35} \right\rfloor = -4\ ,$$

при этом дробный остаток r/b обладает свойством $0 \leq r/b < 1$. Если $r \neq 0$, равенства можно переписать следующим образом:

$$\frac{a}{b} = t + \frac{1}{b/r} \qquad \text{и} \qquad \frac{-124}{35} = -4 + \frac{1}{35/16}\ ,$$

где $b/r > 1$. Алгоритм деления можно применить снова и получить $b = rt' + r'$ и $35 = (16)(2) + 3$. Таким образом, обращая дробный остаток и применяя алгоритм деления снова и снова, указанным способом получаем выражение рационального числа $-124/35$:

$$\begin{aligned}\frac{-124}{35} &= -4 + \frac{16}{35} = -4 + \frac{1}{\left(\dfrac{35}{16}\right)} = \\ &= -4 + \frac{1}{2 + \dfrac{3}{16}} = -4 + \frac{1}{2 + \dfrac{1}{\left(\dfrac{16}{3}\right)}} = \\ &= -4 + \frac{1}{2 + \dfrac{1}{\left(5 + \dfrac{1}{3}\right)}}.\end{aligned}$$

Поскольку дробь $1/3$ имеет числителем 1, нельзя далее сокращать $3/1$, используя алгоритм деления. Выражение

$$-4 + \frac{1}{2 + \dfrac{1}{5 + \dfrac{1}{3}}}$$

называется цепной дробью и представляет собой альтернативный способ записи рационального числа $-124/35$. В связи с тем, что цепная дробь имеет регулярную структуру и числители всегда равны 1, для точного задания цепной дроби необходимо упомянуть только числа -4, 2, 5 и 3, поэтому записываем $-124/35 = [-4; 2, 5, 3]$. Числа -4, 2, 5 и 3 называются элементами цепной дроби или неполными частными, поскольку порождены алгоритмом деления. Точка с запятой отмечают особую природу первого члена -4. Например, $[3; 7] = 3 + 1/7 = 22/7$, но

$$[0; 3, 7] = 0 + \frac{1}{3 + \frac{1}{7}} = 7/22.$$

Возвращаясь к цепной дроби $-124/35 = [-4; 2, 5, 3]$, замечаем, что $3 = (3 - 1) + 1 = (3 - 1) + 1/1$. Таким образом, можно записать

$$\frac{-124}{35} = -4 + \cfrac{1}{2 + \cfrac{1}{5 + \cfrac{1}{2 + \frac{1}{1}}}},$$

так что запись $-124/35 = [-4; 2, 5, 2, 1]$ тоже имеет место, однако в ней на последнем шаге алгоритм деления не использован. Пример подсказывает, что каждое рациональное число имеет два различных представления в виде цепной дроби с целыми числами в качестве элементов: в одном представлении последний элемент больше 1, а в другом — равен 1.

Если x — действительное, но не рациональное число, алгоритм деления не применим; однако, используя функцию целой части, можно получить представление x в виде цепной дроби длины n способом, аналогичным рациональному случаю. Таким образом, для действительного, но не рационального числа x имеется единственное число t_0, где $t_0 \leq x < t_0 + 1$ такое, что $t_0 = \lfloor x \rfloor$. Тогда

$$1 > y_1 = x - \lfloor x \rfloor = x - t_0 > 0$$

есть так называемая дробная часть x. Если $x_1 = 1/y_1$, тогда

$$x = \lfloor x \rfloor + y_1 = t_0 + y_1 = t_0 + \frac{1}{x_1}$$

и $x_1 > 1$. Пусть $t_1 = \lfloor x_1 \rfloor$, $y_2 = x_1 - \lfloor x_1 \rfloor$ и $x_2 = 1/y_2$, т.к. $y_2 > 0$. Тогда $x_1 = t_1 + \frac{1}{x_2}$, где $x_2 > 0$. Сочетание этих равенств дает

$$x = t_0 + \frac{1}{x_1} = t_0 + \cfrac{1}{t_1 + \frac{1}{x_2}}.$$

Продолжая в том же духе, получаем

$$x = t_0 + \cfrac{1}{t_1 + \cfrac{1}{t_2 + \cfrac{1}{t_3 + \cdots + \cfrac{1}{t_{k-1} + \cfrac{1}{x_k}}}}},$$

где t_i — целое число для каждого i, $t_i > 0$ для $i \geq 1$, x_k — иррациональное действительное число и $x_k > 1$. Можно также записать

$$x = [t_0; t_1, t_2, \ldots, t_{k-1}, x_k].$$

Например, пусть $x = \sqrt{3} = 1.7320508\cdots$. Построим представление числа x в виде цепной дроби следующим образом:

$$t_0 = \lfloor x \rfloor = \left\lfloor \sqrt{3} \right\rfloor = 1,$$

где

$$y_1 = x - \lfloor x \rfloor = x - t_0 = \sqrt{3} - 1;$$
$$x_1 = 1/y_1 = 1/(\sqrt{3} - 1) = (\sqrt{3} + 1)/2;$$
$$t_1 = \lfloor x_1 \rfloor = 1,$$

где

$$y_2 = x_1 - \lfloor x_1 \rfloor = x_1 - t_1 = (\sqrt{3} + 1)/2 - 1 = (\sqrt{3} - 1)/2;$$
$$x_2 = 1/y_2 = 2/(\sqrt{3} - 1) = \sqrt{3} + 1;$$
$$t_2 = \lfloor x_2 \rfloor = 2,$$

где

$$y_3 = x_2 - \lfloor x_2 \rfloor = x_2 - t_2 = (\sqrt{3} + 1) - 2 = \sqrt{3} - 1;$$
$$x_3 = 1/y_3 = 1/(\sqrt{3} - 1) = (\sqrt{3} + 1)/2,$$

таким образом,

$$\sqrt{3} = [t_0; x_1] = [1; (\sqrt{3} + 1)/2] =$$
$$= [t_0; t_1, x_2] = [t_0; [t_1; x_2]] = [1; 1, (\sqrt{3} + 1)] =$$
$$= [t_0; t_1, t_2, x_3] = [t_0; t_1, [t_2; x_3]] = [1; 1, 2, (\sqrt{3} + 1)/2].$$

Но поскольку $x_3 = x_1$, структура повторяется , поэтому

$$\sqrt{3} = [1; 1, 2, 1, 2, 1, 2, (\sqrt{3} + 1)/2]$$

и т.д.

Отсюда выводим следующее рекурсивное определение цепной дроби.

ОПРЕДЕЛЕНИЕ 7.9. Для конечной последовательности $t_0, t_1, t_2, \ldots, t_n$ действительных чисел, где $n \geq 0$ и $t_i > 0$ для $i \geq 1$, определим ***конечную цепную дробь*** $[t_0; t_1, t_2, \ldots, t_n]$ следующим образом:

$$[t_0;] = t_0;$$

$$[t_0; t_1] = t_0 + \frac{1}{t_1};$$

$$[t_0; t_1, t_2, \ldots, t_k]; = [t_0; [t_1; t_2, \ldots, t_k]] \text{ для } 1 < k \leq n.$$

Числа $t_0, t_1, \ldots, t_n$ называются ***неполными частными***, или ***элементами*** цепной дроби. Цепная дробь $[t_0; t_1, t_2, \ldots, t_n]$ называется ***простой***, если t_i — целое число для каждого i, т.е. каждый элемент цепной дроби есть целое число. Числа $[t_0;]$, $[t_0; t_1]$, $[t_0; t_1, t_2]$, ..., $[t_0; t_1, t_2, \ldots, t_k]$, ..., $[t_0; t_1, t_2, \ldots, t_n]$ называются ***подходящими дробями*** цепной дроби $[t_0; t_1, t_2, \ldots, t_n]$. $[t_0; t_1, t_2, \ldots, t_k]$, где есть k-я подходящая дробь при $0 \leq k \leq n$. Для удобства обозначения запись $[t_0; t_1, t_2, \ldots, t_k]$, использованная при $k = 0$, будет означать $[t_0;]$. Будем говорить, что две цепные дроби $[t_0; t_1, t_2, \ldots, t_n]$ и $[b_0; b_1, b_2, \ldots, b_m]$ ***равны почленно***, если $n = m$ и $t_i = b_i$ при $0 \leq i \leq n$. Если x — действительное число и $x = [t_0; t_1, t_2, \ldots, t_n]$, тогда будем говорить, что $[t_0; t_1, t_2, \ldots, t_n]$ есть ***представление цепной дробью*** числа x. Два представления называются ***совпадающими***, или ***равными***, если они равны почленно.

Следующая теорема дает иной способ разложения цепной дроби на подходящие дроби.

ТЕОРЕМА 7.10. Если n — положительное целое число и $[t_0; t_1, t_2, \ldots, t_n]$ — цепная дробь, тогда для каждого k, $1 \leq k \leq n$,

$$[t_0; t_1, t_2, \ldots, t_n] = [t_0; t_1, t_2, \ldots, t_{k-1}, [t_k; t_{k+1}, \ldots, t_n]].$$

ДОКАЗАТЕЛЬСТВО. Равенство докажем методом математической индукции. Если $n = 1$, по определению, $[t_0; t_1] = [t_0; [t_1;]]$. Предположим, что утверждение теоремы выполняется при $n = m$, т.е. для любой цепной дроби $[b_0; b_1, \ldots, b_m]$,

$$[b_0; b_1, b_2, \ldots, b_m] = [b_0; b_1, b_2, \ldots, b_{j-1}, [b_j; b_{j+1}, \ldots, b_m]]$$

для $1 \leq j \leq m$. Рассмотрим $n = m + 1$. Пусть $[t_0; t_1, t_2, \ldots, t_{m+1}]$ — цепная дробь и $1 \leq k \leq m + 1$. Тогда

$$\begin{aligned}[t_0; t_1, \ldots, t_{m+1}] &= [t_0; [t_1; t_2, \ldots, t_{m+1}]] \text{ (по определению)} = \\ &= [t_0; [t_1, t_2, \ldots, t_{k-1}, [t_k; t_{k+1}, \ldots, t_{m+1}]]],\end{aligned}$$

полагая по индуктивному предположению $b_i = t_{i+1}$ и $j = k - 1$. Таким образом,

$$[t_0; t_1, \ldots, t_{m+1}] = [t_0; t_1, t_2, \ldots, t_{k-1}, [t_k; t_{k+1}, \ldots, t_{m+1}]]$$

согласно определению цепной дроби. Следовательно, по методу индукции теорема справедлива. ■

■ УПРАЖНЕНИЯ

1. Найдите два представления цепной дробью, $[t_0; t_1, \ldots, t_n]$, для каждого из приведенных ниже рациональных чисел:

а) 37/11; **б)** 48/1003; **в)** -257/2003;
г) 11/37; **д)** $-5/44$; **е)** 5.

2. Вычислите приведенные рациональные числа и выразите их в виде p/q, где p и q — целые числа:

а) $[3; 5, 2]$; **б)** $[0; 3, 5, 2]$; **в)** $[-10; 1, 4, 3]$;
г) $[6; 4, 7, 3, 5]$; **д)** $[2; 5, 3]$; **е)** $[5; 3, 7, 4, 6]$.

3. Найдите представление числа x конечной цепной дробью в виде $x = [t_0; t_1, t_2, t_3, t_4, t_5, x_6]$ для

а) $x = \sqrt{5}$;
б) $x = \sqrt{2}$;
в) $x = \pi$;
г) $x = (1 + \sqrt{5})/2$.

4. Пусть $[t_0; t_1, \ldots, t_n]$ — конечная цепная дробь, где $t_i > 0$ для $1 \leq i \leq n$. Докажите, что если $b > 0$, то

а) $[t_0; t_1, t_2, \ldots, t_n] > [t_0; t_1, t_2, \ldots, t_n + b]$, если n нечетно;
б) $[t_0; t_1, t_2, \ldots, t_n] < [t_0; t_1, t_2, \ldots, t_n + b]$, если n четно.

7.5. ПОДХОДЯЩИЕ ДРОБИ

Вполне очевидно, что $[t_0; t_1, t_2, \ldots, t_k]$ для каждого k представляет собой действительное число. Запишем первые четыре подходящие дроби и упростим их, выразив каждую как отношение двух действительных чисел, где как числитель, так и знаменатель выражены через t_i. Особый интерес будут представлять числитель и знаменатель каждой из подходящих дробей.

$$[t_0;] = t_0 = \frac{t_0}{1} = \frac{p_0}{q_0},$$

где мы положили $p_0 = t_0$ и $q_0 = 1$.

$$[t_0; t_1] = t_0 + \frac{1}{t_1} = \frac{t_0 t_1 + 1}{t_1} = \frac{p_1}{q_1},$$

где $p_1 = t_0 t_1 + 1$ и $q_1 = t_1$.

$$[t_0; t_1, t_2] = [t_0; [t_1; t_2]] = t_0 + \frac{1}{[t_1; t_2]} = t_0 + \frac{1}{t_1 + \dfrac{1}{t_2}} = t_0 + \frac{t_2}{t_1 t_2 + 1} =$$

$$= \frac{t_0 t_1 t_2 + t_0 + t_2}{t_1 t_2 + 1} = \frac{(t_0 t_1 + 1) t_2 + t_0}{t_1 t_2 + 1} = \frac{p_1 t_2 + p_0}{q_1 t_2 + q_0} = \frac{p_2}{q_2},$$

где $p_2 = p_1t_2 + p_0$ и $q_2 = q_1t_2 + q_0$. Подобным образом вычисляем

$$[t_0; t_1, t_2, t_3] = \frac{t_0t_1t_2t_3 + t_0t_1 + t_0t_3 + t_2t_3 + 1}{t_1t_2t_3 + t_1 + t_3} = \\ = \frac{(t_0t_1t_2 + t_0 + t_2)t_3 + (t_0t_1 + 1)}{(t_1t_2 + 1)t_3 + t_1} = \frac{p_2t_3 + p_1}{q_2t_3 + q_1} = \frac{p_3}{q_3},$$

где $p_3 = p_2t_3 + p_1$ и $q_3 = q_2t_3 + q_1$. В каждом случае число $[t_0; t_1, t_2, \ldots, t_k]$ "переформировывается" из представления цепной дробью без какого-либо "сокращения" и образует отношение двух полиномов P и Q в переменных $t_0, t_1, \ldots, t_k$, т.е.

$$[t_0; t_1, t_2, \ldots, t_k] = \frac{P(t_0, t_1, \ldots, t_k)}{Q(t_0, t_1, \ldots, t_k)}.$$

Справедливым является и то, что q_0, q_1, q_2 и q_3 положительны, поскольку получены путем умножения и сложения положительных чисел. Рассмотренные примеры дают основание сформулировать следующую теорему.

ТЕОРЕМА 7.11. Пусть n — неотрицательное целое число и $[t_0; t_1, t_2, \ldots, t_n]$ — конечная цепная дробь, которая рекурсивно определяет конечные последовательности $p_0, p_1, \ldots, p_n$ и $q_0, q_1, \ldots, q_n$ следующим образом:

а) $p_0 = t_0$,
$q_0 = 1$.

б) $p_1 = t_0t_1 + 1$,
$q_1 = t_1$.

в) $p_k = p_{k-1}t_k + p_{k-2}$,
$q_k = q_{k-1}t_k + q_{k-2}$ при $2 \le k \le n$.

Тогда $q_k > 0$ и $[t_0; t_1, t_2, \ldots, t_k] = \dfrac{p_k}{q_k}$ при $0 \le k \le n$.

ДОКАЗАТЕЛЬСТВО. Не представляет труда методом математической индукции показать, что $q_k > 0$ при $0 \le k \le n$. Эту часть теоремы предоставляем доказать читателю. Выше было показано, что теорема справедлива при $k = 0, 1$ и 2, т.е. $[t_0;] = p_0/q_0$, $[t_0; t_1] = p_1/q_1$ и $[t_0; t_1, t_2] = p_2/q_2$. Предположим, что $2 \le k < n$, и для любой цепной дроби $[b_0; b_1, \ldots, b_k]$ и любого j $0 \le j \le k$ справедливо утверждение о том, что $[b_0; b_1, \ldots, b_j] = p'_j/q'_j$, где p'_j и q'_j определены способом, аналогичным (a)-(c), но для цепной дроби $[b_0; b_1, \ldots, b_k]$. Тогда

$$\begin{aligned}[t_0; t_1, \ldots, t_k, t_{k+1}] &= [t_0; t_1, \ldots, t_{k-1}, [t_k; t_{k+1}]] \text{ (по теореме 7.10)} = \\ &= [b_0; b_1, \ldots, b_{k-1}, b_k],\end{aligned}$$

где $b_i = t_i$ при $0 \le i \le k-1$ и $b_k = [t_k; t_{k+1}] = t_k + \dfrac{1}{t_{k+1}}$. Таким образом, в силу индуктивного предположения

$$[t_0; t_1, \ldots, t_{k+1}] = \frac{p'_{k-1}b_k + p'_{k-2}}{q'_{k-1}b_k + q'_{k-2}}.$$

Поскольку $b_i = t_i$ при $0 \le i \le k-1$, имеем, что для таких i $p_i = p'_i$ и $q_i = q'_i$. Подставляя соответствующие выражения для b_k, p'_i и q'_i, получаем

$$[t_0; t_1, \ldots, t_{k+1}] = \frac{p_{k-1}\left(t_k + \dfrac{1}{t_{k+1}}\right) + p_{k-2}}{q_{k-1}\left(t_k + \dfrac{1}{t_{k+1}}\right) + q_{k-2}} =$$

$$= \frac{(p_{k-1}t_k + p_{k-2}) + p_{k-1}/t_{k+1}}{(q_{k-1}t_k + q_{k-2}) + q_{k-1}/t_{k+1}} =$$

$$= \frac{p_k t_{k+1} + p_{k-1}}{q_k t_{k+1} + q_{k-1}} = \frac{p_{k+1}}{q_{k+1}}.$$

Следовательно, по индукции, $[t_0; t_1, \ldots, t_k] = p_k/q_k$ при $0 \le k \le n$. ■

Числа p_k и q_k (теорема 7.11) определены вне зависимости от того, что они могут быть использованы как числитель и знаменатель выражения, равного k-ой подходящей дроби. Рекурсивное отношение, заданное теоремой 7.11, обеспечивает быстрый способ вычисления подходящих дробей для заданной цепной дроби, поскольку числитель p_i и знаменатель q_i можно вычислить одновременно и достаточно просто. Например, для $x = [-4; 2, 5, 2, 1] = [t_0; t_1, t_2, t_3, t_4]$ подходящие дроби можно вычислить согласно приведенной ниже таблице:

k	t_k	p_k	q_k	p_k/q_k
0	-4	-4	1	$-4/1$
1	2	$(-4)(2)+1=-7$	2	$-7/2$
2	5	$(-7)(5)+(-4)=-39$	$(2)(5)+1=11$	$-39/11$
3	2	$-39)(2)+(-7)=-85$	$(11)(2)+2=24$	$-85/24$
4	1	$(-85)(1)+(-39)=-124$	$(24)(1)+11=35$	$-124/35$

Так что $x = -124/35$, т.е. числу, породившему цепную дробь $[-4; 2, 5, 2, 1]$ (см. раздел 5.1).

Благодаря рекурсивному способу задания цепных дробей, имеется значительное количество соотношений, связывающих подходящие дроби через числители и знаменатели, p_k и q_k, введенные теоремой 7.11. Чтобы запись этих соотношений сделать справедливой при $k = 0$, зачастую удобно определять p_k и q_k при $k = -1$, положив

$$p_{-1} = 1;$$
$$q_{-1} = 0.$$

Таким образом, $p_1 = p_0 t_1 + p_{-1} = t_0 t_1 + 1$ и $q_1 = q_0 t_1 + q_{-1} = 1 \cdot t_1 + 0 = t_1$, как в теореме 7.11.

ТЕОРЕМА 7.12. Если $[t_0; t_1, \ldots, t_n]$ — конечная цепная дробь, где t_i — действительные числа, p_k и q_k — заданы теоремой 7.11, тогда

а) $p_k q_{k-1} - p_{k-1} q_k = (-1)^{k-1}$ при $0 \le k \le n$;

б) $\dfrac{p_k}{q_k} - \dfrac{p_{k-1}}{q_{k-1}} = \dfrac{(-1)^{k-1}}{q_k q_{k-1}}$ при $1 \le k \le n$;

в) $\dfrac{p_k}{q_k} - \dfrac{p_{k-2}}{q_{k-2}} = \dfrac{(-1)^k t_k}{q_k q_{k-2}}$ при $2 \le k \le n$.

г) Подходящие дроби четного порядка образуют возрастающую последовательность, т.е.

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots,$$

а подходящие дроби нечетного порядка образуют убывающую последовательность, т.е.

$$\frac{p_1}{q_1} > \frac{p_3}{q_3} > \frac{p_5}{q_5} > \cdots.$$

Кроме того,

$$\frac{p_{2j}}{q_{2j}} \le [t_0; t_1, t_2, \ldots, t_n] \le \frac{p_{2i+1}}{q_{2i+1}}$$

при $0 \le j \le \lfloor n/2 \rfloor$ и $0 \le i \le \lfloor (n-1)/2 \rfloor$, где слева равенство имеет место, когда n четное, а справа равенство имеет место, когда n нечетное.

д) Если дробь $[t_0; t_1, \ldots, t_n]$ — простая, то $q_k \ge k$ при $0 \le k \le n$.

е) Если дробь $[t_0; t_1, \ldots, t_n]$ — простая, то $q_k < q_{k+1}$ при $1 \le k \le n-1$ и $q_0 \le q_1$.

ДОКАЗАТЕЛЬСТВО. а) При $k = 1$ $p_1 q_0 - p_0 q_1 = (t_0 t_1 + 1)(1) - t_0 t_1 = 1 = (-1)^{1-1}$ по теореме 7.11. Пусть $2 \le m < n$ и допустим, что формула в части (а) справедлива при $k = m$. Показано, что формула имеет место при $k = m+1$. Опять, используя теорему 7.11, имеем

$$\begin{aligned} p_{m+1}q_m - p_m q_{m+1} &= (p_m t_{m+1} + p_{m-1})q_m - p_m(q_m t_{m+1} + q_{m-1}) = \\ &= p_{m-1}q_m - p_m q_{m-1} = \\ &= (-1)(p_m q_{m-1} - p_{m-1}q_m) = \\ &= (-1)(-1)^{m-1} = (-1)^{(m+1)-1}. \end{aligned}$$

Таким образом, формула в части (а) имеет место при $1 \le k \le n$. Также $p_0 q_{-1} - q_0 p_{-1} = t_0 \cdot 0 - 1 \cdot 1 = -1$ и формула справедлива при $k = 0$.

б) Утверждение теоремы следует из части (а), если в соответствующем равенстве при $k \ge 1$ произвести деление на $q_k q_{k-1}$.

в) Если $k \ge 2$, то по теореме 7.11

$$\begin{aligned} p_k &= p_{k-1}t_k + p_{k-2}; \\ q_k &= q_{k-1}t_k + q_{k-2}. \end{aligned}$$

Умножение первого из равенств на q_{k-2}, а второго — на p_{k-2} дает

$$\begin{aligned} p_k q_{k-2} &= p_{k-1}q_{k-2}t_k + p_{k-2}q_{k-2}; \\ q_k p_{k-2} &= q_{k-1}p_{k-2}t_k + q_{k-2}p_{k-2}. \end{aligned}$$

Вычитая второе равенство из первого, получаем

$$p_k q_{k-2} - q_k p_{k-2} = (p_{k-1}q_{k-2} - p_{k-2}q_{k-1})t_k.$$

Теперь, учитывая утверждение в части (а), получаем

$$p_k q_{k-2} - q_k p_{k-2} = (-1)^{(k-1)-1} t_k = (-1)^k (-1)^{-2} t_k = (-1)^k t_k.$$

Разделив на $q_k q_{k-2} > 0$, получаем желаемый результат:

$$\frac{p_k}{q_k} - \frac{p_{k-2}}{q_{k-2}} = \frac{(-1)^k t_k}{q_k q_{k-2}}.$$

г) Если $2 \le k \le n$ и k — четное, то $k-2$ тоже четное и $k-2 \ge 0$. Поэтому $(-1)^k = 1$ и из утверждения в части (в) вытекает, что

$$\frac{p_k}{q_k} - \frac{p_{k-2}}{q_{k-2}} = \frac{t_k}{q_k q_{k-2}} > 0,$$

поскольку t_k и $q_k q_{k-2}$ положительные при $k \ge 1$. Если $3 \le k \le n$ и k — нечетное, то $k-2$ также нечетное и $k-2 \ge 1$. В этом случае $(-1)^k = -1$, так что из утверждения в части (в) вытекает, что

$$\frac{p_k}{q_k} - \frac{p_{k-2}}{q_{k-2}} = \frac{(-1)t_k}{q_k q_{k-2}} < 0,$$

так как t_k и $q_k q_{k-2}$ положительны при $k \ge 3$.

Если n нечетное, то согласно (б) $\dfrac{p_n}{q_n} > \dfrac{p_{n-1}}{q_{n-1}}$, так что $[t_0; t_1, t_2, \ldots, t_n] = p_n/q_n$ больше, чем наибольшая подходящая дробь четного порядка, p_{n-1}/q_{n-1}. Если n четное, то $\dfrac{p_n}{q_n} < \dfrac{p_{n-1}}{q_{n-1}}$, так что $[t_0; t_1, t_2, \ldots, t_n] = p_n/q_n$ меньше, чем наименьшая подходящая дробь нечетного порядка, p_{n-1}/q_{n-1}.

Доказательство пунктов (д) и (е) оставляем за читателем. ■

Подходящие дроби для $x = [-4; 2, 5, 2, 1]$ вычислены ранее в данном разделе и приведены в таблице. Подводя итог этим результатам в контексте части (г), находим, что

$$\begin{array}{ccccccccccc} p_0/q_0 & < & p_2/q_2 & < & p_4/q_4 & = & x & < & p_3/q_3 & < & p_1/q_1; \\ -4/1 & < & -39/11 & < & -124/35 & = & x & < & -85/24 & < & -7/2. \end{array}$$

Непосредственным вычислением определяем также, что

$$p_3 q_2 - p_2 q_3 = (-85)(11) - (-39)(24) = (-935) - (-936) = 1.$$

Значение теоремы 7.12 состоит в том, что она точно устанавливает, что последовательные и взаимно исключающие подходящие дроби могут отличаться не более, чем на величину, обратно пропорциональную "знаменателям" подходящих дробей, а также, что подходящие дроби альтернативно больше, чем $[t_0; t_1, t_2, \ldots, t_n]$, и меньше, чем $[t_0; t_1, t_2, \ldots, t_n]$.

Практический интерес представляют некоторые отношения p_k/q_k числителей и знаменателей подходящих дробей. Эти отношения приведены в следующей теореме, доказательство которой предоставляется читателю.

ТЕОРЕМА 7.13. Для конечной цепной дроби $[t_0; t_1, t_2, \ldots, t_n]$

а) $q_k/q_{k-1} = [t_k; t_{k-1}, \ldots, t_2, t_1]$ при $1 \le k \le n$.
б) Если $t_0 \neq 0$, то $p_k/p_{k-1} = [t_k; t_{k-1}, \ldots, t_1, t_0]$ при $1 \le k \le n$.
в) Если $t_0 = 0$, то $p_k/p_{k-1} = [t_k; t_{k-1}, \ldots, t_2, t_1]$ при $2 \le k \le n$.

■ УПРАЖНЕНИЯ

1. Докажите теорему 7.11 при условии, что $q_k > 0$ при $0 \le k \le n$.

2. Вычислите подходящие дроби для

а) $[2; 5, 1, 2, 4]$; **б)** $[2; 5, 1, 2, 3, 1]$; **в)** $[5; 1, 2, 3]$;
г) $[0; 5, 1, 2, 3]$; **д)** $[-5; 3, 8, 2]$.

3. Вычислите подходящие дроби p_k/q_k, определяя p_k и q_k при $0 \le k \le 5$ для $x = [t_0; t_1, t_2, t_3, t_4, t_5, x_6]$ и для следующих значений x (см. упражнения раздела 7.1):

а) $x = \sqrt{5}$;
б) $x = \sqrt{2}$;
в) $x = \pi$;
г) $x = (1 + \sqrt{5})/2$.

4. Докажите утверждения (д) и (е) теоремы 7.12.

5. Докажите теорему 7.13. *Указание*: используйте теорему 7.11.

6. Пусть для цепной дроби $[t_0; t_1, t_2, \ldots, t_n]$ величины p_k и q_k определены теоремой 7.11. Докажите, что

$$\begin{vmatrix} p_k & p_{k-1} \\ q_k & q_{k-1} \end{vmatrix} = (-1)^{k+1} \text{ при } 0 \le k \le n,$$

где левая часть равенства записана с помощью определителя.

7. В теореме 7.11 показано, что если $p_{-1} = 1$ и $q_{-1} = 0$, утверждение (в) справедливо при $k = 1$. Какие значения должны иметь p_{-2} и q_{-2}, чтобы утверждение (в) было справедливо для $k = 0$?

ГЛАВА

8

КОМБИНАТОРИКА И ВЕРОЯТНОСТЬ

8.1. ОСНОВНЫЕ КОМБИНАТОРНЫЕ ПРИНЦИПЫ

Как говорил Альберт Эйнштейн: "Не все, что можно сосчитать, сосчитано, и не все, что сосчитано, можно сосчитать". Хотя это высказывание не очень воодушевляет, все же попытаемся заняться подсчетами. Фактически, очень много проблем в данной книге начинается со слова "сколько". Сколько в этой книге проблем начинается со слова "сколько"? Мы не задаем этот вопрос по нескольким причинам, поскольку не знаем, будет ли ответ содержать сам вопрос? Возможно, главная причина, что вопрос остается без ответа, это отсутствие схемы, по которой на него можно было бы ответить. В действительности, важнее нахождение и использование общих схем для подсчетов, чем решение отдельных задач. Наша цель — избегать простого перечисления элементов с последующим подсчетом, в особенности, если таких элементов очень много. Однако, данный метод можно использовать для небольшого количества объектов, если это поможет найти общую схему.

Представьте себе, что человек, которому нечем больше заняться, подбрасывает монету и кость (в виде игрального кубика). На монете может выпасть "орел" (H) и "решка" (T). При падении кости возможные исходы представляют собой 1, 2, 3, 4, 5 и 6. Возможными исходами обоих событий являются $(H,1)$, $(H,2)$, $(H,3)$, $(H,4)$, $(H,5)$, $(H,6)$, $(H,1)$, $(H,2)$, $(H,3)$, $(T,4)$, $(T,5)$ и $(T,6)$. Один из возможных способов представить эти события — построить ***дерево подсчета***, изображенное на рис. 8.1.

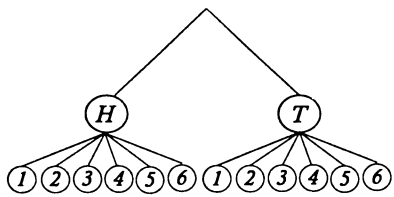


Рис. 8.1

Для нахождения исхода $(H,2)$ идем по левой ветке к H, а затем по ветке к 2. Очевидно, что таким образом представлены все исходы событий, и каждое

из обведенных чисел изображает единственный исход. Если на монете выпал H, то существуют шесть возможных исходов для кубика; и те же шесть возможных исходов, если выпала "решка" T. Таким образом, можно найти общее количество исходов, умножая 2×6, т.е. умножая число возможных исходов для подбрасывания монеты на число возможных исходов для подбрасывания игральной кости. Следует обратить внимание, что исход подбрасывания монеты не влияет на возможный исход подбрасывания кости.

Можно представить, что игрок подбрасывает две кости. Существуют шесть исходов выпадений первого кубика и шесть — второго. Поэтому существует $6 \times 6 = 36$ возможных исходов выпадения при подбрасывании двух кубиков. По желанию можно снова нарисовать дерево подсчета. Аналогичным образом предположить, что выбирается сначала одна буква из 26-символьного алфавита, а затем — вторая, причем вторая буква может совпадать с первой. Для нахождения общего числа буквенных пар важно заметить, что имеются 26 вариантов выбора первой буквы, и для каждой выбранной первой буквы существуют 26 вариантов выбора второй. Следовательно, существуют $26 \times 26 = 676$ возможных способов выбора из алфавита буквенных пар. Если вторая буква не должна совпадать с первой, то после выбора первой буквы имеется только 25 вариантов выбора второй. Поэтому всего имеются $26 \times 25 = 650$ возможных вариантов выбора. Другой способ решения этой же задачи: вычесть из 676, общего числа возможных вариантов выбора буквенных пар из алфавита, число всех вариантов, в которых обе буквы совпадают. Их, понятно, 26, поэтому, вычитая 26 из 676, снова получаем 650.

Предположим, что номерной знак автомобиля содержит три буквы алфавита, за которыми следуют три десятичные цифры. Если повторения разрешены, то существуют 26 вариантов выбора каждой буквы и 10 вариантов выбора каждой цифры. Таким образом, общее число номерных знаков составит

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 = 17{,}576{,}000.$$

Допустим, что по какой-то причине буквы и цифры не могут повторяться. Тогда, после выбора первой буквы имеются 25 вариантов выбора второй и 24 варианта выбора третьей буквы. Аналогично имеем 10 вариантов выбора первой цифры, 9 вариантов выбора второй и 8 вариантов выбора третьей цифры номера. Таким образом, общее число номерных знаков составит

$$26 \times 25 \times 24 \times 10 \times 9 \times 8 = 11{,}232{,}000.$$

Выразим данный метод в приведенной ниже теореме, которую примем в качестве аксиомы.

ТЕОРЕМА 8.1. (Комбинаторный принцип умножения) Пусть задана последовательность событий

$$E_1, E_2, E_3, \ldots, E_m$$

таких, что событие E_1 осуществляется n_1 способами, и если события $E_1, E_2, E_3, \ldots, E_{k-1}$ осуществились, то событие E_k может осуществиться n_k способами. Тогда существует $n_1 \times n_2 \times n_3 \times \cdots \times n_m$ способов осуществления всей последовательности событий.

ПРИМЕР 8.2. Сколько существует функций, задающих взаимно однозначное соответствие между множеством S, содержащим n элементов, и множеством T, содержащим m элементов? Если $n > m$, то такие функции не существуют, поэтому предположим, что $n \leq m$. Не уменьшая общности, допустим, что элементы множества S обозначены $a_1, a_2, a_3, \ldots, a_n$. Существует m вариантов отображения элемента a_1. Если образ a_1 определен, то имеется $m-1$ вариантов отображения элемента a_2, поскольку он не может быть отображен в образ элемента a_1. Аналогично, существует $m-2$ вариантов отображения a_3 и $m-3$ вариантов отображения a_4, и т.д. Руководствуясь этой схемой, получаем, что существует $m-i+1$ вариантов отображения a_i. Поэтому, в соответствии с комбинаторным принципом умножения, имеется

$$(m)(m-1)(m-2)(m-3)\cdots(m-n+1)$$

способов отображения множества S на множество T, при котором никакие два элемента множества S не будут отображены в один и тот же элемент множества T. Таким образом, существует

$$(m)(m-1)(m-2)(m-3)\cdots(m-n+1)$$

взаимно однозначных функций из S в T. □

ПРИМЕР 8.3. Сколько существует функций из множества S, содержащего n элементов, в множество T, содержащее m элементов? На этот раз любой из n элементов множества S может быть отображен в любой из m элементов множества T. Следовательно, существует $(m)(m)(m)(m)\cdots(m) = m^n$ функций из S в T. □

ПРИМЕР 8.4. ***Битовая строка*** — это строка, состоящая из элементов, каждый из которых имеет значение 1 или 0. Сколько существует битовых строк, имеющих длину 5? Сколько существует битовых строк длины k? Поскольку каждый символ строки может иметь значение 1 или 0, то существует два варианта выбора для каждой позиции. Следовательно, существует $2 \times 2 \times 2 \times 2 \times 2 = 2^5$ битовых строк длины 5. По аналогичным соображениям, имеется 2^k битовых строк длины k. □

ПРИМЕР 8.5. Сколько существует подмножеств множества S, содержащего пять элементов? Сколько подмножеств имеет множество, содержащее k элементов? Имеется несколько способов подсчета подмножеств. Укажем два из них. Рассмотрим первый способ. Пусть a_1, a_2, a_3, a_4, a_5 — элементы множества S. Определим взаимно однозначное соответствие между множеством S и множеством битовых строк длины 5, при котором каждому подмножеству s множества S соответствует строка, в которой i-ый бит равен 1, если $a_i \in S$, и равен 0, если $a_i \notin S$. Таким образом, подмножество $\{a_1, a_2, a_4\}$ соответствует строке 11010, подмножество $\{a_2, a_5\}$ соответствует строке 01001, $\varnothing$ соответствует строке 00000, а множество S соответствует строке 11111. Итак, установлено взаимно однозначное соответствие между подмножествами множества S и битовыми строками длины 5. Из предыдущего примера известно, что существует 2^5 битовых строк длины 5. Поэтому существует 2^5 подмножеств множества из пяти элементов. Используя аналогичные рассуждения, можно показать, что существует 2^k подмножеств k-элементного множества.

Второй метод подсчета состоит в том, чтобы для каждого подмножества $s \in S$ определить функцию f_s из S в множество $\{0,1\}$. Функция f_s определена соотношениями $f_s(a_i) = 1$, если $a_i \in s$, и $f_s(a_i) = 0$, если $a_i \notin s$. Обратно, пусть задана функция f из множества S в множество $\{0,1\}$. Если положить $s = \{a_i : f(a_i) = 1\}$, то функция f является функцией f_s. Поэтому имеется взаимно однозначное соответствие между подмножествами множества S и функциями из S в $\{0,1\}$. Но, согласно примеру 8.3, существует 2^5 функций из S в $\{0,1\}$. Подобным образом, если множество S содержит k элементов, то имеется 2^k функций из S в $\{0,1\}$, и поэтому существует 2^k подмножеств множества S. □

Предположим, что на собрании присутствуют 10 мужчин и 15 женщин, и кого-то одного нужно выбрать председателем. Существует $10 + 15 = 25$ способов выбора председателя. Предположим, что человек направляется в ресторан, в котором предлагаются 15 блюд из говядины, 8 блюд из свинины и 12 блюд из морских продуктов. Посетителю ресторана предложен выбор из $15 + 8 + 12 = 35$ различных блюд. Предположим теперь, что студент выбирает книгу на полке, где находятся 25 различных учебников по математике, 30 учебников по информатике и 15 — по химии. Существует $25 + 30 + 15 = 70$ различных вариантов выбора книги студентом. Заметим, что в каждом случае множества, из элементов которых делается выбор, не пересекаются. Эти примеры приводят к следующей теореме, которая формулируется без доказательства.

ТЕОРЕМА 8.6. (Комбинаторный принцип сложения) Пусть $S_1, S_2, S_3, \ldots, S_m$ — попарно непересекающиеся множества (т.е. $S_i \cap S_j = \varnothing$ для всех $i \neq j$), и пусть для каждого i, множество S_i содержит n_i элементов. Количество вариантов выбора из S_1 или S_2 или S_3 или ... или S_m равно $n_1 + n_2 + n_3 + \cdots + n_m$. На языке теории множеств утверждение теоремы имеет вид

$$|S_1 \cup S_2 \cup S_3 \cup \ldots \cup S_m| = |S_1| + |S_2| + |S_3| + \ldots + |S_m|,$$

где $|S|$ обозначает количество элементов множества S.

ПРИМЕР 8.7. Сколько существует целых чисел между 0 и 1000, содержащих ровно одну цифру 6? Пусть S будет множеством целых чисел между 0 и 1000, содержащих ровно одну цифру 6. Пусть S_1 — подмножество множества S, которое содержит число, состоящее из одной цифры, и эта цифра равна 6. Пусть S_2 — подмножество множества S, содержащее двузначные числа ровно с одной цифрой, равной 6. Пусть S_3 — подмножество множества S, содержащее трехзначные числа ровно с одной цифрой, равной 6. Множество S_1 содержит только один элемент — число 6. В S_2 каждый элемент, содержащий 6, имеет ее либо первой, либо второй цифрой. Если 6 — вторая цифра, то существуют 8 различных чисел, поскольку первое число не может быть 0 или 6. Если 6 — первая цифра, то таких чисел 9, поскольку вторая цифра не может быть 6. Таким образом, S_2 содержит $8 + 9 = 17$ элементов. Элемент из S_3 содержит 6 как первую, вторую или третью цифру. Если 6 — первая цифра, то существуют 9 вариантов выбора второй цифры и 9 вариантов выбора третьей цифры. Согласно комбинаторному принципу умножения, S_3 содержит $9 \times 9 = 81$ чисел с 6 как первой цифрой. Если 6 — вторая цифра, то имеются 9 вариантов выбора третьей цифры и 8 вариантов

выбора первой цифры, поскольку первая цифра не может быть нулем. Следовательно, S_3 содержит $9 \times 8 = 72$ чисел, у которых 6 — вторая цифра. Аналогично, S_3 содержит 72 числа, у которых 6 — третья цифра. Следовательно, всего S_3 содержит $81 + 72 + 72 = 225$ элементов. Поскольку $S = S_1 \cup S_2 \cup S_3$, то S содержит $1 + 17 + 225 = 243$ элементов. □

ПРИМЕР 8.8. Сколько существует целых чисел между 0 и 1000, содержащих хотя бы одну цифру 6? Для решения этой задачи положим S множеством целых чисел от 0 до 1000, **не** содержащих цифру 6 ни в одном из разрядов. Пусть S_1 — подмножество множества S, содержащее однозначные числа; S_2 — подмножество множества S, содержащее двузначные числа; S_3 — подмножество множества S, содержащее трехзначные числа; и S_4 — подмножество множества S, содержащее четырехзначные числа. Множество S_1 содержит девять чисел без цифры 6, поскольку только число 6 исключено. Множество S_2 содержит числа, имеющие 8 вариантов выбора первой цифры и 9 вариантов выбора второй цифры, поскольку первая цифра не может быть ни 6, ни 0, а вторая цифра не может быть 6. Поэтому S_2 содержит $8 \times 9 = 72$ элементов без цифры 6. Множество S_3 содержит числа, имеющие 8 вариантов выбора первой цифры, 9 вариантов выбора второй цифры и 9 вариантов выбора третьей цифры. Поэтому множество S_3 содержит $8 \times 9 \times 9 = 648$ элементов без цифры 6. Множество S_4 содержит только число 1000. Следовательно, S содержит $9+72+648+1 = 730$ элементов. Пусть T — множество всех чисел между 0 и 1000. Тогда $T - S$ — множество всех целых чисел между 0 и 1000, содержащих хотя бы одну цифру 6. Более того, S и $T - S$ — непересекающиеся множества. Следовательно, $|S|+|T-S| = |T|$, так что $730+|T-S| = 1001$. Значит, $|T - S| = 271$, поэтому между 1 и 1000 существует 271 целое число, содержащее хотя бы одну цифру 6. □

ПРИМЕР 8.9. Сколькими способами можно выбрать две книги по разным темам, когда на полке находятся 15 книг по информатике, 12 книг по математике и 10 книг по химии? Если выбирать книгу по информатике и книгу по математике, то существуют 15 вариантов выбора книги по информатике и 12 вариантов выбора книги по математике, поэтому существует $12 \times 15 = 180$ возможностей. Если выбирать книги по информатике и по химии, то имеются 15 вариантов выбора книги по информатике и 10 вариантов выбора книги по химии, поэтому существует $15 \times 10 = 150$ возможностей. Если выбирается книга по математике и книга по химии, то имеются 12 способов выбора математической книги и 10 — книги по химии, поэтому имеется $12 \times 10 = 120$ возможностей. Следовательно, существуют $180 + 150 + 120 = 450$ возможных способов выбора двух книг. □

ПРИМЕР 8.10. Сколько нечетных целых чисел находятся между числами 100 и 1000? Один из способов решения этой задачи состоит в следующем. Пусть S — множество нечетных целых чисел между 100 и 1000. Для $i \in \{1, 3, 5, 7, 9\}$ пусть S_i — подмножество множества S такое, что i является последней цифрой его элементов. Для каждого i существуют 9 вариантов выбора первой цифры и 10 вариантов выбора второй цифры, так что каждое множество S_i содержит 90 элементов. Поскольку $S = S_1 \cup S_3 \cup S_5 \cup S_7 \cup S_9$, имеем

$$|S| = |S_1| + |S_3| + |S_5| + |S_7| + |S_9| =$$

$$= 90 + 90 + 90 + 90 + 90 =$$
$$= 450,$$

поэтому между числами 100 и 1000 находятся 450 нечетных целых чисел.

Данную задачу можно решить, используя также комбинаторный принцип умножения. Существуют 5 вариантов выбора последней цифры, поскольку она должна быть нечетной. Существуют 10 вариантов выбора средней цифры и 9 вариантов выбора первой цифры. Таким образом, имеем $9 \times 10 \times 5 = 450$ нечетных чисел между 100 и 1000. Этот пример показывает, каким образом принцип комбинаторного умножения следует из комбинаторного принципа сложения. □

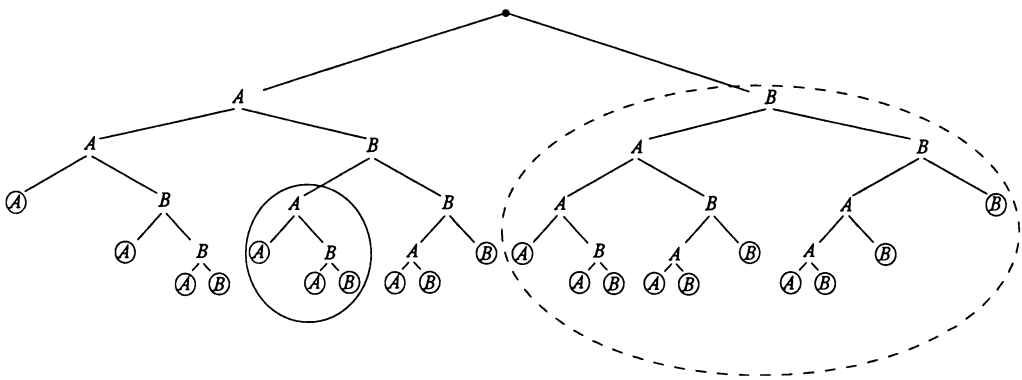


Рис. 8.2

Наконец, вернемся к тому, с чего начали, — к использованию деревьев. Предположим, что команды A и B играют между собой в турнире по бейсболу. Для победы в турнире команде необходимо выиграть три игры из пяти. Все возможные результаты игр представлены деревом на рис. 8.2, где буква в каждой вершине обозначает команду, победившую в игре. Обведенная кружком буква обозначает команду, победившую в турнире. Обратите внимание, что всего имеется 20 возможных исходов: 10 возможных побед у команды A и 10 — у команды B. Предположим, известно, что команда B выиграла первую игру, что оставляет для рассмотрения только правую сторону дерева, начиная с вершины B. Сторона обведена пунктирной линией. Обратите внимание, что имеются четыре исхода, в которых побеждает команда A, и шесть исходов с победой команды B. Предположим, что первую игру выигрывает команда A, вторую — команда B и команда A выигрывает третью игру. Результат игры отображен на дереве (часть дерева, обведенная сплошной линией). Обратите внимание, что имеются два исхода, когда побеждает команда A, и только один исход с победой команды B.

ПРИМЕР 8.11. Предположим, что команды A и B участвуют в ежегодном чемпионате США по бейсболу, для победы в котором командам необходимо выиграть четыре игры из семи возможных. Если команда A выигрывает первые две игры,

то сколько вариантов победы есть у команды A и сколько у команды B? Дерево на рис. 8.3 показывает возможные исходы. Имеются десять исходов с победой команды A и пять возможных исходов с победой команды B. □

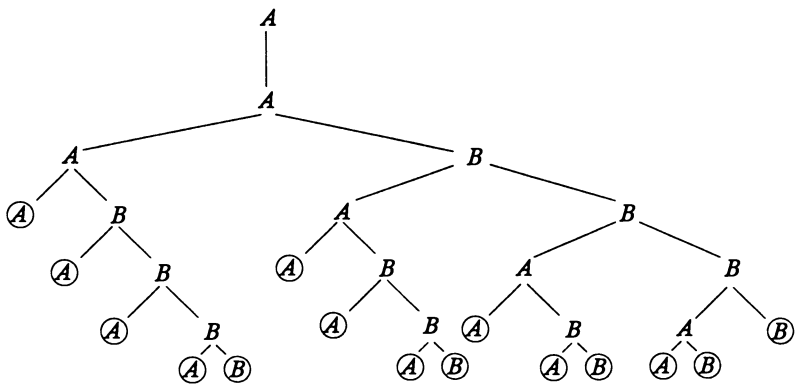


Рис. 8.3

■ УПРАЖНЕНИЯ

1. Сколько положительных целых чисел, меньших 700, делятся на 5? Сколько таких чисел делятся на 3? А сколько таких чисел делятся и на 5, и на 3?
2. Если авиакомпания осуществляет 15 рейсов из Сан-Франциско в Чикаго и 20 рейсов из Чикаго в Нью-Йорк, то сколько всего рейсов из Сан-Франциско в Нью-Йорк проходит транзитом через Чикаго?
3. Сколько существует способов избрания президента, вице-президента, секретаря и казначея среди членов клуба, включающего 8 студентов последнего курса, 10 студентов предпоследнего курса, 15 второкурсников и 20 первокурсников, если
   а) отсутствуют какие-либо ограничения?
   б) президентом должен быть студент последнего курса?
   в) студент последнего курса не может быть вице-президентом?
   г) первокурсники могут быть избраны только на должность секретаря?
4. В ресторане из напитков предлагают кофе, чай, молоко или колу. Предлагают на выбор суп или салат. Имеются 10 различных бифштексов и 5 разнообразных куриных блюд. На гарнир можно выбрать картофель фри, печеный картофель, макароны с сыром или рис. На десерт подают сладкий пирог, мороженое или то и другое вместе.
   а) Сколько можно составить различных меню?
   б) Сколько можно составить различных меню, включающих бифштекс?
   в) Сколько можно составить различных меню, если клиент выбирает бифштекс и картофель или не выбирает ни то, ни другое?

5. Сколько существует различных четырехзначных положительных чисел, если, по крайней мере, две цифры в числе совпадают?
6. Если каждый элемент матрицы размерности 4×5 — неотрицательное однозначное число, то сколько существует таких матриц?
7. Сколько существует вариантов ответа на тест из 30 вопросов, если на каждый вопрос требуется ответ да или нет?
8. Если часы показывают часы, минуты, секунды и AM-PM, то сколько различных моментов времени они могут показать?
9. Позывные американских радиостанций состоят из трех или четырех букв и начинаются с k или w. Сколько существует различных позывных?
10. Сколько существует номерных знаков для автомобилей, состоящих из двух букв и последующих четырех цифр?
11. Если имеются высказывания p, q, r и s, то

а) сколько строк содержит соответствующая таблица истинности?

б) в скольких строках все p, q, r и s имеют истинное значение?

в) в скольких строках хотя бы одно из четырех высказываний ложно?

12. Сколько существует номерных знаков для автомобилей, состоящих из двух букв с последующими четырьмя цифрами, если буквы могут повторяться, а цифры — нет?
13. Сколько существует различных номеров карточек социального страхования?
14. Сколько существует номерных знаков для автомобилей, состоящих из двух букв с последующими четырьмя цифрами, если ни одна из букв не может быть гласной? Если хотя бы одна из букв должна быть гласной?
15. Сколько существует функций из 6-элементного множества в 3-элементное?
16. Сколько существует инъективных функций из 3-элементного множества в 6-элементное?
17. Сколько положительных целых чисел, меньших 700, не делятся на 8?
18. Сколько положительных целых чисел, меньших 1000, не делятся на 5? Которые не делятся на 5?
19. Сколько существует двубуквенных или трехбуквенных инициалов для людей?
20. Сколько существует различных двубуквенных или трехбуквенных инициалов для людей, если никакие буквы не могут повторяться?
21. Сколько существует битовых строк длины 7?
22. Сколько существует битовых строк длины 7, если первый и последний биты совпадают?
23. Сколько существует битовых строк длины 7-элементных строк битов, содержащих две или более единиц?
24. Если телефонный номер не может начинаться с 0, 1 или 8, то сколько существует различных 7-значных телефонных номеров?
25. Сколько существует четырехзначных чисел, содержащих хотя бы одну цифру 7?
26. Сколько существует четырехзначных чисел, которые не содержат цифру 7?
27. Если компьютерный пароль содержит семь символов, которые могут быть цифрой или буквой, то сколько существует паролей, начинающихся с буквы?

8.2. КОМБИНАТОРНЫЙ ПРИНЦИП СЛОЖЕНИЯ

До настоящего момента непересекающиеся множества рассматривались только в связи с комбинаторным принципом сложения. Предположим, что множества S и T не являются непересекающимися и требуется найти $|S \cup T|$. Когда количество элементов множества S суммируется с количеством элементов множества T, элементы, принадлежащие $S \cap T$, учитываются дважды. Поэтому количество элементов, принадлежащих множеству $S \cap T$, нужно из суммы вычесть. Отсюда получаем следующую теорему.

ТЕОРЕМА 8.12. Пусть S и T — множества. Количество элементов, которое можно выбрать из S или T, равно $|S| + |T| - |S \cap T|$. Иными словами, $|S \cup T| = |S| + |T| - |S \cap T|$.

ДОКАЗАТЕЛЬСТВО. Множество $S \cup T = (S - T) \cup (T - S) \cup (S \cap T)$, где $S - T$, $T - S$ и $S \cap T$ — попарно непересекающиеся множества. Поэтому $|S \cup T| = |S - T| + |T - S| + |S \cap T|$. Имеем также, что $S = (S - T) \cup (S \cap T)$ и $T = (T - S) \cup (S \cap T)$, откуда $|S| = |S - T| + |S \cap T|$ и $|T| = |T - S| + |S \cap T|$. Следовательно,

$$\begin{aligned} |S| + |T| - |S \cap T| &= |S - T| + |T - S| + 2|S \cap T| - |S \cap T| = \\ &= |S - T| + |T - S| + |S \cap T| = \\ &= |S \cup T|. \end{aligned}$$

■

ПРИМЕР 8.13. Предположим, что в группе из 100 студентов 60 человек изучают математику, 75 — историю, а 45 человек — и то, и другое.

а) Сколько студентов изучают математику или историю?

б) Сколько студентов не изучают ни математику, ни историю?

а) Пусть универсум $\boldsymbol{U}$ — группа из 100 студентов, M — множество студентов, изучающих математику, H — множество студентов, изучающих историю. Тогда количество студентов, изучающих математику или историю, равно

$$\begin{aligned} |M \cup H| &= |M| + |H| - |M \cap H| = \\ &= 60 + 75 - 45 = \\ &= 90. \end{aligned}$$

б) Количество студентов, не изучающих ни математику, ни историю, равно $|M' \cap H'|$. Но $M' \cap H' = (M \cup H)'$, поэтому

$$\begin{aligned} |M' \cap H'| &= |(M \cup H)'| = \\ &= 100 - 90 = \\ &= 10. \end{aligned}$$

□

ПРИМЕР 8.14. Найдем количество положительных целых чисел, меньших 1001, которые делятся на 3 или на 5. Количество элементов множества S положительных целых чисел, меньших 1001, которые делятся на 3, равно $\lfloor \frac{1001}{3} \rfloor$ или 333. Количество элементов множества T положительных целых чисел, меньших 1001, которые делятся на 5, равно $\lfloor \frac{1001}{5} \rfloor$ или 200. Элементами множества $S \cap T$ являются целые числа, меньшие 1001, которые делятся на 5 и на 3, и поэтому делятся на 15. Следовательно, $|S \cap T| = \lfloor \frac{1001}{15} \rfloor$ или 66. Значит

$$|S \cup T| = 333 + 200 - 66 = 467.$$

□

ПРИМЕР 8.15. Сколько существует способов разместить цифры

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9$$

так, чтобы первая цифра была больше единицы, а последняя — меньше семи? Пусть $\boldsymbol{U}$ — множество всех возможных способов размещений цифр, а S — множество всех размещений цифр таких, что первая цифра больше единицы, а последняя — меньше 7. Для нахождения $|\boldsymbol{U}|$ заметим, что существуют десять способов выбора первой цифры, девять способов выбора второй, восемь способов выбора третьей цифры, ... , два способа выбора девятой и один способ выбора десятой цифры. Поэтому в соответствии с комбинаторным принципом умножения имеются $10! = 3\,628\,800$ возможных способов размещения цифр, значит $|\boldsymbol{U}| = 3\,628\,800$. Как и в первом примере, воспользуемся "остаточным" принципом и найдем количество размещений цифр, которые не входят в S, после чего вычтем это количество из $|\boldsymbol{U}|$. Пусть A — множество всех таких размещений цифр, где первая цифра меньше или равна 1. Подсчитывая количество элементов из A, обнаруживаем, что имеются два варианта выбора первой цифры, а именно, 0 и 1. Существуют девять вариантов выбора второй цифры, восемь вариантов выбора третьей цифры, ... , два варианта выбора девятой цифры и один вариант выбора десятой цифры. Таким образом, множество A содержит $2 \times 9! = 725\,760$ элементов. Пусть B — множество всех таких размещений цифр, что последняя цифра больше или равна 7. Подсчитывая количество элементов B, заметим, что имеются три варианта выбора последней цифры, и, следуя схеме, использованной для множества A, получаем, что существуют 9! вариантов выбора остальных цифр. Таким образом, $|B| = 3 \times 9! = 1\,088\,640$. Множество $A \cap B$ состоит из таких размещений цифр, в которых первая цифра больше или равна 1, а последняя цифра больше или равна 7. Подсчитывая количество элементов множества $A \cap B$, находим, что существуют два способа выбора первой цифры и три способа выбора последней цифры. Используя процедуру, аналогичную предыдущей, получаем 8! способов размещения оставшихся цифр. Таким образом,

$$|A \cap B| = 3 \times 2 \times 8! = 241920$$

и

$$|A \cup B| = 725760 + 1088640 - 241920 = 1572480.$$

Все размещения целых чисел такие, что первая цифра больше единицы, а последняя цифра меньше семи, описывают множество $A' \cap B' = (A \cup B)'$. Поэтому число таких размещений равно $|\boldsymbol{U}| - |(A \cup B)'| = 3628800 - 1572480 = 2056320$. □

Рассмотрев случай двух множеств, переходим к изучению случая трех множеств. Пусть заданы множества A, B и C, требуется определить $|A \cup B \cup C|$, количество элементов множества $A \cup B \cup C$, в случае, когда эти множества могут пересекаться. Если составить сумму количества элементов в каждом множестве A, B и C, то согласно диаграмме Венна, изображенной на рис. 8.4, некоторые подмножества при подсчете будут учтены дважды. Если теперь вычесть $|A \cap B|$, $|A \cap C|$ и $|B \cap C|$, т.е. количество элементов множества $A \cap B$, множества $A \cap C$ и множества $B \cap C$ соответственно, то, как следует из диаграммы Венна, изображенной на рис. 8.5, элементы множества $A \cap B \cap C$ совсем не будут учтены.

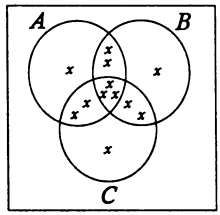


Рис. 8.4

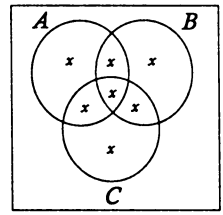


Рис. 8.5

Поэтому добавляем $|A \cap B \cap C|$, после чего каждый элемент множества $A \cup B \cup C$ учтен в сумме ровно один раз. Это дает следующую формулу:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

ПРИМЕР 8.16. Предположим, что, согласно исследованию, из 200 людей, смотрящих телевизор, 110 человек смотрят спортивную передачу, 120 — комедии, 85 предпочитают драмы, 50 смотрят драмы и спорт, 70 — комедии и спорт, 55 смотрят комедии и драмы и 30 человек смотрят все три вида передач. Сколько человек смотрят спорт, комедии или драмы? Сколько человек не смотрят ничего из вышеперечисленного? Пусть $\boldsymbol{U}$ — множество 200 людей, среди которых проводился опрос. Пусть S — множество людей, которые смотрят спорт, D — множество людей, которые смотрят драмы, и C — множество людей, предпочитающих комедии. Тогда первое задание — найти $|S \cup D \cup C|$. Это можно сделать, используя непосредственно формулу, приведенную выше:

$$|S \cup D \cup C| = |S| + |D| + |C| - |S \cap D| - |D \cap C| - |S \cap C| + |S \cap D \cap C| =$$
$$= 110 + 85 + 120 - 50 - 55 - 70 + 30 = 170,$$

так что 170 человек смотрят спорт, комедии или драмы. Второе задание требует найти $|(S \cup D \cup C)'| = |\boldsymbol{U}| - |(S \cup D \cup C)| = 200 - 170 = 30$. Итак, 30 человек не смотрят по телевидению ни спорт, ни драмы, ни комедии. □

Существует альтернативный метод решения приведенной выше задачи, который является более информативным. Известно, что 30 человек смотрят по телевидению спорт, комедии и драмы. Поскольку 50 человек смотрят спорт и драмы, то 20 человек должны смотреть спорт и драмы, но не смотреть комедии. Аналогично, 55 смотрят комедии и драмы, поэтому 25 человек должны смотреть комедии и драмы, но не смотреть спорт. К тому же 70 человек смотрят спорт и комедии, поэтому 40 из них смотрят спорт и комедии, но не смотрят драмы. Таким образом, получаем следующие диаграммы Венна, изображенные на рис. 8.6.

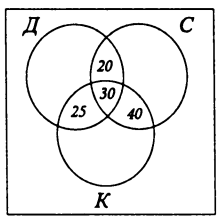


Рис. 8.6

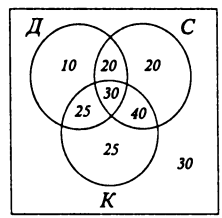


Рис. 8.7

Известно, что 85 человек предпочитают смотреть драмы. Поскольку 75 из них уже перечислены, остальные 10 должны смотреть только драмы. Рассуждая аналогичным образом, получаем, что из 110 тех, кто смотрит спорт, уже перечислены 90 человек, поэтому оставшиеся 20 должны смотреть только спорт. И наконец, замечаем, что из 120 тех, кто смотрит комедии, 95 уже учтены. Таким образом, 25 человек смотрит только комедии. Если подсчитать количество людей в непересекающихся подмножествах, то получим число 170. Поэтому 30 человек не смотрят ни спорт, ни комедии, ни драмы (диаграмма Венна, рис. 8.7).

Используя эту диаграмму, можно, например, установить следующие факты: 35 человек смотрят драмы, но не спорт; 50 человек смотрят спорт или драму, но не комедии; 55 человек предпочитают только одну из трех программ; 85 предпочитают две программы из трех; 85 человек смотрят комедии или спорт, но только не драмы.

ПРИМЕР 8.17. Предположим, что из 100 опрошенных студентов 50 изучают химию, 53 — математику, 42 — физику, 15 — химию и физику, 20 занимаются физикой и математикой, 25 — математикой и химией и 5 студентов изучают все три предмета.

а) Сколько студентов изучают хотя бы один из трех перечисленных предметов?

б) Сколько студентов не изучают ни один из трех перечисленных предметов?

в) Сколько студентов изучают только математику?

г) Сколько студентов изучают физику или химию, но не изучают математику?

д) Сколько студентов не изучают ни математику, ни химию?

Поскольку 5 человек изучают все три предмета, а 15 человек — химию и физику, остаются 10 человек, изучающих химию и физику, но не изучающих математику. Аналогично, $25 - 5 = 20$ человек занимаются математикой и химией, но не физикой, и $20 - 5 = 15$ человек изучают математику и физику, но не изучают химию. Данную ситуацию изображает диаграмма Венна, приведенная на рис. 8.8.

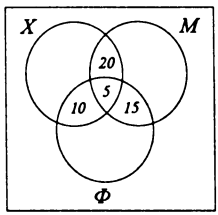


Рис. 8.8

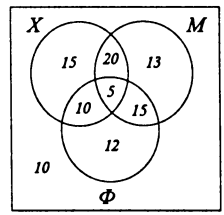


Рис. 8.9

Поскольку 50 студентов изучают химию, и 35 из них уже учтены, то оставшиеся 15 изучают только химию. Аналогично, 53 студента занимаются математикой, и 40 из них уже учтены, поэтому 13 человек изучают только математику. Наконец, 42 студента изучают физику, и 30 из них уже учтены, поэтому 12 человек изучают только физику.

а) Суммируя количество людей, принадлежащих семи непересекающимся подмножествам, получаем 90 тех, кто изучает хотя бы один из трех предметов.

б) Поскольку 90 из 100 студентов изучают хотя бы один предмет, то $100 - 90 = 10$ человек не изучают ни один из этих трех предметов.

в) Из диаграммы Венна следует, что 13 человек изучают только математику.

г) Тридцать семь студентов занимаются химией или физикой, но не изучают математику.

д) Из диаграммы Венна, изображенной на рис. 8.9, следует, что 75 человек изучают математику или физику. Поэтому $100 - 75 = 25$ студентов не изучают ни математику, ни физику. □

ПРИМЕР 8.18. Сколько положительных целых чисел, меньших 1001, делятся на 2, 3 или 5? Сколько положительных целых чисел, меньших 1001, не делятся на 2, 3 или 5? Имеется всего

$$\left\lfloor \frac{1001}{2} \right\rfloor = 500$$

целых чисел, которые делятся на 2. Имеется всего

$$\left\lfloor \frac{1001}{3} \right\rfloor = 333$$

целых числа, которые делятся на 3. Имеется всего

$$\left\lfloor \frac{1001}{5} \right\rfloor = 200$$

целых чисел, которые делятся на 5. Если целое число делится на 2 и на 3, то оно делится на 6, поэтому существует

$$\left\lfloor \frac{1001}{6} \right\rfloor = 166$$

целых чисел, которые делятся на 2 и на 3. Если целое число делится на 2 и 5, то оно делится на 10, поэтому существует

$$\left\lfloor \frac{1001}{10} \right\rfloor = 100$$

целых чисел, которые делятся на 2 и на 5. Если целое число делится на 3 и на 5, то оно делится на 15, поэтому имеется

$$\left\lfloor \frac{1001}{15} \right\rfloor = 66$$

целых чисел, которые делятся на 3 и на 5. Наконец, если целое число делится на 2, 3 и 5, то оно делится на 30, и существует

$$\left\lfloor \frac{1001}{30} \right\rfloor = 33$$

целых числа, которые делятся на 2, 3 и 5. Следовательно, количество чисел, которые делятся на 2, 3 или 5, равно $500 + 333 + 200 - 166 - 100 - 66 + 33 = 734$. Количество положительных чисел, которые не делятся ни на одно из указанных целых чисел, равно $1001 - 734 = 267$. □

ПРИМЕР 8.19. Сколько неотрицательных целых чисел, меньших 100 000, содержат цифры 3, 6 и 9? Пусть универсум $\boldsymbol{U}$ будет множеством всех неотрицательных целых чисел, меньших 100 000. Следовательно, $|\boldsymbol{U}| = 100\,000$. Пусть S_3 — множество всех целых чисел, меньших 100 000, которые не содержат цифру 3. Пусть S_6 — множество всех целых чисел, меньших 100 000, которые не содержат цифру 6. Пусть S_9 — множество всех целых чисел, меньших 100 000, не содержащих цифру 9. Нас интересует $|S_3' \cap S_6' \cap S_9'|$, но

$$|S_3' \cap S_6' \cap S_9'| = |\boldsymbol{U}| - |S_3 \cup S_6 \cup S_9|,$$

а

$$|S_3 \cup S_6 \cup S_9| = \\ = |S_3| + |S_6| + |S_9| - |S_3 \cap S_6| - |S_3 \cap S_9| - |S_6 \cap S_9| + |S_3 \cap S_6 \cap S_9|.$$

Для S_3 существуют девять вариантов выбора каждой из пяти цифр, поскольку только цифра 3 исключена. Поэтому $|S_3| = 9^5$. Аналогично, $|S_6| = |S_9| = 9^5$. Поскольку каждая цифра числа из множества $S_3 \cap S_6$ может быть любой, за исключением 3 и 6, то существуют восемь вариантов выбора каждой цифры, поэтому $|S_3 \cap S_6| = 8^5$. Аналогично, $|S_3 \cap S_9| = |S_6 \cap S_9| = 8^5$. Каждая цифра числа из $S_3 \cap S_6 \cap S_9$ может быть любой, исключая 3, 6 и 9, поэтому существуют семь вариантов выбора каждой цифры. Отсюда $|S_3 \cap S_6 \cap S_9| = 7^5$. Следовательно,

$$|S_3 \cup S_6 \cup S_9| = 9^5 + 9^5 + 9^5 - 8^5 - 8^5 - 8^5 + 7^5 = 95\,650$$

и

$$S_3' \cap S_6' \cap S_9' = |\boldsymbol{U}| - (S_3 \cup S_6 \cup S_9) = 100\,000 - 95\,650 = 4350. \qquad \square$$

■ УПРАЖНЕНИЯ

1. В классе из 200 студентов 120 человек изучают историю Канзаса 1995–96 гг., 110 занимаются тестированием продукта "Йогурт II", а 80 студентов делают и то, и другое. Сколько студентов изучают хотя бы один из этих курсов? Сколько студентов не изучают ни одного из них?
2. Сколько целых чисел между 1 и 401 делятся на 5 или на 7?
3. Сколько целых чисел между 1 и 401 делятся на 7 или на 11?
4. Сколько целых чисел между 1 и 401 делятся на 6 или на 10?
5. Сколько целых чисел между 1 и 401 делятся на 10 или на 15?
6. Сколько целых чисел между 1 и 1001 делятся на 10, но не делятся на 40?
7. Сколько целых чисел между 1 и 1001 делятся на 10, но не делятся на 14?
8. Год является високосным, если (а) количество дней в нем делится на 4, но не делится на 100, или (б) если он делится на 400. Сколько високосных годов было между 1001 и 2001 годами?
9. Сколькими способами можно разместить положительные целые числа, меньшие 10, так чтобы 4 было расположено сразу после 5 или 5 было расположено сразу после 4? Сколько существует размещений, в которых 4 и 5 не стоят рядом?

10. Сколькими способами можно разместить положительные целые числа, меньшие 10, так что 4 было бы расположено сразу после 3 или 7 было бы расположено сразу после 6?

11. Сколько существует положительных целых чисел, содержащих не более пяти цифр, в которых

а) первой цифрой является 3?

б) последней цифрой является 5?

в) первой цифрой является 3 или последней цифрой является 5?

г) ни первая цифра не равна 3, ни последняя цифра не равна 5?

12. Сколько целых чисел между 1 и 2003 делится на 3, 5 или 7?

13. Сколько целых чисел между 1 и 2003 делится на 5, 7 или 11?

14. Сколько целых чисел между 1 и 2003 делится на 4, 5 или 6?

15. Сколько целых чисел между 1 и 2003 делится на 6, 7 или 8?

16. В группе из 200 студентов 75 изучают математику, 70 — историю, 75 — социологию, 35 изучают математику и социологию, 20 — историю и социологию, 25 изучают математику и историю, 15 студентов — все три предмета.

а) Сколько студентов изучают хотя бы один из трех предметов?

б) Сколько студентов изучают только один из трех предметов?

в) Сколько студентов изучают историю или математику, но не изучают социологию?

г) Сколько студентов не изучают ровно два из трех предметов?

д) Сколько студентов не выбрали историю или математику?

17. Согласно опросу 250 телезрителей, 95 из них нравится смотреть новости, 125 предпочитают смотреть спорт, 125 — комедии, 25 — новости и комедии, 45 — спорт и комедии, 35 — новости и спорт, 5 любят все три вида программ.

а) Сколько телезрителей смотрят новости, но не смотрят спорт?

б) Сколько телезрителей смотрят новости или спорт, но не любят комедии?

в) Сколько телезрителей не любят смотреть ни новости, ни спорт?

г) Сколько телезрителей смотрят не только спорт?

д) Сколько телезрителей смотрят спорт и комедии, но не смотрят новости?

18. Сколько пятизначных целых чисел имеют одной из цифр 3, 5 или 7?

19. Сколько пятизначных целых чисел начинаются с 3 и заканчиваются на 5 или содержат цифру 7 ?

20. В группе из 100 студентов 35 изучают французский язык, 42 — испанский, 43 — немецкий, 17 изучают французский и испанский, 15 — испанский и немецкий, 13 — французский и немецкий, и 20 студентов не изучают ни один из трех языков.

а) Сколько студентов изучают французский или немецкий язык, но не изучают испанский?

б) Сколько студентов изучают только один из трех языков?

в) Сколько студентов изучают два из трех языков?

г) Сколько студентов не изучают ни испанский язык, ни французский?

д) Сколько студентов изучают только испанский?

8.3. ПЕРЕСТАНОВКИ И СОЧЕТАНИЯ

Переставляя объекты некоторого набора, мы обычно располагаем их в различном порядке. В этом смысле перестановка — это переупорядочение набора объектов, или функция, которая задает такое переупорядочение (см. главу 4, раздел 4.2.). Теперь будем исследовать, сколько существует способов переупорядочения набора объектов. Все необходимое для этого уже имеется. Для подсчета перестановок достаточно воспользоваться правилом умножения. Рассмотрим количество возможных способов формирования числа, переставляя цифры 1, 2, 3, 4 и 5. Варианты возможных перестановок — это, например, числа 12345, 15342, 32415 и 32415. Для нахождения количества возможных размещений или перестановок заметим, что первую цифру можно выбрать пятью способами, вторую — четырьмя способами, третью — тремя способами, четвертую — двумя, и только один вариант остается для выбора пятой цифры. Поэтому существует 5! возможных перестановок. Точно так же, если необходимо переупорядочить n объектов, то для этого существует $n!$ способов. В перестановках важен порядок. Числа 51342 и 32415, образованные перестановкой цифр 1, 2, 3, 4 и 5, не совпадают. Кроме того, поскольку перестановки рассматриваются как переупорядочения, то каждый объект можно использовать только один раз. Если бы повтор цифр допускался, то при формировании числа для каждой цифры существовало бы пять вариантов выбора, поэтому существовало бы 5^5 возможных чисел.

Мы будем также рассматривать перестановки, когда объектов больше, чем мест для их размещения. Это обобщает сформулированное выше понятие перестановки, так как в этом случае нельзя рассматривать перестановку как переупорядочение. Предположим, например, что в организации — 20 человек и из них требуется выбрать президента, вице-президента, секретаря и казначея. Имеются 20 вариантов выбора президента, 19 вариантов выбора вице-президента, 18 способов выбора секретаря и 17 — казначея. Таким образом, получаем $20\times19\times18\times17$ способов выбора должностных лиц. Заметьте, что порядок все еще остается существенным. Есть разница, является ли Мэри Браун президентом, вице-президентом, секретарем или казначеем. Каждое из размещений на должностях четырех выбранных лиц представляет собой различную организацию руководства. По-прежнему, любого человека можно избрать только один раз. Если Мэри Браун выбрана президентом, то ее нельзя избрать ни на одну из остальных трех должностей. Заметим также, что

$$20 \times 19 \times 18 \times 17 = \frac{20!}{16!} = \frac{20!}{(20-4)!},$$

поскольку все множители в числителе, не превышающие 16, сокращаются с множителями знаменателя, входящими в 16!.

Предположим, имеется n человек. Требуется выбрать r из них и расположить в определенном порядке. Существует n способов выбрать первого человека, $n-1$ способов выбора второго, $n-2$ способов выбора третьего, $n-j+1$ способов выбора j-го и $n-r+1$ способов выбора r-го человека. Следовательно, существует

$$(n)(n-1)(n-2)(n-3)\cdots(n-i+1)\cdots(n-r+1)$$

способов выбрать r человек из n. Но

$$(n)(n-1)(n-2)(n-3)\cdots(n-i+1)\cdots(n-r+1)=\frac{n!}{(n-r)!}.$$

Сформулируем вышесказанное в виде следующей теоремы.

ТЕОРЕМА 8.20. Количество способов выбрать r объектов с учетом порядка из n объектов равно

$$(n)(n-1)(n-2)(n-3)\cdots(n-i+1)\cdots(n-r+1)=\frac{n!}{(n-r)!}.$$

ОПРЕДЕЛЕНИЕ 8.21. Пусть $P(n,r)=\dfrac{n!}{(n-r)!}$. $P(n,r)$ называется числом ***перестановок из n объектов по r***.

Заметим, что если выбирать все n объектов и размещать их в определенном порядке, то $r=n$ и, поскольку $0!=1$, имеем

$$P(n,n)=\frac{n!}{(n-n)!}=\frac{n!}{0!}=n!.$$

Таким образом, приходим к уже известному результату для случая перестановок n элементов.

ПРИМЕР 8.22. Сколько различных четырехзначных чисел можно образовать из цифр 1, 2, 3, ..., 9, если все цифры в каждом четырехзначном числе различны? Для формирования каждого четырехзначного числа выбираем четыре цифры из девяти, поэтому существует

$$P(9,4)=\frac{9!}{(9-4)!}=\frac{9!}{5!}=3024$$

таких различных чисел. □

ПРИМЕР 8.23. Сколькими способами можно расставить в ряд для фотографирования пять мальчиков и шесть девочек, если ни две девочки, ни два мальчика не должны стоять рядом? Мальчиков меньше, чем девочек, поэтому первой должна стоять девочка, а дальше мальчики и девочки должны чередоваться. Итак, ряд должен иметь вид *ДМДМДМДМДМД*. Существуют 6! способов расположить девочек на позициях *Д* и 5! способов расположить мальчиков на позициях *М*. Следовательно, имеется $6!\times 5!$ способов расставить детей. □

ПРИМЕР 8.24. Сколькими способами можно расположить для фотографирования пять мальчиков и пять девочек, если ни две девочки, ни два мальчика не должны стоять рядом? В данной ситуации первым в ряду может быть либо мальчик, либо девочка. Если первой стоит девочка, то ряд имеет вид *ДМДМДМДМДМ*. Имеются 5! способов расставить девочек на позициях *Д* и 5! способов расставить мальчиков на позициях *М*. Поэтому, существуют $5!\times 5!$ способов расположить детей в ряд, если первой стоит девочка. Аналогично, существуют $5!\times 5!$ способов расположить детей в ряд, если первым стоит мальчик. Таким образом, имеются $2\times 5!\times 5!$ способов расположить детей в ряд для фотографирования. □

ПРИМЕР 8.25. Сколькими способами можно рассадить 10 человек за круглым столом, если имеет значение только порядок соседей. Существует несколько вариантов решения данной задачи. Первым делом, отметим, что вращение людей вокруг стола не меняет их взаимного расположения, поскольку соседи справа и слева остаются прежними. Предположим, что место за столом уникально. Тогда существует 10! способов рассадить людей за столом. Считаем, что при вращении места остаются теми же, так как соседи не меняются. Существует десять таких вращений, поэтому делим 10! на 10, что дает 9! способов расположить людей за столом, если имеет значение только порядок соседей.

Иной подход к решению задачи состоит в том, чтобы сначала усадить одного человека. Этим исключается вращение, а оставшихся 9 человек можно рассадить 9! способами. □

ПРИМЕР 8.26. На книжной полке требуется расположить 15 различных книг по математике, 12 различных книг по физике и 16 различных книг по информатике. Сколькими способами это можно сделать, если

а) не существует никаких ограничений?

б) все книги по одному и тому же предмету должны стоять вместе?

в) все книги по одному и тому же предмету должны стоять рядом, но математические книги и книги по информатике не должны стоять рядом?

а) Всего имеется 43 книги, поэтому существуют 43! различных способа расположить их на полке.

б) Предположим, что первыми на полке помещаем книги по математике, книги по физике помещаем вторыми и последними располагаем книги по информатике. Обозначим это расположение *МФИ*. Существуют 15! различных способов расставить книги по математике, 12! — по физике и 16! — по информатике. Следовательно, имеются $15! \times 12! \times 16!$ способов разместить книги. Если первыми поставить на полку книги по физике, затем книги по информатике и, наконец, книги по математике, что можно обозначить как *ФИМ*, то опять получили бы $15! \times 12! \times 16!$ способов разместить книги. Любая перестановка *И*, *М* и *Ф* даст $15! \times 12! \times 16!$ способов размещения книг. Поскольку существуют 3! способа переставить *И*, *М* и *Ф*, то существуют $3! \times 15! \times 12! \times 16!$ различных способов расставить книги на полке.

в) Поскольку книги по физике следует расположить посередине, книги по информатике должны стоять либо первыми, либо последними. Поэтому единственными возможными конфигурациями являются *ИФМ* и *МФИ*. Поскольку для каждой из них существуют $15! \times 12! \times 16!$ различных способов разместить книги, то в итоге получаем $2 \times 15! \times 12! \times 16!$ различных способов разместить книги. □

ПРИМЕР 8.27. Сколько существует перестановок букв $w, e, d, i, g, m, a, t, h$, в которых последовательности букв не образуют слова "we", "dig", и "math"? Например, перестановка $d, g, i, w, e, t, h, a, m$ не подходит, поскольку включает сочетание букв "we". Пусть универсум $\boldsymbol{U}$ будет множеством всех перестановок букв $w, e, d, i, g, m, a, t, h$. Тогда $|\boldsymbol{U}| = 9! = 362\,880$. Пусть S_1 — множество всех перестановок, в которых встречается слово "we". Пусть S_2 — множество всех переста-

новок, в которых встречается слово "dig", и S_3 — множество всех перестановок, которые включают слово "math". В задаче требуется найти $|S_1' \cap S_2' \cap S_3'|$ или $|\boldsymbol{U}| - |S_1 \cup S_2 \cup S_3|$. Но

$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3| - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| + \\ + |S_1 \cap S_2 \cap S_3|.$$

Поскольку буквы "w, e" должны стоять вместе, множество S_1 состоит из всех перестановок восьми символов we, d, i, g, m, a, t, h. Поэтому $|S_1| = 8!$. Множество S_2 состоит из всех перестановок семи символов w, e, dig, m, a, t, h. Следовательно, $|S_2| = 7!$. Множество S_3 состоит из всех перестановок шести символов $w, e, d, i, g, math$. Поэтому $|S_3| = 6!$. Множество $S_1 \cap S_2$ составляют все перестановки шести символов we, dig, m, a, t, h. Поэтому $|S_1 \cap S_2| = 6!$ Множество $S_2 \cap S_3$ составляют все перестановки четырех символов $w, e, dig, math$. Поэтому $|S_2 \cap S_3| = 4!$ Множество $S_1 \cap S_3$ составляют все перестановки пяти символов $we, d, i, g, math$. Поэтому $|S_1 \cap S_2| = 5!$ Наконец, множество $S_1 \cap S_2 \cap S_3$ составляют все перестановки трех символов $we, dig, math$. Поэтому $|S_1 \cap S_2 \cap S_3| = 3!$ и

$$|S_1 \cup S_2 \cup S_3| = 8! + 7! + 6! - 6! - 5! - 4! + 3! = \\ = 45\,222,$$

так что

$$|S_1' \cap S_2' \cap S_3'| = |\boldsymbol{U}| - |S_1 \cup S_2 \cup S_3| = \\ = 362\,880 - 45\,222 = \\ = 317\,658. \qquad \square$$

Вернемся к задаче выбора президента, вице-президента, секретаря и казначея в группе из 20 человек. Известно, что существуют

$$\frac{20!}{16!}$$

различных способов сделать выбор. В этом случае избрание Мэри Браун президентом, Джорджа Смита вице-президентом, Джейн Джонс — секретарем и Джо Джексона — казначеем отличается от избрания Джейн Джонс президентом, Джо Джексона — вице-президентом, Джорджа Смита — секретарем и Мэри Браун — казначеем. Допустим, что вместо выборов президента, вице-президента, секретаря и казначея избирается просто комитет из четырех членов. В этом случае избрание в комитет Мэри Браун, Джорджа Смита, Джейн Джонс и Джо Джексона ничем не отличается от избрания в комитет Джейн Джонс, Джо Джексона, Джорджа Смита и Мэри Браун. Теперь порядок больше не дает различий в способах выбора. При выборе четырех должностных лиц любая перестановка Мэри Браун, Джорджа Смита, Джейн Джонс и Джо Джексона была бы различной, но при избрании комитета они различаются. Поскольку существует 4! перестановки четырех должностных лиц, которые не меняют состав комитета, для нахождения числа комитетов необходимо $\frac{20!}{16!}$ разделить на 4!. Используя те же аргументы при выборе r объектов из n без учета порядка, получаем следующую теорему.

ТЕОРЕМА 8.28. Количество способов выбора r объектов из n объектов без учета порядка равно

$$\frac{n!}{(n-r)!r!}.$$

ОПРЕДЕЛЕНИЕ 8.29. Для $0 \leq r \leq n$ положим $C(n,r) = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$. $C(n,r)$ называется числом ***сочетаний из n объектов по r***.

ТЕОРЕМА 8.30. Для $0 \leq r \leq n$ имеем $C(n,r) = C(n,n-r)$.

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned} C(n,r) &= \frac{n!}{(n-r)!r!} = \\ &= \frac{n!}{(n-r)!(n-(n-r))!} = \\ &= C(n,n-r). \end{aligned}$$

■

Обратите внимание, что в случае сочетаний, как и в случае перестановок, при выборе r объектов каждый объект может быть выбран не более одного раза.

Если выбираются все n объектов без учета порядка, то $r = n$. Поскольку $0! = 1$, имеем

$$C(n,n) = \frac{n!}{(n-n)!(n!)} = \frac{n!}{(0!)(n!)} = 1,$$

поэтому существует только один способ выбрать n элементов: просто взять их все.

ПРИМЕР 8.31. Если множество содержит десять элементов, то сколько оно имеет трехэлементных подмножеств? Поскольку множество не упорядочено, выбираются три элемента из десяти, поэтому всего имеется

$$C(10,3) = \frac{10!}{7!3!} = 120$$

различных подмножеств. □

ПРИМЕР 8.32. Сколько строк длины девять содержат ровно 5 единиц и 4 нуля? Несмотря на то, что строки сами по себе — упорядоченные объекты, эту задачу можно решить, используя сочетания. В строке имеется девять мест для размещения 1 и 0. Можно выбрать любые пять из девяти мест для размещения единицы. Поэтому имеется

$$C(9,5) = \frac{9!}{5!4!} = 126$$

мест для размещения единицы. Как только единицы вставлены, остальные места заполняются нулями. И хотя для строки важен порядок элементов, порядок заполнения пяти мест единицами не существенен; как только эти места выбраны, их порядок уже не имеет значения. □

ПРИМЕР 8.33. Сколькими способами можно выбрать комитет, включающий 6 мужчин и 8 женщин, из группы, состоящей из 12 мужчин и 20 женщин? Существует

$$C(12,6) = \frac{12!}{6!6!}$$

способов выбора мужчин и

$$C(20,8) = \frac{20!}{8!12!}$$

способов выбора женщин. Поэтому, согласно комбинаторному принципу умножения, имеется

$$\frac{12!}{6!6!} \times \frac{20!}{8!12!} = 116396280$$

способов выбрать комитет. □

ПРИМЕР 8.34. Сколько существует вариантов выбора 5 карт из стандартной колоды, содержащей 52 карты? Поскольку порядок карт не имеет значения, речь идет о выборе 5 объектов из 52, поэтому существует

$$C(52,5) = \frac{52!}{5!47!} = 2598960$$

возможных комбинаций. □

ПРИМЕР 8.35. Сколькими способами можно вытянуть 5 карт трефовой масти из стандартной колоды, содержащей 52 карты? В колоде имеется 13 треф, из которых выбираются 5, поэтому существует

$$C(13,5) = \frac{13!}{5!8!} = 1287$$

возможных 5-карточных раскладов пяти треф. □

ПРИМЕР 8.36. Говорят, что 5-карточный расклад содержит каре, если четыре из них являются либо тузами, либо королями, либо дамами и т.д. Эти четыре карты называются картами одного ранга. Сколько существует раскладов, при которых пятерка карт содержит каре? Существуют 13 способов выбрать каре и 48 способов выбрать пятую карту. Поэтому существует $13 \times 48 = 624$ различных раскладов, включающих каре. □

ПРИМЕР 8.37. "Фулл хаус" содержит три карты одного ранга и две карты другого ранга. Например, расклад, содержащий три короля и две шестерки представляет собой "фулл хаус". Сколько существует 5-карточных раскладов с "фулл хаус"? Предположим, что "фулл хаус" составили три короля и две шестерки. Три короля выбираются из четырех, поэтому существуют $C(4,3) = 4$ способа выбрать трех королей. Две шестерки выбираются из четырех, поэтому существуют $C(4,2) = 6$ способов выбрать две шестерки. Поэтому, согласно комбинаторному принципу умножения существуют $4 \times 6 = 24$ способа выбрать трех королей и две шестерки или три карты одного ранга и две карты другого ранга. Существуют 13 способов выбрать три карты одного ранга и 12 способов выбрать две карты одного ранга. Поэтому существуют $13 \times 12 = 156$ различных способов сочетания рангов. Следовательно, существуют $156 \times 24 = 3744$ возможных 5-карточных раскладов с "фулл хаус". □

Рассмотрим разложение

$$(a+b)^5 = (a+b)(a+b)(a+b)(a+b)(a+b).$$

Каждое слагаемое в разложении является результатом выбора a или b в каждом сомножителе $(a+b)$ и последовательного их перемножения. Например, a^5 получено путем выбора a из каждого сомножителя. Если из первого сомножителя выбрано a и b выбрано из остальных сомножителей, то в результате получим ab^4. Предположим, что требуется найти коэффициент при a^3b^2. Слагаемое a^3b^2 получается при выборе трех a и двух b из пяти сомножителей. Поскольку существует

$$C(5,3) = \binom{5}{3} = \frac{5!}{3! \times 2!}$$

способов выбора трех a, то коэффициент при a^3b^2 равен $\binom{5}{3}$.

Обобщая результат, получаем следующую теорему.

ТЕОРЕМА 8.38. (Биномиальная теорема) Для произвольного положительного целого числа n справедливы равенства

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r} = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r.$$

ДОКАЗАТЕЛЬСТВО. Поскольку $a^r b^{n-r}$ получено в результате r-кратного выбора a и $n-r$-кратного выбора b из n сомножителей в выражении $(a+b)^n$, то коэффициент при $a^r b^{n-r}$ равен числу способов r-кратного выбора a из n сомножителей

$$\binom{n}{r}.$$

Второе равенство следует из того факта, что

$$\binom{n}{r} = \binom{n}{n-r}.$$

■

ПРИМЕР 8.39. Построим разложение $(2x+3y^2)^4$. Используя биномиальную теорему, находим

$$\begin{aligned}(2x+3y^2)^4 &= \binom{n}{0}(2x)^4 + \binom{4}{1}(2x)^3(3y^2) + \binom{4}{2}(2x)^2(3y^2)^2 + \\ &\quad + \binom{4}{3}(2x)(3y^2)^3 + \binom{4}{4}(3y^2)^4 = \\ &= (2x)^4 + 4(2x)^3(3y^2) + 6(2x)^2(3y^2)^2 + 4(2x)(3y^2)^3 + (3y^2)^4 = \\ &= 16x^4 + 96x^3y^2 + 216x^2y^4 + 216xy^6 + 81y^8.\end{aligned}$$

□

ПРИМЕР 8.40. Найдем коэффициент при x^4y^6 в разложении $(3x+4y)^{10}$. Используя биномиальную теорему находим, что слагаемое, содержащее x^4y^6, имеет вид

$$\begin{aligned}\binom{10}{4}(3x)^4(4y)^6 &= 210(3x)^4(4y)^6 = \\ &= 210\cdot 3^4\cdot 4^6x^4y^6 = \\ &= 69672960x^4y^6,\end{aligned}$$

поэтому коэффициент при x^4y^6 равен 69672960. □

ПРИМЕР 8.41. Покажем, что для любого положительного целого числа n,

$$2^n = \sum_{r=0}^{n}\binom{n}{r}.$$

Пусть $a = b = 1$, тогда по биномиальной теореме

$$(1+1)^n = \sum_{r=0}^{n}\binom{n}{r}1^r1^{n-r},$$

и результат очевиден. □

Следующая теорема имеет весьма широкое применение. Теорема, в частности, дает возможность построить треугольник Паскаля. В доказательстве, приведенном ниже, использована комбинаторная техника. Читателю предлагается доказать теорему, используя метод математической индукции.

ТЕОРЕМА 8.42. Для всех целых чисел r и n таких, что $1 \le r < n$,

$$C(n,r) = C(n-1,r-1) + C(n-1,r).$$

ДОКАЗАТЕЛЬСТВО. Пусть m — один из n объектов, из которых требуется выбрать r объектов. Из $C(n,r)$ способов, которыми можно выбрать r объектов, рассмотрим то количество случаев, когда m является одним из выбранных объектов, и количество случаев, когда m выбранным объектом не является. Их сумма должна равняться $C(n,r)$. (Почему?) Сначала рассмотрим количество случаев, когда m — один из выбранных объектов. Поскольку m уже выбрано, требуется выбрать $r-1$ объект из $n-1$ объектов. Существует $C(n-1,r-1)$ способов сделать этот выбор. Далее рассмотрим количество способов, при которых m не является одним из выбранных объектов. По-прежнему требуется выбрать r объектов, но теперь из $n-1$ объектов, учитывая, что m не может быть выбран. Другими словами, существует только $n-1$ объектов, из которых выбираются r объектов. Таким образом, существуют $C(n-1,r)$ способов сделать такой выбор. Складывая количество способов выбора в обоих случаях, получаем

$$C(n,r) = C(n-1,r-1) + C(n-1,r).$$ ■

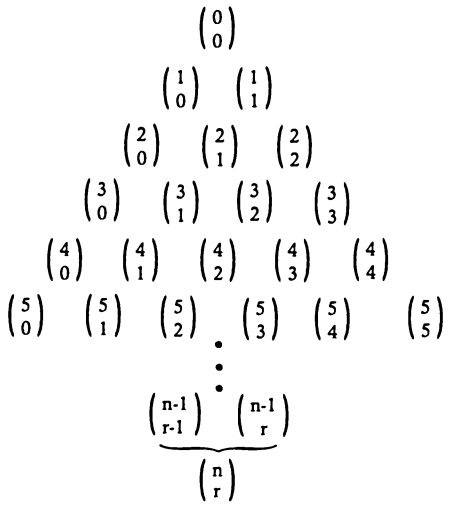


Рис. 8.10

Диаграмма, изображенная на рис. 8.10, известна как ***треугольник Паскаля***. Каждый из внутренних элементов треугольника равен сумме двух элементов, расположенных над ним, что является прямым следствием доказанной выше теоремы. Имеем

$$\binom{2}{1} = \binom{1}{0} + \binom{1}{1} \qquad \text{и} \qquad \binom{3}{2} = \binom{2}{1} + \binom{2}{2}.$$

В первом случае $n = 2$ и $r = 1$, а во втором случае $n = 3$ и $r = 2$. Можно заметить, что $(n + 1)$-ый ряд состоит из коэффициентов разложения $(a + b)^n$. Например,

$$(a+b)^4 = \binom{4}{0}a^4 + \binom{4}{1}a^3b + \binom{4}{2}a^2b^2 + \binom{4}{3}a^3b^3 + \binom{4}{4}b^4.$$

Эти коэффициенты приведены в пятом ряду треугольника. На рис. 8.11 изображен треугольник Паскаля с вычисленными элементами.

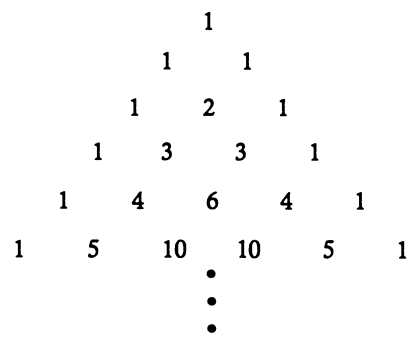


Рис. 8.11

ПРИМЕР 8.43. Найдем разложение $(a+b)^5$. Используя шестой ряд треугольника Паскаля, получаем

$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5. \qquad \square$$

Приведенная ниже теорема важна сама по себе и, кроме того, включает один весьма интересный частный случай.

ТЕОРЕМА 8.44. (Вандермонд) Пусть m, n и r — положительные целые числа такие, что $r \leq \min(m, n)$. Тогда

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}.$$

ДОКАЗАТЕЛЬСТВО. Левая часть равенства выражает количество способов выбора r объектов из $m+n$ объектов. Каким бы образом мы не выбирали r объектов из $m+n$ объектов, для некоторого $0 \leq k \leq r$ всегда k объектов выбираются из m объектов и $r-k$ объектов выбираются из n объектов. Для этого существуют

$$\binom{m}{k}\binom{n}{r-k}$$

способов. Обратно, если k объектов для любого $0 \leq k \leq r$ выбираются из m объектов и $r-k$ объектов выбираются из n объектов, то при этом r объектов выбираются из $m+n$ объектов. Следовательно,

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}. \qquad \blacksquare$$

Следующее утверждение позволяет находить сумму квадратов чисел, образующих строку треугольника Паскаля.

СЛЕДСТВИЕ 8.45. Для любого положительного целого числа n

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2.$$

ДОКАЗАТЕЛЬСТВО. Полагая в предыдущей теореме $n = m = r$, получаем

$$\binom{n+n}{n} = \sum_{k=0}^{n} \binom{n}{k}\binom{n}{n-k}$$

или

$$\binom{n}{k} = \binom{n}{n-k},$$

так что

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}\binom{n}{n-k} = \sum_{k=0}^{n} \binom{n}{k}^2. \qquad \blacksquare$$

■ УПРАЖНЕНИЯ

1. Вычислите

а) $P(8,5)$; **б)** $P(11,8)$; **в)** $C(12,7)$; **г)** $C(14,2)$; **д)** $C(14,12)$.

2. Вычислите

а) $P(8,3)$; **б)** $P(11,4)$; **в)** $C(15,5)$; **г)** $C(12,7)$; **д)** $C(12,5)$.

3. Сколько трехзначных чисел можно образовать, используя цифры 2, 3, 4, 5, 6, 8 и 9? А сколько таких трехзначных чисел меньше 450? Сколько среди них четных чисел? Сколько из них делятся на 4?

4. Сколькими способами девять человек могут расположиться в ряд?

5. К несчастью, судья на выставке цветов не разбирается в орхидеях. Если он выбирает победителей случайным образом среди 18 участниц, то сколько имеется способов вручить первый, второй и третий приз?

6. В скачках участвуют десять лошадей. Сколько существует вариантов призовой тройки лошадей?

7. Пять пар идут в кино. Сколькими способами они могут занять места, если

а) они могут сидеть в любом порядке?

б) все пять пар сидят подряд?

8. Шесть мальчиков и шесть девочек идут на концерт вместе. Сколькими способами они могут занять места, если

а) мальчики не будут сидеть рядом?

б) ни мальчики, ни девочки не будут сидеть все вместе?

в) все мальчики сядут вместе?

г) два мальчика сядут по краям?

д) один мальчик и одна девочка откажутся сесть вместе?

9. Сколько имеется шестизначных чисел, если первая цифра разряд может быть нулем, цифры не должны повторяться и

а) последние две цифры должны быть 7 или 8?

б) первая цифра должна быть 1, а последние цифры не могут быть 7 или 8?

в) цифры 7 и 8 должны стоять рядом?

г) число должно делиться на 4?

д) число должно делиться на 8?

е) в числе должны присутствовать цифры 5 и 6?

10. Сколько существует перестановок букв a, c, f, m, p, r, t и x, если

а) нет никаких ограничений?

б) между a и c должны стоять две или три буквы?

в) буквы a и c не должны быть разделены двумя или тремя буквами?

г) первые четыре буквы должны быть выбраны из a, c, f и m?

д) буквы a, c, f и m должны стоять рядом?

11. Сколько существует способов рассадить за круглым столом пятерых мужчин и пятерых женщин, если двое мужчин не должны сидеть рядом?

12. Сколько существует способов выбрать комитет из 5 человек в клубе, насчитывающем 25 членов?

13. Сколько существует способов составить комитет из 6 мужчин и 7 женщин, если организация состоит из 15 мужчин и 20 женщин?

14. Сколько существует 8-битовых строк, содержащих 3 нуля и 5 единиц?

15. Сколько существует способов вытащить 13 карт из стандартной колоды, содержащей 52 карты?

16. Сколько существует способов вытащить из колоды 13 карт, содержащих 6 карт одной масти?

17. Сколько существует способов вытащить из колоды 13 карт, содержащих 7 карт одной масти?

18. Сколько существует способов вытащить из колоды 13 карт, содержащих 8 карт одной масти?

19. Сколько существует способов вытащить из колоды 13 карт, содержащих 9 карт одной масти?

20. Сколько существует способов получить в 5-карточной раздаче ровно две пары (две карты одного ранга и две карты другого ранга, например, два туза и два короля)?

21. Сколько существует способов получить в 5-карточной раздаче три карты одного ранга (например, три десятки)?

22. Сколько существует способов получить в 5-карточной раздаче пять карт одной масти (флеш)?

23. Сколько существует способов разделить 10 человек на две команды по 5 человек для игры в баскетбол?

24. Пусть $A = \{a, b, c, d, e, f, g, h\}$. Сколько существует

а) трехэлементных подмножеств множества A?

б) пятиэлементных подмножеств множества A, содержащих b?

в) пятиэлементных подмножеств множества A, не содержащих b?

г) пятиэлементных подмножеств множества A, содержащих c, но не содержащих d и e?

д) подмножеств множества A, содержащих хотя бы три элемента?

е) подмножеств множества A, содержащих не более шести элементов?

25. Если монета подброшена 10 раз, то сколько существует способов выпадения четырех "решек" и шести "орлов"? Сколько существует способов выпадения не менее трех "решек"?

26. В зоомагазине продаются 5 черепах, 7 ящериц и 12 мышей. Сколько существует способов выбрать себе 2 черепахи, 3 ящерицы и 5 мышей?

27. Известно, что ответ на тест, состоящий из 30 вопросов, содержит 20 утвердительных ответов и 10 отрицательных. К сожалению, больше ничего не известно. Сколько существует вариантов ответа на тест, содержащих 20 утвердительных ответов на вопросы?

28. Если многоугольник имеет n сторон, то сколько у него диагоналей?

29. В команде из 20 человек каждый игрок одинаково хорошо играет на всех позициях. Сколько существует способов выбрать для начала игры команду из 9 человек?

30. Сколькими способами можно расставить игроков в предыдущей задаче?

31. Найдите разложение $(a + b)^8$, используя треугольник Паскаля.

32. В разложении $(2x + 3y)^{10}$ найдите коэффициент при x^6y^4.
33. В разложении $(3x - 4y)^{13}$ найдите коэффициент при x^8y^5.
34. В разложении $(x + 2y^2)^{13}$ найдите коэффициент при x^5y^{16}.
35. В разложении $(x^3 - 3y^2)^{10}$ найдите коэффициент при x^9y^{14}.
36. Попытайтесь найти числа Фибоначчи в треугольнике Паскаля и описать их местоположение. Указание: посмотрите на второй элемент второго ряда и первый элемент третьего ряда; затем посмотрите на второй элемент третьего ряда и первый элемент четвертого ряда. Наконец, посмотрите на третий элемент третьего ряда вместе со вторым элементом четвертого ряда и первым элементом пятого ряда.

8.4. ФОРМИРОВАНИЕ ПЕРЕСТАНОВОК И СОЧЕТАНИЙ

Пусть множество S, содержащее n элементов, линейно упорядочено, например, как целые числа или буквы алфавита, и требуется найти все перестановки данного множества. Это означает, что нас интересует множество строк, которые включают каждый элемент множества S, причем, только один раз. Для начала введем понятие ***лексикографического порядка***. Лексикографическим порядок называется потому, что он лежит в основе упорядочения слов в словаре. Например, слово *able* в английском словаре стоит перед словом *gray*, так как в алфавите буква a стоит перед буквой g. Слово *goose* стоит перед словом *gray*, потому что, хотя первые буквы совпадают, вторая буква слова *goose* стоит в алфавите перед второй буквой слова *gray*. Слово *seconded* предшествует слову *seconds*, потому что, хотя первые шесть букв совпадают, седьмая буква слова *seconded* стоит в алфавите перед седьмой буквой слова *seconds*. Точно так же число 12346 предшествует числу 12349.

ОПРЕДЕЛЕНИЕ 8.46. Для заданного линейно упорядоченного множества S ***лексикографический порядок*** $\preceq$ на строках элементов множества S определяется следующим образом: $a \preceq b$, если $a = a_1a_2a_3 \ldots a_m$, $b = b_1b_2b_3 \ldots b_n$, и выполнено одно из условий:

а) $a_1 < b_1$;
б) $a_i = b_i$ для всех $1 \leq i \leq k$ и $a_{k+1} < b_{k+1}$;
в) $a_i = b_i$ для всех $1 \leq i \leq m$ и $m < n$.

Теперь будем формировать множество всех перестановок элементов линейно упорядоченного множества. Для простоты рассмотрим только перестановки первых n целых чисел, отметив, что метод справедлив и для прочих случаев. Пусть строка $a_1a_2a_3 \ldots a_n$ обозначает перестановку

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ a_1 & a_2 & a_3 & \ldots & a_n \end{pmatrix}.$$

Таким образом, 51342 представляет перестановку

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 3 & 4 & 2 \end{pmatrix}.$$

Для нахождения всех перестановок необходимо иметь возможность сформировать перестановку, следующую за данной согласно лексикографическому порядку. Поскольку нас интересуют лишь строки одной длины, пункт (в) определения лексикографического порядка можно опустить. При лексикографическом порядке увеличение второй цифры изменит расположение строки в большей мере, чем изменение в пятой цифре. Аналогично, увеличение первой цифры изменит расположение строки в большей степени, чем изменения в седьмой цифре. Например, если начать со слова *dare*, то слово *read* находится от него дальше, чем слово *dear*. Поэтому для формирования следующей строки необходимо менять те символы, которые ближе всего к концу. Предположим, что $a_i > a_{i+1}$ для $m < a_i < n$. Если бы это условие выполнялось для всех неотрицательных m, то мы имели бы лексикографически наибольшую строку, что означало бы завершение процесса. Поэтому допустим, что существует такое m, что $a_i > a_{i+1}$ для $m < i < n$, но $a_m < a_{m+1}$. Если переупорядочить все a_i, где $i > m$, то мы создадим строку, которая будет лексикографически меньше. Например, любая перестановка последних четырех цифр в строке 12348765 создаст лексикографически меньшее число. Таким образом, следует увеличивать цифру a_m, не меняя при этом никакое a_i, где $i < m$, т.е. нужно увеличивать a_m, оставляя без изменения все a_i, где $i < m$. Поэтому выберем наименьшее число a_i такое, что $i > m$ и $a_m < a_i$. Меняем местами a_m и a_i. Затем переставляем в возрастающем порядке все a_i, следующие за новым a_m. Так мы получим наименьшее из чисел, больших исходного числа.

Например, рассмотрим число 12348765. Для получения следующего числа меняем местами 4 и 5, получая 12358764. Затем переставляем 8, 7, 6 и 4 в возрастающем порядке и получаем 12354678.

Пусть задано число 1437652. Для нахождения следующего числа меняем местами 3 и 5, получая 1457632. Теперь переставляем 7, 6, 3, 2 в возрастающем порядке, получая 1452367. Обратите внимание, что перестановка последних цифр свелась к записи их в обратном порядке, поскольку до этого они находились в убывающем порядке. Таким образом, приходим к следующей процедуре:

Процедура **Формирование перестановки** $(a_1, \ldots, a_n)$
Найти m такое, что $a_i > a_{i+1}$ для $m < a_i < n$, но $a_m < a_{m+1}$;
Выбрать наименьшую цифру a_i, такую что $i > m$ и $a_m < a_i$;
Поменять местами a_m и a_i;
Переупорядочить все a_i, стоящие после нового a_m,
в возрастающем порядке;
Конец процедуры.

Теперь рассмотрим метод формирования всевозможных сочетаний из n элементов множества S. Мы не будем использовать лексикографический порядок элементов множества S, а воспользуемся лексикографическим порядком на строках бинарных чисел. Напомним, что любому сочетанию элементов из n элементов соответствует бинарная строка длины n. Если множество S задано в виде

$\{a_1, a_2, a_3, \ldots, a_n\}$, то единица на месте k-го бита в бинарной строке показывает, что a_k присутствует в выбранном подмножестве, в то время как 0 в том же месте бинарной строки показывает, что a_k не включено в выбранное подмножество. Например, если множество S представляет собой $\{a_1, a_2, a_3, a_4, a_5\}$, то 10110 соответствует выбору подмножества $\{a_1, a_3, a_4\}$, а 10001 — выбору подмножества $\{a_1, a_5\}$. Таким образом, для формирования всех сочетаний из множества n элементов достаточно сформировать все бинарные строки длины n. С этой целью проще всего сосчитать от 0 до $2^n - 1$, используя бинарные строки длины n.

Эквивалентный метод состоит в том, чтобы показать, как по данной бинарной строке длины n построить следующую бинарную строку такой же длины. Для этого достаточно найти в строке первый 0 справа, заменить его на 1, а все элементы справа от новой единицы заменить на 0. Как говорилось ранее, формирование сочетаний не является лексикографическим упорядочением объектов множества, а представляет собой лексикографическое упорядочение бинарных строк.

ПРИМЕР 8.47. Пусть $S = \{a, b, c, d, e\}$. Найдем сочетание, следующее за $\{a, d, e\}$. Поскольку сочетанию $\{a, d, e\}$ соответствует бинарная строка 10011, следующей строкой будет 10100. Поэтому следующим сочетанием будет $\{a, c\}$. □

Рассмотрим, наконец, как сформировать все сочетания по r элементов, выбранных из элементов линейно упорядоченного множества S. Снова воспользуемся лексикографическим упорядочением элементов множества S. Как и ранее, не ограничивая общности, можем положить $S = \{1, 2, 3, \ldots n\}$. Будем предполагать, что сочетания перечислены в возрастающем порядке. Как и для перестановок, покажем, как по данному сочетанию найти следующее согласно лексикографическому порядку. Предположим, что $n = 5$ и $r = 3$. Если можно увеличить последнюю цифру, то так и будем делать. Поэтому, если имеем число 123, то его можно заменить на 124. Если имеем 125, то последнюю цифру увеличивать нельзя. Отсюда переходим к следующей цифре и смотрим, можно ли ее увеличить. В данном случае это можно сделать, поменяв 2 на 3. Однако мы стремимся построить наименьшее число из тех, которые больше 125. Поэтому увеличиваем последнюю цифру на 1. Имеем первые две две цифры 1 и 3, значит, следующее число — 134. Предположим, что имеется число 145. Последнюю и предпоследнюю цифры нельзя увеличивать. Однако, 1-ю цифру увеличить можно, поэтому 1 увеличиваем до 2. Чтобы сделать число минимальным, в качестве последних цифр возьмем 3 и 4. В результате получаем число 234.

Если начинать справа, то значение последней цифры будет наибольшим возможным, если оно равно $n = n - r + r$. Если последняя цифра — наибольшая возможная, то предыдущая цифра будет наибольшей возможной, если она равна $n - r + (r - 1)$ или $n - r + i$, где $i = r - 1$ — позиция предыдущей цифры. В общем случае, значение каждой i-ой цифры будет наибольшим возможным, если цифры справа — наибольшие возможные, и это значение равно $n - r + i$. Для формирования максимальной перестановки идем справа налево и проверяем, равно ли значение i-го элемента величине $n - r + i$. Первое значение, которое не удовлетворяет этому условию, можно увеличить. Например, это значение m на j-ом месте. Увеличиваем m на 1 и затем увеличиваем значение каждого элемента, стоящего после j-го, на 1 значения предыдущего элемента. В результате

приходим к следующей процедуре:

Процедура **r-сочетание (n, r)**
Начиная справа, найти первую цифру со значением a_i,
таким что $a_i \neq n - r + i$;
Прибавить единицу к этому значению a_i;
Начиная с a_{i+1}, увеличиваем его значение на 1;
Конец процедуры.

■ УПРАЖНЕНИЯ

1. Определите перестановку, следующую за 21435.
2. Определите перестановку, следующую за 621435.
3. Определите перестановку, следующую за 21453.
4. Определите перестановку, следующую за 416532.
5. Пусть $S = \{a, b, c, d, e\}$. Определите сочетание, следующее за $\{a, c, d\}$.
6. Пусть $S = \{a, b, c, d, e\}$. Определите сочетание, следующее за $\{a, c, d, e\}$.
7. Пусть $S = \{1, 2, 3, 4, 5\}$. Определите сочетание, следующее за $\{1, 4, 5\}$.
8. Пусть $S = \{1, 2, 3, 4, 5, 6\}$. Определите сочетание, следующее за $\{1, 3, 5, 6\}$.
9. Пусть $S = \{1, 2, 3, 4, 5, 6\}$. Определите сочетание, следующее за $\{1, 6\}$.
10. Пусть $S = \{1, 2, 3, 4, 5, 6\}$. Определите трехэлементное сочетание, следующее за $\{1, 5, 6\}$.
11. Пусть $S = \{1, 2, 3, 4, 5, 6\}$. Определите трехэлементное сочетание, следующее за $\{1, 4, 6\}$.
12. Пусть $S = \{1, 2, 3, 4, 5, 6\}$. Определите четырехэлементное сочетание, следующее за $\{1, 4, 5, 6\}$.
13. Пусть $S = \{1, 2, 3, 4, 5, 6\}$. Определите четырехэлементное сочетание, следующее за $\{1, 3, 4, 5\}$.
14. Перечислите все перестановки $1, 2, 3, 4$ в лексикографическом порядке.
15. Перечислите все перестановки x, y, z в лексикографическом порядке.
16. Сформируйте все трехэлементные подмножества множества $\{1, 2, 3, 4, 5\}$.
17. Сформируйте все трехэлементные подмножества множества $\{a, b, c, d, e\}$.
18. Найдите первый и последний элементы в списке перестановок целых чисел $\{1, 2, 3, 4, 5, 6, 7\}$.
19. Найдите первый и последний элементы в списке перестановок букв английского алфавита.
20. Найдите первый и последний элементы в списке трехэлементных подмножеств множества целых чисел $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.
21. Найдите первый и последний элементы в списке пятиэлементных подмножеств множества букв английского алфавита.
22. Найдите первый и последний элементы в списке пятиэлементных подмножеств множества букв английского алфавита.
23. Используя алгоритмы данного раздела, разработайте алгоритм формирования $P(n, r)$ всех перестановок r из n элементов.
24. Используя алгоритмы предыдущего упражнения, сформируйте перестановки трехэлементных подмножеств множества $\{a, b, c, d, e\}$.

8.5. ВВЕДЕНИЕ ВЕРОЯТНОСТИ

Много лет назад студент Стэнфордского университета послал письмо в бюро погоды города Сан-Франциско с вопросом, что они имели в виду, прогнозируя выпадения дождя с вероятностью 60%. Означает ли это, что дождь будет идти на 60% всей площади? Означает ли это, что дождь будет идти 60% времени? "Это означает, — ответило бюро, — что из десяти сотрудников шесть человек утверждали, что пойдет дождь". Вскоре мы увидим, что все на самом деле зависит от выборочного пространства. Многие задачи теории вероятностей и комбинаторики сконцентрированы вокруг азартных игр. Одна из причин заключается в том, что математики, разработавшие значительную часть данной теории, были наняты компаниями, контролировавшими игорный бизнес, для расчетов вероятностей выигрыша или проигрыша в тех или иных азартных играх. Многие из этих игр служат замечательным примером использования вероятности. Вероятность позволит читателю оценить, как много он может потерять, если будет неверно выполнять свое домашнее задание.

Наша концепция вероятности будет зависеть от того, что мы понимаем под экспериментом. Оставив термин "эксперимент" неопределенным, зададимся целью, чтобы он отвечал следующим условиям:

а) эксперимент дает более одного исхода;

б) исход точно не известен;

в) конечное множество всех исходов может быть определено до начала эксперимента;

г) эксперимент можно повторить при тех же условиях.

Примером эксперимента является подбрасывание двух монет. В этом случае множество исходов

$$S = \{(H,H),(H,T),(T,H),(T,T)\},$$

где H обозначает падение монеты вверх "орлом", а T — падение монеты "решкой" вверх. Другим примером эксперимента служит подбрасывание монеты и игральной кости. В этом случае множество результатов

$$S = \{(H,1),(H,2),(H,3),(H,4),(H,5),(H,6),(T,1),(T,2),(T,3),\\ (T,4),(T,5),(T,6)\},$$

где за результатом выпадения монеты обозначен результат выпадения кости.

ОПРЕДЕЛЕНИЕ 8.48. ***Выборочное пространство***, или ***множество элементарных событий*** — это множество всех исходов эксперимента. ***Событие*** — подмножество выборочного пространства.

Таким образом, в каждом из приведенных выше примеров множество S — это выборочное пространство. На языке теории множеств выборочное пространство S является универсом.

Интуитивно, вероятность события — это вероятность того, что событие наступит. Если эксперимент повторяется несколько раз, то вероятность события — это частота наступления события. В этом и следующих разделах будем предполагать, что все исходы эксперимента равновозможны. Заметим, что это вовсе не означает, что все события в мире равновозможны. Просто такое предположение дает нам возможность использовать комбинаторные принципы, поскольку вероятность — одно из основных приложений комбинаторики. С учетом этого предположения сформулируем определение вероятности.

ОПРЕДЕЛЕНИЕ 8.49. Пусть S — множество всех возможных исходов эксперимента. ***Вероятность*** $P(A)$ наступления события A определяется как отношение

$$P(A) = \frac{|A|}{|S|}.$$

ПРИМЕР 8.50. Пусть S — множество всех исходов подбрасывания двух костей. Пусть A — множество всех исходов, при которых сумма показаний двух костей равна 6. Таким образом,

$$A = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}.$$

так что $|A| = 5$. Как было показано выше, если существуют шесть возможных исходов выпадений для первой кости и шесть возможных исходов выпадений для второй кости, то множество S содержит 6×6 возможных исходов. Поэтому

$$P(A) = \frac{|A|}{|S|} = \frac{5}{36}.$$

□

ПРИМЕР 8.51. Пусть S — множество всех исходов при подбрасывании двух монет. Пусть A — множество исходов, когда обе монеты выпадают одинаково. Таким образом, $A = \{(H,H),(T,T)\}$, $S = \{(H,H),(H,T),(T,H),(T,T)\}$ и

$$P(A) = \frac{|A|}{|S|} = \frac{2}{4} = \frac{1}{2}.$$

□

Означает ли это, что если мы подбросим две монеты четыре раза, то на двух монетах выпадут одновременно "орлы" или "решки" точно два раза? Конечно, нет. Однако, если мы подбросим две монеты 400 раз, то можно ожидать, что монеты выпадут надлежащим образом порядка 200 раз.

Предположив опять, что все исходы равновозможны, и воспользовавшись приемами комбинаторики, можем доказать приведенные ниже теоремы.

ТЕОРЕМА 8.52. Если A и B — непересекающиеся множества исходов эксперимента, то $P(A \cup B) = P(A) + P(B)$.

ДОКАЗАТЕЛЬСТВО. Используя принцип сложения, имеем

$$\begin{aligned} P(A \cup B) &= \frac{|A \cup B|}{|S|} = \\ &= \frac{|A| + |B|}{|S|} = \\ &= \frac{|A|}{|S|} + \frac{|B|}{|S|} = \\ &= P(A) + P(B). \end{aligned}$$

■

Следующая теорема определяет некоторые из свойств вероятности события.

ТЕОРЕМА 8.53. Если S — выборочное пространство, A' — дополнение множества A до множества S как универса, то

а) $P(S) = 1$.
б) $P(A') = 1 - P(A)$.
в) $P(\varnothing) = 0$.
г) Для каждого события A имеем $P(A) \geq 0$.
д) Если $A \subseteq B$, то $P(A) \leq P(B)$.
е) Для каждого события A имеем $0 \leq P(A) \leq 1$.

ДОКАЗАТЕЛЬСТВО. а) $P(S) = \dfrac{|S|}{|S|} = 1$.

б) Согласно пункту (а) и предыдущей теореме имеем

$$1 = P(S) = P(A \cup A') = P(A) + P(A').$$

Поэтому $P(A') = 1 - P(A)$.

в) Согласно пункту (б) имеем $P(\varnothing) = 1 - P(\varnothing') = 1 - P(S) = 1 - 1 = 0$.

г) Поскольку должно существовать не менее двух исходов, то $|S| > 0$. Кроме того, $|A| \geq 0$, учитывая, что множество A содержит некоторое число элементов. Следовательно,

$$P(A) = \frac{|A|}{|S|} \geq 0.$$

д) Если $A \subseteq B$, то $B = A + (B - A)$. Поэтому

$$\begin{aligned} P(B) &= P(A + (B - A)) = \\ &= P(A) + P(B - A) \end{aligned}$$

и

$$P(B) - P(A) = P(B - A).$$

Однако, согласно пункту (г),

$$P(B - A) \geq 0.$$

Поэтому

$$P(B) - P(A) \geq 0$$

и

$$P(A) \leq P(B).$$

е) Поскольку $A \subseteq S$, то $P(A) \leq P(S) = 1$. Объединяя с пунктом (г), получаем

$$0 \leq P(A) \leq 1.$$ ■

ПРИМЕР 8.54. Найдем вероятность того, что при подбрасывании двух монет обе они не выпадут "решкой". Пусть A — множество исходов, когда обе монеты выпадают не "решкой". Тогда A' — множество исходов, когда обе монеты выпадают "решкой". Поскольку существуют четыре возможных исхода, и только при одином из них выпадает две "решки", то

$$P(A') = \frac{1}{4} \text{ и } P(A) = 1 - \frac{1}{4} = \frac{3}{4}.$$ □

ПРИМЕР 8.55. Найдем вероятность того, что при подбрасывании двух костей сумма показаний не равна семи. Пусть A — множество результатов, когда сумма показаний не равна семи. Тогда A' — множество исходов, когда сумма показаний равна семи. Общее число исходов при подбрасывании двух костей равно 6×6. Исходы, при которых сумма равна семи, включают $(1,6)$, $(2,5)$, $(3,4)$, $(4,3)$, $(5,2)$ и $(6,1)$. Таким образом $|A'| = 6$, так что

$$P(A') = \frac{6}{36} = \frac{1}{6} \quad \text{и} \quad P(A) = 1 - \frac{1}{6} = \frac{5}{6}.$$ □

Для полноты изложения приведем следующую теорему. Принципы, лежащие в основе этой теоремы, более подробно рассмотрены в разделе, посвященном принципу включения-исключения. Данная теорема представляет собой обобщение теоремы 8.6.

ТЕОРЕМА 8.56. Если A и B — множества исходов эксперимента, то

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

ДОКАЗАТЕЛЬСТВО. На рис. 8.12 показано, что

$$A = (A - B) \cup (A \cap B),$$
$$B = (B - A) \cup (A \cap B),$$
$$\text{и } A \cup B = (A - B) \cup (A \cap B) \cup (B - A),$$

где $(A - B)$, $(A \cap B)$ и $(B - A)$ — попарно непересекающиеся множества.

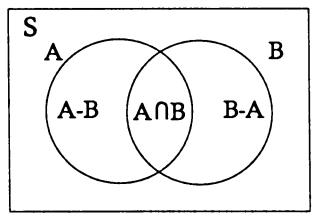


Рис. 8.12

Поэтому, по теореме 8.6 имеем

$$P(A) = P((A - B) \cup (A \cap B)) = P(A - B) + P(A \cap B),$$
$$P(B) = P((B - A) \cup (A \cap B)) = P(B - A) + P(A \cap B),$$

и

$$\begin{aligned} P(A \cup B) &= P((A - B) \cup (A \cap B) \cup (B - A)) = \\ &= P(A - B) + P(A \cap B) + P(B - A). \end{aligned}$$

Следовательно,

$$\begin{aligned} P(A) + P(B) &= (P(A - B) + P(A \cap B)) + (P(B - A) + P(A \cap B)) = \\ &= (P(A - B) + P(A \cap B) + (P(B - A)) + P(A \cap B) = \\ &= P(A \cup B) + P(A \cap B), \end{aligned}$$

так что $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. ■

ПРИМЕР 8.57. Найдем вероятность того, что при случайном выборе номерного знака для автомашины, содержащего три буквы и три цифры, совпадут три буквы или совпадут три цифры. Пусть S — множество всех номерных знаков, содержащих три буквы и три цифры. Пусть A — множество исходов, когда три буквы совпадают, а B — множество исходов, когда совпадают три цифры. Если три буквы совпадают, то существует 26 способов выбора букв и

$$10 \times 10 \times 10 = 1000$$

способов выбора цифр. Следовательно, $|A| = 26000$. Если цифры совпадают, то существует

$$26 \times 26 \times 26 = 17576$$

способов выбора букв и 10 способов выбора цифр. Поэтому $|B| = 175760$. Если все буквы одинаковые и все цифры одинаковые, то существует 26 способов выбора букв и 10 способов выбора цифр. Поэтому $|A \cap B| = 260$. Уже было показано, что существуют

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 = 17576000$$

различных способов создать номерной знак для автомашины, включающий три буквы и три цифры. Следовательно,

$$P(A) = \frac{26{,}000}{17\ 576\ 000}, P(B) = \frac{175{,}760}{17\ 576\ 000} \quad \text{и} \quad P(A \cap B) = \frac{260}{17\ 576\ 000},$$

так что

$$\begin{aligned} P(A \cup B) &= \frac{26\ 000}{17\ 576\ 000} + \frac{175\ 760}{17\ 576\ 000} - \frac{260}{17\ 576\ 000} = \\ &= \frac{201\ 500}{17\ 576\ 000} = \\ &= \frac{2\ 015}{175\ 760} \approx .011465. \end{aligned}$$

□

■ УПРАЖНЕНИЯ

1. Подбрасываются две игральные кости. Найдите вероятность того, что сумма выпавших очков равна

а) 4; **б)** 7; **в)** 8; **г)** 9;

д) 10 или больше.

Найдите вероятность того, что на обеих костях выпадет одинаковое количество очков. Найдите вероятность того, что сумма выпавших очков будет четна.

2. Карта вытягивается из стандартной колоды, содержащей 52 карты. Найдите вероятность вытягивания

а) валета, королевы, короля или туза (картинка);

б) пиковой карты;

в) королевы или пиковой карты;

г) картинки или червовой карты.

3. Из урны, в которой лежат 12 красных, 8 синих, 6 белых и 10 желтых шаров, вытягивается шар. Найдите вероятность вытягивания

а) красного шара;

б) белого или желтого шара;

в) не белого шара;

г) шара, который не будет ни синим, ни белым.

4. Буква случайным образом выбирается из английского алфавита. Какова вероятность выбора гласной буквы?

5. Положительное целое число, меньшее 302, выбирается случайным образом. Найдите вероятность того, что это число

а) делится на 3;

б) делится на 3 и на 7;

в) делится на 3 или на 7;

г) не делится на 3 или на 7.

6. Положительное целое число, меньшее 101, выбирается случайным образом. Найдите вероятность того, что это число

а) делится на простое число;

б) есть произведение двух (не обязательно различных) простых чисел;

в) есть произведение двух различных простых чисел;

г) есть произведение трех различных простых чисел.

7. Пять положительных целых чисел, меньших 10, выбираются случайным образом. Какова вероятность того, что все они различны ?

8. Если номерной знак автомобиля содержит три буквы и последующие три цифры, то какова вероятность, что сумма трех цифр четна? Какова вероятность, что все буквы гласные? Какова вероятность, что хотя бы одна из букв не является гласной?

9. После опроса 250 телезрителей оказалось, что 86 человек любят смотреть новости, 100 человек предпочитают спорт, 90 человек — комедии, 23 человека любят смотреть новости и комедии, 35 смотрят спорт и комедии, 33 человека

смотрят новости и спорт, 10 человек смотрят все три вида передач. Для интервью телезритель выбирается случайным образом.

а) Какова вероятность, что он смотрит новости, но не смотрит спорт?

б) Какова вероятность, что он смотрит новости или спорт, но только не комедии?

в) Какова вероятность, что он не смотрит ни новости, ни спорт?

г) Какова вероятность, что он смотрит новости и спорт, но только не комедии?

10. Если у игрока на руках 13 карт, то какова вероятность того, что он имеет

а) 5 карт одной масти?

б) 6 карт одной масти?

в) 7 карт одной масти?

г) 8 карт одной масти?

д) 9 карт одной масти?

11. У игрока на руках пять карт. Какова вероятность того, что среди них

а) две карты одного ранга (например, два короля)?

б) три карты одного ранга (например, три туза)?

в) на руки стрит-флэш (например, $2, 3, 4, 5, 6$ одной масти)?

г) на руки флэш-рояль (например, $A, K, Q, J, 10$ одной масти)?

д) на руки флэш (например, пять червей)?

е) на руки стрит (например, $2, 3, 4, 5, 6$ одной масти)?

12. В группе из 100 студентов, 40 изучали французский, 50 — испанский, 40 — немецкий, 12 студентов изучали французский и испанский, 13 — испанский и немецкий, 15 — французский и немецкий, и 10 студентов не изучали ни один из трех языков. Если студент выбран случайным образом, то

а) какова вероятность, что он изучал только французский?

б) какова вероятность, что он изучал испанский или немецкий?

в) какова вероятность, что он не изучал французский?

г) какова вероятность, что он изучал немецкий, но не изучал испанский?

д) какова вероятность, что он не изучал ни один из трех языков?

13. В трех урнах находятся: четыре красных шара и пять синих, шесть красных шаров и три синих, пять красных шаров и два синих, соответственно. Если из трех урн случайным образом выбран один шар, то какова вероятность, что этот шар будет красным?

14. Какова вероятность того, что четырехзначное число, не начинающееся с нуля, имеет среди цифр 3, 5 или 7?

15. Какова вероятность того, что четырехзначное число, не начинающееся с нуля, начинается с цифры 3, заканчивается на 5 или одна из его цифр — 7?

16. Произвольно выбирается трехбуквенное слово, буквы не повторяются. Какова вероятность, что буквы расположены в алфавитном порядке?

17. Какова вероятность того, что произвольное четырехзначное число содержит цифру 7? Какова вероятность того, что оно содержит цифру 4 или 9?

8.6. ОБОБЩЕННЫЕ ПЕРЕСТАНОВКИ И СОЧЕТАНИЯ

Рассмотрим теперь разновидность подсчетов, которые представляют собой нечто промежуточное между перестановками и сочетаниями. На самом деле это обобщение и того, и другого. Предположим, что требуется определить различные размещения букв в слове *миссиссиппи*. Проблема состоит в том, что некоторые буквы повторяются. Метод подсчета количества различных размещений букв будет весьма похож на метод подсчета количества сочетаний. Предположим сначала, что все буквы различны. В таком случае существовало бы 11! способов размещения 11 букв. Теперь обратим внимание, что буква "и" встречается четыре раза. Допуская, что все буквы различны, мы предположили, что каждое размещение букв "и", которое оставляет все другие буквы на своих местах, отличается от других таких размещений. Теперь все такие размещения нужно считать одинаковыми. Таким образом, 4! размещения четырех букв "и" на самом деле являются одним размещением. Для подсчета количества размещений 11 букв, когда 4 буквы "и" уже не рассматриваются как различные, нужно 11! разделить на 4!. Точно так же, для подсчета количества размещений, когда четыре буквы "с" не рассматриваются как различные, нужно еще раз разделить на 4!. Осталось рассмотреть две буквы "п". Как и предыдущих случаях, каждое размещение букв "п", которое оставляло все прочие буквы на своих местах и рассматривалось как уникальное, теперь рассматривается как одно и то же размещение. Поскольку существуют 2! способов размещения двух "п", то для подсчета количества размещений, когда две буквы "п" не различаются, нужно разделить еще на 2!. Таким образом, получаем

$$\frac{11!}{4!4!2!}$$

различных размещений.

Заметим, что делая буквы неразличимыми, мы игнорируем их порядок, как это было в сочетаниях.

ТЕОРЕМА 8.58. Пусть множество S содержит n объектов k таких различных типов, что имеется n_1 неразличимых объектов типа 1, n_2 неразличимых объектов типа 2, n_3 неразличимых объектов типа 3 и, вообще, n_i неразличимых объекта типа i. Пусть $P(n; n_1, n_3, n_3, \ldots, n_k)$ — количество различных размещений элементов множества S. Тогда

$$\begin{aligned} P(n; n_1, n_2, \ldots, n_k) &= C(n, n_1)C(n - n_1, n_2)C(n - n_1 - n_2, n_3) \\ &\quad \cdots C(n - n_1 - n_2 - \cdots - n_{k-1}, n_k) \\ &= \frac{n!}{n_1!n_2!n_3! \cdots n_k!}. \end{aligned}$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим первое уравнение. Имеется n мест, которые можно заполнить элементами из S. Существует $C(n, n_1)$ способов выбрать места для n_1 неразличимых объектов типа 1. Если эти места выбраны, то для заполнения останется $n - n_1$ мест, поэтому имеются $C(n - n_1, n_2)$ способов выбрать места для n_2 неразличимых объектов типа 2. Если объекты типа 1 и типа 2 выбраны, то для заполнения остается $n - n_1 - n_2$ мест, поэтому объекты типа 3

можно разместить $C(n-n_1-n_2,n_3)$ способами. Аналогично, объекты типа i для $2 \le i \le k$ можно выбрать $C(n-n_1-n_2-n_3-\cdots-n_{i-1},n_i)$ способами. Используя комбинаторный принцип умножения, имеем

$$C(n,n_1)C(n-n_1,n_2)C(n-n_1-n_2,n_3)\cdots C(n-n_1-n_2-\cdots-n_{k-1},n_k)$$

способов выбора различных размещений элементов из S.

Чтобы показать, что

$$P(n;n_1,n_2,\ldots,n_k)=\frac{n!}{n_1!n_2!n_3!\cdots n_k!},$$

воспользуемся рассуждением, которое использовалось неоднократно. Сначала предположим, что все n объектов из S различны. Если это так, то имеется $n!$ способов разместить данные объекты. Для $1 \le i \le k$, n_i объектов являются неразличимыми. Поэтому способы расположения таких объектов, при которых остальные объекты остаются на своих местах, неразличимы. Поскольку имеется $n_i!$ таких расположений, то для нахождения количества различных размещений, когда все n_i объектов типа i являются неразличимыми, необходимо $n!$ разделить на $n_i!$ для каждого i. Это дает

$$\frac{n!}{n_1!n_2!n_3!\cdots n_k!},$$

что и требовалось доказать. ■

ПРИМЕР 8.59. Предположим, что 12 книг, включающих 4 одинаковых учебника по математике, 6 одинаковых учебников по информатике, 2 одинаковых учебника по химии, следует расставить на полке. Сколькими способами это можно сделать? Используя предыдущую теорему, имеем

$$P(12;4,6,2)=\frac{12!}{4!6!2!}=13860.$$ □

ПРИМЕР 8.60. Сколькими различными способами можно расставить буквы в слове *succeeded*? Девять букв образуют пять типов неразличимых объектов, включающих: (1) букву s, (2) букву u, (3) два вхождения буквы c, (4) три вхождения буквы e и (5) два вхождения буквы d. Следовательно, количество неразличимых объектов равно

$$P(9;1,1,2,3,2)=\frac{9!}{1!1!2!3!2!}=15120.$$

Заметим, что типы, состоящие только из одного объекта, не влияют на конечный результат. Поэтому следует рассматривать типы, содержащие более одного объекта. □

Теперь обобщим биномиальную теорему на случай нахождения коэффициентов разложения $(x_1+x_2+x_3+\cdots+x_m)^n$. Рассмотрим

$$(a+b+c)^4=(a+b+c)(a+b+c)(a+b+c)(a+b+c).$$

Для нахождения коэффициента при ab^2c необходимо найти все способы выбора a, двух b и c, где буквы выбираются из каждого из сомножителей. Возможные

варианты: $abbc$, $cbba$, $abcb$, $dbab$, $babc$, $bcba$, $bbac$, $bbca$, $cabb$, $acbb$, $bacb$, $bcab$. Но это только размещения a, двух b и одного c. Таким образом, количество размещений, которое является коэффициентом при ab^2c, равно

$$P(4;1,2,1) = \frac{4!}{1!2!1!} = 12.$$

Аналогично, коэффициент при b^2c^2 равен

$$P(4;0,2,2) = \frac{4!}{0!2!2!} = 6.$$

ТЕОРЕМА 8.61. Для заданного положительного числа n

$$(x_1 + x_2 + x_3 + \cdots + x_m)^n = \sum \frac{n!}{n_1!n_2!n_3! \cdots n_m!} x_1^{n_1} x_2^{n_2} x_3^{n_3} \cdots x_m^{n_m},$$

где сумма взята по всем неотрицательным целым числам $n_1, n_2, \cdots, n_m$, таким, что $n_1 + n_2 + n_3 + \cdots + n_m = n$.

ДОКАЗАТЕЛЬСТВО. Поскольку $(x_1 + x_2 + \cdots + x_m)^n = (x_1 + x_2 + \cdots + x_m)(x_1 + x_2 + \cdots + x_m) \cdots (x_1 + x_2 + \cdots + x_m)$, то каждое слагаемое в разложении получено путем выбора x_i из каждого сомножителя в произведении и их перемножения. Для нахождения коэффициента при $x_1^{n_1} x_2^{n_2} x_3^{n_3} \cdots x_m^{n_m}$ необходимо найти все возможные способы выбора x_1 в количестве n_1, x_2 в количестве n_2, ... и x_m в количестве n_m. Но это есть не что иное, как число размещений x_1 в количестве n_1, x_2 в количестве n_2, ... и x_m в количестве n_m, которое равно

$$P(n;n_1,n_2,n_3,\cdots,n_m) = \frac{n!}{n_1!n_2!n_3! \cdots n_m!}.$$ ■

ПРИМЕР 8.62. Найдем коэффициент при ab^2c^3 в разложении $(a+b+c)^6$. Этот коэффициент равен

$$P(6;1,2,3) = \frac{6!}{1!2!3!} = 60.$$ □

ПРИМЕР 8.63. Найдем коэффициент при $a^3b^2c^3$ в разложении $(2a+3b+c)^8$. Слагаемое имеет вид

$$P(8;3,2,3)(2a)^3(3b)^2c^3 = 560 \cdot 2^3 \cdot 3^2 a^3 b^2 c^3 = 4320 a^3 b^2 c^3,$$

поэтому коэффициент равен 4320. □

Предположим, что имеются семь различимых шаров, и требуется положить три шара в первую коробку, два шара — во вторую и два шара — в третью коробку. Сколькими способами можно это сделать? Существует $C(7,3)$ способов выбрать три шара для первой коробки. После того, как шары выбраны, остается четыре шара, два из них выбираются для второй коробки. Для этого существуют $C(4,2)$ способов. Наконец, остается два шара, которые мы кладем в последнюю коробку.

Это может быть сделано $C(2,2)$ или одним способом. Таким образом, количество способов размещений шаров равно $C(7,3) \cdot C(4,2) \cdot C(2,2)$. Но по теореме 8.58 это равно

$$P(7;3,2,2) = \frac{7}{3!2!2!}.$$

Как видим, существует связь между размещениями неразличимых объектов и разбиением множества, что, по сути, мы и делали с шарами и коробками. Для того, чтобы выявить эту связь, рассмотрим следующий пример. Предположим, что имеются два красных, три зеленых и четыре синих шара, которые отличаются только цветом. Требуется найти все возможные различные размещения данных шаров. Пусть $s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9$ — позиции, в которые шары могут быть помещены. Допустим, красные шары располагаются в позициях s_2 и s_7. Отождествим такое размещение с помещением s_2 и s_7 в коробку R. Предположим, что зеленые шары находятся в позициях s_1, s_3 и s_9. Отождествим это с помещением s_1, s_3 и s_9 в коробку G. Предположим, наконец, что синие шары располагаются в позициях s_4, s_5, s_6 и s_8. Отождествим это с помещением s_4, s_5, s_6 и s_8 в коробку B. Обратно, если s_1 и s_9 помещены в коробку R, s_2 s_7 и s_8 — в коробку G, а s_3, s_4, s_5 и s_6 — в коробку B, мы отождествим это с помещением красных шаров в позиции s_1 и s_9, зеленых шаров — в позиции s_2, s_7 и s_8, а синих шаров — в позиции s_3, s_4, s_5 и s_6. Таким образом, задача о размещении неразличимых шаров свелась к задаче о помещении шаров в коробки. В общем случае справедлива следующая теорема.

ТЕОРЕМА 8.64. Пусть $C(n; n_1, n_2, n_3, \ldots, n_m)$ — количество разбиений множества S, содержащего n элементов на m множеств S_1, S_2, S_3, ... и S_m, содержащих n_1, n_2, n_3, ... и n_m элементов соответственно. Тогда

$$C(n; n_1, n_2, n_3, \ldots, n_m) = P(n; n_1, n_2, n_3, \ldots, n_m) = \frac{n!}{n_1!n_2!n_3! \cdots n_m!}.$$

ДОКАЗАТЕЛЬСТВО. Существуют $C(n, n_1)$ способов выбора элементов для S_1. Как только эти элементы выбраны, остается $n - n_1$ элементов, из которых необходимо выбрать n_2 элементов для S_2. Последнее можно осуществить $C(n-n_1, n_2)$ способами. Если эти элементы выбраны, остаются $n - n_1 - n_2$ элементов, из которых необходимо выбрать n_3 элементов для S_3. Предположим, что множества S_1, S_2, S_3, ... и S_{i-1} уже выбраны, тогда для выбора n_i элементов для множества S_i остается $n - n_1 - n_2 - \cdots - n_{i-1}$ элементов. Этот выбор можно осуществить $C(n-n_1-n_2-\cdots-n_{i-1}, n_i)$ способами. Используя принцип умножения, получаем

$$\begin{aligned} C(n; n_1, \ldots, n_m) &= C(n, n_1)C(n - n_1, n_2)C(n - n_1 - n_2, n_3) \cdots \\ &\quad \cdots C(n - n_1 - \cdots - n_{k-1}, n_k) = \\ &= \frac{n!}{n_1!n_2!n_3! \cdots n_m!} = \\ &= P(n; n_1, n_2, n_3, \ldots, n_m). \end{aligned}$$

Заметим, что сочетание $C(n,k) = C(n; k, n-k)$ является частным случаем разбиения n-элементного множества на множество, содержащее k элементов (которые

мы выбираем), и множество, содержащее $n-k$ элементов (которые остаются). Важно отметить, что множество разбивается в последовательность подмножеств с заданным количеством элементов. При этом мы подсчитывем количество разбиений множества в последовательность подмножеств заданных размеров. Один из способов убедиться в этом — заметить, что $C(n,k) = C(n;k,n-k)$ — количество способов выбрать k объектов из n объектов. Мы не включаем сюда количество способов выбрать $n-k$ объектов и, следовательно, оставляем k объектов, что также дает разбиение множества S, содержащего n объектов, на подмножества, содержащие k и $n-k$ объектов. Для этого также существуют $C(n;k,n-k)$ способов, поэтому существуют $2C(n;k,n-k)$ способов разбиения множества S на множества, содержащие k объектов и $n-k$ объектов. Другой способ представить себе содержание теоремы — это рассматривать ее как описание раскладывания различимых шаров в разные коробки. ■

ПРИМЕР 8.65. Сколькими способами можно выбрать четыре набора по пять карт из колоды, содержащей 52 карты? По сути, колода разбивается на пять множеств: четыре набора по пять карт и 32 оставшихся карты. Поэтому количество таких наборов равно

$$C(52;5,5,5,5,32) = \frac{52!}{5!5!5!5!32!}.$$ □

■ УПРАЖНЕНИЯ

1. Сколько различных размещений можно образовать, используя буквы в слове *seceeded*?
2. Сколько различных размещений можно образовать, используя буквы в слове *unsuccessful*?
3. Сколько различных размещений можно образовать, используя буквы в слове *Sussex*?
4. Сколько различных размещений можно образовать, используя буквы в слове *Tallahassee*?
5. Предположим, что нужно расставить на полке 26 книг, среди которых 8 одинаковых учебников по математике, 6 одинаковых учебников по информатике, 9 одинаковых учебников по физике и 3 — по химии. Сколькими способами это можно сделать?
6. В урне находятся 6 синих, 5 красных и 14 желтых шаров. Если шары одного цвета неразличимы, то сколько существует способов одновременно вытащить из урны шары одного цвета?
7. Тест содержит 35 вопросов, каждый вопрос имеет пять вариантов ответа: *а*, *б*, *в*, *г* и *д*. Сколько может быть различных ответных листов, в которых выбрано равное количество ответов *а*, *б*, *в*, *г* и *д*.
8. У профессора Кренка в группе 20 студентов. Согласно критерию, известному лишь ему одному, он решил поставить две оценки A, три оценки B, десять C, три D и две оценки F. Сколькими способами он может поставить оценки студентам?
9. В бейсбольной команде 24 игрока. В гостинице они остановились в шести четырехместных номерах. Сколькими способами можно расселить игроков?

10. Чему равен коэффициент при $x^2y^3z^4$ в разложении $(x+y+z)^9$?
11. Чему равен коэффициент при $x^3y^6z^3$ в разложении $(x+2y+3z)^{12}$?
12. Чему равен коэффициент при $w^3x^2y^5z^7$ в разложении $(w+x+y+z)^{17}$?
13. Чему равен коэффициент при $w^4x^5y^6z^3$ в разложении $(2w+4x+y+5z)^{18}$?
14. Чему равен коэффициент при $x^3y^6z^{12}$ в разложении $(x+2y^2+4z^3)^{10}$?
15. Чему равен коэффициент при $w^{10}x^{12}y^4z^3$ в разложении $(4w^5+2x^3+y^2+5z)^{11}$?

8.7. ПЕРЕСТАНОВКИ И СОЧЕТАНИЯ С ПОВТОРЕНИЕМ

Идея перестановки с повторением может показаться бессмысленной, поскольку перестановки ассоциируются с перегруппировками, которые, похоже, не допускают повторений. Если допустить повторение, то необходимо пересмотреть наше представление о перестановке. Главным свойством перестановки является порядок. Строка *cab* отличается от строки *abc*. Если опять рассмотреть задачу о номерных знаках автомобилей с тремя буквами, за которыми следуют три цифры, и не допускать повторений, то получим $26\times25\times24\times10\times9\times8$ различных номерных знаков. Существует $P(25,3)$ способов выбора букв и $P(10,3)$ способов выбора цифр. Таким образом, существуют $P(25,3)\times P(10,3)$ различных номерных знаков. В этом и состоит обычное представление о перестановках. Заметим, что допуская повторение, мы не отказываемся от понятия порядка. Номерной знак ФПФ 199 отличается от номерного знака ПФФ 919. Говоря о перестановках с повторением, имеется в виду, что этот порядок сохраняется. Вспомним, что если в номерном знаке разрешить повторения букв и цифр, то существуют 26 возможных способов выбора для каждой буквы и 10 возможных способов выбора для каждой цифры. Таким образом, если разрешить повторение, то будут существовать $26^3\times10^3$ возможных номерных знаков. В общем случае, если имеется k упорядоченных мест, для каждого из которых можно выбрать любой из n объектов, существуют n^k способов выбора объектов. Таким образом, число перестановок с повторением, когда k объектов выбираются из n объектов, равно n^k. Наверное, идея повторения, использованная здесь, требует разъяснений. Один из способов трактовки повторения — это думать, что повторение предполагает возвращение объекта и повторное его использование. Другой способ трактовки данного типа повторения состоит в том, что имеется столь достаточное количество объектов каждого типа, которое нельзя исчерпать. Например, в случае с номерными знаками требуется теперь только три копии каждой буквы и каждой цифры. Приведенная ниже теорема сформулирована для такого представления о повторении.

ТЕОРЕМА 8.66. Пусть S — множество, содержащее n неразличимых объектов. Тогда количество различных перестановок, образованных выбором k элементов с повторением, равно n^k.

ПРИМЕР 8.67. Лототрон содержит 500 шаров с номерами. Из него выбирают шар, номер которого записывают. Шар возвращают в лототрон и процедура повторяется. Так продолжается до тех пор, пока не наберется комбинация из пяти номеров.

Подсчитаем количество возможных комбинаций чисел. Для каждого из пяти чисел имеется 500 способов выбора. Следовательно, число различных комбинаций составляет 500^5. □

ПРИМЕР 8.68. Сколько существует индивидуальных номеров карточек социального страхования?

Поскольку номер следует выбирать из неотрицательных целых чисел, меньших десяти, то выбираются девять цифр, причем каждая выбирается из десяти цифр с повторением, поэтому имеются 10^9 различных номеров карточек социального страхования. □

Предположим, что комитет состоит из восьми человек. При принятии решения они голосуют "за", "против" или воздерживаются от голосования. Сколько возможных исходов голосования по данному решению? Если интересует вопрос, кто и как голосовал, тогда речь идет о числе перестановок, когда для каждого голосующего имеются три варианта ответа, что дает 3^8 возможных исходов голосования. Допустим, что нас интересует только общий результат голосования. Таким образом, четыре голоса "за", три голоса "против" и один воздержавшийся — это один из возможных исходов. Таким же возможным исходом является вариант: два голоса "за", четыре голоса "против" и два воздержавшихся. Для удобства перечислим сначала голоса "за", затем "против" и, наконец, голоса воздержавшихся. Таким образом, голосование можно изобразить, например, в виде *ЗЗППВВВВ*, где два голоса "за", два "против" и четыре воздержавшихся. Далее можно строить разбиение голосов, например, *ЗЗ|ПП|ВВВВ*. Поскольку порядок расположения голосов понятен, можно перейти к записи $xx|xx|xxxx$, изображающей *ЗЗ|ПП|ВВВВ*. Таким образом, запись $xx||xxxxxx$ будет представлять два голоса "за" и четыре воздержавшихся, а запись $xxxxxxxx||$ будет соответствовать голосованию "за" всех восьми членов комитета. Таким образом, установлено взаимно однозначное соответствие между возможными исходами голосования и различными способами размещения восьми знаков x и двух знаков |. Но это ни что иное, как количество способов выбора двух мест из десяти для размещения знака | или, что эквивалентно, количество способов выбора восьми мест из десяти для размещения знака x. Следовательно, существуют

$$C(10,2) = C(10,8) = \frac{10!}{8!2!}$$

способов разместить x и |. А значит, существуют

$$C(10,2) = C(10,8) = \frac{10!}{8!2!}$$

возможных исходов голосования.

Предположим, что n объектов выбираются из k типов объектов с неограниченным повторением. Пусть a_i — объект типа i, тогда $n_1a_1+n_2a_2+n_3a_3+\cdots+n_ka_k$, где $n_i \geq 0$ для всех i, а $n_1+n_2+n_3+\cdots+n_k = n$ представляет выбор n_i объектов типа i. Как и прежде, это можно записать в виде

$$a_1a_1a_1\cdots a_1|a_2a_2a_2\cdots a_2|\cdots|a_ka_ka_k\cdots a_k,$$

где каждое a_i повторено n_i раз. Поскольку место расположения каждого типа понятно, то выборку можно записать в виде

$$xxx\cdots x|xxx\cdots x|\cdots|xxx\cdots x.$$

Заметим, что разделителей | на один меньше количества типов. Таким образом, имеем n объектов плюс $k-1$ разделителей, образующих $n+k-1$ мест для размещения x или |. Каждое расположение знаков x и | дает новый способ выбора n объектов из k типов объектов с неограниченным повторением. Поскольку существует $C(n+k-1,n)=C(n+k-1,k-1)$ способов выбора места для знака x или, что эквивалентно, для знака |, то существуют

$$C(n+k-1,n)=C(n+k-1,k-1)$$

различных способов выбора n из k типов объектов с неограниченным повторением. Такие выборки будем называть ***сочетаниями из k объектов по n с повторением***. Получаем следующую теорему.

ТЕОРЕМА 8.69. Количество различных сочетаний из k объектов по n равно

$$C(n+k-1,n)=C(n+k-1,k-1)=\frac{(n+k-1)!}{n!(k-1)!}.$$

ПРИМЕР 8.70. Если в булочной продается 10 различных видов пончиков, то сколькими способами можно выбрать дюжину пончиков? Поскольку 12 пончиков выбираются из 10 различных типов с повторением, то имеются

$$\frac{(10+12-1)!}{12!(10-1)!!}=\frac{21!}{12!9!}$$

различных способов выбрать дюжину пончиков. □

ПРИМЕР 8.71. Сколько решений имеет уравнение

$$n_1+n_2+n_3+n_4+n_5=25,$$

где каждое n_i — неотрицательное целое число? Это эквивалентно вопросу о том, сколько существует различных выборок вида

$$n_1a_1+n_2a_2+n_3a_3+n_4a_4+n_5a_5,$$

где имеется n_i объектов типа a_i и

$$n_1+n_2+n_3+n_4+n_5=25,$$

но количество таких выборок — это количество различных сочетаний из 5 элементов по 25 с повторениями. Итак, существуют

$$\frac{(25+5-1)!}{25!(5-1)!!}=\frac{29!}{25!4!}$$

различных решений уравнения $n_1+n_2+n_3+n_4+n_5=25$. □

Было показано, что число различных сочетаний из k объектов по n объектов с повторением равно

$$C(n+k-1,n) = C(n+k-1,k-1) = \frac{(n+k-1)!}{n!(k-1)!}.$$

Если необходимо выбрать хотя бы по одному объекту каждого типа, то имеются $n-k$ выборов из k различных объектов. Следовательно, имеем $C(n-k+k-1,k-1)$ или $C(n-1,k-1)$ способов выбора.

ТЕОРЕМА 8.72. Количество различных сочетаний из k объектов по n объектов с повторением, когда необходимо выбрать хотя бы по одному объекту каждого типа, равно

$$C(n-1,n-k) = C(n-1,k-1) = \frac{(n-1)!}{n!(k-1)!}.$$

ПРИМЕР 8.73. Если в булочной продается 10 различных видов пончиков, то сколькими способами можно выбрать две дюжины пончиков, если необходимо выбрать хотя бы по одному пончику каждого вида? Если бы не было последнего ограничения, то для выбора двух дюжин пончиков из 10 различных видов существовало бы

$$C(24+10-1,24) = C(33,24)$$

различных способов. Однако, учитывая ограничение, можно выбрать только $24-10=14$ пончиков из 10 различных видов, что дает

$$C(14+10-1,14) = C(33,14)$$

различных вариантов выбора. □

ПРИМЕР 8.74. Найдем количество различных целочисленных решений уравнения

$$n_1+n_2+n_3+n_4+n_5 = 25,$$

если $n_1 \geq 1$, $n_2 \geq 1$, $n_3 \geq 2$, $n_4 \geq 2$ и $n_5 \geq 2$. Представим эту задачу, как подсчет количества выборок вида

$$n_1a_1+n_2a_2+n_3a_3+n_4a_4+n_5a_5,$$

где

$$n_1+n_2+n_3+n_4+n_5 = 25,$$

причем следует выбрать, по крайней мере, одно a_1, одно a_2, два a_3, два a_4 и два a_5. Это оставляет для выбора только $25-8=17$ элементов из пяти типов, что дает

$$\frac{(17+5-1)!}{17!} = \frac{21!}{17!}$$

решений для

$$n_1+n_2+n_3+n_4+n_5 = 25.$$

□

■ УПРАЖНЕНИЯ

1. Сколько существует решений уравнения $n_1 + n_2 + n_3 + n_4 + n_5 = 17$ таких, что каждое n_i — неотрицательное целое число?
2. Сколько существует решений уравнения $n_1 + n_2 + n_3 + n_4 = 23$ таких, что каждое n_i — неотрицательное целое число?
3. Сколько существует решений уравнения $n_1 + n_2 + n_3 + n_4 = 23$ таких, что $n_1 \geq 2$, $n_2 \geq 3$, $n_3 \geq 4$, $n_4 \geq 5$?
4. Сколько существует решений уравнения $n_1 + n_2 + n_3 + n_4 = 21$ таких, что $n_1 \geq 2$, $n_2 \geq 3$, $n_3 \geq 4$, $n_4 \geq 5$?
5. Сколько различных коллекций из десяти монет можно собрать из монет стоимостью 1 цент, 5 центов, 10 центов, 25 центов и 50 центов?
6. В киоске продается мороженое 21 вида. Фрэд хочет купить пять порций мороженого. Если мороженого одного вида может быть более одной порции, то сколько существует способов сделать покупку?
7. Если в урне имеются 20 красных, 20 зеленых и 20 синих шаров, то сколькими различными способами можно выбрать 10 шаров?
8. Цветочник продает 10 видов цветов. Сколько различных букетов из 12 цветков он может сделать?
9. Сколько существует чисел, меньших 10000 и таких, что сумма цифр равна 12?
10. Сколько существует бриджевых раздач (из 13 карт), содержащих три трефовые, четыре червовые, пять пиковых карт и одну карту бубей?
11. Если не делать различий между картами одной масти, то сколько может быть типов бриджевых раздач?
12. Человек покупает 12 различных игрушек для своих четырех детей. Сколькими способами он может распределить игрушки? А сколькими способами, если игрушки делить поровну?

8.8. ПРИНЦИП КЛЕТОК

Принцип клеток, известный также как принцип Дирихле, — это очень простая, но тем не менее очень мощная идея. Она особенно полезна в теории чисел. Данный раздел состоит из двух довольно простых теорем и нескольких не совсем простых примеров. Начнем с наиболее известной формы принципа клеток.

ТЕОРЕМА 8.75. Если поместить $n + 1$ голубей в n клеток, то, по крайней мере, в одной клетке будет находится более одного голубя.

ДОКАЗАТЕЛЬСТВО. Если в каждой клетке не более одного голубя, то голубей не может быть больше, чем клеток. Тогда во всех клетках должно быть не более, чем n голубей, что противоречит условию. ■

Из теоремы следует, что если $m > n$ и требуется m голубей разместить в n клетках, то в одной клетке находится более одного голубя.

ПРИМЕР 8.76. В списке, содержащем $m + 1$ положительных целых чисел, по крайней мере, два числа имеют один и тот же остаток при делении на m. Пусть

$a_1, a_2, a_3, \ldots, a_m, a_{m+1}$ — перечень положительных целых чисел. Предположим, что для $1 \leq i \leq m+1$ имеем $a_i = q_i m + r_i$, где $0 \leq r_i < m$. Следовательно, $r_1, r_2, r_3, \ldots, r_m, r_{m+1}$ — список, содержащий $m+1$ положительных целых чисел, меньших m. Но существует только m различных неотрицательных чисел, меньших m. Поэтому два числа из r_i должны быть равны. □

ПРИМЕР 8.77. Если $|A| > |B|$, то не может существовать инъективная функция $f : A \to B$. Пусть $|B| = n$. Существует n множеств $f^{-1}(b) = \{x : f(x) = b\}$ для $b \in B$ и $A = \bigcup_{b \in B} f^{-1}(b)$. Но A содержит больше n элементов. Поэтому не менее двух элементов, например, a и a', должны быть в одном множестве. Но тогда $f(a) = f(a')$ и f не является инъективной функцией. □

ПРИМЕР 8.78. Если в прямоугольнике со сторонами 6 и 8 дюймов помещены пять точек, то существуют две точки, расстояние между которыми не более 5 дюймов. Разделим исходный прямоугольник на четыре прямоугольника, размером 3 на 4 дюйма каждый. Поскольку пять точек должны находиться либо внутри, либо на границах четырех прямоугольников, то хотя бы две точки должны быть либо внутри, либо на границе одного и того же прямоугольника размера 3×4. Но любые такие точки находятся на расстоянии не более 5 дюймов. □

ПРИМЕР 8.79. В последовательности $a_1, a_2, a_3, \ldots, a_{10}$ из десяти целых чисел существует последовательность идущих подряд элементов $a_m, a_{m+1}, a_{m+2}, \ldots, a_{m+n}$, сумма которых делится на 10. Рассмотрим числа $s_1 = a_1$, $s_2 = a_1 + a_2$, $s_3 = a_1 + a_2 + a_3$, $s_4 = a_1 + a_2 + a_3 + a_4$, ..., $s_{10} = a_1 + a_2 + a_3 + a_4 + \cdots + a_{10}$. Если одна из этих сумм делится на 10, то процесс завершен, элементы суммы образуют искомую последовательность. Если нет, то каждое $s_i = 10q_i + r_i$, где $1 \leq r_i \leq 9$. Имеется десять r_i, и только девять значений, которые они могут принимать. Поэтому существуют два r с одинаковыми значениями, например, $r_j = r_k$ для некоторого $j < k$. Тогда

$$\begin{aligned} s_k - s_j &= 10q_k + r_k - (10q_j + r_j) = \\ &= 10q_k - 10q_j = \\ &= 10(q_k - q_j) \end{aligned}$$

кратно 10. Но это не что иное, как сумма $a_{j+1} + a_{j+2} + a_{j+3} + \cdots + a_k$, элементы которой образуют искомую последовательность. □

ПРИМЕР 8.80. (Эрдош и Шекерес) Пусть S — подмножество множества $\{1, 2, 3, \ldots, 2n\}$, содержащее $n+1$ элемент, где n — некоторое положительное целое число. Существуют два элемента, принадлежащие множеству S такие, что один элемент делит другой элемент. Любое положительное целое число можно представить в виде $2^i q_i$, где q_i — положительное нечетное число. Запишем каждый элемент множества S в таком виде. Множество $\{q_1, q_2, \ldots, q_n, q_{n+1}, \}$ содержит $n+1$ положительных нечетных целых чисел, при том что имеется только n нечетных чисел, которые меньше или равны $2n$. Поэтому $q_i = q_j$ при $i \neq j$. Таким образом, множество S содержит два различных целых числа вида $a = 2^r q_i$ и $b = 2^s q_i$ соответственно. Если $r < s$, то a делит b, и если $r > s$, то b делит a. □

А теперь приведем более сильные формы принципа клеток.

ТЕОРЕМА 8.81.

а) Пусть $m > n$ голубей помещены в n клеток, тогда некоторая клетка содержит не менее $\left\lceil \frac{m}{n} \right\rceil$ голубей.

б) Пусть $m = m_1 + m_2 + m_3 + \cdots + m_n - n + 1$ голубей помещены в n клеток, где m_i — положительное целое число для каждого $1 \leq i \leq n$. Тогда, для некоторого i, i-ая клетка содержит не менее m_i голубей.

ДОКАЗАТЕЛЬСТВО. а) Пусть каждая клетка содержит меньше $\left\lceil \frac{m}{n} \right\rceil$ голубей. Тогда, поскольку нас интересуют целые голуби, каждая клетка содержит меньше $\frac{m}{n}$ голубей. Поэтому общее количество голубей меньше $n \times \frac{m}{n} = m$, что противоречит условию.

б) Пусть i-ая клетка содержит меньше, чем m_i голубей. Тогда количество голубей в i-ой клетке меньше или равно $m_i - 1$. Просуммировав по всем клеткам, получаем, что всего голубей не более, чем

$$(m_1 - 1) + (m_2 - 1) + (m_3 - 1) + \cdots + (m_n - 1) = m_1 + m_2 + m_3 + \cdots + m_n - n,$$

что противоречит условию. ■

ПРИМЕР 8.82. Пусть в урне находятся 10 красных, 10 синих и 10 зеленых шаров. Сколько шаров нужно взять из урны, чтобы с уверенностью иметь хотя бы 4 красных или 5 синих, или 6 зеленых шаров? Мы знаем, что 12 шаров недостаточно, поскольку можно выбрать 3 красных, 4 синих и 5 зеленых шаров. Если отождествить шары одного цвета с голубями, находящимися в одной клетке, то, согласно части (б) сильной формы принципа клеток, имеем $n = 3$, $m_1 = 4$, $m_2 = 5$ и $m_3 = 5$. Поэтому искомое количество шаров равно

$$\begin{aligned} m &= m_1 + m_2 + m_3 + \cdots + m_n - n + 1 = \\ &= 4 + 5 + 6 - 3 + 1 = \\ &= 13. \end{aligned}$$

□

ПРИМЕР 8.83. (Эрдош и Шекерес) Конечная последовательность целых чисел $a_1, a_2, a_3, \ldots, a_k$ возрастающая, если каждое число последовательности больше предыдущего. Иначе говоря, если $i > j$, то $a_i > a_j$. Конечная последовательность целых чисел $a_1, a_2, a_3, \ldots, a_k$ убывающая, если каждое число последовательности меньше предыдущего. Иначе говоря, если $i > j$, то $a_i < a_j$. Любая последовательность из $n^2 + 1$ различных целых чисел содержит либо убывающую подпоследовательность, включающую не менее $n + 1$ членов, либо возрастающую подпоследовательность, включающую не менее $n + 1$ членов. Для доказательства предположим, что s_i — количество чисел в самой длинной возрастающей последовательности $a_1, a_2, a_3, \ldots, a_{n^2+1}$, начинающейся с a_i. Если $s_i \geq n + 1$ для некоторого i, то процесс завершен. Если нет, то для каждого i, $1 \leq s_i \leq n$. Поэтому для всех $n^2 + 1$ штук s_i, где $1 \leq i \leq n^2 + 1$, имеется n значений. Согласно части (а) сильной формы принципа клеток получаем, что

$$\left\lceil \frac{n^2 + 1}{n} \right\rceil > n$$

штук s_i равны. Поэтому, по крайней мере, $n+1$ чисел из s_i равны. Пусть все они равны s. Рассмотрим подпоследовательность длины $n+1$, состоящую из таких a_i, что s_i равны s. Эта последовательность — убывающая, потому что если $a_i > a_j$ для $i > j$, то, добавляя a_i в начало последовательности длины s, начинающейся с a_j, можно получить возрастающую последовательность длины $s+1$, начинающуюся с a_i. Но это противоречит тому факту, что самая длинная возрастающая последовательность, начинающаяся с a_i, имеет длину s. Таким образом построена убывающая последовательность длины $n+1$. □

ПРИМЕР 8.84. Предположим, что в комнате находятся шесть человек, и каждые два из них либо друзья, либо враги. Тогда есть три человека, которые дружат между собой, или есть три человека, которые враждуют между собой. Выберем одного человека, назовем его A и поставим в середине комнаты. Разместим всех его врагов вдоль южной стены, а всех друзей — вдоль северной стены. Согласно части (а) сильной формы принципа клеток либо не менее трех человек стоят вдоль южной стены, либо не менее трех человек стоят у северной стены. Если у южной стены стоит не менее трех человек, то или трое из них — взаимные друзья, что и требовалось доказать, или двое из них — враги. Тогда эти люди вместе с A — трое взаимных врагов. Аналогично, если не менее трех человек стоят у северной стены, то или трое из них — взаимные враги, что и требовалось доказать, или двое из них — друзья. Тогда они вместе с A — трое взаимных друзей. □

Рассмотренный выше пример является характерным, хотя и довольно простым примером задач из раздела комбинаторики, который носит название ***теория Рамсея***. Можно считать, что нам повезло, поскольку в теории Рамсея очень мало простых примеров. Почти все ее проблемы по сей день не решены. Поясним теорию Рамсея на примере полного графа с раскрашенными ребрами. Предположим, что в предыдущем примере вершины графа — аналоги людей. Образуем граф следующим образом: если двое людей — друзья, то соответствующие им вершины соединяет красное ребро. Если двое людей — враги, то соответствующие вершины соединяет синее ребро. Поскольку каждые двое — либо друзья, либо враги, то ребро существует между двумя любыми вершинами графа. Следовательно, этот граф — граф K_6, полный граф с шестью вершинами.

В рассмотренном выше примере, утверждалось, что если все ребра графа K_6 раскрашены либо красным, либо синим цветом, то в графе K_6 имеется либо красный треугольник, либо синий треугольник, каждый из которых является подграфом графа K_6, по сути, просто графом K_3. Граф K_5 не обладает таким свойством, что если его ребра раскрашены либо красным, либо синим цветом, то он содержит красный или синий подграф K_3. Это доказывает граф, изображенный на рис. 8.13, где пунктиром изображены красные ребра, а синей линией — синие ребра.

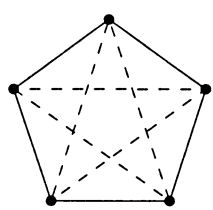


Рис. 8.13

Теперь дадим определение свойства Рамсея для двух цветов на примере графов. Существует обобщение свойства Рамсея для произвольного числа цветов.

ОПРЕДЕЛЕНИЕ 8.85. Полный граф K_n обладает (p, q) ***свойством Рамсея***, если в случае, когда он раскрашен двумя цветами, например, красным и синим, он содержит либо подграф K_p с красными ребрами, либо подграф K_q с синими ребрами.

ОПРЕДЕЛЕНИЕ 8.86. ***Число Рамсея,*** $R(p, q)$ — это наименьшее число n такое, что граф K_n имеет (p, q) свойство Рамсея.

Вполне очевидно, что $R(p, q) = R(q, p)$, поскольку можно заменить красный цвет на синий, а синий — на красный. Также, если $m > R(p, q) = n$, то, поскольку K_n — подграф графа K_m, граф K_m также обладает свойством Рамсея.

Относительно чисел Рамсея будут доказаны только две теоремы. Эти теоремы интересны сами по себе, но более важным является то, что они доказывают существование чисел Рамсея.

ТЕОРЕМА 8.87. Для всех $p \geq 2$ $R(p, 2) = R(2, p) = p$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим граф K_p. Если одно ребро раскрашено красным (или синим) цветом, то существует красный (или синий) граф K_2. Если нет, тогда все ребра раскрашены синим (красным) цветом, и мы получаем синий (красный) граф K_p. ■

ТЕОРЕМА 8.88. (Эрдош и Шекерес) Для всех $p, q \geq 3$

$$R(p, q) \leq R(p-1, q) + R(p, q-1).$$

ДОКАЗАТЕЛЬСТВО. Предположим, что K — полный граф, имеющий $R(p-1, q)+R(p, q-1)$ вершин, ребра которого раскрашены в красный или синий цвет. Пусть v — выбранная вершина графа K. Пусть N — полный подграф графа K, вершины которого V_N связаны с вершиной v красным ребром. Пусть S — полный подграф графа K, вершины которого V_S связаны с вершиной v синим ребром. Поскольку $V_N \cup V_S$ содержит

$$R(p-1, q) + R(p, q-1) - 1 = R(p-1, q) + R(p, q-1) - 2 + 1$$

вершин, разделенных на $n = 2$ множества, то либо V_N имеет $R(p-1, q)$ вершин, либо V_S имеет $R(p, q-1)$ вершин. Поэтому, либо N — полный граф, имеющий $R(p-1, q)$ вершин, либо S — полный граф, имеющий $R(p, q-1)$ вершин. Если N — полный граф, имеющий $R(p-1, q)$ вершин, то он содержит либо подграф K_{p-1} со всеми красными ребрами, либо подграф K_q со всеми синими ребрами. Если он содержит граф K_q со всеми синими ребрами, то доказательство завершено. Если он содержит граф K_{p-1} со всеми красными ребрами, то, добавив вершину v и все красные ребра из v в вершины подграфа K_{p-1}, получим полный граф K_p со всеми красными ребрами, что и требовалось доказать.

Аналогично, если S — полный граф, имеющий $R(p, q-1)$ вершин, то он содержит либо подграф K_p со всеми красными ребрами, либо подграф K_{q-1} со всеми синими ребрами. Если он содержит граф K_p со всеми красными ребрами, то доказательство завершено. Если он содержит граф K_{q-1} со всеми синими вершинами, то если добавить вершину v и все синие ребра из вершины v во все вершины подграфа K_{q-1}, получим полный граф K_q с синими ребрами, что и требовалось доказать. ■

Основываясь на результатах двух приведенных выше теорем и используя индукцию или принцип полного упорядочения для положительных целых чисел, а также технику доказательства второй теоремы, можно доказать еще один результат, сформулированный в виде следующей теоремы

ТЕОРЕМА 8.89. (Рамсей) . Если целые числа p и q больше единицы, то число Рамсея, $R(p, q)$, существует.

■ УПРАЖНЕНИЯ

1. Сколько людей нужно выбрать из группы, включающей 40 человек — женщин и их мужей, чтобы наверняка выбрать супружескую пару?
2. Сколько чисел нужно выбрать из первых 210 положительных целых чисел, чтобы наверняка выбрать четное число?
3. Покажите, что в компании из 30 человек существуют двое, имеющие одинаковое количество взаимных друзей.
4. Если группа людей приехала из пяти различных стран, то насколько велика должна быть эта группа, чтобы она гарантированно включала трех человек, приехавших из одной страны?
5. Если на кухне имеется 7 банок супа с лапшой, 10 банок овощного супа, 14 банок супа с моллюсками, 18 банок лукового супа и 20 банок томатного супа, то сколько банок нужно взять, чтобы наверняка иметь 12 банок супа одного вида?
6. Сколько человек должно быть в группе людей, чтобы, по крайней мере, у двоих дни рождения были в одном месяце?
7. Сколько человек должно быть в группе людей, чтобы, по крайней мере, у троих дни рождения были в одном месяце?
8. Сколько человек должно быть в группе людей, чтобы, по крайней мере, двое родились в один день недели?
9. Докажите, что в произвольном списке из n положительных целых чисел всегда существует последовательность идущих последовательно чисел, сумма которых делится на n.
10. Сколько нужно выбрать слов, чтобы, по крайней мере, два начинались с одной и той же буквы?
11. Сколько нужно выбрать слов, чтобы, по крайней мере, два начинались с одной и той же буквы и оканчивались одной и той же буквой?
12. Анкета разослана 13 первокурсникам, 5 второкурсникам, 15 студентам предпоследнего курса и 20 выпускникам. Сколько анкет нужно собрать, чтобы наверняка получить 9 анкет от студентов одной категории?

13. Сколько чисел нужно выбрать из последовательности $\{1, 2, 3, 4, \ldots, 19, 20\}$, чтобы два из них давали в сумме 21?

14. Покажите, что любое подмножество, содержащее не менее n различных целых чисел, находящихся между числами 2 и $2n$, всегда содержит два взаимно простых числа.

15. Пусть в ресторане 14 столов и 170 стульев, и n — наибольшее количество стульев за столом. Определите наименьшее возможное значение числа n?

16. Докажите, что десятичное разложение рационального числа имеет период.

17. Докажите, что если каждое из $n+1$ различных положительных целых чисел не превышает число $2n$, то два из них отличаются на единицу.

18. Покажите, что если сумма 10 положительных чисел равна 151, то сумма трех из этих чисел должна быть не меньше 46.

19. Покажите, что если m шаров положить в n коробок и

$$m < \frac{n(n-1)}{2},$$

то, по крайней мере, в двух коробках будет находиться одинаковое число шаров.

20. Покажите, что если сумма $n+1$ положительных целых чисел равна $2n$, то для любого положительного целого числа k, меньшего $2n$, существует подмножество этих $n+1$ целых чисел, сумма которых равна k.

8.9. СНОВА О ВЕРОЯТНОСТИ

В предыдущем рассмотрении вероятности предполагалось, что все результаты эксперимента равновозможны. Теперь мы снимем это ограничение и дадим более общее понятие вероятности, для чего наш подход должен быть более аксиоматическим; при этом некоторые теоремы станут аксиомами и определениями.

ОПРЕДЕЛЕНИЕ 8.90. Вероятностной функцией на выборочном пространстве (множестве элементарных событий) S называется такая функция P из подмножеств множества S на действительную числовую ось, что

1. $P(A) \geq 0$ для всех $A \subseteq S$.
2. $P(S) = 1$.
3. Если $A, B \subseteq S$ не имеют общих исходов, то $P(A \cup B) = P(A) + P(B)$.

Доказательства приведенных ниже теорем аналогичны тем, которые даны в разделе 8.2. Единственное отличие состоит в том, что три сформулированных выше условия для вероятностной функции были доказаны в теореме 8.53 в предположении, что все исходы эксперимента равновозможны. Необходимо соблюдать осторожность при использовании такой комбинаторной техники, как подсчет числа сочетаний и перестановок, чтобы не допустить по умолчанию равновозможность исходов эксперимента.

ТЕОРЕМА 8.91. Пусть S — множество исходов, множество S также будет универсом. Пусть A' — дополнение множества A, тогда

а) $P(A') = 1 - P(A)$;
б) $P(\varnothing) = 0$;
в) для любого события A имеем $P(A) \geq 0$;
г) если $A \subseteq B$, то $P(A) \leq P(B)$;
д) для любого события A имеем $0 \leq P(A) \leq 1$.

Читателю предлагается доказать следующую теорему, используя принцип индукции.

ТЕОРЕМА 8.92. Пусть $A_1, A_2, A_3, \ldots, A_m$ — попарно непересекающиеся события. Тогда

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_m) = P(A_1) + P(A_2) + P(A_3) + \cdots + P(A_m).$$

Доказательство следующей теоремы аналогично доказательству соответствующей теоремы 8.56.

ТЕОРЕМА 8.93. Для событий $A, B \subseteq S$ имеем $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Мы имели возможность убедиться, что для вероятности существует теорема, аналогичная комбинаторному принципу сложения. Представляется разумным задать вопрос, существует ли для вероятности теорема, аналогичная комбинаторному принципу умножения. Так оно и есть, но сначала необходимо тщательно разобраться в формулировке комбинаторного принципа умножения. Принцип гласит: "Пусть задана последовательность событий $E_1, E_2, E_3, \ldots, E_m$ таких, что событие E_1 осуществляется n_1 способами, и если события $E_1, E_2, E_3, \ldots, E_{k-1}$ осуществились, то событие E_k может осуществиться n_k способами. Тогда существуют $n_1 \times n_2 \times n_3 \times \cdots \times n_k$ способов осуществления всей последовательности событий". Количество способов осуществления события E_k зависит от того, произошли ли события $E_1, E_2, E_3, \ldots, E_{k-1}$. Таким образом, может оказаться, что число способов осуществления события E_k не реализуется само по себе и может зависеть от предыдущих событий.

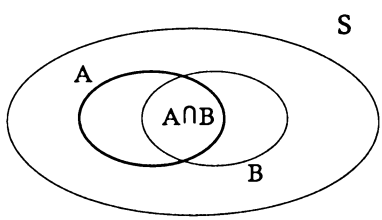


Рис. 8.14

Сходным образом дадим определение условной вероятности. Предположим, что имеются два события A и B. Требуется определить вероятность того, что событие B произойдет, если событие A произошло или наверняка произойдет. Сокращенно будем называть это вероятностью события B при условии A или

относительно события A. Интуитивно понятно, что предположение о том, что событие A произошло или наверняка произойдет, ограничивает выборочное пространство до A, как показано на рис. 8.14, поэтому событие B ограничено теперь исходами из $A \cap B$, и вместо $P(B) = \dfrac{P(B)}{P(S)}$, получаем $\dfrac{P(A \cap B)}{P(A)}$.

ОПРЕДЕЛЕНИЕ 8.94. Вероятность B при условии A, обозначаемая $P(B|A)$, равна $\dfrac{P(A \cap B)}{P(A)}$.

Это можно также записать в виде $P(A \cap B) = P(A)P(B|A)$.

ПРИМЕР 8.95. Предположим, что из колоды вытягивают карту, затем возвращают ее в колоду и вытягивают вторую карту. Какова вероятность того, что обе карты пиковой масти? Какова вероятность того, что хотя бы одна из карт будет пиковой? Пусть A — множество исходов, когда первая карта — пиковая, а B — множество исходов, когда вторая карта — пиковая. Для ответа на первый вопрос необходимо найти $P(A \cap B)$. Заметим, что

$$P(A) = \frac{13}{52} = \frac{1}{4},$$

так как 13 из 52 карт — пиковые. Поскольку первая карта возвращена в колоду, то также имеем

$$P(B|A) = \frac{13}{52} = \frac{1}{4}.$$

Следовательно, вероятность того, что обе карты — пиковые, равна

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) = \\ &= \frac{1}{4} \cdot \frac{1}{4} = \\ &= \frac{1}{16}. \end{aligned}$$

Для ответа на второй вопрос необходимо найти $P(A \cup B)$, но

$$P(A) = P(B) = \frac{13}{52} = \frac{1}{4},$$

поэтому вероятность того, что, по крайней мере, одна из карт — пиковая, равна

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) = \\ &= \frac{1}{4} + \frac{1}{4} - \frac{1}{16} = \frac{7}{16}. \end{aligned}$$

□

ПРИМЕР 8.96. Предположим, что карту вытягивают из колоды, но в колоду *не возвращают*, затем вытягивают вторую карту. Какова вероятность того, что обе карты — пиковые? Какова вероятность того, что хотя бы одна из карт будет пиковой? Пусть A — множество результатов, когда первая карта — пиковая, а B — множество результатов, когда вторая карта — пиковая. Для ответа на первый вопрос необходимо найти $P(A \cap B)$. Заметим, что

$$P(A) = \frac{13}{52} = \frac{1}{4},$$

поскольку 13 из 52 карт — пиковые. Для нахождения $P(B|A)$ заметим, что одну пиковую карту вытянули, но не вернули. Теперь в колоде 51 карта, 12 из которых — пиковые, поэтому

$$P(B|A) = \frac{12}{51}.$$

Следовательно, вероятность того, что обе карты — пиковые, равна

$$P(A \cap B) = P(A)P(B|A) = \frac{1}{4} \cdot \frac{12}{51} = \frac{1}{17}.$$

Чтобы ответить на второй вопрос нужно найти $P(A \cup B)$. Известно, что

$$P(A) = \frac{1}{4}.$$

Теперь задача состоит в том, чтобы найти $P(B)$. Если событие B произошло, то либо две пиковые карты были вытянуты подряд, либо первая карта не была пиковой, а вторая карта была пиковой. Поэтому

$$P(B) = P(A \cap B) + P(A' \cap B).$$

Согласно предыдущей теореме

$$P(A' \cap B) = P(A')P(B|A').$$

Поскольку 39 карт не являются пиковыми, то

$$P(A') = \frac{39}{52} = \frac{3}{4}.$$

После выбора первой карты среди 51 карты все еще остается 13 пиковых, так что

$$P(B|A') = \frac{13}{51}.$$

Следовательно,

$$P(A' \cap B) = \frac{3}{4} \cdot \frac{13}{51} = \frac{13}{68}$$

и

$$P(B) = \frac{1}{17} + \frac{13}{68} = \frac{17}{68} = \frac{1}{4}.$$

Теперь видно, что когда A неизвестно, то $P(B)$ является таким же, как если бы B осуществилось без реализации события A. Можем определить теперь, что

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \\ = \frac{1}{4} + \frac{1}{4} - \frac{1}{17} = \frac{15}{34}.$$

□

ПРИМЕР 8.97. Предположим, что из стандартной колоды, содержащей 52 карты, последовательно вытягиваются 6 карт. Какова вероятность того, что шестая карта — это третья вытянутая карта пиковой масти? Пусть A — событие, при котором две пиковые карты вытянуты за первые пять шагов, а B — событие, при котором пиковая карта вытянута на шестом шаге. Требуется найти $P(A \cap B)$. Опять воспользуемся тем фактом, что

$$P(A \cap B) = P(A)P(B|A).$$

Количество способов осуществления события A равно количеству способов осуществления того, что вытянуты 2 пиковые карты из 13 возможных, а 3 карты вытянуты из 39 карт, которые не являются пиковыми. Таким образом,

$$|A| = \binom{13}{2}\binom{39}{3}.$$

Всего имеется $\binom{52}{5}$ способов выбора 5 карт из 52. Поэтому

$$P(A) = \frac{\binom{13}{2}\binom{39}{3}}{\binom{52}{5}}.$$

Для определения $P(B|A)$ заметим, что поскольку две пиковые карты были выбраны, то существуют 11 способов выбора пиковой карты. Для выбора шестой карты в колоде осталось 47 карт, поэтому

$$P(B|A) = \frac{11}{47}$$

и

$$P(A \cap B) = \frac{\binom{13}{2}\binom{39}{3}}{\binom{52}{5}} \cdot \frac{11}{47} \approx .064.$$

□

Ранее было отмечено, что если первую карту вытянули и вернули в колоду, то вероятность, что вторая вытянутая карта — пиковая, не зависит от выбора первой карты. Когда исход первого события не влияет на исход второго события, говорят, что второе событие независимо от первого события.

ОПРЕДЕЛЕНИЕ 8.98. Если $P(B) = P(B|A)$, то событие B называется ***независимым*** от события A.

Поскольку $P(A \cap B) = P(A)P(B|A)$, то для случая $P(B|A) = P(B)$ получаем следующую теорему.

ТЕОРЕМА 8.99. Если событие B независимо от события A, то $P(A \cap B) = P(A)P(B)$.

ПРИМЕР 8.100. Предположим, что монету подбрасывают десять раз. Пусть A — событие, при котором монета падает "решкой" вверх первые девять раз, а B — событие, при котором на десятом шаге монета падает "решкой" вверх. Покажем, что событие B является независимым от события A. Вероятность $P(A \cap B)$ — это вероятность, что десять подбрасываний монеты дадут десять выпадений "решкой" вверх. Всего существует $2^{10} = 1024$ равновозможных исходов, и только в одном из них "решка" выпадает вверх десять раз. Поэтому

$$P(A \cap B) = \frac{1}{1024}.$$

Событие A может осуществиться двумя способами. Все десять выпадений монеты будут "решкой" вверх или первые девять выпадений будут "решкой" вверх, а последнее выпадение будет "орлом" вверх. Следовательно,

$$P(A) = \frac{2}{1024}$$

и

$$P(B|A) = \frac{\frac{1}{1024}}{\frac{2}{1024}} = \frac{1}{2} = p(B).$$

Значит, событие B не зависит от события A. □

Если монета подбрасывается n раз, можно считать, что результат каждого подбрасывания не зависит от результатов прочих подбрасываний. Например, на результат третьего подбрасывания никоим образом не влияют результаты первых двух подбрасываний. Такое событие будем называть ***повторением независимых испытаний***. Предположим, что в эксперименте, состоящем из повторения n независимых испытаний, каждое испытание имеет два возможных исхода: успех и неудача. Пусть p — вероятность успеха и $q = 1 - p$ — вероятность неудачи. Требуется найти вероятность k успехов в n независимых испытаниях. Если вернуться к нашему примеру с монетой и рассматривать выпадение "решки" как успех, то $p = q = \frac{1}{2}$. Если $n = 5$ и $k = 3$, то вопрос стоит о вероятности получить три "решки" после пяти подбрасываний монеты.

В случае заданного эксперимента с повторением n независимых испытаний для нахождения вероятности k успехов предположим сначала, что они осуществляются в определенном порядке. Например, первыми реализуются k успехов, а вторыми — $n - k$ неудач. Поскольку испытания независимые, то вероятность осуществления k успехов с последующей реализацией $n - k$ неудач равна $p^k q^{n-k}$.

Фактически, каждое осуществление k успехов в n независимых испытаниях будет иметь вероятность $p^k q^{n-k}$. Поскольку существуют $\binom{n}{k}$ способов получения k успехов в n независимых испытаниях, то вероятность k успехов в n испытаниях равна $\binom{n}{k}p^k q^{n-k}$. Таким образом, приходим к следующей теореме.

ТЕОРЕМА 8.101. В эксперименте с n независимыми испытаниями, называемыми испытаниями Бернулли, где каждое имеет два исхода, успех с вероятностью p и неудачу с вероятностью $q = 1 - p$, вероятность осуществления k успехов равна $\binom{n}{k}p^k q^{n-k}$.

Говорят, что такой эксперимент имеет ***биномиальное распределение*** поскольку

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n.$$

ПРИМЕР 8.102. Дана монета с равновозможным выпадением обеих сторон. Какова вероятность выпадения трех "решек" в результате пяти подбрасываний монеты? Поскольку $p = q = \frac{1}{2}$, вероятность выпадения трех "решек" после пяти подбрасываний равна

$$\binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{10}{32}.$$

□

ПРИМЕР 8.103. Предположим, что правильная игральная кость подбрасывается десять раз. Какова вероятность трехкратного выпадения шестерки? Вероятность p выпадения шестерки равна $\frac{1}{6}$. Следовательно, вероятность q выпадения не шестерки равна $1 - \frac{1}{6} = \frac{5}{6}$. Поэтому вероятность выпадения шестерки равно три раза равна

$$\binom{10}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7.$$

□

Предположим, что в выборочном пространстве каждому исходу поставлено в соответствие действительное число. Например, если подбрасываются две игральные кости, то таким числом может быть сумма показаний на обеих костях. Если монета подбрасывается десять раз, то таким числом может быть общее количество выпадений "решки". Такое число назовем случайной величиной и определим следующим образом.

ОПРЕДЕЛЕНИЕ 8.104. ***Случайной величиной*** называется функция из выборочного пространства во множество действительных чисел.

Если задана случайная величина на выборочном пространстве, можно рассматривать вероятность того, что случайная величина будет равна заданному значению. Например, пусть монета подбрасывается семь раз и $R(x)$ — число выпаданий "решки". Тогда вероятность, что $R(x) = 5$, равна

$$\binom{7}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^2.$$

Если игральная кость подбрасывается десять раз и $R(x)$ — количество выпадений пятерки, то вероятность, что $R(x) = 4$, равна

$$\binom{10}{4}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^6.$$

Предположим, что мы играем в игру, в которой, подбрасывая игральную кость, получаем количество долларов, равное выпавшему числу. Таким образом, если выпадает число 5, мы получаем \$5. Сколько нам следовало бы заплатить за участие в игре? Это количество называется ожидаемым значением. Оно находится суммированием произведений количества долларов, которое должно быть получено, на вероятности получения таких сумм. В данном случае эта вероятность равна $\frac{1}{6}$ для каждого значения, поэтому ожидаемое значение равно

$$1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=\frac{21}{6}=3.5$$

и справедливая цена за участие в игре должна быть \$3.50.

В рассмотренном выше примере случайная величина $R(x)$ принимает значение x, которое выпадает на кости. Таким образом, $R(1) = \$1$, $R(2) = \$2$, ..., $R(6) = \$6$. В общем случае, пусть R — случайная величина, определенная на результатах $x_1, x_2, x_3, \ldots, x_n$, которые осуществляются с соответствующими вероятностями $P_1, P_2, P_3, \ldots, P_n$. ***Ожидаемое значение***, или ***математическое ожидание*** $E(R) = \sum_{i=1}^{n} R(x_i)P_i$.[1] Пусть

$$a_1, a_2, a_3, \ldots, a_n$$

— значения, которые может принимать случайная величина R (т.е. множество результатов функции R). Множество всех результатов, для которых R имеет фиксированное значение a_i, обозначается $R = a_i$. Вероятность того, что X принимает значение $P(R = a_i)$, обозначается через p_i. Если использовать данное обозначение, то эквивалентной, но более удобной формой записи математического ожидания $E(R)$ будет $E(R) = \sum_{i=1}^{m} a_i p_i$.

В приведенных ниже примерах будут определены две случайные величины для рулетки.

ПРИМЕР 8.105. Предположим, что колесо рулетки в равномерно расположенных по колесу ячейках имеет числа от 0 до 36. Заплатив \$1, игрок выбирает число. Если он выигрывает, то получает \$36. В случае проигрыша — не получает ничего. Пусть R_1 — случайная величина, равная количеству денег, которые получит игрок. Пусть e_1 — событие $R = 36$, а e_2 — событие $R = 0$. Тогда p_1, вероятность выигрыша, равна $\frac{1}{37}$, а p_2 — вероятность проигрыша, равна $\frac{36}{37}$. Поэтому,

$$\begin{aligned} E(R_1) &= 36\cdot\frac{1}{37}+0\cdot\frac{36}{37}= \\ &= \frac{36}{37}\approx \\ &\approx .973. \end{aligned}$$

Поскольку игрок платит \$1, то в среднем, за попытку он теряет около \$0.027. □

[1] В отечественной литературе принято обозначение $M(R)$. — *Прим. ред.*

ПРИМЕР 8.106. Предположим, что мы играем в ту же самую игру, но теперь R_2 — случайная величина, равная сумме выигрыша или проигрыша. Пусть e_1 — событие $R = 35$, а e_2 — событие $R_2 = -1$. Тогда p_1, вероятность выигрыша, равна $\frac{1}{37}$, а p_2, вероятность проигрыша, равна $\frac{36}{37}$. Поэтому

$$\begin{aligned} E(R_2) &= 35 \cdot \frac{1}{37} + (-1) \cdot \frac{36}{37} = \\ &= -\frac{1}{37} \approx \\ &\approx -.027. \end{aligned}$$

Поскольку R_2 представляет сумму выигрыша или проигрыша, то в среднем за попытку игрок теряет \$0.027, что согласуется с результатом первого примера. □

ПРИМЕР 8.107. Представим себе ту же самую рулетку, на которой нечетные числа от 1 до 36 — черные, а четные числа от 1 до 36 — красные. Игрок платит \$1 и выбирает цвет. Если он выигрывает, то получает \$2. Если проигрывает, то не получает ничего. Пусть R_3 — случайная величина, равная выигрышу. Пусть e_1 — событие $R_3 = 2$, а e_2 — событие $R_3 = 0$. Тогда p_1, вероятность выигрыша, равна $\frac{18}{37}$, а p_2, вероятность проигрыша, равна $\frac{19}{37}$. Поэтому,

$$\begin{aligned} E(R_3) &= 2 \cdot \frac{18}{37} + 0 \cdot \frac{19}{37} = \\ &= \frac{36}{37} \approx \\ &\approx .973\,. \end{aligned}$$

Поскольку игрок платит \$1, то в среднем за попытку он теряет около \$0.027. □

ПРИМЕР 8.108. Предположим, что мы играем в ту же самую игру, но теперь R_4 — случайная величина, равная выигрышу или проигрышу. Пусть e_1 — событие $R = 1$, а e_2 — событие $R_2 = -1$. Тогда p_1, вероятность выигрыша, равна $\frac{18}{37}$, а p_2, вероятность проигрыша, равна $\frac{19}{37}$. Поэтому

$$\begin{aligned} E(R_4) &= 1 \cdot \frac{18}{37} + (-1) \cdot \frac{19}{37} = \\ &= -\frac{1}{37} \approx \\ &\approx -.027. \end{aligned}$$

Поскольку R_4 представляет сумму выигрыша или проигрыша, то в среднем игрок теряет за попытку \$0.027, что согласуется с результатом первого примера. □

Снова вернемся к испытаниям Бернулли, где случайная величина дает число успехов в n испытаниях. Покажем, что в данном случае математическое ожидание есть произведение количества испытаний на вероятность успеха.

ТЕОРЕМА 8.109. Пусть в эксперименте с n независимыми испытаниями каждое имеет два исхода: успех с вероятностью p и неудачу с вероятностью $q = 1-p$; случайная величина R представляет число успехов. Тогда математическое ожидание $E(R) = np$.

ДОКАЗАТЕЛЬСТВО. Поскольку вероятность k успехов равна

$$\binom{n}{k}p^k q^{n-k} = \frac{n!}{k!(n-k)!}p^k q^{n-k}$$

и первое слагаемое в приведенной ниже сумме равно 0, математическое ожидание равно

$$\begin{aligned} E(R) &= \sum_{k=0}^{n} k\frac{n!}{k!(n-k)!}p^k q^{n-k} = \\ &= \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!}p^k q^{n-k} = \\ &= np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}q^{n-k}. \end{aligned}$$

Положив $m = k-1$, получаем

$$\begin{aligned} E(R) &= np\sum_{m=0}^{n-1}\frac{(n-1)!}{(m)!(n-(m+1))!}p^m q^{n-(m+1)} = \\ &= np\sum_{m=0}^{n-1}\frac{(n-1)!}{(m)!((n-1)-m)!}p^m q^{(n-1)-m} = \\ &= np(p+q)^{n-1} = \\ &= np, \end{aligned}$$

поскольку $p+q=1$. ■

ПРИМЕР 8.110. Монету подбрасывают 20 раз. Игрок получает \$1 за каждое выпадание "решки". Справедливая цена за участие в игре будет $np = 20 \cdot \frac{1}{2} = 10$ долларов. □

Теперь рассмотрим дальнейшие свойства математического ожидания, которые вскоре нам пригодятся.

ТЕОРЕМА 8.111. Пусть R и S — случайные величины, определенные на исходах $x_1, x_2, x_3, \ldots, x_n$, которые осуществляются с соответствующими вероятностями $P_1, P_2, P_3, \ldots, P_n$.

а) Пусть c — константа. Тогда $E(cR) = c \cdot E(R)$.
б) Пусть c — константа. Тогда $E(c) = c$.
в) $E(R+S) = E(R) + E(S)$.

ДОКАЗАТЕЛЬСТВО. Для исходов $x_1, x_2, x_3, \ldots, x_n$, которые осуществляются с соответствующими вероятностями $P_1, P_2, P_3, \ldots, P_n$, математичесткое ожидание $E(R) = \sum_{i=1}^{n} R(x_i)P_i$ и $E(S) = \sum_{i=1}^{n} S(x_i)P_i$. Поэтому

а) $E(cR) = \sum_{i=1}^{n} cR(x_i)P_i = c\sum_{i=1}^{n} R(x_i)P_i = c \cdot E(R)$.

б) $E(c) = \sum_{i=1}^{n} cP_i = c\sum_{i=1}^{n} P_i = c$, поскольку $\sum_{i=1}^{n} P_i = 1$.

в) $E(R+S) = \sum_{i=1}^{n}(R(x_i) + S(x_i))P_i =$

$$= \sum_{i=1}^{n} R(x_i)P_i + S(x_i)P_i = \sum_{i=1}^{n} R(x_i)P_i + \sum_{i=1}^{n} S(x_i)P_i =$$

$$= E(R) + E(S).$$ ■

В случайных выборках мы интуитивно думаем о среднем значении как о некоторой середине или о наиболее вероятном значении случайной величины. В некотором смысле, это ее усредненное значение. Можно заметить, что если каждое значение случайной величины равновероятно, тогда математическое ожидание будет "средним арифметическим" значений случайной величины.

Приведенное ниже определение интуитивно выглядит вполне логичным.

ОПРЕДЕЛЕНИЕ 8.112. ***Среднее значение*** μ случайной величины R — это ее математическое ожидание $E(R)$.

Дисперсия случайной величины R показывает, насколько велик разброс или насколько широко распределены значения случайной величины.

ОПРЕДЕЛЕНИЕ 8.113. ***Дисперсия*** σ^2 случайной величины есть $E((R\text{-}\mu)^2) = E((R - E(R))^2)$.

Используя теорему 8.111, получаем важное свойство дисперсии.

ТЕОРЕМА 8.114. Пусть R — случайная величина, тогда $\sigma^2 = E(R^2) - \mu^2 = E(R^2) - (E(R))^2$.

ДОКАЗАТЕЛЬСТВО.

$\sigma^2 = E((R-\mu)^2) =$

$= E(R^2 - 2R\mu + \mu^2) =$

$= E(R^2) - 2\mu E(R) + E(\mu^2) =$ по теореме 8.111, поскольку μ — константа.

$= E(R^2) - 2\mu^2 + \mu^2 =$ по теореме 8.111, поскольку μ — константа.

$= E(R^2) - \mu^2.$ ■

Важно отметить, что и среднее значение, и дисперсия — теоретические понятия, и их не следует путать со средним значением и разбросом множества данных.

Для примера выберем опять биномиальное распределение. Будем также предполагать, что случайная величина R указывает количество успешных испытаний. Известно уже, что

$$\mu = E(R) = np.$$

Поскольку $\sigma^2 = E(R^2) - (E(R))^2$, то теперь попытаемся найти $E(R^2)$. По определению,

$$E(R^2) = \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^k q^{n-k}.$$

Расписывая $k^2 = k(k-1) + k$ и учитывая тот факт, что первые два слагаемых в приведенной ниже сумме равны 0, получаем

$$\begin{aligned}
E(R^2) &= \sum_{k=0}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k q^{n-k} + \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} = \\
&= \sum_{k=2}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k q^{n-k} + \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} = \\
&= \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)!} p^k q^{n-k} + E(R) = \\
&= n(n-1) \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!(n-k)!} p^k q^{n-k} + E(R) = \\
&= n(n-1)p^2 \sum_{k=2}^{n-2} \frac{(n-2)!}{(k-2)!(n-k)!} p^{k-2} q^{n-2-(k-2)} + E(R) = \\
&= n(n-1)p^2 \sum_{k=0}^{n-2} \frac{(n-2)!}{k!(n-(k+2)!} p^k q^{n-2-k} + E(R) = \\
&= n(n-1)p^2 \sum_{k=0}^{n-2} \frac{(n-2)!}{k!(n-2-k)!} p^k q^{n-2-k} + E(R) = \\
&= n(n-1)p^2 (p+q)^{n-2} + np = \\
&= n(n-1)p^2 + np = \\
&= n^2 p^2 - np^2 + np = \\
&= n^2 p^2 + np(1-p) = \\
&= n^2 p^2 + npq.
\end{aligned}$$

Таким образом,

$$\begin{aligned}
E(R^2) - (E(R))^2 &= n^2 p^2 + npq - (np)^2 = \\
&= npq,
\end{aligned}$$

что приводит к следующей теореме.

ТЕОРЕМА 8.115. Пусть в эксперименте с n независимыми испытаниями каждое имеет два исхода: успех с вероятностью p и неудачу с вероятностью $q = 1 - p$; случайная величина R описывает число успехов. Тогда дисперсия $\sigma^2 = npq$.

ПРИМЕР 8.116. В эксперименте игральную кость подбросили 72 раза. Пусть случайная величина R представляет количество выпаданий числа 1, так что $p = \frac{1}{6}$ и $q = \frac{5}{6}$. Тогда

$$E(R) = \mu = np = 72 \cdot \frac{1}{6} = 12$$

и

$$\sigma^2 = npq = 72 \cdot \frac{1}{6} \cdot \frac{5}{6} = 10.$$

□

Уже было отмечено, что дисперсия — мера разброса значений случайной величины. Неравенство Чебышева дает более точное соотношение между дисперсией и распределением случайной величины. Перед тем как сформулировать и доказать неравенство Чебышева, рассмотрим следующую лемму.

ЛЕММА 8.117. Пусть R — случайная величина, определенная на выборочном пространстве S. Тогда

$$P(|R| \geq c) \leq \frac{E(R^2)}{c^2} = \frac{\sigma^2}{c^2}.$$

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned} P(|R| \geq c) &= \sum_{|R(x_i)| \geq c} P_i \leq \\ &\leq \frac{1}{c^2} \sum_{|R(x_i)| \geq c} |R(x_i)|^2 P_i \leq \\ &\leq \frac{1}{c^2} \sum_{x_i \in S} |R(x_i)|^2 P_i = \\ &= \frac{1}{c^2} E(R^2). \end{aligned}$$

■

Подставляя $R - \mu$ вместо R, получаем ***неравенство Чебышева***.

ТЕОРЕМА 8.118. (Неравенство Чебышева) Пусть R — случайная величина, определенная на выборочном пространстве S. Тогда

$$P(|R - \mu| \geq c) \leq \frac{E((R-\mu)^2)}{c^2} = \frac{\sigma^2}{c^2}.$$

■ УПРАЖНЕНИЯ

1. В партии из пяти автомобилей один имеет дефект. Два человека собирались приобрести автомобили. Какова вероятность того, что вторая из покупаемых машин будет исправна, если известно, что первая машина была без дефекта.
2. Подбрасываются две кости. Какова вероятность того, что на первой выпало 5, если известно, что сумма равна 8?
3. Подбрасываются две кости. Какова вероятность того, что сумма значений на кубике равна 7, если известно, что сумма нечетная?

4. Подбрасываются две кости. Какова вероятность того, что сумма значений на кубике равна 8, если известно, что сумма больше или равна 7?
5. В урне содержится восемь белых и семь черных шаров. Вынимается четыре шара без возвращения. Какова вероятность, что первый шар белый, если известно, что было вытянуто два белых шара? Какова вероятность, что первый шар белый, если известно, что было вытянуто два белых шара (шары вытягивались с возвращением)?
6. Четыре туза из стандартной колоды перетасованы и разложены на столе картинкой вниз. Две карты, не раскрывая, убирают. Какова вероятность, что третья карта — туз пик?
7. Монету подбрасывают три раза. Какова вероятность того, что "решка" выпадет три раза, если известно, что, по крайней мере, один раз выпала "решка"?
8. Пусть A и B — независимые события. Пусть $P(A) = .7$ и $P(B) = .6$, найдите $P(B|A)$.
9. Из трех урн вынимают по одному шару. В первой урне находятся три белых и два черных шара. Во второй — два белых и три черных шара. Третья урна содержит три белых и три черных шара. Если известно, что два из трех вынутых шаров белые, то какова вероятность, что белый шар был вынут из первой урны?
10. Из стандартной колоды с 52 картами вытягивают две карты. Какова вероятность того, что вторая карта — король, если известно, что первая карта — туз? Какова вероятность того, что вторая карта — король, если известно, что первая карта — туз пик?
11. Шесть шаров пронумерованы числами 0, 1, 1, 2, 5 и 10 и помещены в урну. Произвольным образом выбирается один шар, и если он имеет номер n, то игрок получает n долларов. Каково математическое ожидание выигрыша?
12. Шесть шаров пронумерованы числами 0, 0, 0, 0, 10 и 20 и помещены в урну. Произвольным образом выбирается один шар, и если этот шар имеет номер n, то игрок получает n долларов. Каково математическое ожидание выигрыша?
13. Восемь шаров пронумерованы числами 0, 1, 1, 2, 2, 2, 5 и 10 соответственно. Шары помещены в урну, и три из них произвольно вытягиваются. Игрок получает сумму в долларах, равную сумме чисел на трех шарах. Каково математическое ожидание выигрыша?
14. Восемь шаров пронумерованы числами 0, 0, 0, 2, 3, 3, 7 и 10, соответственно. Шары помещены в урну, и три из них произвольно вытягиваются. Игрок получает сумму в долларах, равную сумме чисел на трех шарах. Каково математическое ожидание выигрыша?
15. В игре подбрасывают два кубика. Игрок получает сумму в долларах, равную сумме значений на кубиках. Сколько может стоить участие в этой игре?
16. Предположим, что картам присвоены такие стоимости: туз имеет стоимость 1, двойка — 2, ..., десятка — 10, валет — 11, дама — 12, король — 13. Игрок вытягивает карту. Если вытянута карта червовой масти, то игрок получает удвоенную стоимость карты. Если это карта бубей, он получает стоимость карты. Если карта черной масти, то игрок платит \$10. Каково математическое ожидание в этой игре?

17. Фирма "Ace Cleaning House" учредила главный приз в размере $10 миллионов участнику конкурса. Если человек платит 33 цента за марку (фирма "Ace Cleaning House" предоставляет конверт и анкету участника) и 50 миллионов человек участвуют в тотализаторе, то каково математическое ожидание победы в конкурсе?

18. Допустим, что вероятность рождения девочки 0.51, а мальчика — 0.49. Если в семье Джонс пятеро детей, то какова вероятность, что в ней три девочки? Какова вероятность, что в семье не менее четырех девочек?

19. Если бейсбольная команда выигрывает с вероятностью 60%, то какова вероятность, что команда одержит победу в трех из последующих пяти игр?

20. Две бейсбольные команды, участвующие в турнире, имеют равные шансы на выигрыш. Если команда A проиграла две первые игры, то какова вероятность, что команда A выиграет турнир?

21. Если монету подбрасывают семь раз, то какова вероятность выпадания четырех "решек"?

22. Если игральную кость подбрасывают семь раз, то какова вероятность того, что шестерка выпадет равно два раза?

23. Если 10% продукции составляют брак, то какова вероятность получить два бракованных изделия среди выбранных десяти изделий?

24. Сэм попадает по мячу один раз из трех. Какова вероятность, что он попадет три раза из десяти?

25. Если десять карт последовательно вытягиваются из колоды, то какова вероятность, что две карты будут червовыми?

26. Если две игральные кости подбрасывают 12 раз, то какова вероятность, что 6 выпадет ровно 3 раза?

27. Предположим, что мы играем в рулетку, и рассматриваемая случайная величина такая же, как в примере 8.105, за исключением того, что имеется дополнительная ячейка, обозначенная 00. Найдите математическое ожидание случайной величины R_1.

28. Предположим, что мы играем в рулетку, и рассматриваемая случайная величина такая же, как в примере 8.106, за исключением того, что имеется дополнительная ячейка, обозначенная 00. Найдите математическое ожидание случайной величины R_2.

29. Предположим, что мы играем в рулетку, и рассматриваемая случайная величина такая же, как в примере 8.107, за исключением того, что имеется дополнительная ячейка, обозначенная 00. Найдите математическое ожидание случайной величины R_3.

30. Предположим, что мы играем в рулетку, и рассматриваемая случайная величина такая же, как в примере 8.108, за исключением того, что имеется дополнительная ячейка, обозначенная 00. Найдите математическое ожидание случайной величины R_4.

31. Пусть в эксперименте с 10 независимыми испытаниями каждое имеет два исхода: успех с вероятностью 0.3 и неудачу; случайная величина R описывает число успехов. Каково математическое ожидание эксперимента? Какова дисперсия?

8.10. ТЕОРЕМА БАЙЕСА

Теорема Байеса названа в честь Реверенда Томаса Байеса, министра и математика. Его теория, изложенная в работе *Essay towards Solving a Problem in the Doctrine of Chances*, была опубликована в 1764 году, через четыре года после его смерти, в журнале ***Philosophical Transactions***, издаваемом Лондонским Королевским обществом. У многих людей возникали проблемы с пониманием теоремы Байеса. Казалось, что она меняет местами причину и следствие. Изучая следствие, можно определять вероятность причины. Теорема Байеса весьма полезна для определения причин заболеваний, таких как рак. Она также полезна при изучении воздействия лечения на заболевание.

Предположим, что имеется n непересекающихся событий $B_1, B_2, B_3, \ldots, B_n$, которые разбивают выборочное пространство, и известно, что событие A осуществилось. По определению, $P(B_i|A)$ равна $\frac{P(A \cap B_i)}{P(A)}$. Но

$$A = (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \cup \cdots \cup (A \cap B_n),$$

так что

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + \cdots + P(A \cap B_n)$$

и

$$P(B_i|A) = \frac{P(A \cap B_i)}{P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + \cdots + P(A \cap B_n)}.$$

Поскольку $P(A \cap B_i) = P(A|B_i)P(B_i)$, имеем

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)},$$

откуда следует, что если известны вероятности $P(B_i)$ и $P(A|B_i)$ для каждого i, то можно определить $P(B_i|A)$. Таким образом, мы "обратили" условную вероятность.

Например, предположим, что Сэм вышел из дому, чтобы купить замок для гаража. Вероятность того, что он пойдет в "Andy's Automatic Unopeners" равна 0.3; вероятность того, что он пойдет в "Ben's Best Buys" равна 0.2; и вероятность того, что за покупкой он пойдет в "Cheap Charley's Cash and Credit" равна 0.5. Если он пойдет в "Andy's Automatic Unopeners", то вероятность, что он купит высококачественный автоматический гаражный замок, равна 0.2. Если он пойдет в "Ben's Best Buys", то вероятность сделать такую же покупку равна 0.4. Если он пойдет в "Cheap Charley's Cash and Credit", то вероятность купить высококачественный автоматический гаражный замок равна 0.3. Если Сэм приобрел такой замок, то какова вероятность, что он купил его в "Ben's Best Buys"?

Пусть B_1 — событие, что необходимый замок был куплен в "Andy's Automatic Unopeners", B_2 — событие, что этот замок был куплен в "Ben's Best Buys", и B_3 — событие, что этот замок был куплен в "Cheap Charley's Cash and Credit". Пусть A — событие, которое состоит в том, что Сэм купил необходимый замок. Тогда $P(B_1) = .3$, $P(A|B_1) = .2$, $P(B_2) = .2$, $P(A|B_2) = .4$, $P(B_3) = .5$ и $P(A|B_3) = .3$. Поэтому

$$P(B_2|A) = \frac{(.2)(.4)}{(.3)(.2) + (.2)(.4) + (.5)(.3)} = \frac{8}{29} \approx .28.$$

Если события $B_1, B_2, B_3, \ldots, B_n$ равновозможны, то

$$P(B_1) = P(B_2) = P(B_3) = \cdots = P(B_n).$$

Тогда соотношение для $P(B_i|A)$ сводится к

$$P(B_i|A) = \frac{P(A|B_i)}{P(A|B_1) + P(A|B_2) + P(A|B_3) + \cdots + P(A|B_n)}.$$

Например, предположим, что урна 1 содержит три красных, четыре синих и шесть белых шаров, урна 2 содержит пять красных, три синих и два белых шара, а урна 3 содержит четыре красных, два синих и четыре белых шара. Из произвольным образом выбранной урны выбран шар. Если этот шар красный, то какова вероятность, что выбрана урна 1? Пусть B_1 — событие, что была выбрана урна 1, B_2 — событие, что была выбрана урна 2, а B_3 — событие, что была выбрана урна 3. Пусть R — событие, что был выбран красный шар. Поскольку выбор урны носит случайный характер, то предположим, что $P(B_1) = P(B_1) = P(B_1)$. Следовательно, можно использовать упрощенную формулу. Имеем $P(R|B_1) = .3$, $P(R|B_2) = .5$ и $P(R|B_3) = .4$. Поэтому

$$P(B_1|R) = \frac{.3}{.3 + .5 + .4} = .25.$$

■ УПРАЖНЕНИЯ

1. Компания производит автопокрышки на трех заводах. Сорок пять процентов покрышек производится на заводе A, 25% — на заводе B и 30% производятся на заводе C. Предположим, что 5% автопокрышек, произведенных на заводе A, бракованные; на заводе B брак составляет 10%; а на заводе C бракованные автопокрышки составляют 2%. Если купленная автопокрышка оказалась бракованной, то какова вероятность, что она была произведена на заводе B?
2. Коробка 1 содержит три красных, четыре белых и один синий шар. Коробка 2 содержит два красных, четыре белых и два синих шара. Коробка 3 содержит два красных, три белых и три синих шара. Если произвольным образом выбранный шар оказался синим, то какова вероятность, что он был вынут из коробки 3?
3. В коробке лежат три монеты. У одной монеты на обеих сторонах — "решки", одна монета правильная, и еще одна монета выпадает "решкой" вверх в 70% случаев. Монету выбирают случайным образом и подбрасывают. Если монета выпадет "решкой", то какова вероятность того, что была выбрана правильная монета?
4. Анализ крови является на 90% эффективным при диагнозе некоторого заболевания. В 4% случаев анализ ошибочно устанавливает болезнь у здорового человека. Если 10% людей имеют это заболевание, то какова вероятность, что человек, получивший результат анализа, подтверждающий болезнь, в действительности имеет это заболевание?
5. Студент с одинаковой вероятностью может выбрать один из трех курсов из блока социально-гуманитарных дисциплин. Если он выберет "Историю

Канзаса, 1995–1997", то вероятность сдачи экзамена составит 40%. Если он выберет "Курс устного счета повышенного типа II", то вероятность сдать экзамен — 30%. При выборе курса под названием "Телевидение III: выбор каналов", вероятность сдачи экзамена составит 60%. Если студент сдаст экзамен, то какова вероятность того, что он выбрал "Историю Канзаса, 1995–1997"?

6. В процессе производства 80% изделий получаются высококачественными. Из тех, что проверены, 10% изделий признаны бракованными и не подлежащими продаже. Только 5% бракованных изделий улучшены и отправлены в продажу. Если изделие поступило в продажу, то какова вероятность того, что изделие имеет брак?
7. Допустим, что страховая компания классифицирует водителей как умелых, неумелых и нормальных. Предположим, что 25% водителей являются умелыми, 50% — нормальными и 25% — неумелыми. Предположим также, что у умелого водителя вероятность аварии в следующем году составит 10%, вероятность аварии у нормального водителя — 20%, а вероятность аварии у неумелого водителя составит 30%? Если в прошлом году у Сэма была авария, то какова вероятность того, что он умелый водитель?
8. Подбрасывают правильную монету. Если она падает "решкой" вверх, то подбрасывают две игральные кости. Игрок получает сумму в долларах, равную выпавшей на кубиках. Если монета падает вверх "орлом", то подбрасывают три монеты, и за каждую выпавшую "решку" игрок получает \$4. Если игрок выиграл \$8, то какова вероятность, что первая монета выпала "решкой"?

8.11. ЦЕПИ МАРКОВА

Предположим, что проводится серия экспериментов с возможными исходами $s_1, s_2, s_3, \ldots, s_n$. Назовем эти исходы ***состояниями***. Предположимь, что мы начинаем в одном из этих состояний и что $p_i^{(0)}$ — вероятность того, что мы начинаем в состоянии s_i. Далее будем предполагать, что вероятность данного исхода в текущем эксперименте зависит только от исхода предыдущего эксперимента. Пусть p_{ij} — вероятность того, что в результате эксперимента состояние было изменено от состояния s_i к состоянию s_j. Таким образом, эксперимент начался в состоянии s_i и закончился в состоянии s_j. Пусть $p_i^{(1)}$ — вероятность того, что исходом эксперимента будет состояние s_i. Тогда

$$p_i^{(1)} = p_1^{(0)}p_{1i} + p_2^{(0)}p_{2i} + p_3^{(0)}p_{3i} + \cdots + p_1^{(0)}p_{1i},$$

а это означает, что вероятность исхода в состоянии s_i равна сумме вероятностей начать эксперимент в некотором другом состоянии и окончить в состоянии s_i. Если эксперимент был начат в состоянии s_j, то в результате эксперимента исходом должно стать некоторое состояние. Поэтому для каждого $0 \leq j \leq n$ имеем

$$p_{j1} + p_{j2} + p_{j3} + \cdots + p_{jn} = 1.$$

Пусть $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)}, \ldots, p_n^{(0)})$ и $p^{(1)} = (p_1^{(1)}, p_2^{(1)}, p_3^{(1)}, \ldots, p_n^{(1)})$. Предположим для простоты, что существуют три возможных состояния, поэтому $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)})$. Пусть T — матрица

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}.$$

Матрица T называется ***матрицей перехода***. В общем случае, матрица T имеет вид

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} \end{bmatrix}.$$

Пусть

$$(a, b, c) = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)}) \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix},$$

тогда

$$\begin{aligned} a &= p_1^{(0)} p_{11} + p_2^{(0)} p_{21} + p_3^{(0)} p_{31} = p_1^{(1)}, \\ b &= p_1^{(0)} p_{12} + p_2^{(0)} p_{22} + p_3^{(0)} p_{32} = p_2^{(1)}, \\ c &= p_1^{(0)} p_{13} + p_2^{(0)} p_{23} + p_3^{(0)} p_{33} = p_3^{(1)}, \end{aligned}$$

так что

$$(p_1^{(1)}, p_2^{(1)}, p_3^{(1)}) = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)}) \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}.$$

ПРИМЕР 8.119. Погода классифицируется в прогнозах как ясная, умеренно пасмурная и пасмурная. Если погода ясная, то вероятность, что она будет ясной на следующий день, составляет 0.5; вероятность, что она будет умеренно пасмурной, составляет 0.4; а вероятность пасмурной погоды на следующий день составляет 0.1. Если погода умеренно пасмурная, то вероятность, что на следующий день она будет ясной, равна 0.3; вероятность, что погода останется умеренно пасмурной, равна 0.5; и вероятность, что погода станет пасмурной, равна 0.2. Если погода пасмурная, то вероятность, что на следующий день она останется пасмурной, составляет 0.4; вероятность, что она станет умеренно пасмурной, составляет 0.4; и вероятность, что она станет ясной, составляет 0.2. Если вероятность ясной погоды в воскресенье равна 0.6, а вероятность умеренной пасмурности — 0.4, то какова вероятность, что погода в понедельник будет ясной? Какова вероятность, что во вторник погода будет умеренно пасмурной?

Если порядок, в котором перечисляются погодные условия, таков: ясно, умеренно пасмурно и пасмурно, то

$$p^{(0)} = (.6, .4, 0)$$

и

$$T = \begin{bmatrix} .5 & .4 & .1 \\ .3 & .5 & .2 \\ .2 & .4 & .4 \end{bmatrix}.$$

Следовательно,

$$p^{(1)} = (.6, .4, 0) \begin{bmatrix} .5 & .4 & .1 \\ .3 & .5 & .2 \\ .2 & .4 & .4 \end{bmatrix} = (.42, .44, .14),$$

и вероятность, что в понедельник будет ясная погода, равна 0.42.

Пусть $p_1^{(2)}$ — вероятность того, что во вторник будет ясная погода, $p_2^{(2)}$ — вероятность того, что во вторник будет умеренно пасмурно и $p_3^{(2)}$ — вероятность того, что во вторник будет пасмурно. Пусть $p^{(2)} = (p_1^{(2)}, p_2^{(2)}, p_3^{(2)})$. Тогда

$$p^{(2)} = (.42, .44, .14) \begin{bmatrix} .5 & .4 & .1 \\ .3 & .5 & .2 \\ .2 & .4 & .4 \end{bmatrix} = (.37, .444, .186).$$

Следовательно, вероятность того, что во вторник будет умеренно пасмурная погода, равна 0.444. □

Пусть $p_i^{(m)}$ — вероятность, что исходом m-го проведения эксперимента будет состояние s_i, и $p^{(m)} = (p_1^{(m)}, p_2^{(m)}, p_3^{(m)}, \ldots, p_n^{(m)})$.

ТЕОРЕМА 8.120. Для любого положительного целого числа m имеем $p^{(m)} = p^{(0)}T^m$.

ДОКАЗАТЕЛЬСТВО. Докажем теорему, используя индукцию. Было показано, что для $m = 1$ утверждение справедливо. Предположим, что оно справедливо для $n = k$, так что $p^{(k)} = p^{(0)}T^k$. Поскольку

$$p_j^{(k+1)} = p_1^{(k)}p_{1j} + p_2^{(k)}p_{2j} + p_3^{(k)}p_{3j} + p_n^{(k)}p_{n3j},$$

то

$$p^{(k+1)} = p^{(k)}T = p^{(0)}T^kT = p^{(0)}T^{k+1}.$$

■

ПРИМЕР 8.121. В процессе опроса владельцев автомобилей трех американских марок: марки A, марки B и марки C, им был задан вопрос о том, какую торговую марку они бы выбрали для следующей покупки. Среди владельцев автомобилей

марки A 20% в сказали, что они опять выбрали бы марку A, в то время как 50% ответили, что они бы перешли на марку B, а 30% заявили, что предпочли бы марку C. Среди владельцев автомобилей марки B 20% сказали, что перейдут на марку A, в то время как 70% заявили, что приобрели бы опять автомобиль марки B, а 10% заявили, что в следующий раз предпочли бы марку C. Среди владельцев автомобилей марки C, 30 процентов ответили, что перешли бы на марку A, 30% сказали, что перешли бы на марку B, а 40% заявили, что остались бы верны той же торговой марке C.

(а) Если для первой машины некто приобрел автомобиль марки A, то какова вероятность, что его второй машиной будет автомобиль марки C?

(б) Если при покупке первой машины покупатель подбросил монету, выбирая между маркой B и маркой C, то какова вероятность, что его третьей машиной будет автомобиль марки A?

(в) Если все три марки автомобилей равновозможны для покупки первого автомобиля, то какова вероятность, что третьей машиной станет автомобиль марки B?

Матрица перехода для этого события имеет вид

$$\begin{bmatrix} .2 & .5 & .3 \\ .2 & .7 & .1 \\ .3 & .3 & .4 \end{bmatrix}.$$

Для случая (а) имеем $p^{(0)} = (1, 0, 0)$, поэтому

$$p^{(1)} = (1, 0, 0) \begin{bmatrix} .2 & .5 & .3 \\ .2 & .7 & .1 \\ .3 & .3 & .4 \end{bmatrix} = (.2, .5, .3).$$

Вероятность того, что вторая машина будет марки C, равна 0.3. Для случаев (б) и (в) требуется найти

$$T^2 = \begin{bmatrix} .23 & .54 & .23 \\ .21 & .62 & .17 \\ .24 & .48 & ..28 \end{bmatrix}.$$

Для (б) имеем $p^{(0)} = (0, .5, .5)$ и

$$p^{(2)} = (0, .5, .5) \begin{bmatrix} .23 & .54 & .23 \\ .21 & .62 & .17 \\ .24 & .48 & ..28 \end{bmatrix} = (.225, .55, .225),$$

поэтому вероятность того, что третий автомобиль будет марки A, равна 0.225. □

ПРИМЕР 8.122. Для случая (в) имеем $p^{(0)} = (.333, .333, .333)$ и

$$p^{(2)} = (.333, .333, .333) \begin{bmatrix} .23 & .54 & .23 \\ .21 & .62 & .17 \\ .24 & .48 & ..28 \end{bmatrix} = (.227, .547, .227),$$

где все числа округлены до трех цифр после десятичной точки. Поэтому вероятность выбора марки C для третьей машины равна 0.227. □

■ УПРАЖНЕНИЯ

1. Пусть осуществление событий A, B и C представлено состояниями s_1, s_2 и s_3 соответственно. Пусть переход осуществляется каждый день, и матрица перехода имеет вид

$$\begin{bmatrix} .2 & .5 & .3 \\ .3 & .6 & .1 \\ .2 & .3 & ..5 \end{bmatrix}.$$

а) Какова вероятность того, что если событие A произошло в первый день, то оно произойдет снова на третий день?

б) Какова вероятность того, что если событие B произошло в первый день, то событие C произойдет на четвертый день?

в) Какова вероятность того, что если событие A произошло в первый день, то событие B произойдет на пятый день?

2. Пусть осуществление событий A, B и C представлено состояниями s_1, s_2 и s_3, соответственно. Пусть переход осуществляется каждый день, и матрица перехода имеет вид

$$\begin{bmatrix} .3 & .4 & .3 \\ .4 & .2 & .4 \\ .2 & .5 & .3 \end{bmatrix}.$$

а) Какова вероятность того, что если событие A произошло в первый день, то событие B произойдет на третий день?

б) Какова вероятность того, что если событие A произошло в первый день, то событие C произойдет на четвертый день?

в) Какова вероятность того, что если событие C произошло в первый день, то событие C произойдет на пятый день?

3. Аптека продает средство от изжоги трех торговых марок: A, B и C. Покупатели меняют марки препарата согласно матрице перехода

$$\begin{bmatrix} .7 & .2 & .1 \\ .2 & .6 & .2 \\ .1 & .1 & .8 \end{bmatrix}.$$

Если препарат приобретается ежемесячно, то:

а) Какова вероятность того, что если в этом месяце приобретен препарат марки A, то через три месяца будет приобретен препарат марки C?

б) Какова вероятность того, что если в этом месяце приобретен препарат марки B, то через два месяца будет приобретен препарат марки ?

в) Какова вероятность того, что если в этом месяце приобретен препарат марки C, то через три месяца будет опять приобретен препарат марки C?

4. Предположим, что Конгресс объявил о возможном сокращении налогов, и это сообщение распространяется от человека к человеку. Если сокращение

налогов состоится, то вероятность того, что сообщение будет передано следующему человеку правильно, составляет 0.6. Если сокращения налогов не будет, то вероятность правильности передачи информации составит 0.7.

а) Если сокращение налогов состоится, то какова вероятность, что четвертый человек, услышавший эту новость, услышал ее правильно?

б) Если сокращение налогов не состоится, то какова вероятность, что третий человек, услышавший эту новость, услышит ее правильно?

в) Если сокращение налогов не состоится, то какова вероятность, что пятый человек, услышавший эту новость, поймет, что сокращение налогов состоится?

5. Команда A и команда B участвуют в ежегодном чемпионате США по бейсболу. Каждая команда имеет равные шансы выиграть первую игру. Если команда A выиграет первую игру, то вероятность ее победы в следующей игре равна 0.6. Если первую игру выиграет команда B, то вероятность ее победы в следующей игре равна 0.7. Какова вероятность, что команда B выиграет третью игру? Какова вероятность, что команда A выиграет четвертую игру?

6. У самолета четыре мотора. Для выполнения боевого задания ему необходимо оставаться в воздухе в течение часа. Если из-за вражеского обстрела вероятность выхода из строя одного из моторов в любой 15-минутный период времени равна 0.25 и для продолжения полета самолету нужны, по крайней мере, два мотора, то какова вероятность, что самолет вернется на базу? Предполагается, что вероятность того, что самолет с тремя поврежденными моторами будет оставаться в воздухе настолько долго, что может выйти из строя четвертый мотор, равна 0.

7. В колледже 80% первокурсников становятся второкурсниками. 10% бросают учебу или переводятся. 90% в второкурсников на следующий год становятся студентами предпоследнего года обучения. 5% второкурсников бросают учебу или переводятся. 95% студентов предпоследнего курса становятся на следующий год старшекурсниками. 2% студентов предпоследнего курса бросают учебу или переводятся. 1% старшекурсников бросают учебу или переводятся, а 98% заканчивают колледж. Те, кто не заканчивают, бросают учебу, переводятся или остаются на второй год обучения. Какова вероятность, что поступивший в колледж в течение четырех лет бросит учебу или переведется? Если 1000 новичков стали студентами, то сколько из них закончат обучение через четыре года? Сколько закончат обучение через пять лет?

ГЛАВА

9

АЛГЕБРАИЧЕСКИЕ СТРУКТУРЫ

9.1. ВНОВЬ О ЧАСТИЧНО УПОРЯДОЧЕННЫХ МНОЖЕСТВАХ

Согласно определению 2.44, отношение R на множестве A есть отношение частичного порядка, если оно рефлексивно, антисимметрично и транзитивно. Множество A в этом случае называется ***частично упорядоченным множеством*** или ***ЧУ-множеством*** с порядком R. Отношение частичного порядка будем обозначать символом $\leq$ (или $\geq$, где $a \geq b$ если $a \leq b$). Частично упорядоченное множество A с порядком $\leq$ будем обозначать через $(A, \leq)$.

Согласно определению 2.47, два элемента a и b частично упорядоченного множества (ЧУ-множества) $(S, \leq)$ ***сравнимы***, если $a \leq b$ или $b \leq a$. Если каждые два элемента частично упорядоченного множества $(S, \leq)$ сравнимы, тогда $(S, \leq)$ называется ***вполне упорядоченным множеством***, или ***цепью***.

ОПРЕДЕЛЕНИЕ 9.1. Для подмножества B ЧУ-множества A элемент a из A называется ***верхней гранью*** B, если $b \leq a$ (или $a \geq b$) для всех b из B. Элемент a называется ***наименьшей верхней гранью*** (нвг) подмножества B, если (а) a — верхняя грань B и (б) если любой другой элемент a' множества A является верхней гранью B, то $a \leq a'$. Наименьшая верхняя грань для всего ЧУ-множества A (если она существует) называется ***наибольшим элементом*** A. Для подмножества B ЧУ-множества A элемент a из A называется ***нижней гранью*** B, если $a \leq b$ (или $b \geq a$) для всех b из B. Элемент a называется ***наибольшей нижней гранью*** (ннг) подмножества B, если (а) a — нижняя грань B; (б) если любой другой элемент a' из A является нижней гранью B, то $a \geq a'$. Наибольшая нижняя грань всего ЧУ-множества A (если она существует) называется ***наименьшим элементом*** A.

ОПРЕДЕЛЕНИЕ 9.2. Элемент a подмножества B ЧУ-множества A называется ***максимальным элементом*** B, если для каждого элемента $b \in B$ из того, что $b \geq a$, следует $b = a$. Другими словами, во множестве B нет элемента, который

был бы "больше", чем a. Элемент a подмножества B ЧУ-множества A называется ***минимальным элементом*** B, если для каждого элемента $b \in B$ из того, что $b \leq a$, следует $b = a$. Другими словами, в B нет элемента, который был бы "меньше", чем a. Обычно термины "минимальный" и "максимальный" элемент относятся ко всему множеству, т.е. $A = B$.

ПРИМЕР 9.3. Пусть, как в примере 2.45, $C = \{1, 2, 3\}$ и X — булеан множества C:

$$X = P(C) = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

Определим отношение $\leq$ на множестве X таким образом: $T \leq V$, если $T \subseteq V$. По определению, $\{1,2\}$ есть наибольшая нижняя грань для $\{\varnothing, \{1\}, \{2\}\}$, а также для $\{\varnothing, \{1\}, \{2\}, \{1,2\}\}$. Множество $\{1,2,3\}$ — наименьшая верхняя грань для X. Элемент $\varnothing$ является наибольшей нижней гранью для всех трех множеств. □

ПРИМЕР 9.4. Рассмотрим множество $\hat{X} = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\}$ с таким же отношением порядка, как и в примере 9.3. В этом случае множества $\{1,2\}$, $\{1,3\}$, и $\{2,3\}$ являются максимальными элементами множества $\hat{X}$, но в $\hat{X}$ нет наибольшего элемента. Такая ситуация возможна, потому что данные три максимальных элемента не сравнимы между собой. □

ПРИМЕР 9.5. Рассмотрим множество $\ddot{X} = \{\{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$ с тем же отношением порядка, как и в примере 9.3. В этом случае $\{1\}$, $\{2\}$ и $\{3\}$ — минимальные элементы множества $\ddot{X}$, но наименьшего элемента множество $\ddot{X}$ не имеет. □

ПРИМЕР 9.6. Пусть T — множество всех положительных делителей числа 30 и $\leq_1$ есть отношение порядка, которое задается таким образом: $m \leq_1 n$, если m делит n. Наименьшая верхняя грань для $\{3,5\}$ — число 15. Понятно, что наименьшая верхняя грань, 30, и наибольшая нижняя грань, 1, множества T сравнимы. Понятно, что $5 \leq_1 15$, поскольку 15 делится на 5, но 5 и 6 не сравнимы. □

ПРИМЕР 9.7. Наименьшая верхняя грань и наибольшая нижняя грань множества не обязательно являются элементами этого множества. Пусть Q — множество рациональных чисел и пусть $\leq$ — обычное отношение "меньше или равно", и $<$ — отношение "меньше" для рациональных чисел. Тогда для множества $A = \{x \in Q : 0 < x < 1\}$ с частичным порядком $\leq$ имеем $0 = \text{ннг}(A)$ и $1 = \text{нвг}(A)$, однако $0, 1 \notin A$. □

Если для каждого двухэлементного подмножества ЧУ-множества A существует наименьшая верхняя грань, принадлежащая A, то на множестве A можно определить следующее бинарное отношение. Если a и b принадлежат A, то пусть $a \vee b = \text{нвг}\{a, b\}$. Если используются обозначения булевой алгебры, то вместо $a \vee b$ записывают $a + b$.

ОПРЕДЕЛЕНИЕ 9.8. ЧУ-множество A, все двухэлементные подмножества которого имеют наименьшую верхнюю грань, принадлежащую A, называется ***верхней полурешеткой*** и обозначается $(\mathbf{A}, \vee)$ или $(A, +)$.

Если для каждого двухэлементного подмножества ЧУ-множества A существует наибольшая нижняя грань, принадлежащая A, то на множестве A можно определить следующее бинарное отношение. Если a и b принадлежат A, то пусть $a \wedge b = \text{ннг}\{a, b\}$. Если используются обозначения булевой алгебры, то вместо $a \wedge b$ записывают $a \cdot b$.

ОПРЕДЕЛЕНИЕ 9.9. ЧУ-множество A, все двухэлементные подмножества которого имеют наибольшую нижнюю грань, принадлежащую A, называется ***нижней полурешеткой*** и обозначается $(\mathbf{A}, \wedge)$ или $(A, \cdot)$.

В примере 9.3 $\{1,2\} \vee \{1,3\} = \{1,2,3\}$ и $\{1,2\} \wedge \{1,3\} = \{1\}$. В этом случае наименьшая верхняя грань и наибольшая нижняя грань — объединение и пересечение множеств соответственно, и обозначения совпадают с обозначениями теории множеств.

В примере 9.6 $6 \vee 15 = 30$ и $6 \wedge 15 = 3$. В этом случае наименьшая верхняя грань и наибольшая нижняя грань — это наименьшее общее кратное и наибольший общий делитель соответственно.

Доказательство приведенной ниже теоремы, в которой сформулированы алгебраические свойства полурешеток, предоставляется читателю.

ТЕОРЕМА 9.10.

а) Пусть A — верхняя полурешетка. Тогда для всех $a, b, c \in A$ $a \vee (b \vee c) = (a \vee b) \vee c$, $a \vee a = a$ и $a \vee b = b \vee a$.

б) Пусть A — нижняя полурешетка. Тогда для всех $a, b, c \in A$ $a \wedge (b \wedge c) = (a \wedge b) \wedge c$, $a \wedge a = a$ и $a \wedge b = b \wedge a$.

■ УПРАЖНЕНИЯ

1. Постройте диаграмму Гессе (см. раздел 2.6) для ЧУ-множества из примера 9.3.
2. Постройте диаграмму Гессе для ЧУ-множества из примера 9.6.
3. Постройте диаграмму Гессе для ЧУ-множества из примера 9.4.
4. Постройте диаграмму Гессе для ЧУ-множества из примера 9.5.
5. По диаграмме Гессе для ЧУ-множества, изображенной на рис. 9.1, определите:
   а) наибольший элемент ЧУ-множества (если он существует);
   б) наименьший элемент ЧУ-множества (если он существует);
   в) максимальный и минимальный элементы ЧУ-множества;
   г) является ли рассматриваемое ЧУ-множество верхней или нижней полурешеткой (или и тем, и другим)?

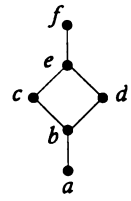


Рис. 9.1

6. По диаграмме Гессе для ЧУ-множества, изображенной на рис. 9.2, определите:

а) наибольший элемент ЧУ-множества (если он существует);

б) наименьший элемент ЧУ-множества (если он существует);

в) максимальный и минимальный элементы ЧУ-множества;

г) является ли рассматриваемое ЧУ-множество верхней или нижней полурешеткой (или и тем, и другим)?

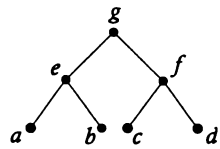


Рис. 9.2

7. По диаграмме Гессе для ЧУ-множества, изображенной на рис. 9.3, определите:

а) наибольший элемент ЧУ-множества (если он существует);

б) наименьший элемент ЧУ-множества (если он существует);

в) максимальный и минимальный элементы ЧУ-множества;

г) является ли рассматриваемое ЧУ-множество верхней или нижней полурешеткой (или и тем, и другим)?

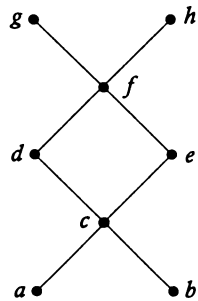


Рис. 9.3

8. По диаграмме Гессе для ЧУ-множества, изображенной на рис. 9.4, определите:

а) наибольший элемент ЧУ-множества (если он существует);

б) наименьший элемент ЧУ-множества (если он существует);

в) максимальный и минимальный элементы ЧУ-множества;

г) является ли рассматриваемое ЧУ-множество верхней или нижней полурешеткой (или и тем, и другим)?

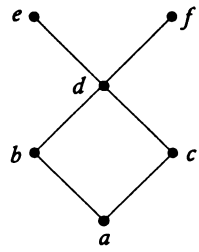


Рис. 9.4

9. По диаграмме Гессе для ЧУ-множества, изображенной на рис. 9.5, определите:

а) наибольший элемент ЧУ-множества (если он существует);

б) наименьший элемент ЧУ-множества (если он существует);

в) максимальный и минимальный элементы ЧУ-множества;

г) является ли рассматриваемое ЧУ-множество верхней или нижней полурешеткой (или и тем, и другим)?

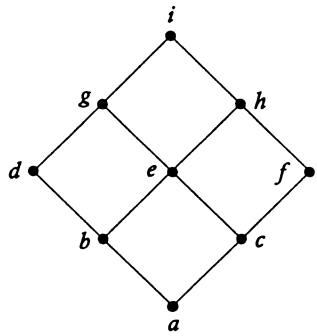


Рис. 9.5

10. Основываясь на диаграммах Гессе из предыдущих упражнений, опишите их характерные свойства, которые гарантируют, что

а) ЧУ-множество имеет максимальный элемент;

б) ЧУ-множество имеет наименьшую верхнюю грань (или наибольший элемент);

в) ЧУ-множество имеет минимальный элемент;

г) ЧУ-множество имеет наибольшую нижнюю грань (или наименьший элемент).

11. Докажите следующее:

а) пусть A — верхняя полурешетка. Тогда $a \vee (b \vee c) = (a \vee b) \vee c$ для всех $a, b, c \in A$, $a \vee a = a$ для всех $a \in A$ и $a \vee b = b \vee a$ для всех $a, b \in A$;

б) пусть A — нижняя полурешетка. Тогда $a \wedge (b \wedge c) = (a \wedge b) \wedge c$ для всех $a, b, c \in A$, $a \wedge a = a$ для всех $a \in A$ и $a \wedge b = b \wedge a$ для всех $a, b \in A$.

12. Докажите, что в каждом конечном ЧУ-множестве имеются максимальный и минимальный элементы.

13. Докажите, что в каждой конечной верхней полурешетке имеется наименьшая верхняя грань (или наибольший элемент).

14. Докажите, что в каждой конечной нижней полурешетке имеется наибольшая нижняя грань (или наименьший элемент).

15. В главе 2 булевы операции $\vee$ и $\wedge$ на множестве $\{0,1\}$ определены таблицами:

$\vee$	0	1
0	0	1
1	1	1

и

$\wedge$	0	1
0	0	0
1	0	1

.

Была введена также булева матрица как матрица, у которой все элементы 0 или 1. Если $A = [A_{ij}]$ и $B = [B_{ij}]$ — булевы матрицы размерности $m \times n$, то

$U = A \vee B$ определяется как $U_{ij} = A_{ij} \vee B_{ij}$ при $1 \leq i \leq m, 1 \leq j \leq n$;
$V = A \wedge B$ определяется как $V_{ij} = A_{ij} \wedge B_{ij}$ при $1 \leq i \leq m, 1 \leq j \leq n$.

Пусть S — множество всех булевых матриц размерности $n \times n$. На множестве S определим отношение $\leq$ следующим образом: $A \leq B$ для $A, B \in S$, если $A_{ij} \leq B_{ij}$ при $1 \leq i, j \leq n$.

а) Докажите, что множество S с этим отношением представляет собой ЧУ-множество.

б) Докажите, что $A \vee B$ — наименьшая верхняя грань для $\{A,B\}$.

в) Докажите, что $A \wedge B$ — наибольшая нижняя грань для $\{A,B\}$.

г) Докажите, что S является и верхней, и нижней полурешеткой.

д) Что является наибольшим и наименьшим элементом для множества S?

9.2. ПОЛУГРУППЫ И ПОЛУРЕШЕТКИ

В определении 4.21 бинарная операция на множестве S определена как функция $b : S \times S \to S$. Поскольку множество значений бинарной операции на множестве S является подмножеством множества S, то, по определению, бинарная операция обладает свойством замыкания, когда результат операции над двумя элементами r и s множества S также является элементом S. Существует достаточно много примеров бинарных операций.

Например, если S — множество положительных целых чисел, то произведение и сумма — бинарные операции, поскольку и произведение, и сумма положительных целых чисел есть положительное целое число. Во многих случаях такие операции являются еще и ассоциативными, например, как для сложения положительных целых чисел $(a+b)+c = a+(b+c)$ и умножения положительных целых чисел $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. Пусть S — множество матриц размерности $n \times n$, элементы которых — целые числа. Сумма и произведение элементов из S снова являются элементами S. Более того, если A, B и C — матрицы размерности $n \times n$, то $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ и $(A + B) + C = A + (B + C)$. В приведенном ниже определении бинарным операциям с такой структурой даны специальные названия.

ОПРЕДЕЛЕНИЕ 9.11. Множество S с такой бинарной операцией $*$ на $S \times S$, что для всех a, b и c из S имеет место $(a * b) * c = a * (b * c)$, называется ***полугруппой*** и обозначается $(S, *)$, или просто S, если понятно, какая операция имеется в виду. Если, в дополнение к этому, для всех a и b из S выполняется $a * b = b * a$, то множество S с оператором $*$ называется ***абелевой***, или ***коммутативной полугруппой***. Если в $(S, *)$ существует элемент I такой, что $I * a = a * I = a$ для всех a из A, то такое I называется ***единицей*** полугруппы $(S, *)$, а $(S, *)$ называется ***полугруппой с единицей***, или ***моноидом***. Если $(S, *)$ — полугруппа, и $\overline{S} \subseteq S$, то $\overline{S}$ называется ***подполугруппой*** полугруппы S, если $*$ — бинарная операция на $\overline{S}$. Это эквивалентно следующему: $(\overline{S}, *)$ — подполугруппа полугруппы $(S, *)$, если $\overline{S} \subseteq S$, и для каждых $a, b \in \overline{S}$ имеем $a * b \in \overline{S}$.

Отметим еще раз: тот факт, что $*$ является бинарной операцией на множестве $S \times S$, означает, что если $s, s' \in S$, то $s * s' \in S$. Указанное свойство называется ***замкнутостью***.

ПРИМЕР 9.12. Множество S матриц размерности $n \times n$ является моноидом $(S, \cdot)$. Единицей в данном случае будет матрица I с единицей в качестве элементов главной диагонали и нулем в качестве прочих элементов, так что

$$I_{ij} = \begin{cases} 1, & \text{если } i = j; \\ 0, & \text{если } i \neq j. \end{cases}$$

Следовательно, единичная матрица размерности 3×3 имеет вид

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

□

ПРИМЕР 9.13. Если $(S, \cdot)$ — полугруппа матриц размерности $n \times n$, элементами которых являются рациональные числа с операцией матричного умножения и если $(\overline{S}, \cdot)$ — полугруппа матриц размерности $n \times n$, элементы которых — целые числа, тогда $(\overline{S}, \cdot)$ — подполугруппа полугруппы $(S, \cdot)$. □

ПРИМЕР 9.14. Для положительного целого числа n положим $S_n^0 = \{x : x \text{ целое число и } x \geqslant n\} \cup \{0\}$. Полугруппа S_n^0 — коммутативный моноид с операцией сложения целых чисел и нулем в качестве единицы. Пусть $S_n^1 = \{x : x \text{ целое число и } x \geqslant n\} \cup \{1\}$. Полугруппа S_n^1 — коммутативный моноид с операцией умножения целых чисел с единицей в качестве единицы. Если $m \geq n$, то S_m^0 — подполугруппа полугруппы S_n^0 и S_m^1 — подполугруппа полугруппы S_n^1 □

ПРИМЕР 9.15. Пусть S — множество всех функций из непустого множества A в само множество A с бинарной операцией композиции. Таким образом, для $f, g \in S$ результат операции $f \circ g$ определяется как $(f \circ g)(a) = f(g(a))$ для всех $a \in A$. Поскольку для $f, g, h \in S$ имеет место $((f \circ g) \circ h)(a) = (f \circ (g \circ h))(a) = f(g(h(a)))$ для всех $a \in A$, то $(f \circ g) \circ h = f \circ (g \circ h)$, поэтому S — полугруппа. Далее функция $I : A \to A$, определенная как $I(a) = a$ для всех $a \in A$, является единицей в S и, следовательно, S — моноид. □

ПРИМЕР 9.16. Пусть A — конечный набор символов. Такое множество символов называется ***алфавитом***. Например, A могло бы быть подмножеством английского алфавита или просто множеством $\{0,1\}$. ***Строка*** или ***слово*** символов из множества A имеет вид $a_1a_2a_3a_4\ldots a_n$, где $a_i \in A$. Таким образом, если $A = \{0,1\}$, то 1, 100, 101, 11011 и 0001 — возможные строки символов из множества A. Кроме того, определим пустую строку, которая обозначается λ и не содержит никаких символов. Пусть A^* обозначает множество всех строк алфавита A. Введем бинарную операцию $\circ$, называемую ***конкатенацией*** на A^*, следующим образом: если $a_1a_2a_3a_4\ldots a_n$ и $b_1b_2b_3b_4\ldots b_m \in A^*$, то

$$a_1a_2a_3a_4\ldots a_n \circ b_1b_2b_3b_4\ldots b_m = a_1a_2a_3a_4\ldots a_nb_1b_2b_3b_4\ldots b_m.$$

Например, если $A = \{0,1\}$, то $11011 \circ 100010 = 11011100010$. В частности, если ω — строка в A^*, то $\lambda \circ \omega = \omega \circ \lambda = \omega$, и это говорит о том, что если строке предшествует пустая строка или пустая строка следует за ней, то в результате конкатенации получаем ту же самую строку.

Пусть $a_1a_2\ldots a_n$, $b_1b_2\ldots b_m$ и $c_1c_2\ldots c_p \in A^*$. Тогда

$$\begin{aligned}(a_1a_2\ldots a_n \circ b_1b_2\ldots b_m) \circ c_1c_2\ldots c_p &= (a_1a_2\ldots a_n \circ b_1b_2\ldots b_m) \circ c_1c_2\ldots c_p = \\ &= a_1a_2\ldots a_nb_1b_2\ldots b_m \circ c_1c_2\ldots c_p = \\ &= a_1a_2\ldots a_nb_1b_2\ldots b_mc_1c_2\ldots c_p = \\ &= a_1a_2\ldots a_n \circ b_1b_2\ldots b_mc_1c_2\ldots c_p = \\ &= a_1a_2\ldots a_n \circ (b_1b_2\ldots b_m \circ c_1c_2\ldots c_p),\end{aligned}$$

и бинарная операция конкатенации ассоциативна на A^*, так что множество A^* вместе с операцией конкатенации представляет собой полугруппу. Далее, поскольку λ — это единица, то A^* — моноид. Заметим, что этот моноид с очевидностью некоммутативен. Если $\overline{A} \subseteq A$, то $\overline{A}^*$ — подполугруппа полугруппы A^*. Полугруппа A^*называется ***свободной полугруппой*** над алфавитом A. □

ПРИМЕР 9.17. Пусть $Z_n = \{[0],[1],[2],\ldots,[n-1]\}$ — множество классов вычетов по модулю n (см. раздел 3.6). Согласно теореме 3.54 $(Z_n,+)$ и $(Z_n,\cdot)$ вполне определены. Предоставляем читателю доказать, что $(Z_n,+)$ и $(Z_n,\cdot)$ являются полугруппами. □

Пусть $(S,\cdot)$ — полугруппа, и $a \in S$. Определим a^n рекурсивно, полагая $a^0 = 1$ и $a^n = a \cdot a^{n-1}$. Отсюда непосредственно следует, что $a^1 = a$. Предлагаем читателю, используя метод индукции, показать, что $a^k \cdot a^m = a^{k+m}$ для всех целых чисел $k, m \geq 0$. Множество

$$\langle a\rangle = \{a^n : n \geq 0\} = \{1, a, a^2, a^3, \ldots\}$$

является подполугруппой полугруппы S.

ОПРЕДЕЛЕНИЕ 9.18. Полугруппа $\langle a\rangle$ называется циклической полугруппой. Точнее, она называется циклической полугруппой, порожденной a.

Доказательство следующей теоремы оставляем читателю.

ТЕОРЕМА 9.19. Пусть $(S, \cdot)$ — полугруппа и $a_1, a_2, a_3, \ldots, a_k \in S$. Пусть $A = \{a_1, a_2, a_3, \ldots, a_k\}$ и $A^* = \langle a_1, a_2, a_3, \ldots, a_k \rangle$ — множество, состоящее из всех конечных произведений a_1, a_2, a_3, a_k. Тогда A^* — полугруппа. Более того, A^* — наименьшая подполугруппа полугруппы S, содержащая A.

ОПРЕДЕЛЕНИЕ 9.20. Полугруппа A^* называется полугруппой, порожденной множеством A. Если для каждого собственного подмножества B множества A имеем $B^* \neq A^*$, то A называется ***минимальным порождающим множеством*** для полугруппы A^*.

ОПРЕДЕЛЕНИЕ 9.21. Пусть $(S, \cdot)$ и $(T, \circ)$ — полугруппы и $f : S \to T$ — такая функция, что $f(s \cdot s') = f(s) \circ f(s')$. Функция f называется ***гомоморфизмом*** из S в T.

ТЕОРЕМА 9.22. Пусть $(S, \cdot)$ и $(T, \circ)$ — полугруппы и $f : S \to T$ — гомоморфизм из S в T. Если $\bar{S}$ — подполугруппа полугруппы S, то $f(\bar{S})$ — подполугруппа полугруппы T.

ДОКАЗАТЕЛЬСТВО. Пусть $t, t' \in f(\bar{S})$. Существуют такие $s, s' \in \bar{S}$, что $f(s) = t$ и $f(s') = t'$. По определению полугруппы $s \cdot s' \in \bar{S}$. По определению гомоморфизма $f(s \cdot s') = f(s) \circ f(s') = t \cdot t'$, т.е. $t \cdot t' \in f(\bar{S})$ и $f(\bar{S})$ — полугруппа. ■

Доказательство приведенной ниже теоремы оставляем читателю.

ТЕОРЕМА 9.23. Пусть $(S, \cdot)$ и $(T, \circ)$ — полугруппы и $f : S \to T$ — гомоморфизм из S в T. Если $\bar{T}$ — подполугруппа полугруппы T, то $f^{-1}(\bar{T})$ — подполугруппа полугруппы S.

ОПРЕДЕЛЕНИЕ 9.24. Пусть $(S, \cdot)$ — полугруппа и R — отношение эквивалентности на S. Если R обладает свойством, что если $s_1 R s_2$ и $s_3 R s_4$, то выполняется $s_1 s_3 R s_2 s_4$ для всех $s_1, s_3, s_2, s_4 \in S$, тогда R называется ***отношением конгруэнтности***.

ТЕОРЕМА 9.25. Пусть $(S, \cdot)$ и $(T, \circ)$ — полугруппы и $f : S \to T$ — гомоморфизм из S в T. Определим отношение R на множестве S таким образом: sRs', если $f(s) = f(s')$. Тогда отношение R — отношение конгруэнтности.

ДОКАЗАТЕЛЬСТВО. Для всех $s \in S$ $f(s) = f(s)$, поэтому R — рефлексивно. Если $f(s) = f(s')$, то $f(s') = f(s)$, и R — симметрично. Если $f(s) = f(s')$ и $f(s') = f(s'')$, то $f(s) = f(s'')$, поэтому R транзитивно. Допустим, $s_1 R s_2$ и $s_3 R s_4$. Следовательно, $f(s_1) = f(s_2)$ и $f(s_3) = f(s_4)$. Учитывая, что f — гомоморфизм, получаем

$$f(s_1 \cdot s_3) = f(s_1) \circ f(s_3)$$

и

$$f(s_1 \cdot s_3) = f(s_1) \circ f(s_3) = f(s_2) \circ f(s_4) = f(s_2 \cdot s_4),$$

поэтому $s_1 \cdot s_3 R s_2 \cdot s_4$. ■

ОПРЕДЕЛЕНИЕ 9.26. Коммутативная полугруппа $(S, *)$ называется ***полурешеткой***, если $a * a = a$ для всех $a \in S$. Элемент a полугруппы называется ***идемпотентом***, если $a * a = a$. Таким образом, полурешетка — это коммутативная полугруппа, в которой каждый элемент является идемпотентом. Если $(S, *)$ — полурешетка и $\overline{S} \subseteq S$, то $\overline{S}$ будет подполурешеткой полурешетки $(S, *)$, если $*$ — бинарная операция на $\overline{S}$. Эквивалентно, $(\overline{S}, *)$ — подполурешетка полурешетки $(S, *)$, если $\overline{S} \subseteq S$, и для всех $a, b \in \overline{S}$ имеем $a * b \in \overline{S}$.

Из теоремы 9.10 следует, что верхние и нижние полурешетки являются полурешетками. Более того, каждую полурешетку можно рассматривать как верхнюю или как нижнюю полурешетку.

ТЕОРЕМА 9.27. Пусть S — полурешетка. Определим на S отношение $\leq$ таким образом: $a \leq b$, если $a * b = b$ для $a, b \in S$. Тогда $(S, \leq)$ — это ЧУ-множество, и $a * b$ — наименьшая верхняя грань для a и b. Следовательно, $(S, *)$ — верхняя полурешетка. Аналогично, $(S, *)$ можно рассматривать как нижнюю полурешетку.

ДОКАЗАТЕЛЬСТВО. Поскольку $a * a = a$ для всех $a \in S$, $a \leq a$, поэтому $(S, \leq)$ рефлексивна. Если $a \leq b$ и $b \leq a$, то $a * b = b$ и $a * b = b * a = a$. Поэтому $b = a$ и $(S, \leq)$ антисимметрична. Допустим, что $a \leq b$ и $b \leq c$. Имеем $a * b = b$ и $b * c = c$. Получаем

$$\begin{aligned} a * c &= a * (b * c) = \\ &= (a * b) * c = \\ &= b * c = \\ &= c \end{aligned}$$

и $a \leq c$. Следовательно, $(S, \leq)$ транзитивна и является ЧУ-множеством.

Теперь покажем, что $a * b$ — наименьшая верхняя грань для a и b. Поскольку

$$\begin{aligned} a * (a * b) &= (a * a) * b = \\ &= a * b, \end{aligned}$$

то $a \leq a * b$. Точно так же,

$$\begin{aligned} b * (a * b) &= b * (b * a) = \\ &= (b * b) * a = \\ &= b * a = \\ &= a * b, \end{aligned}$$

так что $b \leq a * b$. Поэтому $a * b$ — верхняя грань для a и b. Предположим, что c — тоже верхняя грань для a и b, тогда $a * c = c$ и $b * c = c$. Следовательно,

$$\begin{aligned} (a * b) * c &= a * (b * c) = \\ &= a * c = \\ &= c \end{aligned}$$

и $a * b \leq c$. Поэтому $a * b$ — наименьшая верхняя грань для a и b, и $(S, \leq)$ — верхняя полурешетка. ■

■ УПРАЖНЕНИЯ

1. Пусть $(S, \cup)$ — полурешетка всех подмножеств множества $\{a, b, c\}$ с операцией объединения. Найдите три разные подполурешетки полурешетки $(S, \cup)$.
2. Пусть $(S, \cdot)$ — полурешетка на множестве всех матриц, элементы которых — вещественные числа, а операция — матричное умножение. Найдите пять подполугрупп полурешетки $(S, \cdot)$.
3. Докажите, что множество положительных целых чисел вида $4k + 1$, где k — неотрицательное целое число, образует полугруппу с операцией умножения.
4. Найдите пять подполугрупп множества положительных рациональных чисел с операцией сложения.
5. Докажите, что $(Z_n, +)$ и $(Z_n, \cdot)$ являются полугруппами для любого целого числа $n \geq 1$.
6. Докажите по методу индукции, что если S — полугруппа и $a \in S$, то $a^k \cdot a^m = a^{k+m}$ для всех целых чисел $m, n \geq 0$.
7. Докажите теорему 9.19. Пусть $(S, \cdot)$ — полугруппа и $a_1, a_2, a_3, \ldots, a_k \in S$. Пусть $A = \{a_1, a_2, a_3, \ldots, a_k\}$ и $A^* = \langle a_1, a_2, a_3, \ldots, a_k \rangle$ — множество всех конечных произведений a_1, a_2, a_3, a_k. Тогда A^* — полугруппа. Более того, A^* — наименьшая подполугруппа полугруппы S, содержащая A.
8. Докажите теорему 9.23. Пусть $(S, \cdot)$ и $(T, \circ)$ — полугруппы, и $f : S \to T$ — гомоморфизм из S в T. Если $\bar{T}$ — подполугруппа полугруппы T, то $f^{-1}(\bar{T})$ — подполугруппа полугруппы S.
9. Пусть $(S, \circ)$ — полугруппа всех функций на множестве A с операцией композиции.
   а) Покажите, что множество взаимно однозначных соответствий образует подполугруппу полугруппы S.
   б) Если $f \in S$ — идемпотент, то $f \circ f = f$. Такая функция называется ***ретракционным отображением***. Образ f называется ***ретрактором***. Покажите, что ретракционное отображение является взаимно однозначным отображением на свой образ.
   в) Покажите, что функции $f(x) = |x|$, $g(x) = \lfloor x \rfloor$ и $h(x) = \lceil x \rceil$, заданные на множестве действительных чисел, являются ретракционными отображениями. Что представляют собой соответствующие ретракторы?
   г) Коммутируют ли функции из пункта (в)? Иначе, верно ли, что $f \circ g = g \circ f$?
   д) Образуют ли функции из пункта (в) полурешетку?
10. Пусть S — множество действительных чисел и $*$ — бинарный оператор, определенный соотношением $a * b = \max(a, b)$. Покажите, что $(S, *)$ — полугруппа. Является ли $(S, *)$ полурешеткой?
11. Пусть $(S, *)$ — множество матриц размерности $n \times n$ с операцией умножения в качестве бинарной операции. Известно, что для $(n \times n)$-матриц A и B выполняется $\det(A) \cdot (\det B) = \det(AB)$. Используя этот факт, покажите, что матрицы, у которых $\det \neq 0$, образуют моноид. Что можно сказать о матрицах, у которых определитель равен 0?
12. Приведите пример бинарной операции $*$ на множестве вещественных чисел, относительно которой они не образуют полугруппу.

13. Покажите, что единица в моноиде единственна в том смысле, что существует только один элемент, например, I такой, что $a * I = I * a = a$ для всех a в моноиде.

14. Пусть S — множество. Определим операцию $*$ на S таким образом: $a * b = b$ для всех $a, b \in S$. Докажите, что $(S, *)$ — полугруппа, в которой каждый элемент является идемпотентом. Почему $(S, *)$ не является полурешеткой?

9.3. РЕШЕТКИ

В предыдущем разделе были рассмотрены множества с одной бинарной операцией. В данном разделе рассматриваются множества с двумя бинарными операциями.

ОПРЕДЕЛЕНИЕ 9.28. ЧУ-множество $(S, \leq)$, которое является и нижней, и верхней полурешеткой, называется ***решеткой***.

В решетке $(S, \leq)$ будем обозначать наименьшую верхнюю грань $a, b \in S$ через $a \vee b$ (называемое ***соединением*** a и b), а наибольшую нижнюю грань $a, b \in S$ через $a \wedge b$ (называемое ***пересечением*** a и b). Решетка будет обозначаться через $(S, \vee, \wedge)$, чтобы подчеркнуть используемые бинарные операции. На основе свойств наименьшей верхней грани и наибольшей нижней грани докажем следующие свойства решетки, которые можно использовать в качестве альтернативного определения.

ТЕОРЕМА 9.29. Пусть $(S, \vee, \wedge)$ — решетка, тогда следующие свойства имеют место для всех $a, b, c \in S$.

а) Коммутативность

$$a \wedge b = b \wedge a;$$
$$a \vee b = b \vee a.$$

б) Ассоциативность

$$(a \wedge b) \wedge c = a \wedge (b \wedge c);$$
$$(a \vee b) \vee c = a \vee (b \vee c).$$

в) Поглощение

$$a \wedge (a \vee b) = a;$$
$$a \vee (a \wedge b) = a.$$

ДОКАЗАТЕЛЬСТВО. Для доказательства пунктов (а) и (б) см. теорему 9.10 и упражнения в конце раздела. Покажем, что $a \wedge (a \vee b) = a$:

$a \leq a$	по определению частичного порядка ;
$a \leq a \vee b$	по определению наименьшей верхней грани.

Поэтому a есть нижняя грань для $\{a, a \vee b\}$. Допустим, что c — нижняя грань для $\{a, a \vee b\}$. Тогда $c \leq a$, так что a — наибольшая нижняя грань для $\{a, a \vee b\}$ и, следовательно, $a = a \wedge (a \vee b)$. Доказательство того, что $a \vee (a \wedge b) = a$ оставляем читателю. ■

ПРИМЕР 9.30. Пусть S — булеан (множество всех подмножеств) множества A с частичным порядком $U \leq V$, если $U \subseteq V$ для всех $U, V \in S$. Тогда нвг$\{U, V\} = U \cup V$ и ннг$\{U, V\} = U \cap V$, поэтому $(S, \cup, \cap)$ является решеткой. □

ПРИМЕР 9.31. Пусть S — множество положительных целых чисел с частичным порядком: $n \leq m$, если m делит n. Тогда нвг$\{m, n\}$ = НОК(m, n) и ннг$\{m, n\}$ = НОД(m, n) для всех $m, n \in S$. Поэтому S с бинарными операциями НОК и НОД является решеткой. □

ПРИМЕР 9.32. Рациональные, действительные и целые числа с обычным частичным порядком образуют решетку, где нвг$\{m, n\} = \max(m, n)$ и ннг$\{m, n\} = \min(m, n)$. □

ОПРЕДЕЛЕНИЕ 9.33. Непустое подмножество $\overline{S}$ решетки $(S, \vee, \wedge)$ называется ***подрешеткой*** решетки S, если для всех $a, b \in \overline{S}$ $a \wedge b \in \overline{S}$ и $a \vee b \in \overline{S}$.

Следовательно, в примере 9.30, если B — подмножество множества A и $\overline{S}$ — булеан множества B, то $(\overline{S}, \cup, \cap)$ будет подрешеткой решетки $(S, \cup, \cap)$. В примере 9.31, если $\overline{S}$ — множество всех положительных делителей фиксированного целого числа n, то $\overline{S}$ с бинарными операциями НОК и НОД будет подрешеткой решетки S. В примере 9.32 целые числа образуют подрешетку рациональных чисел, а рациональные числа, в свою очередь, образуют подрешетку действительных чисел.

ОПРЕДЕЛЕНИЕ 9.34. Решетка $(S, \vee, \wedge)$ называется ***ограниченной***, если множество S, как ЧУ-множество, имеет наименьшую верхнюю грань и наибольшую нижнюю грань. Наименьшая верхняя грань обозначается через 1, а наибольшая нижняя грань — через 0. Эквивалентно, решетка ограничена, если существуют элементы $0, 1 \in S$ такие, что $0 \wedge a = 0$ и $1 \vee a = 1$ для всех $a \in S$.

Поскольку 0 — наименьший элемент множества, и 1 — наибольший элемент множества, получаем следующий результат.

ТЕОРЕМА 9.35. В ограниченной решетке $1 \wedge a = 1$ и $0 \vee a = a$ для всех a из решетки.

ПРИМЕР 9.36. Решетка в примере 9.30 ограничена, A — максимальный элемент решетки и $\varnothing$ — минимальный элемент. □

ПРИМЕР 9.37. В примере 9.31, если $\overline{S}$ — полурешетка всех положительных делителей фиксированного целого числа n, то $\overline{S}$ ограничена с максимальным элементом n и минимальным элементом 1. □

В главе 2 отмечалось, что для множеств A, B и C выполняются соотношения $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ и $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. Эти соотношения были названы свойством дистрибутивности для множеств. Далее будет дано определение аналогичных свойств для решеток.

ОПРЕДЕЛЕНИЕ 9.38. Решетка $(S, \vee, \wedge)$ называется ***дистрибутивной решеткой***, если для всех $a, b, c \in S$ имеем

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c);$$
$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

В качестве упражнения предлагаем читателю доказать, что решетки, изображенные на рис. 9.6, не дистрибутивны.

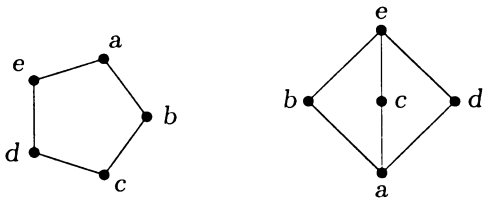


Рис. 9.6

Хотя подобное утверждение выходит за рамки данной книги, однако можно показать, что решетка дистрибутивна тогда и только тогда, когда она не содержит ни одну из этих двух решеток в качестве подрешеток. Следовательно, решетки на рис. 9.7 не являются дистрибутивными.

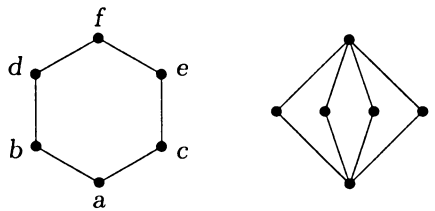


Рис. 9.7

Булева алгебра, о которой шла речь в разделе 2.3, является специальным случаем ограниченной дистрибутивной решетки. Теперь перейдем к системе обозначений булевой алгебры из раздела 2.3, заменив бинарную операцию $\vee$ на $+$, а бинарную операцию $\wedge$ на $\cdot$. Далее решетка будет обозначаться как $(S, +, \cdot)$. Для удобства приведем определение булевой алгебры еще раз. Более подробно данный вопрос рассмотрен в разделе 2.4.

ОПРЕДЕЛЕНИЕ 9.39. ***Булева алгебра*** — это ограниченная дистрибутивная решетка $(S, +, \cdot)$, в которой имеется унарная операция $'$ на множестве S такая, что $x + x' = 1$ и $x \cdot x' = 0$. Для $x \in S$ x' называется ***дополнением*** x. Следовательно, булева алгебра, обозначаемая как $(S, +, \cdot, ', 1, 0)$, обладает следующими свойствами, которые выполняются для всех $x, y, z \in S$.

1. Законы ассоциативности

$$x + (y + z) = (x + y) + z;$$
$$x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$

2. Законы коммутативности

$$x + y = y + x;$$
$$x \cdot y = y \cdot x.$$

3. Законы дистрибутивности

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z);$$
$$x + (y \cdot z) = (x + y) \cdot (x + z).$$

4. Законы дополнения

$$x + x' = 1;$$
$$x \cdot x' = 0.$$

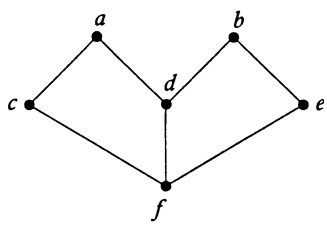
5. Законы тождества

$$x + 0 = x;$$
$$x \cdot 1 = x.$$

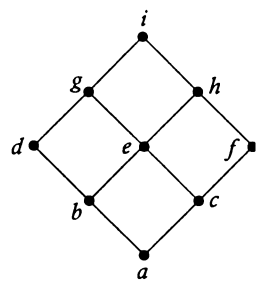
■ УПРАЖНЕНИЯ

1. Какое изображение из приведенных ниже представляет собой решетку?

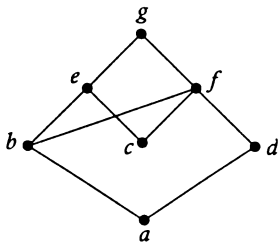
а)



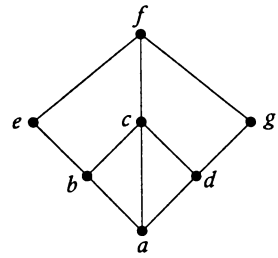
б)



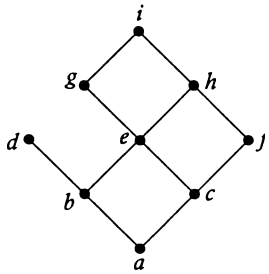
в)



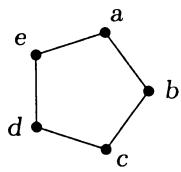
г)

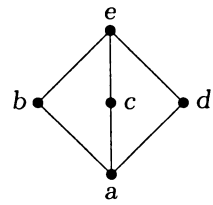


д)



2. Покажите, что приведенные ниже решетки не дистрибутивны.





3. Покажите, что решетка, изображенная на рис. 9.8, не дистрибутивна.
4. Покажите, что решетка, изображенная на рис. 9.9, не дистрибутивна.

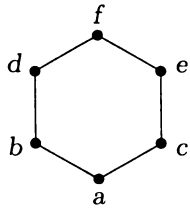


Рис. 9.8

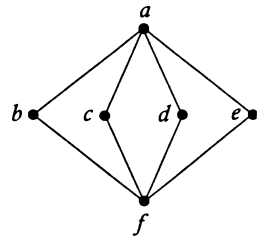
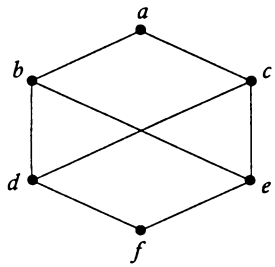


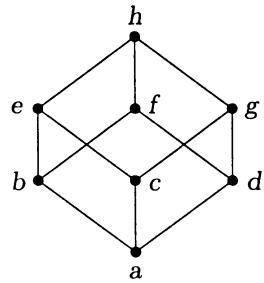
Рис. 9.9

5. Какие из приведенных ниже решеток являются дистрибутивными?

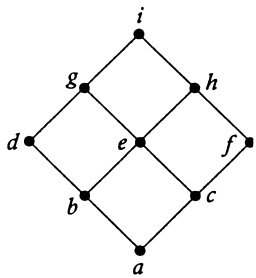
а)



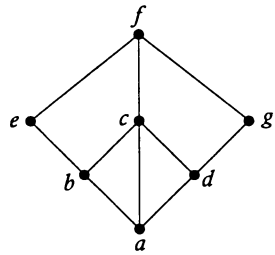
б)



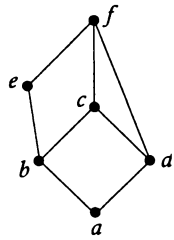
в)



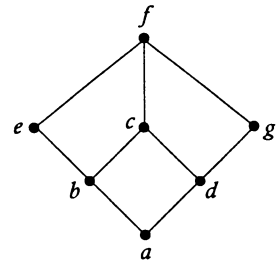
г)



д)



е)



6. Докажите, что для всех a, b, принадлежащих решетке, $a \vee (a \wedge b) = a$.

7. ***Атомом*** в булевой алгебре называется ненулевой элемент x, обладающий таким свойством, что для каждого y из алгебры либо $xy = x$, либо $xy = 0$.

а) Пусть S — булеан (множество всех подмножеств) множества $\{a, b, c\}$. Пусть $(S, \cup, \cap, ', \varnothing)$ — булева алгебра на S с обычным определением пересечения и объединения. Что является атомами данного множества?

б) Пусть S — множество всех высказываний, полученных из p, q и r с использованием обычных пропозициональных связок $\wedge$, $\vee$ и $\sim$. Покажите, что в булевой алгебре $(S, \vee, \wedge, \sim, T, F)$ атомы — это элементарные конъюнкции, определенные в разделе 1.5.

в) Найдите атомы булевой алгебры, изображенной на рис. 9.10.

г) Найдите атомы булевой алгебры, изображенной на рис. 9.11.

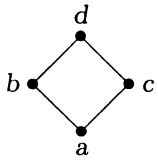


Рис. 9.10

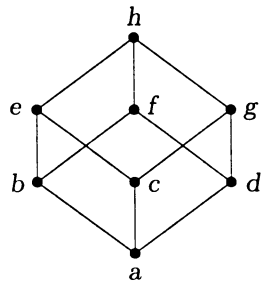


Рис. 9.11

9.4. ГРУППЫ

Ранее отмечалось, что множество неотрицательных целых чисел с операцией $+$ образуют моноид. Однако, если включить в рассмотрение множество всех целых чисел с той же операцией $+$, то опять получим моноид, но обладающий свойством, которое отсутствовало в первоначальном моноиде, а именно: для каждого целого числа a существует целое число $-a$ такое, что $a + (-a) = (-a) + a = 0$, единице моноида. Аналогичным образом, если рассмотреть множество положительных рациональных чисел с операцией умножения, то снова получим моноид, и для каждого рационального числа $\frac{p}{q}$ существует рациональное число $\frac{q}{p}$ такое, что $\frac{p}{q} \cdot \frac{q}{p} = 1$, единице моноида. Рассмотрим свойства, фигурирующие при решении уравнения $2x + 3 = 7$. Сначала прибавляем -3, — элемент, обратный к 3 относительно сложения, к обеим частям уравнения, и получаем $(2x+3)+(-3) = 7+(-3)$. После использования ассоциативности получаем уравнение $2x+(3-3) = 4$. Затем применяем свойство обратного элемента и получаем $2x + 0 = 4$. Используя свойство нейтральности 0 относительно сложения, имеем $2x = 4$. Затем для получения окончательного решения применяем свойства нейтральности и ассоциативности.

Группы — это нечто большее, чем только абстракция множества действительных чисел с операцией сложения или множества положительных целых чисел с операцией умножения. Далее показано, что множество перестановок n элементов образует группу. Группы имеют важное применение в физике и во многих других областях. Они используются в теории чисел, теории кодов, в комбинаторике.

Любой моноид $(S, \circ)$, обладающий свойством, что для каждого элемента моноида s существует элемент моноида s^{-1} такой, что $s \circ s^{-1} = s^{-1} \circ s = I$, единице моноида, называется ***группой***. Более формально понятие группы может быть определено следующим образом.

ОПРЕДЕЛЕНИЕ 9.40. ***Группа*** представляет собой множество G вместе с бинарной операцией (или произведением) $\circ$ на $G \times G$, обладающей следующими свойствами:

1. $a \circ (b \circ c) = (a \circ b) \circ c$ для всех a, b и c из G (т.е. операция $\circ$ на S ассоциативна).
2. В G существует элемент 1, называемый ***единицей***, который обладает свойством $a \circ 1 = 1 \circ a = a$ для всех a из G.
3. Для каждого элемента a из множества G существует элемент a^{-1} из множества G, который называется ***обратным элементом*** для a и такой, что $a \circ a^{-1} = a^{-1} \circ a = 1$.

Если группа G обладает таким свойством, что $a \circ b = b \circ a$ для всех a, b из G, то она называется ***коммутативной***, или ***абелевой*** группой.

Снова наблюдается ситуация, что поскольку $\circ$ — бинарная операция на G, то отсюда вытекает, что если $g, g' \in G$, то $g \circ g' \in G$. Это свойство называется ***замкнутостью***.

ПРИМЕР 9.41. Примеры групп включают следующее:

1. Четные целые числа с операцией сложения.
2. Множество $(n \times m)$-матриц с действительными элементами и бинарной операцией сложения.
3. Классы вычетов по модулю некоторого положительного целого числа n, образующие Z_n с бинарной операцией сложения. □

ОПРЕДЕЛЕНИЕ 9.42. Если G — группа с n элементами, то n называется ***порядком*** группы G.

Пусть $S = \{i, -i, -1, 1\}$, где $i = \sqrt{-1}$. Легко убедиться, что это множество вместе с операцией умножения образует группу порядка 4.

Заметим, что всякая группа является полугруппой. Обратное, вообще говоря, неверно.

ТЕОРЕМА 9.43. Единица группы G единственна.

ДОКАЗАТЕЛЬСТВО. Допустим, что 1 и e — единицы группы G. Тогда $1 = 1 \circ e = e$ по определению единицы. ■

ТЕОРЕМА 9.44. В каждой группе обратный элемент для каждого элемента единственный.

ДОКАЗАТЕЛЬСТВО. Пусть a — элемент группы G, и оба элемента b и c — обратные к элементу a, тогда

$$b = b \circ 1 = b \circ (a \circ c) = (b \circ a) \circ c = 1 \circ c = c.$$ ■

ТЕОРЕМА 9.45. Для каждого элемента a группы G имеем $(a^{-1})^{-1} = a$.

ДОКАЗАТЕЛЬСТВО. Поскольку $a^{-1} \circ a = a \circ a^{-1} = 1$, то a удовлетворяет определению обратного элемента для a^{-1}. Поскольку обратный элемент единственный, то a и является элементом, обратным к a^{-1}. ■

ТЕОРЕМА 9.46. Для элементов a и b группы G имеем $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.

ДОКАЗАТЕЛЬСТВО. $(a \circ b) \circ (b^{-1} \circ a^{-1}) = (a \circ (b \circ b^{-1})) \circ a^{-1} = (a \circ 1) \circ a^{-1} = a \circ a^{-1} = 1$. Аналогично, $(b^{-1} \circ a^{-1}) \circ (a \circ b) = 1$, и по определению обратного элемента, и вследствие его единственности имеем $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$. ■

Если a — элемент группы G, то обозначим $(a^{-1})^k = a^{-k}$. Пусть a^0 обозначает единицу 1.

Доказательство приведенных ниже теорем оставляем читателю.

ТЕОРЕМА 9.47. Пусть G — группа и a — элемент группы G.

а) $a^n \circ a^{-n} = 1$ для всех положительных целых чисел n.
б) $a^{(m+n)} = a^m \circ a^n$ для всех целых чисел n и m.
в) $(a^m)^n = a^{mn}$ для всех целых чисел m и n.
г) $(a^{-n})^{-1} = a^n$ для всех целых чисел n.

ТЕОРЕМА 9.48. Если a — элемент группы $(G, \circ)$ и $a \circ a = a$, то $a = 1$, единице группы G.

ЛЕММА 9.49. Если G — конечная группа и a — элемент G, то $a^s = 1$ для некоторого положительного целого числа s.

ДОКАЗАТЕЛЬСТВО. Если G — конечная группа и a — элемент G, то известно, что $a^j = a^k$ для некоторых положительных целых чисел j и k, поскольку только конечное число степеней может быть различно. Пусть $j < k$, тогда $a^{-j}a^k = a^{-j}a^j = 1$ и $a^{k-j} = 1$. Пусть $s = k - j$. ■

Пусть p — наименьшее целое число, которое больше 1 и такое, что $a^p = 1$. Элементы $1, a, a^2, \ldots, a^{p-1}$ различны, т.к. если $a^r = a^s$ для $0 \leq r$, $s < p$ и $r > s$, то $a^{r-s} = 1$, что противоречит тому факту, что p — наименьшее положительное целое число такое, что $a^p = 1$. Также, если $n > p$, то $a^n = a^r$ для некоторого $0 \leq r < p$, поскольку если $n = pq + r$ для некоторого $0 \leq r < p$, то

$$a^n = a^{pq+r} = a^{pq} \circ a^r = (a^p)^q \circ a^r = 1^q \circ a^r = 1 \circ a^r = a^r.$$

ТЕОРЕМА 9.50. Пусть G — группа и a — элемент G такой, что $a^s = 1$ для некоторого s. Если p — наименьшее положительное целое число такое, что $a^p = 1$, то $p \mid s$. Целое число p называется ***порядком*** a.

ДОКАЗАТЕЛЬСТВО. Пусть p — наименьшее положительное целое число такое, что $a^p = 1$. Как это было сделано выше, положим $s = pq + r$ для некоторого $0 \leq r < p$, тогда $a^s = a^{pq+r} = a^{pq}a^r = (a^p)^q a^r = (1)^q a^r = a^r$. Следовательно, $a^s = 1$ тогда и только тогда, когда $r = 0$, потому что p — наименьшее положительное целое число, обладающее свойством $a^p = 1$. ■

Снова рассмотрим множество $S = \{i, -i, -1, 1\}$ с операцией умножения. Каков порядок i? Каков порядок -1?

ОПРЕДЕЛЕНИЕ 9.51. Подмножество H группы G является ***подгруппой*** G, если H с той же самой операцией, что и G, также является группой.

ПРИМЕР 9.52. Пусть $(Z,+)$ — группа целых чисел с операцией сложения. Для положительного целого числа m пусть $(mZ,+)$ — группа всех чисел, кратных m. Понятно, что это группа, поскольку если сложить два числа, кратных m, то снова получим число, кратное m. Элемент, обратный mi, это $-mi = m(-i)$. □

ПРИМЕР 9.53. Пусть $(R,+)$ — группа действительных чисел с операцией сложения. Тогда группа $(Q,+)$ — рациональные числа со сложением, является подгруппой группы $(R,+)$. □

ПРИМЕР 9.54. Пусть $(R^+,\cdot)$ — группа положительных действительных чисел с умножением. Группа $(Q^+,\cdot)$ положительных рациональных чисел с умножением является подгруппой группы $(R^+,\cdot)$. □

ПРИМЕР 9.55. Легко проверить, что $(S,+)$, где $S = \{[0],[2],[4]\}$ — подгруппа группы $(Z_6,+)$. □

Чтобы определить, является ли подмножество группы подгруппой, не требуется проверять ассоциативность, поскольку она наследуется от исходной группы. Необходимо проверить наличие обратного элемента и то, что произведение двух элементов из множества принадлежит множеству. Необходимо убедиться в том, что единица содержится во множестве, но это, как показывает доказательство приведенной ниже теоремы, следует из свойства замкнутости и свойства обратного элемента. Теорема дает альтернативный метод проверки того, что некоторое подмножество группы является подгруппой.

ТЕОРЕМА 9.56. Непустое подмножество H группы $(G,\cdot)$ будет подгруппой тогда и только тогда, когда для всех $h_1, h_2 \in H$ имеем $h_1 \cdot h_2^{-1} \in H$.

ДОКАЗАТЕЛЬСТВО. Сначала предположим, что H — группа и $h_1, h_2 \in H$. Тогда h_2^{-1} принадлежит группе, поскольку элемент, обратный элементу группы, тоже принадлежит группе. Следовательно, по замыканию, $h_1 h_2^{-1}$ принадлежит H. Обратно предположим, что для всех $h_1, h_2 \in H$ имеем $h_1 \cdot h_2^{-1} \in H$. Если $h_1 \in H$, то $h_1 h_1^{-1} \in H$, поэтому единица $I \in H$. Следовательно, для всех $h_2 \in H$ имеет место $I \cdot h_2^{-1} \in H$ и $h_2^{-1} \in H$. Наконец, если $h_1, h_2 \in H$, то $h_1 \cdot (h_2^{-1})^{-1} = h_1 \cdot h_2 \in H$. Поэтому H — группа. ■

Доказательство приведенной ниже теоремы оставляем читателю.

ТЕОРЕМА 9.57. Если g — элемент группы G такой, что $g^n = 1$ для некоторого n, и p — наименьшее положительное целое число такое, что $g^p = 1$, тогда множество $\{g, g^2, \ldots, g^p\}$ является подгруппой группы G.

ОПРЕДЕЛЕНИЕ 9.58. Множество $\{g, g^2, \ldots, g^p\}$ называется ***циклической группой*** или, точнее, циклической группой, порожденной g. Она обозначается через $\langle g \rangle$.

Еще раз рассмотрим множество $S = \{i, -i, -1, 1\}$ с умножением. Это циклическая группа. Что может быть использовано в качестве g? Найдите две подгруппы группы S. Являются ли они циклическими?

Доказательство приведенной ниже теоремы оставляем читателю.

ТЕОРЕМА 9.59. Пусть $(G, \cdot)$ — группа и $a_1, a_2, a_3, \ldots, a_k \in G$. Пусть $A = \{a_1, a_2, a_3, \ldots, a_k\}$ и $A^* = \langle a_1, a_2, a_3, \ldots, a_k \rangle$ — множество всех конечных произведений элементов $a_1, a_2, a_3, \ldots, a_k$ и обратных к ним. Тогда A^* — группа. Более того, A^* — наименьшая подгруппа группы G, содержащая A.

ОПРЕДЕЛЕНИЕ 9.60. Подгруппа A^* называется группой, ***порожденной*** множеством A. Если для каждого собственного подмножества B множества A имеем $B^* \neq A^*$, тогда A называется ***порождающим множеством*** для A^*. Если множество A порождает группу G и никакое собственное подмножество множества A не порождает G, тогда A называется ***минимальным порождающим множеством*** для группы G.

В циклической группе порождающее множество может быть образовано одним элементом. Возможно ли иметь минимальное порождающее множество более чем из одного элемента?

ОПРЕДЕЛЕНИЕ 9.61. Для подгруппы H группы G и произвольного a из G $a \circ H = \{x : x = a \circ h$ для некоторого h из $H\}$ называется ***левым смежным классом*** подгруппы H группы G.

ПРИМЕР 9.62. Ранее было показано, что множество Z целых чисел с бинарной операцией сложения, $+$, является группой. Пусть

$$H = \{n : n \in Z \text{ и } n = 5k \text{ для некоторого целого числа } k\} = \\ = \{\ldots, -10, -5, 0, 5, 10, 15, 20, \ldots\}.$$

Покажем, что H является подгруппой группы Z. Понятно, что $H \subseteq Z$. Множество H замкнуто относительно сложения, поскольку если $s \in H$ и $t \in H$, то $s = 5i$ и $t = 5j$ для соответствующих целых чисел i и j. Таким образом, $s + t = 5i + 5j = 5(i + j) = 5k$ для $k = i + j$, так что $s + t$ принадлежит H. Единичный элемент H является единичным элементом G, а именно, 0. Каждый элемент H имеет обратный элемент относительно $+$ в H, потому что если $t = 5k \in H$, то $t^{-1} = 5(-k)$.

Теперь рассмотрим левые смежные классы подгруппы H, генерируемые различными элементами множества Z. Пусть $a \circ H = a + H = \{x : x = a + h$ для некоторого h $H\}$, тогда

$$0 + H = \{x : x = 0 + h \text{ для некоторого } h \text{ из } H\} = \\ = \{\ldots, -10, -5, 0, 5, 10, 15, 20, \ldots\};$$

$$1+H=\{x : x=1+h \text{ для некоторого } h \text{ из } H\}=$$
$$=\{\ldots,-9,-4,1,6,11,16,21,\ldots\};$$

$$2+H=\{x : x=2+h \text{ для некоторого } h \text{ из } H\}=$$
$$=\{\ldots,-8,-3,2,7,12,17,22,\ldots\};$$

$$3+H=\{x : x=3+h \text{ для некоторого } h \text{ из } H\}=$$
$$=\{\ldots,-7,-2,3,8,13,18,23,\ldots\};$$

$$4+H=\{x : x=4+h \text{ для некоторого } h \text{ из } H\}=$$
$$=\{\ldots,-6,-1,4,9,14,19,24,\ldots\}.$$

Таким образом, снова получен класс вычетов Z_5, поскольку $a+H=[a]$, и пять левых смежных классов подгруппы H образуют разбиение исходной группы Z. □

ЛЕММА 9.63. Для фиксированной подгруппы H группы G левые смежные классы подгруппы H группы G образуют разбиение группы G.

ДОКАЗАТЕЛЬСТВО. Каждый левый смежный класс непустой, поскольку для левого смежного класса $a\circ H$; элемент $a=a\circ 1$ находится в $a\circ H$. Предположим, что пересечение $a\circ H$ и $b\circ H$ непустое; пусть элемент c принадлежит пересечению. Следовательно, $c=a\circ h=b\circ h'$ для некоторых h и h' из H. Умножая обе части уравнения на h^{-1}, получаем $a\circ h\circ h^{-1}=b\circ h'\circ h^{-1}$; поэтому по определению обратного элемента имеем $a=b\circ(h'\circ h^{-1})$. Поскольку $h'\circ h^{-1}$ принадлежит H, то элемент a содержится в $b\circ H$. Следовательно, $a\circ h$ содержится в $b\circ H$ для всех h из H, поэтому $a\circ H\subseteq b\circ H$. Аналогично, $b\circ H\subseteq a\circ H$ и $b\circ H=a\circ H$. Следовательно, левые смежные классы образуют разбиение группы G. ■

ЛЕММА 9.64. Если G — конечная группа и H — подгруппа группы G, то все левые смежные классы подгруппы H группы G содержат одно и то же количество элементов, а именно, количество элементов, которое находится в подгруппе H.

ДОКАЗАТЕЛЬСТВО. Пусть $a\circ H$ — левый смежный класс подгруппы H из G. Определим $f : H\to a\circ H$ таким образом: $f(h)=a\circ h$. Читателю остается показать, что f представляет собой взаимно однозначное соответствие, или биекцию. ■

ТЕОРЕМА 9.65. (Лагранж) Если G — конечная группа и H — подгруппа группы G, то порядок H делит порядок G.

ДОКАЗАТЕЛЬСТВО. Если p — порядок H, q — количество левых смежных классов H в G и n — порядок G, то согласно двум предыдущим леммам $n=pq$. ■

ТЕОРЕМА 9.66. Если G — группа порядка n и g приналежит G, то $g^n=1$.

ДОКАЗАТЕЛЬСТВО. Пусть p — наименьшее положительное число такое, что $g^p=1$ и $H=\{g,g^2,\ldots,g^p\}$. Тогда H — подгруппа с p элементами и, следовательно, p делит n. Пусть $n=pq$, таким образом, $g^n=g^{pq}=(g^p)^q=1^q=1$. ■

■ УПРАЖНЕНИЯ

1. Составьте таблицу сложения и умножения для Z_4.
2. Составьте таблицу сложения и умножения для Z_7.
3. Вычислите $[1] \circ [2] \circ [3]$ в Z_4.
4. Вычислите $[1] \circ [2] \circ [3] \circ [4]$ в Z_7.
5. В Z_{10} найдите элемент, обратный $[7]$ относительно умножения.
6. В Z_{12} найдите все такие a, что $[a] \circ [6] = [0]$.
7. Пусть $Z_n = \{[0], [1], [2], \ldots, [n-1]\}$ для некоторого целого числа n. Пусть $S = \{[k] : k$ положительное и взаимно простое с $n\}$. Является ли S группой с операцией умножения?
8. При выполнении каких условий Z_n будет группой с операцией умножения?
9. Докажите, что числа, кратные 3, образуют подгруппу целых чисел с операцией сложения. Опишите соответствующие смежные классы этой подгруппы.
10. Докажите, что числа, кратные 3, образуют полугруппу целых чисел с операцией умножения. Почему они не образуют группу?
11. Докажите, что множество классов вычетов по модулю 6 образует группу относительно сложения. Будут ли они образовывать группу относительно умножения?
12. Докажите теорему 9.47. Пусть G — группа и a — элемент группы G.
    а) $a^n \circ a^{-n} = 1$ для всех положительных целых чисел n.
    б) $a^{(m+n)} = a^m \circ a^n$ для всех целых чисел n и m.
    в) $(a^m)^n = a^{mn}$ для всех целых чисел m и n.
    г) $(a^{-n})^{-1} = a^n$ для всех целых чисел n.
13. Докажите теорему 9.57. Если g — элемент группы G такой, что $g^n = 1$ для некоторого n, и p — наименьшее положительное целое число такое, что $g^p = 1$, тогда множество $\{g, g^2, \ldots, g^p\}$ является подгруппой группы G.
14. Докажите теорему 9.48. Если a — элемент группы $(G, \circ)$ и $a \circ a = a$, то $a = 1$, единице группы G.

9.5. ГРУППЫ И ГОМОМОРФИЗМЫ

Существуют специальные функции на множестве групп такие, что некоторые свойства области определения сохраняются в образе функции. Например, образы подгрупп из области определения являются подгруппами в области значений. Такие функции называются гомоморфизмами. Обратите внимание, что группы являются полугруппами, и это определение гомоморфизма пригодно также для полугрупп.

ОПРЕДЕЛЕНИЕ 9.67. Пусть $(G, \bullet)$ и $(H, *)$ — группы, где $\bullet$ и $*$ — операции на G и H соответственно. Пусть $f : G \to H$ — функция. Функция f называется ***гомоморфизмом***, если $f(g \bullet g') = f(g) * f(g')$ для всех g и g' из G. Гомоморфизм f называется ***мономорфизмом***, если функция f — инъекция, ***эпиморфизмом***, если функция f — сюръекция, и ***изоморфизмом***, если функция f — биекция.

ПРИМЕР 9.68. В примере 9.62, где $H = \{n : n \in Z \text{ и } n = 5k \text{ для некоторого целого числа } k\}$ является подгруппой группы Z, множество левых смежных классов подгруппы H имеет вид

$$W = \{0 + H, 1 + H, 2 + H, 3 + H, 4 + H\}.$$

Определим сложение левых смежных классов следующим образом:

$$(a + H) \oplus (b + H) = (a + b) + H,$$

где сложение $(a + b)$ есть сложение целых чисел a и b. Несложно показать, что определение операции $\oplus$ не зависит от представления $a + H$ и $b + H$.

Легко показать, что для группы $(Z_5, +)$ классов вычетов по модулю 5 функция $f : Z_5 \to W$, определенная соотношением

$$f([k]) = k + H,$$

является гомоморфизмом между группами $(Z_5, +)$ и $(W, \oplus)$ и также является изоморфизмом.

Рассмотрим две группы $(Z, +)$ и $(Z_5, +)$ и функцию $g : Z \to Z_5$, определенную таким образом:

$$g(a) = [a].$$

Функция g является гомоморфизмом, поскольку $g(a + b) = [a + b] = [a] + [b] = g(a) + g(b)$. Гомоморфизм g является эпиморфизмом, поскольку ясно, что для любого заданного элемента $[c]$ из группы Z_5 имеем $g(c) = [c]$. □

ПРИМЕР 9.69. В одной из задач, рассмотренных в разделе 5.8, показано, что $(n \times n)$-матрица имеет обратную тогда и только тогда, когда ее детерминант не равен 0. Хотя этот вопрос выходит за рамки данной книги, можно показать, что функция вычисления детерминанта представляет собой гомоморфизм из множества $(n \times n)$-матриц с действительными (или рациональными) элементами с операцией умножения на множество действительных (или рациональных) чисел с умножением. Следовательно, если определитель каждой из матриц не равен нулю, то определитель их произведения тоже не равен нулю. Отсюда следует, что произведение двух матриц, каждая из которых имеет обратную, также обратимо. Следовательно, множество таких $(n \times n)$-матриц образует группу с операцией умножения, которая называется группой несингулярных матриц. Если определитель матрицы M имеет ненулевое значение r, то чему равен определитель матрицы M^{-1}? □

Доказательство приведенных ниже теорем оставляем читателю.

ТЕОРЕМА 9.70. Пусть $f : G \to H$ — гомоморфизм из группы G в группу H и 1 — единица группы G. Тогда $f(1)$ — единица группы H.

ТЕОРЕМА 9.71. Пусть $f : G \to H$ — гомоморфизм из группы G в группу H и g' — элемент, обратный элементу g из G. Тогда $f(g')$ есть элемент, обратный элементу $f(g)$ из H.

ТЕОРЕМА 9.72. Если $f : G \to H$ — гомоморфизм из группы G в группу H и K — подгруппа группы H, то $f^{-1}(K)$ — подгруппа группы G.

ТЕОРЕМА 9.73. Если $f : G \to H$ — гомоморфизм из группы G в группу H и K — подгруппа G, то $f(K)$ — подгруппа группы H.

ОПРЕДЕЛЕНИЕ 9.74. Для подмножеств H и K группы $(G, \circ)$ пусть $H \circ K = \{h \circ k : h \in H, k \in K\}$. Если $H = \{h\}$, то $H \circ K$ обычно записывают как hK.

Доказательство приведенной ниже теоремы оставляем читателю.

ТЕОРЕМА 9.75. Если H, J и K — подмножества группы $(G, \circ)$, то $(H \circ J) \circ K = H \circ (J \circ K)$.

В примере 9.68 показано, что группа делится на классы и эти смежные классы можно "перемножать" (в том случае умножением групп было сложение целых чисел). В общем случае это не всегда возможно. Тем не менее, будет показано, что если подгруппа, которая используется для формирования смежных классов, обладает определенными свойствами, то имеется возможность перемножения смежных классов.

ОПРЕДЕЛЕНИЕ 9.76. Если H — подгруппа группы $(G, \circ)$, обладающая свойством $gHg^{-1} = H$ для всех $g \in G$, то такая группа H называется ***нормальной подгруппой***.

ОПРЕДЕЛЕНИЕ 9.77. Пусть $f : G \to H$ — гомоморфизм из группы G в группу H. ***Ядром*** гомоморфизма f называется множество $\{x : x \in G \text{ и } f(x) = 1\} = f^{-1}(\{1\})$, где 1 — единица группы H.

Доказательство приведенной ниже теоремы оставляем читателю.

ТЕОРЕМА 9.78. Ядро гомоморфизма $f : G \to H$ есть нормальная подгруппа группы G.

ТЕОРЕМА 9.79. Подгруппа H группы $(G, \circ)$ является нормальной подгруппой тогда и только тогда, когда $gH = Hg$ для всех $g \in G$.

ДОКАЗАТЕЛЬСТВО. Предположим, что подгруппа H нормальна в группе G, так что $gHg^{-1} = H$ для всех $g \in G$. Пусть $gh \in gH$. Поскольку $gHg^{-1} = H$, то $ghg^{-1} = h'$ для некоторого $h' \in H$. Следовательно, $gh = h'g$ и $gh \in Hg$, поэтому $gH \subseteq Hg$. Аналогично, $Hg \subseteq gH$. Доказать утверждение в обратную сторону предоставляется читателю. ■

ТЕОРЕМА 9.80. Если H — подгруппа группы $(G, \circ)$, то $H \circ H = H$.

ДОКАЗАТЕЛЬСТВО. Предположим, что H — группа. Тогда по определению замыкания $H \circ H \subseteq H$. Поскольку H — группа, то для $h \in H$ имеем $h = h \circ I$, где I — единица группы G, а значит, и группы H. Поэтому $h \in H \circ H$ и $H \subseteq H \circ H$. Таким образом, $H = H \circ H$. ■

Обратная теорема неверна. Однако, если H — множество положительных целых чисел с обычным умножением, то $H = H \circ H$, но H не является группой.

ТЕОРЕМА 9.81. Если H — нормальная подгруппа группы $(G, \circ)$, то $abH = (aH)(bH)$ для всех $a, b \in G$.

ДОКАЗАТЕЛЬСТВО. Для $a, b \in G$,

$$\begin{aligned} abH &= a(bH) = && \text{по теореме 9.75} \\ &= a(b(HH) = && \text{по теореме 9.80} \\ &= a(bH)H = && \text{по теореме 9.75} \\ &= a(Hb)H = && \text{по теореме 9.79} \\ &= (aH)(bH). && \text{по теореме 9.75} \end{aligned}$$

■

СЛЕДСТВИЕ 9.82. Если H — нормальная подгруппа группы $(G, \circ)$, то смежные классы подгруппы H в группе G образуют группу относительно операции $(aH)(bH) = abH$. Эта группа называется ***фактор-группой*** и обозначается G/H.

СЛЕДСТВИЕ 9.83. Если $f : G \to G/H$ определить соотношением $f(a) = aH$, то f — гомоморфизм.

Доказательство приведенных ниже теорем предоставляется читателю.

ТЕОРЕМА 9.84. Если $f : G \to G'$ — гомоморфизм из группы $(G, \circ)$ в группу $(G', \circ)$, то $H = \{x : f(x) = 1$, единице группы $G'\}$ будет подгруппой группы G.

ТЕОРЕМА 9.85. Если $f : G \to H$ — гомоморфизм из группы G в группу H и K — подгруппа группы H, то $f^{-1}(\{1\})$, где 1 — единица группы H, есть нормальная подгруппа группы G.

В разделе 4.2 показано, что перестановки на конечном множестве замкнуты относительно бинарной операции композиции. Было также показано, как найти перестановку, обратную данной. Как и другие функции, эти перестановки ассоциативны относительно композиции. Следовательно, множество всех перестановок фиксированного конечного множества образует группу относительно бинарной операции композиции. Множество всех перестановок n элементов называется симметрической группой и обозначается S_n.

ПРИМЕР 9.86. Вспомним, что если f перестановка на множестве $\{1, 2, 3, \ldots, n\}$, то ее можно представить в виде

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ f(1) & f(2) & \ldots & f(n) \end{pmatrix}.$$

Рассмотрим следующую подгруппу симметрической группы S_4, основанную на квадрате, изображенном на рис. 9.12.

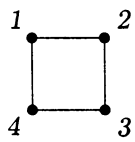


Рис. 9.12

Перестановка

$$\delta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

меняет местами 1 и 3. Геометрически — это отражение квадрата относительно диагонали между вершинами 2 и 4. Аналогично, перестановка

$$\delta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

меняет местами 2 и 4. Геометрически — это отражение квадрата относительно диагонали между вершинами 1 и 3. Перестановка

$$\phi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

геометрически отражает квадрат относительно горизонтальной оси, проходящей через центр квадрата. Перестановка

$$\phi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

геометрически отражает квадрат относительно вертикальной оси, проходящей через центр квадрата. Пусть ρ_1, ρ_2, и ρ_3 — вращения квадрата по часовой стрелке на 90°, 180° и 270° соответственно. Следовательно, $\rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$, $\rho_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$ и $\rho_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$. Пусть I — тождественная перестановка. Все результаты умножения этих элементов содержит приведенная ниже таблица.

$\circ$	I	ρ_1	ρ_2	ρ_3	δ_1	δ_2	ϕ_1	ϕ_2
I	I	ρ_1	ρ_2	ρ_3	δ_1	δ_2	ϕ_1	ϕ_2
ρ_1	ρ_1	ρ_2	ρ_3	I	ϕ_1	ϕ_2	δ_2	δ_1
ρ_2	ρ_2	ρ_3	I	ρ_1	δ_2	δ_1	ϕ_2	ϕ_1
ρ_3	ρ_3	I	ρ_1	ρ_2	ϕ_2	ϕ_1	δ_1	δ_2
δ_1	δ_1	ϕ_2	δ_2	ϕ_1	I	ρ_2	ρ_3	ρ_1
δ_2	δ_2	ϕ_1	δ_1	ϕ_2	ρ_2	I	ρ_1	ρ_3
ϕ_1	ϕ_1	δ_1	ϕ_2	δ_2	ρ_1	ρ_3	I	ρ_2
ϕ_2	ϕ_2	δ_2	ϕ_1	δ_1	ρ_3	ρ_1	ρ_2	I

Эта группа называется ***октической группой***, или ***группой симметрий квадрата***. Множество $H = \{I, \rho_1, \rho_2, \rho_3\}$ — подгруппа октической группы. Поскольку H содержит четыре элемента, и число элементов в смежном классе должно делить порядок группы, то существует только один смежный левый класс, aH для любого $a \notin H$. Этот класс состоит из всех элементов октической группы, которые не попали в H. Но он должен совпадать с Ha для любого $a \notin H$, поскольку существует только один другой правый смежный класс. Следовательно, H — нормальная подгруппа октической группы. Читателю остается показать, что $\{I, \rho_2\}$ — нормальная подгруппа октической группы. □

Пусть A — конечное множество из n элементов и пусть S_n — группа перестановок на A. Поскольку имеется $n!$ перестановок элементов множества A, то S_n содержит $n!$ элементов. Пусть, для удобства, 1, 2, ..., n — элементы из A. Для фиксированной перестановки σ определим отношение R на A таким образом: aRb, если $b = \sigma^k(a)$ для некоторого неотрицательного целого числа k.

ТЕОРЕМА 9.87. Отношение R есть отношение эквивалентности на A.

ДОКАЗАТЕЛЬСТВО. Поскольку группа S_n конечна, то $\sigma^m = 1$, единице группы S_n для некоторого неотрицательного целого числа m. Следовательно, $\sigma^m(a) = a$ и aRa. Пусть aRb, тогда $b = \sigma^k(a)$ для некоторого неотрицательного целого числа k. Поскольку $\sigma^k\sigma^{m-k} = \sigma^m = 1$, то σ^{m-k} — обратный элемент для σ^k, так что $\sigma^{m-k}(b) = a$ и bRa. Следовательно, R симметрично. Пусть aRb и bRc. Тогда $b = \sigma^j(a)$ для некоторого неотрицательного целого числа j и $c = \sigma^k(b)$ для некоторого неотрицательного целого числа k. Следовательно, $c = \sigma^{j+k}(a)$, и R транзитивно. Поэтому R — отношение эквивалентности. ■

Поскольку R является отношением эквивалентности, оно делит множество $\{1, 2, 3, \ldots, n\}$ на классы эквивалентности. Например, пусть A=$\{1, 2, 3, \ldots, 8\}$ и

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 3 & 2 & 7 & 8 & 6 & 1 & 5 \end{pmatrix}.$$

Легко видеть, что множество классов эквивалентности имеет вид $\{\{1, 4, 7\}, \{2, 3\}, \{5, 8\}, \{6\}\}$. Множество классов эквивалентности, созданное таким образом, называется ***орбитой*** перестановки σ.

Теперь определим специальный вид перестановки, называемый ***циклом***. Если обозначить $\sigma_c = (a_1, a_2, a_3, \ldots, a_k)$ как цикл, тогда $\sigma_c(a_i) = a_{i+1}$ для $i < k$, и $\sigma_c(a_k) = a_1$. Например, если $\sigma_c = (1, 3, 5, 7)$ — цикл в S_8, то

$$\sigma_c = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 2 & 5 & 4 & 7 & 6 & 1 & 8 \end{pmatrix}.$$

Данную перестановку σ можно записать как произведение циклов, которые не имеют общих элементов. Для этого сначала необходимо разделить множество A на классы эквивалентности. Выбираем элемент a и образуем цикл $(a, \sigma(a), \sigma^2(a), \sigma^3(a), \ldots, \sigma^k(a))$, где $\sigma^{k+1}(a) = a$. Эти циклы не содержат общих компонент, поскольку классы эквивалентности не пересекаются. Назовем эти циклы ***непересекающимися***.

Например, в перестановке

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 3 & 2 & 7 & 8 & 6 & 1 & 5 \end{pmatrix}$$

с множеством классов эквивалентности $\{\{1, 4, 7\}, \{2, 3\}, \{5, 8\}, \{6\}\}$ циклами являются $(1, 4, 7)$, $(2, 3)$, $(5, 8)$ и (6). Поскольку каждый цикл двигает только числа из цикла, а циклы непересекающиеся, то $\sigma = (1, 4, 7)(2, 3)(5, 8)(6)$. Поскольку циклы не пересекаются, то порядок циклов не меняет перестановку. По желанию, одноэлементные циклы можно опускать, поскольку они реально ничего не меняют. Поэтому можно записать $\sigma = (1, 4, 7)(2, 3)(5, 8)$. Если использовать одноэлементные циклы, то тождественную перестановку в S_8 можно записать как $(1)(2)(3)(4)(5)(6)(7)(8)$.

■ УПРАЖНЕНИЯ

1. Докажите теорему 9.70. Пусть $f : G \to H$ — гомоморфизм из группы G в группу H и 1 — единица группы G. Тогда $f(1)$ — единица группы H.
2. Докажите теорему 9.71. Пусть $f : G \to H$ — гомоморфизм из группы G в группу H и g' — элемент, обратный элементу g из G. Тогда $f(g')$ есть элемент, обратный элементу $f(g)$ из H.
3. Следуя доказательству теоремы 9.64, пусть $a \circ H$ — левый смежный класс подгруппы H группы G. Определим $f : H \to a \circ H$ соотношением $f(h) = a \circ h$. Покажите, что функция f — биекция.
4. Докажите теорему 9.78. Ядро гомоморфизма $f : G \to H$ есть нормальная подгруппа группы G.
5. Пусть $Z_5^R = \{[1], [2], [3], [4]\}$ — множество элементов Z_5 без элемента $[0]$. Эти элементы соответствуют целым числам, которые дают при делении на 5, соответственно, остатки $1, 2, 34$. Вместе с операцией умножения, $\cdot$, введенной для Z_5, получаем группу $(Z_5^R, \cdot)$. Покажите, что $(Z_5^R, \cdot)$ изоморфна $(Z_4, +)$, непосредственно указав изоморфизм $f : Z_5^R \to Z_4$.
6. Покажите, что $\{I, \rho_2\}$ — нормальная подгруппа октической группы.
7. Найдите двухэлементную подгруппу октической группы, которая не является нормальной подгруппой.
8. Докажите теорему 9.72. Если $f : G \to H$ — гомоморфизм из группы G в группу H и K — подгруппа группы H, то $f^{-1}(K)$ — подгруппа группы G.
9. Докажите теорему 9.73. Если $f : G \to H$ — гомоморфизм из группы G в группу H и K — подгруппа G, то $f(K)$ — подгруппа группы H.
10. Запишите приведенные ниже перестановки как произведение непересекающихся циклов:
    а) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 9 & 4 & 7 & 8 & 6 & 1 & 5 & 2 \end{pmatrix}$;
    б) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 7 & 8 & 3 & 6 & 9 & 1 & 5 \end{pmatrix}$;
    в) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 1 & 7 & 3 & 6 & 5 \end{pmatrix}$;
    г) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 1 & 2 \end{pmatrix}$.
11. Найдите $\sigma \circ \tau$, где
    а) $\sigma = (1,3,4)(2,5)$ и $\tau = (1,2)(3,4)(5)$;
    б) $\sigma = (1,2)(3,4)(5,6)$ и $\tau = (1,2)(3,4)(5,6)$;
    в) $\sigma = (1,2,3)(4,5,6)$ и $\tau = (1,2)(3,4)(5,6)$;
    г) $\sigma = (1,2,3,8)(4,5,6)(7)$ и $\tau = (1,2)(3,8)(4,5)(6,7)$.
12. Докажите теорему 9.75. Если H, J и K — подмножества группы $(G, \circ)$, то $(H \circ J) \circ K = (H \circ (J \circ K)$
13. Докажите теорему, обратную теореме 9.79. Если для заданной подгруппы H группы $(G, \circ)$ имеет место $gH = Hg$ для всех $g \in G$, то H — нормальная подгруппа.

ГЛАВА

10 НЕКОТОРЫЕ СПЕЦИАЛЬНЫЕ ВОПРОСЫ ТЕОРИИ ЧИСЕЛ

10.1. ЦЕЛОЧИСЛЕННЫЕ РЕШЕНИЯ ЛИНЕЙНЫХ УРАВНЕНИЙ

В данной главе рассмотрены некоторые положения теории чисел, которые являются в большей степени алгебраическими. Начнем с нахождения решений алгебраических уравнений.

Используя алгоритм Евклида, можно найти целочисленные решения уравнения $ax+by=c$, если они существуют. Приведенное ниже утверждение было ранее доказано как теорема 3.38 и приводится здесь еще раз для удобства изложения.

ТЕОРЕМА 10.1. Уравнение $ax+by=c$, где a, b и c — целые числа, имеет целочисленное решение (т.е. существуют целые числа x и y такие, что $ax+by=c$) тогда и только тогда, когда c делится на НОД(a,b).

Если c делится на НОД(a,b), то решением уравнения $ax+by=c$ является пара чисел

$$x_0 = \frac{u \cdot c}{\text{НОД}(a,b)}, \qquad y_0 = \frac{v \cdot c}{\text{НОД}(a,b)},$$

где u и v — любое решение уравнения НОД$(a,b) = au + bv$.

ПРИМЕР 10.2. Найти решение уравнения $85x+34y=51$. В примере 7.5 показано, что НОД$(85,34)=17$ и $(85)(1)+(34)(-2)=17$. Поэтому решение имеет вид

$$x_0 = \frac{u \cdot c}{\text{НОД}(a,b)} = \frac{1 \cdot 51}{17} = 3$$

и

$$y_0 = \frac{v \cdot c}{\text{НОД}(a,b)} = \frac{(-2) \cdot 51}{17} = -6.$$

Для проверки вычислим

$$ax_0 + by_0 = 85 \cdot 3 + 34 \cdot (-6) = 255 + (-204) = 51.$$

Иной способ построения решения состоит в непосредственном использовании уравнения $au + bv = \text{НОД}(a, b)$. Поскольку

$$au + bv = \text{НОД}(a, b)$$

или

$$a \cdot (1) + b \cdot (-2) = 17,$$

то, умножая на 3, получаем

$$a \cdot (3) + b \cdot (-6) = 51.$$

Замечая, что при $x = 5$ и $y = -11$

$$85 \cdot 5 + 34 \cdot (-11) = 425 + (-374) = 51,$$

приходим к выводу, что может существовать более одного решения. □

ПРИМЕР 10.3. Решить уравнение $252x + 580y = 20$.В примере 7.6 показано, что $\text{НОД}(252, 580) = 4$ и $(252)(-23) + (580)(10) = 4$. Умножая каждое слагаемое на 5, получаем $(252)(-115) + (580)(50) = 20$. Следовательно, $x = -115$ и $y = 50$ являются решением. □

Теперь у нас есть возможность определять, существует ли решение, и находить частное решение для уравнения $ax + by = c$, если решение существует. Теорема, приведенная ниже, дает возможность находить все решения уравнения.

ТЕОРЕМА 10.4. Если a и b — ненулевые целые числа и (x_0, y_0) — решение уравнения $ax + by = c$, тогда любое другое решение (x, y) имеет вид

$$x = x_0 + \frac{b}{d}t;$$

$$y = y_0 - \frac{a}{d}t,$$

где t — произвольное целое число, а $d = \text{НОД}(a, b)$.

ДОКАЗАТЕЛЬСТВО. Если (x, y) и (x_0, y_0) являются решениями уравнения $ax + by = c$, то $ax + by = ax_0 + by_0$. Поэтому $ax - ax_0 = by_0 - by$ и $a(x - x_0) = b(y_0 - y)$. Разделив обе части соотношения на $d = \text{НОД}(a, b)$, имеем $\frac{a}{d}(x - x_0) = \frac{b}{d}(y_0 - y)$. Поскольку по теореме 7.8 $\text{НОД}\left(\frac{a}{d}, \frac{b}{d}\right) = 1$, получаем, что $\frac{a}{d}(y_0 - y)$ и $\frac{b}{d}(x - x_0)$, скажем, $x - x_0 = u\left(\frac{b}{d}\right)$ и $(y_0 - y) = v\left(\frac{a}{d}\right)$. Тогда $\left(\frac{a}{d}\right)u\left(\frac{b}{d}\right) = \left(\frac{b}{d}\right)v\left(\frac{a}{d}\right)$, и отсюда после сокращения $u = v$. Положив $t = u = v$, получаем $x = x_0 + \left(\frac{b}{d}\right)t$ и $y = y_0 - \left(\frac{a}{d}\right)t$. Необходимо еще показать, что пара $\left(x_0 + \left(\frac{b}{d}\right)t,\ y_0 - \left(\frac{a}{d}\right)t\right)$ является решением уравнения $ax + by = c$. Но

$$\begin{aligned} a\left(x_0 + \left(\frac{b}{d}\right)t\right) + b\left(y_0 - \left(\frac{a}{d}\right)t\right) &= ax_0 + a\left(\frac{b}{d}\right)t + by_0 - b\left(\frac{a}{d}\right)t = \\ &= ax_0 + by_0 = \\ &= c. \end{aligned}$$

■

Возвращаясь к примеру 10.2, мы видим, что общее решение уравнения $85x + 34y = 51$ имеет вид $x = 3 + 2t$ и $y = -6 - 5t$. В примере 10.3 общим решением уравнения $252x + 580y = 20$ является $x = -115 + 145t$ и $y = 50 - 63t$.

■ УПРАЖНЕНИЯ

1. Найдите решение каждого из приведенных ниже уравнений:
   а) $24x + 81y = 6$; **б)** $803x + 154y = 33$; **в)** $73x + 151y = 3$;
   г) $165x + 418y = 121$; **д)** $27x + 78y = 12$.
2. Найдите общее решение каждого из уравнений упражнения 1.
3. Найдите (если оно существует) решение каждого из приведенных ниже уравнений:
   а) $23x + 18y = 4$; **б)** $299x + 533y = 52$; **в)** $39x + 299y = 27$;
   г) $272x + 102y = 68$; **д)** $27x + 180y = 33$.
4. Найдите общее решение каждого из уравнений упражнения 3, если оно существует.
5. Напишите в псевдокодах процедуру для нахождения целочисленных решений линейных уравнений.

10.2. РЕШЕНИЯ СРАВНЕНИЙ

В предыдущем разделе рассматривались уравнения вида $ax + by = c$, имеющие в качестве решений целые числа x и y. В частном случае $b = 0$ можно находить решения уравнения $ax = c$. Теперь будем искать решения сравнений $ax \equiv c \pmod{n}$ в том смысле, что требуется найти такое целое число x, при котором целое число ax сравнимо с c по модулю n. Или, на языке классов эквивалентности, для классов эквивалентности $[a]$ и $[c]$ по модулю n, найти класс эквивалентности $[x]$ такой, что $[a] \odot [x] = [c]$, где равенство означает равенство множеств.

Используя таблицу умножения для Z_5 (см. раздел 3.6),

$[a] \odot [b]$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[2]$	$[0]$	$[2]$	$[4]$	$[1]$	$[3]$
$[3]$	$[0]$	$[3]$	$[1]$	$[4]$	$[2]$
$[4]$	$[0]$	$[4]$	$[3]$	$[2]$	$[1]$

можно найти решение сравнения

$$3x \equiv 1 \pmod{5},$$

рассматривая (по модулю 5)

$$[3] \odot [x] = [1].$$

Анализ таблицы умножения показывает, что $[x] = [2]$ будет решением, т.к. $[3] \odot [2] = [1]$. Поэтому можно положить $x = 2$. Поскольку $[2] = \{\ldots, -8, -3, 2, 7, 12, \ldots\}$, отмечаем, что каждое число из совокупности $x = -3$, $x = 7$ и $x = -8$ является тем значением x, которое удовлетворяет

$$3x \equiv 1 \pmod 5.$$

В теореме, приведенной ниже, сформулированы важные свойства сравнимости. Части (а) и (г) доказаны в теореме 3.54. Доказательство остальных пунктов предоставляется читателю.

ТЕОРЕМА 10.5.

а) Если $a \equiv b \pmod n$ и $c \equiv d \pmod n$, то $a + c \equiv b + d \pmod n$ и $ac = bd \pmod n$.
б) Если $ac \equiv bc \pmod n$ и НОД$(c, n) = 1$, то $a \equiv b \pmod n$.
в) Если $a \equiv b \pmod n$, то $a^m \equiv b^m \pmod n$ для всех целых положительных чисел m.
г) Если $a \equiv b \pmod{mn}$, то $a \equiv b \pmod m$ и $a \equiv b \pmod n$.
д) Для $c \neq 0$ соотношение $ac \equiv bc \pmod n$ имеет место тогда и только тогда, когда $a \equiv b \left(\mathrm{mod}\ \dfrac{n}{\text{НОД}(c,n)}\right)$.
е) Если $a \equiv b \pmod m$, $a \equiv b \pmod n$ и НОД$(m, n) = 1$, то $a \equiv b \pmod{mn}$.

В следующих двух теоремах формулируются условия, при которых решение $ax \equiv c \pmod n$ существует, и указывается явный вид решения.

ТЕОРЕМА 10.6. Сравнение $ax \equiv c \pmod m$ имеет решением целое число x тогда и только тогда, когда НОД$(a, m) \mid c$. Все целочисленные решения имеют вид

$$x = x_0 + \frac{t \cdot m}{\text{НОД}(a, m)},$$

где t — любое целое число, и для x_0 существует такое y_0, что (x_0, y_0) является решением уравнения $ax + my = c$.

ДОКАЗАТЕЛЬСТВО. По определению, $ax \equiv c \pmod m$ тогда и только тогда, когда $ax - c$ делится на m, т.е. существует целое число j такое, что $ax - c = jm$. Это справедливо тогда и только тогда, когда

$$ax + my = c$$

имеет решение. Целые числа a, m и c фиксированы, и требуется найти целые числа x и y такие, что $ax + my = c$. По теореме 10.1 $ax + my = c$ имеет целочисленное решение тогда и только тогда, когда НОД$(a, m) \mid c$. Также согласно этой теореме, решение имеет вид

$$x_0 = \frac{u \cdot c}{\text{НОД}(a, m)}, \qquad y_0 = \frac{v \cdot c}{\text{НОД}(a, m)},$$

где u и v выбраны таким образом, что $au + mv = \text{НОД}(a, m)$. По теореме 10.4 все решения имеют вид

$$x = x_0 + \frac{t \cdot m}{\text{НОД}(a, m)}, \qquad y = y_0 - \frac{t \cdot m}{\text{НОД}(a, m)},$$

где t — произвольное целое число. В данном случае необходимо решение только для x. Таким образом, все целочисленные решения $ax \equiv c \pmod{m}$ имеют вид

$$x = x_0 + \frac{t \cdot m}{\text{НОД}(a, m)},$$

где t — любое целое число. ■

Следующая теорема предоставляет различные решения сравнений $ax \equiv c \pmod{m}$. Поскольку существует конечное число классов эквивалентности по модулю m, может существовать только конечное число различных решений по модулю m. Все они даны в приведенной ниже теореме. Доказательство теоремы предоставляется читателю.

ТЕОРЕМА 10.7. Если $\text{НОД}(a, m) \mid c$, то $ax \equiv c \pmod{m}$ имеет конечное число различных решений по модулю m. Эти решения имеют вид

$$x_0 + \frac{t \cdot m}{\text{НОД}(a, m)} \text{ по модулю } m = \left[\left[x_0 + \frac{t \cdot m}{\text{НОД}(a, m)}\right]\right]_m$$

для $t = 1, 2, 3, \ldots, \text{НОД}(a, m)$, где для x_0 существует такое y_0, что (x_0, y_0) является решением уравнения $ax + my = c$.

ПРИМЕР 10.8. В силу того, что $\text{НОД}(35, 84) = 7$ и $7 \mid 14$, сравнение

$$35 \cdot x \equiv 14 \pmod{84}$$

имеет в точности 7 различных решений по модулю 84, которые имеют вид

$$x_0 + \frac{84 \cdot t}{7} = x_0 + 12 \cdot t,$$

где $t = 1, 2, 3, \ldots, 7$ и (x_0, y_0) является решением

$$35x + 84y = 14,$$

которое равносильно

$$5x + 12y = 2.$$

Проверка дает в качестве решения $x_0 = -2$ и $y_0 = 1$. Семь различных решений по модулю 84 имеют следующий вид.

t	$x_0 + 12t$
1	$-2 + 12 \cdot 1 = 10$
2	$-2 + 12 \cdot 2 = 22$
3	$-2 + 12 \cdot 3 = 34$
4	$-2 + 12 \cdot 4 = 46$
5	$-2 + 12 \cdot 5 = 58$
6	$-2 + 12 \cdot 6 = 70$
7	$-2 + 12 \cdot 7 = 82$

Когда НОД$(a, m) = 1$, существует единственное решение сравнения $ax \equiv c \pmod{m}$. Например, рассмотрим сравнение

$$6x \equiv 7 \pmod{55},$$

НОД$(6, 55) = 1$, и, очевидно, 1 делит 7. Поэтому существует только одно решение по модулю 55, которое имеет вид

$$x_0 + \frac{t \cdot m}{\text{НОД}(6, 55)} = x_0 + \frac{1 \cdot 55}{1} = x_0 + 55 \equiv x_0 \pmod{55},$$

где (x_0, y_0) — решение уравнения $ax + my = c$ или $6x + 55y = 7$. Для нахождения x_0 и y_0 начнем перебор с возвратом по алгоритму Евклида, как показано в примерах, следующих за теоремой 7.2, получая при этом $6(-9) + 55(1) = 1 =$ НОД$(6, 55)$. Умножая каждое слагаемое на 7, получаем $6(-63) + 55(7) = 7$, так что $x_0 = -63$ и $x = -63 + 55 = -8$. □

ПРИМЕР 10.9. Решить сравнение

$$623x \equiv -406 \pmod{84}.$$

Число 623 больше, чем модуль сравнения 84, а -406 является отрицательным. Поскольку разыскиваются решения по модулю 84, выбираем целые числа в диапазоне $0, 1, 2, \ldots, 83$, т.к. они являются возможными остатками при делении на 84 и простейшими представителями классов эквивалентности, порожденных сравнимостью по модулю 84. Используя алгоритм деления, получаем

$$623 = 84 \cdot 7 + 35, \text{ так что } 623 \equiv 35 \pmod{84};$$
$$-406 = 84(-5) + 14, \text{ так что } -406 \equiv 14 \pmod{84}.$$

Таким образом,

$$35x \equiv 14 \pmod{84}$$

равносильно исходному $623x \equiv -406 \pmod{84}$. Решение сравнения $35x \equiv 14 \pmod{84}$ было найдено в предыдущем примере. □

Доказательство следующей теоремы предоставляется читателю.

ТЕОРЕМА 10.10. Если $a \equiv b \pmod{n_1}$, $a \equiv b \pmod{n_2}$, ..., $a \equiv b \pmod{n_k}$ $n =$ НОК$(n_1, n_2, \ldots, n_k)$, то $a \equiv b \pmod{n}$, и обратно.

■ УПРАЖНЕНИЯ

1. Для каких значений n справедливо $75 \equiv 35 \pmod{n}$?
2. Найдите решения следующих сравнений:
   а) $4x \equiv 3 \pmod{7}$; **б)** $27x \equiv 12 \pmod{15}$; **в)** $28x \equiv 56 \pmod{49}$;
   г) $24x \equiv 6 \pmod{81}$; **д)** $91x \equiv 26 \pmod{169}$.
3. Докажите, что если a — целое нечетное число, то $a^2 \equiv 1 \pmod{8}$.
4. Перечислите все положительные целые числа, которые одновременно сравнимы с 5 по модулю 3 и сравнимы с 5 по модулю 2.

5. Докажите, что если $a \equiv b \pmod{n_1}$, $a \equiv b \pmod{n_2}$, ..., $a \equiv b \pmod{n_k}$, и $n = \text{НОК}(n_1, n_2, \ldots, n_k)$, то $a \equiv b \pmod{n}$.
6. Докажите, что если $ac \equiv bc \pmod{n}$ и $\text{НОД}(c, n) = 1$, то $a \equiv b \pmod{n}$.
7. Докажите, что если $a \equiv b \pmod{n}$, то $a^m \equiv b^m \pmod{n}$ для всех положительных целых чисел m.
8. Докажите, что если $a \equiv b \pmod{m}$, $a \equiv b \pmod{n}$, и $\text{НОД}(m, n) = 1$, то $a \equiv b \pmod{mn}$.
9. Докажите, что для $c \neq 0$ $ac \equiv bc \pmod{n}$ тогда и только тогда, когда $a \equiv b \left(\text{mod } \dfrac{n}{\text{НОД}(c, n)}\right)$.

10.3. КИТАЙСКАЯ ТЕОРЕМА ОБ ОСТАТКАХ

В предыдущем разделе рассматривались сравнения вида

$$ax \equiv c \pmod{m},$$

где целые числа a и c, а также положительное целое число m были заданы, и требовалось найти целое число x, удовлетворяющее сравнению. Теперь переходим к рассмотрению системы сравнений:

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\vdots \\ x &\equiv a_n \pmod{m_n}, \end{aligned}$$

где числа m_i — попарно взаимно простые. То есть, требуется найти целое число x, которое при делении на m_i дает остаток a_i, если $\text{НОД}(m_i, m_j) = 1$ при $i \neq j$.

Еще в древние времена люди рассматривали системы сравнений и успешно их решали. Очень часто ставились задачи на устный счет, наподобие следующей. Представьте, что группа обезьян пытается разложить по кучкам груду кокосовых орехов. Если обезьяны разложат орехи в кучки по пять штук, то останется четыре ореха. Если разложат в кучки по четыре, останутся три ореха. Кучки по семь дадут остаток два. Кучки по девять — остаток шесть. Каково минимальное возможное число орехов?

Если x — возможное число орехов в кучке, тогда наличие четырех в остатке при раскладке в кучки по пять можно выразить как

$$x \equiv 4 \pmod{5}.$$

Аналогично, другие условия имеют вид

$$\begin{aligned} x &\equiv 3 \pmod{4}; \\ x &\equiv 2 \pmod{7}; \\ x &\equiv 6 \pmod{9}. \end{aligned}$$

Наименьшее целое положительное число x, удовлетворяющее четырем сравнениям, и является искомым решением. Решение таких задач дает следующая теорема.

ТЕОРЕМА 10.11. (Китайская теорема об остатках) Пусть $m_1, m_2, \ldots, m_n$ — попарно взаимно простые числа , т.е. НОД$(m_i, m_j) = 1$ для всех i и j, меньших или равных n, где $i \neq j$. Тогда система сравнений

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\vdots \\ x &\equiv a_n \pmod{m_n} \end{aligned}$$

имеет решение, которое единственно по модулю, равному целому числу $m_1 m_2 \cdots m_n$ Далее, если

$$M_j = \frac{\prod_{i=1}^{n} m_i}{m_j}$$

и z_j — решение сравнения $M_j z_j \equiv a_j \pmod{m_j}$ для каждого j, тогда решение имеет вид

$$x = \left[\left[\sum_{j=1}^{n} M_j z_j\right]\right]_{m_1 m_2 \cdots m_n}.$$

ДОКАЗАТЕЛЬСТВО. Пусть x определено согласно теореме. Тогда при любом k, $1 \leq k \leq n$,

$$x = \left[\left[\sum_{j=1}^{n} M_j z_j\right]\right]_{m_1 m_2 \cdots m_n},$$

так что

$$\begin{aligned} x &\equiv \sum_{j=1}^{n} M_j z_j \pmod{\prod_{i=1}^{n} m_i} \\ &\equiv \sum_{j=1}^{n} M_j z_j \pmod{m_k}, \text{ по теореме 10.5(г)} \\ &\equiv M_k z_k \pmod{m_k} \\ &\equiv a_k \pmod{m_k}, \end{aligned}$$

поэтому x удовлетворяет n сравнениям, $x \equiv a_k \pmod{m_k}$ при $1 \leq k \leq n$. Если x' также удовлетворяет n сравнениям, тогда

$$x - x' \equiv 0 \pmod{m_i} \quad \text{при } 1 \leq i \leq n.$$

Поскольку НОД$(m_i, m_j) = 1$ при $i \neq j$, получаем

$$x \equiv x' \pmod{\prod_{i=1}^{n} m_i},$$

т.е. решение x единственно по модулю $\prod_{i=1}^{n} m_i$. ■

ПРИМЕР 10.12. Найти решение системы сравнений

$$x \equiv 5 \pmod 4;$$
$$x \equiv 7 \pmod{11}.$$

Поскольку числа 4 и 11 — взаимно простые, существует целое число, а именно, 10 такое, что $(4)(10) \equiv 7 \pmod{11}$, и существует целое число (а именно, 3) такое, что $(11)(3) \equiv 5 \pmod 4$. Следовательно, $(4)(10) + (11)(3) = 73$, которое сравнимо с 29 по модулю 44, удовлетворяет обоим вышеприведенным сравнениям. □

ПРИМЕР 10.13. Найдем ответ на вопрос об обезьянах и орехах, решая систему сравнений

$$x \equiv 4 \pmod 5;$$
$$x \equiv 3 \pmod 4;$$
$$x \equiv 2 \pmod 7;$$
$$x \equiv 6 \pmod 9.$$

Имеем $M_1 = 4 \cdot 7 \cdot 9 = 252$, $M_2 = 5 \cdot 7 \cdot 9 = 315$, $M_3 = 180$ и $M_4 = 140$. Поскольку числа 5 и 252 — взаимно простые, существует целое число z_1 такое, что $252z_1 \equiv 4 \pmod 5$ или, что то же самое, $2z_1 \equiv 4 \pmod 5$, или $z_1 \equiv 2 \pmod 5$. Следовательно, z_1 может быть равно 7.

Поскольку числа 4 и 315 — взаимно простые, существует целое число z_2 такое, что $315z_2 \equiv 3 \pmod 4$ или, что то же самое, $3z_2 \equiv 3 \pmod 4$. Следовательно, z_2 может быть равно 1.

Поскольку числа 7 и 180 — взаимно простые, существует целое число z_3 такое, что $180z_3 \equiv 2 \pmod 7$ или, что то же самое, $5z_3 \equiv 2 \pmod 7$. Следовательно, z_3 может быть равно 6.

Поскольку числа 9 и 140 — взаимно простые, существует целое число z_4 такое, что $140z_4 \equiv 6 \pmod 9$ или, что то же самое, $5z_4 \equiv 6 \pmod 9$. Следовательно, z_4 может быть равно 3.

Отсюда $x = (7)(252) + (1)(315) + (6)(180) + (3)(140) \pmod{5 \cdot 4 \cdot 7 \cdot 9}$ или $x \equiv 3579 \pmod{1260}$ и $x = 1059$ — наименьшее положительное целочисленное решение. □

Китайскую теорему об остатках можно следующим образом обобщить на случай модулей $m_1, m_2, \ldots, m_n$, которые не являются взаимно простыми.

ТЕОРЕМА 10.14. Система сравнений

$$x \equiv a_1 \pmod{m_1}$$
$$x \equiv a_2 \pmod{m_2}$$
$$\vdots$$
$$x \equiv a_n \pmod{m_n}$$

имеет решение тогда и только тогда, когда НОД(m_i, m_j) делит $a_i - a_j$ при всех i и j, где $1 \leq i < j \leq n$. Если решение существует, оно единственно по модулю НОК$(m_1, m_2, \ldots, m_n)$.

ДОКАЗАТЕЛЬСТВО. Для $n = 1$ справедливость теоремы очевидна. Покажем технику доказательства, рассматривая случай $n = 2$. Пусть заданы уравнения

$$x \equiv a_1 \pmod{m_1};$$
$$x \equiv a_2 \pmod{m_2}.$$

Имеем

$$x = a_1 + km_1.$$

Подставляя это значение во второе уравнение, получаем

$$a_1 + km_1 \equiv a_2 \pmod{m_2}$$

или

$$km_1 \equiv a_2 - a_1 \pmod{m_2}.$$

Поскольку НОД(m_1, m_2) делит $a_1 - a_2$, существует решение для k, а $x + km_1$ является решением для обоих уравнений. Поскольку m_1 и m_2 делят НОК$((m_1, m_2))$, то

$$x + km_1 + \text{НОК}((m_1, m_2))$$

также является решением. Если как x_1, так и x_2 — решения обоих сравнений, то $x_1 - x_2$ делится как на m_1, так и на m_2, и поэтому делится на НОК$((m_1, m_2))$.

Теперь предположим, что имеется система сравнений

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\vdots \\ x &\equiv a_k \pmod{m_k} \\ x &\equiv a_{k+1} \pmod{m_{k+1}} \end{aligned}$$

и есть значение $\bar{x}$ такое, что $\bar{x} \equiv a_i \pmod{m_i}$ при $1 \leq i \leq k$. Поэтому каждое решение сравнения $x \equiv a_i \pmod{m_i}$ при $1 \leq i \leq k$ имеет вид $\bar{x} + uM_k$, где

$$M_k = \text{НОК}(m_1, m_2, \ldots, m_k).$$

Подставляя это значение в последнее сравнение, имеем

$$\bar{x} + uM_k \equiv a_{k+1} \pmod{m_{k+1}}.$$

Если существует решение для u, тогда $\bar{x} + uM_k$ есть решение для всех $k + 1$ сравнений. Решение этого сравнения существует, если НОД$(M_k, m_{k+1}) | (\bar{x} - a_{k+1})$. Но для всех i

$$\bar{x} - a_{k+1} = \bar{x} - a_i + a_i - a_{k+1},$$

а поскольку $\bar{x} - a_i$ делится на m_i и $a_i - a_{k+1}$ делится на НОД(m_i, m_{k+1}), то НОД$(m_i, m_{k+1}) | (\bar{x} - a_{k+1})$ для всех i. Таким образом, НОД$(M_k, m_{k+1}) | (\bar{x} - a_{k+1})$. (Чтобы убедиться в этом, подберите наивысшую степень простого числа, которая делит M_k и m_{k+1}. Покажите, что это число делит m_i и m_{k+1} для некоторого i.) Мы показали, что решение существует. Предоставляем читателю доказательство единственности решения по модулю НОК$(m_1, m_2, \ldots, m_n)$. ■

Приведенный ниже пример демонстрирует алгоритм решения таких задач.

ПРИМЕР 10.15. Найти решение:

$$x \equiv 5 \pmod 6;$$
$$x \equiv 3 \pmod{10};$$
$$x \equiv 8 \pmod{15}.$$

Решение сравнения

$$x \equiv 5 \pmod 6$$

имеет вид $5+6t$. Подставляя эту величину во второе сравнение, получаем $5+6t \equiv 3 \pmod{10}$. Поэтому $6t \equiv 8 \pmod{10}$ и

$$t \equiv 3 \pmod{10}.$$

Поэтому, $t = 3 + 10u$. Подставляя t в третье уравнение, имеем

$$3 + 10u \equiv 8 \pmod{10}.$$

Поэтому $10u \equiv 5 \pmod{10}$ и $u \equiv 2 \pmod{15}$. Отсюда $u = 2 + 15v$ для некоторого v, и

$$t = 3 + 10(2 + 15v) = 150v + 23,$$

а

$$x = 5 + 6(150v + 23) = 203 + 6 \cdot 10 \cdot 15v,$$

так что $x = 203$ является решением. □

■ УПРАЖНЕНИЯ

1. Решите приведенные ниже системы сравнений:

а) $x \equiv 9 \pmod{12}$;
$x \equiv 6 \pmod{25}$.

б) $x \equiv 3 \pmod 4$;
$x \equiv 5 \pmod 9$;
$x \equiv 10 \pmod{35}$.

в) $x \equiv a \pmod{15}$;
$x \equiv b \pmod{16}$.

г) $x \equiv 5 \pmod 7$;
$x \equiv 12 \pmod{15}$;
$x \equiv 18 \pmod{22}$.

2. Решите приведенные ниже системы сравнений:

а) $x \equiv 2 \pmod{13}$;
$x \equiv 5 \pmod{21}$.

б) $x \equiv 7 \pmod{17}$;
$x \equiv 9 \pmod{13}$;
$x \equiv 3 \pmod{12}$.

в) $x \equiv a \pmod{15}$;
$x \equiv b \pmod{16}$.

г) $x \equiv 5 \pmod 9$;
$x \equiv 13 \pmod{11}$;
$x \equiv 7 \pmod 5$.

3. Если в игре с разноцветными шариками в мешочке располагать их по 15 в ряд, в мешочке останутся 4 шарика. Если шарики располагать по 8 в ряд, то 3 шарика останутся в мешочке. Если каждый ряд будет содержать 23 шарика, то 10 шариков останется в мешочке. Какое наименьшее количество шариков первоначально могло находиться в мешочке? Какое было при этом количество рядов по 15, 8 и 23 шарика в каждом?
4. Предположим, имеются две шестерни: одна (шестерня А) — с 25 зубьями, другая (шестерня В) — с 54 зубьями. У шестерни А поврежден зубец, а у шестерни В повреждена секция между двумя зубцами. Шестерни сцеплены вместе, при этом шестерня А находится слева. Шестерня А вращается по часовой стрелке, а шестерня В — против часовой стрелки. Перед началом вращения зубец шестерни А точно захвачен шестерней В, поврежденный зубец шестерни А находится за три зубца до места сцепления, а поврежденная секция — на 20 секций вперед от места сцепления. После какого количества захватов зубьев поврежденный зубец впервые попадет в поврежденную секцию? Как часто после этого поврежденные части будут вновь сталкиваться?
5. Найдите, если это возможно, решение каждой из приведенных ниже систем сравнений:

а) $x \equiv 21 \pmod{36}$; $x \equiv 5 \pmod{8}$.

б) $x \equiv 8 \pmod{12}$; $x \equiv 5 \pmod{9}$; $x \equiv 14 \pmod{15}$.

в) $x \equiv 19 \pmod{49}$; $x \equiv 10 \pmod{14}$.

10.4. СВОЙСТВА ФУНКЦИИ ϕ

Пусть $n = p_1^{\alpha_1} p_2^{\alpha_1} \cdots p_k^{\alpha_k}$ — разложение на простые множители числа n. Каждый положительный делитель числа n либо равен 1, либо делится на p_i при некотором i, и каждое целое число, взаимно простое с n, не имеет ни одного из указанных чисел в качестве делителя. Некоторые свойства числа n зависят от количества целых чисел s, $1 \leq s \leq n$, *не* содержащих ни одно из p_i в качестве делителя. Например, если $n = 40 = 2^3 \cdot 5$, тогда целые числа s, $1 \leq s \leq n$, и их разложения на множители имеют вид

$1 = 1$	**$11 = 11$**	**$21 = 3 \cdot 7$**	**$31 = 31$**
$2 = 2$	$12 = 2^2 \cdot 3$	$22 = 2 \cdot 11$	$32 = 2^5$
$3 = 3$	**$13 = 13$**	**$23 = 23$**	**$33 = 3 \cdot 11$**
$4 = 2^2$	$14 = 2 \cdot 7$	$24 = 2^3 \cdot 3$	$34 = 2 \cdot 17$
$5 = 5$	$15 = 3 \cdot 5$	$25 = 5^2$	$35 = 5 \cdot 7$
$6 = 2 \cdot 3$	$16 = 2^4$	$26 = 2 \cdot 13$	$36 = 2^2 \cdot 3^2$
$7 = 7$	**$17 = 17$**	**$27 = 3^3$**	**$37 = 37$**
$8 = 2^3$	$18 = 2 \cdot 3^2$	$28 = 2^2 \cdot 7$	$38 = 2 \cdot 19$
$9 = 3^2$	**$19 = 19$**	**$29 = 29$**	**$39 = 3 \cdot 13$**
$10 = 2 \cdot 5$	$20 = 2^2 \cdot 5$	$30 = 2 \cdot 3 \cdot 5$	$40 = 2^3 \cdot 5$

Каждое из целых чисел, выделенных жирным шрифтом, не делится ни на 2, ни на

5, и является числом, взаимно простым с $n = 40$. Количество чисел s, $1 \leq s \leq n$, взаимно простых с n, обозначается $\phi(40) = 16$.

ОПРЕДЕЛЕНИЕ 10.16. Пусть $\phi(n)$ — количество положительных целых чисел, меньших n и взаимно простых с n, т.е. $\phi(n)$ — количество приведенных вычетов по модулю n. Функция ϕ называется ***тотиент-функцией Эйлера***, или ***функцией Эйлера*** ϕ.

Из приведенной выше таблицы разложения на множители следует, что

$$\begin{array}{lll} \phi(1) = 1; & \phi(5) = 4; & \phi(9) = 6; \\ \phi(2) = 1; & \phi(6) = 2; & \phi(10) = 4; \\ \phi(3) = 2; & \phi(7) = 6; & \phi(11) = 10; \\ \phi(4) = 2; & \phi(8) = 4; & \phi(12) = 4. \end{array}$$

Любое положительное целое число n может быть выражено с помощью положительных целых чисел, не превосходящих и взаимно простых с каждым делителем числа n. Например, $6 = 2 \cdot 3$ имеет четыре делителя: 1, 2, 3 и 6. Из приведенной выше таблицы следует, что

$$\phi(1) + \phi(2) + \phi(3) + \phi(6) = 1 + 1 + 2 + 2 = 6.$$

Это свойство сформулировано в следующей теореме.

ТЕОРЕМА 10.17. (Гаусс) Если n — положительное целое число, то

$$\sum_{d|n} \phi(d) = n,$$

где делители d являются положительными делителями числа n.

ДОКАЗАТЕЛЬСТВО. Пусть d — положительный делитель числа n. Пусть $C(d)$ — множество положительных целых чисел $1 \leq m \leq n$, для которых НОД$(m, n) = d$. $C(d)$ и $C(d')$ не пересекаются, если $d \neq d'$, поскольку произвольное целое число может иметь с числом n только один наибольший общий делитель. Согласно теореме 7.8, $C(d)$ является также множеством положительных целых чисел m, $1 \leq m \leq n$, для которых НОД$(m/d, n/d) = 1$. Но количество положительных целых чисел, меньших n/d и взаимно простых с n/d, совпадает со значением функции $\phi(n/d)$. Поскольку объединение этих множеств есть множество целых чисел между 1 и n, то $n = \sum_{d|n} \phi(n/d)$. Но для каждого d, которое делит n, имеется соответствующее n/d, которое делит n. Следовательно, $\sum_{d|n} \phi(n/d) = \sum_{d|n} \phi(d) = n$, что и доказывает теорему. ■

ПРИМЕР 10.18. Пусть $n = 12$. Делителями 12 являются 1, 2, 3, 4, 6 и 12. Согласно приведенной выше таблице значений функции Эйлера ϕ

$$\phi(1) + \phi(2) + \phi(3) + \phi(4) + \phi(6) + \phi(12) = 1 + 1 + 2 + 2 + 2 + 4 = 12.$$

Для иллюстрации доказательства теоремы 10.17 в частном случае $d = 1$, 2, 3, 4, 6 и 12 находим, что соответствующие значения n/d равны, соответственно, $n/d = 12$, 6, 4, 3, 2 и 1, так что две упомянутые суммы равны. □

Теперь перейдем к способам вычисления $\phi(n)$ для любого целого положительного числа n. Решить указанную задачу помогут три следующие теоремы.

ТЕОРЕМА 10.19. Если числа m и n — взаимно простые, то

$$\phi(mn) = \phi(m)\phi(n).$$

ДОКАЗАТЕЛЬСТВО. Пусть числа m и n — взаимно простые. Целое число является взаимно простым с mn, тогда и только тогда, когда оно взаимно простое и с m, и с n. Пусть число a — взаимно простое с m и пусть $a < m$. Рассмотрим последовательность $a, a+m, a+2m, \ldots, a+(n-1)m$. Никакие два из этих чисел не сравнимы по модулю n, поскольку, если $a + jm \equiv a + km \pmod{n}$, то $n \mid (jm - km)$. Отсюда $n \mid m(j-k)$. Поскольку числа m и n — взаимно простые, то $n \mid (j-k)$, что невозможно. Следовательно, рассматриваемая последовательность представляет собой полную систему вычетов по модулю n, и каждый элемент данной последовательности сравним по модулю n с целым положительным числом, меньшим n. Следовательно, количество этих элементов, взаимно простых с n, равно $\phi(n)$. Поскольку существуют $\phi(m)$ таких последовательностей, то существует $\phi(m)\phi(n)$ чисел, взаимно простых одновременно и с m, и с n, которые меньше mn и взаимно простые с mn. Поэтому $\phi(mn) = \phi(m)\phi(n)$. ■

Например, пусть $m = 8$ и $n = 15$. Тогда $\phi(8) = 4$, поскольку только 1, 3, 5 и 7 — положительные целые числа, которые меньше 8 и взаимно простые с 8. Также $\phi(15) = 8$, поскольку только 1, 2, 4, 7, 8, 11, 13 и 14 — положительные целые числа, которые меньше 15 и взаимно простые с 15. Следовательно, $\phi(120) = \phi(8)\phi(15) = 32$, что можно проверить непосредственно. В соответствии с утверждением теоремы 10.19 говорят, что ϕ — мультипликативна относительно взаимно простых множителей. В следующей главе рассматриваются другие мультипликативные функции. Теперь покажем, как вычислять $\phi(n)$, когда n представляет собой степень единственного простого числа.

ТЕОРЕМА 10.20. Если p — простое число, то $\phi(p^k) = p^k - p^{k-1}$.

ДОКАЗАТЕЛЬСТВО. Числа, не превышающие p^k и не являющиеся взаимно простыми с p^k, имеют вид $p, 2p, 3p, \ldots, (p^{k-1})p$. Поскольку имеются p^{k-1} таких целых чисел, то существует $p^k - p^{k-1}$ целых чисел, взаимно простых с p^k. Следовательно, $\phi(p^k) = p^k - p^{k-1}$. ■

СЛЕДСТВИЕ 10.21. Целое положительное число p является простым тогда и только тогда, когда $\phi(p) = p - 1$.

ДОКАЗАТЕЛЬСТВО. Если p — простое число, то из теоремы 10.20 непосредственно следует, что $\phi(p) = p - 1$. С другой стороны, если число p не является простым, у него есть делитель d, отличный от p и от 1. Поскольку, по определению, $\phi(p) \leq p - 1$ и d являются одним из $p - 1$ положительных целых чисел, меньших p, имеем $\phi(p) \leq p - 2$, что составляет противоречие. ■

СЛЕДСТВИЕ 10.22. $\phi(2^k) = 2^{k-1}$.

На основе свойства мультипликативности $\phi(mn) = \phi(m)\phi(n)$ для взаимно простых чисел m и n в сочетании с соотношением $\phi(p^k) = p^k - p^{k-1}$ можно для любого целого положительного числа n получить явное выражение для $\phi(n)$, используя разложение n на простые множители. Эта формула приведена в следующей теореме, доказательство которой предоставляется читателю в качестве упражнения.

ТЕОРЕМА 10.23. Если n — целое положительное число с разложением на простые множители вида

$$n = p_1^{\alpha_1} p_2^{\alpha_1} \cdots p_t^{\alpha_t},$$

то

$$\phi(n) = \prod_{i=1}^{t} [p_i^{\alpha_i - 1}(p_i - 1)] = n \prod_{i=1}^{t} \left(1 - \frac{1}{p_i}\right).$$

ПРИМЕР 10.24. Поскольку $n = 40 = 2^3 \cdot 5$, $\phi(40) = 40(1 - 1/2)(1 - 1/5) = 40(1/2)(4/5) = 16$, что совпадает со значением, полученным в одном из примеров в начале этого раздела. Кроме этого, в главе 3 показано, что $n = 39616304 = 2^4 \cdot 7^2 \cdot 13^3 \cdot 23^1$, поэтому

$$\begin{aligned}\phi(39616304) &= 2^3(2-1)7^1(7-1)13^2(13-1)23^0(23-1) = \\ &= 8 \cdot 1 \cdot 7 \cdot 6 \cdot 169 \cdot 12 \cdot 1 \cdot 22 = 14990976.\end{aligned}$$ □

Существуют ограничения, касающиеся количества целых чисел s, $1 \leq s \leq n$, взаимно простых с n. Одну из таких ограничивающих связей устанавливает следующая теорема.

ТЕОРЕМА 10.25. Если целое число n больше 2, то $\phi(n)$ — четное.

ДОКАЗАТЕЛЬСТВО. Если $n = 2^k m$, где m — целое нечетное число и $k > 1$, то $\phi(2^k m) = \phi(2^k)\phi(m) = 2^{k-1}\phi(m)$ и, следовательно, $\phi(n)$ — четное. Если $n = p^k m$, где p — нечетное простое число и числа p^k и m — взаимно простые, тогда $\phi(p^k m) = \phi(p^k)\phi(m) = (p^k - p^{k-1})\phi(m)$. Но $p^k - p^{k-1} = p^{k-1}(p-1)$ и $p - 1$ — четное число, поскольку p — нечетное. Следовательно, $\phi(n)$ — четное. ■

Приведенный ниже результат нами уже упоминался, но здесь он излагается формально.

ТЕОРЕМА 10.26. Если n — целое число, тогда ненулевой класс приведенных вычетов образует группу относительно умножения по модулю n.

ДОКАЗАТЕЛЬСТВО. Если числа a и b — взаимно простые с n, то, очевидно, число ab — также взаимно простое с n, поэтому наш класс вычетов замкнут относительно умножения. Если число a — взаимно простое с n, то сравнение $ax \equiv 1 \pmod{n}$ имеет единственное решение, и, следовательно, существует величина, обратная к a. ■

Пусть число p — простое. Поскольку $\{1, 2, \ldots, p-1\}$ — множество приведенных вычетов по модулю p, то $[1], [2], \ldots, [p-1]$ образуют, как это было показано в разделе 3.6, группу относительно умножения. Следующая теорема показывает, что произведение $[1] \odot [2] \odot \cdots \odot [p-1]$ всех ненулевых классов вычетов всегда является классом вычетов $[p-1] = [-1]$. На языке сравнимости это эквивалентно утверждению, что $1 \cdot 2 \cdot \cdots \cdot (p-1) \equiv -1 \pmod{p}$.

ТЕОРЕМА 10.27. (Уилсон) Целое положительное число p является простым тогда и только тогда, когда $(p-1)! \equiv -1 \pmod{p}$.

ДОКАЗАТЕЛЬСТВО. Если число p — простое, то $p \equiv 0 \pmod{p}$ и $p - 1 \equiv -1 \pmod{p}$. Когда p — простое, ненулевой класс вычетов по модулю p образует группу относительно умножения, так что каждый класс вычетов является спаренным со своим обращением и дает в произведении $[1]$. Таким образом, если $1 \leq u \leq p-1$, то существует единственное целое число u^{-1}, $1 \leq u^{-1} \leq p-1$ такое, что $u \cdot u^{-1} \equiv 1 \pmod{p}$. Имеем $u = u^{-1}$ или $u \neq u^{-1}$. Очевидно, что для $u = 1$ имеем $u^{-1} = 1$, а также $u^2 = uu^{-1} \equiv 1 \pmod{p}$. Если существует целое число a, $1 < a \leq p-1$ такое, что $a^2 \equiv 1 \pmod{p}$, то $a^2 - 1 = (a-1)(a+1) \equiv 0 \pmod{p}$ и $p \mid (a-1)(a+1)$. Таким образом, $p \mid (a-1)$ или $p \mid (a+1)$. Поскольку $a - 1 \neq 0$ и $a < p$, получаем, что $p \nmid (a-1)$. Таким образом, $p \mid (a+1)$, откуда $p \leq a + 1$; и поскольку $a \leq p - 1$, это имеет следствием $a + 1 \leq p$. Получаем, что $p = a + 1$ или $a = p - 1$. Таким образом, для $1 \leq u \leq p-1$ только $u = 1$ и $u = p - 1$ обладают таким свойством, что $u^2 \equiv 1 \pmod{p}$. Итак, получаем $(p-1)! = 1(u_1 u_1^{-1})(u_2 u_2^{-1}) \cdots (u_k u_k^{-1})(p-1) \equiv 1 \cdot 1 \cdot 1 \cdot \cdots \cdot 1 \cdot (p-1) = p - 1 \equiv -1 \pmod{p}$, где u_j — одно из целых чисел 2, 3, ..., $(p-2)$, и $k = (p-3)/2$.

Если число p не является простым, то $p = r \cdot s$, где $1 < r, s < p$. Поскольку $(p-1)!$ делится на r, то $(p-1)! \equiv 0 \pmod{r}$, так что $(p-1)! \not\equiv -1 \pmod{r}$. Поэтому p должно быть простым. ■

Например, пусть $p = 5$. Тогда $(p-1)! = 4! = 24 \equiv -1 \pmod{5}$. Заметим, что в теореме говорится о том, что произведение $(p-1)!$ не может быть сравнимо с -1, если число p не является простым. Посредством теоремы можно проверять простоту числа p, устанавливая, справедливо ли сравнение $(p-1)! \equiv -1 \pmod{p}$. Однако, такой критерий не используется для больших значений p, т.к. вычисление $(p-1)! \pmod{p}$ практически нецелесообразно.

Теперь рассмотрим теорему Уилсона с алгебраической точки зрения. Уже известно, что $Z_p - \{[0]\}$ образует группу относительно умножения. Поэтому каждый ненулевой элемент из Z_p имеет мультипликативную инверсию — обратный элемент относительно умножения. Доказательство приведенной выше теоремы Уилсона показывает, что только $[1]$ и $[p-1]$ совпадают со своими обратными элементами. Следовательно, в произведении $[1][2][3] \cdots [p-1]$ каждый элемент является спаренным со своим обратным элементом, так что $[1][2][3] \cdots [p-1] = [1][p-1] = [p-1]$ или, что то же самое,

$$1 \cdot 2 \cdot 3 \cdot \cdots \cdot (p-1) \equiv p - 1 \pmod{p} \equiv -1 \pmod{p}.$$

Легко показать, что в циклической группе четного порядка существуют только два элемента, которые совпадают со своими обратными элементами. Этими элементами в данном случае являются $[1]$ и $[p-1]$.

Рассматриваемая нами функция ϕ названа в честь Леонарда Эйлера, перу которого принадлежит наибольшее количество математических трудов. Многие из доказанных им теорем встречаются на страницах этой книги. Творческое наследие Эйлера могло бы составить более 75 объемных томов. Ему принадлежат открытия практически во всех областях математики. Только в теории чисел ему принадлежит более 140 оригинальных работ, включая доказательства целого ряда малых теорем Ферма. Он считается основоположником топологии, а также целых разделов математического анализа. Престижную премию Парижской Академии наук, присуждаемую раз в два года, Эйлер получал 12 раз. Многие из ныне существующих систем математических обозначений введены Эйлером.

Леонард Эйлер (Léonard Euler, 1707–1783) был сыном лютеранского священника в Швейцарии. Его отец был его первым учителем. Он хотел, чтобы сын также пошел в духовенство, но Эйлеру посчастливилось стать учеником Иоганна Бернулли (Jean Bernoulli), одного из лучших европейских математиков того времени. Не найдя в Швейцарии условий для научной деятельности, он вместе с другими европейскими математиками уехал работать в Россию, где незадолго до этого была основана Санкт-Петербургская Академия наук. Во время пребывания в России он ослеп на один глаз. Вскоре после прибытия Эйлера в Россию в стране начались политические репрессии, и после 14 лет жизни в России он уехал в Германию, чтобы возглавить математический отдел Берлинской Академии наук. По-видимому, в этот период жизни в Германии на вопрос королевы-матери о том, почему он такой неразговорчивый, Эйлер ответил, что только что приехал из страны, где вешают всякого, кто разговаривает.

Тем не менее, Эйлер оставался весьма уважаемым в России. Когда в 1760 году началась война между Россией и Германией, дом, где жил Эйлер, был разрушен, и когда об этом узнали в России, материальный ущерб был моментально возмещен, а императрица сделала ему подарок. После 25 лет пребывания в Германии, по причине разногласий с королем Фридрихом II, Эйлер вернулся в Россию, приняв великодушное приглашение Екатерины II. Через четыре года после возвращения Эйлер ослеп на второй глаз и оставался слепым последние 17 лет своей жизни, однако занятий математикой не прекращал. В 1771 году в доме Эйлера случился пожар. Его слуга-швейцарец, Петр Гримм, храбро ринулся в пылающий дом и вынес слепого Эйлера. Императрица Екатерина вскоре выстроила Эйлеру новый дом. Говорят, что 18 сентября 1783 года, проведя вечер за вычислением законов для описания подъема воздушных шаров и наметив в общих чертах вычисление орбиты открытой незадолго до того планеты Уран, Эйлер курил трубку и играл со своим внуком, когда его хватил удар. Трубка вывалилась изо рта, он произнес слова: "Я умираю", и эпоха Эйлера закончилась.

Существует целый ряд занимательных историй об Эйлере. Говорят, он мог декламировать слово в слово *Энеиду* Вергилия, хотя читал ее лишь в детстве. Тьебо (Thiébault) рассказывает, что к Российскому двору был пригашен Дидро, который, будучи атеистом, стал распространять свои идеи. Чтобы заставить его замолчать, был придуман план. Эйлер, истинно верующий христианин, подошел к Дидро и провозгласил по-французски: "$(a + b^n)/n = x$, следовательно, Бог существует! Отвечай!" Дидро не знал математики, поэтому молчал. Поднялся хохот, отчего Дидро смутился настолько, что немедленно вернулся во Францию.

Другая история повествует о том, как Эйлер, будучи в почтенном возрасте, слепой, был приглашен княгиней Дашковой посетить ее выступление в связи со вступлением княгини в должность директора Императорской Академии наук в Санкт-Петербурге. Эйлер, в сопровождении сына и внука, приехал вместе с княгиней в ее экипаже. После выступления, в котором Эйлеру была воздана высокая хвала, княгиня села на место, предполагая, что Эйлер займет почетное место рядом с ней. Высокомерный профессор Штелин занял это место до того, как Эйлера к нему подвели. Дашкова обратилась к Эйлеру с предложением занять любое место, сделав его тем самым почетным. Этот жест княгини произвел благоприятное впечатление на всех присутствовавших.

■ УПРАЖНЕНИЯ

1. Найдите такие целые положительных числа m и n, что $\phi(mn) \neq \phi(m)\phi(n)$.
2. Докажите, что если число p — простое и $p > 2$, то $(p-2)! \equiv 1 \pmod{p}$.
3. Докажите, что $1 \cdot 2 \cdot 3 \cdot \cdots \cdot 1007 \equiv 1 \pmod{1009}$.
4. Пусть
$$S(n) = \sin\left[\pi \cdot \frac{(n-1)!+1}{n}\right].$$
Покажите, что число n — простое тогда и только тогда, когда $S(n) = 0$.
5. Докажите теорему 10.23.
6. Постройте таблицу значений $\phi(n)$ при $1 \leq n \leq 50$.
7. Вычислите значение $\phi(2025)$.
8. Докажите, что если число n — составное и $n > 4$, то $(n-1)! \equiv 0 \pmod{n}$.
9. Докажите, что число n — простое тогда и только тогда, когда n делит $(n-1)!+1$.

10.5. ПОРЯДОК ЦЕЛОГО ЧИСЛА

В этом разделе будут рассмотрены такие целые числа j, что $a^j \equiv 1 \pmod{m}$. В частности, нас будет интересовать наименьшее из этих положительных чисел j.

ТЕОРЕМА 10.28. (Эйлер) Если m — целое положительное число и НОД(a, m)=1, то $a^{\phi(m)} \equiv 1 \pmod{m}$.

ДОКАЗАТЕЛЬСТВО. Пусть m — целое положительное число и a — число, взаимно простое с m. Если $\{x_1, x_2, \ldots, x_k\}$ — приведенная система вычетов по модулю m, то, поскольку числа a и m — взаимно простые, $\{ax_1, ax_2, \ldots, ax_k\}$ также является системой приведенных вычетов. Следовательно, каждое x_i сравнимо по модулю m только с одним ax_j. Поэтому

$$x_1 x_2 \cdots x_k \equiv a x_1 a x_2 \cdots a x_k \pmod{m}$$

или

$$a^{\phi(m)} x_1 x_2 \cdots x_k \equiv x_1 x_2 \cdots x_k \pmod{m},$$

и, поскольку числа m и $x_1 x_2 \cdots x_k$ взаимно просты, то на x_i можно сокращать, что дает $a^{\phi(m)} \equiv 1 \pmod{m}$. ■

Например, при $a = 3$, $m = 4$ имеем $\phi(4) = 2$, так что $3^2 = 9 \equiv 1 \pmod 4$.

Если в теореме 10.28 m — простое число, то каждое целое положительное число, которое меньше m, является взаимно простым с m, так что $\phi(m) = m - 1$. Случай простого числа m рассмотрен как следствие к теореме 10.20. Таким образом, как частный случай, справедлива следующая теорема.

ТЕОРЕМА 10.29. (Малая теорема Ферма) Если p — простое число, то для каждого такого целого числа a, что $0 < a < p$, имеем $a^{p-1} \equiv 1 \pmod p$.

Например, если $p = 7$, то $p - 1 = 6$. В таком случае шестая степень каждого целого положительного числа, которое меньше $p = 7$, должна быть сравнима с 1 по модулю 7:

$$1^6 = 1 \equiv 1 \pmod 7,$$
$$2^6 = 64 \equiv 1 \pmod 7,$$
$$3^6 = 729 \equiv 1 \pmod 7,$$
$$4^6 = 4096 \equiv 1 \pmod 7,$$
$$5^6 = 15625 \equiv 1 \pmod 7,$$
$$6^6 = 46656 \equiv 1 \pmod 7,$$
$$7^6 = 117649 \equiv 0 \pmod 7.$$

Утверждение, обратное малой теореме Ферма, неверно. Например, $3^{90} \equiv 1 \pmod{91}$ однако, $91 = 7 \cdot 13$ — составное число. С другой стороны, если p — целое положительное число и $0 < a < p$ таково, что $a^{p-1} \not\equiv 1 \pmod p$, тогда p не может быть простым. Таким образом, малая теорема Ферма содержит частичный критерий простоты числа, поскольку с ее помощью можно показать, что целое положительное число не является простым без определения нетривиального делителя числа p. Составные положительные числа n таковы, что $a^{n-1} \equiv 1 \pmod n$ для некоторого a, $1 < a < n$, до определенной степени схожи с простыми числами; по этой причине такого рода составное число n называют ***псевдопростым*** числом по основанию a. Таким образом, число $n = 91$ — псевдопростое по основанию $a = 3$. Однако, если выбрать $a = 2$, то получим $2^{90} \equiv 64 \not\equiv 1 \pmod{91}$; то есть, число $n = 91$ не является псевдопростым по основанию 2. Итак, 91 — псевдопростое число по основанию 3, но не является псевдопростым по основанию 2.

ОПРЕДЕЛЕНИЕ 10.30. Пусть n — целое положительное число и a — целое число такое, что НОД$(a, n) = 1$. ***Порядком*** числа a по модулю n называется наименьшее целое положительное число k такое, что $a^k \equiv 1 \pmod n$. Это число обозначается через $\operatorname{ord}_n a$.

ТЕОРЕМА 10.31. Пусть n — целое положительное число, НОД$(a, n) = 1$ и $k = \operatorname{ord}_n a$. Тогда

а) если $a^m \equiv 1 \pmod n$, где m — целое положительное число, то $k \mid m$;
б) $k \mid \phi(n)$;
в) для целых r и s, $a^r \equiv a^s \pmod n$ тогда и только тогда, когда $r \equiv s \pmod k$;

г) никакие два из целых чисел $a, a^2, a^3, \ldots, a^k$ не являются сравнимыми по модулю k;

д) если m — целое положительное число, то порядок числа a^m по модулю n равен $\dfrac{k}{\text{НОД}(k,m)}$;

е) порядок числа a^m по модулю n равен k тогда и только тогда, когда числа m и k — взаимно простые.

ДОКАЗАТЕЛЬСТВО.

а) Если $a^m \equiv 1 \pmod{n}$ для целого положительного числа m, то согласно алгоритму деления $m = kq + r$, где $0 \leq r < k$. Следовательно, $a^m = a^{kq+r} = a^{kq}a^r$, так что $a^r \equiv 1 \pmod{n}$. Но это противоречит определению порядка числа a, если не $r = 0$. Следовательно, $k \mid m$.

б) Поскольку по теореме 10.28 имеем $a^{\phi(n)} \equiv 1 \pmod{n}$, из части (а) следует, что $k \mid \phi(n)$.

в) Предположим, что $r > s$. Поскольку числа a и n — взаимно простые, $a^r \equiv a^s \pmod{n}$ тогда и только тогда, когда $a^{r-s} \equiv 1 \pmod{n}$; следовательно, согласно части (а) k делит $r - s$, и $r \equiv s \pmod{k}$.

г) Справедливость утверждения непосредственно следует из части (в).

д) Пусть $d = \text{НОД}(k, m)$, так что $k = ud$ и $m = vd$. $(a^m)^{k/\,\text{НОД}(k,m)} = (a^m)^{ud/d} = a^{um} = a^{uvd} = a^{(ud)v} = a^{kv} \equiv 1 \pmod{n}$. Допустим, что число t таково, что $(a^m)^t \equiv 1 \pmod{n}$. Тогда $a^{mt} \equiv 1 \pmod{n}$, так что $k \mid mt$, потому что $\text{ord}_n\, a = k$. Следовательно, $ud \mid vdt$ и, поскольку числа u и v — взаимно простые, $u \mid t$. В силу того, что $k = ud$, $(k/d) = k/\,\text{НОД}, (k, m)$ делит t, поэтому $k/\,\text{НОД}(k, m)$ и является порядком числа a^m по определению.

е) Это утверждение следует непосредственно из части (д).

■

ПРИМЕР 10.32. Для иллюстрации теоремы 10.31 предположим, что $n = 14 = 2 \cdot 7$, так что $\phi(n) = (2-1)(7-1) = 6$. Первичная приведенная система вычетов для $n = 14$ есть множество $\{1, 3, 5, 9, 11, 13\}$. Рассмотрим приведенную ниже таблицу вычетов для степеней числа $a = 5$,

m	$[[a^m]]_n$	m	$[[a^m]]_n$
1	5	8	11
2	11	9	13
3	13	10	9
4	9	11	3
5	3	12	1
6	1	13	5
7	5		

из которой видно, что после $m = 6$ идет повторение одной и той же схемы. Таким образом, $k = \text{ord}_{14}\, 5 = 6$. Для $m = 12$, $a^m = 5^{12} \equiv 1 \pmod{14}$ и $k \mid m$, что согласуется с утверждением теоремы 10.31(а). Также $\text{ord}_{14}\, 5 \mid \phi(14)$, поскольку $6 \mid 6$ [теорема 10.31(б)]. Кроме того, $2 \equiv 8 \equiv 14 \pmod{6}$ и $5^2 \equiv 5^8 \equiv 5^{14} \equiv 11$

(mod 14) [теорема 10.31(в)]. Согласно таблице никакие два из чисел 5^1, 5^2, 5^3, 5^4, 5^5 и 5^6 не являются сравнимыми по модулю 14 [теорема 10.31(г)]. Поскольку $\mathrm{ord}_n\, b \mid \phi(n)$ для любого целого числа b и $\phi(n) = 6$ для $n = 14$, порядок каждого b в $\{1, 3, 5, 9, 11, 13\}$ можно легко подсчитать, как это было сделано для $a = 5$.

b	$\mathrm{ord}_n\, b$
1	1
3	6
5	6
9	3
11	3
13	2

Если $m = 4$, то $5^m \equiv 9 \pmod{14}$, но $\mathrm{ord}_{14}\, 5/\mathrm{НОД}(\mathrm{ord}_{14}\, 5, 4) = 6/\mathrm{НОД}(6, 4) = 6/2 = 3$. Согласно таблице порядков $\mathrm{ord}_{14}\, 5^4 = 3$ [теорема 10.31(д)].

Только $b = 3$ и $b = 5$ имеют порядок 6 по модулю 14. Показателями степени m в таблице, приведенной выше, определяющими значение a^m, которое сравнимо либо с числом 3, либо с числом 5, являются $m = 1, 5, 7, 11$, и 13. Это только такие значения m, которые являются взаимно простыми с числом $n = 14$ [теорема 10.31(е)]. □

ТЕОРЕМА 10.33. Если $\mathrm{НОД}(a, n) = \mathrm{НОД}(b, n) = 1$ и $\mathrm{ord}_n\, a$ является взаимно простым с $\mathrm{ord}_n\, b$, то $\mathrm{ord}_n\,(ab) = (\mathrm{ord}_n\, a) \cdot (\mathrm{ord}_n\, b)$.

ДОКАЗАТЕЛЬСТВО. Пусть $\mathrm{ord}_n\, a = R$ и $\mathrm{ord}_n\, b = S$, тогда

$$(ab)^{RS} = a^{RS}b^{RS} = (a^R)^S(b^S)^R = 1 \cdot 1 \equiv 1 \pmod{n}.$$

По теореме 10.31, $\mathrm{ord}_n\,(ab) \mid RS$. Поскольку числа R и S — взаимно простые, то существуют целые числа r и s, для которых $\mathrm{ord}_n\,(ab) = rs$, $r \cdot w = R$ и $s \cdot x = S$. Покажем теперь, что $r = R$ и $s = S$. По определению r и s,

$$\begin{aligned}(ab)^{rs} &= a^{rs}b^{rs} \equiv 1 \pmod{n};\\ (a^{rs}b^{rs})^w &\equiv 1^w = 1 \pmod{n};\\ (a^{rw})^s \cdot (b^{rw})^s &\equiv 1 \pmod{n}.\end{aligned}$$

Но поскольку $a^{rw} \equiv 1 \pmod{n}$ и $rw = R$, имеем

$$b^{Rs} \equiv 1 \pmod{n}.$$

По теореме 10.31(а) имеем $\mathrm{ord}_n\, b \mid Rs$ или, что то же самое, $S \mid Rs$. В силу того, что $\mathrm{НОД}(R, S) = 1$, имеем $S \mid s$, но также $s \mid S$, так что $S = s$. Аналогично, $R = r$. Таким образом, $\mathrm{ord}_n\,(ab) = (\mathrm{ord}_n\, a) \cdot (\mathrm{ord}_n\, b)$. ■

ПРИМЕР 10.34. Если $n = 11$, то все приведенные вычеты являются взаимно простыми с n. Таблица порядков по модулю 11 имеет вид

Вычет	Порядок	Вычет	Порядок
1	1	6	10
2	10	7	10
3	5	8	10
4	5	9	5
5	5	10	2

Если $a = 3$ и $b = 10$, то $ab = 30 \equiv 8 \pmod{11}$. Таким образом, $\mathrm{ord}_{11}(ab) = \mathrm{ord}_{11}(30) = \mathrm{ord}_{11}\, 8 = 10 = (\mathrm{ord}_{11}\, 3) \cdot (\mathrm{ord}_{11}\, 10)$. Безусловно, НОД$(3, 11) =$ НОД$(10, 11) = 1$, $\mathrm{ord}_{11}\, 3 = 5$ и $\mathrm{ord}_{11}\, 10 = 2$ — взаимно простые. Заметим, что если $a = 3$ и $c = 7$, то $\mathrm{ord}_{11}\, 3 = 5$ не является взаимно простым с $\mathrm{ord}_{11}\, 7 = 10$. В этом случае $\mathrm{ord}_{11}(ac) = \mathrm{ord}_{11}\, 21 = \mathrm{ord}_{11}\, 10 = 2 \neq (\mathrm{ord}_{11}\, 3) \cdot (\mathrm{ord}_{11}\, 7) = 50$. □

ПРИМЕР 10.35. Порядок $a = 5$ по модулю $n = 14$ был получен в примере 10.32 при вычислении a^m для $m = 1, 2, 3, \ldots, \phi(n)$ вплоть до нахождения того, что $a^m \equiv 1 \pmod{n}$. Из теоремы 10.31(б) следует, что порядок a по модулю n должен делить $\phi(n)$; поэтому вместо проверки по одному каждого m, $1 \leq m \leq \phi(n)$, тестируются только те значения m, которые делят $\phi(n)$. Для $n = 14$ $\phi(n) = 6$, единственными положительными делителями которого являются 1, 2, 3 и 6. В данном случае работа по определению $\mathrm{ord}_{14}\, 5$ сокращается незначительно, поскольку в примере 10.32 значения $m = 4$ и 5 были уже протестированы.

Однако, для $n = 58$ и $a = 25$, используя теорему 10.31(б), легко получаем $\mathrm{ord}_{58}\, 25 = 7$. $\phi(58) = \phi(2 \cdot 29) = (2 - 1)(29 - 1) = 28 = 2^2 \cdot 7$. Единственными положительными делителями числа $2^2 \cdot 7$ являются 1, 2, 4, 7, 14 и 28. Приведенную ниже таблицу легко построить:

m	$[[25^m]]_{58}$
1	25
2	45
4	53
7	1

Поэтому $\mathrm{ord}_{58}\, 25 = 7$, и нет необходимости проверять $n = 14$ и 28. □

Результаты, полученные в теоремах 10.31 и 10.36, приводят к формулировке критерия, названного критерием простоты числа Лукаса.

ТЕОРЕМА 10.36. (Лукас) Если n — целое положительное число, и существует такое целое число a, что

$$a^{n-1} \equiv 1 \pmod{n}$$

и

$$a^{\frac{n-1}{p}} \not\equiv 1 \pmod{n}$$

для каждого простого числа p, которое делит $n - 1$, тогда n — простое число.

ДОКАЗАТЕЛЬСТВО. Из соотношения $a^{n-1} \equiv 1 \pmod{n}$ следует, что НОД$(a, n) = 1$, а также, по теореме 10.31(а), что $\mathrm{ord}_n\, a \mid (n-1)$. Если число p — простое и такое, что $p \mid (n-1)$, тогда из сравнимости $a^{(n-1)/p} \not\equiv 1 \pmod{n}$ следует, что $\mathrm{ord}_n\, a \nmid [(n-1)/p]$, потому что, если $\mathrm{ord}_n\, a \mid [(n-1)/p]$, это противоречило бы $a^{\mathrm{ord}_n\, a} \equiv 1 \pmod{n}$. Но $\mathrm{ord}_n\, a \mid (n-1)$ и $\mathrm{ord}_n\, a \nmid [(n-1)/p]$ для всех p, которые делят $n-1$, откуда следует, что $\mathrm{ord}_n\, a = n-1$. По теореме 10.31(б), $\phi(n) = n-1$. Поэтому, согласно следствию 10.21, n — простое число. ■

Для того чтобы использовать критерий Лукаса для проверки n, необходимо уметь раскладывать на множители число $n-1$, что само по себе может представлять трудности. Более того, требуется находить соответствующее a. Целое число a, введенное в теореме 10.36, называется ***примитивным корнем*** числа n. Используя критерий с числом $a = 7$, можно показать, что число Мерсенна $n = 2^{31} - 1$ является простым, поскольку $n - 1 = 2 \cdot 3^2 \cdot 7 \cdot 11 \cdot 31 \cdot 151 \cdot 331$. Смотрите упражнения в конце данного раздела.

Если НОД$(a, n) = 1$ и число n — простое, теорема Ферма утверждает, что $a^{n-1} \equiv 1 \pmod{n}$. Ее обобщение, теорема Эйлера, для любого положительного числа n дает $a^{\phi(n)} \equiv 1 \pmod{n}$. Исключая эти и некоторые другие случаи, вычисление a^e по модулю n или, более точно, вычисление $[[a^e]]_n$, т.е. остатка от деления a^e на n, для большого значения e может представлять значительные трудности, поскольку само вычисление a^e в таких случаях и деление его на n практически нецелесообразно.

В предыдущих примерах мы использовали запись $e = e_1 + e_2 + \cdots + e_k$ с соответствующими e_i, вычисляли $[[a^{e_i}]]_n$, перемножали результаты и приводили произведение по модулю n. Этот метод работает, поскольку

$$[[st]]_n = [[\,[[s]]_n \cdot [[t]]_n\,]]_n\,.$$

При этом значения e_i были выбраны специальным образом.

Более эффективный алгоритм вычисления подобен рассмотренному выше, но использует двоичное представление показателя степени e, т.е. представление на основе числа 2. А именно,

$$\begin{aligned} e &= b_m 2^m + b_{m-1} 2^{m-1} + \cdots + b_1 2^1 + b_0 = \\ &= [b_m b_{m-1} \cdots b_1 b_0]_{\textit{двоичное}}\,, \end{aligned}$$

где $b_i = 0$ или 1 и $b_m = 1$. Таким образом,

$$a^e = a^{b_m 2^m + b_{m-1} 2^{m-1} + \cdots + b_1 2^1 + b_0}.$$

Если в выражении для показателя степени сделать перегруппировку с тем, чтобы сократить количество сомножителей, получим представление в соответствии с правилом Горнера:

$$e = (\cdots((b_m \cdot 2 + b_{m-1}) \cdot 2 + b_{m-2}) \cdot 2 + \cdots + b_1) \cdot 2 + b_0,$$

так что

$$a^e = (\cdots((a^{b_m \cdot 2} \cdot a^{b_{m-1}})^2 \cdot a^{b_{m-2}})^2 \cdot \ldots \cdot a^{b_1})^2 \cdot a^{b_0}.$$

Мы можем найти значение $[[a^e]]_n$, последовательно вычисляя выражения в скобках в порядке их вложения, начиная с внутренних, и приводя каждое произведение по модулю n. Таким образом, для $e = [b_m b_{m-1} \cdots b_1 b_0]_{\text{двоичное}}$, начинаем с

$$p_m = [[a]]_n .$$

Затем для $k = m-1, m-2, \ldots, 2, 1$ и 0 вычисляем

$$p_k = \begin{cases} [[p_{k+1}^2]]_n & \text{если } b_k = 0; \\ [[p_{k+1}^2 \cdot a]]_n & \text{если } b_k = 1. \end{cases}$$

Окончательным результатом является $p_0 = [[a^e]]_n$. Более подробно, начиная с $p_m = [[a]]_n$, получаем следующее произведение p_k, возводя в квадрат предыдущее произведение и приводя полученное по модулю n, когда $b_k = 0$; возводя в квадрат предыдущее произведение, умножая его на a и приводя полученное по модулю n, когда $b_k = 1$. Алгоритм работает, поскольку если a возводится в квадрат k раз, результатом будет a^{2^k}; а если $a^{2^k} b$ возводится в квадрат j раз, в результате получится $a^{2^{k+j}} b^{2^j}$.

ПРИМЕР 10.37. Предположим, что необходимо вычислить $[[3^{103}]]_{41}$. Поскольку

$$\begin{aligned} 103 &= 2^6 + 2^5 + 2^2 + 2^1 + 1 = \\ &= 1100111 \end{aligned}$$

и $m = 6$, получаем

k	b_k	$p_k = [[p_{k+1}^2 \cdot a^{b_k}]]_n$
6	1	$3 \equiv 3$
5	1	$3^2 \cdot 3 = 27$
4	0	$(27)^2 = 729 \equiv 32$
3	0	$(32)^2 = 1024 \equiv 4$
2	1	$(40)^2 \cdot 3 = 4800 \equiv 3$
1	1	$(3)^2 \cdot 3 = 27$
0	1	$(27)^2 \cdot 3 = 2187 \equiv 14$

Поэтому $[[3^{103}]]_{41} = 14$. Используя сравнимость по модулю 41 специальным образом, получаем

$$\begin{aligned} 3^{10} &= 57049 \equiv 9; \\ 3^{50} &= (3^{10})^5 \equiv 9^5 = 59049 \equiv 9; \\ 3^{103} &= 3^{50} \cdot 3^{50} \cdot 3^3 \equiv 9 \cdot 9 \cdot 27 = 2187 \equiv 14. \end{aligned}$$

□

■ УПРАЖНЕНИЯ

1. Докажите малую теорему Ферма в такой формулировке. Если p — простое число и $a \not\equiv 0 \pmod{p}$, то $a^{p-1} \equiv 1 \pmod{p}$.

2. Докажите, что если НОД$(a, m) = 1$, то сравнение $ax \equiv b \pmod{m}$ имеет решение $x \equiv a^{\phi(m)-1}b \pmod{m}$.

3. Пусть p — простое число, $p > 2$ и $J \equiv 0 \pmod{(p-1)}$. Докажите, что

$$1^J + 2^J + 3^J + \cdots + (p-1)^J \equiv -1 \pmod{p}.$$

4. Пусть НОД$(m, n) = 1$ и пусть множества $\{r_1, r_2, \ldots, r_m\}$ и $\{s_1, s_2, \ldots, s_n\}$ — полные системы вычетов по модулю m и n соответственно. Докажите, что множество mn целых чисел $\{n \cdot r_i + m \cdot s_j : 1 \leq i \leq m \text{ и } 1 \leq j \leq n\}$ представляет собой полную систему вычетов по модулю mn.

5. Восполните все детали доказательства малой теоремы Ферма, используя формулу бинома Ньютона и индукцию по a.

6. Используйте метод упражнения 2 для решения таких сравнений:

а) $5x \equiv 8 \pmod{11}$; **б)** $7x \equiv 8 \pmod{25}$; **в)** $9x \equiv 13 \pmod{25}$.

7. Докажите, что для каждого простого числа p

$$(a+b)^p \equiv a^p + b^p \pmod{p}.$$

8. Докажите теорему, обратную к теореме 10.31(а): если n — целое положительное число, НОД$(a, n) = 1$, $k = \operatorname{ord}_n a$ и $k \mid m$, то $a^m \equiv 1 \pmod{m}$.

9. Определите $\operatorname{ord}_n a$ для $1 \leq a \leq n-1$, если

а) $n = 9$; **б)** $n = 20$; **в)** $n = 27$.

10. Покажите, что

а) $b^{10} - 1$ делится на 11, если b и 11 — взаимно простые числа;

б) $b^{10k} - 1$ делится на 11, если НОД$(b, 11) = 1$;

в) $b^7 - b$ делится на 42 при любом целом b.

11. Покажите, что

а) $7^4 \equiv 1 \pmod{5}$;

б) $7^4 \equiv 1 \pmod{2}$;

в) $7^4 \equiv 1 \pmod{10}$;

г) $7^{4k} \equiv 1 \pmod{10}$ для любого положительного целого числа k.

Что представляет собой последний десятичный разряд числа 7^{4000}?

12. Покажите, что

а) $7^{20} \equiv 1 \pmod{25}$; **б)** $7^2 \equiv 1 \pmod{4}$;

в) $7^{20} \equiv 1 \pmod{4}$; **г)** $7^{20} \equiv 1 \pmod{100}$.

Что представляет собой последний десятичный разряд числа 7^{500}?

13. Подсчитайте следующие вычеты :

а) $\left[\left[3^{275}\right]\right]_{100}$ **б)** $\left[\left[6^{5000}\right]\right]_{1000}$

в) $\left[\left[11^{24681}\right]\right]_{83}$ **г)** $\left[\left[3497^{100000}\right]\right]_{1234}$

д) $\left[\left[7^{2\cdot 3^2 \cdot 7 \cdot 11 \cdot 31 \cdot 331}\right]\right]_{2^{31}-1} = \left[\left[7^{14221746}\right]\right]_{2147483647}$ (Для выполнения этого задания вам, вероятно, потребуется использовать компьютер.)

14. Покажите при помощи критерия Лукаса, что следующие числа являются простыми:

а) 37; **б)** 199.

15. Докажите следующие утверждения.

а) Пусть n — нечетное положительное целое число. Если существует такое положительное целое число a, что

(i) $a^{\frac{n-1}{2}} \equiv -1 \pmod{n}$ и

(ii) $a^{\frac{n-1}{p}} \not\equiv 1 \pmod{n}$ для каждого нечетного простого числа p, которое делит $(n-1)$,

то n — простое число.

б) Воспользовавшись критерием части (а) вместо критерия Лукаса, докажите, что целые числа в упражнении 14(а) и (б) являются простыми.

16. Докажите, что если $F(n) = 2^{2^n} + 1$ — число Ферма и существует целое положительное целое число a такое, что $a^{2^{2^n}} \equiv 1 \pmod{F(n)}$ и $a^{2^{(2^n-1)}} \not\equiv 1 \pmod{F(n)}$, то $F(n)$ — простое число.

17. Используйте упражнение 16 для доказательства, что приведенные ниже целые числа являются простыми:

а) $F(3) = 257$; **б)** $F(4) = 65537$.

ГЛАВА

11

НЕКОТОРЫЕ СПЕЦИАЛЬНЫЕ ВОПРОСЫ ТЕОРИИ РЕКУРСИИ

11.1. ОДНОРОДНЫЕ ЛИНЕЙНЫЕ РЕКУРРЕНТНЫЕ ОТНОШЕНИЯ

В главе 5 было начато рассмотрение в общих чертах методов исключения рекурсии, или решения рекурсивных функций. Под решением рекурсивной функции мы понимали следующее: рассматривали функцию $a(n)$ или a_n, заданную рекурсивно на множестве положительных или неотрицательных целых чисел, и заменяли ее функцией, заданной в явном виде, когда функция a_n была выражена непосредственно как функция от n. Например, функция, заданная рекурсивно согласно соотношениям

$$a_1 = 1,$$
$$a_k = a_{k-1} + k,$$

может быть задана непосредственно в виде

$$a_n = \frac{n(n+1)}{2}.$$

Использовался метод, который заключался в вычислении функции a_n для нескольких первых значений n из ее области определения с последующим поиском выражения для непосредственного задания a_n как функции n. Вообще, решение произвольных рекуррентных отношений представляет большие трудности, иногда фактически непреодолимые.

В этом разделе рассматривается широкий класс рекурсивных функций, для которых решения известны.

Сначала определим класс рекурсивных функций, известный как класс линейных рекурсивных функций.

ОПРЕДЕЛЕНИЕ 11.1. Рекурсивное отношение вида

$$a_n = c_1(n)a_{n-1} + c_2(n)a_{n-2} + c_3(n)a_{n-3} + \cdots + c_p(n)a_p + f(n)$$

называется ***линейным рекуррентным отношением порядка p.***

ОПРЕДЕЛЕНИЕ 11.2. Линейное рекуррентное отношение вида

$$a_n = c_1(n)a_{n-1} + c_2(n)a_{n-2} + c_3(n)a_{n-3} + \cdots + c_p(n)a_{n-p}$$

называется ***линейным однородным рекуррентным отношением порядка p.***

Рассматриваемое отношение названо линейным рекуррентным отношением, поскольку показатель степени каждого a_i равен единице. Другими словами, ни одно из a_i не возводится в какую-либо степень, кроме первой. Таким образом, отношение $a_n = 3a_{n-1}^3 + 4a_{n-2}$ не будет линейным, т.к. a_{n-1} возведено в третью степень. Однако, отношение $a_n = 3n^3a_{n-1} + na_{n-2}$ линейно. К сожалению, рассматриваемое множество линейных рекуррентных отношений необходимо ограничить еще в большей степени. Дополнительно требуется, чтобы коэффициенты $c_i(n)$ для каждого i были константами.

ОПРЕДЕЛЕНИЕ 11.3. Линейное рекурсивное отношение вида

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p} + f(n), \qquad c_p \neq 0$$

с постоянными коэффициентами c_i при $1 \leq i \leq p$ называется ***линейным рекуррентным отношением с постоянными коэффициентами порядка p.***

Далее будем рассматривать некоторые линейные рекурсии такого вида, но даже при указанных выше ограничениях возможности решения задачи будут зависеть от выбора функции $f(n)$. Линейные рекуррентные отношения с постоянными коэффициентами общего вида рассматриваются в следующем разделе. Но прежде, чем решать задачи такого типа, требуется ввести еще одно ограничение на рассматриваемый класс линейных рекурсий.

ОПРЕДЕЛЕНИЕ 11.4. Линейное рекуррентное отношение вида

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p}, \qquad c_p \neq 0$$

с постоянными коэффициентами c_i при $1 \leq i \leq p$ называется ***линейным однородным рекуррентным отношением с постоянными коэффициентами порядка p.***

Такое отношение является частным случаем линейного рекуррентного отношения с постоянными коэффициентами, когда $f(n) = 0$. Один из известных примеров — последовательность Фибоначчи: $1, 1, 2, 3, 5, 8, 13, 21, \ldots$, в которой каждый элемент после первых двух равен сумме двух предшествующих элементов последовательности. Рекурсивно это условие определено следующим образом.

$$\text{Фиб}(1) = 1;$$
$$\text{Фиб}(2) = 1;$$
$$\text{Фиб}(n) = \text{Фиб}(n-1) + \text{Фиб}(n-2) \text{ для } n > 2.$$

Таким образом, в этом случае $p = 2$, $c_1 = 1$ и $c_2 = 1$. Далее в разделе приводится решение этой рекурсивной функции.

Для начала давайте рассмотрим случай $p = 1$. Допустим, что рекурсивная функция имеет вид

$$a_0 = c;$$
$$a_n = ba_{n-1} \qquad \text{при } n > 0.$$

Отсюда получаем последовательность

$$\begin{aligned} a_0 &= c \\ a_1 &= bc \\ a_2 &= b^2c \\ a_3 &= b^3c \\ a_4 &= b^4c \\ &\vdots \\ a_n &= b^nc, \end{aligned}$$

в которой значение c определяется через значение, присвоенное a_0. Легко заметить, что отношение $a_n = ba_{n-1}$ имеет решение вида $a_n = b^nc$. Далее убедимся, что для любого $n > 1$ решение можно представить в виде $a_n = b^nc$. Но если предположить, что мы с самого начала рассматривали рекурсивную функцию вида

$$a_1 = c;$$
$$a_n = ba_{n-1} \qquad \text{при } n > 1,$$

то получим последовательность вида

$$\begin{aligned} a_1 &= c \\ a_2 &= bc \\ a_3 &= b^2c \\ a_4 &= b^3c \\ a_5 &= b^4c \\ &\vdots \\ a_n &= b^{n-1}c. \end{aligned}$$

Можно ли по-прежнему использовать запись $a_n = c'b^n$, если это действительно существенно? Ответ утвердительный, поскольку $a_n = b^{n-1}c = b^n(c/b)$. Если положить $c' = (c/b)$, то $a_n = c'b^n$. Допустим, мы начали действовать иначе и, в силу таинственных причин, например, посоветовавшись с оракулом, предположили, что a_n имеет вид $a_n = r^n$ для некоторого значения r. Тогда, учитывая соотношение $a_n = ba_{n-1}$, получаем $r^n = br^{n-1}$. Если разделить обе части равенства на r^{n-1}, имеем $r = b$, так что $a_n = b^n$. Заметим далее, что если умножить

обе части равенства $r^n = br^{n-1}$ на константу c, получим $cr^n = bcr^{n-1}$, так что если $a_n = r^n$ есть решение для $a_n = ba_{n-1}$, то решением является и $a_n = cr^n$. Поэтому получаем, что общее решение имеет вид cb^n. Но в таком случае, задав значение a_1, можно определить значение c.

Например, если

$$a_0 = 4;$$
$$a_n = 5a_{n-1} \qquad \text{при } n > 1,$$

то $a_n = c5^n$. Но $a_0 = c5^0 = 4$, поэтому $c = 4$. Вообще говоря, отношение $a_n = ba_{n-1}$ имеет решение cb^n, где $c = a_0$.

Если же

$$a_1 = 4;$$
$$a_n = 5a_{n-1} \qquad \text{при } n > 1$$

и $a_n = c5^n$, то $a_1 = c5^1 = 4$ и $c = 4/5$.

Причиной морочить читателю голову очевидными вещами и придавать им вид непонятных, рассматривая случай $p = 1$, является переход к обсуждению ситуации $p = 2$, когда

$$a_n = c_1a_{n-1} + c_2a_{n-2}.$$

Опять предположим, что имеется решение вида $a_n = r^n$, и в этом случае получаем

$$r^n = c_1r^{n-1} + c_2r^{n-2}.$$

Если разделить обе части равенства на r^{n-2}, получим

$$r^2 = c_1r + c_2.$$

Таким образом, величину r можно найти, решая квадратное уравнение $r^2 = c_1r + c_2$. Это характеристическое уравнение описывает ***характеристический многочлен*** $r^2 - c_1r - c_2$ отношения $a_n = c_1a_{n-1} + c_2a_{n-2}$.

Например, предположим, что $a_n = a_{n-1} + 6a_{n-2}$. Тогда характеристический многочлен задает характеристическое квадратное уравнение $r^2 = r - 6$ или $r^2 - r - 6 = 0$. Раскладывая на множители, получаем $(r - 3)(r + 2) = 0$, так что решениями характеристического уравнения являются $r = -2$ и $r = 3$. Поскольку решения данного рекуррентного отношения имеют вид $a_n = r^n$, где r — решение уравнения $r^2 - r - 6 = 0$, то отсюда следует, что $a_n = 3^n$ и $a_n = (-2)^n$ — решения рассматриваемого отношения.

Предположим, найдены два различных действительных решения уравнения $r^2 = c_1r + c_2$, например, $r = a$ и $r = b$, так что для уравнения $a_n = c_1a_{n-1} + c_2a_{n-2}$ имеем решения $a_n = a^n$ и $a_n = b^n$, поэтому $a^n = c_1a^{n-1} + c_2a^{n-2}$ и $b^n = c_1b^{n-1} + c_2b^{n-2}$. Но если умножить $a^n = c_1a^{n-1} + c_2a^{n-2}$ на константу c, то получим $ca^n = c_1ca^{n-1} + c_2ca^{n-2}$, так что $a_n = ca^n$ также является решением для $a_n = c_1a_{n-1} + c_2a_{n-2}$.

Далее, если сложить $a^n = c_1a^{n-1} + c_2a^{n-2}$ и $b^n = c_1b^{n-1} + c_2b^{n-2}$, получим

$$\begin{aligned} a^n + b^n &= c_1a^{n-1} + c_1b^{n-1} + c_2a^{n-2} + c_2b^{n-2} = \\ &= c_1(a^{n-1} + b^{n-1}) + c_2(a^{n-2} + b^{n-2}), \end{aligned}$$

так что $a_n = a^n + b^n$ является решением для $a_n = c_1 a_{n-1} + c_2 a_{n-2}$. Подводя итог сказанному выше, получаем, что поскольку умножение на константу решения для $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ также дает решение, поэтому если a^n и b^n — решения, то для некоторых констант c и d решениями являются также ca^n и db^n. Но поскольку сумма решений для $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ есть решение, то $ca^n + db^n$ для констант c и d также является решением. Таким образом, общее решение для $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ имеет вид $ca^n + db^n$, где c и d — константы, а $r = a$ и $r = b$ — решения уравнения $r^2 = c_1 r + c_2$.

Например, известно, что $a_n = a_{n-1} + 6a_{n-2}$ имеет решения $a_n = 3^n$ и $a_n = (-2)^n$. Поэтому $a_n = c3^n + d(-2)^n$ общее решение для $a_n = a_{n-1} + 6a_{n-2}$. Если было бы необходимо полностью определить решение $a_n = a_{n-1} + 6a_{n-2}$ как рекурсивной функции, то следовало бы задать первые два значения последовательности. Например, если $a_0 = 1$ и $a_1 = 8$, то $a_0 = c3^0 + d(-2)^0 = c + d = 1$ и $a_1 = c3^1 + d(-2)^1 = 3c - 2d = 8$. Решая систему уравнений

$$\begin{aligned} c + d &= 1, \\ 3c - 2d &= 8, \end{aligned}$$

находим, что $c = 2$ и $d = -1$, поэтому решением рекурсивной функции

$$\begin{aligned} a_0 &= 1, \\ a_1 &= 8, \\ a_n &= a_{n-1} + 6a_{n-2} \qquad \text{при } n > 2 \end{aligned}$$

является $a_n = 2 \cdot 3^n - (-2)^n$.

ПРИМЕР 11.5. Решим рекурсивную функцию

$$\begin{aligned} a_1 &= 2; \\ a_2 &= 10; \\ a_n &= 5a_{n-1} - 6a_{n-2} \qquad \text{при } n > 2. \end{aligned}$$

Сначала запишем характеристическое уравнение $r^2 = 5r - 6$ или $r^2 - 5r + 6 = 0$. Раскладывая на множители, получаем $(r - 2)(r - 3) = 0$, так что $r = 2$ и $r = 3$. Поэтому общее решение отношения $a_n = 5a_{n-1} - 6a_{n-2}$ имеет вид $a_n = c2^n + d3^n$. Но $a_1 = c2^1 + d3^1 = 2c + 3d = 2$ и $a_2 = c2^2 + d3^2 = 4c + 9d = 10$. Решая систему уравнений

$$\begin{aligned} 2c + 3d &= 2, \\ 4c + 9d &= 10, \end{aligned}$$

получаем $c = -2$ и $d = 2$, поэтому $a_n = (-2) \cdot 2^n + 2 \cdot 3^n$. □

ПРИМЕР 11.6. Найти решение для последовательности Фибоначчи, заданной соотношениями

$$\begin{aligned} &\text{Фиб}(1) = 1; \\ &\text{Фиб}(2) = 1; \\ &\text{Фиб}(n) = \text{Фиб}(n-1) + \text{Фиб}(n-2) \qquad \text{при } n > 2. \end{aligned}$$

Предупреждение: в некоторых книгах последовательность Фибоначчи определяется иначе, и поэтому имеет другое решение.

Характеристическим уравнением для $\text{Фиб}(n) = \text{Фиб}(n-1) + \text{Фиб}(n-2)$ при $n > 2$ будет $r^2 = r + 1$ или $r^2 - r - 1 = 0$. Используя формулу корней квадратного уравнения, находим, что

$$r = \frac{1 \pm \sqrt{5}}{2},$$

поэтому общее решение для $\text{Фиб}(n) = \text{Фиб}(n-1) + \text{Фиб}(n-2)$ имеет вид

$$\text{Фиб}(n) = c\left(\frac{1+\sqrt{5}}{2}\right)^n + d\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Чтобы упростить решение уравнения, положим $a_0 = 0$. По-прежнему, каждый член последовательности равен сумме двух предшествующих ему, если они существуют. Решая относительно c и d, получаем

$$\text{Фиб}(0) = c\left(\frac{1+\sqrt{5}}{2}\right)^0 + d\left(\frac{1-\sqrt{5}}{2}\right)^0 = c + d = 0$$

и

$$\text{Фиб}(1) = c\left(\frac{1+\sqrt{5}}{2}\right)^1 + d\left(\frac{1-\sqrt{5}}{2}\right)^1 = 1.$$

Поскольку $c+d = 0$, то $c = -d$. Поэтому второе уравнение может быть переписано как

$$-d\left(\frac{1+\sqrt{5}}{2}\right)^1 + d\left(\frac{1-\sqrt{5}}{2}\right)^1 = 1.$$

Так что

$$d = -\frac{1}{\sqrt{5}} \quad \text{и} \quad c = \frac{1}{\sqrt{5}}$$

и

$$\text{Фиб}(n) = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

есть последовательность Фибоначчи. □

Изучив ситуацию, в которой имелось два различных действительных корня, рассмотрим случай совпадающих корней. Итак, пусть задано рекурсивное отношение $a_n = c_1 a_{n-1} + c_2 a_{n-2}$, и характеристическое уравнение $r^2 = c_1 r + c_2$ имеет равные корни, например, $r = a$. Будем утверждать, что $a_n = a^n$ и $a_n = na^n$ — оба являются решениями для $a_n = c_1 a_{n-1} + c_2 a_{n-2}$. Известно, что $a_n = a^n$ — решение, но необходимо убедиться в том, что $a_n = na^n$ также является решением. Поскольку оба корня уравнения $r^2 = c_1 r + c_2$ равны a, то уравнение $r^2 = c_1 r + c_2$ в действительности имеет вид $(r-a)^2 = 0$ или $r^2 = 2ar - a^2$, так что $c_1 = 2a$ и $c_2 = -a^2$. Поэтому рассматриваемое рекуррентное отношение есть

$a_n = 2aa_{n-1} - a^2a_{n-2}$. Необходимо показать, что $a_n = na^n$ является решением. Подставив это выражение в отношение, получаем

$$\begin{aligned} na^n &= 2a(n-1)a^{n-1} - a^2(n-2)a^{n-2} = \\ &= 2na^n - 2a^n - na^n + 2a^n = \\ &= na^n, \end{aligned}$$

так что $a_n = na^n$ в действительности удовлетворяет рекуррентному отношению.

ПРИМЕР 11.7. Найдем общее решение рекуррентного отношения $a_n = 6a_{n-1} - 9a_{n-2}$. Характеристическое уравнение имеет вид $r^2 = 6r - 9$ или $r^2 - 6r + 9 = 0$. Раскладывая на множители, получаем $(r-3)^2 = 0$, поэтому $r = 3$ — корень кратности два. Следовательно, общим решением данного рекуррентного отношения является $a_n = c3^n + dn3^n$. □

ПРИМЕР 11.8. Найдем решение рекурсивной функции

$$\begin{aligned} a_0 &= 2, \\ a_1 &= 6, \\ a_n &= 4a_{n-1} - 4a_{n-2} \qquad \text{при } n > 2. \end{aligned}$$

Характеристический многочлен для $a_n = 4a_{n-1} - 4a_{n-2}$ имеет вид $r^2 - 4r + 4$, он задает уравнение $r^2 - 4r + 4 = 0$. Раскладывая на множители, получаем $(r-2)^2 = 0$, так что $r = 2$ — двойной корень. Поэтому общим решением для рекурсивной функции является $a_n = c2^n + dn2^n$, $a_0 = c2^0 + d(0)2^0 = c = 2$, $a_1 = c2^1 + d(1)2^1 = 2c + 2d = 6$. Поэтому $c = 2$ и $d = 1$. Следовательно, $a_n = 2 \cdot 2^n + n2^n = (n+2)2^n$. □

Осталось рассмотреть случай, когда оба корня характеристического уравнения — комплексные числа. Для этого нам понадобится одно важное свойство комплексных чисел, которое выводится из свойств тригонометрических функций.

ТЕОРЕМА 11.9. Для углов α и β

$$(\cos(\alpha) + i\sin(\alpha))(\cos(\beta) + i\sin(\beta)) = (\cos(\alpha+\beta) + i\sin(\alpha+\beta)).$$

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned} &(\cos(\alpha) + i\sin(\alpha))(\cos(\beta) + i\sin(\beta)) = \\ &\quad = \cos(\alpha)\cos(\beta) + i^2\sin(\alpha)\sin(\beta) + i\sin(\alpha)\cos(\beta) + i\sin(\beta)\cos(\alpha) = \\ &\quad = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) + i(\sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)) = \\ &\quad = \cos(\alpha+\beta) + i\sin(\alpha+\beta). \end{aligned}$$

■

Из этой теоремы следует, что $(\cos(\theta) + i\sin(\theta))^2 = \cos(2\theta) + i\sin(2\theta)$. Попробуйте доказать следующую теорему, используя метод индукции.

ТЕОРЕМА 11.10. (Муавр) Для произвольного угла θ имеет место равенство $(\cos(\theta) + i\sin(\theta))^k = \cos(k\theta) + i\sin(k\theta)$.

Нам необходимо еще одно свойство комплексных чисел. Рассмотрим комплексное число $a + bi$, как изображено на рис. 11.1.

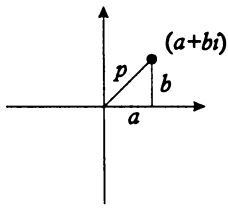


Рис. 11.1

По теореме Пифагора $\rho = \sqrt{a^2 + b^2}$. По определению $\cos(\theta) = \frac{a}{\rho}$, а $\sin(\theta) = \frac{b}{\rho}$, поэтому

$$a + bi = \rho(\frac{a}{\rho} + i\frac{b}{\rho}) = \rho(\cos(\theta) + i\sin(\theta)).$$

Таким образом, любое комплексное число может быть преобразовано к виду $\rho(\cos(\theta) + i\sin(\theta))$. Из приведенной выше теоремы далее следует, что $(a + bi)^n = (\rho(\cos(\theta) + i\sin(\theta)))^n = \rho^n(\cos(n\theta) + i\sin(n\theta))$.

Рассмотрим рекуррентное отношение $a_n = a_{n-1} - a_{n-2}$, которое имеет характеристический многочлен $r^2 - r + 1$, и соответствующее характеристическое уравнение $r^2 - r + 1 = 0$. По формуле корней квадратного уравнения находим, что

$$\begin{aligned} r &= 1 \pm i\sqrt{1-4} = \\ &= 1 \pm \sqrt{3}i. \end{aligned}$$

Как и ранее, предположим, что искомое общее решение имеет вид $cr_1^2 + dr_2^2$, где r_1 и r_2 — корни характеристического уравнения. Таким образом,

$$\begin{aligned} a_n &= c(1 + \sqrt{3}i)^n + d(1 - \sqrt{3}i)^n = \\ &= 2^n c\left(\frac{1}{2} + \frac{\sqrt{3}}{2}\right)^n + 2^n d\left(\frac{1}{2} - \frac{\sqrt{3}}{2}\right)^n = \\ &= 2^n c\left(\cos(\frac{\pi}{3}) + i\sin(\frac{\pi}{3})\right)^n + 2^n d\left(\cos(\frac{\pi}{3}) - i\sin(\frac{\pi}{3})\right)^n = \\ &= 2^n\left(c\left(\cos(n\cdot\frac{\pi}{3}) + i\sin(n\cdot\frac{\pi}{3})\right) + d\left(\cos(n\cdot\frac{\pi}{3}) - i\sin(n\cdot\frac{\pi}{3})\right)\right) = \\ &= 2^n((c+d)\cos(n\cdot\frac{\pi}{3}) + i(c-d)\sin(n\cdot\frac{\pi}{3})). \end{aligned}$$

Если положить $k_1 = c + d$ и $k_2 = i(c - d)$, то

$$a_n = 2^n(k_1\cos(n\cdot\frac{\pi}{3}) + k_2\sin(n\cdot\frac{\pi}{3})).$$

Здесь может показаться, что ловкостью рук удален какой-либо след комплексных чисел. И это действительно так. Предположим, например, что $a_0 = 1$ и $a_1 = 4$. Тогда

$$a_0 = 2^0(k_1\cos(0\cdot\frac{\pi}{3}) + k_2\sin(0\cdot\frac{\pi}{3})) = k_1\cos(0) + k_2\sin(0) = k_1$$

и $a_0 = k_1 = 1$. К тому же

$$a_1 = 2^1(k_1 \cos(1 \cdot \frac{\pi}{3}) + k_2 \sin(1 \cdot \frac{\pi}{3})) = 2((\cos(\frac{\pi}{3}) + k_2 \sin(\frac{\pi}{3})) =$$
$$= 2\left(\frac{1}{2} + k_2 \frac{\sqrt{3}}{2}\right) = 1 + k_2\sqrt{3}$$

и $a_1 = 1 + k_2\sqrt{3} = 4$. Решая относительно k_2, получаем

$$k_2\sqrt{3} = 3,$$

так что

$$k_2\sqrt{3} = \frac{3}{\sqrt{3}} = \sqrt{3}$$

и

$$a_n = 2^n(\cos(n \cdot \frac{\pi}{3}) + \sqrt{3}\sin(n \cdot \frac{\pi}{3})).$$

Вообще, если корни характеристического уравнения равны $r_1 = a + bi$ и $r_2 = a - bi$, то

$$a_n = \rho^n(k_1 \cos(n\theta) + k_2 \sin(n\theta)),$$

где $\rho = \sqrt{a^2 + b^2}$, $\cos(\theta) = \frac{a}{\rho}$ и $\sin(\theta) = \frac{b}{\rho}$.

ПРИМЕР 11.11. Решить рекурсивную функцию, заданную следующим образом:

$$a_0 = 1;$$
$$a_1 = 2;$$
$$a_n = 2\sqrt{2}a_{n-1} - 4a_{n-2}.$$

Характеристический многочлен имеет вид $r^2 - 2\sqrt{2}r + 4$. Используя формулу корней квадратного уравнения, находим $r = \sqrt{2} \pm \sqrt{2}i$, так что $\rho = 2$и $\cos(\theta) = \sin(\theta) = \frac{\sqrt{2}}{2}$. Поэтому $\theta = \frac{\pi}{4}$ и

$$a_n = 2^n(k_1 \cos(n \cdot \frac{\pi}{4}) + k_2 \sin(n \cdot \frac{\pi}{4})).$$

Следовательно,

$$a_0 = 2^0(k_1 \cos(0 \cdot \frac{\pi}{4}) + k_2 \sin(0 \cdot \frac{\pi}{4})) = k_1 \cos(0) + k_2 \sin(0) = k_1$$

и $a_0 = k_1 = 1$. К тому же,

$$a_1 = 2^1(k_1 \cos(1 \cdot \frac{\pi}{4}) + k_2 \sin(1 \cdot \frac{\pi}{4})) = 2((\cos(\frac{\pi}{4}) + k_2 \sin(\frac{\pi}{4})) =$$
$$= 2\left(\frac{\sqrt{2}}{2} + k_2 \frac{\sqrt{2}}{2}\right) = \sqrt{2} + k_2\sqrt{2}$$

и $a_1 = 2 = \sqrt{2} + k_2\sqrt{2}$. Решая относительно k_2, получаем

$$k_2 = \frac{2 - \sqrt{2}}{\sqrt{2}} = \sqrt{2} - 1$$

и

$$a_n = 2^n(\cos(n \cdot \frac{\pi}{4}) + (\sqrt{2} - 1)\sin(n \cdot \frac{\pi}{4})). \qquad \square$$

Теперь рассмотрим линейные рекуррентные отношения вида

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + c_3 a_{n-3} + \cdots + c_p a_{n-p},$$

где $p \geq 2$. Например, рассмотрим рекуррентное отношение

$$a_n = -6a_{n-1} - 5a_{n-2} + 24a_{n-3} + 36a_{n-4}.$$

Ему соответствует характеристическое уравнение

$$r^4 + 6r^3 + 5r^2 - 24r - 36 = 0.$$

Раскладывая на множители, получаем

$$(r - 2)(r + 2)(r + 3)^2 = 0.$$

В таком случае общее решение имеет вид

$$a_n = a \cdot (2)^n + b \cdot (-2)^n + c \cdot (-3)^n + d \cdot n(-3)^n.$$

Следующая формула общего решения приводится без доказательства.

ТЕОРЕМА 11.12. Пусть рекуррентное отношение

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + c_3 a_{n-3} + \cdots + c_p a_p$$

имеет характеристическое уравнение

$$r^n - c_1 r^{n-1} - c_2 r^{n-2} - c_3 r^{n-3} - \cdots - c_p r^{n-p} = 0.$$

Если r_i — корень характеристического уравнения кратности q_i, так что $(r - r_i)^{q_i}$ является множителем характеристического многочлена, то

$$a_{i1} r_i^n + a_{i2} n r_i^n + a_{i3} n^2 r_i^n + a_{i4} n^3 r_i^n + \cdots + a_{iq_i} n^{q_i - 1} r_i^n$$

включается в общее решение для a_n.

ПРИМЕР 11.13. Пусть рекуррентное отношение для a_n имеет характеристическое уравнение

$$(r - 1)^3(r - 2)^2(r - 3)^2 = 0.$$

Тогда

$$a_n = a + bn + cn^2 + d \cdot (2)^n + e \cdot n(2)^n + f \cdot 3^n + g \cdot n(3)^n$$

есть общее решение для a_n. $\square$

Определение чисел Каталана в разделе 5.2 является примером задания однородного линейного рекуррентного отношения, хотя его коэффициенты не являются постоянными. Напомним, что они были определены следующим образом.

$$\text{Кат}(0) = 1;$$
$$\text{Кат}(n+1) = \frac{2(2n+1)}{(n+2)}\,\text{Кат}(n).$$

■ УПРАЖНЕНИЯ

1. Найдите среди приведенных ниже линейные рекуррентные отношения:

а) $a_n = n^2 a_{n-1} - \sqrt{n}a_{n-2}$;

б) $a_n = a_{n-1}^2 - a_{n-2}$;

в) $a_n = a_{n-1} + \sqrt{3}a_{n-2} + \sin(n)$;

г) $a_n = a_{n-1} + 3a_{n-2}a_{n-3}$;

д) $a_n = a_{n-1} + 3a_{n-2} - na_{n-3}$.

2. Найдите среди рекуррентных отношений предыдущего упражнения однородные линейные рекуррентные отношения.

3. Найдите общее решение для приведенных ниже рекуррентных отношений:

а) $a_n - 3a_{n-1} = 0$;

б) $a_n + 3a_{n-1} = 0$;

в) $a_n = -a_{n-1} + 6a_{n-2}$;

г) $a_n = a_{n-1} + 3a_{n-2}$.

4. Найдите общее решение для приведенных ниже рекуррентных отношений:

а) $a_n + 7a_{n-1} = 0$;

б) $a_n - 5a_{n-1} = 0$;

в) $a_n = +2a_{n-1} + 8a_{n-2}$;

г) $a_n = +3a_{n-1} - 2a_{n-2}$.

5. Решите приведенные ниже рекуррентные функции:

а) $a_0 = 1$;
$a_n = -4a_{n-1}$ при $n > 1$;

б) $a_1 = 1$;
$a_n = 6a_{n-1}$;

в) $a_0 = 2$;
$a_1 = 5$;
$a_n = 5a_{n-1} - 6a_{n-2}$ при $n > 2$;

г) $a_0 = 2$;
$a_1 = 4$;
$a_n = 7a_{n-1} - 12a_{n-2}$ при $n > 2$.

6. Решите приведенные ниже рекуррентные функции:

а) $a_0 = 0$;
$a_1 = 5$;
$a_n = 9a_{n-1} - 20a_{n-2}$ при $n > 2$;

б) $a_0 = 2$;
$a_1 = 5$;
$a_n = 7a_{n-1} - 12a_{n-2}$ при $n > 2$;

в) $a_0 = 2$;
$a_1 = 1$;
$a_n = 2a_{n-1} + 2a_{n-2}$ при $n > 2$;

г) $a_0 = 4$;
$a_1 = 2$;
$a_n = -a_{n-1} + a_{n-2}$ при $n > 2$.

7. Решите приведенные ниже рекуррентные функции:

а) $a_0 = 3$;
$a_1 = 21$;
$a_n = 6a_{n-1} - 9a_{n-2}$ при $n > 2$;

б) $a_0 = 2$;
$a_1 = 6$;
$a_n = 4a_{n-1} - 4a_{n-2}$ при $n > 2$;

в) $a_0 = 1$;
$a_1 = -8$;
$a_n = -4a_{n-1} - 4a_{n-2}$ при $n > 2$;

г) $a_0 = -3$;
$a_1 = 1$;
$a_n = 2a_{n-1} - a_{n-2}$ при $n > 2$;

д) $a_0 = 2$;
$a_1 = 6$;
$a_n = -2a_{n-1} - a_{n-2}$ при $n > 2$.

8. Решите приведенные ниже рекуррентные функции:

а) $a_0 = 3$;
$a_1 = 4$;
$a_n = -4a_{n-2}$ при $n > 2$;

б) $a_0 = 1$;
$a_1 = 2$;
$a_n = a_{n-1} - a_{n-2}$ при $n > 2$;

в) $a_0 = 5$;
$a_1 = 4$;
$a_n = -16a_{n-2}$ при $n > 2$;

г) $a_0 = 1$;
$a_1 = 0$;
$a_n = -2\sqrt{2}a_{n-1} - 4a_{n-2}$ при $n > 2$;

д) $a_0 = 2$;
$a_1 = 1 - \sqrt{3}$;
$a_2 = 1 - 2\sqrt{3}$;
$a_n = a_{n-3}$ при $n > 3$.

9. Найдите общее решение для приведенных ниже рекуррентных отношений:

а) $a_n - 5a_{n-2} + 4a_{n-4} = 0$;

б) $a_n - 2a_{n-1} + 2a_{n-2} - 2a_{n-3} + a_{n-4} = 0$;

в) $a_n = a_{n-4}$;

г) $a_n = +3a_{n-1} - 3a_{n-2} + a_{n-3} - a_{n-4} + 3a_{n-5} - 3a_{n-6} + a_{n-7}$.

10. Покажите, что последовательность

$$F(n) = \binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots$$

удовлетворяет рекуррентному отношению для чисел Фибоначчи и, следовательно, является альтернативным представлением чисел Фибоначчи.

11. Общая ***последовательность Лукаса*** L_n задается рекурсивной функцией

$$\begin{aligned} L_1 &= a, \\ L_2 &= b, \\ L_n &= L_{n-1} + L_{n-2} \qquad \text{при } n > 2, \end{aligned}$$

где p и q — целые числа. Заметим, что если $L_1 = L_2 = 1$, то имеем последовательность Фибоначчи Фиб(n). Покажите, что $L_n = b\,\text{Фиб}(n-1) + a\,\text{Фиб}(n-2)$ для всех $n > 2$.

12. Используя теорему 11.9 и индукцию, докажите теорему 11.10. Для заданного угла θ $(\cos(\theta) + i\sin(\theta))^k = \cos(k\theta) + i\sin(k\theta)$.

11.2. НЕОДНОРОДНЫЕ ЛИНЕЙНЫЕ РЕКУРРЕНТНЫЕ ОТНОШЕНИЯ

Рассмотрим рекурсивную функцию a_n, заданную соотношениями

$$a_1 = c,$$
$$a_n = a \cdot a_{n-1} + b,$$

где $a \neq 1$. Сразу получаем, что

$$\begin{aligned}
a_1 &= c;\\
a_2 &= ac + b;\\
a_3 &= a^2c + ab + b;\\
a_4 &= a^3c + a^2b + ab + b;\\
a_5 &= a^4c + a^3b + a^2b + ab + b;\\
&\vdots\\
a_n &= a^{n-1}c + a^{n-2}b + a^{n-3}b + \cdots + a^2b + ab + b =\\
&= a^{n-1}c + b(a^{n-2} + a^{n-3} + \cdots + a^2 + a + 1 =\\
&= a^{n-1}c + b\left(\frac{a^{n-1} - 1}{a - 1}\right) =\\
&= \left(c + \frac{b}{a-1}\right)a^{n-1} - \frac{b}{a-1} =\\
&= \left(c - \frac{b}{1-a}\right)a^{n-1} + \frac{b}{1-a}.
\end{aligned}$$

Заметим, что данную рекурсивную функцию можно записать в более общем виде

$$a_n = c_1 a_{n-1} + f(n),$$

где $c_1 = a$ и $f(n) = b$. Будем стремиться решить эту задачу с помощью методов, использованных для однородных рекурсий. Для этого необходима следующая теорема.

ТЕОРЕМА 11.14. Пусть $a_n^0 = P_n$ удовлетворяет уравнению

$$a_n^0 = c_1 a_{n-1}^0 + c_2 a_{n-2}^0 + c_3 a_{n-3}^0 + \cdots + c_p a_{n-p}^0,$$

и $a_n = Q_n$ есть частное решение уравнения

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + c_3 a_{n-3} + \cdots + c_p a_{n-p} + f(n).$$

Тогда $a_n = P_n + Q_n$ также является решением уравнения

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + c_3 a_{n-3} + \cdots + c_p a_{n-p} + f(n).$$

ДОКАЗАТЕЛЬСТВО. Поскольку $a_n^0 = P_n$ удовлетворяет уравнению

$$a_n^0 = c_1 a_{n-1}^0 + c_2 a_{n-2}^0 + c_3 a_{n-3}^0 + \cdots + c_p a_{n-p}^0,$$

имеем

$$P_n = c_1P_{n-1} + c_2P_{n-2} + c_3P_{n-3} + \cdots + c_pP_{n-p}.$$

Поскольку $a_n = Q_n$ удовлетворяет уравнению

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p} + f(n),$$

имеем

$$Q_n = c_1Q_{n-1} + c_2Q_{n-2} + c_3Q_{n-3} + \cdots + c_pQ_{n-p} + f(n).$$

Складывая

$$P_n = c_1P_{n-1} + c_2P_{n-2} + c_3P_{n-3} + \cdots + c_pP_{n-p}$$

и

$$Q_n = c_1Q_{n-1} + c_2Q_{n-2} + c_3Q_{n-3} + \cdots + c_pQ_{n-p} + f(n),$$

получаем

$$P_n + Q_n = c_1(P_{n-1} + Q_{n-1}) + c_2(P_{n-2} + Q_{n-2}) + \cdots + c_p(P_{n-p} + Q_{n-p}) + f(n),$$

так что $a_n = P_n + Q_n$ удовлетворяет уравнению

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p} + f(n). \quad \blacksquare$$

ТЕОРЕМА 11.15. Пусть Q_n есть решение уравнения

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p} + f(n).$$

Тогда каждое решение уравнения

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p} + f(n)$$

имеет вид $P_n + Q_n$, где P_n является решением уравнения

$$a_n^0 = c_1a_{n-1}^0 + c_2a_{n-2}^0 + c_3a_{n-3}^0 + \cdots + c_pa_{n-p}^0.$$

ДОКАЗАТЕЛЬСТВО. Известно, что $P_n + Q_n$ — решение уравнения

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p} + f(n).$$

Предположим, что Q_n и R_n являются решениями уравнения

$$a_n = c_1a_{n-1} + c_2a_{n-2} + c_3a_{n-3} + \cdots + c_pa_{n-p} + f(n).$$

Поэтому

$$Q_n = c_1Q_{n-1} + c_2Q_{n-2} + c_3Q_{n-3} + \cdots + c_pQ_{n-p} + f(n)$$

и

$$R_n = c_1R_{n-1} + c_2R_{n-2} + c_3R_{n-3} + \cdots + c_pR_{n-p} + f(n).$$

Вычитая первое уравнение из второго, получаем

$$R_n - Q_n = c_1(R_{n-1} - Q_{n-1}) + c_2(R_{n-2} - Q_{n-2}) + \cdots + c_p(R_{n-p} - Q_{n-p}),$$

так что $R_n - Q_n$ — решение уравнения

$$a_n^0 = c_1 a_{n-1}^0 + c_2 a_{n-2}^0 + c_3 a_{n-3}^0 + \cdots + c_p a_{n-p}^0.$$

Обозначим $R_n - Q_n = P_n$. Тогда $R_n = Q_n + P_n$, где P_n удовлетворяет уравнению

$$a_n^0 = c_1 a_{n-1}^0 + c_2 a_{n-2}^0 + c_3 a_{n-3}^0 + \cdots + c_p a_{n-p}^0.$$

и теорема доказана. ■

Известно, что рекуррентное отношение $a_n = a \cdot a_{n-1}$ имеет общее решение $a_n = k_1 a^n$, где k_1 — константа. Поэтому, если можно было бы найти частное решение Q_n рекуррентного отношения $a_n = a \cdot a_{n-1} + b$, тогда его общее решение имело бы вид $R_n = Q_n + k_1 a^n$.

ЗАМЕЧАНИЕ 11.16. *Будем предполагать, что частное решение Q_n для*

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + c_3 a_{n-3} + \cdots + c_p a_{n-p} + f(n)$$

имеет такой же вид, как и функция $f(n)$.

Например, если $f(n) = a \cdot b^n$, то предполагаем, что частное решение Q_n имеет вид $k_2 \cdot b^n$, а если $f(n)$ — многочлен третьей степени, то предполагаем, что Q_n имеет вид $k_2 \cdot n^3 + k_3 \cdot n^2 + k_4 \cdot n + k_5$.

В рассматриваемом частном случае $f(n)$ — константа, поэтому предполагаем, что константа k_2 является решением для $a_n = a \cdot a_{n-1} + b$. Поэтому, подставляя k_2 в уравнение, получаем

$$k_2 = a \cdot k_2 + b,$$

так что

$$k_2 = \frac{b}{1-a}.$$

Поэтому рекуррентное отношение имеет вид

$$a_n = k_1 a^n + \frac{b}{1-a}.$$

Поскольку в частном случае $a_1 = c$, то имеем

$$c = k_1 a + \frac{b}{1-a} k_1 = \frac{1}{a}\left(c - \frac{b}{1-a}\right),$$

так что

$$a_n = \frac{1}{a}\left(c - \frac{b}{1-a}\right) a^n + \frac{b}{1-a} = \left(c - \frac{b}{1-a}\right) a^{n-1} + \frac{b}{1-a},$$

что согласуется с решением, приведенным в начале этого раздела.

В начале раздела была рассмотрена рекурсивная функция a_n вида

$$a_1 = c,$$
$$a_n = a \cdot a_{n-1} + b,$$

где $a \neq 1$. Предположим, что $a = 1$, поэтому

$$a_1 = c,$$
$$a_n = a_{n-1} + b.$$

Уравнение $a_n = a_{n-1}$ имеет характеристическое уравнение $r - 1 = 0$, поэтому общее решение для $a_n = a_{n-1}$ имеет вид

$$a_n = k_1(1)^n = k_1.$$

Следует искать частное решение для $a_n = a_{n-1} + b$ в виде $a_n = k_2$, поскольку $f(n) = b$, т.е. равно константе. Но $a_n = k_2$ удовлетворяет уравнению $a_n = a_{n-1}$, так что если $b \neq 0$, оно не может удовлетворять $a_n = a_{n-1} + b$. Поэтому повторим рассмотренную выше процедуру для кратных корней и положим $a_n = n \cdot k_2$. Если подставить это выражение в уравнение $a_n = a_{n-1} + b$, получим

$$n \cdot k_2 = (n-1)k_2 + b,$$

или

$$k_2 = b.$$

Поэтому

$$a_n = k_1 + nb.$$

Поскольку $a_1 = c$,

$$c = k_1 + b$$

и

$$k_1 = c - b.$$

Следовательно,

$$a_n = c - b + nb =$$
$$= c + (n-1)b.$$

ЗАМЕЧАНИЕ 11.17. *Допустим, имеется рекуррентное отношение*

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + c_3 a_{n-3} + \cdots + c_p a_{n-p} + f(n).$$

Если $f(n) = aP(n)$ *и общим решением для однородного рекуррентного отношения*

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + c_3 a_{n-3} + \cdots + c_p a_{n-p}$$

является $b_1 f(n) + b_2 n f(n) + b_3 n^2 f(n) + \cdots + b_j n^j f(n)$, *но* $b_j n^j f(n)$ *решением не является, то* $Q_n = kn^{j+1} f(n)$.

Например, если $a_n = 4a_{n-1} - 4a_{n-2} + 3 \cdot 2^n$, то характеристическим многочленом является $r^2 - 4r + 4$, и однородное рекуррентное отношение имеет вид $a_n = a \cdot 2^n + b \cdot n2^n$, поэтому предположим, что частное решение для $a_n = 4a_{n-1} - 4a_{n-2} + 3 \cdot 2^n$ имеет вид $k \cdot n^2 2^n$.

ПРИМЕР 11.18. Пусть $a_n = 5a_{n-1} - 4a_{n-2} + 3 \cdot 2^n$. Характеристическим многочленом является $r^2 - 5r + 4$, и однородное рекуррентное отношение имеет вид $a_n = a + b \cdot 4^n$. Поэтому предположим, что Q_n имеет вид $k \cdot 2^n$. Подставляя это выражение в рекуррентное отношение, получаем

$$k \cdot 2^n = 5k \cdot 2^{n-1} - 4k \cdot 2^{n-2} + 3 \cdot 2^n$$

или

$$4k \cdot 2^{n-2} = 10k \cdot 2^{n-2} - 4k \cdot 2^{n-2} + 12 \cdot 2^{n-2},$$

так что

$$4k = 10k - 4k + 12$$

и $k = -6$. Следовательно, общее решение для $a_n = 5a_{n-1} - 4a_{n-2} + 3 \cdot 2^n$ имеет вид $a_n = a + b \cdot 4^n - 6 \cdot 2^n$. □

ПРИМЕР 11.19. Пусть $a_n = 5a_{n-1} - 6a_{n-2} + 6 \cdot 3^n$. Характеристическим многочленом будет $r^2 - 5r + 6$, и однородное рекуррентное отношение имеет вид $a_n = a \cdot 2^n + b \cdot 3^n$. Поэтому предположим, что Q_n имеет вид $k \cdot n3^n$. Подставляя это выражение в рекуррентное отношение, получаем

$$kn3^n = 5k(n-1)3^{n-1} - 6k(n-2)3^{n-2} + 6 \cdot 3^n$$

или

$$9kn3^{n-2} = 15k(n-1)3^{n-2} - 6k(n-2)3^{n-2} + 54 \cdot 3^{n-2},$$

так что

$$9kn = 15k(n-1) - 6k(n-2) + 54$$

и $k = 18$. Поэтому общим решением рекуррентного отношения $a_n = 5a_{n-1} - 6a_{n-2} + 6 \cdot 3^n$ будет $a_n = a \cdot 2^n + b \cdot 3^n + 18n3^n$. □

ПРИМЕР 11.20. Решим рекурсивную функцию:

$$a_1 = 1,$$
$$a_n = a_{n-1} + n.$$

Учитывая, что a_n равно сумме первых n положительных целых чисел, фактически требуется найти формулу для суммы первых n положительных целых чисел. Характеристическим уравнением будет

$$r - 1 = 0,$$

поэтому решение однородного отношения имеет вид $a_n = k_1$. Поскольку $f(n) = n$, естественно предположить, что частное решение для $a_n = a_{n-1} + n$ имеет вид $k_2 n + k_3$. Поскольку константа является решением однородного отношения,

частное решение для $a_n = a_{n-1} + n$ следует искать в виде $a_n = k_2 n^2 + k_3 n$. Подставляя это выражение в рекуррентное отношение $a_n = a_{n-1} + n$, получаем

$$k_2 n^2 + k_3 n = k_2 (n-1)^2 + k_3 (n-1) + n$$

или

$$k_2 n^2 + k_3 n = k_2 (n^2 - 2n + 1) + k_3 (n-1) + n.$$

Приравнивая коэффициенты при n, получаем

$$k_3 = -2k_2 + k_3 + 1$$

или

$$k_2 = \frac{1}{2}.$$

Приравнивая константы, получаем

$$k_2 - k_3 = 0,$$

так что

$$k_3 = \frac{1}{2}.$$

Поэтому общее решение для $a_n = a_{n-1} + n$ имеет вид

$$a_n = k_1 + \frac{1}{2} n^2 + \frac{1}{2} n.$$

Поскольку $a_1 = 1$, имеем

$$a_1 = 1 = k_1 + \frac{1}{2} + \frac{1}{2},$$

так что $k_1 = 0$ и

$$a_n = \frac{1}{2} n^2 + \frac{1}{2} n = \frac{n(n+1)}{2}. \qquad \square$$

ПРИМЕР 11.21. Найдем решение рекуррентного отношения

$$a_n = 3a_{n-1} - 2a_{n-2} + 3 \sin\left(\frac{n\pi}{2}\right)$$

при $n \geq 2$. Однородное отношение легко решается. Характеристическое уравнение имеет вид

$$r^2 - 3r + 2 = 0$$

или

$$(r-1)(r-2) = 0,$$

так что общим решением для однородного отношения будет

$$a_n^0 = a + b \cdot 2^n.$$

Учитывая, что $f(n)=3\sin\left(\frac{n\pi}{2}\right)$ и решение однородного рекуррентного отношения с квадратичным характеристическим многочленом, имеющим комплексные корни $a+bi$ и $a-bi$, имеет вид $b_1\cos(n\theta)+b_2\sin(n\theta)$, частное решение для

$$a_n = 3a_{n-1} - 2a_{n-2} + 3\sin\left(\frac{n\pi}{2}\right)$$

будем искать в виде $b_1\cos\left(\frac{n\pi}{2}\right)+b_2\sin\left(\frac{n\pi}{2}\right)$. После подстановки этого выражения в рекуррентное отношение имеем

$$b_1\cos\left(\frac{n\pi}{2}\right)+b_2\sin\left(\frac{n\pi}{2}\right) = 3b_1\cos\left(\frac{(n-1)\pi}{2}\right)+3b_2\sin\left(\frac{(n-1)\pi}{2}\right)-$$
$$-2\cos b_1\left(\frac{(n-2)\pi}{2}\right)-2b_2\sin\left(\frac{(n-2)\pi}{2}\right)+3\sin\left(\frac{n\pi}{2}\right).$$

Так как это должно выполняться для всех $n\geq 2$, подставляем $n=2$ и получаем

$$b_1\cos(\pi)+b_2\sin(\pi) = 3b_1\cos\left(\frac{\pi}{2}\right)+3b_2\sin\left(\frac{\pi}{2}\right)-2b_1\cos(0)-2b_2\sin(0)+3\sin(\pi),$$

или

$$-b_1 = +3b_2 - 2b_1,$$

или

$$b_1 = 3b_2.$$

Если подставить $n=3$, получим

$$b_1\cos\left(\frac{3\pi}{2}\right)+b_2\sin\left(\frac{3\pi}{2}\right) =$$
$$= 3b_1\cos(\pi)+3b_2\sin(\pi)-2\cos b_1\left(\frac{\pi}{2}\right)-2b_2\sin\left(\frac{\pi}{2}\right)+3\sin\left(\frac{3\pi}{2}\right),$$

или

$$-b_2 = -3b_1 - 2b_2 - 3,$$

или

$$3b_1 + b_2 = -3.$$

Подставляя в это уравнение $b_1=3b_2$, получаем

$$9b_2 + b_2 = -3$$

или

$$b_2 = -\frac{3}{10},$$

так что

$$b_1 = -\frac{9}{10},$$

и искомое общее решение имеет вид $a_n = a + b \cdot 2^n - \frac{9}{10}\cos\left(\frac{n\pi}{2}\right) - \frac{3}{10}\sin\left(\frac{n\pi}{2}\right)$. Заметьте, мы показали, что это общее решение, если оно существует, поскольку справедливость этого показана только для $n = 2$ и $n = 3$. В общем случае будем предполагать, что решение существует, и использовать изложенный метод. Однако, для того чтобы правильно показать, что это решение, следует взять исходное уравнение

$$b_1\cos\left(\frac{n\pi}{2}\right) + b_2\sin\left(\frac{n\pi}{2}\right) = 3b_1\cos\left(\frac{(n-1)\pi}{2}\right) + 3b_2\sin\left(\frac{(n-1)\pi}{2}\right) - \\ - 2\cos b_1\left(\frac{(n-2)\pi}{2}\right) - 2b_2\sin\left(\frac{(n-2)\pi}{2}\right) + 3\sin\left(\frac{n\pi}{2}\right)$$

и, используя тригонометрические тождества, выразить все слагаемые в виде $\cos\left(\frac{(n-2)\pi}{2}\right)$ или $\sin\left(\frac{(n-2)\pi}{2}\right)$.

Рассматривая каждое слагаемое отдельно, находим

$$b_1\cos\left(\frac{n\pi}{2}\right) = b_1\cos\left(\frac{(n-2)\pi}{2}\right)\cos(\pi) - b_1\sin\left(\frac{(n-2)\pi}{2}\right)\sin(\pi) = \\ = -b_1\cos\left(\frac{(n-2)\pi}{2}\right);$$

$$b_2\sin\left(\frac{n\pi}{2}\right) = b_2\cos\left(\frac{(n-2)\pi}{2}\right)\sin(\pi) + b_2\sin\left(\frac{(n-2)\pi}{2}\right)\cos(\pi) = \\ = -b_2\sin\left(\frac{(n-2)\pi}{2}\right);$$

$$3b_1\cos\left(\frac{(n-1)\pi}{2}\right) = 3b_1\cos\left(\frac{(n-2)\pi}{2}\right)\cos\left(\frac{\pi}{2}\right) - 3b_1\sin\left(\frac{(n-2)\pi}{2}\right)\sin\left(\frac{\pi}{2}\right) = \\ = -3b_1\sin\left(\frac{(n-2)\pi}{2}\right);$$

$$3b_2\sin\left(\frac{(n-1)\pi}{2}\right) = 3b_2\cos\left(\frac{(n-2)\pi}{2}\right)\sin\left(\frac{\pi}{2}\right) + \sin\left(\frac{(n-2)\pi}{2}\right)\cos\left(\frac{\pi}{2}\right) = \\ = 3b_2\cos\left(\frac{(n-2)\pi}{2}\right);$$

$$3\sin\left(\frac{n\pi}{2}\right) = 3\cos\left(\frac{(n-2)\pi}{2}\right)\sin(\pi) + 3\sin\left(\frac{(n-2)\pi}{2}\right)\cos(\pi) = \\ = -3\sin\left(\frac{(n-2)\pi}{2}\right).$$

Подставляя эти слагаемые в

$$b_1\cos\left(\frac{n\pi}{2}\right) + b_2\sin\left(\frac{n\pi}{2}\right) = 3b_1\cos\left(\frac{(n-1)\pi}{2}\right) + 3b_2\sin\left(\frac{(n-1)\pi}{2}\right) - \\ - 2\cos b_1\left(\frac{(n-2)\pi}{2}\right) - 2b_2\sin\left(\frac{(n-2)\pi}{2}\right) + 3\sin\left(\frac{n\pi}{2}\right),$$

имеем

$$-b_1\cos\left(\frac{(n-2)\pi}{2}\right)-b_2\sin\left(\frac{(n-2)\pi}{2}\right)=$$
$$=-3b_1\sin\left(\frac{(n-2)\pi}{2}\right)+3b_2\cos\left(\frac{(n-2)\pi}{2}\right)-2b_1\cos\left(\frac{(n-2)\pi}{2}\right)-$$
$$-2b_2\sin\left(\frac{(n-2)\pi}{2}\right)-3\sin\left(\frac{(n-2)\pi}{2}\right).$$

Приравнивая коэффициенты при $\cos\left(\frac{(n-2)\pi}{2}\right)$, имеем

$$-b_1=3b_2-2b_1 \qquad \text{при} \qquad b_1=3b_2.$$

Приравнивая коэффициенты при $\sin\left(\frac{(n-2)\pi}{2}\right)$, имеем

$$-b_2=-3b_1-2b_2-3 \qquad \text{при} \qquad 3b_1+b_2=3,$$

что совпадает с уравнениями, полученными ранее. Следовательно, теперь мы убедились, что решение действительно существует. □

■ УПРАЖНЕНИЯ

1. Найдите общее решение для приведенных ниже рекуррентных отношений:

а) $a_n=2a_{n-1}+5$; **б)** $a_n=-3a_{n-1}+n$;

в) $a_n=-a_{n-1}+12a_{n-2}+2^n$; **г)** $a_n=-3a_{n-1}-a_{n-2}+3^n$;

д) $a_n=4a_{n-1}-4a_{n-2}+n^2$.

2. Найдите общее решение для приведенных ниже рекуррентных отношений:

а) $a_n=-2a_{n-1}+n^2$; **б)** $a_n=4a_{n-1}+4^n$;

в) $a_n=6a_{n-1}+9a_{n-2}+(-2)^n$; **г)** $a_n=-9a_{n-2}+3^n$;

д) $a_n=-4a_{n-1}-2a_{n-2}+5$.

3. Найдите общее решение для приведенных ниже рекуррентных отношений:

а) $a_n=2a_{n-1}+2^n$; **б)** $a_n=-6a_{n-1}-9a_{n-2}+(-3)^n$;

в) $a_n=3_{n-1}-2a_{n-2}+5$; **г)** $a_n=a_{n-1}+6a_{n-2}+3^n$;

д) $a_n=2\sqrt{3}a_{n-1}-4a_{n-2}+3$.

4. Найдите общее решение для приведенных ниже рекуррентных отношений:

а) $a_n=a_{n-2}+\cos\left(\frac{n\pi}{2}\right)$; **б)** $a_n=-a_{n-2}+\cos\left(\frac{n\pi}{2}\right)$;

в) $a_n=6a_{n-1}-12a_{n-2}+8a_{n-3}+3^n$; **г)** $a_n=6a_{n-1}-12a_{n-2}+8a_{n-3}+2^n$;

д) $a_n=-a_{n-3}+n$.

5. Найдите приведенные ниже суммы, используя рекуррентные отношения:

а) $1^2+2^2+3^2+\cdots+n^2$;
Используйте соотношение
$a_n=a_{n-1}+n^2$.

б) $1+4+7+\cdots+3n-2$;
Используйте соотношение
$a_n=a_{n-1}+3n-2$.

в) $2^1 + 2^2 + 2^3 + 2^4 + \cdots + 2^n$;
Используйте соотношение $a_n = a_{n-1} + 2^n$.

г) $1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + \cdots + n \cdot (n+1)$;
Используйте соотношение $a_n = a_{n-1} + n \cdot (n+1)$.

6. Найдите приведенные ниже суммы, используя рекуррентные отношения:

а) $1^3 + 2^3 + 3^3 + \cdots + n^3$;
Используйте соотношение $a_n = a_{n-1} + n^3$.

б) $1^2 \cdot 2 + 2^2 \cdot 3 + 3^2 \cdot 4 + \cdots + n^2 \cdot (n+1)$;
Используйте соотношение $a_n = a_{n-1} + n^2 \cdot (n+1)$.

в) $1 \cdot 2^0 + 2 \cdot 2^1 + 3 \cdot 2^2 + \cdots + n \cdot 2^{(n-1)}$;
Используйте соотношение $a_n = a_{n-1} + n \cdot 2^{(n-1)}$.

11.3. КОНЕЧНЫЕ РАЗНОСТИ

Для студентов, прослушавших курс математического анализа, многие результаты, изложенные в данном разделе, покажутся хорошо известными. Фактически, конечные разности можно рассматривать как бесконечно малые без предельных переходов. Аналогично исчислению бесконечно малых, теория конечных разностей имеет широкое применение в таких разнообразных областях, как информатика, актуарная математика, экономика, психология и социология. Сфера их использования включает обнаружение случайных погрешностей, построение полиномов, аппроксимирующих функциональную зависимость по результатам измерений, экстраполирование и интерполирование функций, суммирование функций, дифференцирование комбинаторных функций, аппроксимацию площадей и многое другое. Конечные разности — одно из базовых средств численного анализа. Следует, однако, подчеркнуть, что изучение данного раздела не требует знания математического анализа. В этом разделе мы только вводим понятие о конечных разностях и показываем некоторые их элементарные свойства.

Первая разность, или ***разность первого порядка*** функции f, обозначаемая $\triangle f$, определена следующим образом

$$\triangle f(x) = f(x+1) - f(x).$$

Вторая разность, или ***разность второго порядка*** функции f, обозначаемая $\triangle^2 f$, определена следующим образом

$$\triangle^2 f(x) = \triangle(\triangle f(x)).$$

Поэтому,

$$\begin{aligned}\triangle^2 f(x) &= \triangle(f(x+1) - f(x)) = \\ &= (f(x+2) - f(x+1)) - (f(x+1) - f(x)) = \\ &= f(x+2) - 2f(x+1) + f(x).\end{aligned}$$

В общем случае n-ая разность функции f, обозначаемая $\triangle^n f$, индуктивно определена выражением

$$\triangle^n f(x) = \triangle(\triangle^{n-1} f(x)).$$

Далее приводится формула, выражающая $\triangle^n f(x)$ через $f(x)$. Приведенная ниже таблица иллюстрирует разностную функцию:

x	$f(x)$	$\triangle f(x)$	$\triangle^2 f(x)$	$\triangle^3 f(x)$	$\triangle^4 f(x)$
1	1	7	12	6	0
2	8	19	18	6	0
3	27	37	24	6	
4	64	61	30		
5	125	91			
6	216				

Заметим, что в данном случае $f(x) = x^3$ и $\triangle^4 f(x) = 0$ для всех значений x.

Оператором называется функция, которая отображает функции в функции. Следовательно, $\triangle$ является оператором. Определим также другой оператор E согласно выражению $E(f(x)) = f(x+1)$. Таким образом,

$$\triangle f(x) \;=\; E(f(x)) - f(x) \;=\; (E - I)(f(x)),$$

где использовано обозначение $(F+G)(u) = F(u)+G(u)$. Оператор I представляет собой тождественный оператор. Поэтому $\triangle = (E - I)$ и $E = I + \triangle$.

Используя запись $\triangle = (E - I)$ и полагая $E^0 = I$, получаем

$$\triangle^n(f(x)) = (E-I)^n(f(x)) = E^n(f(x)) - n \cdot E^{n-1}(f(x)) + \cdots$$
$$\cdots + (-1)^k \binom{n}{k} \cdot E^{n-k}(f(x)) + \cdots + (-1)^n E^0(f(x)) =$$
$$= f(x+n) - n \cdot f(n-1) + \cdots + (-1)^k \binom{n}{k} (f(x+n-k)) + \cdots + (-1)^n f(x).$$

ПРИМЕР 11.22. $\triangle^3(f(x)) = f(x+3) - 3f(x+2) + 3f(x+1) - f(x)$. □

ТЕОРЕМА 11.23. Операторы $\triangle$ и E обладают следующими свойствами.

Для действительного числа a и функций f и g

а) $\triangle(f+g) = \triangle(f) + \triangle(g)$; $E(f+g) = E(f) + E(g)$;
б) $\triangle(af) = a\triangle(f)$; $E(af) = aE(f)$;
в) $E\triangle = \triangle E$;
г) $\triangle(a) = 0$; $E(a) = a$.

ДОКАЗАТЕЛЬСТВО. Докажем пункт (а).

$$\begin{aligned}
\triangle((f+g)(x)) &= \triangle(f(x) + g(x)) = \\
&= f(x+1) + g(x+1) - (f(x) + g(x)) = \\
&= f(x+1) - f(x) + g(x+1) - g(x) = \\
&= \triangle f(x) + \triangle g(x) = \\
&= (\triangle f + \triangle g)(x)
\end{aligned}$$

и

$$\begin{aligned}\triangle((af)(x)) &= \triangle(af(x)) = \\ &= af(x+1) + af(x) = \\ &= a(f(x+1) - f(x)) = \\ &= a\triangle f(x).\end{aligned}$$

Доказательство пунктов (б) и (в) предоставляется читателю. ■

ПРИМЕР 11.24.

$$\begin{aligned}\triangle(3x^4 + 2x^3 + 5x + 4) &= \triangle(3x^4) + \triangle(2x^3) + \triangle(5x) + \triangle(4) = \\ &= 3\triangle(x^4) + 2\triangle(x^3) + 5\triangle(x).\end{aligned}$$ □

Теперь необходимо найти $\triangle(x^n)$. К сожалению, здесь все не так просто. Имеем

$$\triangle(x) = (x+1) - x = 1$$

и

$$\triangle(x^2) = (x+1)^2 - x^2 = x^2 + 2x + 1 - x^2 = 2x + 1.$$

В общем случае имеем

$$\begin{aligned}\triangle(x^n) = (x+1)^n - x^n = x^n + nx^{n-1} + \cdots + \binom{n}{k}x^k + \cdots + 1 - x^n = \\ = nx^{n-1} + \cdots + \binom{n}{k}x^k \cdots + 1.\end{aligned}$$

Теперь введем два свойства оператора $\triangle$, которые знакомы студентам, прослушавшим курс математического анализа. Они называются соответственно ***правило произведения*** и ***правило частного***.

ТЕОРЕМА 11.25. Для функций f и g

$$\triangle(fg) = f \cdot \triangle(g) + E(g) \cdot \triangle(f)$$

и

$$\triangle\left(\frac{f}{g}\right) = \frac{g \cdot \triangle(f) - f \cdot (\triangle(g))}{g \cdot (E(g))}.$$

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned}\triangle(fg(x)) &= \triangle(f(x) \cdot g(x)) = \\ &= f(x+1) \cdot g(x+1) - f(x) \cdot g(x) = \\ &= f(x+1) \cdot g(x+1) - f(x) \cdot g(x+1) + f(x) \cdot g(x+1) - f(x) \cdot g(x) = \\ &= (f(x+1) - f(x)) \cdot g(x+1) + f(x) \cdot (g(x+1) - \cdot g(x)) = \\ &= \triangle f(x) \cdot E(g(x)) + f(x) \cdot \triangle g(x) \\ &= f(x) \cdot \triangle g(x) + Eg(x) \cdot \triangle f(x) = \\ &= (f \cdot \triangle g + Eg \cdot \triangle f)(x).\end{aligned}$$

Доказательство второй части предоставляется читателю. ■

ПРИМЕР 11.26.

$$\begin{aligned}\triangle((x^2+6x)(2x^2+5) &= (x^2+6x)\triangle(2x^2+5) + E(2x^2+5)\triangle(x^2+6x) = \\ &= (x^2+6x)(2\triangle(x^2)+\triangle(5)) + E(2x^2+5)(\triangle(x^2)+6\triangle(x)) = \\ &= (x^2+6x)(4x+2) + E(2x^2+5)(2x+1+6) = \\ &= (x^2+6x)(4x+2) + (2((x+1)^2+5)(2x+7) = \\ &= (x^2+6x)(4x+2) + (2x^2+4x+7)(2x+7).\end{aligned}$$ □

ПРИМЕР 11.27.

$$\begin{aligned}\triangle\left(\frac{x^2+2}{2x^2+6x+5}\right) &= \frac{(2x^2+6x+5)\cdot\triangle(x^2+2) - (x^2+2)\cdot\triangle(2x^2+6x+5)}{(2x^2+6x+5)\cdot E(2x^2+6x+5)} = \\ &= \frac{(2x^2+6x+5)\cdot(\triangle(x^2)+\triangle(2)) - (x^2+2)\cdot(2\triangle(x^2)+6\triangle(x)+\triangle(5))}{(2x^2+6x+5)\cdot(2(x+1)^2+6(x+1)+5)} = \\ &= \frac{(2x^2+6x+5)\cdot(2x+1) - (x^2+2)\cdot(2(2x+1)+6)}{(2x^2+6x+5)\cdot(2(x^2+2x+1)+6x+11)} = \\ &= \frac{(2x^2+6x+5)\cdot(2x+1) - (x^2+2)\cdot(4x+8)}{(2x^2+6x+5)\cdot(2x^2+10x+13)}.\end{aligned}$$ □

Как показывает приведенная ниже теорема, находить $\triangle f(x)$ для $f(x)=a^x$ гораздо легче.

ТЕОРЕМА 11.28. Если $f(x)=a^x$, то $\triangle f(x) = a^x(a-1)$. В частности, если $f(x)=2^x$, то $\triangle f(x)=2^x$.

ДОКАЗАТЕЛЬСТВО. $\triangle(a^x) = a^{x+1}-a^x = a^x(a-1)$. ■

ПРИМЕР 11.29.

$$\begin{aligned}\triangle(x^2+3x+4+2^x+4^x) &= \triangle(x^2)+3\triangle(x)+\triangle(4)+\triangle(2^x)+\triangle(4^x) = \\ &= 2x+1+3+2^x+3\cdot 4^x = \\ &= 2x+4+2^x+3\cdot 4^x.\end{aligned}$$ □

■ УПРАЖНЕНИЯ

1. Найдите $\Delta^2(x^3)$, вычисленное при $x=1$.
2. Найдите $\Delta^3(2^x)$, вычисленное при $x=1$.
3. Найдите $\Delta^2E^2(x^4)$, вычисленное при $x=4$.
4. Найдите $\Delta^2E(n!)$, вычисленное при $n=1$.
5. Найдите $\Delta((3x+2)(x-3))$.
6. Найдите $\Delta((4x-2)(x+5))$.
7. Найдите $\Delta\left(\dfrac{x-6}{3x+4}\right)$.
8. Найдите $\Delta\left(\dfrac{2x+3}{4x+1}\right)$.
9. Покажите, что $\Delta(x!) = x\cdot x!$

10. Докажите части (б), (в) и (г) теоремы 11.23 для действительного числа a и функций f и g

а) $\triangle(af) = a\triangle(f)$; **б)** $E(af) = aE(f)$; **в)** $E\triangle = \triangle E$;

г) $\triangle(a) = 0$; **д)** $E(a) = a$.

11. Докажите вторую часть теоремы 11.25.

$$\triangle\left(\frac{f}{g}\right) = \frac{g \cdot \triangle(f) - f \cdot (\triangle(g))}{g \cdot (E(g))}.$$

11.4. ФАКТОРИАЛЬНЫЕ МНОГОЧЛЕНЫ

Ранее была возможность убедиться в том, что нахождение конечных разностей функции x^n довольно запутанно. Это — неблагоприятный момент, поскольку полиномиальные функции весьма распространены. Для решения данной проблемы рассмотрим факториальную задачу. Пусть при $x \geq n \geq 0$

$$x^{(n)} = \begin{cases} x(x-1)(x-2)\cdots(x-n+1), & \text{если } n > 0; \\ 1, & \text{если } n = 0. \end{cases}$$

К счастью,

$$\begin{aligned} \triangle(x^{(n)}) &= (x+1)(x)(x-1)\cdots(x-n+2) - x(x-1)(x-2)\cdots(x-n+1) = \\ &= (x)(x-1)(x-2)\cdots(x-n+2)((x+1)-(x-n+1) = \\ &= (x)(x-1)(x-2)\cdots(x-n+2)(n) = \\ &= n(x)(x-1)(x-2)\cdots(x-(n-1)+1) = nx^{(n-1)}, \end{aligned}$$

так что теперь имеется функция с очень хорошими разностями. В действительности те, кто изучал математический анализ, возможно, надеялись на такой хороший результат для $x^{(n)}$. Заметим также, что $x^{(1)} = x$. Таким образом, если

$$f(x) = 5x^{(4)} + 6x^{(3)} - 3x^{(2)} + 2x^{(1)} + 7,$$

то

$$\triangle f(x) = (5 \cdot 4)x^{(3)} + (6 \cdot 3)x^{(2)} - 3 \cdot 2x^{(1)} + 2 = 20x^{(3)} + 18x^{(2)} - 6x^{(1)} + 2.$$

Если $f(x) = a_n x^{(n)} + a_{n-1}x^{(n-1)} + \cdots a_2 x^{(2)} + a_1 x + a_0$ для некоторого n, то f называется ***факториальным многочленом***. Очевидно, что вычислять разности факториального многочлена очень просто. Задача состоит в том, чтобы обычный многочлен записать в виде факториального. Ниже будет дан алгоритм решения этой задачи. Читатель может заметить, что здесь используется техника, подобная одной из тех, которые применялись в разделе 5.6 для перехода от представления числа на основе 10 к представлению числа на основе n.

Начнем с примера. Предположим, что обычный многочлен

$$3x^4 - 19x^3 + 34x^2 - 21x + 5$$

необходимо записать в виде факториального многочлена

$$a_4x^{(4)} + a_3x^{(3)} + a_2x^{(2)} + a_1x + a_0.$$

Напомним, что оба многочлена задают одну и ту же функцию. Они являются только ее разными представлениями.

Если разделить

$$3x^4 - 19x^3 + 34x^2 - 21x - 5$$

на x, то получим частное $3x^3 - 19x^2 + 34x - 21$ с остатком -5. Если разделить

$$a_4x^{(4)} + a_3x^{(3)} + a_2x^{(2)} + a_1x + a_0$$

на x, то получим

$$a_4(x-1)(x-2)(x-3) + a_3(x-1)(x-2) + a_2(x-1) + a_1$$

с остатком a_0. Следовательно, $a_0 = -5$ и

$$3x^3 - 19x^2 + 34x - 21 = \\ = a_4(x-1)(x-2)(x-3) + a_3(x-1)(x-2) + a_2(x-1) + a_1.$$

Если

$$3x^3 - 19x^2 + 34x - 21$$

разделить на $x-1$, то получим частное $3x^2 - 16x + 18$ с остатком -3.

Если

$$a_4(x-1)(x-2)(x-3) + a_3(x-1)(x-2) + a_2(x-1) + a_1$$

разделить на $x-1$, получим частное

$$a_4(x-2)(x-3) + a_3(x-2) + a_2$$

с остатком a_1. Поэтому $a_1 = -3$ и

$$3x^2 - 16x + 18 = a_4(x-2)(x-3) + a_3(x-2) + a_2.$$

Если разделить

$$3x^2 - 16x + 18$$

на $x-2$, то получим частное $3x - 10$ с остатком -2. Если разделить

$$a_4(x-2)(x-3) + a_3(x-2) + a_2$$

на $x-2$, то получим частное $a_4(x-3) + a_3$ и остаток a_2. Следовательно,

$$a_2 = -2$$

и

$$3x - 10 = a_4(x-3) + a_3.$$

Разделив $3x - 10$ на $x - 3$, получаем 3 и остаток -1.

Разделив $a_4(x - 3) + a_3$ на $x - 3$, получаем a_4 и остаток a_3. Поэтому

$$a_4 = 3 \qquad \text{и} \qquad a_3 = -1.$$

Итак, искомый факториальный многочлен имеет вид

$$3x^{(4)} - x^{(3)} - 2x^{(2)} - 3x + -5.$$

Окончательно для нахождения коэффициента a_0 факториального многочлена делим на x и берем остаток. Для нахождения коэффициента a_1 факториального многочлена делим частное на $x-1$ и берем остаток. Продолжаем процесс деления полученного каждый раз частного на $x-2$, $x-3$ и $x-4$ для нахождения остатков a_2, a_3 и a_4 соответственно.

Теперь представим алгоритм нахождения факториального многочлена $a_n x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_2 x^{(2)} + a_1 x^{(1)} + a_0 x$.

Алгоритм **Факториальный многочлен**:

Шаг 1. Для заданного многочлена $f(x)$ степени n положить $k = 0$.

Шаг 2. Разделить $f(x)$ на $x - k$, получив остаток r и частное $q(x)$.

Шаг 3. Положить $a_k = r$ и $f(x) = q(x)$.

Шаг 4. Если $k = n$, то процесс завершен. В противном случае положить $k = k + 1$ и вернуться к шагу 2.

Простейший способ реализовать этот процесс — использовать сокращенное деление. Проиллюстрируем его на примере, положив $f(x) = x^4 - 8x^3 + 21x^2 - 6x + 3$.

0	1	−8	21	−6	3
		0	0	0	0
1	1	−8	21	−6	[3]
		1	−7	14	
2	1	−7	14	[8]	
		2	−10		
3	1	−5	[4]		
		3			
4	1	[−2]			
	[1]				

Это дает факториальный многочлен $x^{(4)} - 2x^{(3)} + 4x^{(2)} + 8x + 3$.

Существуют несколько способов перехода от факториального многочлена к обычному. Один из них — просто раскрыть каждое слагаемое и собрать члены с одинаковыми степенями x. Другой метод мог бы состоять в вычислении факториального многочлена в $n+1$ точке, где n — степень многочлена, с последующей аппроксимацией этих значений многочленом степени n.

Третий метод представляет собой обратную форму сокращенного деления. Чтобы показать, как работает этот метод, рассмотрим факториальный многочлен

$$3x^{(4)} - x^{(3)} - 2x^{(2)} - 3x - 5.$$

Раскрыв каждое слагаемое, имеем

$$3x(x-1)(x-2)(x-3) - x(x-1)(x-2) - 2x(x-1) - 3x - 5.$$

Раскрыв $(x-3)$ в первом члене, оставив x с первым членом и присоединив -3 ко второму члену, получаем

$$3x^2(x-1)(x-2) + (3(-3)-1)x(x-1)(x-2) - 2x(x-1) - 3x - 5.$$

Теперь, раскрыв $(x-2)$ во втором члене, оставив x со вторым членом и присоединив -2 к третьему члену, получаем

$$3x^2(x-1)(x-2) - 10x^2(x-1) + ((-10)(-2)-2)x(x-1) - 3x - 5.$$

Теперь раскрытие $(x-1)$ в третьем члене, присоединение x к третьему члену, а -1 — к четвертому члену дает

$$3x^2(x-1)(x-2) - 10x^2(x-1) + 18x^2 + (-18-3)x - 5$$

или

$$3x^2(x-1)(x-2) - 10x^2(x-1) + 18x^2 - 21x - 5,$$

что совпадает с рассматриваемым многочленом после стадии 1. Заметим, что, используя приведенный ниже процесс, мы бы получили тот же самый результат в таблице

	3	−1	−2	−3	−5
		−9	20	−18	
3	3	−10	18	−21	
2					
1					

Заметим, что числа в верхней строке являются коэффициентами факториального многочлена, а числа в первом столбце — целые числа, начинающиеся с 3, на единицу меньше, чем степень многочлена. Переносим вниз 3, как при сокращенном делении, и умножаем на 3 первое число столбца. Но вместо сложения с -1, следующим числом строки, вычитаем 9 из -1, получая -10. Теперь берем -10, умножаем на 2, следующее число в столбце, и вычитаем из -2, следующего числа в строке. Получаем 18. Продолжаем процесс, умножая 18 на 1 и вычитая результат из -3, получаем -21. Последнее число в строке не используется и остается неизменным. Легко видеть, что числа в третьей строке вместе с последним числом в первой строке являются коэффициентами многочлена на стадии 1.

Теперь берем многочлен на первой стадии

$$3x^2(x-1)(x-2) - 10x^2(x-1) + 18x^2 - 21x - 5$$

и повторяем процесс. Раскрываем $x-2$, оставляем x с первым членом, присоединяем -2 ко второму члену, после чего получаем

$$3x^3(x-1)+((3)(-2)-10)x^2(x-1)+18x^2-21x-5.$$

Теперь раскрываем $x-1$ во втором члене, оставляя x со вторым членом, присоединяя -1 к третьему члену, что дает

$$3x^3(x-1)-16x^3+(16+18)x^2-21x-5$$

или

$$3x^3(x-1)-16x^3(x-1)+34x^2-21x-5$$

для многочлена второй стадии. Обратите внимание, что каждая стадия повышает степень x в том слагаемом, в котором раскрывается множитель.

Снова, возвращаясь к таблице, повторяем процесс за исключением того, что сносим 3, первый элемент строки, умножаем на 2 и вычитаем результат из -10, числа во втором столбце. Это дает -16. Умножаем -16 на 1 и вычитаем результат из 18, получая 34. Опять -21, последнее число в строке, не используется. Снова замечаем, что оставшаяся строка вместе с числами -21 и -5 формирует многочлен второй стадии:

	3	-1	-2	-3	-5
		-9	20	-18	
3	3	-10	18	-21	
		-6	16		
2	3	-16	34		
1					

Теперь берем многочлен второй стадии

$$3x^3(x-1)-16x^3+34x^2-21x-5$$

и раскрываем $x-1$ в первом члене. Это дает

$$3x^4-19x^3+34x^2-21x-5,$$

что является искомым многочленом.

Снова, глядя на таблицу, сносим 3 и умножаем на 1. Вычитаем результат из -16. Это оставляет неиспользованным -19. Наконец, сносим 3. Замечаем, что

числа в скобках являются искомыми коэффициентами.

	3	−1	−2	−3	[−5]
		−9	20	−18	
3	3	−10	18	[−21]	
		−6	16		
2	3	−16	[34]		
		−3			
1	3	[−19]			
	[3]				

Продемонстрировав этот процесс, предоставляем читателю четко сформулировать алгоритм перехода от факториального многочлена к обычному.

ПРИМЕР 11.30. Заменим многочлен $x^{(4)} - 2x^{(3)} + 4x^{(2)} + 6x + 3$ обычным многочленом. Используя обратное сокращенное деление, имеем

	1	−2	4	6	[3]
		−3	10	−14	
3	1	−5	14	[−8]	
		−2	7		
2	1	−7	[21]		
		−1			
1	1	[−8]			
	[1]				

В результате получаем многочлен $f(x) = x^4 - 8x^3 + 21x^2 - 8x + 3$. □

Было показано, что

$$\begin{aligned} x^{(n)} &= x(x-1)(x-2)\cdots(x-(n-1)+1)(x-n+1) = \\ &= (x-n+1)x(x-1)(x-2)\cdots(x-(n-1)+1) = \\ &= (x-n+1)x^{(n-1)}. \end{aligned}$$

Решая относительно $x^{(n-1)}$, имеем

$$x^{(n-1)} = \frac{x^{(n)}}{x-n+1}.$$

Теперь имеется способ определить $x^{(n)}$, когда n — отрицательное целое число.

Положив $n = 1$, имеем

$$x^{(0)} = \frac{x^{(1)}}{x - 1 + 1} = 1.$$

Положив $n = 0$, имеем

$$x^{(-1)} = \frac{x^{(0)}}{x - 0 + 1} = \frac{1}{x + 1}.$$

Теперь воспользуемся индукцией и соотношением

$$x^{(n-1)} = \frac{x^{(n)}}{x - n + 1},$$

чтобы показать, что

$$x^{(-m)} = \frac{1}{(x + m)^{(m)}}$$

для всех $m \geq 1$, при которых знаменатель не равен нулю.

Было показано, что

$$x^{(-1)} = \frac{1}{x + 1} = \frac{1}{(x + 1)^{(1)}}.$$

Предположим, что

$$x^{(-k)} = \frac{1}{(x + k)^{(k)}}.$$

Тогда

$$\begin{aligned} x^{(-(k+1))} = x^{(-k-1))} &= \frac{x^{(-k)}}{x - (-k) + 1} = \\ &= \frac{1}{(x + k)^{(k)}} \cdot \frac{1}{x + k + 1} = \\ &= \frac{1}{(x + k)(x + k - 1) \cdots (x + 1)(x + k + 1)} = \\ &= \frac{1}{(x + k + 1)(x + k)(x + k - 1) \cdots (x + 1))} = \\ &= \frac{1}{(x + k + 1)^{(k+1)}}. \end{aligned}$$

Теперь необходимо найти $\triangle(x^{-n})$. По определению

$$\begin{aligned} \triangle(x^{-n}) &= \triangle\left(\frac{1}{(x + n)^{(n)}}\right) = \frac{1}{(x + n + 1)^{(n)}} - \frac{1}{(x + n)^{(n)}} = \\ &= \frac{1}{(x + n + 1)(x + n)(x + n - 1) \cdots (x + 2)} - \frac{1}{(x + n)(x + n - 1) \cdots (x + 1)} = \\ &= \frac{x + 1 - (x + n + 1)}{(x + n + 1)(x + n)(x + n - 1) \cdots (x + 1)} = \frac{-n}{(x + n + 1)^{(n+1)}} = -nx^{-n-1}, \end{aligned}$$

поэтому для каждого целого числа n, не равного нулю, имеем $\Delta(x^{(n)}) = nx^{(n-1)}$.

Теперь есть возможность обобщить одну из комбинаторных теорем. Для положительного целого числа n пусть

$$\binom{x}{n} = \frac{x^{(n)}}{n!}.$$

Когда x — целое число, это обычное определение числа сочетаний.

Взяв разность, находим

$$\Delta\binom{x}{n} = \frac{nx^{(n-1)}}{n!} = = \frac{x^{(n-1)}}{(n-1)!} = = \binom{x}{n-1}.$$

Но

$$\Delta\binom{x}{n}\binom{x+1}{n} - \binom{x}{n},$$

поэтому имеем

$$\binom{x-1}{n-1} = \binom{x+1}{n} - \binom{x}{n}$$

или

$$\binom{x+1}{n} = \binom{x}{n} + \binom{x}{n-1}.$$

Положив $y = x + 1$, получаем

$$\binom{y}{n} = \binom{y-1}{n} + \binom{y-1}{n-1}$$

в более привычной форме, которая ранее была доказана для случая, когда y — целое число.

ОПРЕДЕЛЕНИЕ 11.31. ***Числа Стирлинга первого рода*** определены соотношениями

$$s_0^{(n)} = 0 \qquad \text{при всех } n \geq 1;$$
$$s_n^{(n)} = 1 \qquad \text{при всех } n \geq 0.$$
$$s_k^{(n+1)} = s_{k-1}^{(n)} + n s_k^{(n)}$$

ОПРЕДЕЛЕНИЕ 11.32. ***Числа Стирлинга второго рода*** определены соотношениями

$$S_0^{(n)} = 0 \qquad \text{при всех } n \geq 1;$$
$$S_n^{(n)} = 1 \qquad \text{при всех } n \geq 0.$$
$$S_k^{(n+1)} = S_{k-1}^{(n)} + k S_k^{(n)}$$

Покажем, что числа Стирлинга первого и второго рода могут быть описаны на основе многочленов. Числа Стирлинга первого рода равны абсолютным значениям коэффициентов многочлена, равного $x^{(n)}$, для $n = 1, 2, 3, \ldots$

ТЕОРЕМА 11.33. Пусть

$$x^{(n)} = s_n^{(n)} x^n - s_{n-1}^{(n)} x^{n-1} + s_{n-2}^{(n)} x^{n-2} - \cdots \\ \cdots + (-1)^{n-2} s_2^{(n)} x^2 + (-1)^{n-1} s_1^{(n)} x + (-1)^n s_0^{(n)}.$$

Коэффициенты $s_i^{(n)}$ являются ***числами Стирлинга первого рода***.

ДОКАЗАТЕЛЬСТВО. По определению, $s_0^{(n)} = 0$ для всех $n \geq 1$ и $s_n^{(n)} = 1$ для всех $n \geq 0$. Чтобы показать, что $s_i^{(n)}$ — числа Стирлинга первого рода, требуется показать, что $s_i^{(n+1)} = s_{i-1}^{(n)} + n \cdot s_i^{(n)}$ для $1 \leq i \leq n$.

По определению, $x^{(n+1)} = (x - n)x^{(n)} = x \cdot x^{(n)} - n \cdot x^{(n)}$. Следовательно,

$$\sum_{i=0}^{n+1} (-1)^{n+1-i} s_i^{(n+1)} x^i = x \cdot \sum_{i=0}^{n} (-1)^{n-i} s_i^{(n)} x^i - n \cdot \sum_{i=0}^{n} (-1)^{n-i} s_i^{(n)} x^i = \\ = \sum_{i=0}^{n} (-1)^{n-i} s_i^{(n)} x^{i+1} - n \cdot \sum_{i=0}^{n} (-1)^{n-i} s_i^{(n)} x^i = \\ = \sum_{i=1}^{n+1} (-1)^{n-(i-1)} s_{i-1}^{(n)} x^i - n \cdot \sum_{i=0}^{n} (-1)^{n-i} s_i^{(n)} x^i.$$

Увеличивая i на 1 в первой сумме, получаем

$$\sum_{i=1}^{n+1} (-1)^{n-i+1} s_{i-1}^{(n)} x^i + n \cdot \sum_{i=0}^{n} (-1)^{n-i+1} s_i^{(n)} x^i.$$

Приравнивая коэффициенты при x^i для $1 \leq i \leq n$, имеем $s_i^{(n+1)} = s_{i-1}^{(n)} + n s_i^{(n)}$, что и требовалось доказать. ■

Покажем, что числа Стирлинга второго рода есть не что иное, как коэффициенты факториальных многочленов, равных x^n для $n = 1, 2, 3, \cdots$.

ТЕОРЕМА 11.34. Пусть

$$x^n = S_n^{(n)} x^{(n)} + S_{n-1}^{(n)} x^{(n-1)} + S_{n-2}^{(n)} x^{(n-2)} + \cdots S_2^{(n)} x^{(2)} + S_1^{(n)} x + S_0^{(n)}.$$

Коэффициенты $s_i^{(n)}$ являются ***числами Стирлинга второго рода***.

ДОКАЗАТЕЛЬСТВО. По определению, $S_0^{(n)} = 0$ для всех $n \geq 1$ и $S_n^{(n)} = 1$ для всех $n \geq 0$. Требуется показать, что для $1 \leq i \leq n$

$$S_i^{(n+1)} = S_{i-1}^n + i \cdot S_i^n.$$

По определению,

$$\begin{aligned}
\sum_{i=0}^{n+1} S_i^{(n+1)} x^{(i)} = x^{n+1} = x \cdot x^n = x \cdot \sum_{i=0}^{n} S_i^{(n)} x^{(i)} = \\
= (i + x - i) \cdot \sum_{i=0}^{n} S_i^{(n)} x^{(i)} = \\
= i \cdot \sum_{i=0}^{n} S_i^{(n)} x^{(i)} + (x - i) \cdot \sum_{i=0}^{n} S_i^{(n)} x^{(i)} = \\
= i \cdot \sum_{i=0}^{n} S_i^{(n)} x^{(i)} + \sum_{i=0}^{n} S_i^{(n)} (x - i) \cdot x^{(i)} = \\
= i \cdot \sum_{i=0}^{n} S_i^{(n)} x^{(i)} + \sum_{i=0}^{n} S_i^{(n)} x^{(i+1)} = \\
= i \cdot \sum_{i=0}^{n} S_i^{(n)} x^{(i)} + \sum_{i=1}^{n+1} S_{i-1}^{(n)} x^{(i)}.
\end{aligned}$$

Приравнивая коэффициенты при x^i для $1 \leq i \leq n$, получаем $S_i^{(n+1)} = i \cdot S_i^{(n)} + S_{i-1}^{(n)}$, что и требовалось доказать. ■

■ УПРАЖНЕНИЯ

1. Найдите $\Delta^2(3x^{(3)} - 5x^{(2)} + 4)$.

2. Найдите $\Delta^4(4x^{(5)} - 3x^{(4)} + 6x^{(3)} + x^{(2)} - 3x + 4$.

3. Найдите $\Delta^3(x^{(5)} + 3x^{(4)} - 3x^{(3)} - 2x^{(2)} + 4x - 1$.

4. Найдите $\Delta(3x^{(2)} + 2x - 5)(x^{(4)} + 4x^{(3)} + 2x^{(2)})$.

5. Найдите $\Delta(x^{(3)} - 2x^{(2)} - 4)(x^{(4)} + 2x^{(3)} - 6x^{(2)} + x - 3)$.

6. Найдите $\Delta\left(\dfrac{x^{(3)} + 6x^{(2)} + 3x}{x^{(5)} + 3x^{(3)}}\right)$.

7. Найдите $\Delta\left(\dfrac{x^{(4)} - 3x^{(2)} + 3x}{x^{(4)} - 3x^{(3)} + 6x^{(2)}}\right)$.

8. Упростите выражение

$$\binom{y+1}{n+1}\Delta^n u_{x+1} - \binom{y}{n+1} u_{x+1} - \binom{y}{n}\Delta^n u_x.$$

9. Найдите $\Delta^3((x-3)(x-5)(x-7))$.

10. Найдите $\Delta^3\binom{x}{8}$.

11. Найдите $\Delta^4\binom{x}{5}$ при $x = 0$.
12. Преобразуйте $x^4 - 2x^3 + 3x^2 + 5$ в факториальный многочлен.
13. Преобразуйте $x^4 - 6x^3 + 4x^2 - 4x + 5$ в факториальный многочлен.
14. Преобразуйте $x^4 - x^3 + x^2 - x + 1$ в факториальный многочлен.
15. Преобразуйте $x^5 - x^4 + x^3 - x^2 + x - 1$ в факториальный многочлен.
16. Преобразуйте $x^{(4)} - x^{(3)} + x^{(2)} - x + 1$ в обычный многочлен.
17. Преобразуйте $x^{(4)} + 3x^{(3)} - 4x^{(2)} - 2x + 1$ в обычный многочлен.
18. Преобразуйте $4x^{(4)} - 2x^{(3)} + 3x^{(2)} - 4x + 1$ в обычный многочлен.
19. Преобразуйте $x^{(5)} - x^{(4)} + x^{(3)} - x^{(2)} + x - 1$ в обычный многочлен.
20. Постройте алгоритм перехода от факториальных многочленов к обычным.
21. Найдите $\Delta^3\binom{8}{x}$, вычисленное при $x = 0$.

11.5. СУММИРОВАНИЕ РАЗНОСТЕЙ

Мы знаем, что

$$\triangle^n(f(x)) = f(x+n) - n \cdot f(n-1) + \cdots + (-1)^k \binom{n}{k} (f(x+n-k)) + \cdots - f(x).$$

Поэтому, если x принимает только целочисленные значения, любое разностное отношение может быть представлено в виде рекуррентного отношения. Например,

$$\triangle^3(f(x)) - 3\triangle^2(f(x)) = 2\triangle f(x)$$

может быть записано как

$$f(x+3) - 3f(x+2) + 3f(x+1) + f(x) - \\ - 3(f(x+2) - 2f(x+1) + f(x)) = 2(f(x+1) - f(x)),$$

так что

$$f(x+3) = 6f(x+2) - 7(f(x+1) - 4f(x),$$

что является записью в рекурсивной форме. Поэтому для решения этого отношения можно использовать результаты первых трех разделов данной главы.

В этом разделе мы используем несколько иной подход. Определим $\sum$ как оператор, обратный разностному. То есть, желательно, чтобы из $\Delta F(x) = f(x)$ следовало $\sum f(x) = F(X)$. Мы установим, однако, что на самом деле, если $\Delta F(x) = f(x)$, то $\sum f(x) = F(X) + c$, где c — константа, так как $\Delta(F(x) + c) = \Delta F(x) + \Delta(c) = \Delta F(x)$. Те, кто знаком с математическим анализом, без труда узнают в этом аналог первообразной функции или неопределенного интеграла.

Из того, что $\Delta(F(x)) = 0$, вообще говоря, не следует, что $F(x)$ — константа. Рассмотрим, например, $F(x) = \sin(2\pi x)$:

$$\Delta F(x) = \sin(2\pi(x+1)) - \sin(2\pi x) =$$

$$= \sin(2\pi x + 2\pi) - \sin(2\pi x) =$$
$$= \sin(2\pi x) - \sin(2\pi x) = 0.$$

Это объясняется тем, что $\sin(x)$ — периодическая функция. Но в дальнейшем мы ограничимся рассмотрением многочленов, рациональных и экспоненциальных функций. Для таких функций, если $\Delta(F(x)) = 0$, то $F(x)$ — константа.

ПРИМЕР 11.35. Используя разностные формулы предыдущего раздела, находим, что если $f(x) = \dfrac{x^{(n+1)}}{n+1}$, то

$$\begin{aligned}
\Delta f(x) &= \Delta\frac{(x^{(n+1)})}{n+1} = \\
&= \frac{1}{n+1}\Delta(x^{(n+1)}) = && \text{по теореме 11.23} \\
&= \frac{1}{n+1}(n+1)x^{(n)} = && \text{поскольку } \Delta(x^{(n)}) = nx^{(n-1)} \\
&= x^{(n)},
\end{aligned}$$

так что

$$\sum x^{(n)} = \frac{x^{(n+1)}}{n+1} + c.$$

Пусть

$$\Delta f(x) = 15x^{(4)} + 8x^{(3)} + 6x^{(2)} + 4x^{(1)} + 5,$$

тогда

$$f(x) = \frac{15x^{(5)}}{5} + \frac{8x^{(4)}}{4} + \frac{6x^{(3)}}{3} + \frac{4x^{(2)}}{2} + 5x + c =$$
$$= 3x^{(5)} + 2x^{(4)} + 2x^{(3)} + 2x^{(2)} + 5x + c. \qquad \square$$

Далее, мы видим, что если $f(x) = \left(\dfrac{a^n}{a-1}\right)$, то

$$\begin{aligned}
\Delta f(x) &= \Delta\left(\frac{a^x}{a-1}\right) = \\
&= \frac{1}{a-1}\Delta a^x = && \text{по теореме 11.23} \\
&= \frac{1}{a-1}(a^x)(a-1) = && \text{показано в предыдущем разделе} \\
&= (a^x),
\end{aligned}$$

так что $\sum(a^x) = \left(\dfrac{a^x}{a-1}\right) + c.$

ПРИМЕР 11.36. Пусть $\Delta f(x) = 15x^{(4)} + 3x^{(2)} + 2^x + 8 \cdot 5^x$. Тогда

$$f(x) = \frac{15x^{(5)}}{5} + \frac{3x^{(3)}}{3} + 2^x + \frac{8 \cdot 5^x}{5-1} + c =$$

$$= 3x^{(5)} + x^{(3)} + 2^x + 2 \cdot 5^x + c. \qquad \square$$

Также пусть

$$f(x) = \binom{x}{n+1}, \qquad \text{тогда} \qquad \Delta f(x) = \binom{x}{n},$$

как показано в предыдущем разделе, так что

$$\sum \binom{x}{n} = \binom{x}{n+1} + c.$$

ПРИМЕР 11.37. Пусть $\Delta f(x) = 5x^{(4)} + 6 \cdot 4^x + 6 \cdot \binom{10}{5}$. Тогда

$$f(x) = x^{(5)} + 2 \cdot 4^x + 6 \cdot \binom{10}{6} + c. \qquad \square$$

Наконец, рассмотрим $f(x) = \dfrac{1}{1-n} x^{(-n+1)}$.

$$\begin{aligned} \Delta f(x) &= \Delta \frac{x^{(-n+1)}}{1-n} = \\ &= \frac{1}{1-n} \Delta x^{(-n+1)} = \qquad \text{по теореме 11.23} \\ &= \frac{1}{1-n}(-n+1)x^{(-n+1-1)} = x^{(-n)}. \end{aligned}$$

Следовательно, $\sum x^{(-n)} = -\dfrac{x^{(-n+1)}}{-n+1} + c.$

ПРИМЕР 11.38. Пусть $\Delta f(x) = x^{(2)} + 2^x + 5^x + x^{(-3)} + x^{(-7)}$, тогда

$$f(x) = \frac{1}{3}x^{(3)} + 2^x + \frac{1}{4}5^x - \frac{1}{2}x^{(-2)} + \frac{1}{6}x^{(-6)} + c. \qquad \square$$

Теперь рассмотрим метод, который называется ***суммированием по частям***.

ТЕОРЕМА 11.39. $\sum f(x)\Delta g(x) = f(x)g(x) - \sum E(g(x))\Delta f(x) + c.$

ДОКАЗАТЕЛЬСТВО. Из теоремы 11.25 имеем

$$\triangle(fg) = f \cdot \triangle(g) + E(g) \cdot \triangle(f).$$

Поэтому

$$\sum \triangle(fg) = \sum f \cdot \triangle(g) + \sum E(g) \cdot \triangle(f),$$

так что

$$\sum f \cdot \triangle(g) = (fg) - \sum E(g) \cdot \triangle(f). \qquad \blacksquare$$

Рассмотрим теперь суммы, которые являются аналогом определенных интегралов в математическом анализе. Пусть

$$f(x) = \Delta F(x) = F(x+1) - F(x),$$

тогда

$$\sum_{x=1}^{n} f(x) = \sum_{x=1}^{n} (F(x+1) - F(x)) = F(n+1) - F(1).$$

Обозначим $F(n+1) - F(1)$ через $F(x)\big|_1^{n+1}$.

ПРИМЕР 11.40. Пусть $f(x) = x^{(3)} + 3^x + x^{(-4)}$, тогда

$$\sum_{x=1}^{7} f(x) = \sum_{x=1}^{7} x^{(3)} + 3^x + x^{(-4)} = \left(\frac{x^{(4)}}{4} + \frac{1}{2}3^x - \frac{1}{3}x^{(-3)}\right)\Bigg|_1^8 =$$
$$= \frac{8^{(4)}}{4} + \frac{1}{2}3^8 - \frac{1}{3}8^{(-3)} - \left(\frac{1^{(4)}}{4} + \frac{1}{2}3 + \frac{1}{3}3^{(-3)}\right). \qquad \square$$

ПРИМЕР 11.41. Найти сумму $1^2 + 2^2 + 3^2 + \cdots + n^2$.

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \sum_{x=1}^{n} x^2 =$$
$$= \sum_{x=1}^{n} x^{(2)} + x =$$
$$= \left(\frac{x^{(3)}}{3} + \frac{x^{(2)}}{2}\right)\Bigg|_1^{n+1} =$$
$$= \frac{1}{6}\left(2x^{(3)} + 3x^{(2)}\right)\Big|_1^{n+1} =$$
$$= \frac{1}{6}\left(2x^3 - 3x^2 + x\right)\Big|_1^{n+1} =$$
$$= \frac{1}{6}(2(n+1)^3 - 3(n+1)^2 + (n+1)) - 0 =$$
$$= \frac{1}{6}n(n+1)(2n+1). \qquad \square$$

ПРИМЕР 11.42. $\sum_{x=0}^{n} r^x = \frac{r^x}{r-1}\Big|_0^{n+1} = \frac{r^{n+1}-1}{r-1}.$ $\square$

ПРИМЕР 11.43. $\sum_{x=1}^{n} x = \frac{x^{(2)}}{2}\Big|_1^{n+1} = \frac{x(x-1}{2}\Big|_1^{n+1} = \frac{(n+1)n}{2}.$ $\square$

ПРИМЕР 11.44. Как упоминалось выше,

$$\Delta(fg) = f \cdot \Delta(g) + E(g) \cdot \Delta(f).$$

Поэтому

$$\sum_{x=1}^{n} \Delta(f(x)g(x)) = \sum_{x=1}^{n} f(x) \cdot \Delta(g(x)) + \sum_{x=1}^{n} E(g(x)) \cdot \Delta(f(x))$$

или

$$(f(x)g(x))\Big|_1^{n+1} = \sum_{x=1}^{n} f(x) \cdot \Delta(g(x)) + \sum_{x=1}^{n} g(x+1) \cdot \Delta(f(x)).$$

Следовательно,

$$\sum_{x=1}^{n} f(x) \cdot \Delta(g(x)) = (f(x)g(x))\Big|_1^{n+1} - \sum_{x=1}^{n} g(x+1) \cdot \Delta(f(x)). \qquad \square$$

ПРИМЕР 11.45.

$$\begin{aligned}
\sum_{x=1}^{n} x3^x &= \sum_{x=1}^{n} x\Delta\frac{3^x}{2} = \\
&= \frac{x3^x}{2}\Big|_1^{n+1} - \sum_{x=1}^{n} \frac{3^{x+1}}{2} \cdot \Delta x = \\
&= \frac{(n+1)3^{n+1}}{2} - \frac{3}{2} - \left(\frac{3^{x+1}}{4}\right)\Big|_1^{n+1} = \\
&= \frac{(n+1)3^{n+1}}{2} - \frac{3}{2} - \left(\left(\frac{3^{n+2}}{4}\right) - \frac{9}{4}\right) = \\
&= \frac{(2n-1)3^{n+1}+3}{4}.
\end{aligned} \qquad \square$$

■ УПРАЖНЕНИЯ

1. Найдите сумму разностей первого порядка для функции $3^x - x^3 + x^2 - x + 5$ от $x = 1$ до $x = 9$.
2. Найдите сумму разностей первого порядка для функции $2^x - 3x^3 + 2x^2$ от $x = 1$ до $x = 5$.
3. Найдите $\sum_{i=4}^{8} 2^x - 12x^{(3)}$.
4. Найдите $\sum_{i=6}^{10} 3^x + 10x^{(4)} - 8x^{(3)}$.
5. Найдите $\sum_{i=4}^{7} \binom{x}{2} + 4^x + 12x^{(5)}$.
6. Найдите $\sum_{i=6}^{11} \binom{x}{4} + 3 \cdot 5^x + 14x^{(6)}$.

7. Покажите, что $\sum_{i=1}^{n} 2i = n(n+1)$.

8. Покажите, что $\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$.

9. Покажите, что $\sum_{i=1}^{n} i(i+1) = \frac{n(n+1)(n+2)}{3}$.

10. Покажите, что $\sum_{i=1}^{n} 4i - 2 = 2n^2$.

11. Используя суммирование по частям, найдите $\sum_{x=1}^{8} (2x+3)3^x$.

12. Используя суммирование по частям, найдите $\sum_{x=2}^{10} 4(3x-6)4^x$.

13. Используя суммирование по частям, найдите $\sum_{x=0}^{8} (x)(x-1)2^x$.

14. Используя суммирование по частям, найдите $\sum_{x=0}^{8} (x)(x-1)2^x$.

15. Используя суммирование по частям, найдите $\sum_{x=0}^{8} x^2 2^x$.

16. Используя суммирование по частям, найдите $\sum_{x=0}^{10} x^{(2)} 2^x$.

17. Используя суммирование по частям, найдите $\sum_{x=0}^{10} x^{(4)} 4^x$.

18. Используя суммирование по частям, найдите $\sum_{x=0}^{9} x^{(4)} x^{(6)}$.

19. Используя суммирование по частям, найдите $\sum_{x=0}^{9} x^{(2)} x^{(8)}$.

20. Докажите, что $\Delta \sum_{i=c}^{x} f(i) = f(x+1)$.

21. Докажите, что $\Delta \sum_{i=x}^{c} f(i) = -f(x)$.

ГЛАВА

12

СНОВА О КОМБИНАТОРНЫХ ПОДСЧЕТАХ

12.1. ЗАДАЧИ О РАЗМЕЩЕНИИ

В этом разделе мы займемся подсчетом числа способов размещения объектов в ящики или урны. Подобные задачи нами уже рассматривались. Так, в разделе 8.3 было показано, что число способов поместить m из n различимых объектов в один ящик, а остальные $n - m$ — в другой, равно

$$C(n, m) = \frac{n!}{m!(n-m)!}.$$

Более того, мы показали, что если имеется n различимых объектов и k различимых ящиков, то число способов поместить n_i объектов в i-ый ящик для всех $1 \leq i \leq k$, где $n = n_1 + n_2 + n_3 + \cdots + n_k$, определяется по формуле

$$C(n; n_1, n_2, n_3, \cdots, n_k) = \frac{n!}{n_1! n_2! n_3! \cdots n_k!}.$$

Отметим, что в каждом из этих случаев ящик может быть пуст, а объекты и ящики являются различимыми.

Рассмотрим теперь количество способов размещения n различимых объектов в k различимых ящиках при условии, что некоторые из ящиков могут быть пустыми. Отметим, что количество объектов, помещаемых в каждый ящик, не фиксировано, поэтому существуют k возможностей выбора ящика для размещения каждого из n объектов. Согласно комбинаторному принципу умножения существуют k^n способов разместить объекты в ящиках. Это число совпадает с количеством функций из множества S, содержащего n объектов, во множество T, содержащее k объектов, что нетрудно понять, если отождествить отображение объекта s из множества S в объект t из множества T с размещением объекта s в ящик t.

ТЕОРЕМА 12.1. Существуют k^n способов размещения n различимых объектов в k различимых ящиках, когда некоторые из ящиков могут быть пустыми.

Задача значительно усложняется в ситуации, когда ни один ящик не может быть пустым, поэтому мы отложим ее рассмотрение на потом.

В разделе 8.6 было показано, что число способов размещения n неразличимых объектов в k различимых ящиках, некоторые из которых могут быть пустыми, равно

$$C(n+k-1,n) = C(n+k-1,k-1) = \frac{(n+k-1)!}{n!(k-1)!}.$$

Кроме того, показано, что число способов размещения n неразличимых объектов в k различимых ящиках, ни один из которых не может быть пустым, равно

$$C(n-1,n-k) = C(n-1,k-1) = \frac{(n-1)!}{(n-k)!(k-1)!}.$$

Подсчитаем теперь количество способов размещения n различимых шаров в k неразличимых ящиках. Сначала рассмотрим случай, когда ни один из ящиков не может быть пустым. *Это также будет числом способов разбиения множества, содержащего n объектов на k непересекающихся подмножеств.*

ТЕОРЕМА 12.2. Существуют $S_k^{(n)}$ способов размещения n различимых шаров в k неразличимых ящиках, когда ни один из ящиков не может быть пустым, где $\{S_k^{(n)} : 0 \le k \le n\}$ — множество ***чисел Стирлинга второго рода***.

ДОКАЗАТЕЛЬСТВО. Требуется показать, что если $f(n,k)$ — количество способов размещения n различимых шаров в k неразличимых ящиках, когда ни один из ящиков не может быть пустым, то $f(n,n) = 1$ для всех $n \ge 0$, $f(n,0) = 0$ для всех $n \ge 1$ и $f(n+1,k) = f(n,k-1) + kf(n,k)$. Очевидно, что для $n \ge 1$ имеется только один способ размещения n шаров в n ящиках, так что ни один из ящиков не будет пуст. Поэтому $f(n,n) = 1$ для всех $n \ge 1$. Случай $n = 0$ несколько надуманный, но представляется разумным предположить, что существует только один способ разместить 0 объектов в 0 ящиках. Поэтому положим $f(0,0) = 1$.

Очевидно, что для $n \ge 1$ не существует способа разместить 0 объектов в n ящиках и не оставить ни один ящик пустым, поэтому $f(n,0) = 0$ для всех $n \ge 1$.

Чтобы доказать соотношение $f(n+1,k) = f(n,k-1) + kf(n,k)$, предположим, что n объектов размещены в k ящиках и нужно поместить в ящик $(n+1)$-ый объект. Имеются две возможности: либо положить объект в пустой ящик, либо — в уже занятый. В первом случае остальные n объектов должны быть размещены в оставшихся ящиках. Поскольку имеется $k-1$ таких ящиков, то n объектов следует поместить в $k-1$ ящик. Существует $f(n,k-1)$ способов такого размещения. Если $(n+1)$-ый объект помещается в уже занятый ящик, то предыдущие n объектов были размещены в k ящиках. Это возможно сделать $f(n,k)$ способами. Следующий, $(n+1)$-ый, объект может быть помещен только в один из k ящиков, поэтому существуют $k \cdot f(n,k)$ способов размещения объектов при условии, что ни один из k ящиков не пуст. Следовательно, существуют $f(n,k-1) + kf(n,k)$ способов разместить $n+1$ объектов в k ящиках, и

$$f(n+1,k) = f(n,k-1) + kf(n,k).$$

Отсюда, $f(n,k) = S_k^{(n)}$, где $\{S_k^{(n)} : 0 \le k \le n\}$ — множество чисел Стирлинга второго рода. ■

Теперь, используя свойство $S_k^{(n+1)} = S_{k-1}^{(n)} + kS_k^{(n)}$, можно сформировать таблицу чисел Стирлинга второго рода аналогично тому, как это было сделано в случае треугольника Паскаля.

Треугольник для чисел Стирлинга второго рода

n	$S_0^{(n)}$	$S_1^{(n)}$	$S_2^{(n)}$	$S_3^{(n)}$	$S_4^{(n)}$	$S_5^{(n)}$	$S_6^{(n)}$	$S_7^{(n)}$	$S_8^{(n)}$	$S_9^{(n)}$	$S_{10}^{(n)}$
0	1										
1	0	1									
2	0	1	1								
3	0	1	3	1							
4	0	1	7	6	1						
5	0	1	15	25	10	1					
6	0	1	31	90	65	15	1				
7	0	1	63	301	350	140	21	1			
8	0	1	127	966	1701	1050	266	28	1		
9	0	1	255	3025	7770	6951	2646	462	36	1	
10	0	1	511	9330	34105	42525	22827	5880	750	45	1

ТЕОРЕМА 12.3. Снимем ограничение на то, что ни один из ящиков не может быть пустым. Возможны ситуации: (1) все n объектов помещены в один ящик, что можно осуществить $S_1^{(n)}$ способами; (2) все n объектов помещены в два ящика, что можно осуществить $S_2^{(n)}$ способами, ..., (k) все n объектов помещены в k ящиков, что можно осуществить $S_k^{(n)}$ способами, и т.д. Итого, имеется

$$S_1^{(n)} + S_2^{(n)} + S_3^{(n)} + \cdots + S_k^{(n)}$$

способов размещения n объектов в k ящиках при условии, что ящики могут быть пустыми.

Используя тот факт, что $S_k^{(n)}$ — число возможных способов размещения n различимых объектов в k неразличимых ящиках, ни один из которых не может быть пустым, мы можем найти количество способов размещения n различимых объектов в k различимых ящиках, ни один из которых не может быть пустым. По сути, нам нужно перейти от неупорядоченных ящиков к упорядоченным. При переходе от упорядоченных ящиков к неупорядоченным нужно делить соответствующее количество способов размещения на $k!$, как было сделано при переходе от перестановок к сочетаниям. Поэтому при обратном переходе нужно умножать на $k!$. Таким образом, существуют $k!S_k^{(n)}$ способов размещения n различимых объектов в k неразличимых ящиках, ни один из которых не может быть пустым.

Пусть S — множество, содержащее n объектов, и T — множество, содержащее k объектов. Ранее было подсчитано количество всех функций из S в T, а также количество биективных функций из S в T. Теперь мы можем подсчитать количество сюръективных функций из S в T. Представляя k объектов множества T как ящики, отображение элемента s из S в элемент t из T можно отождествить с размещением объекта s в ящик t. Поскольку функция сюръективна, то ни один

из ящиков не будет пустым. *Следовательно, существуют* $k!S_k^{(n)}$ *сюръективных функций из* n*-элементного множества в* k*-элементное множество.*

Теперь мы можем построить следующую таблицу:

	Размещение n объектов в k ящиках	**Количество способов**
(1)	объекты — различимые, ящики — различимые, n_i объектов в i-ом ящике	$C(n; n_1, n_2, n_3, \cdots, n_k)$
(2)	объекты — различимые, ящики — различимые, ящики могут быть пустыми	k^n
(3)	объекты — различимые, ящики — различимые, ящики не могут быть пустыми	$k!S_k^{(n)}$
(4)	объекты — неразличимые, ящики — различимые, ящики могут быть пустыми	$C(n+k-1, n)$
(5)	объекты — неразличимые, ящики — различимые, ящики не могут быть пустыми	$C(n-1, k-1)$
(6)	объекты — различимые, ящики — неразличимые, ящики могут быть пустыми	$S_1^{(n)} + S_2^{(2)} + \cdots + S_k^{(n)}$
(7)	объекты — различимые, ящики — неразличимые, ящики не могут быть пустыми	$S_k^{(n)}$

ТЕОРЕМА 12.4. Завершив рассмотрение размещений объектов в ящики, перейдем к нахождению количества способов разделения n объектов на k-циклы при условии, что никакие два из k-циклов не имеют общих элементов. Цикл из m объектов имеет вид $a_1a_2a_3\cdots a_m$, где a_m считается соседом a_1, так что

$$a_1a_2a_3\cdots a_j\cdots a_m = a_j\cdots a_ma_1a_2a_3\cdots a_{j-1}.$$

Следовательно,

$$a_1a_2a_3 = a_3a_1a_2 = a_2a_3a_1.$$

Можно представить себе n-цикл как n человек, сидящих за круглым столом. В главе 8 было установлено, что существуют $(m-1)!$ способов рассадить m человек за круглым столом. Следовательно, если дано m объектов, то существуют $(m-1)!$ способов построить m-цикл.

Теперь подсчитаем количество способов разделения n объектов на k-циклы при условии, что никакие два из k-циклов не имеют общих элементов.

ТЕОРЕМА 12.5. Количество возможных способов разделения n объектов на k-циклы так, что никакие два из k-циклов не имеют общих элементов, равно $s_k^{(n)}$, где $\{s_k^{(n)} : 0 \le k \le n\}$ — множество чисел Стирлинга первого рода.

ДОКАЗАТЕЛЬСТВО. Пусть $f(n,k)$ — количество возможных способов разделения n объектов на k-циклы, когда никакие два из k-циклов не имеют общих элементов. Чтобы доказать справедливость соотношения $f(n,k) = s_k^{(n)}$, необходимо сначала показать, что $f(n,0) = 0$ для всех $n \ge 1$. Но если $n \ge 1$, то невозможно

сформировать 0 циклов и использовать все n объектов. Следовательно, $f(n,0)=0$ для всех $n \geq 1$. Теперь покажем, что $f(n,n)=1$ для всех $n \geq 0$. Если $n \geq 1$, то можно построить n циклов длины 1, т.е. 1-циклов. Очевидно, сделать это можно единственным образом, поскольку порядок записи цикла несущественен. Если $n=0$, то представляется разумным считать, что существует единственный способ построить цикл, не используя объекты, — просто не строить его. Поэтому $f(n,n)=1$ для всех $n \geq 0$. Далее необходимо показать, что

$$f(n+1,k)=f(n,k-1)+nf(n,k).$$

Каждое размещение $n+1$ объектов сводится либо к размещению $n+1$-го объекта в новый цикл, либо к добавлению его в уже существующий цикл. Если $n+1$-ый объект размещен в новый цикл, то остальные n объектов размещены в $k-1$ циклах. Последнее можно осуществить $f(n,k-1)$ способами. Если $n+1$-ый объект размещен в уже существующий цикл, содержащий, например, m объектов, то $n+1$-ый объект может быть помещен вслед за любым из m объектов в цикле. Следовательно, можно построить m различных циклов, если $n+1$-ый объект размещать в цикле, содержащем m объектов. Допустим, что имеются циклы c_1, c_2, ..., c_k, где c_i является n_i циклом и $n_1+n_2+\cdots+n_k=n$. Тогда, если a_{n+1} размещен в цикл c_i, то можно построить n_i новых циклов. Следовательно, путем добавления $n+1$-го элемента в фиксированное множество k-циклов, содержащих n элементов, можно образовать $n_1+n_2+\cdots+n_k=n$ новых циклов. Поскольку существуют $f(n,k)$ способов размещения n элементов в k-циклах, имеется $nf(n,k)$ способов образовать k-цикл с $n+1$ объектами, если в цикл добавлен $n+1$-ый объект. Следовательно, существуют

$$f(n,k-1)+nf(n,k)$$

способов разбиения $n+1$ объектов по k-циклам и $f(n,k-1)=s_k^{(n)}$. ■

Используя соотношение $s_k^{(n+1)}=s_{k-1}^{(n)}+ns_k^{(n)}$, можно построить таблицу чисел Стирлинга первого рода аналогично тому, как была построена таблица чисел Стирлинга второго рода.

ТЕОРЕМА 12.6. Треугольник для чисел Стирлинга первого рода

n	$s_0^{(n)}$	$s_1^{(n)}$	$s_2^{(n)}$	$s_3^{(n)}$	$s_4^{(n)}$	$s_5^{(n)}$	$s_6^{(n)}$	$s_7^{(n)}$	$s_8^{(n)}$	$s_9^{(n)}$	$s_{10}^{(n)}$
0	1										
1	0	1									
2	0	1	1								
3	0	2	3	1							
4	0	6	11	6	1						
5	0	24	50	35	10	1					
6	0	120	274	225	85	15	1				
7	0	720	1764	1624	735	175	21	1			
8	0	5040	13068	13132	6769	1960	322	28	1		
9	0	40320	109584	118124	67284	22449	4536	546	36	1	
10	0	362880	1026576	1172700	723680	269325	63273	9450	870	45	1

■ УПРАЖНЕНИЯ

1. Сколькими способами можно разместить 7 объектов в 3 ящиках, если
   а) объекты различимы, ящики различимы и ящики могут быть пустыми?
   б) объекты различимы, ящики различимы и ящики не могут быть пустыми?
   в) объекты неразличимы, ящики различимы и ящики могут быть пустыми?
   г) объекты неразличимы, ящики различимы и ящики не могут быть пустыми?
   д) объекты различимы, ящики неразличимы и ящики могут быть пустыми?
   е) объекты различимы, ящики неразличимы и ящики не могут быть пустыми?
2. Сколькими способами можно разместить 10 объектов в 4 ящиках, если
   а) объекты различимы, ящики различимы и ящики могут быть пустыми?
   б) объекты различимы, ящики различимы и ящики не могут быть пустыми?
   в) объекты неразличимы, ящики различимы и ящики могут быть пустыми?
   г) объекты неразличимы, ящики различимы и ящики не могут быть пустыми?
   д) объекты различимы, ящики неразличимы и ящики могут быть пустыми?
   е) объекты различимы, ящики неразличимы и ящики не могут быть пустыми?
3. Заполните следующие два ряда треугольника для чисел Стирлинга второго рода.
4. Заполните следующие два ряда треугольника для чисел Стирлинга первого рода.
5. Сколькими способами можно разместить 12 объектов в 4 ящиках, если
   а) объекты различимы, ящики различимы и ящики могут быть пустыми?
   б) объекты различимы, ящики различимы и ящики не могут быть пустыми?
   в) объекты неразличимы, ящики различимы и ящики могут быть пустыми?
   г) объекты неразличимы, ящики различимы и ящики не могут быть пустыми?
   д) объекты различимы, ящики неразличимы и ящики могут быть пустыми?
   е) объекты различимы, ящики неразличимы и ящики не могут быть пустыми?
6. Сколькими способами можно разместить 14 объектов в 6 ящиках, если
   а) объекты различимы, ящики различимы и ящики могут быть пустыми?
   б) объекты различимы, ящики различимы и ящики не могут быть пустыми?
   в) объекты неразличимы, ящики различимы и ящики могут быть пустыми?
   г) объекты неразличимы, ящики различимы и ящики не могут быть пустыми?
   д) объекты различимы, ящики неразличимы и ящики могут быть пустыми?
   е) объекты различимы, ящики неразличимы и ящики не могут быть пустыми?

7. Пусть $S(n,k) = \frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$.

а) Покажите, что $S(n,0) = 0$ для всех $n \geq 1$.

б) Покажите, что $S(n,k)$ удовлетворяет рекурсивному отношению $S(n+1,k) = S(n,k-1) + kS(n,k)$.

12.2. ЧИСЛА КАТАЛАНА

В этом разделе рассмотрены три различные задачи, но в процессе их решения понятно, что это, по сути, одна и та же задача.

12.2.1 Задача 1

Человек идет на место работы, которое находится в десяти кварталах к северу и десяти кварталах к востоку от его дома. По диагонали, соединяющей его дом и место работы, течет река. К сожалению, на этом участке реки нет мостов, так что при достижении диагонали человек также достигает и реки. Следовательно, он может попасть на диагональ, но после этого должен повернуть направо. Если улицы проходят по квадратной сетке, так, что в каждом квартале есть улицы в направлении север-юг и улицы в направлении восток-запад, то сколькими способами человек может попасть на работу, не возвращаясь и не пересекая реку. Примеры его возможных маршрутов показаны на рис. 12.1.

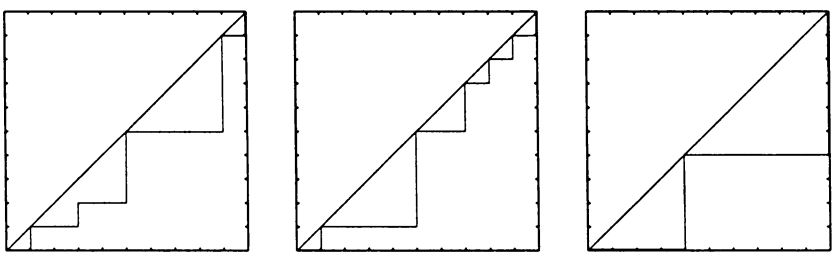


Рис. 12.1

Если обозначить продвижение человека на один квартал на восток через E, а его продвижение на один квартал на север через N, то первый из приведенных маршрутов можно обозначить как

$$ENEENEENNNEEEENNNNEN.$$

Второй маршрут можно обозначить как

$$ENEEEENNNNEENNENENEN$$

и третий — как

$$EEEENNNNEEEEEENNNNNN.$$

Заметим, что любой маршрут должен начинаться с E, и в любой точке маршрута количество продвижений E должно быть не меньше количества продвижений N, иначе человек попадет в реку.

12.2.2 Задача 2

Сколькими способами можно сложить последовательность из десяти чисел 1 и десяти чисел -1 так, что если S_k — сумма первых k элементов последовательности, то $S_k \geq 0$ для всех $1 \leq k \leq 20$. Возможными последовательностями являются

$$1, -1, 1, 1, -1, 1, 1, -1, -1, -1, 1, 1, 1, 1, -1, -1, -1, -1, 1, -1;$$

$$1, -1, 1, 1, 1, 1, -1, -1, -1, -1, 1, 1, -1, -1, 1, -1, 1, -1, 1, -1$$

и

$$1, 1, 1, 1, -1, -1, -1, -1, 1, 1, 1, 1, 1, 1, -1, -1, -1, -1, -1, -1.$$

Заметим опять, что в любой точке вдоль любой из последовательностей количество $+1$ должно быть не меньше количества -1.

12.2.3 Задача 3

Задано 11 чисел и бинарные операции между ними. Сколькими способами в этом выражении можно вставить 10 пар круглых скобок? Эта проблема выглядит похожей на предыдущую, поскольку при движении слева направо закрывающих (правых) скобок должно быть не больше, чем открывающих (левых) скобок. Однако, проблемой остается их размещение. Оказывается, что лучше использовать левые (открывающие) скобки и первые 10 чисел, чем левые и правые скобки, поскольку чисел не может быть больше, чем левых скобок. Например, рассмотрим

$$3 + 4 - 8 \times 2 \div 4 - 7 + 4 \times 5\hat{}6 - 2 \times 4.$$

Согласно последовательности

$$1, -1, 1, 1, -1, 1, 1, -1, -1, -1, 1, 1, 1, 1, -1, -1, -1, -1, 1, -1$$

разместим левые скобки там, где имеется 1, а символ — там, где имеется -1, после чего получим

$$(3 + ((4 - ((8 \times 2) \div 4)) - ((((7 + 4) \times 5)\hat{}6) - (2 \times 4)))).$$

Левые скобки размещены в паре с правыми скобками. Теперь рассмотрим

$$1, -1, 1, 1, 1, 1, -1, -1, -1, -1, 1, 1, -1, -1, 1, -1, 1, -1, 1, -1,$$

что соответствует

$$(3 + ((((4 - 8) \times 2) \div 4) - ((7 + 4) \times (5\hat{}(6 - (2 \times 4)))))),$$

и наконец

$$1, 1, 1, 1, -1, -1, -1, -1, 1, 1, 1, 1, 1, 1, -1, -1, -1, -1, -1, -1,$$

что соответствует

$$((((3 + 4) - 8) \times 2) \div ((((((4 - 7) + 4) \times 5)\hat{}6) - 2) \times 4)).$$

Формулировке задач мы уделили достаточно много времени. Теперь приступим к их решению. Для начала найдем решение при $n = 10$, чтобы окончательно разобраться с нашим примером, а затем выведем формулу для произвольного n. Вернемся к человеку, идущему на работу. Для начала рассмотрим все пути, по которым он может добраться до работы мокрым или сухим, не возвращаясь. Ему следует в некотором порядке пройти 10 кварталов на север и 10 на восток. Поскольку рассматриваются все пути, включая те, которые пересекают реку, то десять E и десять N могут быть в любом порядке. Если рассмотреть строку из десяти E и десяти N как строку из 20 пробелов, то необходимо выбрать 10 из них, в которых разместить E. Таким образом, общее количество путей к месту работы равно количеству способов выбрать 10 пробелов из 20, в которых будут размещены E. Это число равно C_{10}^{20}, или $\binom{20}{10} = \frac{20!}{10!10!}$. Теперь воспользуемся довольно распространенным в подсчетах приемом. Подсчитаем количество путей, при прохождении которых наш персонаж будет мокрым, и вычтем это число из общего количества путей. Иными словами, найдем количество путей, которые нам не подходят, и вычтем его из общего количества путей, чтобы получить количество путей, по которым наш персонаж может безопасно добраться до места работы. Наш герой оказывался в реке всякий раз, когда, достигнув какой-либо точки маршрута, он проходил на север больше кварталов, чем на восток. Иными словами, если в некоторой точке продвижений N у него на одно больше, чем продвижений E, то он попадает в реку. Например, в строке

$$EENN\boldsymbol{N}EENENENEENNNENE \tag{1}$$

знак $\boldsymbol{N}$ показывает точку, в которой человек попадает в реку. К моменту попадания в реку было сделано два продвижения E и три продвижения N, после чего осталось еще 8 E и 7 N. Заменим теперь все E, расположенные в строке после $\boldsymbol{N}$, на N, а все N — на E, после чего получаем строку

$$EENN\boldsymbol{N}NNENENENNEEENEN. \tag{2}$$

На новом пути (который не приведет человека к месту работы) имеется 2 E и 3 N, за которым следует 8 N и 7 E, что в сумме даст 9 E и 11 N. Это понятно, если учесть, что при попадании в реку в пройденном пути было на одно N больше, поэтому в оставшейся части пути на одно E было больше. После изменения остатка пути в нем оказалось на одно N больше. Следовательно, во всем пути продвижений N стало на два больше, чем продвижений E. Попробуем еще раз. В строке

$$ENENENEN\boldsymbol{N}ENENEENENNE \tag{3}$$

человек попадает в реку после 5 продвижений N и 4 продвижений E. После $\boldsymbol{N}$ осталось 5 N и 6 E продвижений. Опять заменяем каждое E, стоящее в строке после $\boldsymbol{N}$, на N, а каждое N — на E, после чего получаем 6 N и 5 E, стоящие в строке после $\boldsymbol{N}$,

$$ENENENEN\boldsymbol{N}NENENNENEEN. \tag{4}$$

Поэтому новая строка имеет в сумме 9 E и 11 N продвижений.

Обратно, если имеется строка, в которой 9 E и 11 N, то в некоторой точке строки продвижений N должно быть больше, чем продвижений E. Выбираем

первую такую точку и меняем первое такое N на $\boldsymbol{N}$. Далее снова изменяем оставшуюся часть строки, меняя каждое N на E и каждое E — на N. Например, пусть дана строка

$$ENENEENENENN\boldsymbol{N}ENENNNE, \tag{5}$$

которая содержит 9 E и 11 N. Выберем первое место в последовательности, в котором количество продвижений N больше, чем количество продвижений E, и заменим N в этой точке на $\boldsymbol{N}$. На данный момент в строке имеется 6 E и 7 N, включая $\boldsymbol{N}$. Остаток строки содержит 3 E и 4 N. Меняем все N, стоящие после $\boldsymbol{N}$, на E, а все E, стоящие после $\boldsymbol{N}$, на N. В результате получаем

$$ENENEENENENN\boldsymbol{N}NENEENN, \tag{6}$$

в которой после $\boldsymbol{N}$ стоит 4 E и 3 N. В сумме получаем 10 E и 10 N. Опять получили путь, который приводит нашего героя в реку, и это произойдет в точке $\boldsymbol{N}$. Это понятно, поскольку мы начали с момента 9 E и 11 N, и до момента $\boldsymbol{N}$, и включая его, у нас продвижений N было на одно больше, чем продвижений E. Следовательно, оставшаяся часть строки, стоящая после $\boldsymbol{N}$, также имеет на одно N больше, чем E. После изменения остатка строки в ней продвижений E на одно больше, чем продвижений N. Следовательно, во всей новой строке всегда будет равное количество E и N, а начальная часть строки, включая $\boldsymbol{N}$, всегда будет содержать на одно N больше, чем E. При этом $\boldsymbol{N}$ — то лишнее продвижение на север, которое и приводит нашего героя в реку. Легко заметить, что каждая строка, которая "окунает" наш персонаж в реку, может быть преобразована в строку с 9 E и 11 N, и обратно, каждая строка, содержащая 9 E и 11 N, может быть преобразована в строку, которая приводит нашего героя в реку. Отсюда количество строк, приводящих в реку, равно количеству способов сформировать строку, содержащую 9 E и 11 N. Таким образом, 9 мест для E выбираются из 20 мест, и для этого существуют $C(20,9) = \binom{20}{9} = \frac{20!}{9!11!}$ способов. Поэтому количество способов добраться до места работы сухим равно числу, полученному при вычитании из общего числа способов количества путей, при которых персонаж попадает в реку. Обозначим это число C_{10}, так что $C_{10} = \frac{20!}{10!10!} - \frac{20!}{9!11!}$. Это число можно было бы подсчитать, но настоящие математики никогда не используют числа, когда можно оперировать символами, чтобы избежать арифметических ошибок.

Предположим, что имеется n кварталов на север и n на восток, т.е. в рассмотренном выше примере $n = 10$. Теперь используем строку длины $2n$, половину которой составляют символы E, а вторую половину — символы N. Во избежание попадания человека в реку, нежелательно, чтобы число символов N превышало число символов E. Как и ранее, рассмотрим все возможные пути попасть к месту работы сухим или мокрым без возвращений. Выбираем из $2n$ возможных мест n мест для размещения E. Как было показано, существуют $C(2n,n) = \binom{2n}{n} = \frac{(2n)!}{n!n!}$ таких размещений. Далее опять находим пути, которые нам *не* подходят, поскольку приводят нашего героя в реку. Используя процедуру, описанную выше, находим, что количество таких путей равно количеству строк, в которых символ E появляется $n-1$ раз, символ N появляется $n+1$ раз. Число таких строк равно количеству способов выбрать $n-1$ место из $2n$ возможных для размещения символов E. Таких размещений будет $C(2n,n-1) = \binom{2n}{n-1} = \frac{(2n)!}{(n-1)!(n+1)!}$, так что C_n,

количество безопасных путей добраться до места работы, равно $\frac{(2n)!}{n!n!} - \frac{(2n)!}{(n-1)!(n+1)!}$. Теперь можно осуществить подсчет. Итак, получаем

$$C_n = \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n-1)!(n+1)!} = \frac{(2n)!}{n!n!} - \left(\frac{n}{n+1}\right)\left(\frac{(2n)!}{n!n!}\right) =$$
$$= \left(1 - \left(\frac{n}{n+1}\right)\right)\left(\frac{(2n)!}{n!n!}\right) = \left(\frac{1}{n+1}\right)\left(\frac{(2n)!}{n!n!}\right) = \left(\frac{1}{n+1}\right)\binom{2n}{n}.$$

Прежде чем оставить это путешествие, снова вернемся к случаю $n = 10$. Рассмотрим иной способ подсчета подходящих путей, т.е. C_n, а также случай, когда наш герой достигает реки, пройдя один километр на восток и один на север. Считаем, что дом нашего героя находится в начале координат $(0,0)$, так что он достигает точки $(1,1)$ и продолжает путь к месту работы. Этот путь изображен на рис. 12.2.

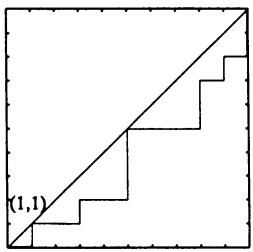


Рис. 12.2

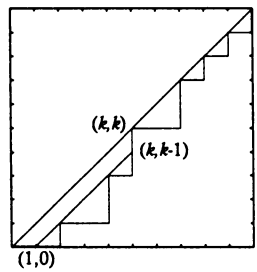


Рис. 12.3

Сколько существует таких путей? Очевидно, попасть в точку $(1,1)$ по суше можно только одним способом. Легко видеть, что задача о том, как попасть из точки $(1,1)$ в точку $(10,10)$, т.е. к месту работы, ничем не отличается от исходной, за исключением того, что теперь нужно, не угодив в реку, пройти 9 кварталов на восток и 9 на север. Это можно сделать C_9 способами. Поскольку $C_0 = 1$, мы говорим, что имеются $C_0 \cdot C_9$ способов попасть к месту работы, проходя через точку $(1,1)$. Это соответствует обозначениям, которые мы введем в дальнейшем.

Теперь предположим, что наш герой по пути на работу первый раз достигает берега реки в точке (k,k), где $1 < k < 10$. Для начала подсчитаем, сколько путей, не достигающих берега реки, вплоть до точки (k,k), ведут из точки $(0,0)$ в точку (k,k). Любой из путей должен начинаться с движения на 2 квартала на восток, иначе персонаж достигнет реки в точке $(1,1)$. Закончить путь персонаж должен движением на 2 квартала на север, иначе ему следовало бы достигнуть реки в точке $(k-1, k-1)$. Как видно из рис. 12.3, если провести линию, параллельную реке и проходящую на один километр южнее между точками $(1,0)$ и $(k, k-1)$, то путь не пересечет эту линию. Следовательно, можно использовать любой путь, который начинается в точке $(1,0)$, проходит $k-1$ блоков на север и $k-1$ блоков на восток и достигает точки $(k, k-1)$, *не пересекая новую диагональ*. Существуют C_{k-1} способов провести такой путь. Следовательно, имеются C_{k-1} путей из точки $(0,0)$ в точку (k,k), которые до достижения точки (k,k) не приводят к реке.

Теперь рассмотрим пути, которые по суше ведут из точки (k,k) в точку места работы. Но эта задача эквивалентна задаче о перемещении посуху на $10-k$ кварталов на север и $10-k$ кварталов на восток, а значит, число путей равно C_{10-k}. Поскольку существуют C_{k-1} способов попасть в точку (k,k) и C_{10-k} способов попасть из точки (k,k) к месту работы, то существуют $C_{k-1}\cdot C_{10-k}$ способов попасть на работу, достигнув реки первый раз в точке (k,k). Более того, для каждого пути существует только одно значение k, поскольку место, где он первый раз достигает реки, только одно. Рассмотрим случай, когда наш герой попадает к месту работы, не достигая реки. Иными словами, рассмотрим случай $k=10$. Используем вышеприведенные аргументы, чтобы показать, что количество путей на работу в этом случае равно C_{k-1}, т.е. C_9, но в этом случае, в отличие от предыдущих, нет второй части путешествия. Для удобства записи, исходя из того, что $C_0=1$, запишем, что число способов дойти до работы, не достигая реки, равно $C_9\cdot C_0$. Если просуммировать количество путей для всех возможных значений k, $1\le k\le 10$, то получим количество всех возможных путей от дома к месту работы. Следовательно,

$$C_{10}=\sum_{k=1}^{10} C_{k-1}\cdot C_{10-k}.$$

Как легко видеть, рассуждения не связаны с использованием числа 10. Если взять n вместо 10, то получим

$$C_n=\sum_{k=1}^{n} C_{k-1}\cdot C_{n-k}.$$

Пусть задан $n+2$-сторонний выпуклый многоугольник; требуется разделить многоугольник на треугольники, проводя прямые из вершины в вершину, причем прямые не должны пересекаться. Сколько существуют способов такого разбиения? Обозначим это число через T_n. Например, для $n=4$ возможная триангуляция (разбиение на треугольники) показана на рис. 12.4. Для удобства обозначим вершины через $v_0, v_a, v_1, v_2, v_3, \ldots, v_n$. Предположим, что проведена линия из вершины v_0 в вершину v_1. Сколько треугольников можно начертить? (См. рис. 12.5.)

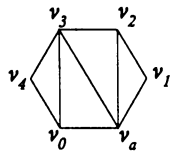


Рис. 12.4

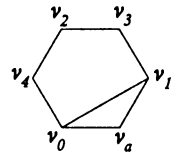


Рис. 12.5

Теперь для триангуляции остался многоугольник $v_0v_1v_2v_3\ldots v_n$. Поскольку у него $n+1$ сторон, то имеется T_{n-1} способов триангуляции этого многоугольника. Предположим, что мы начали триангулировать исходный $n+2$-сторонний многоугольник, нарисовав прямую из вершины v_0 в вершину v_k, где $1<k<n$,

и запретили рисовать прямые из вершины v_0 в любую другую вершину v_i, где $i < k$, так что k — наименьшее целое число, для которого существует отрезок из вершины v_0 в вершину с таким нижним индексом. Но тогда должен существовать отрезок из вершины v_k в вершину v_a, образующий треугольник $v_k v_a v_0$. В результате для триангуляции остается многоугольник $v_a v_1 v_2 v_3 \ldots v_k$. Поскольку в этом многоугольнике $k+1$ сторон, то существуют T_{k-1} способов разбиения его на треугольники. (См. рис. 12.6.)

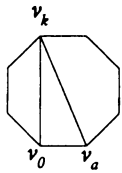


Рис. 12.6

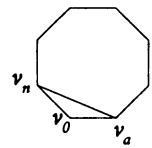


Рис. 12.7

Для триангуляции остается также многоугольник $v_0 v_k v_{k+1} v_{k+2} v_{k+3} \ldots v_n$. В нем $n-k+2$ сторон, поэтому его можно разбить на треугольники T_{n-k} способами. Следовательно, многоугольник справа от линии, соединяющей вершины v_0 и v_k, можно триангулировать T_{k-1} способами, а многоугольник слева от этой линии можно триангулировать T_{n-k} способами. Поэтому весь многоугольник вместе с линией от вершины v_0 к вершине v_k можно триангулировать $T_{k-1} \cdot T_{n-k}$ способами. Рассмотрим случай, когда первой линией будет линия из вершины v_0 в вершину v_n, что можно интерпретировать как отсутствие прямых, выходящих из v_0 к любым другим вершинам. Но в таком случае должна быть прямая, выходящая из вершины v_n в вершину v_a, и многоугольник $v_a v_1 v_2 v_3 \ldots v_n$ триангулирован. (См. рис. 12.7.)

Но этот многоугольник имеет $n+1$ сторон и поэтому может быть триангулирован T_{k-1} способами. Принимая во внимание, что суммирование по k от 1 до n учитывет все возможные разбиения на треугольники исходного $n+2$-стороннего многоугольника и, полагая $T_0 = 1$, получаем

$$T_n = \sum_{k=1}^{n} T_{k-1} \cdot T_{n-k},$$

что совпадает с определением чисел Каталана, поэтому $T_n = C_n$.

■ УПРАЖНЕНИЯ

1. Вычислите C_5, C_6 и C_7.
2. Во время перерыва 20 человек выстроились у автомата по продаже напитков. Любой напиток стоит 50 центов. В автомате нет монет для сдачи, но как только автомат получает монету, он может возвращать ее как сдачу. Если 10 человек могут заплатить без сдачи, а 10 человек имеют только купюры в один доллар, то сколькими способами они могут выстроиться в очередь, чтобы каждый, кто нуждается в сдаче, мог ее получить.

3. Сколькими способами можно расставить в очереди 15 мужчин и 15 женщин, если в любом месте очереди женщин впереди должно быть не меньше, чем мужчин?
4. Пусть имеется последовательность из 12 чисел. Сколькими способами можно перемножить эти числа, не меняя их порядка записи.

12.3. ОБЩЕЕ ВКЛЮЧЕНИЕ-ИСКЛЮЧЕНИЕ И РАЗУПОРЯДОЧЕНИЯ

Ранее мы уже обсуждали принцип включения-исключения для трех или менее множеств. В этом разделе мы рассмотрим принцип включения-исключения для конечного числа конечных множеств. Для начала докажем следующую теорему.

ТЕОРЕМА 12.7. Пусть $A_1, A_2, A_3, \ldots, A_n$ — совокупность конечных множеств. Количество элементов во множестве $A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n$ определяется формулой

$$|A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \\ + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \cdots\cdots + (-1)^{n+1} |A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n|.$$

ДОКАЗАТЕЛЬСТВО. Необходимо показать, что в правой части приведенного выше равенства каждый элемент множества $A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n$ подсчитан и учтен ровно один раз. Предположим, что элемент a принадлежит в точности p множествам совокупности $A_1, A_2, A_3, \ldots, A_n$. В сумме $\sum_i |A_i|$ элемент a учтен p раз. В сумме $\sum_{i<j} |A_i \cap A_j|$ элемент a учтен всякий раз, когда выбраны два множества, содержащие элемент a. Существуют $\binom{p}{2}$ способов выбрать такие два множества. Следовательно, в сумме $\sum_{i<j} |A_i \cap A_j|$ элемент a учтен $\binom{p}{2}$ раз. В сумме $\sum_{i<j<k} |A_i \cap A_j \cap A_k|$ элемент a учтен всякий раз, когда выбраны три множества, содержащие элемент a. Существуют $\binom{p}{3}$ способов выбрать такие три множества. Следовательно, в сумме $\sum_{i<j<k} |A_i \cap A_j \cap A_k|$ элемент a учтен $\binom{p}{3}$ раз. В сумме элементов всех возможных пересечений r множеств, где $r \leq p$, элемент a учтен $\binom{p}{r}$ раз. Следовательно, в сумме

$$\sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \\ + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n|$$

элемент a учтен $p - \binom{p}{2} + \binom{p}{3} - \cdots + (-1)^{i+1}\binom{p}{i} + \cdots + (-1)^{p+1}$ раз. Но

$$0 = (1-1)^p = 1 - p + \binom{p}{2} - \binom{p}{3} + \cdots + (-1)^p,$$

так что

$$1 = p - \binom{p}{2} + \binom{p}{3} - \cdots + (-1)^{i+1}\binom{p}{i} + \cdots + (-1)^{p+1},$$

и получаем, что элемент a учтен в точности один раз. ■

На основании этой теоремы докажем утверждение, известное как теорема о включении-исключении.

ТЕОРЕМА 12.8. (о включении-исключении) Пусть $A_1, A_2, A_3, \ldots, A_n$ — набор конечных множеств. Количество элементов множества $A_1' \cap A_2' \cap A_3' \cap \ldots \cap A_n'$ определяется формулой

$$|A_1' \cap A_2' \cap \ldots \cap A_n'| = |\mathbf{U}| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \\ - \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \cdots + (-1)^n |A_1 \cap A_2 \cap \ldots \cap A_n|.$$

ДОКАЗАТЕЛЬСТВО. Из теории множеств известно, что

$$A_1' \cap A_2' \cap A_3' \cap \ldots \cap A_n' = (A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n)' = \\ = \mathbf{U} - (A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n),$$

так что имеем соотношение

$$|A_1' \cap A_2' \cap A_3' \cap \ldots \cap A_n'| = |\mathbf{U}| - |A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n|,$$

подстановка которого в

$$|A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \\ + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n|$$

дает требуемый результат. ■

ПРИМЕР 12.9. Сколько положительных чисел, меньших 2003, делятся на 2, 3, 5 или 7? Сколько положительных чисел, меньших 2001, не делятся на 2, 3, 5 или 7? Имеются $\lfloor \frac{2002}{2} \rfloor = 1001$ целых чисел, делящихся на 2; $\lfloor \frac{2002}{3} \rfloor = 666$ целых чисел, делящихся на 3, $\lfloor \frac{2002}{5} \rfloor = 400$ целых чисел, делящихся на 5, и $\lfloor \frac{2002}{7} \rfloor = 286$ целых чисел, делящихся на 7. Выбирая по два, получаем, что $\lfloor \frac{2002}{6} \rfloor = 333$ целых числа делятся на 2 и 3, $\lfloor \frac{2002}{10} \rfloor = 200$ целых чисел делятся на 2 и 5, $\lfloor \frac{2002}{15} \rfloor = 133$ целых числа делятся на 3 и 5, $\lfloor \frac{2002}{14} \rfloor = 143$ целых числа делятся на 2 и 7, $\lfloor \frac{2002}{21} \rfloor = 95$ целых чисел делятся на 3 и 7, и $\lfloor \frac{2002}{35} \rfloor = 57$ целых чисел делятся на 5 и 7. Выбирая по три, получаем, что $\lfloor \frac{2002}{30} \rfloor = 66$ целых числа делятся на 2, 3 и 5, $\lfloor \frac{2002}{42} \rfloor = 47$ целых чисел делятся на 2, 3 и 7, $\lfloor \frac{2002}{70} \rfloor = 28$ целых чисел делятся на 2, 5 и 7, и $\lfloor \frac{2002}{105} \rfloor = 19$ целых чисел делятся на 3, 5 и 7. Наконец, имеются $\lfloor \frac{2002}{210} \rfloor = 9$ целых чисел, которые делятся на 2, 3, 5 и 7. Итак, существуют

$$1001+666+400+286-(333+200+133+143+95+57)++66+47+28+19-9$$

или $2353-961+160-9=1543$ целых числа, которые делятся на 2, 3, 5 или 7. Соответственно, существуют

$$2001-1543=458$$

целых чисел, которые не делятся на 2, 3, 5 или 7. □

Воспользуемся теперь теоремой о включении-исключении для подсчета количества разупорядочений.

ОПРЕДЕЛЕНИЕ 12.10. ***Разупорядочением*** называется перестановка n различных упорядоченных символов, при которой ни один из символов не остается на своем месте. Количество разупорядочений на n различных упорядоченных символах обозначается через D_n.

Для удобства обозначим рассматриваемые n символов как $1,2,3,\ldots,n$. Для $n=1$ величина $D_1=0$, поскольку перестановка одного элемента оставляет его на месте. Для $n=2$ единственным разупорядочением будет

$$\begin{pmatrix}1&2\\2&1\end{pmatrix},$$

поэтому $D_2=1$. Для $n=3$ разупорядочениями будут перестановки

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} \quad \text{и} \quad \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},$$

так что $D_3=2$. Существует 9 разупорядочений, если $n=4$. Предлагаем читателю найти эти перестановки самостоятельно.

Попробуем найти количество разупорядочений для $n=4$, используя технику подсчетов. Пусть $\boldsymbol{U}$ — множество всех перестановок на множестве $\{1,2,3,4\}$. Таким образом, $|\boldsymbol{U}|=4!$. Пусть A_1 — множество перестановок, оставляющих на месте символ 1, A_2 — множество перестановок, оставляющих на месте символ 2, A_3 — множество перестановок, оставляющих на месте символ 3 и A_4 — множество перестановок, оставляющих на месте символ 4. В таком случае, множество $A_1'\cap A_2'\cap A_3'\cap A_4'$ — множество перестановок, не оставляющих на месте ни один из символов. По теореме о включении-исключении

$$|A_1'\cap A_2'\cap A_3'\cap A_4'|=|\boldsymbol{U}|-\sum_{i\le 4}|A_i|+\sum_{i<j\le 4}|A_i\cap A_j|-$$
$$-\sum_{i<j<k\le 4}|A_i\cap A_j\cap A_k|+|A_1\cap A_2\cap A_3\cap A_4|.$$

Пусть $A_{ij}=A_i\cap A_j$, $A_{ijk}=A_i\cap A_j\cap A_k$ и $A_{1234}=A_1\cap A_2\cap A_3\cap A_4$. Таким образом, A_{ij} — множество всех перестановок, оставляющих на месте символы i и j, A_{ijk} — множество перестановок, оставляющих на месте символы i, j и k, и

множество A_{1234} содержит только тождественную перестановку, которая оставляет на месте все четыре символа. Для нахождения суммы $\sum_{i\leq 4}|A_i|$ требуется найти количество перестановок, оставляющих на месте символ i.

Рассмотрим, в частности, A_1. Любая перестановка из A_1 оставляет символ 1 на месте и переставляет символы 2, 3 и 4. Следовательно, множество A_1 содержит 3! перестановки. По аналогичным соображениям, множество A_i содержит 3! элемента. Следовательно,

$$|A_1|+|A_2|+|A_3|+|A_4| = 4\times 3! = 4! = 4!\times\frac{1}{1!}.$$

Найдем теперь

$$\sum_{i<j\leq 4}|A_i\cap A_j| = \sum_{i<j\leq 4}|A_{ij}|.$$

Рассмотрим частный случай $|A_{12}|$. Чтобы принадлежать множеству A_{12}, перестановка должна оставлять на месте символы 1 и 2 и переставлять символы 3 и 4. Для этого существует 2! способа. Таким образом, $|A_{12}| = 2$. Аналогично, $|A_{ij}| = 2$ при всех $1\leq i<j\leq 4$. Теперь возникает вопрос, сколько имеется таких множеств A_{ij}? Поскольку символы i и j выбираются из 4 элементов, то существуют $\binom{4}{2}$ множеств вида A_{ij}. Следовательно,

$$\sum_{i<j\leq 4}|A_{ij}| = \binom{4}{2}2! = \frac{4!}{2!2!}2! = 4!\frac{1}{2!}.$$

Далее найдем

$$\sum_{i<j<k\leq 4}|A_i\cap A_j\cap A_k| = \sum_{i<j<k\leq 4}|A_{ijk}|.$$

Сначала рассмотрим A_{123}. Поскольку символы 1, 2 и 3 не меняют своих мест, то остается только 4. Отсюда, множество A_{123} содержит только 1 элемент. Аналогично, множество A_{ijk} при $1\leq i<j<k\leq 4$ содержит только 1 элемент. Поскольку символы i, j и k выбираются из 4 элементов, то существуют $\binom{4}{3}$ множеств вида A_{ijk}. Следовательно,

$$\sum_{i<j<k\leq 4}|A_{ijk}| = \frac{4!}{3!1!}\times 1 = 4!\frac{1}{3!}.$$

Наконец, множество A_{1234}, как уже отмечалось, состоит из 1 элемента, но

$$1 = \frac{4!}{4!} = 4!\times\frac{1}{4!}.$$

Сводя вместе полученные результаты, получаем

$$|A_1'\cap A_2'\cap A_3'\cap A_4'| = 4! - 4!\times\frac{1}{1!} + 4!\frac{1}{2!} - 4!\frac{1}{3!} + 4!\times\frac{1}{4!} =$$
$$= 4!(1-\frac{1}{1!}+\frac{1}{2!}-\frac{1}{3!}+\frac{1}{4!}).$$

В более общем виде можно сформулировать следующую теорему.

ТЕОРЕМА 12.11. Для $n > 1$ величина D_n, количество разупорядочений на n символах $1, 2, 3, \ldots, n$, вычисляется по формуле

$$D_n = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right).$$

ДОКАЗАТЕЛЬСТВО. Для доказательства этого утверждения рассмотрим $A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_k$ при $1 \le k \le n$. Любой элемент множества $A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_k$ оставляет на месте символы $1, 2, 3, \ldots k$, поэтому он переставляет $n - k$ символов. Следовательно, $|A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_k| = (n-k)!$. Аналогично, для любых фиксированных $m_1, m_2, m_3, \ldots m_k$ имеем

$$|A_{m_1} \cap A_{m_2} \cap A_{m_3} \cap \ldots \cap A_{m_k}| = (n-k)!.$$

Поскольку k символов выбираются из n символов, то существуют $\binom{n}{k}$ множеств вида $A_{m_1} \cap A_{m_2} \cap A_{m_3} \cap \ldots \cap A_{m_k}$. Следовательно,

$$\sum |A_{m_1} \cap A_{m_2} \cap A_{m_3} \cap \ldots \cap A_{m_k}| = \binom{n}{k} \times (n-k)! =$$

$$= \frac{n!}{k!(n-k)!} \times (n-k)! = n! \times \frac{1}{k!}.$$

Снова предположив, что $\boldsymbol{U}$ — множество всех перестановок n объектов, так что $|\boldsymbol{U}| = n!$, получаем

$$D_n = |A_1' \cap A_2' \cap A_3' \cap \ldots \cap A_n'| = |\boldsymbol{U}| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| -$$

$$- \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \cdots + (-1)^n |A_1 \cap A_2 \cap \ldots \cap A_n| =$$

$$= n! - \frac{1}{1!}n! + \frac{1}{2!}n! - \frac{1}{3!}n! + \cdots + (-1)^n \frac{1}{n!}n! =$$

$$= n!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}).$$

■

В математическом анализе показано, что это разложение в ряд Тейлора функции e^{-1} имеет вид

$$e^{-1} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \cdots.$$

Следовательно, $D_n \approx n!e^{-1}$.

ПРИМЕР 12.12. Неопытный официант подает 10 различных блюд, обслуживая 10 клиентов. Сколькими способами он может сделать так, чтобы ни один клиент не получил свой заказ? В данном случае имеет место хаотическая перестановка 10 объектов, поэтому количество таких перестановок равно

$$10!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} + \frac{1}{10!}).$$

□

ПРИМЕР 12.13. Матрицей перестановок размерности 8×8 является матрица, в каждой строке и в каждом столбце которой только один элемент равен 1, а все остальные равны 0. Сколько существует матриц перестановок размерности 8×8, содержащих на главной диагонали только нули? Такие матрицы являются матричными представлениями перестановок 8 символов, которые и являются разупорядочениями. Следовательно, имеется

$$8!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!})$$

таких матриц. □

ПРИМЕР 12.14. В шахматах две ладьи (туры) могут атаковать друг друга, если они расположены на одной вертикали (столбце) или на одной горизонтали (строке). Следовательно, неатакующие ладьи — это те, которые расположены в разных столбцах или в разных строках. Очевидно, что восемь неатакующих ладей можно расположить на стандартной (8×8)-клеточной шахматной доске. Сколькими способами это можно сделать? Поскольку размещение на шахматной доске неатакующих ладей эквивалентно размещению единиц в перестановочной матрице, и существуют 8! перестановок на 8 объектах, то существуют 8! различных перестановочных матриц размерности 8×8.

Если не привязываться к перестановочным матрицам, то расположение ладьи в i-ой строке и в j-ом столбце можно идентифицировать с упорядоченной парой чисел (i, j). Если это сделано для каждой ладьи, то в результате имеем 8 упорядоченных пар,что дает перестановку, в которой первый элемент каждой упорядоченной пары отображен на второй элемент. В данном случае имеем стандартное представление функции как множества упорядоченных пар. □

ПРИМЕР 12.15. Сколькими различными способами можно расположить на стандартной (8×8)-клеточной шахматной доске 8 неатакующих ладей, чтобы у каждой номер строки и номер столбца не совпадали? Итак, ладья, расположенная в i-ой строке, не может находиться в i-ом столбце. Количество таких размещений равно количеству перестановочных матриц, у которых отсутствуют единицы на главной диагонали (там только нули). В терминах перестановок это означает, что ни один элемент не отображается сам на себя. Следовательно, существуют

$$8!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!})$$

различных способов разместить ладьи. □

■ УПРАЖНЕНИЯ

1. Докажите, что для любого положительного целого числа n

$$\binom{n}{1} - \binom{n}{2} + \binom{n}{3} - \binom{n}{4} + \cdots + (-1)^n \binom{n}{1} = 0.$$

2. У преподавателя обучаются 25 студентов. Преподаватель проводит тест, а затем просит студентов обменяться бумагами так, чтобы никто не проверял свой ответ. Сколькими способами это можно сделать?

3. Студент сдает экзамен, в процессе которого на каждый из 15 вопросов он должен выбрать один из 15 ответов. Если никакие два вопроса не имеют совпадающих ответов, то сколькими способами студент может дать неправильный ответ?

4. Пусть $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ и $B = \{-1, -2, -3, -4, -5, -6, -7, -8-, 9\}$. Каждому элементу из первого множества нужно поставить в соответствие единственный элемент из второго множества. Элементы каждой пары суммируются. Сколькими способами можно образовать пары, так чтобы ни одна из сумм не равнялась нулю?

5. Неграмотный почтальон должен доставить по одному письму каждому из 7 своих клиентов.

а) Сколькими способами он может доставить письма, так чтобы ни один из клиентов не получил свою почту?

б) Сколькими способами он может доставить письма, так чтобы хоть один клиент получил свою почту?

в) Сколькими способами он может доставить письма, так чтобы только один клиент получил свою почту?

г) Сколькими способами он может разнести письма, так чтобы только один клиент не получил свою почту?

6. Пусть множество A содержит 12 элементов, и φ — перестановка элементов множества A. Пусть $S = \{\theta : \theta$ — перестановка на множестве A и $\theta(a) \neq \varphi(a)$ дя всех $a \in A\}$. Найдите $|S|$.

7. Семь джентльменов отправляются на вечеринку и сдают там свои шляпы. Сколькими способами непутевый гардеробщик может вернуть шляпы, так чтобы

а) ни один из джентльменов не получил свою шляпу?

б) только один из джентльменов получил свою шляпу?

в) хоть один из джентльменов получил свою шляпу?

г) хотя бы двое джентльменов получили свои шляпы?

8. Сколько существует перестановок элементов множества $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ таких, что

а) в каждой перестановке ни один элемент не остается на своем месте, т.е. перестановка не отображает ни одно из чисел само в себя?

б) перестановка не оставляет ни одно четное число на своем месте?

в) перестановка оставляет на своем месте хотя бы одно четное число?

г) в результате перестановки в точности четыре числа остаются на своем месте?

9. Используя предыдущую задачу в качестве схемы, определите количество перестановок на n элементах, в результате которых в точности k элементов остаются на своих местах.

10. Найдите количество положительных целых чисел, меньших или равных числу 2300, взаимно простых с числом 700.

11. Найдите количество положительных целых чисел, меньших или равных числу 5460, взаимно простых с числом 700.

12.4. ЛАДЕЙНЫЕ ПОЛИНОМЫ И ЗАПРЕЩЕННЫЕ ПОЗИЦИИ

В приведенных выше примерах на (8×8)-клеточной шахматной доске было изучено расположение неатакующих ладей, а также расположение ладей на шахматной доске, когда они не могут находиться на пересечении строки и столбца, номера которых совпадают. Клетки шахматной доски, на которых ладья не может быть расположена, можно представить себе как **запрещенные позиции**. Рассмотрим теперь шахматную доску или часть шахматной доски с запрещенными позициями. При использовании такой $(n \times n)$-клеточной шахматной доски ситуация также описывается как ***перестановки с запрещенными позициями*** по аналогии с перестановками неатакующих ладей. До того, как у читателей, не играющих в шахматы, возникнет желание пропустить этот раздел или отложить книгу в сторону, заметим, что шахматы и шахматная доска используются только потому, что они хорошо иллюстрируют то, что мы собираемся делать.

Начнем с введения ладейного полинома.

ОПРЕДЕЛЕНИЕ 12.16. Пусть C — произвольная доска с m квадратами, которая для некоторого целого числа n является частью $n \times n$-клеточной шахматной доски. Для $0 \leq k \leq m$ целое число $r_k(C)$ равно количеству способов, которыми k ладей могут быть размещены на доске C в неатакующих позициях. ***Ладейным полиномом*** $R(x, C)$ на C называется производящая функция для чисел $r_k(C)$, так что

$$R(x, C) = r_0(C) + r_1(C)x + r_2(C)x^2 + r_3(C)x^3 + \cdots + r_m(C)x^m.$$

ПРИМЕР 12.17. Пусть C — доска, изображенная на рис. 12.8. Всегда имеется способ разместить 0 ладей на доске, поэтому $r_0(C) = 1$ для любой доски C. В этом случае имеются три способа разместить на C одну ладью, поскольку ее можно расположить в любой клетке доски C. Расположить две ладьи в неатакующей позиции можно только одним способом. Следовательно,

$$R(x, C) = 1 + 3x + x^2. \qquad \square$$

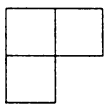


Рис. 12.8

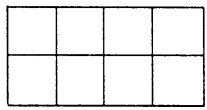


Рис. 12.9

ПРИМЕР 12.18. Пусть C — доска, изображенная на рис. 12.9. Доска состоит из 8 блоков, поэтому имеются 8 способов расположить одну ладью на доске C. Для размещения двух ладей в неатакующей позиции имеются 4 способа выбора на верхней горизонтали и три места для размещения ладьи на второй горизонтали, поскольку ладьи не могут располагаться друг под другом. Значит, существует всего 12 возможных способов расположения двух неатакующих ладей. Следовательно,

$$R(x, C) = 1 + 8x + 12x^2. \qquad \square$$

ПРИМЕР 12.19. Пусть C — доска, изображенная на рис. 12.10. Чтобы упростить задачу, заметим, что доску C можно представить состоящей из двух частей, C_1 и C_2, как показано на рисунке. Поскольку обе части не имеют ни общих строк, ни общих столбцов, то размещение ладей в одной части не влияет на размещение ладей в другой части.

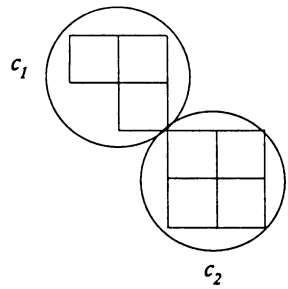


Рис. 12.10

Для размещения одной ладьи на C нужно разместить ее либо на C_1, либо на C_2. Имеются 3 способа размещения одной ладьи на C_1 и 4 — на C_2, поэтому есть 7 способов размещения ладьи на доске C. Для размещения двух неатакующих ладей на доске C можно размещать их в части C_1, одну ладью — в C_1 и одну ладью — в C_2 или обе — в C_2. Имеется 1 способ разместить две ладьи в части C_1 и 3 способа размещения ладьи в части C_1 и 4 — в части C_2, что дает 12 способов размещения 1 ладьи в части C_1 и одной ладьи в части C_2. Есть 2 способа размещения двух неатакующих ладей в части C_2. Таким образом, существуют $1 + 12 + 2 = 15$ способов размещения двух неатакуюших ладей на доске C. Чтобы разместить 3 ладьи на доске C, можно две ладьи разместить в части C_1 и одну ладью — в части C_2, или наоборот. Имеется 1 способ размещения двух неатакующих ладей в части C_1 и 4 способа размещения одной ладьи в части C_2, что дает 4 способа размещения двух ладей в части C_1 и одной ладьи — в части C_2. Есть 3 способа размещения одной ладьи в части C_1 и 2 способа размещения двух неатакущих ладей в части C_2, поэтому существуют 6 способов размещения двух ладей в части C_2 и одной ладьи — в части C_1. Это даст всего $4 + 6 = 10$ способов размещения 3 ладей на доске C. И наконец, чтобы разместить четыре неатакующие ладьи на доске C, можно две неатакующие ладьи разместить в части C_1 и две — в части C_2. Имеется 1 способ размещения двух неатакующих ладей в части C_1 и 2 способа размещения двух неатакующих ладей в части C_2. Поэтому есть 2 способа размещения четырех неатакующих ладей на доске C. Таким образом, ладейный полином для доски C имеет вид

$$R(x, C) = 1 + 7x + 15x^2 + 10x^3 + 2x^4. \qquad \square$$

ПРИМЕР 12.20. Пусть C — доска, изображенная на рис. 12.11. Как и в предыдущем примере, доску C можно представить состоящей из двух частей, C_1 и C_2, как показано на рисунке. Поскольку обе части не имеют ни общих строк, ни общих столбцов, то размещение ладей в одной части не влияет на размещение ладей в другой части.

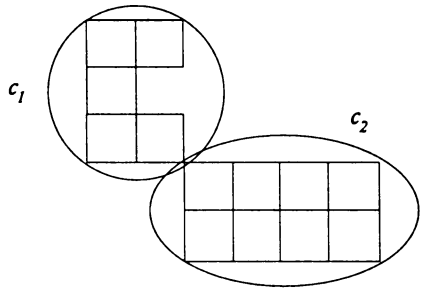


Рис. 12.11

Имеются 5 способов размещения ладьи в части C_1 и 8 — в части C_2. Поэтому существуют 13 способов размещения ладьи на доске C. Отметим, что $r_1(C) = r_1(C_1) + r_1(C_2)$. Учитывая, однако, что $r_0(C_1) = r_0(C_2) = 1$, можно положить

$$r_1(C) = r_1(C_1)r_0(C_2) + r_0(C_1)r_1(C_2).$$

При выборе двух неатакующих ладей можно выбрать две ладьи из части C_1, одну ладью из части C_1 и одну — из части C_2 или выбрать две ладьи из части C_2. Имеются 4 способа выбора двух неатакующих ладей из части C_1. Есть 5 способов выбора одной ладьи из части C_1 и 8 способов выбора одной ладьи из части C_2. Следовательно, имеются $5 \times 8 = 40$ способов выбора 1 ладьи из части C_1 и 1 — из части C_2. Имеются 12 способов выбора двух неатакующих ладей из части C_2. Следовательно, существуют $5 + 40 + 12 = 56$ способов расположения двух неатакующих ладей на доске C. Согласно методу выбора имеем, что

$$r_2(C) = r_2(C_1) + r_1(C_1)r_1(C_2) + r_2(C_2),$$

и это можно переписать в виде

$$r_2(C) = r_2(C_1)r_0(C_2) + r_1(C_1)r_1(C_2) + r_0(C_1)r_2(C_2).$$

Далее, чтобы получить 3 неатакующие ладьи на доске C, можно попробовать получить 3 неатакующие ладьи из части C_1, но такой возможности нет. Можно получить одну ладью из части C_1 и две неатакующие ладьи из части C_2. Имеются 5 способов получить одну ладью из части C_1 и 12 способов получить две неатакующие ладьи из части C_2. Это дает $5 \times 12 = 60$ способов получить три неатакующие ладьи, располагая одну ладью в части C_1 и две неатакующие ладьи в части C_2. Также можно получить две неатакующие ладьи из части C_1 и одну ладью из части C_2. Имеются 4 способа получить две неатакующие ладьи из части C_1 и 8 способов получить одну ладью из части C_2. Следовательно, существуют $4 \times 8 = 32$ способа получить три неатакующие ладьи, располагая две неатакующие ладьи в части C_1 и одну ладью в части C_2. Снова можно попробовать расположить три неатакующие ладьи в части C_2, но такого способа нет. Следовательно, существуют $r_3(C) = 60 + 32 = 92$ способа расположить три неатакующие ладьи на доске C. В соответствии с методом выбора можно также записать

$$r_3(C) = r_2(C_1)r_1(C_2) + r_1(C_1)r_2(C_2).$$

Наконец, чтобы расположить четыре неатакующие ладьи на доске C, нужно две неатакующие ладьи получить из части C_1 и две неатакующие ладьи — из части C_2. Имеются 4 способа получить две неатакующие ладьи из части C_1 и 12 способов получить две неатакующие ладьи из части C_2. Таким образом, есть $4 \times 12 = 48$ способов получить четыре неатакующие ладьи на доске C. Можно записать, что

$$r_4(C) = r_2(C_1)r_2(C_2).$$

Таким образом,

$$R(x, C) = 1 + 13x + 56x^2 + 92x^3 + 48x^4.$$

А также

$$\begin{aligned} R(x, C) = {} & r_0(C_1)r_0(C_2) + (r_1(C_1)r_0(C_2) + r_0(C_1)r_1(C_2))x + \\ & + (r_2(C_1)r_0(C_2) + r_1(C_1)r_1(C_2) + r_0(C_1)r_2(C_2))x^2 + \\ & + (r_2(C_1)r_1(C_2) + r_1(C_1)r_2(C_2))x^3 + r_2(C_1)r_2(C_2)x^4 = \\ = {} & (r_0(C_1) + (r_1(C_1)x + (r_2(C_1)x^2)((r_0(C_2) + (r_1(C_2)x + (r_2(C_2)x^2) = \\ = {} & R(x, C_1)R(x, C_2). \end{aligned}$$

Следовательно, можно было бы получить ладейный полином для C, перемножая ладейный полином для части C_1 и ладейный полином для части C_2. Действительно,

$$R(x, C_1) = 1 + 5x + 4x^2$$

и

$$R(x, C_2) = 1 + 8x + 12x^2,$$

а их произведение

$$\begin{aligned} R(x, C_1)R(x, C_2) &= (1 + 5x + 4x^2)(1 + 8x + 12x^2) = \\ &= 1 + 13x + 56x^2 + 92x^3 + 48x^4 = \\ &= R(x, C). \end{aligned}$$

□

На основе изложенного выше примера приходим к следующей теореме.

ТЕОРЕМА 12.21. Если доска C состоит из двух частей, C_1 и C_2, не имеющих ни общих строк, ни общих столбцов, то $R(x, C) = R(x, C_1)R(x, C_2)$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим коэффициент $r_k(C)$ при x^k в полиноме $R(x, C)$. Для выбора k неатакующих ладей на доске C следует выбрать l из них из части C_1 и $k - l$ — из части C_2. Существуют $r_l(C_1)$ способов выбора l неатакующих ладей из части C_1 и $r_{k-l}(C_2)$ способов выбора $k - l$ неатакующих ладей из части C_2. Это дает $r_l(C_1)r_{k-l}(C_2)$ способов выбора l неатакующих ладей из части C_1 и $k - l$ неатакующих ладей из части C_2. Если предположить, что l принимает все значения от 0 до k, получим

$$\begin{aligned} r_k(C) = {} & r_0(C_1)r_k(C_2) + r_1(C_1)r_{k-1}(C_2) + \cdots \\ & \cdots + r_l(C_1)r_{k-l}(C_2) + \cdots + r_k(C_1)r_0(C_2). \end{aligned}$$

Но это не что иное, как коэффициент при x^k в произведении

$$R(x, C_1)R(x, C_2) = (r_0(C_1) + r_1(C_1)x + \cdots + r_m(C_1)x^m)(r_0(C_2) + r_1(C_2)x + \cdots + r_m(C_2)x^m). \quad \blacksquare$$

ПРИМЕР 12.22. Снова рассмотрим доску, изображенную на рис. 12.12.

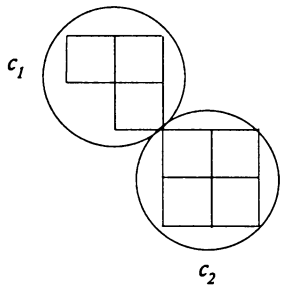


Рис. 12.12

Имеем

$$R(x, C_1) = 1 + 3x + x^2$$

и

$$R(x, C_2) = 1 + 4x + 2x^2.$$

Следовательно,

$$\begin{aligned} R(x, C_1) &= R(x, C_1)R(x, C_2) = \\ &= (1 + 3x + x^2)(1 + 4x + 2x^2) = \\ &= 1 + 7x + 15x^2 + 10x^3 + 2x^4. \end{aligned}$$ □

Существует еще один способ нахождения $R(x, C)$ путем разбиения C на доски. В этом случае доски могут иметь общие строки и столбцы.

ТЕОРЕМА 12.23. Пусть задана доска C, пусть s — квадрат доски C. Пусть C_s — доска C, в которой удален квадрат s. Пусть $C_s^{\#}$ — доска C, в которой удалены строка и столбец, содержащие удаленную клетку s. Тогда $R(x, C) = xR(x, C_s^{\#}) + R(x, C_s)$.

ДОКАЗАТЕЛЬСТВО. Пусть k — положительное целое число. Если k неатакующих ладей помещены на доске C и она имеет m клеток, то на клетке s либо есть ладья, либо ее нет. Если ладья помещена на клетке s, то остальные $k - 1$ ладей должны быть выбраны из доски $C_s^{\#}$, что можно осуществить $r_{k-1}(C_s^{\#})$ способами. Если ладья не помещена на клетке s, то k ладей должны быть выбраны из доски C_s, что можно осуществить $r_k(C_s)$ способами. Следовательно,

$$r_k(C) = r_{k-1}(C_s^{\#}) + r_k(C_s),$$

и поскольку $r_0(C) = r_0(C_s) = 1$,

$$\sum_{k=0}^{m} r_k(C)x^k = \sum_{k=1}^{m} r_{k-1}(C_s^{\#})x^k + \sum_{k=0}^{m} r_k(C_s)x^k.$$

Но

$$\sum_{k=1}^{m} r_{k-1}(C_s^{\#})x^k = \sum_{k=0}^{m-1} r_k(C_s^{\#})x^{k+1} = x\sum_{k=0}^{m-1} r_k(C_s^{\#})x^k = x\sum_{k=0}^{m} r_k(C_s^{\#})x^k,$$

поскольку $r_m(C_s^{\#}) = 0$, что должно равняться 0, т.к. $C_s^{\#}$ содержит менее m клеток. Следовательно,

$$R(x, C) = xR(x, C_s^{\#}) + R(x, C_s). \qquad \blacksquare$$

ПРИМЕР 12.24. Пусть C — доска, изображенная на рис. 12.13, где s — отмеченная клетка. В таком случае C_s — доска, изображенная на рис. 12.14, и доска $C_s^{\#}$ — фигура, изображенная на рис. 12.15.

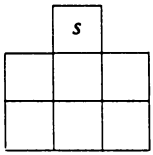


Рис. 12.13

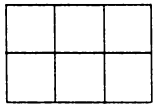


Рис. 12.14

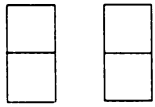


Рис. 12.15

Ладейные полиномы имеют вид

$$R(x, C_s) = 1 + 6x + 6x^2$$

и

$$R(x, C_s^{\#}) = 1 + 4x + 2x^2.$$

Следовательно,

$$\begin{aligned} R(x, C) &= xR(x, C_s^{\#}) + R(x, C_s) = \\ &= x(1 + 4x + 2x^2) + 1 + 6x + 6x^2 = \\ &= 1 + 7x + 10x^2 + 2x^3. \end{aligned}$$ □

ПРИМЕР 12.25. Пусть C — доска, изображенная на рис. 12.16, где s — отмеченная клетка. Тогда C_s — доска, изображенная на рис. 12.17, и $C_s^{\#}$ — доска, изображенная на рис. 12.18.

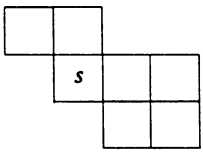


Рис. 12.16

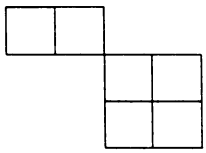


Рис. 12.17

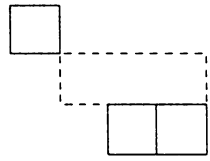


Рис. 12.18

$$R(x, C_s) = R(x, C'_s)R(x, C''_s),$$

где C'_s — верхний левый прямоугольник и C''_s — нижний правый квадрат.

$$R(x, C'_s) = 1 + 2x$$

и

$$R(x, C''_s) = 1 + 4x + 2x^2.$$

Следовательно,

$$R(x, C_s) = (1 + 2x)(1 + 4x + 2x^2) = 1 + 6x + 10x^2 + 4x^3.$$

Легко заметить, что

$$R(x, C^{\#}_s) = 1 + 3x + 2x^2.$$

Следовательно,

$$\begin{aligned} R(x, C) &= xR(x, C^{\#}_s) + R(x, C_s) = \\ &= x(1 + 3x + 2x^2) + 1 + 6x + 10x^2 + 4x^3 = \\ &= 1 + 7x + 13x^2 + 6x^3. \end{aligned}$$

□

Изучению ладейных полиномов было уделено достаточно времени, поэтому представляется разумным теперь указать на некоторое их применение. Вспомним перестановочные матрицы. Каждая из них имеет только по одной единице в каждой строке и в каждом столбце, и нули — во всех остальных позициях. Если ладьи отождествлять с единицами, то перестановку n элементов можно отождествить с размещением n неатакующих ладей на шахматной доске размером $n \times n$. Если рассматривать перестановку ϕ на n элементах как функцию, то ее можно представить как множество упорядоченных пар $\{(k, \phi(k)) : k = 1, 2, 3, \ldots n\}$. Если отождествить упорядоченную пару $(k, \phi(k))$ с клеткой шахматной доски, лежащей на пересечении строки k и столбца $\phi(k)$, то опять приходим к размещению n неатакующих ладей на шахматной доске размером $n \times n$. Поэтому различные перестановки на n элементах можно отождествлять с различными расположениями n неатакующих ладей на шахматной доске размером $n \times n$.

Предположим, что студент хочет прослушать 5 различных дисциплин в течение пяти различных лекционных часов. Если изобразить ситуацию с помощью шахматной доски, то ее заштрихованные области могут изображать лекционные часы, когда требуемые дисциплины не читаются. Пусть ϕ отображает каждую дисциплину на период, когда ее читают. Как и ранее, будем рассматривать ϕ как перестановку. В таком случае, у нас имеются те значения k, которые не должна принимать функция $\phi(k)$. Например, предположим, что $n = 5$, и $\phi(1)$ не должно принимать значения 2, 3 или 5; $\phi(2)$ не должно равняться 1, 2 или 3; $\phi(3)$ не должно равняться 1, 3 или 5; $\phi(4)$ не должно принимать значения 1, 2 или 5 и $\phi(5)$ не должно быть равно 1, 3 или 4. Перестановки, удовлетворяющие этим ограничениям, называются ***допустимыми размещениями*** на доске. Запрещенная область, или запрещенные позиции, — это то, что принадлежит заштрихованной области шахматной доски, изображенной на рис. 12.19.

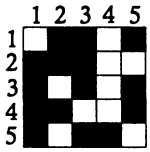


Рис. 12.19

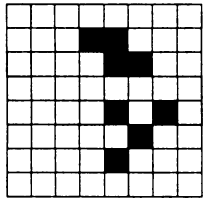


Рис. 12.20

Вполне очевидно, что допустимыми размещениями являются

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 3 & 5 \end{pmatrix}.$$

Предположим теперь, что клеток в запрещенной области сравнительно мало. Пусть имеется доска, запрещенная область которой изображена на рисунке 12.20. Работать с запрещенной областью, которая весьма мала, гораздо проще. Предположим, что имеется доска размером $n \times n$, у которой запрещенная область A относительно мала. Пусть A_i — множество всех перестановок ϕ таких, что $\phi(i)$ принадлежит запрещенной области A. В таком случае нас интересует $|A'_1 \cap A'_2 \cap \ldots \cap A'_n|$. Но

$$|A'_1 \cap A'_2 \cap \ldots \cap A'_n| = n! - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \\ - \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \cdots + (-1)^n |A_1 \cap A_2 \cap \ldots \cap A_n|,$$

поэтому необходимо найти значения всех членов, входящих в правую часть уравнения. Рассмотрим $|A_i|$ для фиксированного значения i. Пусть n_i — количество клеток запрещенной области, которые находятся в строке i. Тогда существуют n_i способов выбора $\phi(i)$ таких, что $\phi(i)$ принадлежит A_i. Но для $k \neq i$ функция $\phi(k)$ может принимать любое значение, поскольку ϕ является перестановкой. Следовательно, имеются $(n-1)!$ способов выбора других значений функции ϕ. Следовательно, и в области A_i существуют $n_i(n-1)!$ перестановок. Суммируя, получаем

$$\sum_i |A_i| = \sum_i n_i(n-1)! = \left(\sum_i n_i\right)(n-1)!.$$

Но $(\sum_i n_i) = r_1(A)$, количеству клеток в запрещенной области A. Следовательно,

$$\sum_i |A_i| = r_1(A)(n-1)!.$$

Теперь рассмотрим $\sum_{i<j} |A_i \cap A_j|$. Пусть i и j — фиксированы. Чтобы ϕ принадлежало $A_i \cap A_j$, требуется, чтобы $\phi(i)$ и $\phi(j)$ — принадлежали области A и находились в различных столбцах. Пусть Let n_{ij} — количество способов таких,

что $\phi(i)$ и $\phi(j)$ — принадлежат области A и не находятся в одном и том же столбце. Поэтому существуют n_{ij} способов выбора $\phi(i)$ и $\phi(j)$. Для $k \neq i, j$ значение $\phi(k)$ может быть любым, поскольку ϕ является перестановкой. Следовательно, имеются $(n-2)!$ способов выбора другого значения функции ϕ. Следовательно, $A_i \cap A_j$ содержит $n_{ij}(n-2)!$ перестановок. Суммируя по i и j, получаем

$$\sum_{i<j} |A_i \cap A_j| = \sum_{i<j} n_{ij}(n-2)! = \left(\sum_{i<j} n_{ij} \right) (n-2)!.$$

Но $\sum_{i<j} n_{ij}$ — всего лишь количество способов, которыми можно расположить две неатакующие ладьи в области A. Поэтому $\sum_{i<j} n_{ij} = r_2(A)$ и

$$\sum_i |A_i| = r_2(A)(n-2)!.$$

Аналогично, для нахождения $\sum_{i_1<i_2<\cdots<i_m} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}|$ положим $n_{i_1 i_2 \ldots i_m}$ равным количеству способов таких, что $\phi(i_1), \phi(i_2), \ldots, \phi(i_m)$ все принадлежат области A и не находятся в одном столбце. Поэтому имеются $n_{i_1 i_2 \ldots i_m}$ способов выбора $\phi(i_1), \phi(i_2), \ldots, \phi(i_m)$. Для $k \neq i_1, i_2, \ldots, i_m$ значение $\phi(k)$ может быть любым, поскольку ϕ — перестановка. Следовательно, имеются $(n-m)!$ способов выбора другого значения для ϕ, и $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}$ содержит $n_{i_1 i_2 \ldots i_m}(n-m)!$ перестановок. Суммируя по i_1, i_2, ..., i_m, имеем

$$\sum_{i_1<i_2<\cdots<i_m} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}| = \sum_{i_1<i_2<\cdots<i_m} n_{i_1 i_2 \ldots i_m}(n-m)! =$$

$$= \left(\sum_{i_1<i_2<\cdots<i_m} n_{i_1 i_2 \ldots i_m} \right) (n-m)!.$$

Но $\sum_{i_1<i_2<\cdots<i_m} n_{i_1 i_2 \ldots i_m}$ — всего лишь количество способов размещения m неатакующих ладей в области A. Следовательно, $\sum_{i_1<i_2<\cdots<i_m} n_{i_1 i_2 \ldots i_m} = r_m(A)$ и

$$\sum_{i_1<i_2<\cdots<i_m} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}| = r_m(A)(n-m)!.$$

Поэтому

$$|A'_1 \cap A'_2 \cap \ldots \cap A'_n| = n! - r_1(A)(n-1)! + r_2(A)(n-2)! + \cdots + (-1)^n r_n(A),$$

где ладейный полином для области A, $R(x, A) = \sum_{k=0}^{n} r_k(A) x^k$.

ПРИМЕР 12.26. Рассмотрим доску, изображенную на рис. 12.21, где заштрихованная область A — запрещенная зона. Найдем количество допустимых размещений.

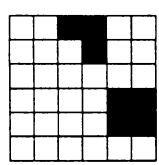


Рис. 12.21

Требуется найти

$$|A_1' \cap A_2' \cap A_3' \cap A_4' \cap A_5' \cap A_6'| = 6! - r_1(A) \cdot 5! + r_2(A\cdot) \cdot 4! + \cdots + (-1)^6 r_6(A).$$

Уже известно, что полином ладьи для области A имеет вид

$$R(x, A) = 1 + 7x + 15x^2 + 10x^3 + 2x^4,$$

поэтому количество допустимых размещений определяет соотношение

$$|A_1' \cap A_2' \cap A_3' \cap A_4' \cap A_5' \cap A_6'| = 6! - 7 \cdot 5! + 15 \cdot 4! - 10 \cdot 3! + 2 \cdot 2! = 184. \quad \square$$

ПРИМЕР 12.27. Рассмотрим доску, изображенную на рис. 12.22, где заштрихованная область A — запрещенная зона. Найдем количество возможных размещений.

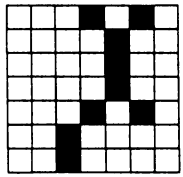


Рис. 12.22

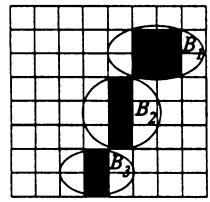


Рис. 12.23

Требуется найти

$$|A_1' \cap A_2' \cap \cdots \cap A_7'| = 7! - r_1(A) \cdot 6! + r_2(A\cdot) \cdot 5! + \cdots + (-1)^7 r_7(A),$$

поэтому нужен ладейный полином для области A. Заметим, что на доске, изображенной на рис. 12.23, заштрихованная область B имеет такой же ладейный полином, как и область A. Заштрихованная область B состоит из областей B_1, B_2 и B_3, у которых нет общих строк и столбцов. Поэтому

$$R(x, A) = R(x, B) = R(x, B_1) \cdot R(x, B_2) \cdot R(x, B_3).$$

Поскольку

$$\begin{aligned} R(x, B_1) &= 1 + 4x + 2x^2; \\ R(x, B_2) &= 1 + 3x; \\ R(x, B_3) &= 1 + 2x, \end{aligned}$$

то имеем

$$\begin{aligned} R(x, A) &= (1 + 4x + 2x^2)(1 + 3x)(1 + 2x) = \\ &= 1 + 9x + 28x^2 + 34x^3 + 12x^4. \end{aligned}$$

Следовательно,

$$|A_1' \cap A_2' \cap \cdots \cap A_7'| = 7! - 9 \cdot 6! + 28 \cdot 5! - 34 \cdot 4! + 12 \cdot 3! = 1176,$$

так что количество возможных размещений равно 1176. $\square$

■ УПРАЖНЕНИЯ

1. Найдите ладейный полином для доски, изображенной на рис. 12.24.
2. Найдите ладейный полином для доски, изображенной на рис. 12.25.

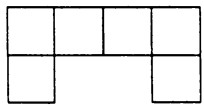


Рис. 12.24

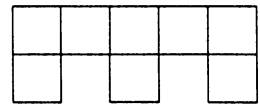


Рис. 12.25

3. Найдите ладейный полином для доски, изображенной на рис. 12.26.
4. Найдите ладейный полином для доски, изображенной на рис. 12.27.

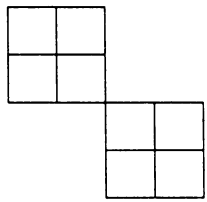


Рис. 12.26

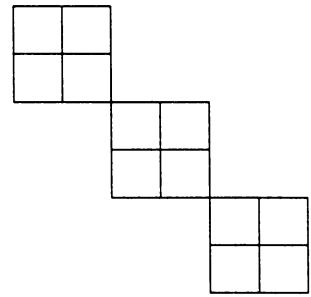


Рис. 12.27

5. Найдите ладейный полином для доски, изображенной на рис. 12.28.
6. Найдите ладейный полином для доски, изображенной на рис. 12.29.

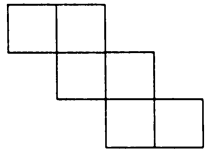


Рис. 12.28

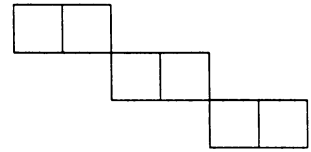


Рис. 12.29

7. Найдите ладейный полином для доски, изображенной на рис. 12.30.
8. Найдите ладейный полином для доски, изображенной на рис. 12.31.

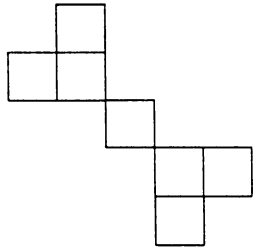


Рис. 12.30

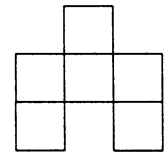


Рис. 12.31

9. Найдите ладейный полином для доски, изображенной на рис. 12.32.
10. Найдите ладейный полином для доски, изображенной на рис. 12.33.

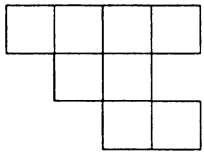


Рис. 12.32

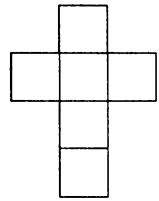


Рис. 12.33

11. Найдите ладейный полином для доски, изображенной на рис. 12.34.
12. Найдите ладейный полином для доски, изображенной на рис. 12.35.

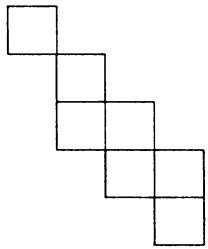


Рис. 12.34

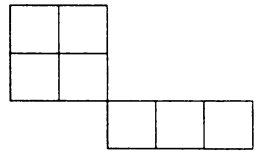


Рис. 12.35

13. Для заданной доски, изображенной на рис. 12.36, где заштрихованная область A — запрещенная зона, найдите количество допустимых размещений.
14. Для заданной доски, изображенной на рис. 12.37, где заштрихованная область A — запрещенная зона, найдите количество допустимых размещений.

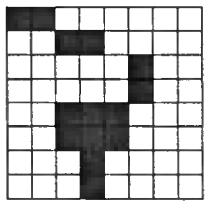


Рис. 12.36

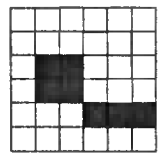


Рис. 12.37

15. Для заданной доски, изображенной на рис. 12.38, где заштрихованная область A — запрещенная зона, найдите количество допустимых размещений.
16. Для заданной доски, изображенной на рис. 12.39, где заштрихованная область A — запрещенная зона, найдите количество допустимых размещений.

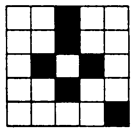


Рис. 12.38

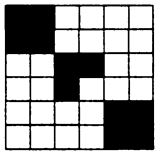


Рис. 12.39

17. Для заданной доски, изображенной на рис. 12.40, где заштрихованная область A — запрещенная зона, найдите количество допустимых размещений.

18. Для заданной доски, изображенной на рис. 12.41, где заштрихованная область A — запрещенная зона, найдите количество допустимых размещений.

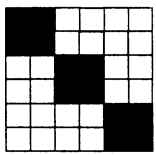


Рис. 12.40

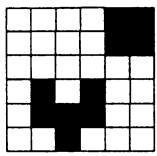


Рис. 12.41

19. Найдите количество допустимых перестановок на доске размером 4×4, когда запрещенные позиции — клетки вдоль обеих диагоналей.

20. В футбольной команде, состоящей из шести игроков, игрок A не может играть на позициях 1 и 3, игрок B — на позициях 2 и 4, игрок C — на позиции 6, игрок D не может играть на позициях 5 и 6, игрок E не может играть на позициях 5 и 6, а игрок F не может играть на позициях 1 и 3. Найдите количество возможных способов заполнить все позиции игроками.

21. Шестеро мужчин арендуют дом с шестью спальнями. Первый мужчина не хочет спать в спальнях 4 и 6. Второй не хочет спать в спальнях 4, 5 и 6. Третий мужчина не может спать в спальне 3. Четвертому мужчине не подходят спальни 1 и 2. Пятый не хочет спать в спальнях 1 и 2. Шестой мужчина не может спать в спальнях 1, 2 и 3. Сколькими способами жильцы могут распределить между собой спальни?

22. Футбольный тренер купил шесть новых автомобилей для шести лучших игроков. Он приобрел Volvo, Mercedes, BMW, Lexus, Lincoln и Cadillac. Первый игрок не хочет управлять Cadillac или Lincoln. Второй игрок не будет управлять Cadillac или Lincoln. Третий не будет управлять Lexus. Чевертый игрок не будет управлять Mercedes. Пятый игрок не будет управлять Volvo или BMW. Шестой не будет управлять Volvo или BMW. Сколькими способами тренер может раздать машины?

23. ***Латинский квадрат*** размером $n \times n$ — это квадратная таблица, в которой элементы каждой строки и элементы каждого столбца являются перестановкой первых n положительных целых чисел. Таким образом, каждое число появляется один раз в каждой строке и в каждом столбце. Предположим, что первые две строки чисел латинского квадрата размером 4×4 имеют вид $1, 4, 3, 2$ и $4, 1, 2, 3$. Сколько можно образовать различных третьих строк?

Указание: в первой строке таблицы вычеркните позицию, на которой присутствует 1, и т.д.

24. Предположим, что первые две строки латинского квадрата размером 4×4 имеют вид $4, 3, 1, 2$ и $3, 4, 2, 1$. Сколько можно образовать различных третьих строк?

25. Предположим, что первые две строки латинского квадрата размером 5×5 имеют вид $1, 3, 2, 4, 5$ и $2, 1, 3, 5, 4$. Сколько можно образовать различных третьих строк?

26. Предположим, что первые две строки чисел латинского квадрата размером 5×5 имеют вид $1, 2, 3, 4, 5$ и $3, 4, 5, 2, 1$. Сколько можно образовать различных третьих строк?

27. Предположим, что первые две строки чисел латинского квадрата размером 5×5 имеют вид $1, 2, 3, 4, 5$ и $2, 3, 4, 5, 1$. Сколько можно образовать различных третьих строк?

ГЛАВА

13

ПРОИЗВОДЯЩИЕ ФУНКЦИИ

13.1. ОПРЕДЕЛЕНИЕ ПРОИЗВОДЯЩЕЙ ФУНКЦИИ

В этом разделе вводится абстрактное определение производящей функции. Читатель, не желающий выходить на такой уровень абстракции, может пропустить этот раздел, не теряя нити изложения.

В дальнейшем мы убедимся, что производящие функции, которые мы сейчас определим, являются основой эффективных методов решения рекуррентных отношений и выполнения комбинаторных подсчетов.

Пусть $a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots$ — последовательность действительных чисел. Запишем ее в виде $\langle a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots\rangle$. Сумма двух последовательностей определяется следующим соотношением:

$$\langle a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots\rangle + \langle b_0, b_1, b_2, b_3, b_4, \ldots, b_n, \ldots\rangle = \\ = \langle a_0 + b_0, a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4, \ldots, a_n + b_n, \ldots\rangle.$$

Для действительного числа c введем операцию

$$c \circ \langle a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots\rangle = \langle ca_0, ca_1, ca_2, ca_3, ca_4, \ldots, ca_n, \ldots\rangle.$$

Кроме того, определим операцию

$$x \circ \langle a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots\rangle = \langle 0, a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots\rangle,$$

где x сдвигает каждый элемент последовательности на одну позицию вправо, а в оставшуюся свободной помещает 0. Наконец, определим

$$x^{k+1} \circ \langle a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots\rangle = x \circ (x^k \circ \langle 0, a_0, a_1, a_2, a_3, a_4, \ldots, a_n, \ldots\rangle),$$

так что

$$x \circ \langle 1, 0, 0, 0, 0, \ldots\rangle = \langle 0, 1, 0, 0, 0, \ldots\rangle;$$
$$x^2 \circ \langle 1, 0, 0, 0, 0, \ldots\rangle = x \circ \langle 0, 1, 0, 0, 0, \ldots\rangle = \langle 0, 0, 1, 0, 0, \ldots\rangle;$$

$$x^3 \circ \langle 1,0,0,0,0,\ldots\rangle = x \circ (x^2 \circ \langle 1,0,0,0,0,\ldots\rangle) = \\ = x \circ \langle 0,0,1,0,0,\ldots\rangle = \langle 0,0,0,1,0,\ldots\rangle.$$

Методом индукции можно доказать, что $x^k \circ \langle 1,0,0,0,0,\ldots\rangle$ — последовательность, в которой после k нулей стоит единица, а оставшаяся часть последовательности состоит из нулей. Поскольку $\langle 1,0,0,0,0,\ldots\rangle$ — это последовательность $1,0,0,0,\ldots$, ее можно отождествить с числом 1. Таким образом,

$$x = x \cdot 1 = x \circ \langle 1,0,0,0,0,\ldots\rangle = \langle 0,1,0,0,0,\ldots\rangle;$$
$$x^2 = x^2 \cdot 1 = x^2 \circ \langle 1,0,0,0,0,\ldots\rangle = \langle 0,0,1,0,0,\ldots\rangle;$$
$$x^3 = x^3 \cdot 1 = x^3 \circ \langle 1,0,0,0,0,\ldots\rangle = \langle 0,0,0,1,0,0,\ldots\rangle;$$
$$x^4 = x^4 \cdot 1 = x^4 \circ \langle 1,0,0,0,0,\ldots\rangle = \langle 0,0,0,0,1,0,0,\ldots\rangle,$$

и $x^k = x^k \circ \langle 1,0,0,0,0,\ldots\rangle$ — последовательность, в которой после k нулей стоит единица, а оставшаяся часть последовательности состоит из нулей. Понятно, что

$$\langle a_0,a_1,a_2,a_3,a_4,0,0,0,0,\ldots\rangle = a_0 \circ \langle 1,0,0,0,0,0,\ldots\rangle + \\ + a_1 \circ \langle 0,1,0,0,0,0,\ldots\rangle + \\ + a_2 \circ \langle 0,0,1,0,0,0,\ldots\rangle + \\ + a_3 \circ \langle 0,0,0,1,0,0,\ldots\rangle + \\ + a_4 \circ \langle 0,0,0,0,1,0,\ldots\rangle = \\ = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$$

и

$$\langle a_0,a_1,a_2,a_3,\ldots,a_n,0,0,0,0,\ldots\rangle = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n.$$

В общем случае

$$\langle a_0,a_1,a_2,a_3,a_4,\ldots,a_n,\ldots\rangle = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n + \cdots.$$

Выражение $a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n \cdots$ называется ***производящей функцией*** последовательности $a_0, a_1, a_2, a_3, \ldots$

В разделе 13.5 используются производящие функции несколько иного типа, называемые экспоненциальными. Для обычной производящей функции

$$\langle 1,1,1,1,1,1,\ldots\rangle = 1 + x + x^2 + x^3 + x^4 + x^5 + \cdots,$$
$$\langle a_0,a_1,a_2,a_3,a_4,\cdots\rangle = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots$$

и

$$\langle a_0,a_1,a_2,a_3,\ldots,a_n,0,0,0,0,\ldots\rangle = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n,$$

а для экспоненциальной производящей функции мы полагаем

$$\langle 1,1,1,1,1,1,\ldots\rangle = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots.$$

Тогда

$$\langle a_0, a_1, a_2, a_3, a_4, \cdots \rangle = a_0 + a_1\frac{x}{1!} + a_2\frac{x^2}{2!} + a_3\frac{x^3}{3!} + a_4\frac{x^4}{4!} + \cdots$$

и

$$\langle a_0, a_1, a_2, a_3, \ldots, a_n, 0, 0, 0, 0, \ldots \rangle = a_0 + a_1\frac{x}{1!} + a_2\frac{x^2}{2!} + a_3\frac{x^3}{3!} + \cdots + a_n\frac{x^n}{n!}.$$

13.2. ПРОИЗВОДЯЩИЕ ФУНКЦИИ И РЕКУРРЕНТНЫЕ ОТНОШЕНИЯ

В этом разделе рассматривается использование производящей функции для решения рекуррентных отношений. В следующих трех разделах рассматривается применение производящих функций для комбинаторных подсчетов. Назовем $a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n + \cdots$ ***производящей функцией*** последовательности a_0, a_1, a_2, a_3, ...

Пусть $f(x)$ и $g(x)$ — многочлены. Определим отношение $\dfrac{f(x)}{g(x)}$ как

$$\frac{f(x)}{g(x)} = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n \cdots$$

тогда и только тогда, когда

$$f(x) = g(x)(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n \cdots).$$

ОПРЕДЕЛЕНИЕ 13.1. Если $\dfrac{f(x)}{g(x)} = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n \cdots$, то $a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n + \cdots$ называется ***разложением*** $\dfrac{f(x)}{g(x)}$.

ТЕОРЕМА 13.2. Для любого m имеет место разложение

$$\frac{1}{(1-ax)^m} = 1 + \binom{m}{1}ax + \binom{m+1}{2}a^2x^2 + \\ + \binom{m+2}{3}a^3x^3 + \cdots + \binom{m+n-1}{n}a^nx^n + \cdots.$$

ДОКАЗАТЕЛЬСТВО. Используя метод индукции, покажем для $m = 1$, что

$$\frac{1}{(1-ax)} = 1 + ax + a^2x^2 + a^3x^3 + \cdots + a^nx^n + \cdots,$$

что эквивалентно

$$1 = (1-ax)(1 + ax + a^2x^2 + a^3x^3 + \cdots + a^nx^n + \cdots).$$

Но

$$(1-ax)(1+ax+a^2x^2+a^3x^3+\cdots+a^nx^n+\cdots)=$$
$$=1+ax+a^2x^2+\cdots+a^nx^n+\cdots-ax(1+ax+a^2x^2+\cdots+a^nx^n+\cdots)=$$
$$=1+ax+a^2x^2+\cdots+a^nx^n+\cdots-ax-a^2x^2-\cdots-a^nx^n-\cdots=1.$$

Таким образом, для $m=1$ теорема справедлива.

Предположим, что теорема справедлива для $m=k$, то есть, что

$$\frac{1}{(1-ax)^k}=1+\binom{k}{1}ax+\binom{k+1}{2}a^2x^2+$$
$$+\binom{k+2}{3}a^3x^3+\cdots+\binom{k+n-1}{n}a^nx^n+\cdots$$

Требуется доказать, что

$$\frac{1}{(1-ax)^{k+1}}=1+\binom{k+1}{1}ax+\binom{k+2}{2}a^2x^2+$$
$$+\binom{k+3}{3}a^3x^3+\cdots+\binom{k+n}{n}a^nx^n+\cdots$$

Поскольку

$$\frac{1}{(1-ax)^{k+1}}=\frac{1}{(1-ax)}\frac{1}{(1-ax)^k},$$

имеем

$$\frac{1}{(1-ax)^k}=(1-ax)\frac{1}{(1-ax)^{k+1}}.$$

Поэтому приведенное выше выражение для производящей функции последовательности $\frac{1}{(1-ax)^{k+1}}$ верно тогда и только тогда, когда после умножения на $(1-ax)$ мы получаем производящую функцию последовательности $\frac{1}{(1-ax)^k}$.

Но

$$(1-ax)(1+\tbinom{k+1}{1}ax+\tbinom{k+2}{2}a^2x^2+\tbinom{k+3}{3}a^3x^3+\cdots+\tbinom{k+n}{n}a^nx^n+\cdots)=$$
$$=1+\tbinom{k+1}{1}ax+\tbinom{k+2}{2}a^2x^2+\tbinom{k+3}{3}a^3x^3+\cdots+\tbinom{k+n}{n}a^nx^n+\cdots$$
$$\cdots-(ax)(1+\tbinom{k+1}{1}ax+\tbinom{k+2}{2}a^2x^2+\tbinom{k+3}{3}a^3x^3+\cdots+\tbinom{k+n}{n}a^nx^n+\cdots)=$$
$$=1+\tbinom{k+1}{1}ax+\tbinom{k+2}{2}a^2x^2+\tbinom{k+3}{3}a^3x^3+\cdots+\tbinom{k+n}{n}a^nx^n+\cdots-$$
$$-ax-\tbinom{k+1}{1}a^2x^2-\tbinom{k+2}{2}a^3x^3-\cdots-\tbinom{k+n-1}{n-1}a^nx^n-\cdots=$$
$$=1+\tbinom{k}{1}ax+\tbinom{k+1}{2}a^2x^2+\tbinom{k+2}{3}a^3x^3+\cdots+\tbinom{k+n-1}{n}a^nx^n+\cdots,$$

поскольку, согласно тождеству Паскаля, имеем

$$\binom{k+n-1}{n}+\binom{k+n-1}{n-1}=\binom{k+n}{n},$$

так что

$$\binom{k+n}{n} - \binom{k+n-1}{n-1} = \binom{k+n-1}{n}.$$

Следовательно,

$$(1-ax)^{k+1}(1+\tbinom{k+1}{1}ax+\tbinom{k+2}{2}a^2x^2+\tbinom{k+3}{3}a^3x^3+\cdots+\tbinom{k+n}{n}a^nx^n+\cdots) =$$
$$=(1-ax)^k(1+\tbinom{k}{1}ax+\tbinom{k+1}{2}a^2x^2+\tbinom{k+2}{3}a^3x^3+\cdots+\tbinom{k+n-1}{n}a^nx^n+\cdots).$$

Но, по индуктивному предположению,

$$(1-ax)^k(1+\tbinom{k}{1}ax+\tbinom{k+1}{2}a^2x^2+\tbinom{k+2}{3}a^3x^3+\cdots+\tbinom{k+n-1}{n}a^nx^n+\cdots)=1,$$

что и требовалось доказать. ■

Из доказанной теоремы следует, что

$$\frac{1}{(1-ax)^2} = 1+2ax+3a^2x^2+4a^3x^3+\cdots+(n+1)a^nx^n+\cdots$$

и

$$\frac{1}{(1-ax)^3} = 1+3ax+6a^2x^2+10a^3x^3+\cdots+\frac{(n+1)(n+2)}{2}a^nx^n+\cdots.$$

Из математического анализа известно, что любая дробь вида

$$\frac{Ax+B}{(1-Cx)(1-Dx)},$$

где константы C и D различны, равна

$$\frac{a}{1-Cx}+\frac{b}{1-Dx}$$

для некоторых постоянных a и b. Такое представление называется ***разложением на элементарные дроби***.

Более того, любая дробь вида

$$\frac{Ax^2+Bx+C}{(1-Dx)(1-Ex)(1-Fx)},$$

где константы D, E и F различны, равна

$$\frac{a}{1-Dx}+\frac{b}{1-Ex}+\frac{c}{1-Fx}$$

для некоторых постоянных a, b и c. Также, если константы D и E различны, то

$$\frac{Ax^2+Bx+C}{(1-Dx)^2(1-Ex)} = \frac{a}{1-Dx}+\frac{b}{(1-Dx)^2}+\frac{c}{1-Ex}$$

для некоторых постоянных a, b и c, и если константы E и F различны, то

$$\frac{Ax^3+Bx^2+Cx+D}{(1-Ex)^3(1-Fx)}=\frac{a}{1-Ex}+\frac{b}{(1-Ex)^2}+\frac{c}{(1-Ex)^3}+\frac{d}{1-Fx}$$

для некоторых постоянных a, b, c и d.

Предположим, что требуется найти разложение производящей функции

$$\frac{15x}{1+3x-4x^2}=\frac{15x}{(1+4x)(1-x)}.$$

Пусть

$$\frac{15x}{(1+4x)(1-x)}=\frac{a}{1+4x}+\frac{b}{1-x}$$

для некоторых a и b. Умножая обе части равенства на $(1+4x)(1-x)$, получаем

$$15x=a(1-x)+b(1+4x).$$

Полагая $x=1$, приходим к уравнению

$$15(1)=a(1-1)+b(1+4),$$

решая которое, получаем $b=3$. Полагая $x=-\frac{1}{4}$, приходим к уравнению

$$15(-\frac{1}{4})=a(1-\frac{1}{4})+b(1+4(-\frac{1}{4})),$$

решая которое, получаем $a=-5$, поэтому

$$\frac{15x}{(1+4x)(1-x)}=\frac{-5}{1+4x}+\frac{3}{1-x}.$$

Применяя формулу

$$\frac{1}{(1-ax)}=1+ax+a^2x^2+a^3x^3+\cdots+a^nx^n+\cdots,$$

при $a=-4$ и $a=1$, получаем

$$\frac{1}{1+4x}=1+(-4)x+(-4)^2x^2+(-4)^3x^3+\cdots+(-4)^nx^n+\cdots$$

и

$$\frac{1}{(1-x)}=1+x+x^2+x^3+\cdots+x^n+\cdots,$$

так что

$$\begin{aligned}\frac{-5}{1+4x}+\frac{3}{1-x}&=-5(1+(-4)x+(-4)^2x^2+(-4)^3x^3+\cdots\\&\cdots+(-4)^nx^n+\cdots)+3(1+x+x^2+x^3+\cdots+x^n+\cdots)=\\&=-2+23x-77x^2+323x^3+\cdots+(-5(-4)^n+3)x^n+\cdots.\end{aligned}$$

В дальнейшем нам понадобится еще одна теорема о производящих функциях.

ТЕОРЕМА 13.3. При всех $n\geq 0$ имеем $\dfrac{1-x^{n+1}}{1-x}=1+x+x^2+x^3+\cdots+x^n$.

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned}(1-x^{n+1})\frac{1}{1-x} &= (1-x^{n+1})(1+x+x^2+\cdots+x^n+\cdots) = \\ &= (1)(1+x+x^2+\cdots+x^n+\cdots) - x^{n+1}(1+x+x^2+\cdots+x^n+\cdots) = \\ &= (1+x+x^2+\cdots+x^n+x^{n+1}+\cdots) - (x^{n+1}+x^{n+2}+x^{n+2}+\cdots) = \\ &= 1+x+x^2+x^3+\cdots+x^n. \end{aligned}$$

■

Теперь рассмотрим применение производящих функций для решения рекуррентных отношений. Сначала проиллюстрируем метод на простом примере. Рассмотрим рекуррентное отношение $a_n = 3a_{n-1}$ или $a_n - 3a_{n-1} = 0$. Пусть

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots .$$

Умножим обе части равенства на $3x$, помня о том, что умножение на x сдвигает каждый элемент последовательности вправо, так что

$$3xf(x) = 3a_0x + 3a_1x^2 + 3a_2x^3 + 3a_3x^4 + \cdots + 3a_{n-1}x^n + 3a_nx^{n+1} + \cdots .$$

Вычитая $3xf(x)$ из $f(x)$, получаем

$$\begin{aligned}f(x) - 3xf(x) = a_0 + (a_1 - 3a_0)x + (a_2 - 3a_1)x^2 + \\ + (a_3 - 3a_2)x^3 + \cdots + (a_n - 3a_{n-1})x^n + \cdots .\end{aligned}$$

Но $a_n - 3a_{n-1} = 0$ при всех $n \geq 1$, поэтому

$$f(x) - 3xf(x) = a_0.$$

Таким образом,

$$(1-3x)f(x) = a_0$$

и

$$\begin{aligned}f(x) = \frac{a_0}{(1-3x)} &= \\ &= a_0(1+3x+3^2x^2+3^3x^3+\cdots+3^nx^n+\cdots) = \\ &= a_0 + 3a_0x + 3^2a_0x^2 + 3^3a_0x^3 + \cdots + 3^na_0x^n + \cdots .\end{aligned}$$

Следовательно, $a_n = 3^na_0$. Заметим, что мы умножали обе части равенства на $3x$, поэтому, вычитая, должны получить $a_n - 3a_{n-1}$. Но, по определению данного рекуррентного отношения, это выражение равно 0.

ПРИМЕР 13.4. Найдем функцию, заданную рекуррентными отношениями

$$\begin{aligned}a_0 &= 5; \\ a_k &= a_{k-1} + 3.\end{aligned}$$

Второе уравнение можно переписать в виде $a_k - a_{k-1} - 3 = 0$.

Снова положим

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots .$$

Поскольку коэффициент при a_{k-1} равен 1, то теперь требуется найти

$$xf(x) = a_0x + a_1x^2 + a_2x^3 + a_3x^4 + \cdots + a_nx^{n+1} + \cdots .$$

Кроме того, нам потребуется производящая функция, у которой при каждом члене стоит коэффициент 3, т.е.

$$\begin{aligned}\frac{3}{1-x} &= 3(1 + x + x^2 + x^3 + \cdots + x^n + \cdots) = \\ &= 3 + 3x + 3x^2 + 3x^3 + \cdots + 3x^n + \cdots .\end{aligned}$$

Теперь

$$\begin{aligned}f(x) - xf(x) - \frac{3}{1-x} = a_0 - 3 + (a_1 - a_0 - 3)x + \\ + (a_2 - a_1 - 3)x^2 + \cdots + (a_n - a_{n-1} - 3) + \cdots .\end{aligned}$$

Но $a_k - a_{k-1} - 3 = 0$ при всех $k \geq 1$, так что

$$f(x) - xf(x) - \frac{3}{1-x} = a_0 - 3 = 2.$$

Решая это уравнение относительно $f(x)$, получаем

$$\begin{aligned}f(x) &= \frac{3}{(1-x)^2} + \frac{2}{1-x} = \\ &= 3(1 + 2x + 3x^2 + \cdots + (n+1)x^n + \cdots) + \\ &\quad + 2(1 + x + x^2 + \cdots + x^n + \cdots) = \\ &= 5 + 8x + 11x^2 + 14x^3 + \cdots + (3n+5)x^n + \cdots ,\end{aligned}$$

так что $a_n = 3n + 5$. □

ПРИМЕР 13.5. Найдем функцию, заданную рекуррентными отношениями

$$\begin{aligned}a_0 &= 1; \\ a_1 &= 4; \\ a_k &= a_{k-1} + 6a_{k-2}.\end{aligned}$$

Как и ранее, положим

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots .$$

Поскольку коэффициент при a_{k-1} равен 1 и коэффициент при a_{k-2} равен 6, нужно найти

$$xf(x) = a_0x + a_1x^2 + a_2x^3 + a_3x^4 + \cdots + a_nx^{n+1} + \cdots$$

и

$$6x^2f(x) = 6a_0x^2 + 6a_1x^3 + 6a_2x^4 + 6a_3x^5 + \cdots + 6a_nx^{n+2} + \cdots .$$

Поэтому

$$f(x) - xf(x) - 6x^2f(x) = a_0 + (a_1 - a_0)x + (a_2 - a_1 - 6a_0)x^2 + \\ + (a_3 - a_2 - 6a_1)x^3 + \cdots + (a_n - a_{n-1} - 6a_{n-2})x^n + \cdots .$$

Но $a_nx^n - a_{n-1} - 6a_{n-2} = 0$ для всех $n \geq 1$, так что

$$f(x) - xf(x) - 6x^2f(x) \;=\; a_0 + (a_1 - a_0)x \;=\; 1 + 3x.$$

Решая уравнение относительно $f(x)$, получаем

$$f(x) = \frac{1 + 3x}{1 - x - 6x^2} = \frac{1 + 3x}{(1 - 3x)(1 + 2x)}.$$

Полагая

$$\frac{1 + 3x}{(1 - 3x)(1 + 2x)} = \frac{a}{(1 - 3x)} + \frac{b}{(1 + 2x)}$$

и умножая обе части равенства на $(1 - 3x)(1 + 2x)$, получаем

$$1 + 3x = a(1 + 2x) + b(1 - 3x).$$

Полагая $x = -\dfrac{1}{2}$, получаем

$$1 + \left(-\frac{3}{2}\right) = a\left(1 + 2\left(-\frac{1}{2}\right)\right) + b\left(1 - 3\left(-\frac{1}{2}\right)\right),$$

так что $b = -\dfrac{1}{5}$. Полагая $x = \dfrac{1}{3}$, получаем

$$1 + 1 = a\left(1 + 2\left(\frac{1}{3}\right)\right) + b\left(1 - 3\left(\frac{1}{3}\right)\right),$$

так что $a = \dfrac{6}{5}$. Следовательно, имеем

$$f(x) = \left(\frac{1}{5}\right)\left(\frac{6}{(1 - 3x)} - \frac{1}{(1 + 2x)}\right) = \\ = \left(\frac{6}{5}\right)(1 + 3x + 3^2x^2 + 3^3x^3 + \cdots + 3^nx^n + \cdots) - \\ - \left(\frac{1}{5}\right)1 + (-2)x + (-2)^2x^2 + (-2)^3x^3 + \cdots + (-2)^nx^n + \cdots ,$$

так что

$$a_n = \left(\frac{1}{5}\right)(6 \cdot (3)^n - (-2)^n).$$

□

ПРИМЕР 13.6. Найдем функцию, заданную рекуррентными отношениями

$$a_0 = 1;$$
$$a_k = 3a_{k-1} + 4^n.$$

Как и ранее, пусть

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots .$$

Поскольку коэффициент при $a_{k-1} = 3$, найдем

$$3xf(x) = 3a_0x + 3a_1x^2 + 3a_2x^3 + 3a_3x^4 + \cdots + 3a_nx^{n+1} + \cdots .$$

Так как

$$\frac{1}{(1-4x)} = (1 + 4x + 4^2x^2 + 4^3x^3 + \cdots + 4^nx^n + \cdots),$$

это даст 4^nx^n. Поэтому

$$f(x) - 3xf(x) - \frac{1}{(1-4x)} =$$
$$= a_0 - 1 + (a_1 - 3a_0 - 4)x + (a_2 - 3a_1 - 4^2)x^2 +$$
$$+ (a_3 - 3a_2 - 4^2)x^3 + \cdots + (a_n - a_{n-1} - 4^n)x^n + \cdots .$$

Поскольку $a_n - 3a_{k-1} - 4^n = 0$ при всех $n \geq 1$, имеем

$$f(x) - 3xf(x) - \frac{1}{(1-4x)} = a_0 - 1 = 0.$$

Решая уравнение относительно $f(x)$, имеем

$$f(x) = \frac{1}{(1-4x)(1-3x)}.$$

Пусть

$$\frac{1}{(1-4x)(1-3x)} = \frac{a}{(1-4x)} + \frac{b}{(1-3x)}.$$

Умножая обе части равенства на $(1-4x)(1-3x)$, имеем

$$1 = (1-3x)a + (1-4x)b.$$

Полагая $x = \frac{1}{3}$, получаем

$$1 = \left(1 - 3 \cdot \frac{1}{3}\right)a + \left(1 - 4 \cdot \frac{1}{3}\right)b,$$

так что $b = -3$. Полагая $x = \frac{1}{4}$, имеем

$$1 = \left(1 - 3 \cdot \frac{1}{4}\right)a + \left(1 - 4 \cdot \frac{1}{4}\right)b,$$

поэтому $a = 4$. Следовательно, имеем

$$f(x) = \frac{4}{(1-4x)} - \frac{3}{(1-3x)} = \\ = 4(1 + 4x + 4^2x^2 + 4^3x^3 + \cdots + 4^nx^n + \cdots) - \\ - 3(1 + 3x + 3^2x^2 + 3^3x^3 + \cdots + 3^nx^n + \cdots,$$

поэтому

$$a_n = 4 \cdot 4^n - 3 \cdot 3^n = 4^{n+1} - 3^{n+1}. \qquad \square$$

ПРИМЕР 13.7. Найдем функцию, заданную рекуррентными отношениями

$$a_0 = 3;$$
$$a_n = 2a_{n-1} + n.$$

Второе уравнение также можно переписать в виде $a_n - 2a_{n-1} - n = 0$.

Опять положим

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots.$$

Поскольку коэффициент при a_{k-1} равен -2, найдем

$$2xf(x) = 2a_0x + 2a_1x^2 + 2a_2x^3 + 2a_3x^4 + \cdots + 2a_nx^{n+1} + \cdots.$$

Кроме того, нам нужна производящая функция, каждый член которой имеет множителем n, но это есть

$$\frac{x}{(1-x)^2} = x(1 + 2x + 3x^2 + \cdots + nx^{n-1} + \cdots) = \\ = x + 2x^2 + 3x^3 + \cdots + nx^n + \cdots.$$

Теперь

$$f(x) - 2xf(x) - \frac{x}{(1-x)^2} = a_0 + (a_1 - 2a_0 - 1)x + \\ + (a_2 - 2a_1 - 2)x^2 + \cdots + (a_n - 2a_{n-1} - n) + \cdots.$$

Но $a_k - 2a_{k-1} - k = 0$ для всех $k \geq 1$, поэтому

$$f(x) - 2xf(x) - \frac{x}{(1-x)^2} = a_0 = 3.$$

Решая уравнение относительно $f(x)$, получаем

$$f(x) = \frac{x}{(1-2x)(1-x)^2} + \frac{3}{(1-2x)} = \frac{3 - 5x + 3x^2}{(1-2x)(1-x)^2} = \\ = \frac{A}{(1-x)} + \frac{B}{(1-x)^2} + \frac{C}{(1-2x)},$$

так что

$$3 - 5x + 3x^2 = A(1-x)(1-2x) + B(1-2x) + C(1-x)^2 = \\ = A(1 - 3x + 2x^2) + B(1-2x) + C(1 - 2x + x^2).$$

Приравнивая коэффициенты при x^2, x и 1, получаем соответственно

$$3 = 2A + C; \\ -5 = -3A - 2B - 2C; \\ 3 = A + B + C.$$

Решая эти уравнения, получаем $A = -1$, $B = -1$ и $C = 5$. Следовательно,

$$f(x) = \frac{-1}{(1-x)} - \frac{1}{(1-x)^2} + \frac{5}{(1-2x)} = \\ = -(1 + x + x^2 + \cdots + x^n + \cdots) - (1 + 2x + 3x^2 + \cdots \\ \cdots + (n+1)x^n + \cdots) + 5(1 + 2x + (2x)^2 + \cdots + (2x)^n \cdots).$$

и

$$a_n = -n - 2 + 5(2^n). \qquad \square$$

■ УПРАЖНЕНИЯ

1. Выразите приведенные ниже дроби через элементарные дроби:

а) $\dfrac{1}{(x+4)(x+3)}$;

б) $\dfrac{2x-3}{(x-2)(x-3)}$;

в) $\dfrac{x^2+2x+3}{(x-1)^2(x-2)}$;

г) $\dfrac{x}{(x-1)(x-2)(x-3)}$;

д) $\dfrac{3x^2+1}{(x-1)^2(x-2)^2}$.

2. Выразите приведенные ниже дроби через элементарные дроби:

а) $\dfrac{2x+1}{(x-4)(x+2)}$;

б) $\dfrac{2x-3}{(x-2)(x+4)}$;

в) $\dfrac{x^2+2x+3}{(x+1)(x-2)^2}$;

г) $\dfrac{x^2}{(x-1)(x+3)(x-5)}$;

д) $\dfrac{3x^2+1}{(x-1)^3}$.

3. Найдите разложения приведенных ниже производящих функций:

а) $\dfrac{1}{x+4}$;

б) $\dfrac{1}{(x-2)^2}$;

в) $\dfrac{1}{(x-1)^3}$;

г) $\dfrac{x}{(x-1)(x-2)}$;

д) $\dfrac{3x^2+1}{(x-1)^2(x-2)}$.

4. Найдите разложения приведенных ниже производящих функций:

а) $\dfrac{1}{x-3}$;

б) $\dfrac{x}{(x-1)^2}$;

в) $\dfrac{x}{(x-1)^3}$;

г) $\dfrac{x}{(x+1)(x-2)(x+3)}$;

д) $\dfrac{2x^2+6}{(x-1)^2(x+2)^2}$.

5. Используя производящие функции, решите следующие рекуррентные отношения:

а) $a_0 = 1$;
$a_n = 2a_{n-1} + 3^n$ при $n > 0$.

б) $a_0 = 2$;
$a_n = 3a_{n-1} + n$ при $n > 0$.

в) $a_0 = 8$;
$a_1 = 16$;
$a_n = 2a_{n-1} + 3a_{n-2}$ при $n > 1$.

г) $a_0 = 1$;
$a_1 = 0$;
$a_n = 4a_{n-1} - 4a_{n-2}$ при $n > 1$.

д) $a_0 = 4$;
$a_n = 2a_{n-1} - 3$ при $n > 0$.

6. Используя производящие функции, решите следующие рекуррентные отношения:

а) $a_0 = 1$;
$a_n = 2a_{n-1}$ при $n > 0$.

б) $a_0 = 3$;
$a_n = 4a_{n-1} + 2n$ при $n > 0$.

в) $a_0 = 1$;
$a_1 = 3$;
$a_n = 5a_{n-1} - 6a_{n-2}$ при $n > 1$.

г) $a_0 = 1$;
$a_1 = 3$;
$a_n = 7a_{n-1} - 10a_{n-2}$ при $n > 1$.

д) $a_0 = 1$;
$a_n = 2a_{n-1} + n - 1$ при $n > 0$.

7. Используя производящие функции, решите следующие рекуррентные отношения:

а) $a_0 = 1$;
$a_n = a_{n-1} + n$ при $n > 0$.

б) $a_0 = 2$;
$a_n = 3a_{n-1} + 2^n$ при $n > 0$.

в) $a_0 = 1$;
$a_n = 2a_{n-1} + 2^n$ при $n > 0$.

г) $a_0 = 1$;
$a_1 = 8$;
$a_n = 6a_{n-1} - 9a_{n-2} + 2^n$ при $n > 1$.

13.3. ПРОИЗВОДЯЩИЕ ФУНКЦИИ И КОМБИНАТОРНЫЕ ПОДСЧЕТЫ

В настоящем разделе речь пойдет о применении производящих функций для комбинаторных подсчетов. Это превосходный метод, который оценит каждый, кто столкнется с его использованием. Ранее мы рассматривали производящие функции как формальные ряды, не касаясь таких понятий анализа, как сходимость, ряды Тейлора, интегрирование и дифференцирование.

Для начала рассмотрим функцию $(1+x)^n$. Известно, что по формуле бинома Ньютона

$$(1+x)^n = \binom{n}{0}x^0 + \binom{n}{1}x^1 + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n,$$

так что коэффициент при x^k равен количеству способов выбора k объектов из n объектов без учета порядка. Следовательно, производящая функция вида $(1+x)^n$ является производящей функцией для числа сочетаний из n объектов по k.

По теореме 13.2

$$\frac{1}{(1-x)^m} = 1 + \binom{m}{1}x + \binom{m+1}{2}x^2 + \binom{m+2}{3}x^3 + \cdots + \binom{m+n-1}{n}x^n + \cdots,$$

поэтому коэффициент при x^k есть число сочетаний из m объектов по k с повторениями. Следовательно, производящая функция $\frac{1}{(1-x)^m}$ является производящей функцией, которая применяется для определения числа сочетаний из m объектов по k с повторениями.

Теперь рассмотрим произведение

$$(1_a + x_a)(1_b + x_b)(1_c + x_c + x_c^2)(1_d + x_d + x_d^2),$$

где нижние индексы отслеживают только "происхождение" каждого x и никак не влияют на его значение. Таким образом, x_a, x_b, x_c и x_d — это x, где нижний индекс — просто метка. Если положить $x_i^0 = 1_i$ при $i = a, b, c$ и d и рассмотреть слагаемое x^4, то мы увидим, что оно представляет собой сумму

$$x_a^0x_b^0x_c^2x_d^2 + x_a^0x_b^1x_c^2x_d^1 + x_a^1x_b^0x_c^2x_d^1 + x_a^1x_b^1x_c^2 1x_d^0 + \\ + x_a^0x_b^1x_c^1x_d^2 + x_a^1x_b^0x_c^1x_d^2 + x_a^1x_b^1x_c^0x_d^2 + x_a^1x_b^1x_c^1x_d^1$$

Сумма степеней в каждом слагаемом равна четырем, значит, имеются все возможные решения уравнения

$$e_a + e_b + e_c + e_d = 4$$

при $0 \le e_a \le 1$, $0 \le e_b \le 1$, $0 \le e_c \le 2$ и $0 \le e_d \le 2$. В общем случае, при $0 \le r \le 6$ коэффициент при x^r равен количеству решений уравнения

$$e_a + e_b + e_c + e_d = r$$

при $0 \le e_a \le 1$, $0 \le e_b \le 1$, $0 \le e_c \le 2$ и $0 \le e_d \le 2$.

Предположим, имеется множество A, содержащее 1 объект типа a, 1 объект типа b, 2 объекта типа c и 2 объекта типа d. Пусть x_i^j представляет j объектов типа i. Таким образом, 1_a обозначает 0 элементов типа a, x_b обозначает 1 элемент типа b и x_c^2 обозначает 2 элемента типа c. Например, $x_a^1x_b^1x_c^2 1x_d^0$ представляет выбор 1 объекта типа a, 1 объекта типа b и 2 объектов типа c. Тогда коэффициент при x^4 равен количеству способов выбора 4 объектов среди объектов 4 типов, принадлежащих множеству A. Легко представить, что $1_c + x_c + x_c^2$ означает выбор 0, 1 или 2 объектов типа c.

$$(1+x)(1+x+x^2)(1+x+x^2+x^3)(1+x+x^2+x^3+x^4+x^5).$$

Коэффициент при x^r представляет количество способов выбора r объектов из множества, содержащего 1 элемент типа a, 2 элемента типа b, 3 элемента типа c и 5 элементов типа d, если нижние индексы в произведении опущены, а также количество решений уравнения

$$e_a + e_b + e_c + e_d = r$$

при $0 \le e_a \le 1$, $0 \le e_b \le 2$, $0 \le e_c \le 3$ и $0 \le e_d \le 5$.

ПРИМЕР 13.8. В урне находятся 4 красных, 5 синих и 2 зеленых шара. (а) Сколько существует способов выбора 7 шаров из урны? (б) Сколько существует способов выбора из урны 7 шаров, если 1 шар красный и 2 синие.

Для части (а) производящая функция имеет вид

$$(1 + x + x^2 + x^3 + x^4)(1 + x + x^2 + x^3 + x^4 + x^5)(1 + x + x^2),$$

и количество способов выбора 7 шаров равно коэффициенту при x^7 в разложении этой производящей функции.

В части (б) нужно учесть, что 1 вытянутый шар красный, соответствующий многочлен имеет вид $(x + x^2 + x^3 + x^4)$, что представляет вытягивание 1, 2, 3 или 4 красных шаров. Точно так же, учитывая, что 2 вытянутых шара синие, соответствующий многочлен должен иметь вид $(x^2+x^3+x^4+x^5)$, что представляет вытягивание 2, 3, 4 или 5 синих шаров. Таким образом, производящий многочлен имеет вид

$$(x + x^2 + x^3 + x^4)(x^2 + x^3 + x^4 + x^5)(1 + x + x^2),$$

и количество способов выбора 7 шаров равно коэффициенту при x^7 в разложении производящей функции. □

ПРИМЕР 13.9. Допустим, в урне находятся 3 красных, 8 зеленых, 9 оранжевых и 2 белых шара. Сколько существует способов выбора 12 шаров из урны, если 1 выбранный шар — красный, количество выбранных зеленых шаров — четное, а количество оранжевых — нечетное? Учитывая, что хотя бы 1 выбранный шар должен быть красным, то могут быть выбраны 1, 2 или 3 таких шара, поэтому многочлен, представляющий красные шары, имеет вид

$$x + x^2 + x^3.$$

Поскольку количество выбранных зеленых шаров четное, выбранными могут быть 0, 2, 4, 6 или 8 таких шаров, поэтому многочлен, представляющий зеленые шары, имеет вид

$$1 + x^2 + x^4 + x^6 + x^8.$$

Количество выбранных оранжевых шаров должно быть нечетным, поэтому выбранными могут быть 1, 3, 5, 7 или 9 таких шаров, так что многочлен, представляющий оранжевые шары, имеет вид

$$x + x^3 + x^5 + x^7 + x^9.$$

Следовательно, производящей функцией для рассматриваемой задачи будет

$$(x + x^2 + x^3)(1 + x^2 + x^4 + x^6 + x^8)(x + x^3 + x^5 + x^7 + x^9)(1 + x + x^2),$$

а ответ на вопрос даст коэффициент при x^{12}. □

Для нахождения коэффициентов некоторых полиномиальных производящих функций введем в рассмотрение производящие функции из предыдущего раздела.

ПРИМЕР 13.10. Найдем коэффициент при x^{24} в разложении производящей функции $(x^3+x^4+x^5+x^6+\cdots)^4$.

$$\begin{aligned}(x^3+x^4+x^5+x^6+\cdots)^4 &= (x^3)^4(1+x+x^2+x^3+x^4+\cdots)^4 = \\ &= (x^{12})\left(\frac{1}{1-x}\right)^4 = \qquad \text{по теореме 13.2} \\ &= (x^{12})\frac{1}{(1-x)^4} = \\ &= (x^{12})(1+4x+\tbinom{5}{2}x^2+\cdots+\tbinom{4+n-1}{n}x^n+\cdots) \\ &\qquad \text{по теореме 13.2.}\end{aligned}$$

Следовательно, чтобы найти коэффициент при x^r в разложении этой производящей функции, необходимо найти коэффициент при x^{r-12} в разложении производящей функции

$$1+4x+\binom{5}{2}x^2+\cdots+\binom{4+n-1}{n}x^n+\cdots,$$

так что коэффициент при x^r равен

$$\binom{4+r-12-1}{r-12}=\binom{r-9}{r-12},$$

а коэффициент при x^{24} равен

$$\binom{24-9}{24-12}=\binom{15}{12}.$$ □

ПРИМЕР 13.11. Выясним, сколько существует способов выбора 12 объектов из совокупности объектов пяти типов, если необходимо выбрать не более 2 объектов первых трех типов и неограниченное количество объектов остальных двух типов.

Производящая функция имеет вид

$$(1+x+x^2)^3(1+x+x^2+x^3+\cdots)^2.$$

Но по теореме 13.3

$$(1+x+x^2)^3=\left(\frac{1-x^3}{1-x}\right)^3.$$

Поэтому

$$\begin{aligned}(1+x+x^2)^3(1+x+x^2+x^3+\cdots)^2 &= \left(\frac{1-x^3}{1-x}\right)^3(1+x+x^2+x^3+\cdots)^2 = \\ &= \frac{(1-x^3)^3}{(1-x)^3}\cdot\frac{1}{(1-x)^2}=(1-x^3)^3\cdot\frac{1}{(1-x)^5}=(1-x^3)^3\cdot\frac{1}{(1-x)^5}.\end{aligned}$$

Поскольку

$$(1-x^3)^3 = 1 - 3x^3 + 3x^6 - x^9$$

и

$$\frac{1}{(1-x)^5} = 1 + 5x + \binom{6}{2}x^2 + \cdots + \binom{5+n-1}{n}x^n \cdots,$$

то, согласно правилу умножения полиномов, коэффициент при x^{12} в произведении

$$(1 - 3x^3 + 3x^6 - x^9)\left(1 + 5x + \binom{6}{2}x^2 + \cdots + \binom{5+n-1}{n}x^n \cdots\right)$$

равен

$$1 \cdot \binom{5+12-1}{12} - 3\binom{5+9-1}{9} + 3\binom{5+6-1}{6} - \binom{5+3-1}{3}. \qquad \square$$

ПРИМЕР 13.12. Сколько существует способов выбора 20 объектов из множества объектов пяти типов, если количество выбранных объектов первого типа кратно 5, количество выбранных объектов второго типа кратно 3, объектов третьего типа следует выбрать не более 4, объектов четвертого типа — не менее 3, а объектов пятого типа — не более 2?

Производящая функция имеет вид

$$(1+x^5+x^{10}+\cdots)(1+x^3+x^6+\cdots)(1+x+\cdots+x^4)(x^3+x^4+\cdots)(1+x+x^2)$$

и может быть представлена в виде выражения

$$\frac{1}{1-x^5}\cdot\frac{1}{1-x^3}\cdot\frac{1-x^5}{1-x}\cdot\frac{x^3}{1-x}\cdot\frac{1-x^3}{1-x} = \frac{x^3}{(1-x)^3},$$

исходя из этого, производящая функция имеет вид

$$x^3\left(1 + 3x + \binom{4}{2}x^2 + \cdots \binom{3+n-1}{n}x^n + \cdots\right).$$

Коэффициент при x^r равен $\binom{r-1}{r-3}$, а коэффициент при x^{20} равен $\binom{19}{17}$. $\square$

ПРИМЕР 13.13. Найдем производящую функцию, n-ый коэффициент которой дает количество неотрицательных целочисленных решений уравнения $e_1 + 4e_2 + 5e_3 + 3e_4 = n$, что эквивалентно нахождению количества способов выбора n объектов, когда объекты второго типа выбираются по 4 за один раз, объекты третьего типа выбираются по 5, а объекты четвертого типа выбираются по 3 за раз. Следовательно, производящая функция имеет вид

$$(1 + x + x^2 + \cdots)(1 + x^4 + x^8 + \cdots)(1 + x^5 + x^{10} + \cdots)(1 + x^3 + x^6 + \cdots),$$

но это равно

$$\frac{1}{(1-x)(1-x^4)(1-x^5)(1-x^3)}. \qquad \square$$

■ УПРАЖНЕНИЯ

1. Найдите производящую функцию, в которой коэффициент при x^r описывает количество решений уравнения

а) $e_1 + e_2 + e_3 + e_4 = r$, где $0 \le e_2 \le 2$ и $0 \le e_4 \le 4$;

б) $e_1 + e_2 + e_3 = r$, где $0 \le e_1 \le 3$ и e_3 четно;

в) $e_1 + e_2 + e_3 + e_4 = r$, где e_1 нечетно, e_2 четно и $e_4 \ge 4$;

г) $e_1 + e_2 + e_3 + e_4 = r$, где $e_i \ge i$ для всех i;

д) $e_1 + 2e_2 + 3e_3 + 4e_4 = r$.

2. Найдите производящую функцию, r-ый коэффициент которой дает количество решений уравнения

а) $e_1 + e_2 + e_3 + e_4 = r$, где $e_1 \ge 2$ и $0 \le e_3 \le 3$;

б) $e_1 + e_2 + e_3 = r$, где $0 \le e_i \le 3$ для всех i;

в) $e_1 + e_2 + e_3 + e_4 = r$, где e_1 нечетно, e_2 и e_3 четны и $e_4 \ge 2$;

г) $e_1 + e_2 + e_3 + e_4 = r$, где $e_i \ge 3$ для всех i;

д) $e_1 + 2e_2 + e_3 + 2e_4 = r$.

3. Найдите производящую функцию, которую можно использовать для нахождения количества способов выбора r элементов из

а) 4 красных, 3 синих, 6 оранжевых и 2 зеленых шаров;

б) 2 красных, 5 зеленых, 4 оранжевых и 3 зеленых шаров, если необходимо выбрать не менее 1 шара каждого цвета;

в) 6 красных, 12 черных, 7 белых и 10 синих шаров, если необходимо выбрать не менее 4 черных шаров и четное количество синих шаров;

г) 6 красных, 12 черных, 7 белых и 10 синих шаров, если необходимо выбрать не менее 1 шара каждого типа, четное количество красных шаров и нечетное количество белых шаров.

4. Найдите производящую функцию для нахождения количества способов выбора r элементов из

а) 6 красных, 5 синих, 4 оранжевых и 3 зеленых шаров;

б) 7 красных, 5 зеленых, 8 оранжевых и 4 белых шаров, если необходимо выбрать нечетное количество зеленых и красных шаров и четное количество оранжевых и белых шаров;

в) 5 красных, 4 фиолетовых, 6 белых и 8 черных шаров, если необходимо выбрать не менее 2 черных шаров, 1 оранжевого, 3 красных и четное количество синих шаров;

г) 9 красных, 7 черных, 6 белых и 11 синих шаров, если необходимо выбрать не менее 2 шаров каждого типа, нечетное количество красных шаров и нечетное количество белых шаров.

5. Найдите производящую функцию для определения количества способов составить k центов, используя монеты стоимостью 1, 5, 10 и 25 центов.

6. В маленьком городке 50 зарегистрированных избирателей, каждый из них либо голосует дважды, либо остается дома, за исключением мэра города, который голосует либо три раза, либо пять. Используя производящую функцию, определите, сколько существует способов собрать n голосов.

7. Для всех $n \geq 8$, почтовое отправление стоимостью n центов оплачивается марками стоимостью 5 и 3 цента. Используйте производящую функцию для описания количества способов оплаты почтового отправления стоимостью k центов, используя исключительно марки по цене 5 и 3 центов.

8. Найдите производящую функцию, которая используется для нахождения количества способов разложения k различных типов шоколадок в 6 коробок, если в первую коробку помещать четное количество шоколадок, во вторую — нечетное количество, не менее 3 шоколадок помещать в третью коробку, не более 3 — в четвертую коробку и любое их количество — в оставшиеся две коробки.

9. Подбрасывание k игральных кубиков дает в сумме n. Используйте производящую функцию для определения количества способов подбрасывания.

10. Найдите коэффициент при x^7 в разложении производящей функции

$$f(x) = (x + x^2 + x^3)^5.$$

11. Найдите коэффициент при x^{10} в разложении производящей функции

$$f(x) = \frac{x^4}{(1-x)^5}.$$

12. Найдите коэффициент при x^{17} в разложении производящей функции

$$f(x) = \frac{x^4}{1-x^5}.$$

13. Найдите коэффициент при x^{19} в разложении производящей функции

$$f(x) = \frac{x^4}{1-x^5}.$$

14. Найдите коэффициент при x^{20} в разложении производящей функции

$$f(x) = (x^4 + x^8 + x^{16} + \cdots)^3.$$

15. Найдите коэффициент при x^{12} в разложении производящей функции

$$f(x) = (x^2 + x^3 + x^4 + \cdots)^4.$$

16. Найдите коэффициент при x^{10} в разложении производящей функции

$$f(x) = (x^2 + x^3 + x^4 + x^5)(x + x^5 + x^{10}).$$

17. Найдите коэффициент при x^{12} в разложении производящей функции

$$f(x) = \frac{x+3}{1-2x+x^2}.$$

18. Найдите коэффициент при x^{12} в разложении производящей функции

$$f(x) = \frac{x^3 - 3x^2}{(1-x)^4}.$$

19. Найдите коэффициент при x^{12} в разложении $f(x) = (1-3x)^{-4}$.

20. Найдите коэффициент при x^{12} в разложении производящей функции

$$f(x) = \frac{1}{(1-x^3)^4}.$$

21. Используйте производящую функцию для нахождения количества способов выбора 10 шаров из 20 красных, 20 белых и 20 синих шаров, если

а) следует выбрать не менее одного шара каждого цвета;

б) следует выбрать четное количество красных шаров и четное количество синих шаров.

22. Используя производящую функцию, определите, сколько существует способов распределения 15 игрушек между 5 детьми, если ни один ребенок не может получить более 4 игрушек?

23. Используя производящую функцию, определите, сколькими способами можно разложить в коробки шестнадцать конфет, если имеется 4 вида конфет и нужно выбирать не более 4 конфет каждого из первых трех видов.

24. Используя производящую функцию, определите, сколько имеется способов раскладывания 16 одинаковых предметов в 5 различных ящиков, если каждый ящик может содержать не менее 2 и не более 6 предметов?

25. Пусть $F(x)$ — производящая функция, коэффициенты которой являются числами Фибоначчи. Используя рекуррентное отношение для $F(x)$, покажите, что

$$F(x) = \frac{x}{(1-(x+x^2))}.$$

Используя это выражение, постройте разложение производящей функции $F(x)$. Покажите, что коэффициент при x^{n+1} равен

$$\binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots,$$

поэтому

$$\text{Фиб}(n+1) = \binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots.$$

13.4. РАЗБИЕНИЯ

В этом разделе при помощи производящих функций описывается число распределений множества, содержащего n неразличимых объектов, по заданному числу неразличимых ящиков. Это эквивалентно описанию количества способов разбиения положительного целого числа n на группы, содержащие заданное количество положительных целых чисел, сумма которых равна n; при этом группы, которые отличаются только порядком элементов, являются неразличимыми.

Если противное не оговорено, то под числом в этом разделе мы будем подразумевать положительное целое число.

Например, число 3 можно выразить как сумму $1+1+1$ и как $1+2$. Число 4 можно выразить как $1+1+1+1$, $1+1+2$, $1+3$, $2+2$ и 4. Способы выражения числа 4 в виде суммы двух чисел сводятся к $1+3$ и $2+2$. Таких способов два, поэтому существует два способа разложить 4 неразличимых объекта по двум неразличимым ящикам так, чтобы ни один ящик не был пустым. Способы выражения числа 4 в виде суммы двух или менее чисел таковы: $1+3$, $2+2$ и 4. Таких способов три, поэтому существует три способа разложить 4 неразличимых объекта по 2 неразличимым ящикам, так что некоторые из этих ящиков могут быть пустыми.

Сначала рассмотрим количество способов, благодаря которым n неразличимых объектов раскладывается по n ящикам при условии, что некоторые ящики могут быть пустыми. Это то же самое, что искать число способов разбиения числа n на группы из n или менее чисел, сумма которых равна n. Последнее представляет собой число неотрицательных целочисленных решений уравнения

$$e_1 + 2e_2 + 3e_3 + \cdots + ne_n = n.$$

Если n фиксировано, то количество решений находится как коэффициент при x^n в разложении производящей функции

$$\frac{1}{(1-x)(1-x^2)(1-x^3)\cdots(1-x^n)}.$$

Последний пример предыдущего раздела показывает, что количество решений эквивалентно нахождению количества способов выбора n объектов при условии, что объекты первого типа выбираются по одному, объекты второго типа — по два, объекты третьего типа — по три и т.д. Именно это то мы и делаем, поскольку каждая выбранная единица вносит в сумму показателей степеней 1, и эти единицы суммируются вплоть до n в x^n; каждая выбранная двойка вносит в сумму степеней 2, и эти двойки суммируются вплоть до n и т.д. По мере увеличения n требуется все больше членов в произведении. Поскольку мы хотим, чтобы наша производящая функция описывала число решений для всех n, то искомая производящая функция имеет вид

$$\frac{1}{(1-x)(1-x^2)(1-x^3)\cdots(1-x^k)\cdots},$$

в результате приходим к следующей теореме.

ТЕОРЕМА 13.14. Производящая функция, n-ый коэффициент которой равен количеству способов, которыми можно разложить n неразличимых объектов по n ящикам при условии, что некоторые ящики могут быть пустыми, или, что эквивалентно, равен количеству разбиений числа n на группы, содержащие n или менее чисел, сумма которых равна n, имеет вид

$$\frac{1}{(1-x)(1-x^2)(1-x^3)\cdots(1-x^k)\cdots}.$$

ПРИМЕР 13.15. Покажем, что любое число можно выразить следующим образом

$$2^0 + 2^1 + 2^2 + 2^3 + \cdots + 2^k.$$

Из данного выражения следует, что любое число можно записать в двоичной системе счисления, что является весьма важным моментом при использовании целых чисел в компьютерных приложениях.

(а) Убедимся сначала, что для всех n

$$(1-x)(1+x)(1+x^2)\cdots(1+x^{2^n})\cdots = \\ = (1-x^{2^n})(1+x^{2^n})(1+x^{2^{n+1}})(1+x^{2^{n+2}})\cdots .$$

Используем метод индукции по n. Вполне очевидно, что утверждение справедливо при $n=0$. Предположим, что оно выполняется для $n=k$, так что

$$(1-x)(1+x)(1+x^2)\cdots(1+x^{2^n})\cdots = (1-x^{2^k})(1+x^{2^k})(1+x^{2^{k+1}})(1+x^{2^{n+2}})\cdots ,$$

но

$$(1-x^{2^k})(1+x^{2^k}) = (1-x^{2^{k+1}}).$$

Отсюда

$$(1-x)(1+x)(1+x^2)\cdots(1+x^{2^n})\cdots = (1-x^{2^{k+1}})(1+x^{2^{k+1}})(1+x^{2^{n+2}})\cdots ,$$

поэтому утверждение справедливо для $n=k+1$ и для всех n.

(б) Покажем далее, что $(1-x)(1+x)(1+x^2)\cdots(1-x^{2^n})\cdots = 1$. Согласно части (а) для любого $k \geq 1$ в левой части равенства коэффициент при x^k равен 0. Поэтому производящая функция $1+0x+0x^2+\cdots = 1$.

(в) Согласно части (б) получаем, что

$$(1+x)(1+x^2)\cdots(1+x^{2^n})\cdots = \frac{1}{1-x} = 1+x+x^2+x^3+\cdots .$$

(г) Левая часть равенства в части (в) есть производящая функция последовательности, состоящей из степеней числа 2, где каждая степень появляется только раз. Правая часть равенства представляет собой производящую функцию множества всех положительных целых чисел. Следовательно, любое число есть сумма степеней числа 2. □

ПРИМЕР 13.16. Рассмотрим количество способов разбиения числа n на группы, содержащие n или менее различных чисел. Таким образом, каждое число может войти в заданную сумму только один раз. Например, число 4 можно представить либо как 4, либо как $1+3$. В каждом представлении числа n в виде суммы различных чисел любое число, меньшее или равное n, может входить в сумму или не входить. Поэтому производящая функция имеет вид

$$(1+x)(1+x^2)(1+x^3)\cdots(1+x^k)\cdots . \qquad \square$$

ПРИМЕР 13.17. Количество способов разбиения числа n на группы, содержащие n или менее различных чисел, равно количеству способов разбиения числа n на группы, содержащие n или менее нечетных чисел, сумма которых равна n.

Сначала найдем производящую функцию, при помощи которой сможем определить количество способов разбиения числа n на группы, содержащие n или менее нечетных чисел, сумма которых равна n. Исходя из теоремы 13.14, находим, что производящая функция имеет вид

$$\frac{1}{(1-x)(1-x^3)(1-x^5)\cdots(1-x^{2k-1})\cdots}.$$

Из предыдущего примера следует, что производящая функция, используемая для нахождения количества способов разбиения числа n на группы, содержащие n или менее различных чисел, имеет вид

$$(1+x)(1+x^2)(1+x^3)\cdots(1+x^k)\cdots.$$

Но

$$(1+x^k)=\frac{1-x^{2k}}{1-x^k},$$

так что

$$(1+x)(1+x^2)(1+x^3)\cdots(1+x^k)\cdots = \\ =\frac{1-x^2}{1-x}\cdot\frac{1-x^4}{1-x^2}\cdot\frac{1-x^6}{1-x^3}\cdot\frac{1-x^8}{1-x^4}\cdot\frac{1-x^{10}}{1-x^5}\cdots.$$

Но k-ый член в числителе сокращается с $2k$-ым членом в знаменателе, а в знаменателе остается только $(2k+1)$-ый член. Поэтому имеем

$$(1+x)(1+x^2)(1+x^3)\cdots(1+x^k)\cdots=\frac{1}{1-x}\cdot\frac{1}{1-x^3}\cdot\frac{1}{1-x^5}\cdot\frac{1}{1-x^7}\cdot\frac{1}{1-x^9}\cdots$$

что является производящей функцией, используемой для определения количества способов разбиения числа n на группы, содержащие n или менее нечетных чисел, сумма которых равна n. □

Следующая цель — описать количество способов разложения n неразличимых предметов по k неразличимым ящикам. Для этого понадобится простой, но удобный метод, состоящий в использовании так называемых ***диаграмм Ферре***. Будем использовать одну из модификаций этого метода, где простоте отдается предпочтение над строгим алгебраическим описанием. В данной версии начнем с описания распределения предметов по ящикам, используя строки возрастающей длины. Например, распределение 12 предметов по 4 ящикам можно изобразить так, как показано на рис. 13.1.

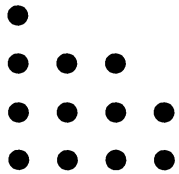


Рис. 13.1

Если судить по строкам, то видно, что в первый ящик помещен 1 предмет, во второй — 3 предмета, в третий ящик — 4 предмета и в четвертый — 4 предмета. Если же судить по столбцам, то в первом ящике находится 4 предмета, во втором — 3 предмета, в третьем — 3 предмета и в четвертом — два предмета. В столбцах ящики расположены в порядке убывания их размера. Но это несущественно, т.к. ящики неразличимы. В частности, легко заметить, что каждому распределению по строкам соответствует некоторое распределение по столбцам, и наоборот. Следовательно, существует взаимно однозначное соответствие между распределениями по строкам, содержащим ящики возрастающего размера, и распределениями по столбцам, содержащим ящики уменьшающегося размера. Рассмотрим произвольное распределение 12 предметов по 4-м ящикам. Допустим, что в первом ящике — 1 предмет, во втором — 2 предмета, 4 предмета — в третьем ящике и 6 предметов — в четвертом ящике. Соответствующая диаграмма Ферре приведена на рис. 13.2.

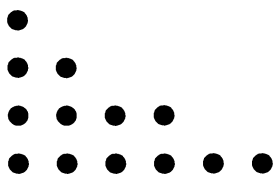


Рис. 13.2

Если теперь посмотреть на распределение предметов по столбцам, то можно увидеть, что в первом ящике находятся 4 предмета, во втором ящике — 3 предмета, в третьем и четвертом ящиках — по 2 предмета и по 1 предмету — в пятом и шестом ящиках. Однако, поскольку в распределении по строкам с самого начала использовались 4 ящика, то в распределении по столбцам нельзя поместить в каждый ящик более 4 предметов. И наоборот, если в распределении по столбцам класть в каждый ящик не более 4-х предметов, то при переходе к строкам понадобятся не более четырех ящиков. Более того, в распределении по строкам 4 непустых ящика будут только тогда, когда в распределении по столбцам будет один ящик, содержащий в точности 4 предмета. Таким образом, существует взаимно однозначное соответствие между распределениями 12 предметов по 4 ящикам, где некоторые пусты, и такими распределениями 12 предметов, когда в каждом ящике находятся не более 4 предметов. Кроме того, существует взаимно однозначное соответствие между распределениями 12 предметов по 4 ящикам, при которых в каждый ящик попадает хотя бы один предмет, и такими распределениями 12 предметов, при которых хотя бы 1 ящик содержит 4 предмета. Такие же рассуждения с n предметами и k ящиками приводят к следующей теореме.

ТЕОРЕМА 13.18. Количество способов размещения n предметов в k ящиках при условии, что некоторые ящики могут быть пустыми, равно количеству размещений n предметов по ящикам при условии, что в каждом ящике будет находится не более k предметов. Кроме того, количество способов размещения n предметов в k ящиках, где ни один ящик не может быть пустым, равно количеству размещений n предметов по ящикам при условии, что в каждом ящике может находиться не более k предметов и хотя бы один ящик содержит k предметов.

Сформулируем эквивалентную теорему в терминах разбиений числа n.

ТЕОРЕМА 13.19. Количество способов разбиения числа n на группы, содержащие k или менее чисел, равно количеству способов разбиения числа n на группы, содержащие n или менее чисел, при которых ни одно число из группы не превышает k. Кроме того, количество способов разбиения числа n на группы, содержащие ровно k чисел, равно количеству способов разбиения числа n на группы, содержащие n или менее неотрицательных целых чисел, когда наибольшее число в каждой группе равно k.

Поскольку количество способов разбиения числа n на группы, содержащие n или менее чисел при условии, что никакое число в группе не превышает k, равно количеству неотрицательных целочисленных решений уравнения

$$e_1 + 2e_2 + 3e_3 + \cdots + ke_k = n,$$

то производящая функция имеет вид

$$(1 + x + x^2 + \cdots)(1 + x^2 + x^4 + \cdots)(1 + x^3 + x^6 + \cdots)\cdots(1 + x^k + x^{2k} + \cdots)$$

или

$$\frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^3} \cdot \cdots \cdot \frac{1}{1-x^k}.$$

ТЕОРЕМА 13.20. Производящая функция, описывающая количество способов разбиения числа n на группы из k или менее неотрицательных целых чисел, имееет вид

$$(1 + x + x^2 + \cdots)(1 + x^2 + x^4 + \cdots)(1 + x^3 + x^6 + \cdots)\cdots(1 + x^k + x^{2k} + \cdots)$$

или

$$\frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^3} \cdot \cdots \cdot \frac{1}{1-x^k}.$$

Теперь сформулируем эквивалентную теорему в терминах распределения неразличимых объектов.

ТЕОРЕМА 13.21. Производящая функция, описывающая количество способов размещения n неразличимых объектов по k ящикам при условии, что некоторые из ящиков пустые, имеет вид

$$(1 + x + x^2 + \cdots)(1 + x^2 + x^4 + \cdots)(1 + x^3 + x^6 + \cdots)\cdots(1 + x^k + x^{2k} + \cdots),$$

или

$$\frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^3} \cdot \cdots \cdot \frac{1}{1-x^k}.$$

Теперь рассмотрим случай с количеством способов разбиения числа n на группы из k чисел, сумма которых равна n. По теореме 13.19 это есть количество способов разбиения числа n на группы из n или менее неотрицательных целых чисел, когда наибольшее число в каждой группе равно k. В этом случае производящая функция имеет вид

$$(1 + x + x^2 + \cdots)(1 + x^2 + x^4 + \cdots)(1 + x^3 + x^6 + \cdots)\cdots(x^k + x^{2k} + \cdots),$$

что равно

$$x^k(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)(1+x^3+x^6+\cdots)\cdots(1+x+x^{2k}+\cdots)$$

или

$$\frac{x^k}{(1-x)(1-x^2)(1-x^3)\cdots(1-x^k)}.$$

Этот результат можно получить иначе, воспользовавшись производящей функцией, описывающей количество способов разбиения числа n на группы из k или менее чисел, сумма которых равна n, и вычесть из нее производящую функцию, описывающую количество способов разбиения числа n на группы из $k-1$ или менее чисел, сумма которых равна n. В результате получим производящую функцию, где число n разбито на группы, содержащие ровно k чисел, сумма которых равна n. Исходя из вышеизложенного, имеем

$$\begin{aligned}(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)\cdots(1+x^{k-1}+x^{2(k-1)}+\cdots)(1+x^k+x^{2k}+\cdots) &= \\ = -(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)\cdots & \\ \cdots(1+x^{k-1}+x^{2(k-1)}+\cdots) &= \\ = (1+x+x^2+\cdots)(1+x^2+x^4+\cdots)\cdots & \\ \cdots(1+x^{k-1}+x^{2(k-1)}+\cdots)((1+x^k+x^{2k}+\cdots)-1) &= \\ = (1+x+x^2+\cdots)(1+x^2+x^4+\cdots)\cdots & \\ \cdots(1+x^{k-1}+x^{2(k-1)}+\cdots)((x^k+x^{2k}+\cdots). &\end{aligned}$$

Таким образом, мы получили следующую теорему.

ТЕОРЕМА 13.22. Производящая функция, описывающая количество способов разбиения числа n на группы, содержащие ровно k неотрицательных целых чисел, сумма которых равна n, имеет вид

$$x^k(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)(1+x^3+x^6+\cdots)\cdots(1+x+x^{2k}+\cdots)$$

или

$$\frac{x^k}{(1-x)(1-x^2)(1-x^3)\cdots(1-x^k)}.$$

Следующая теорема содержит равносильное утверждение.

ТЕОРЕМА 13.23. Производящая функция, описывающая количество способов размещения n неразличимых предметов в k ящиках таким образом, что никакой из ящиков не будет пустым, равна

$$x^k(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)(1+x^3+x^6+\cdots)\cdots(1+x^k+x^{2k}+\cdots)$$

или

$$\frac{x^k}{(1-x)(1-x^2)(1-x^3)\cdots(1-x^k)}.$$

■ УПРАЖНЕНИЯ

1. Найдите количество разбиений числа 7.
2. Найдите количество разбиений числа 5.
3. Найдите производящую функцию, описывающую количество способов размещения 10 неразличимых объектов в 6 неразличимых ящиках.
4. Найдите производящую функцию, описывающую количество способов разбиения числа n на различные нечетные числа.
5. Найдите производящую функцию, описывающую количество способов разбиения числа n на различные четные числа.
6. Предположим, что атлет участвует в соревнованиях по 3 видам спорта. В каждом из состязаний он может выиграть 1, 2, 3, 4 или 5 очков. Найдите производящую функцию, показывающую, сколько существует различных способов набрать n очков.
7. При помощи диаграммы Ферре покажите, что количество разбиений числа на четные части равно количеству разбиений этого числа на части так, что каждая часть присутствует четное количество раз.
8. При помощи диаграммы Ферре покажите, что количество разбиений числа n на 3 части равно количеству разбиений числа $2n$ на 3 части меньшего размера, чем n. Что можно сказать о количестве разбиений числа n на 4 части и числа $3n$ на 4 части меньшего размера, чем n? Можно ли это утверждение обобщить?
9. При помощи диаграммы Ферре покажите, что число разбиений числа n равно количеству разбиений числа $2n$ на n частей.
10. При помощи диаграммы Ферре покажите, что число разбиений числа $2n+k$ на $n+k$ частей одинаково для всех k.
11. Пусть несчастные случаи классифицированы в соответствии с днем недели, в который они произошли. Найдите производящую функцию, описывающую количество различных способов, которыми можно классифицировать n несчастных случаев.

13.5. ЭКСПОНЕНЦИАЛЬНЫЕ ПРОИЗВОДЯЩИЕ ФУНКЦИИ

В данном разделе будут использованы производящие функции вида

$$a_0 + a_1\frac{x}{1!} + a_2\frac{x^2}{2!} + a_3\frac{x^3}{3!} + \cdots + a_n\frac{x^n}{n!}.$$

Для начала положим

$$e^x = e(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots.$$

Единственное свойство этой функции, которое нам понадобится, сформулировано в следующей теореме. Ее доказательство мы предоставляем читателю.

ТЕОРЕМА 13.24. Для любого неотрицательного числа n справедливо $(e^x)^n = e^{nx}$.

Известно, что

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n,$$

поэтому при использовании обычных производящих функций $(1+x)^n$ является производящей функцией для числа сочетаний. Однако,

$$\begin{aligned}(1+x)^n &= \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n = \\ &= \frac{n!}{0!(n)!} + \frac{n!}{1!(n-1)!}x + \frac{n!}{2!(n-2)!}x^2 + \cdots + \frac{n!}{k!(n-k)!}x^k + \cdots + \frac{n!}{0!(n)!}x^n = \\ &= \frac{n!}{n!} + \frac{n!}{(n-1)!}\cdot\frac{x}{1!} + \frac{n!}{(n-2)!}\cdot\frac{x^2}{2!} + \cdots + \frac{n!}{(n-k)!}\cdot\frac{x^k}{k!} + \cdots + \frac{n!}{0!}\cdot\frac{x^n}{n!} = \\ &= P(n,0) + P(n,1)\frac{x}{1!} + P(n,2)\frac{x^2}{2!} + \cdots + P(n,k)\frac{x^k}{k!} + \cdots + P(n,n)\frac{x^n}{n!},\end{aligned}$$

так что $(1+x)^n$ является экспоненциальной производящей функцией, описывающей количество перестановок.

Итак, коэффициенты экспоненциальной производящей функции не равны количеству сочетаний, как в случае обычной производящей функции, а равны количеству перестановок. Теперь переходим к подсчетам объектов с учетом их порядка.

Предположим, что имеется неограниченное количество красных, зеленых, белых и синих предметов. Известно, что производящая функция, описывающая количество способов выбора n предметов при условии, что необходимо выбрать хотя бы 2 зеленых и 3 синих предмета, имеет вид

$$(1+x+x^2+\cdots)(x^2+x^3+x^4+\cdots)(1+x+x^2+\cdots)(x^3+x^4+x^5+\cdots).$$

Например, если выбираются 12 предметов, то $x^2x^3x^4x^3$ описывает выбор 2 красных, 3 зеленых, 4 белых и 3 синих предметов. Предположим, однако, что выбираются не красные, зеленые, белые и синие предметы, а буквы *к*, *з*, *б* и *с*, и спрашивается, сколько слов можно из них составить. Нам нужно получить не один из возможных результатов — 2*к*, 3*з*, 4*б* и 3*с*, а число

$$\frac{12!}{2!3!4!2!},$$

выражающее количество способов перегруппировки 2*к*, 3*з*, 4*б* и 3*с*. Заметим, что, если вместо использования приведенного выше произведения производящих функций использовать

$$\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\cdots\right)\left(\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\right)\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\cdots\right)\left(\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\cdots\right),$$

то вместо $x^2x^3x^4x^3$, как одного из сомножителей, дающих x^{12}, можно получить

$$\frac{x^2}{2!}\frac{x^3}{3!}\frac{x^4}{4!}\frac{x^3}{3!} = \frac{x^{12}}{2!3!4!2!} = \frac{12!}{2!3!4!2!}\cdot\frac{x^{12}}{12!}.$$

Коэффициент при $\frac{x^{12}}{12!}$ равен $\frac{12!}{2!3!4!2!}$. Таким образом, если требуется количество способов, которыми можно переупорядочить предметы, а не выбрать их, то вместо обычных производящих функций необходимо использовать экспоненциальные производящие функции.

При использовании обычных производящих функций коэффициент при x^n даст общее количество способов, которыми можно выбрать n предметов. Если использовать экспоненциальную производящую функцию, то коэффициент при $\frac{x^n}{n!}$ даст количество способов, которыми n предметов можно выбрать и переупорядочить. Напомним, что в разделе 8.6 приводилось равенство

$$P(n; r_1, r_2, \cdots, r_k) = \frac{n}{r_1! r_2! \cdots r_k!} = C(n; r_1, r_2, \cdots, r_k),$$

определяющее количество перестановок n предметов при условии, что имеется r_i предметов типа i, при этом

$$r_1 + r_2 + \cdots + r_k = n,$$

и количество разбиений множества, содержащего n неразличимых предметов, по k неразличимым ящикам, когда r_i предметов попадает в i-ый ящик, и

$$r_1 + r_2 + \cdots + r_k = n.$$

Следовательно, экспоненциальную производящую функцию можно использовать для решения задач обоих типов.

ПРИМЕР 13.25. Сколько последовательностей длины n можно сформировать из целых чисел $1, 2, 3$ и 4, если должно быть не менее одной цифры 1, двух цифр 2, нечетное количество цифр 3 и четное количество цифр 4?

Если бы мы просто выбирали объекты, производящая функция имела бы вид

$$(x + x^2 + x^3 \cdots)(x^2 + x^3 + x^4 + \cdots)(x + x^3 + x^5 \cdots)(x^2 + x^4 + x^6 + \cdots).$$

Учитывая, что требуется количество последовательностей, порядок элементов имеет значение, поэтому соответствующая экспоненциальная функция имеет вид

$$\left(\frac{x}{1!} + \frac{x^2}{2!} + \cdots\right)\left(\frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\right)\left(\frac{x}{1!} + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots\right)\left(\frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots\right).$$

□

ПРИМЕР 13.26. Найдем производящую функцию, описывающую количество способов размещения n человек в трех комнатах при условии, что в каждой комнате должно быть не менее двух и не более десяти человек. Используя основное свойство обычных экспоненциальных функций, определяем, что искомая экспоненциальная производящая функция имеет вид

$$f(x) = \left(\frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{10}}{10!}\right)\left(\frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{10}}{10!}\right)\left(\frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{10}}{10!}\right)$$

или

$$f(x) = \left(\frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{10}}{10!}\right)^3. \qquad \square$$

ПРИМЕР 13.27. Из объектов n типов требуется найти количество различных перестановок k объектов при условии, что имеется неограниченное количество объектов каждого типа. Из вышеизложенного исследования перестановок с неограниченным числом повторений известно, что ответом является число n^k (см. раздел 8.7). Попытаемся применить экспоненциальную производящую функцию. В данном случае экспоненциальная производящая функция имеет вид

$$f(x) = \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\right)^n.$$

Но

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots,$$

так что

$$(e^x)^n = \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\right)^n,$$

и

$$(e^x)^n = e^{nx} = \left(1 + \frac{nx}{1!} + \frac{(nx)^2}{2!} + \frac{(nx)^3}{3!} + \cdots\right),$$

так что коэффициент при $\dfrac{x^r}{k!}$, как и ожидалось, равен n^k. $\square$

Доказательство следующей теоремы предоставляем читателю.

ТЕОРЕМА 13.28.

а) $e^x - 1 = \dfrac{x}{1!} + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \cdots;$

б) $\dfrac{e^x + e^{-x}}{2} = 1 + \dfrac{x^2}{2!} + \dfrac{x^4}{4!} + \dfrac{x^6}{6!} + \cdots;$

в) $\dfrac{e^x - e^{-x}}{2} = \dfrac{x}{1!} + \dfrac{x^3}{3!} + \dfrac{x^5}{5!} + \dfrac{x^7}{7!} + \cdots.$

ПРИМЕР 13.29. Найдем экспоненциальную функцию для описания количества способов размещения n гостей в трех комнатах при условии, что в первой комнате находится хотя бы один гость, нечетное количество гостей находится во второй комнате и четное количество — в третьей комнате. Производящая функция задается выражением

$$f(x) = \left(\frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\right)\left(\frac{x}{1!} + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots\right)\left(\frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots\right) =$$

$$= (e^x - 1)\frac{e^x - e^{-x}}{2} \cdot \frac{e^x + e^{-x}}{2} = (e^x - 1)\frac{e^{2x} - e^{-2x}}{4} =$$

$$= \frac{1}{4}(e^{3x} - e^{-x} - e^{2x} + e^{-2x}),$$

поэтому коэффициент при $\dfrac{x^n}{n!}$ равен $\dfrac{1}{4}((3)^n - (-1)^n - (2)^n + (-2)^n)$. $\square$

Теперь имеются средства для нахождения количества способов размещения n различимых объектов в k различимых ящиках при условии, что ни один из ящиков не будет пустым, и количество способов размещения n различимых объектов в k неразличимых ящиках, когда ни один из ящиков не будет пустым.

ТЕОРЕМА 13.30. Количество способов размещения n различимых объектов в k различимых ящиках при условии, что ни один из ящиков не будет пустым и $1 \le k \le n$, равно

$$T_k^{(n)} = \sum_{i=1}^{k} (-1)^i \binom{k}{i} (k-i)^n.$$

ДОКАЗАТЕЛЬСТВО. Экспоненциальная производящая функция, описывающая количество размещений n различимых объектов в k различимых ящиках при условии, что ни один из ящиков не будет пустым, имеет вид

$$f(x) = \left(\frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \right)^k.$$

Но

$$\left(\frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \right)^k = (e^x - 1)^k = \qquad \text{по теореме 13.28}$$

$$= \sum_{i=1}^{k} \binom{k}{i} (-1)^i e^{(k-i)x} \qquad \text{по формуле бинома Ньютона.}$$

По определению,

$$e^{(k-i)x} = 1 + \frac{(k-i)x}{1!} + \frac{(k-i)^2x^2}{2!} + \frac{(k-i)^3x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{(k-i)^n x^n}{n!},$$

так что

$$f(x) = \sum_{i=1}^{k} \binom{k}{i} (-1)^k \sum_{n=0}^{\infty} \frac{(k-i)^n x^n}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{i=1}^{k} \binom{k}{i} (-1)^k (k-i)^n,$$

и коэффициент при $\dfrac{x^n}{n!}$ равен

$$T_k^{(n)} = \sum_{i=1}^{k} \binom{k}{i} (-1)^i (k-i)^n. \qquad \blacksquare$$

Теперь мы готовы доказать следующую теорему.

ТЕОРЕМА 13.31. Количество способов размещения n различимых объектов в k неразличимых ящиках при условии, что ни один из ящиков не будет пустым и $1 \le k \le n$, равно

$$S_k^{(n)} = \frac{1}{k!} \sum_{i=1}^{k} (-1)^i \binom{k}{i} (k-i)^n,$$

где $\{S_k^{(n)} : 0 \le k \le n\}$ — множество ***чисел Стирлинга второго рода***.

ДОКАЗАТЕЛЬСТВО. Поскольку $T_k^{(n)}$ равно количеству способов размещения n неразличимых объектов в k различимых ящиков, то при переходе от множества с различимыми объектами ко множеству с неразличимыми объектами приходится делить на $k!$, т.к. все $k!$ перестановок k ящиков тождественны. Следовательно,

$$S_k^{(n)} = \frac{1}{k!} T_k^{(n)} = \frac{1}{k!} \sum_{i=1}^{k} (-1)^i \binom{k}{i} (k-i)^n. \qquad \blacksquare$$

■ УПРАЖНЕНИЯ

1. Найдите разложение соответствующей производящей функции, чтобы определить, сколько кодов из десяти цифр в каждом можно образовать из цифр 0, 1 и 2, если
   а) цифра 0 должна встречаться в коде хотя бы дважды;
   б) цифра 1 может встречаться не более трех раз, а цифра 2 — не более четырех раз;
   в) цифра 1 может встречаться четное, а цифра 0 — нечетное число раз.
2. Двенадцать воздушных шаров должны быть привязаны к 12 фонарным столбам по маршруту парада. Найдите разложение соответствующей производящей функции для определения количества способов развешивания красных, синих, желтых и зеленых шаров при условии, что
   а) использован хотя бы один шар каждого цвета;
   б) использовано не более 4 красных, 3 желтых, 7 зеленых и 6 синих шаров;
   в) использовано не более 2 красных шаров, не более 3 зеленых шаров, от 1 до 5 зеленых шаров и четное количество синих шаров.
3. Найдите производящую функцию последовательности факториалов $0!, 1!, 2!, 3!\ldots$.
4. Найдите производящую функцию, описывающую количество способов раскрашивания 8×8-клеточной шахматной доски, используя красный, белый и синий цвета, при условии, что четное количество клеток — красные.
5. Используя производящую функцию, определите количество чисел, состоящих из n цифр, все цифры в которых нечетные, а цифры 1 и 3 присутствуют в них не менее одного раза.
6. Воспользуйтесь производящей функцией для нахождения количества чисел, состоящих из n цифр, все цифры которых нечетные, при условии, что все цифры присутствуют хотя бы один раз, а цифра 5 появляется четное число раз.
7. Воспользуйтесь производящей функцией для нахождения количества способов размещения 30 человек в 3 комнатах гостиницы, если в каждой комнате находится хотя бы по одному человеку.
8. Воспользуйтесь производящей функцией для нахождения количества способов размещения 30 человек в 3 комнатах гостиницы, если в каждой комнате находится четное количество людей.

9. Используйте производящую функцию для определения количества способов размещения 25 объектов четырех типов по порядку, если количество объектов первого типа — четное, количество объектов второго типа — нечетное, и хотя бы по одному объекту третьего и четвертого типа.

10. Используя производящую функцию, определите, сколько существует способов распределения 9 кусков торта между 4 детьми при условии, что каждый ребенок из первых двух получает кусок торта.

11. Найдите производящую функцию для вычисления количества упорядочений монет стоимостью 1 цент, 5 центов, 10 центов и 25 центов, которые в сумме дадут n центов.

12. Покажите, что $\frac{e^n}{(1-x)^n}$ — экспоненциальная производящая функция, описывающая количество способов выбора некоторого, возможно пустого, подмножества из k объектов. Распределите их по n ящикам, учитывая порядок в подмножестве.

13. Найдите экспоненциальную производящую функцию, описывающую количество способов разбиения числа n на k частей, учитывая порядок.

14. Используя метод индукции, докажите теорему 13.24. $e^{nx} = (e^x)^n$ для любого неотрицательного целого числа n.

15. Докажите теорему 13.28.

а) $e^x - 1 = \dfrac{x}{1!} + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \cdots;$

б) $\dfrac{e^x + e^{-x}}{2} = 1 + \dfrac{x^2}{2!} + \dfrac{x^4}{4!} + \dfrac{x^6}{6!} + \cdots;$

в) $\dfrac{e^x - e^{-x}}{2} = \dfrac{x}{1!} + \dfrac{x^3}{3!} + \dfrac{x^5}{5!}\dfrac{x^7}{7!} + \cdots.$

ГЛАВА

14 НЕКОТОРЫЕ СПЕЦИАЛЬНЫЕ ВОПРОСЫ ТЕОРИИ ГРАФОВ

14.1. АЛГЕБРАИЧЕСКИЕ СВОЙСТВА ГРАФОВ

В этом разделе рассматриваются отношения между графами. Если два графа находятся в данном отношении, то свойства, которыми обладает один граф, присущи и другому графу. Материал раздела представлен так, что большая часть понятий может быть обобщена на случай мультиграфов и псевдографов.

ОПРЕДЕЛЕНИЕ 14.1. Функция f из графа $G(V, E)$ в граф $G'(V', E')$ называется ***гомоморфизмом*** из G в G' и обозначается $f : G \to G'$, если обладает следующими свойствами.

а) Если $e \in E$, то $f(e) \in E'$. $(f(E) \subseteq E')$.

б) Если $v \in V$, то $f(v) \in V'$. $(f(V) \subseteq V')$.

в) Если вершины u и v инцидентны ребру e графа G, то вершины $f(u)$ и $f(v)$ инцидентны ребру $f(e)$ графа G'.

Доказательство приведенных ниже теорем предоставляется читателю.

ТЕОРЕМА 14.2. Если функция f — гомоморфизм из G в G', то $f(G)$ — подграф $(f(V), f(E))$ графа G'.

ТЕОРЕМА 14.3. Если граф G связный и f — гомоморфизм, то граф $f(G)$ связный.

ТЕОРЕМА 14.4. Если граф G полный и f — гомоморфизм, то $f(G)$ полный.

Из примеров увидим, что многие свойства графа G не являются инвариантными относительно f.

ОПРЕДЕЛЕНИЕ 14.5. Гомоморфизм $f : G \to G'$ является ***изоморфизмом***, если $f : V \to V'$ и $f : E \to E'$ представляют собой взаимно однозначные соответствия. Если $f : G \to G'$ — изоморфизм, то G и G' называются ***изоморфными***.

Таким образом, изоморфизм, по сути, является переименованием вершин и ребер графа V, которое сохраняет свойство гомоморфности, так что если вершины u и v инцидентны ребру e графа G, то вершины $f(u)$ и $f(v)$ инцидентны ребру $f(e)$ графа G'. Инварианты изоморфизмов приведены в задачах.

Вполне очевидно, что практически все свойства графов инвариантны относительно изоморфизма. Отсюда следует, что простейший способ показать неизоморфность двух графов — установить свойство, которым обладает один граф и не обладает другой.

ОПРЕДЕЛЕНИЕ 14.6. Если граф $G(V, E)$ содержит ребро $e = \{v_1, v_2\}$ и граф $G'(V', E')$ получен из графа $G(V, E)$ добавлением новой вершины v в множество V и заменой ребра $\{v_1, v_2\}$ ребрами $\{v_1, v\}$ и $\{v, v_2\}$, то граф $G'(V', E')$ называется ***расширением*** графа $G(V, E)$. Если графы $G_1, G_2, G_3, \ldots G_n$ таковы, что G_{i+1} является расширением графа G_i, то граф G_n назовем ***производным*** от графа G_1.

Если граф $G'(V', E')$ — расширение графа $G(V, E)$, то посредине одного из ребер множества V появляется вершина, а исходное ребро делится на два новых ребра, которые соединяют вершины, инцидентные исходному ребру, и новую вершину.

ОПРЕДЕЛЕНИЕ 14.7. Графы G и G' называются ***гомеоморфными***, если существует граф G'' такой, что оба графа, G и G', являются производными от графа G''.

ПРИМЕР 14.8. Граф, изображенный на рис. 14.1, является расширением графа, изображенного на рис. 14.2.

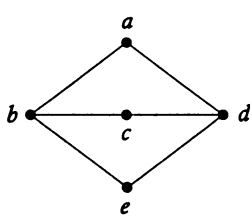


Рис. 14.1

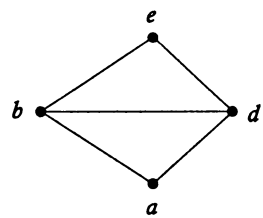


Рис. 14.2

□

ПРИМЕР 14.9. Граф, изображенный на рис. 14.3, является производным от графа, изображенного на рис. 14.4.

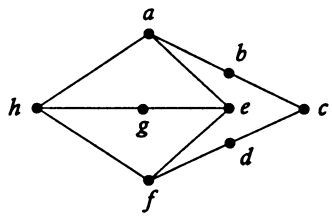


Рис. 14.3

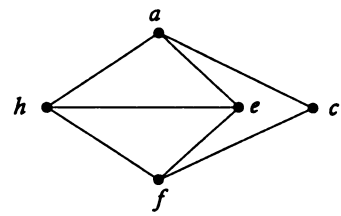


Рис. 14.4 □

ПРИМЕР 14.10. Граф, изображенный на рис. 14.5, является производным от графа, изображенного на рис. 14.6.

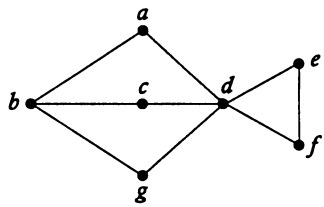


Рис. 14.5

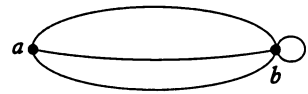


Рис. 14.6 □

ПРИМЕР 14.11. Граф, изображенный на рис. 14.7, является гомеоморфным графу, изображенному на рис. 14.8.

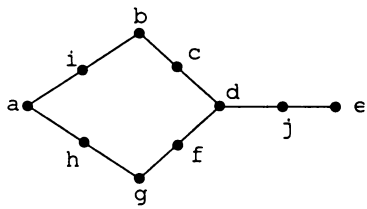


Рис. 14.7

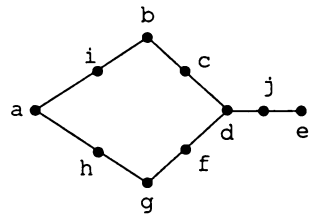


Рис. 14.8 □

Приведенные ниже определения введены, главным образом, для рассмотрения планарных графов в следующем разделе. Они демонстрируют, каким образом свойства графов остаются инвариантными для отношений между графами.

Если граф $G'(V', E')$ — расширение графа $G(V, E)$, то новая добавленная вершина имеет степень 2. Степени других вершин не изменились. Отсюда вытекает следующая теорема.

ТЕОРЕМА 14.12. Если графы G и G' — гомеоморфны, то у них одинаковое количество вершин нечетной степени.

Из данной теоремы и теорем 6.45 и 6.49 непосредственно вытекает следующая теорема.

ТЕОРЕМА 14.13. Если графы G и G' гомеоморфны, то граф G имеет эйлеров цикл (собственный путь) тогда и только тогда, когда граф G' имеет эйлеров цикл (собственный путь).

Если G' — подграф графа G, то это обозначается как $G' \preceq G$.

ОПРЕДЕЛЕНИЕ 14.14. Пусть $G = G(V, E)$ — граф и $G_1, G_2, G_3, \ldots, G_n$ — подграфы графа G. Подграф G' графа G называется ***объединением*** графов $G_1, G_2, G_3, \ldots, G_n$ и обозначается $\bigcup_{i=1}^{n} G_i$, если

1. Вершина $v \in G'$ тогда и только тогда, когда $v \in G_i$ для некоторого $1 \leq i \leq n$.
2. Ребро $e \in G'$ тогда и только тогда, когда $e \in G_i$ для некоторого $1 \leq i \leq n$.

ОПРЕДЕЛЕНИЕ 14.15. Пусть $G = G(V, E)$ — граф и $G_1, G_2, G_3, \ldots, G_n$ — подграфы графа G. Подграф G' графа G называется ***пересечением*** графов $G_1, G_2, G_3, \ldots, G_n$ и обозначается $\bigcap_{i=1}^{n} G_i$, если

1. Вершина $v \in G'$ тогда и только тогда, когда $v \in G_i$ для всех $1 \leq i \leq n$.
2. Ребро $e \in G'$ тогда и только тогда, когда $e \in G_i$ для всех $1 \leq i \leq n$.

ОПРЕДЕЛЕНИЕ 14.16. Пусть $G = G(V, E)$ — граф и $G_1, G_2, G_3, \ldots, G_n$ подграфы графа G. Подграфы $G_1, G_2, G_3, \ldots, G_n$ являются ***попарно непересекающимися***, если $G_i \cap G_j = \varnothing$ для всех $1 \leq i < j \leq n$.

Доказательство приведенных ниже теорем предоставляется читателю.

ТЕОРЕМА 14.17. Если G_1 и G_2 — различные компоненты графа G, то G_1 и G_2 — попарно непересекающиеся.

ТЕОРЕМА 14.18. Граф G является объединением попарно непересекающихся компонент.

ОПРЕДЕЛЕНИЕ 14.19. Пусть $G(V, E)$ — граф. ***Дополнением*** графа G, обозначаемым $G^c(V, E')$, называется граф такой, что для всех вершин $u, v \in V$ ребро между u и v в графе G^c существует тогда и только тогда, когда в графе G отсутствует ребро, соединяющее u и v.

ОПРЕДЕЛЕНИЕ 14.20. Подграф $G'(V', E')$ является ***остовным графом*** для графа $G = (V, E)$, если $V' = V$.

ПРИМЕР 14.21. Для заданного графа, изображенного на рис. 14.9,

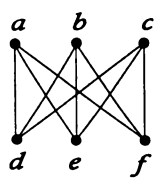


Рис. 14.9

графы, изображенные на рис. 14.10–14.12, — остовные графы.

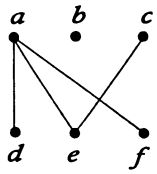


Рис. 14.10

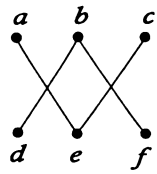


Рис. 14.11

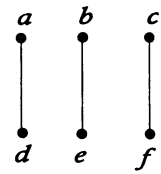


Рис. 14.12

В то же время графы, изображенные на рис. 14.13–14.15, остовными не являются.

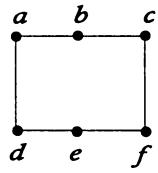


Рис. 14.13

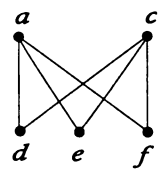


Рис. 14.14

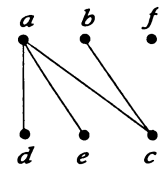


Рис. 14.15 □

Определение остовного дерева дано в главе 6 и приводится здесь для удобства изложения. В главе 6 было показано также, что каждый связный граф имеет остовное дерево.

ОПРЕДЕЛЕНИЕ 14.22. Дерево называется ***остовным деревом*** графа G, если оно является остовным графом графа G.

Доказательство приведенных ниже теорем предоставляется читателю.

ТЕОРЕМА 14.23. Если $T(V, E')$ — остовное дерево графа $G = (V, E)$, то для любого цикла $v_0v_1v_2v_3v_4, \ldots v_0$, по крайней мере, одно из ребер принадлежит $E - E'$.

Предположим, что компания создает коммуникационную сеть (рис. 14.16).

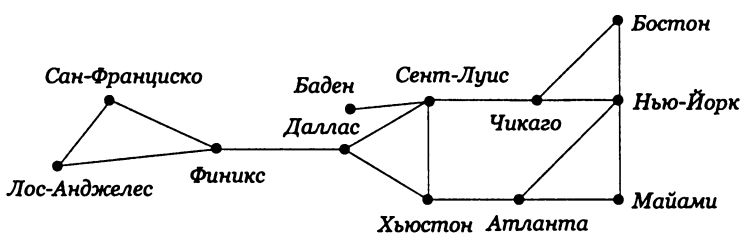


Рис. 14.16

Для компании важно знать, сколько городов окажутся не связанными между собой, если удалить некоторые линии. В частности, сколько линий требуется удалить, чтобы города оказались не связанными.

ОПРЕДЕЛЕНИЕ 14.24. Множество ребер C связного графа $G = (V, E)$ называется ***разрезающим множеством***, если удаление ребер из множества C нарушает связность графа, а удаление собственного подмножества множества C оставляет граф связным. Если множество C состоит из одного ребра, то это ребро называется ***разрезающим ребром***.

Иной способ описания разрезающего множества — это определение его как минимального множества ребер, удаление которых нарушает связность графа.

ПРИМЕР 14.25. Для графа, изображенного на рис. 14.17, e_5 и e_6 — разрезающие ребра.

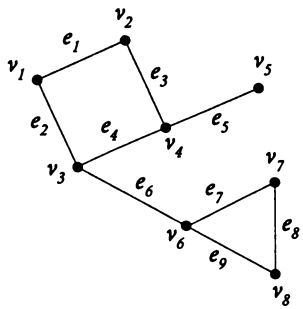


Рис. 14.17 □

ПРИМЕР 14.26. Для графа, изображенного на рис. 14.18, $\{\{v_1, v_2\}, \{v_5 v_6\}\}$ и $\{\{v_2, v_3\}, \{v_6, v_7\}\}$ — разрезающие множества.

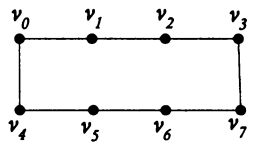


Рис. 14.18 □

Доказательства приведенных ниже теорем предоставляются читателю.

ТЕОРЕМА 14.27. Если $T(V, E')$ — остовное дерево графа $G = (V, E)$ и C — разрезающее множество графа G, то $C \cap E' \neq \varnothing$.

ТЕОРЕМА 14.28. Ребро e графа G является разрезающим ребром графа G тогда и только тогда, когда оно не входит в цикл графа G.

Другая важная задача: сколько городов лишится связи, если коммуникационная сеть выйдет из строя в определенном городе (города представлены вершинами). Вопрос сводится к следующему: "Что произойдет, если удалить вершину графа"?

ОПРЕДЕЛЕНИЕ 14.29. Вершина $a \in V$ связного графа $G = (V, E)$ является ***разрезающей вершиной***, или ***точкой сочленения***, если удаление этой вершины и инцидентных ей ребер приводит к нарушению связности графа.

ОПРЕДЕЛЕНИЕ 14.30. Граф $G(V, E)$ называется ***двусвязным***, если не содержит точек сочленения.

ТЕОРЕМА 14.31. Вершина a графа $G = (V, E)$ является точкой сочленения тогда и только тогда, когда существуют различные вершины v и w такие, что каждый путь из v в w проходит через a.

ДОКАЗАТЕЛЬСТВО. Сначала докажем достаточность. Предположим, что каждый путь из вершины v в w проходит через вершину a. Тогда, если вершина a удалена, то не существует более пути из вершины v в w, и граф G становится несвязным. Следовательно, вершина a — точка сочленения.

Для доказательства необходимости в утверждении теоремы используется метод доказательства от противного. Предположим, что для каждой пары различных вершин v и w существует путь, который не проходит через a. В этом случае при удалении вершины a для всех $v, w \in V$ всегда остается путь из вершины v в w, и граф G остается связным. Следовательно, вершина a не является точкой сочленения. ■

ПРИМЕР 14.32. В графе, изображенном на рис. 14.19, вершины v_3, v_4 и v_6 — разрезающие вершины.

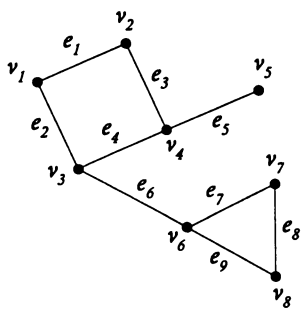


Рис. 14.19 □

ТЕОРЕМА 14.33. Для связного графа $G = (V, E)$ определим отношение R на E: $e_1 R e_2$, если $e_1 = e_2$ или в графе G существует простой цикл, содержащий e_1 и e_2 в качестве ребер. Тогда отношение R является отношением эквивалентности.

ДОКАЗАТЕЛЬСТВО. Свойства рефлексивности и симметричности очевидны. Предположим, что $e_1 R e_2$ и $e_2 R e_3$. Если $e_1 = e_2$ или $e_2 = e_3$, то результат очевиден. В противном случае существует цикл, содержащий в качестве ребер e_1 и e_2, и цикл, содержащий в качестве ребер e_2 и e_3. Если оба цикла содержат e_1 и e_3, доказательство завершено. Если нет, то существуют различные циклы, содержащие ребро e_2 и, возможно, другие ребра вместе с ним. Пусть $v_0 v_1 v_2 v_3 \cdots v_m v_0$ — цикл, содержащий e_1, и $v'_0 v'_1 v'_2 v'_3 \cdots v'_n v'_0$ — цикл, содержащий e_3. Пусть $v_i v_{i+1} v_{i+2} \cdots v_j$ и $v'_k v'_{k+1} v'_{k+2} \cdots v'_l$ — пути максимальной длины в циклах $v_0 v_1 v_2 v_3 \cdots v_m v_0$ и $v'_0 v'_1 v'_2 v'_3 \cdots v'_n v'_0$, соответственно такие, что e_1 — ребро пути $v_i v_{i+1} v_{i+2} \cdots v_j$, а e_3 — ребро пути $v'_k v'_{k+1} v'_{k+2} \cdots v'_l$, причем $v_i = v'_k$ и $v_j = v'_l$, но пути не имеют других общих точек. Значит путь $v_i v_{i+1} v_{i+2} \cdots v_j v'_{l-1} v'_{l-2} \cdots v'_k$ является простым циклом, содержащим ребра e_1 и e_3. Следовательно, отношение R транзитивно. ■

ОПРЕДЕЛЕНИЕ 14.34. Пусть для каждого класса эквивалентности E_i и отношения эквивалентности R V_i — множество вершин, инцидентных ребрам из множества E_i, и $G_i(V_i, E_i)$ — подграф графа $G(V, E)$ с вершинами V_i и ребрами E_i. Подграф $G_i(V_i, E_i)$ называется ***компонентой двусвязности*** графа $G(V, E)$.

ТЕОРЕМА 14.35. Если (a, b) и (c, d) — различные ребра из компоненты двусвязности $G_i(V_i, E_i)$, то в подграфе $G_i(V_i, E_i)$ существует простой цикл, содержащий в качестве ребер (a, b) и (c, d).

ДОКАЗАТЕЛЬСТВО. Если (a, b) и (c, d) — различные ребра из компоненты двусвязности $G_i(V_i, E_i)$, то в графе $G(V, E)$ существует простой цикл, содержащий в качестве ребер (a, b) и (c, d). Но если существует ребро (c, d) из этого цикла, не принадлежащее классу эквивалентности E_i, то существуют два ребра из разных классов эквивалентности, которые входят в один простой цикл, что противоречит принципу построения классов эквивалентности. ■

Доказательство приведенных ниже теорем предоставляется читателю.

ТЕОРЕМА 14.36. Если компонента двусвязности $G_i(V_i, E_i)$ состоит из единственного ребра e_i, то e_i — разрезающее ребро графа G.

ТЕОРЕМА 14.37. Если каждые два различных ребра входят в общий простой цикл графа $G(V, E)$, то граф $G(V, E)$ — двусвязный.

ДОКАЗАТЕЛЬСТВО. Предположим, что каждые два различных ребра входят в один и тот же цикл. Если $G(V, E)$ содержит единственное ребро e_i, то результат очевиден. В противном случае, пусть u и v — вершины графа G и a — вершина на пути из вершины u в вершину v. Пусть $\{c, a\}$ и $\{a, b\}$ — различные ребра из этого пути. Тогда $\{c, a\}$и $\{a, b\}$ — ребра, входящие в один и тот же простой цикл, поэтому существует путь из b в c, который не проходит через вершину a. Согласно теореме 14.28 вершина a не может быть точкой сочленения. Следовательно, граф $G(V, E)$ двусвязный. ■

СЛЕДСТВИЕ 14.38. Подграф $G_i(V_i, E_i)$ — двусвязный.

ТЕОРЕМА 14.39. Если $G_i(V_i, E_i)$ и $G_j(V_j, E_j)$ — компоненты двусвязности графа $G(V, E)$, то $V_i \cap V_j$ содержит не более одной вершины.

ДОКАЗАТЕЛЬСТВО. Если $V_i \cap V_j$ содержит две вершины a и b, то существуют два ребра $\{a, c\}$ и $\{b, d\}$, принадлежащие классу эквивалентности E_i. По определению E_i, в $G_i(V_i, E_i)$ существует простой цикл, содержащий $\{a, c\}$ и $\{b, d\}$. Следовательно, в графе $G_i(V_i, E_i)$ существует путь из вершины a в b. Аналогично, в графе $G_j(V_j, E_j)$ существует путь из b в a. Но поскольку графы $G_i(V_i, E_i)$ и $G_j(V_j, E_j)$ не имеют общих ребер, то путь из вершины a в b в графе $G_i(V_i, E_i)$, за которым следует путь из вершины b в a в графе $G_j(V_j, E_j)$, является циклом, который содержит простой цикл. Но тогда в общем простом цикле существуют ребра из графов $G_i(V_i, E_i)$ и $G_j(V_j, E_j)$, что противоречит принципу построения графа $G_i(V_i, E_i)$ и графа $G_j(V_j, E_j)$. ■

ТЕОРЕМА 14.40. Вершина a является точкой сочленения тогда и только тогда, когда для некоторого $i \neq j$ эта вершина принадлежит $V_i \cap V_j$ для компонент двусвязности $G_i(V_i, E_i)$ и $G_j(V_j, E_j)$

ДОКАЗАТЕЛЬСТВО. Если a — точка сочленения, то существуют ребра $\{a, b\}$ и $\{a, c\}$ такие, что не существует пути из вершины b в c, не проходящего через a. (См. доказательство теоремы 14.37.) Следовательно, ребра $\{a, b\}$ и $\{a, c\}$ не входят в один и тот же простой цикл и, значит, принадлежат различным компонентам двусвязности, например, $G_i(V_i, E_i)$ и $G_j(V_j, E_j)$. Поэтому $a \in V_i \cap V_j$. Обратно, если $a \in V_i \cap V_j$, то существует ребро $\{a, b\}$ из класса эквивалентности E_i и ребро $\{a, c\}$ из класса еквивалентности E_j. Но при этом не существует другого пути из b в c, кроме как через a. Иначе $\{a, b\}$ и $\{a, c\}$ были бы ребрами одного и того же простого цикла, что противоречит построению множеств E_i и E_j. Следовательно, a — точка сочленения. ■

ТЕОРЕМА 14.41. Граф $G(V, E)$ является двусвязным тогда и только тогда, когда любые два различных ребра входят в один и тот же простой цикл графа $G(V, E)$.

ДОКАЗАТЕЛЬСТВО. Половина утверждения теоремы является переформулировкой теоремы 14.37. Предположим, что в графе $G(V, E)$ имеются два различных ребра, которые не входят в общий простой цикл графа $G(V, E)$. Тогда они принадлежат различным компонентам двусвязности. Следовательно, соединяющий их путь должен проходить через вершину, которая, согласно предыдущей теореме, является точкой сочленения, общей для этих компонент двусвязности. Поэтому граф $G(V, E)$ не является двусвязным. ■

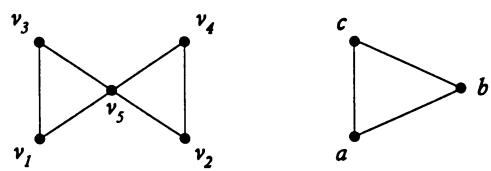
ТЕОРЕМА 14.42. Если $G_i(V_i, E_i)$ — компонента двусвязности графа $G(V, E)$ и $G_i(V_i, E_i) \neq G(V, E)$, то существует, по крайней мере, одна несовпадающая компонента двусвязности $G_j(V_j, E_j)$ такая, что $V_i \cap V_j$ содержит ровно одну вершину.

ДОКАЗАТЕЛЬСТВО. Согласно теореме 14.39, для любых двух различных компонент двусвязности V_i и V_j, пересечение множеств $V_i \cap V_j$ содержит не более одной вершины. Однако, если $a \in V_i$ и $b \notin V_i$, существует путь из вершины a в b, и последняя вершина v по ходу пути, принадлежащая V_i, является элементом пересечения множеств V_i и V_j для некоторого j. ■

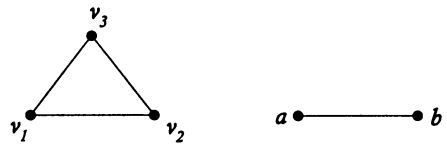
■ УПРАЖНЕНИЯ

1. Для каждой приведенной ниже пары графов найдите гомоморфизм из первого графа во второй, если он существует.

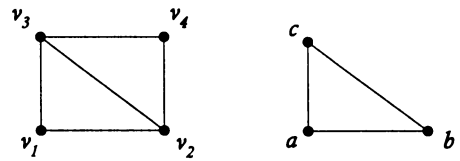
   а)



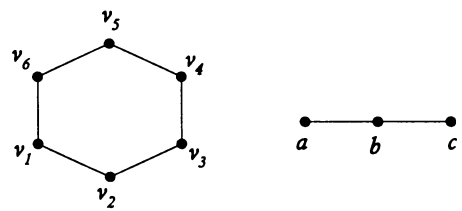
б)


v_3
v_1
v_2
a
b

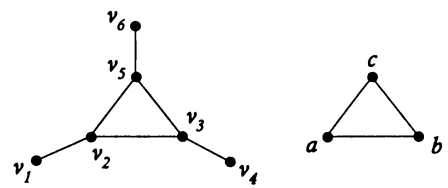
в)


v_3
v_4
v_1
v_2
c
a
b

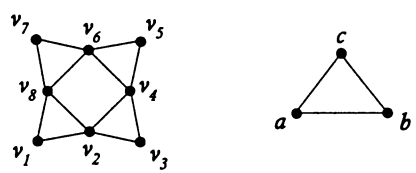
г)


v_5
v_6
v_4
v_1
v_3
v_2
a
b
c

д)

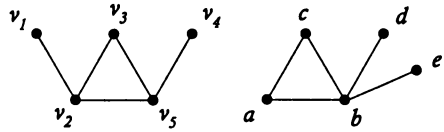

v_6
v_5
v_3
v_1
v_2
v_4
c
a
b

е)

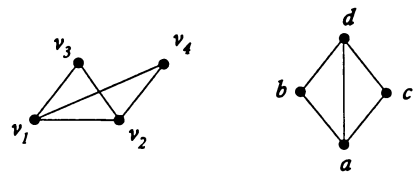

v_7
v_6
v_5
v_8
v_4
v_1
v_2
v_3
c
a
b

2. Для каждой приведенной ниже пары графов опишите изоморфизм или покажите, что вследствие нарушения инвариантности графы не изоморфны.

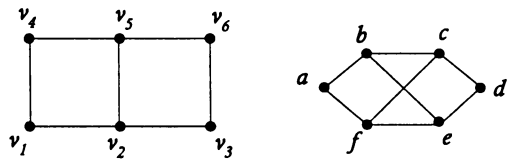
а)



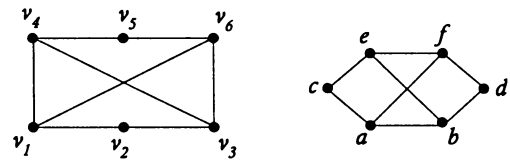
б)



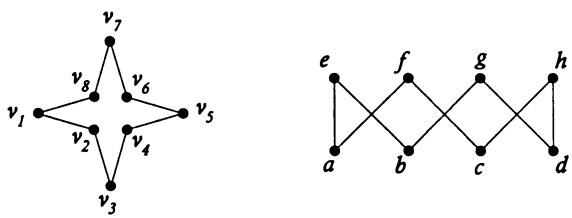
в)



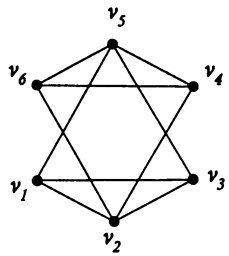
г)

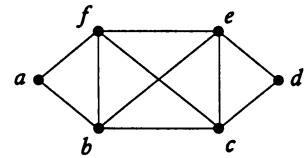


д)



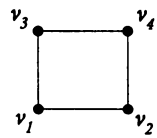
е)

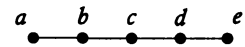




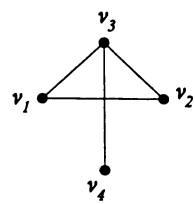
3. Для каждой приведенной ниже пары графов опишите изоморфизм или покажите, что вследствие нарушения инвариантности графы не изоморфны.

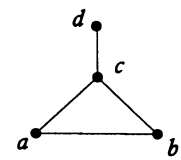
а)



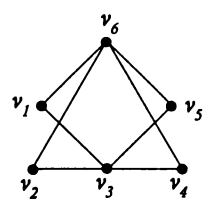


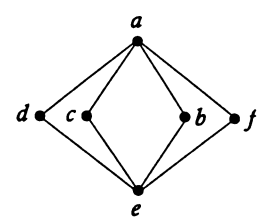
б)



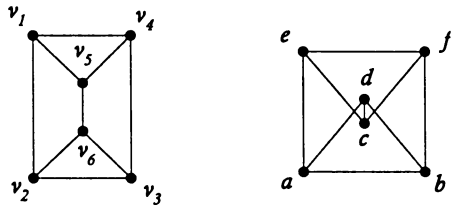


в)

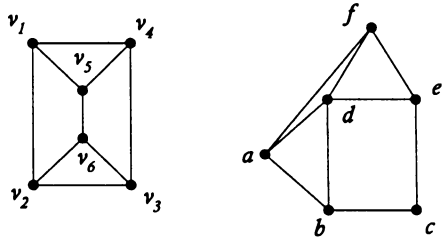




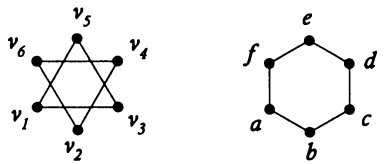
г)



д)

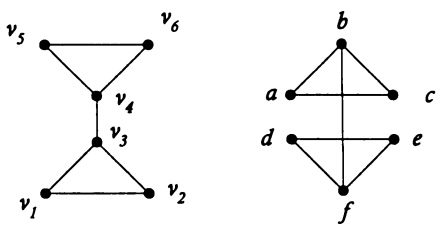


е)

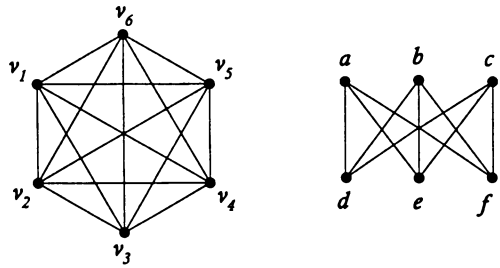


4. Для каждой приведенной ниже пары графов опишите изоморфизм или покажите, что вследствие нарушения инвариантности графы не изоморфны.

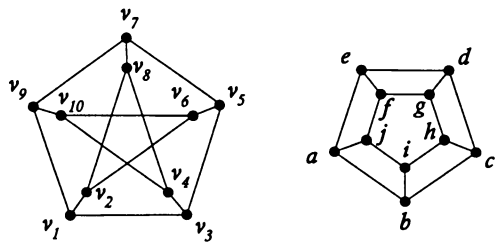
а)



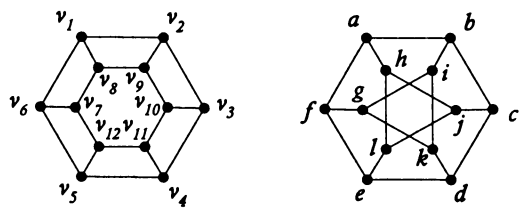
б)


v6
v1
v5
v2
v4
v3
a
b
c
d
e
f

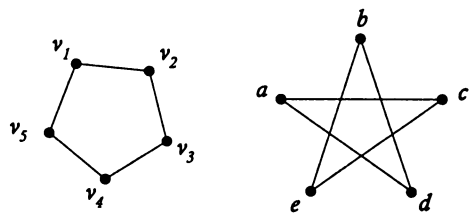
в)


v7
v8
v9
v10
v6
v5
v4
v2
v1
v3
e
d
f
g
j
h
i
a
c
b

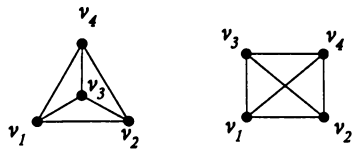
г)


v1
v2
v8
v9
v6
v7
v10
v3
v12
v11
v5
v4
a
b
h
i
g
f
j
c
l
k
e
d

д)


v1
v2
v5
v3
v4
b
a
c
e
d

е)



5. Какой из приведенных ниже графов является производным от графа, изображенного на рис. 14.20?

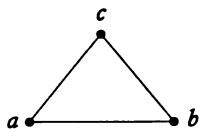
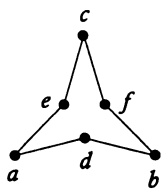
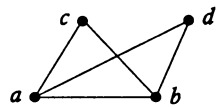


Рис. 14.20

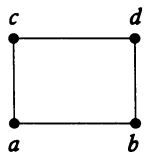
а)



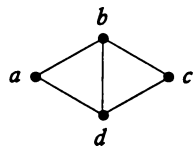
б)



в)



г)



6. Какой из приведенных ниже графов является производным от графа, изображенного на рис. 14.21?

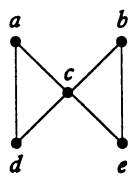
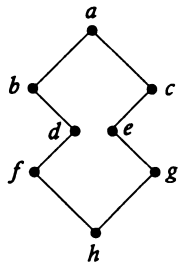
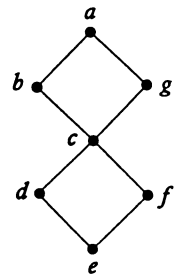


Рис. 14.21

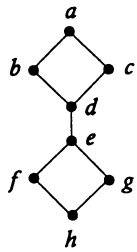
а)



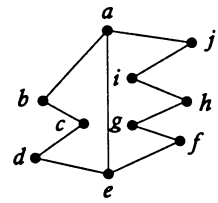
б)



в)

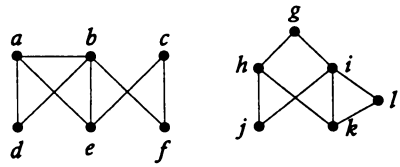


г)

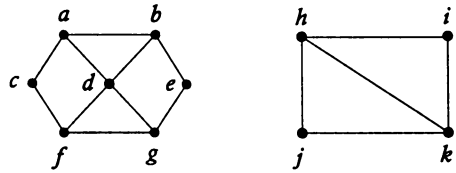


7. Какие из приведенных ниже пар графов гомеоморфны?

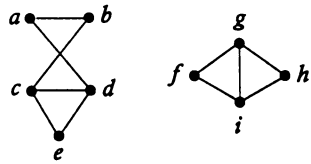
а)



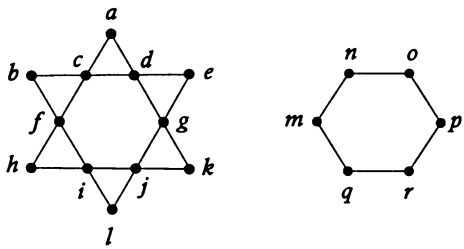
б)



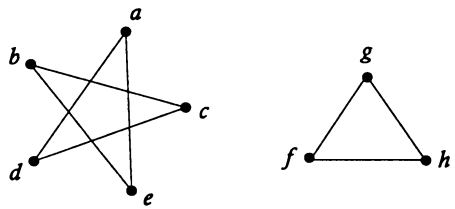
в)



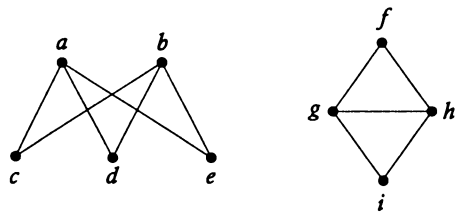
г)



д)

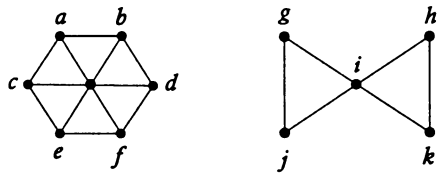


е)

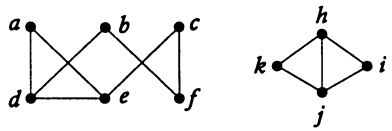


8. Какие из приведенных ниже пар графов гомеоморфны?

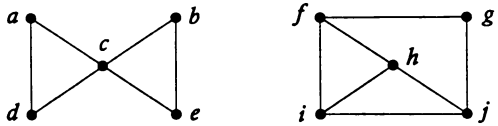
а)



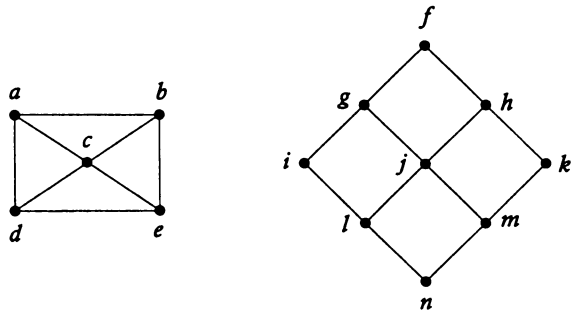
б)



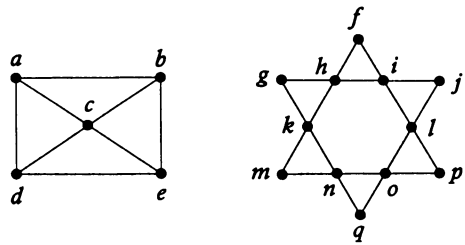
в)



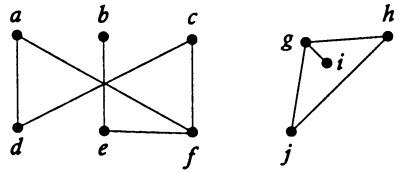
г)



д)

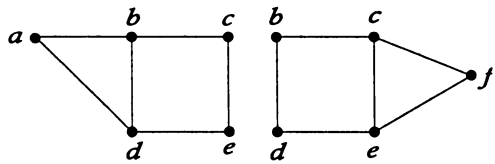


е)

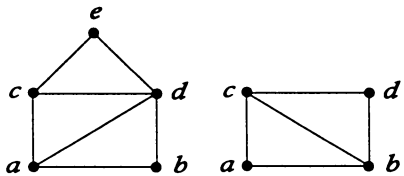


9. Найдите объединение и пересечение приведенных ниже множеств графов.

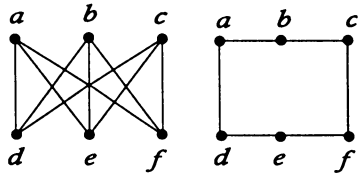
а)



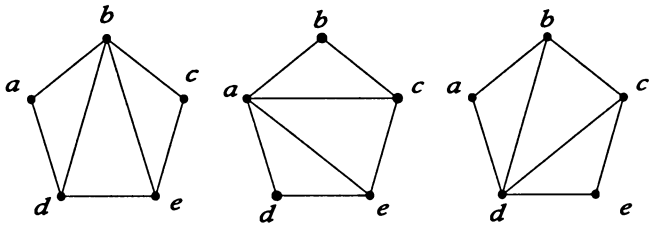
б)


e
c
d
a
b
c
d
a
b

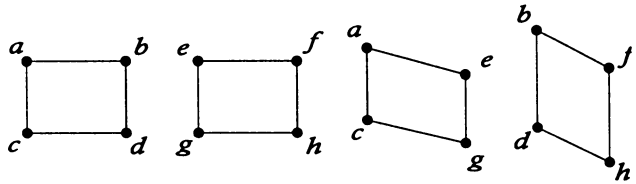
в)


a
b
c
d
e
f
a
b
c
d
e
f

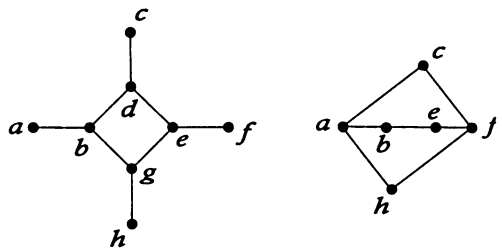
г)


b
a
c
d
e
b
a
c
d
e
b
a
c
d
e

д)

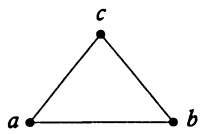

a
b
c
d
e
f
g
h
a
e
c
g
b
f
d
h

е)

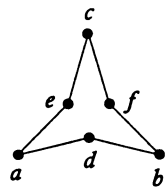

c
d
a
f
b
e
g
h
c
e
a
f
b
h

10. Найдите дополнения приведенных ниже графов.

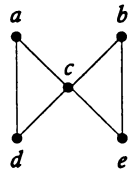
а)



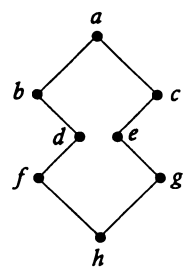
б)



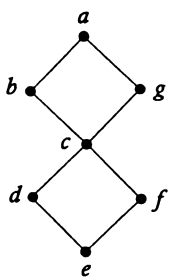
в)



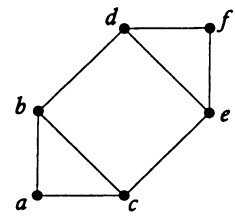
г)



д)

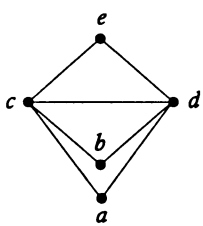


е)

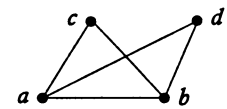


11. Найдите дополнения приведенных ниже графов.

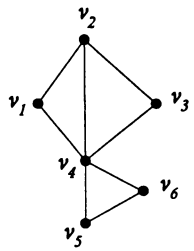
а)



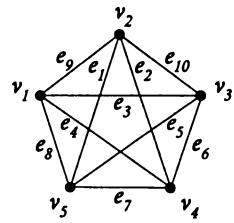
б)



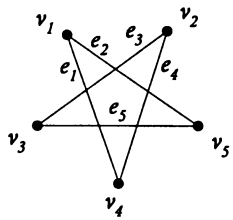
в)



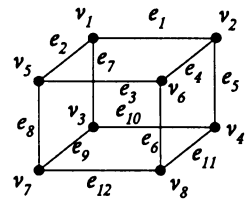
г)



д)



е)



12. Задан граф G, изображенный на рис. 14.22. Какие из приведенных ниже графов являются его остовными графами?

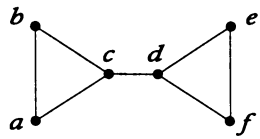
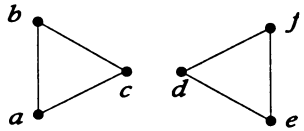
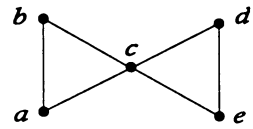


Рис. 14.22

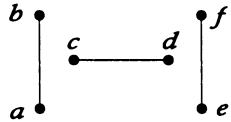
а)



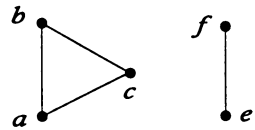
б)



в)



г)



13. Задан граф G, изображенный на рис. 14.23. Какие из приведенных ниже графов являются его остовными графами?

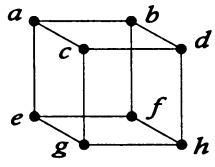
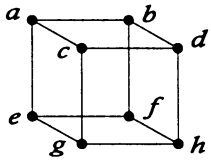
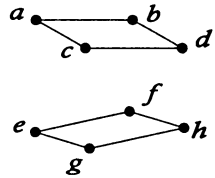


Рис. 14.23

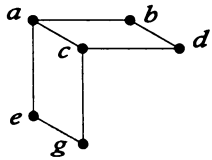
а)



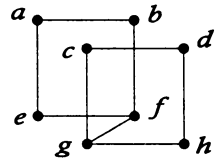
б)



в)



г)



14. Покажите, что любой гомоморфизм из графа, изображенного на рис. 14.24, на граф G является изоморфизмом.

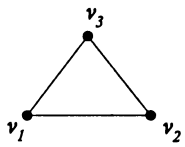


Рис. 14.24

15. Покажите, что если G и G' — графы, $f : G \to G'$ — гомоморфизм, а граф G связный, то граф G' может быть несвязным.

16. Докажите теорему 14.3. Если граф G связный и f — гомоморфизм, то граф $f(G)$ тоже связный.

17. Докажите теорему 14.4. Если граф G полный и f — гомоморфизм, то граф $f(G)$ тоже полный.

18. Докажите теорему 14.17. Если графы G_1 и G_2 являются различными компонентами графа G, то графы G_1 и G_2 попарно непересекающиеся.

19. Докажите теорему 14.18. Граф G — объединение попарно непересекающихся компонент.

20. Докажите теорему 14.23. Если $T(V, E')$ — остовное дерево графа $G = (V, E)$, то для любого цикла $v_0v_1v_2v_3v_4, \ldots v_0$ в графе G, по крайней мере, одно ребро должно принадлежать множеству $E - E'$.

21. Докажите, что если граф G связный и f — гомоморфизм, то граф $f(G)$ тоже связный.

22. Покажите, что если $G(V, E)$ и $G'(V', E')$ — графы, $f : G \to G'$ — изоморфизм, и вершина $v \in V$ имеет степень n, то в графе $f(G)$ вершина $f(v)$

необязательно имеет степень n. Что можно сказать о степени вершины $f(v)$ как о вершине графа $f(G)$? Что можно сказать о степени вершины $f(v)$ как о вершине графа G'?

23. Покажите, что если графы G и G' — изоморфны, то они имеют одинаковое количество вершин и ребер.

24. Покажите, что если f — изоморфизм из графа G в граф G', то граф G связный тогда и только тогда, когда граф G' связный.

25. Покажите, что если f — изоморфизм из графа G в граф G' и H — компонента графа G, то $f(H)$ — компонента графа G'.

26. Покажите, что если f — изоморфизм из графа G в граф G' и вершина $v \in V$ имеет степень n в графе G, то вершина $f(v)$ имеет степень n в графе G'.

27. Покажите, что если f — изоморфизм из графа G в граф G' и граф G имеет эйлеров цикл, то граф G' тоже имеет эйлеров цикл.

28. Покажите, что если f — изоморфизм из графа G в граф G' и граф G имеет собственный эйлеров путь, то граф G' тоже имеет собственный эйлеров путь.

29. Покажите, что если f — изоморфизм из графа G в граф G' и граф G является двудольным, то граф G' тоже является двудольным.

30. Покажите, что если f — изоморфизм из G в G' и граф G является полным двудольным графом K_{mn}, то граф G' является полным двудольным графом K_{mn}.

31. Докажите, что пересечение графа $G = (V, E)$ и его дополнения представляет собой граф, состоящий только из вершин графа G и не содержащий ребер.

32. Докажите, что объединение графа $G = (V, E)$ и его дополнения является полным графом.

33. Покажите, что если f — изоморфизм из графа G в граф G' и e — разрезающее ребро графа G, то $f(e)$ — разрезающее ребро графа G'.

34. Граф G — ***самодополнительный*** граф, если он изоморфен своему дополнению (см. определение 14.19). Приведите пример самодополнительного графа.

35. Какая связь между матрицами смежности графа и его дополнения?

36. Покажите, что граф G_1 изоморфен графу G тогда и только тогда, когда граф G_1^c изоморфен графу G^c.

37. *Пусть G и G' — графы, а B и B' — соответственно их матрицы смежности. Докажите, что граф G изоморфен графу G' тогда и только тогда, когда существует перестановка P такая, что $B' = PBP^t = PBP^{-1}$.

38. Найдите три разрезающих множества в графе, изображенном на рис. 14.25. Найдите точку сочленения. Имеются ли разрезающие ребра?

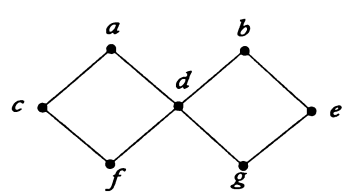


Рис. 14.25

39. Найдите три разрезающих множества в графе, изображенном на рис. 14.26. Найдите точки сочленения. Найдите разрезающие ребра.

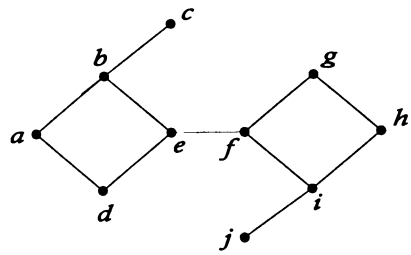


Рис. 14.26

40. Найдите три разрезающих множества в графе, изображенном на рис. 14.27. Найдите точки сочленения. Найдите разрезающие ребра.

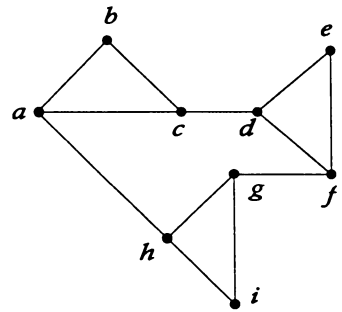


Рис. 14.27

41. Докажите теорему 14.28. Ребро e графа G является разрезающим ребром графа G тогда и только тогда, когда оно не входит в цикл графа G.

42. Докажите теорему 14.27. Если $T(V, E')$ — остовное дерево графа $G = (V, E)$ и C — разрезающее множество графа G, то $C \cap E' \neq \varnothing$.

43. Докажите теорему 14.36. Если компонента двусвязности $G_i(V_i, E_i)$ состоит из единственного ребра e_i, то e_i — разрезающее ребро графа G.

44. Дано следующее определение. Пусть $G = G(V, E)$ — граф и $G_1, G_2, G_3, \ldots, G_n$ — подграфы графа G. Подграф G' графа G называется ***пересечением*** графов $G_1, G_2, G_3, \ldots, G_n$ и обозначается $\bigcap\limits_{i=1}^{n} G_i$, если

1. Для всех i $G' \preceq G_i$.
2. Для всех i $G'' \preceq G_i \preceq G$, то $G'' \preceq G'$.

Покажите, что это определение эквивалентно определению пересечения, приведенному в данном разделе.

45. Дано следующее определение. Пусть $G = G(V, E)$ — граф и $G_1, G_2, G_3, \ldots, G_n$ — подграфы графа G. Подграф G' графа G называется ***объединением*** графов $G_1, G_2, G_3, \ldots, G_n$ и обозначается $\bigcup\limits_{i=1}^{n} G_i$, если

1. Для всех i $G_i \preceq G'$.
2. Для всех i $G_i \preceq G'' \preceq G$, то $G' \preceq G''$.

Покажите, что данное определение эквивалентно определению объединения, приведенному в данном разделе.

14.2. ПЛАНАРНЫЕ ГРАФЫ

Интегральная микросхема состоит из слоев миниатюрных микросхем, впечатанных в пластину. В такой ситуации крайне важно исключить пересечение проводов в местах, не предназначенных для соединений. Если изобразить места указанных соединений вершинами графа, то возникнет задача построения графа с непересекающимися ребрами. Важно отметить, что нас интересует *возможность* построения графа с непересекающимися ребрами. Например, граф, изображенный на рис. 14.28, изображается также другим способом (рис. 14.29). Граф, изображенный на рис. 14.30, может быть изображен, как показано на рис. 14.31.

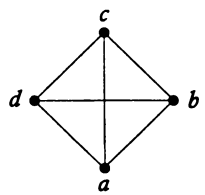


Рис. 14.28

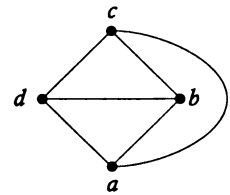


Рис. 14.29

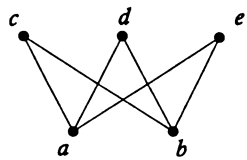


Рис. 14.30

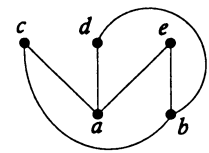


Рис. 14.31

Если бы проблема, рассмотренная нами в разделе 6.1, заключалась в предоставлении двух видов коммунальных услуг трем домам или трех видов коммунальных услуг двум домам, то ее можно было бы решить, используя в каждом случае граф, изображенный на рис. 14.32, согласно которому линии снабжения не будут пересекаться.

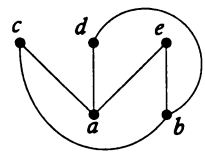


Рис. 14.32

Чтобы решить задачу о предоставлении коммунальных услуг трем домам, потребуются некоторые определения и методы. Начнем с наименования графов, обладающих указанными свойствами.

ОПРЕДЕЛЕНИЕ 14.43. ***Планарным графом*** называется граф, который может быть изображен в плоскости, так что его ребра не пересекаются. Граф, который не является планарным, называется ***непланарным***.

Рассмотрим граф как рисунок, изображенный на листе бумаги. Если граф планарен и нарисован так, что никакие линии не пересекаются, и его необходимо разрезать вдоль ребер, то граф окажется разделенным на части, включая внешнюю часть. Такие части называются ***гранями***. Заметим, что граница каждой грани является циклом.

ОПРЕДЕЛЕНИЕ 14.44. Грань планарного графа — максимальный участок плоскости такой, что любые две точки этого участка могут быть соединены кривой, не пересекающей ребро графа.

Теперь докажем теорему, называемую формулой Эйлера, с помощью которой можно определять, является ли граф планарным.

ТЕОРЕМА 14.45. Если G — связный планарный граф, содержащий v вершин, e ребер и f граней, то $v - e + f = 2$.

ДОКАЗАТЕЛЬСТВО. Докажем утверждение, используя индукцию по числу ребер. Если $e = 0$, то имеется одна вершина, нет ребер и есть одна грань, которая составляет всю плоскость. Поэтому имеем $1 - 0 + 1 = 2$, и теорема верна, если $e = 0$. Отметим также, что если имеется одно ребро, то существуют две вершины и одна грань. Поэтому имеем $2 - 1 + 1 = 2$; и теорема верна для $e = 1$.

Предположим, что теорема верна для произвольного связного планарного графа, у которого k ребер, и имеется планарный граф G_{k+1}, у которого $k + 1$ ребер. Теперь удалим ребро, так что останется связный планарный граф G_k, у которого k ребер, поэтому формула $v - e + f = 2$ для него верна. После этого докажем, что формула также верна и для связного планарного графа G_{k+1}.

Предположим, что граф G_{k+1} не имеет циклов. Движемся вдоль пути до тех пор, пока не достигнем вершины, из которой нет другого выходящего ребра. Это произойдет, если граф G_{k+1} не содержит циклы. Найденная вершина будет иметь степень 1. Удалим эту вершину и ребро, инцидентное ей. Количество граней не изменится. При этом граф останется связным и планарным и будет иметь k ребер, значит формула $v - e + f = 2$ верна. Поскольку количество ребер и количество вершин уменьшилось на 1, то значение выражения $v - e + f$ после процедуры удаления останется неизменным, поэтому $v - e + f = 2$ для связного планарного графа G_{k+1}.

Далее предположим, что граф G_{k+1} содержит цикл. Удалим из цикла ребро e_i с инцидентными вершинами u_i и v_i, но сами вершины оставим. Согласно теореме 14.28 известно, что все еще имеется путь из вершины u_i в вершину v_i, так что граф остается связным. Граф также остается планарным и будет иметь k ребер, так что выполняется формула $v - e + f = 2$. Поскольку ребро e_i входит в цикл, оно разделяет две грани. Следовательно, удаление ребра равноценно удалению грани. Таким образом, удалены одно ребро и одна грань. Поэтому значение выражения $v - e + f$ не изменилось, и формула $v - e + f = 2$ справедлива для графа G_{k+1}. ■

Теперь имеется все необходимое для доказательства того, что граф, описывающий проблему предоставления трех видов коммунальных услуг трем домам, является полным двудольным графом $K_{3,3}$ и его нельзя изобразить без пересекающихся ребер.

ТЕОРЕМА 14.46. Полный двудольный граф $K_{3,3}$ *не* является планарным.

ДОКАЗАТЕЛЬСТВО. Используем метод приведения к абсурду и предположим, что граф $K_{3,3}$ — планарный. Если $K_{3,3}$ планарный граф, и поскольку имеется девять ребер и шесть вершин, то $6 - 9 + f = 2$, поэтому $f = 5$. Пусть A и B — непересекающиеся трехэлементные множества вершин, формирующие множество V вершин графа $K_{3,3}$. Если начать путь из одного из непересекающихся множеств, например, A, и не повторять ребра, то можно попасть в вершину из множества B, вернуться в вершину из множества A, вернуться в вершину из множества B и, наконец, вернуться в вершину из множества A, прежде чем завершить цикл. Каждый цикл в $K_{3,3}$ представляет собой путь, длина которого, по меньшей мере, равна 4. Поэтому каждая грань определена циклом, в котором не менее четырех ребер. Следовательно, сумма ребер всех граней больше, чем $4f$. Но каждое ребро подсчитывается не более двух раз, поскольку оно может служить границей только для двух граней. Значит, сумма ребер всех граней должна быть меньше, чем $2e$. Объединяя эти неравенства, получаем, $4f \leq 2e$, поэтому $4f \leq 18$. Но это противоречит тому, что $f = 5$. Следовательно, мы пришли к противоречию, и граф $K_{3,3}$ *не* является планарным. ■

Теперь докажем, что полный граф K_5, изображенный на рис. 14.33, не является планарным. Для этого нам необходима лемма (так называется теорема, которую используют для доказательства другой теоремы).

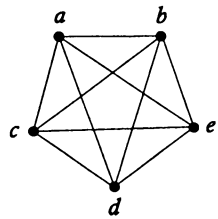


Рис. 14.33

ЛЕММА 14.47. В произвольном связном планарном графе G с количеством вершин не менее трех имеет место неравенство $3v - e \geq 6$.

ДОКАЗАТЕЛЬСТВО. Если граф G не имеет циклов, то в нем количество ребер меньше количества вершин: $e \leq v$. Учитывая также, что $v \geq 3$, $2v - 6 \geq 0$, получаем, $e \leq v + 2v - 6 = 3v - 6$ или $3v - e \geq 6$. Если граф G содержит цикл, опять суммируем ребра, ограничивающие грани. Поскольку граница каждой грани включает не менее трех ребер, то сумма ребер всех граней должна быть больше $3f$. Но каждое ребро может быть границей не более двух граней (см. предыдущее доказательство), поэтому количество ребер должно быть меньше $2e$. Следовательно, $3f \leq 2e$ и $2 = v - e + f$, откуда $6 = 3v - 3e + 3f \leq 3v - 3e + 2e = 3v - e$ и $3v - e \geq 6$. ■

ТЕОРЕМА 14.48. Полный граф K_5 не является планарным.

ДОКАЗАТЕЛЬСТВО. Граф K_5 имеет пять вершин и десять ребер, $3v - e = 3 \cdot 5 - 10 = 5$, поэтому, согласно предыдущей лемме, граф K_5 не является планарным. ■

Используя предыдущую лемму, докажем еще одну теорему, которая будет использована в следующем разделе для доказательства того, что планарный граф можно раскрасить, используя не более пяти цветов.

ТЕОРЕМА 14.49. Каждый планарный граф G содержит вершину степени 5 или менее.

ДОКАЗАТЕЛЬСТВО. Для каждого ребра имеется две вершины (см. главу 2), каждое ребро вносит число 2 в сумму степеней вершин. Поэтому сумма степеней вершин равна $2e$. Если в графе G не более двух вершин, то, очевидно, вершины имеют степень 5 или менее. Отсюда предположим, что имеются три и более вершин. Тогда по предыдущей лемме имеем $e \leq 3v - 6$, поэтому $2e \leq 6v - 12$. Но если степень каждой вершины больше или равна 6, то сумма степеней вершин больше или равна $6v$. По условию сумма степеней вершин равна $2e$, поэтому $2e \geqslant 6v$. А это невозможно, поскольку $2e \leq 6v - 12$. Следовательно, существует вершина со степенью не более 5. ■

Доказательство следующей теоремы предоставляется читателю.

ТЕОРЕМА 14.50. Если два связных графа гомеоморфны, то они либо оба планарны, либо оба непланарны.

Из этой теоремы вытекает непосредственно следующая теорема.

ТЕОРЕМА 14.51. Произвольный граф, гомеоморфный графу $K_{3,3}$ или K_5, не является планарным.

Поскольку любой граф, содержащий непланарный подграф, является непланарным, это доказывает часть приведенной ниже теоремы. Остальную часть утверждения примем без доказательства.

ТЕОРЕМА 14.52. (Куратовский) Граф является планарным тогда и только тогда, когда он не содержит подграф, гомеоморфный $K_{3,3}$ или K_5.

ПРИМЕР 14.53. Покажем, что граф, изображенный на рис. 14.34, является планарным. Учитывая, что граф состоит из двух компонент, можем отделить их и образовать граф, изображенный на рис. 14.35, который, очевидно, является планарным.

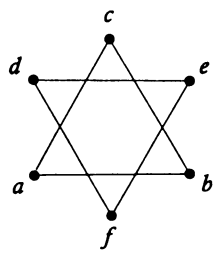


Рис. 14.34

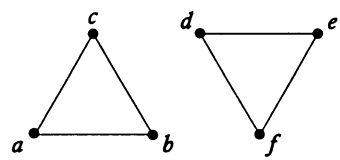


Рис. 14.35

□

ПРИМЕР 14.54. Покажем, что граф, изображенный на рис. 14.36, является планарным. Передвигая вершину d, получаем более простой граф, изображенный на рис. 14.37. Передвигая вершину c, получаем граф (рис. 14.38), который, очевидно, является планарным.

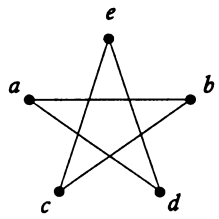


Рис. 14.36

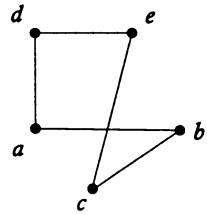


Рис. 14.37

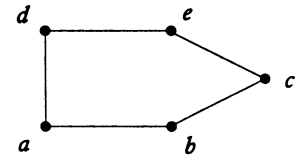


Рис. 14.38 □

ПРИМЕР 14.55. Покажем, что граф Петерсена, изображенный на рис. 14.39, не является планарным. Покажем, что граф содержит подграф, гомеоморфный $K_{3,3}$. Единственная трудность — выбрать правильное множество вершин. Если в качестве верхних вершин выбрать e, f и g, а в качестве нижних — j, a, и i и соединить каждую верхнюю вершину с каждой нижней вершиной, получим граф, изображенный на рис. 14.40. Отсюда видим, что вершина c не нужна, поэтому вершину и все ребра, инцидентные ей, удаляем. Все остальные ребра пока присутствуют. Теперь можем удалить вершины h, b и d, формируя ребра $\{f, j\}$, $\{a, g\}$, $\{e, i\}$, соответственно, и получить граф, изображенный на рис. 14.41, который является гомеоморфным предыдущему графу. Но этот граф может быть изображен как $K_{3,3}$. Следовательно, граф Петерсена содержит подграф, гомеоморфный $K_{3,3}$, и поэтому не является планарным.

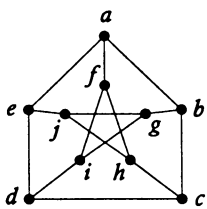


Рис. 14.39

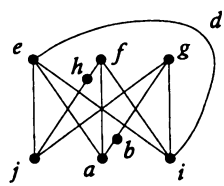


Рис. 14.40

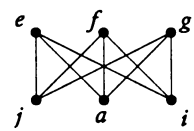
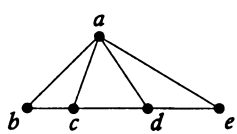


Рис. 14.41 □

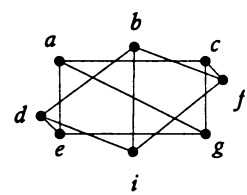
■ УПРАЖНЕНИЯ

1. Каждый из приведенных ниже графов проверить на планарность. Ответ аргументируйте.

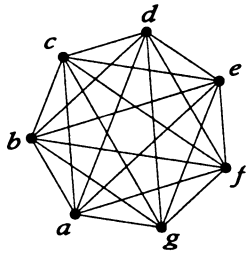
а)



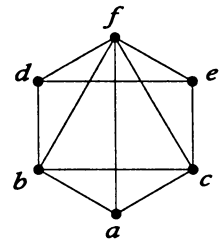
б)



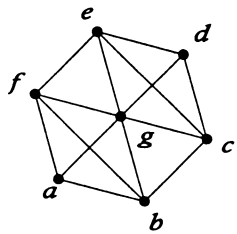
в)



г)



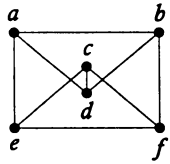
д)



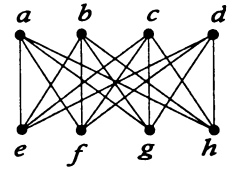
2. Для каждого планарного графа из предыдущего упражнения проверить, что $v - e + f = 2$.

3. Каждый из приведенных ниже графов проверить на планарность. Аргументировать решение.

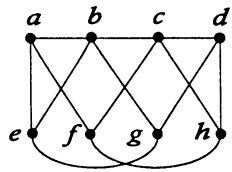
а)



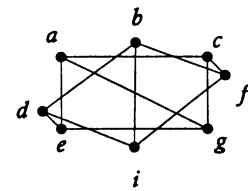
б)



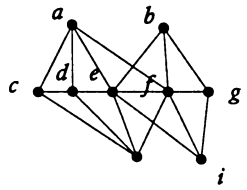
в)



г)



д)



4. Для каждого планарного графа из предыдущего упражнения проверить, что $v - e + f = 2$.
5. Если планарный граф содержит 12 вершин со степенью 3, то сколько у него ребер и граней ?
6. Если планарный граф имеет вершины со степенями 2, 2, 2, 3, 3, 3, 4, 4, 4 и 5, соответственно, то сколько у него ребер и граней?
7. Найдите значение величины $v - e + f$ для графа, у которого
   а) две компоненты;
   б) три компоненты;
   в) k компонент.
8. Если G — планарный граф, у которого n вершин и каждая грань графа G ограничена циклом длины 3, найдите зависимость количества граней и ребер от числа n.
9. Докажите, что произвольный граф с четырьмя вершинами является планарным.
10. Докажите, что граф с пятью вершинами, одна из которых имеет степень 2, является планарным.
11. Используя метод индукции, докажите, что дерево является планарным графом.
12. Докажите теорему 14.50. Если два связанных графа гомеоморфны, то они или оба планарны, или оба непланарны.

14.3. РАСКРАСКА ГРАФОВ

Одной из наиболее известных нерешенных проблем на протяжение многих лет оставалась так называемая проблема четырех красок. Она возникла в связи с раскрашиванием географических карт и заключается в следующем: карту нужно раскрасить, используя только четыре цвета, так, чтобы любые две граничащие страны были раскрашены разными цветами. Известно, что для раскрашивания карты пяти красок достаточно, а трех — нет. Для карты, нанесенной на тор, минимальное количество красок было определено, но проблема четырех красок оставалась нерешенной. На сегодняшний день она решена, но, возможно, не самым лучшим способом. Решение потребовало длительной проверки компьютером огромного числа случаев, которые иначе проанализировать не удавалось. Хотя сам по себе метод является выдающимся достижением и вполне достаточен, чтобы прекратить поиски контрпримера, было бы неплохо, если бы кому-то удалось

найти более элегантное доказательство гипотезы. Самое замечательное в этом доказательстве — то, что оно расширило наше представление о математическом доказательстве.

Рассмотрим карту, изображенную на рис. 14.42.

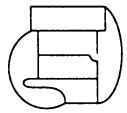


Рис. 14.42

Если левая часть карты закрашена красным цветом, то при закрашивании примыкающих центральных граней чередуются синий и зеленый цвета. Правую грань снова можно закрасить красным цветом, поскольку, как следует из рис. 14.43, она не граничит с левой гранью.

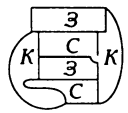


Рис. 14.43

При изображении карты с помощью графа естественно использовать ребра в качестве границ, а грани — в качестве стран, так что карта, изображенная на рис. 14.44,

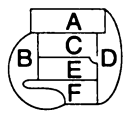


Рис. 14.44

превращается в граф, изображенный на рис. 14.45.

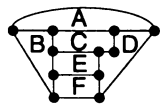


Рис. 14.45

Однако, вместо использования такого графа лучше сделать некоторые изменения. Гораздо удобнее работать с **двойственным графом** G' планарного графа G, который формируется путем замены граней (или стран) вершинами и соединения их ребрами, если грани в исходном графе смежные. Следовательно, в двойственном графе G' вершины являются смежными тогда и только тогда, когда соответствующие грани являются смежными в исходном графе G. Таким образом, двойственным к рассматриваемому графу становится граф, изображенный на рис. 14.46. Внешняя грань будет включаться как вершина в том случае, если она раскрашена.

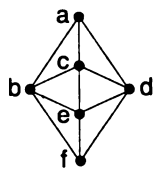


Рис. 14.46

ПРИМЕР 14.56. Для графа G, изображенного на рисунке 14.47, двойственным является граф G' (рис. 14.48), поэтому в данном случае граф G изоморфен графу G'.

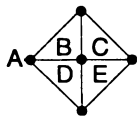


Рис. 14.47

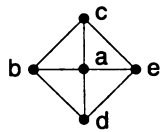


Рис. 14.48 □

Поскольку грани графа G стали вершинами G', раскрашиваем теперь вершины графа G'. Исходная проблема раскрашивания стран на карте при условии, чтобы никакие две страны, имеющие общую границу, не были окрашены в один цвет, стала теперь проблемой раскрашивания вершин графа при том же условии.

А теперь рассмотрим новую задачу. Зафиксируем для графа количество цветов и попытаемся определить, сколько существует способов раскрашивания вершин при условии, что никакие две смежные не будут одного цвета. Эта концепция принадлежит Г.Д. Биркгофу (G. D. Birkhoff) [11]. Предположим, например, что имеется пять цветов и необходимо раскрасить граф, изображенный на рис. 14.49.



Рис. 14.49

Для раскрашивания первой вершины имеется пять цветов, и только четыре цвета для раскашивания второй. Необходимо учесть, что вторая вершина отличается по цвету от первой. Следовательно, существует $5 \cdot 4 = 20$ способов раскрашивания графа. Заметим, что данный метод использовался в разделе 8.1 для подсчета количества перестановок, а также и в других случаях, при использовании комбинаторного принципа умножения.

Предположим, что имеются четыре цвета и необходимо раскрасить граф G, изображенный на рис. 14.50

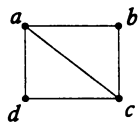


Рис. 14.50

Для раскрашивания первой вершины a имеются четыре цвета. Далее, имеются три цвета для раскрашивания вершины b.Только два цвета остается для раскрашивания вершины c, поскольку она смежная с a и b, цвет ее вершины должен отличаться от цвета вершин a и b. Для раскрашивания вершины d также имеются два цвета. Цвет вершины d может совпадать с цветом вершины b, но должен отличаться от цвета смежных вершин a и c. Таким образом, имеем $4 \cdot 3 \cdot 2 \cdot 2 = 48$ способов раскраски графа G. Важно не раскрашивать вершины b и d до определения количества цветов для вершины c, иначе потребуется разбить подсчет на два

случая: когда b и d — одного цвета и когда b и d по цвету не совпадают. Предположим, что имеется λ цветов для раскрашивания графа G. Имеются λ цветов для раскрашивания вершины a, $\lambda-1$ цветов для раскрашивания вершины b, $\lambda-2$ цветов для раскрашивания c и $\lambda-2$ цветов для раскрашивания d. Следовательно, имеются $(\lambda)(\lambda-1)(\lambda-2)(\lambda-2)$ способов раскраски графа G. Обозначим эту величину через $C_G(\lambda)$. Таким образом, $C_G(\lambda)$ — число способов раскраски графа G с использованием λ цветов. Заметим, что $C_G(0)=C_G(1)=C_G(2)=0$.

Функция $C_G(\lambda)$ предполагается полиномиальной при формулировке нижеприведенного определения. На интуитивном уровне это кажется верным и будет доказано позже.

ОПРЕДЕЛЕНИЕ 14.57. Пусть G — граф. ***Раскраской*** графа G называется окрашивание вершин графа G такое, что никакие две смежные вершины не имеют один цвет. Пусть $C_G(\lambda)$ обозначает количество способов раскраски графа G с использованием λ цветов, так что никакие две смежные вершины не имеют один цвет, т.е. $C_G(\lambda)$ — количество способов раскраски графа G. Для фиксированного графа G функция $C_G(\lambda)$ является полиномиальной функцией от λ, называемой ***хроматическим многочленом*** графа G. ***Хроматическое число*** графа — это наименьшее число цветов, которое используется для раскраски графа. Это наименьшее положительное число n такое, что $C_G(n)\neq 0$.

ПРИМЕР 14.58. Пусть граф G состоит из пяти изолированных вершин, так что в нем нет ребер, поэтому нет смежных вершин. Если раскрашивать граф λ цветами, то будем иметь λ вариантов для каждой вершины, поэтому будет существовать $\lambda\cdot\lambda\cdot\lambda\cdot\lambda\cdot\lambda=\lambda^5$ способов раскраски графа и $C_G(\lambda)=\lambda^5$. Очевидно, что если граф G состоит из k изолированных вершин, то $C_G(\lambda)=\lambda^k$, а хроматическое число равно 1. □

ПРИМЕР 14.59. Пусть G — граф K_5, так что каждая из пяти его вершин является смежной с остальными четырьмя вершинами. Если раскрашивать этот граф λ цветами, то для первой вершины имеется λ вариантов цвета. Поскольку вторая вершина смежная с первой, то имеем $\lambda-1$ вариантов выбора цвета для следующей вершины. Третья вершина является смежной с первой и со второй, поэтому для нее имеется $\lambda-2$ вариантов выбора. Четвертая вершина — смежная с каждой из первых трех раскрашенных, поэтому для нее имеется $\lambda-3$ вариантов выбора цвета. И, наконец, пятая вершина является смежной с каждой из первых четырех окрашенных вершин, поэтому имеется $\lambda-4$ вариантов выбора цвета для этой вершины. Таким образом, $C_G(\lambda)=(\lambda)(\lambda-1)(\lambda-2)(\lambda-3)(\lambda-4)$. Отметим, что этот граф не может быть раскрашен четырьмя цветами, поскольку $C_G(4)=0$, но учитывая, что граф G не является планарным, этот факт не противоречит утверждению, что планарный граф может быть окрашен четырьмя цветами. Используя рассмотренный выше частный случай графа K_5 в качестве ориентира для изучения графа $G=K_n$, находим, что если $G=K_n$, то $C_G(\lambda)=(\lambda)(\lambda-1)(\lambda-2)(\lambda-3)\cdots(\lambda-(n-1))$. Таким образом, хроматическое число для графа G равно n. □

Раскраска одной компоненты графа не ограничивает раскрашивание другой компоненты, поскольку никакая из вершин одной компоненты не является смежной с вершиной другой компоненты. Следовательно, имеем следующую теорему.

ТЕОРЕМА 14.60. Если $G = G_1 \cup G_2 \cup G_3 \cup \cdots \cup G_n$, где $G_1, G_2, G_3, \ldots, G_n$ — компоненты графа G, то $C_G(\lambda) = C_{G_1}(\lambda) \cdot C_{G_2}(\lambda) \cdot C_{G_3}(\lambda) \cdot \cdots \cdot C_{G_n}(\lambda)$.

Из теоремы следует, что $C_G(\lambda) = 0$ тогда и только тогда, когда $C_{G_i}(\lambda) = 0$ для некоторого $1 \leq i \leq n$. Отсюда вытекает приведенное ниже следствие.

СЛЕДСТВИЕ 14.61. Если для раскраски графа G требуется k цветов, то и для одной из его компонент требуется для раскраски k цветов.

Учитывая этот результат, рассмотрим раскрашивание только связных графов.

Теперь на основе заданного графа $G(V, E)$ сформируем два специальных графа. Пусть $e = \{a, b\}$ — ребро графа G. Пусть $G_{\hat{e}}$ — граф G, в котором удалено ребро e, но сохранены вершины a и b. Пусть $G_{/e}$ — граф $G_{\hat{e}}$ с отождествленными (склеенными) вершинами a и b.

Отметим, что граф $G_{/e}$, фактически, является гомоморфным образом графа $G_{\hat{e}}$, где гомоморфизм $f : G_{\hat{e}} \to G_{/e}$ определен на вершинах графа $G_{\hat{e}}$ таким образом: $f(a) = f(b)$ и $f(v) = v$ для всех v из $V - \{a, b\}$. Для ребер имеем $f(\{u, v\}) = \{f(u), f(v)\}$. Также отметим, что $f = f \circ f$, поскольку для вершин и ребер графа $G_{/e}$ функция f представляет собой тождественное отображение. Функция с такими свойствами называется стягивающим отображением.

ПРИМЕР 14.62. Пусть G — граф, изображенный на рис. 14.51. Тогда граф $G_{\hat{e}}$ изображен на рис. 14.52, а $G_{/e}$ — на рис. 14.53.

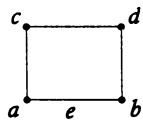


Рис. 14.51

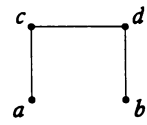


Рис. 14.52

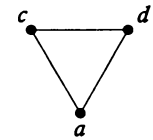


Рис. 14.53 □

ПРИМЕР 14.63. Пусть G — граф, изображенный на рис. 14.54. Тогда $G_{\hat{e}}$ — граф на рис. 14.55, а $G_{/e}$ — на рис. 14.56.

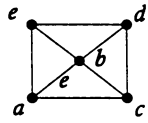


Рис. 14.54

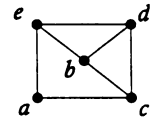


Рис. 14.55

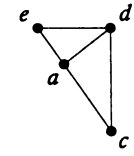


Рис. 14.56 □

Введенные специальные графы используем для доказательства следующей теоремы.

ТЕОРЕМА 14.64. Для произвольного планарного графа G с ребром e имеет место равенство $C_{G_{\hat{e}}}(\lambda) = C_G(\lambda) + C_{G_{/e}}(\lambda)$.

ДОКАЗАТЕЛЬСТВО. Предположим, что имеется раскраска графа $G_{\hat{e}}$. Если цвет вершины a не совпадает с цветом вершины b, то эта раскраска будет также раскраской графа G и наоборот. Если цвет вершины a совпадает с цветом вершины b, в таком случае имеем раскраску графа $G_{/e}$ и наоборот. Поэтому произвольная раскраска графа $G_{\hat{e}}$ является либо раскраской графа G, либо раскраской графа $G_{/e}$, но не обоих одновременно. Следовательно, количество раскрасок графа $G_{\hat{e}}$ равно количеству раскрасок графа G плюс количество раскрасок графа $G_{/e}$, или $C_{\hat{e}}(\lambda) = C_G(\lambda) + C_{/e}(\lambda)$. ■

Интуитивно понятно, что для произвольного планарного графа G с n вершинами $C_G(\lambda)$ — многочлен степени n, поскольку всякий раз в примерах при построении $C_G(\lambda)$ формировалась для каждой вершины величина $\lambda - k$ для некоторого целого числа $0 \leq k \leq n$. Фактически, это предположение было использовано при определении хроматических многочленов. Теперь докажем это, строго используя предыдущую теорему и метод индукции по числу ребер.

ТЕОРЕМА 14.65. Для произвольного планарного графа G с n вершинами функция $C_G(\lambda)$ представляет собой многочлен n-ой степени.

ДОКАЗАТЕЛЬСТВО. Пусть m — количество ребер в планарном графе G с n вершинами. Если $m = 0$, тогда согласно примеру 14.58 функция $C_G(\lambda) = \lambda^n$, т.е. является многочленом n-ой степени. Предположим, что теорема верна для всех планарных графов, у которых n вершин и k ребер. Напомним, что k — фиксированное число, а n — нет. Пусть граф G имеет n вершин и $k+1$ ребер. По теореме 14.64, функция $C_G(\lambda) = C_{G_{\hat{e}}}(\lambda) - C_{G_{/e}}(\lambda)$. Но граф $G_{\hat{e}}$ имеет n вершин и k ребер, и граф $G_{/e}$ имеет $n-1$ вершину и k ребер. К тому же оба графа планарны. Следовательно, по индуктивному предположению функции $C_{G_{\hat{e}}}(\lambda)$ и $C_{G_{/e}}(\lambda)$ — многочлены степени n и $n-1$ соответственно. Поэтому $C_G(\lambda) = C_{G_{\hat{e}}}(\lambda) - C_{G_{/e}}(\lambda)$ — многочлен степени n. ■

Из доказанной теоремы вытекает следующая теорема.

ТЕОРЕМА 14.66. Для произвольного непустого планарного связного графа G постоянный член в $C_G(\lambda)$ равен 0. Если граф G имеет две или более вершин, то сумма коэффициентов многочлена $C_G(\lambda)$ равна 0.

ДОКАЗАТЕЛЬСТВО. Поскольку непустой планарный граф G не может быт раскрашен 0 цветами, то $C_G(0) = 0$. Если граф G имеет две или более вершин, то поскольку граф G связный, существуют две смежные вершины и граф G не может быть раскрашен одним цветом. Поэтому $C_G(1) = 0$. Но значение $C_G(1)$ равно сумме коэффициентов многочлена $C_G(\lambda)$, поэтому сумма коэффициентов многочлена $C_G(\lambda)$ равна 0. ■

Использование равенства $C_{G_{\hat{e}}}(\lambda) = C_G(\lambda) + C_{G_{/e}}(\lambda)$ используется двумя способами. Предположим, имеется граф $G_{\hat{e}}$, тогда можно найти многочлен $C_{G_{\hat{e}}}(\lambda)$, выраженный как сумма многочленов $C_G(\lambda)$ и $C_{G_{/e}}(\lambda)$. Но граф G по сравнению с $G_{\hat{e}}$ имеет дополнительное ребро и, таким образом, увеличивает количество ребер, а граф $G_{/e}$, в котором две вершины графа $G_{\hat{e}}$ отождествлены, уменьшает количество вершин. В конце концов получим полные графы. Теоретически — это ответ, но, возможно, не самый простой.

Предположим, например, что требуется найти хроматический многочлен для графа, изображенного на рис. 14.57. Если положить $e = \{a, b\}$, а $G_{\hat{e}}$ будет вышеупомянутым графом, в этом случае G — граф, изображенный на рис. 14.58, а $G_{/e}$ — граф, изображенный на рис. 14.59.

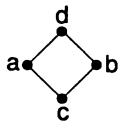


Рис. 14.57

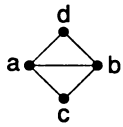


Рис. 14.58



Рис. 14.59

Следовательно, $C_{G_{\hat{e}}}(\lambda) = C_G(\lambda) + G_{/e}$. Если f — ребро $\{c, d\}$ и $G = G'_{\hat{f}}$, тогда G' — граф, изображенный на рис. 14.60, $G_{/f}$ — граф, изображенный на рис. 14.61,

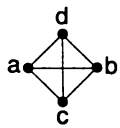


Рис. 14.60

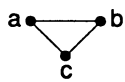


Рис. 14.61

и $C'_{G'_{\hat{f}}}(\lambda) = C_{G'}(\lambda) + C_{G'_{/f}}(\lambda)$. Но G' является графом K_4, поэтому

$$C_{G'}(\lambda) = \lambda(\lambda - 1)(\lambda - 2)(\lambda - 3).$$

Граф $G_{/f}$ является графом K_3, поэтому

$$C_{G_{/f}}(\lambda) = \lambda(\lambda - 1)(\lambda - 2).$$

Следовательно,

$$\begin{aligned} C_G(\lambda) = C'_{G'_{\hat{f}}}(\lambda) &= C_{G'}(\lambda) + C_{G_{/f}}(\lambda) = \\ &= \lambda(\lambda - 1)(\lambda - 2)(\lambda - 3) + \lambda(\lambda - 1)(\lambda - 2) = \\ &= \lambda(\lambda - 1)(\lambda - 2)(\lambda - 2). \end{aligned}$$

Пусть $g = \{c, d\}$ и $G_{/e}$ — граф $G''_{\hat{g}}$. Тогда граф G'' изображен на рис. 14.62, граф $G_{/f}$ — на рис. 14.63,



Рис. 14.62

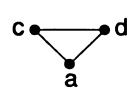


Рис. 14.63

и $C'_{G''_g}(\lambda) = C_{G''}(\lambda) + C_{G/g}(\lambda)$. Но G'' является графом K_3, поэтому

$$C_{G''}(\lambda) = \lambda(\lambda-1)(\lambda-2).$$

G/g является графом K_2, поэтому

$$C_{G/f}(\lambda) = \lambda(\lambda-1).$$

Следовательно,

$$\begin{aligned} C_{G/e} = C'_{G''_g}(\lambda) &= C_{G''}(\lambda) + C_{G/g}(\lambda) = \\ &= \lambda(\lambda-1)(\lambda-2) + \lambda(\lambda-1) = \\ &= \lambda(\lambda-1)(\lambda-1). \end{aligned}$$

Таким образом,

$$\begin{aligned} C_{G\hat{e}}(\lambda) &= C_G(\lambda) + G_{/e} = \\ &= \lambda(\lambda-1)(\lambda-2)(\lambda-2) + \lambda(\lambda-1)(\lambda-1) = \\ &= \lambda(\lambda-1)[(\lambda-2)(\lambda-2) + (\lambda-1)] = \\ &= \lambda(\lambda-1)(\lambda^2-3\lambda+3). \end{aligned}$$

Второй метод использования равенства $C_{G\hat{e}}(\lambda) = C_G(\lambda) + C_{G/e}(\lambda)$ является рекурсивным, в котором данное равенство записывется в виде $C_G(\lambda) = C_{G\hat{e}}(\lambda) - C_{G/e}(\lambda)$. Согласно этому методу удаляем ребра и склеиваем вершины, уменьшая количество ребер и вершин в графе до тех пор, пока это возможно, и, если необходимо, получаем изолированные вершины. Рассмотрим тот же граф на рис. 14.64. Если предположить, что G является этим графом, и удалить ребро $\{a, d\}$, то $G_{\hat{e}}$ — граф, изображенный на рис. 14.65, а $G_{/e}$ — на рис. 14.66.

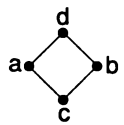


Рис. 14.64

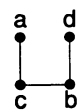


Рис. 14.65

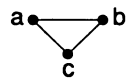


Рис. 14.66

Если для раскраски графа $G_{\hat{e}}$ используется λ цветов, то и для раскраски вершины a можно использовать λ цветов, а для раскраски любой другой вершины можно использовать $\lambda-1$ цветов, поскольку каждая из оставшихся вершин может быть окрашена в любой цвет, отличный от цвета предыдущей вершины. Таким образом,

$$C_{G\hat{e}}(\lambda) = \lambda(\lambda-1)^3.$$

Граф $G_{/e} = K_3$, так что

$$C_{G/e}(\lambda) = \lambda(\lambda-1)(\lambda-2)$$

и $C_{G_{\hat{e}}}(\lambda) - C_{G_{/e}}(\lambda)$,

$$\begin{aligned} C_G(\lambda) &= C_{G_{\hat{e}}}(\lambda) - C_{G_{/e}}(\lambda) = \\ &= \lambda(\lambda-1)^3 + \lambda(\lambda-1)(\lambda-2) = \\ &= \lambda(\lambda-1)((\lambda-1)^2 - (\lambda-2)) = \\ &\quad \lambda(\lambda-1)(\lambda^2 - 3\lambda + 3), \end{aligned}$$

и этот результат, к счастью, совпадает с вычисленным ранее.

Проблема четырех красок эквивалентна утверждению, что для каждого планарного графа G значение $C_G(4) \neq 0$. Как уже отмечалось, проблема четырех красок в течение долгого времени оставалась нерешенной. На это были две весомые причины. С одной стороны, проблема столь проста, что ее можно было объяснить любому, и все же она оставалась нерешенной. Проблема впервые была сформулирована в 1852 Фрэнсисом Гутерье (Francis Gutherie), учеником Огюста де Моргана (Augustus De Morgan). Некоторые математики посвятили жизнь решению этой проблемы. Существовало достаточно много "доказательств" проблемы четырех красок, найденных любителями и профессиональными математиками. В 1880 году казалось, что Альфред Бэй Кэмпе (Alfred Bay Kempe) [60] решил проблему. Однако, в 1890, П. Дж. Хивуд (P. J. Heawood) [42] нашел в его доказательстве ошибку. Модифицировав доказательство Кемпе, П. Дж. Хивуд смог доказать, что произвольную плоскую карту можно раскрасить пятью цветами.

ТЕОРЕМА 14.67. Произвольный планарный граф G можно раскрасить, используя только пять цветов.

ДОКАЗАТЕЛЬСТВО. В соответствии с доводами, принятыми ранее, предположим, что граф G связный. Помимо этого будем считать, что граф G изображен так, что никакие из его ребер не пересекаются. В доказательстве используется индукция по количеству вершин. Если граф включает только одну вершину (фактически пять или менее), то, несомненно, такой граф можно раскрасить, используя только пять цветов. Предположим, что при раскрашивании произвольного графа с k вершинами используются только пять цветов. Пусть G — граф с $k+1$ вершиной. По теореме 14.49 граф имеет вершину степени 5 или менее. Пусть v — такая вершина. Пусть G' — подграф графа G, в котором удалена вершина v и все инцидентные ей ребра. Поскольку в графе G' только k вершин, то по индуктивному предположению граф G' можно раскрасить, используя только пять цветов. Если степень вершины v меньше пяти, тогда в графе G она является смежной с четырьмя или менее вершинами и, следовательно, может быть раскрашена цветом, отличным от цветов смежных вершин. Раскраска завершена. Поэтому предположим, что степень вершины v равна 5. В этом случае вершина v смежная с пятью вершинами графа G, назовем их v_1, v_2, v_3, v_4 и v_5. Если какие-либо вершины из этих пяти имеют одинаковый цвет, то для их окраски использовано только четыре цвета, и можно снова выбрать неиспользованный цвет для окрашивания вершины v, в результате раскраска завершена. Поэтому предположим, что вершины v_1, v_2, v_3, v_4 и v_5 окрашены различными цветами. Пусть $1, 2, 3, 4$ и 5 — номера красок, которыми окрашены вершины v_1, v_2, v_3, v_4 и v_5 соответственно. Предположим для определенности, что вершины v_1, v_2, v_3, v_4 и v_5

расположены по часовой стрелке вокруг вершины v. Начиная с вершины v_1 графа G', будем строить подграф G_{13} графа G' следующим образом : множество вершин подграфа V_{13} графа G_{13} состоит из вершины v_1 и всех вершин графа G', которые могут быть связаны с v_1 путями, проходящими только через вершины с цветом 1 или 3. Предположим сначала, что $v_3 \notin V_{13}$. По построению графа G_{13} граф G' не содержит вершины, которые имеют цвет 1 или 3, не принадлежат множеству V_{13} и являются смежными с вершинами из V_{13}. Поэтому в графе G_{13} цвета 1 и 3 можно поменять местами, не меняя цвета остальных вершин. Теперь вершины v_1 и v_3 окрашены одним цветом и, согласно результатам, полученным ранее, граф G можно раскрасить. Если $v_3 \in V_{13}$, то существует путь из v_1 в $v_3$6, который проходит только через вершины, окрашенные цветом 1 или 3. Назовем этот путь P_{13}. Тогда путь $vv_1P_{13}v_3v$ является циклом и имеет форму, изображенную на рис. 14.67 или рис. 14.68. В первом случае вершина v_2 находится внутри цикла, а вершина v_4 — вне цикла.

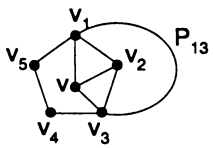


Рис. 14.67

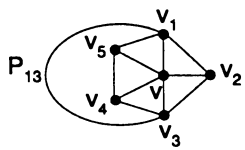


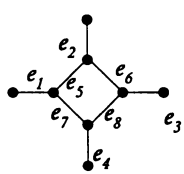
Рис. 14.68

Во втором случае вершина v_4 находится внутри цикла, а вершина v_2 — вне цикла. В каждом из случаев в графе G' не существует путь от вершины v_2 к вершине v_4, проходящий через вершины цвета 2 или 4. Дело в том, что граф G планарный, поэтому его вершины не пересекаются, и переход из вершины v_2 в вершину v_4 требует прохождения через вершину цикла $vv_1P_{13}v_3v$, а все эти вершины окрашены цветом 1 или 3, за исключением, возможно, вершины v, которая не принадлежит графу G'. Поэтому можно повторить изложенную выше процедуру, формируя граф G_{24} вместо G_{13}, и раскрасить вершину v. ■

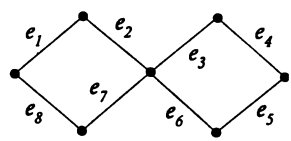
■ УПРАЖНЕНИЯ

1. Найдите графы, двойственные данным :

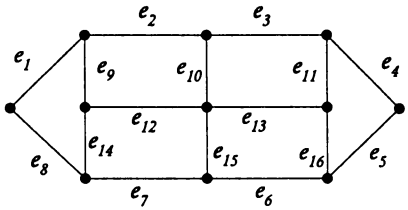
а)



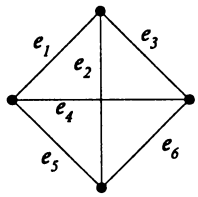
б)



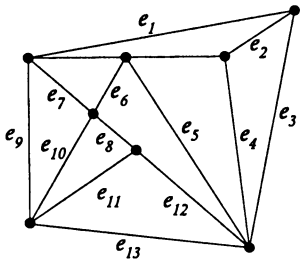
в)



г)

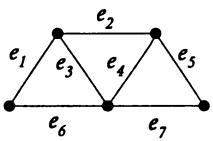


д)

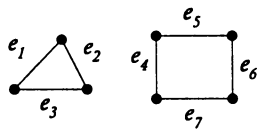


2. Найдите графы, двойственные данным:

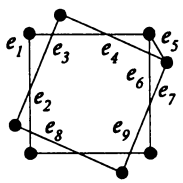
а)



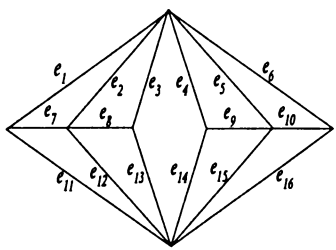
б)



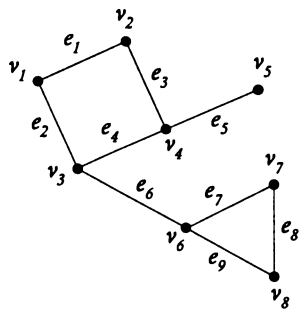
в)



г)

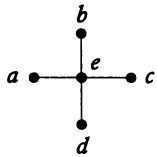


д)

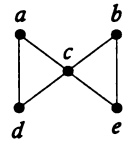


3. Определите хроматическое число приведенных ниже графов.

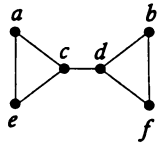
а)



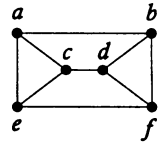
б)



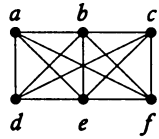
в)



г)

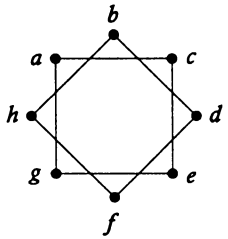


д)

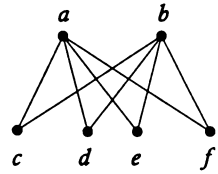


4. Определите хроматическое число приведенных ниже графов.

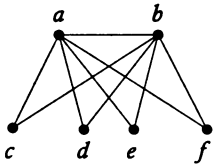
а)



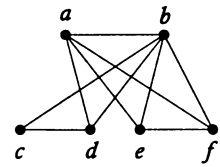
б)



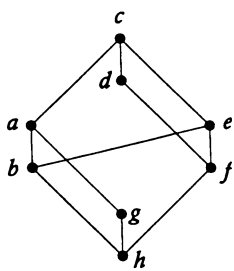
в)



г)

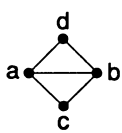


д)

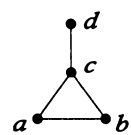


5. Покажите, что двудольный граф $K_{m,n}$ всегда можно раскрасить двумя цветами. Определите хроматический многочлен графа $K_{m,n}$?
6. Чему равно хроматическое число гиперкуба?
7. Используя теорему14.64, найдите хроматический многочлен каждого из приведенных графов:

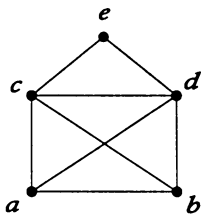
а)



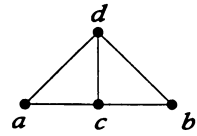
б)



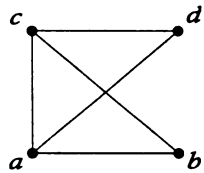
в)



г)

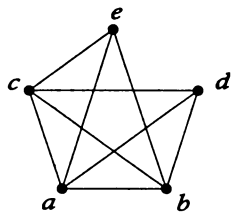


д)

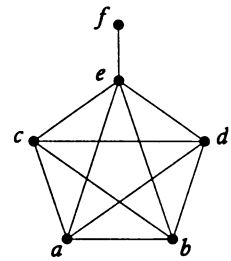


8. Используя теорему 14.64, найдите хроматический многочлен каждого из приведенных графов:

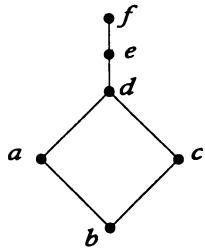
а)



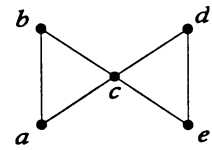
б)



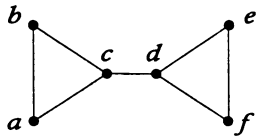
в)



г)



д)



9. Докажите, что если произвольный планарный граф с n вершинами изоморфен своему двойственному графу, то он имеет $2n - 2$ ребер.

10. Докажите, что если хроматическое число связного графа, имеющего, по крайней мере, одно ребро, равно 2, то он является двудольным.

14.4. ПУТИ И ЦИКЛЫ ГАМИЛЬТОНА

В 1857 году математик Уильям Роуэн Гамильтон (William Rowan Hamilton) придумал игру. Существует несколько версий того, как это произошло. По одной из версий он описал эту игру в письме к другу. Согласно другой, он действительно изобрел игру и продал ее производителю игрушек. В любом случае она, по-видимому, включала додекаэдр, т.е. правильный многогранник, 12 граней которого представляли собой конгруэнтные правильные пятиугольники. В каждом из 20 углов, или вершин тела, просверливалась дырка, в которую вставлялся колышек, изображавший город. Используя веревку, требовалось найти путь через города, посетив каждый город один раз, и вернуться в исходный город. Додекаэдр на плоскости изображается так, как показано на рис. 14.69.

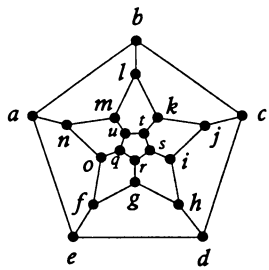


Рис. 14.69

Проблема в таком случае сводится к нахождению цикла в графе, проходящего через каждую вершину, исключая начальную, только один раз. Отсюда любой цикл графа, обладающий таким свойством, называется гамильтоновым циклом. Этот цикл в некотором смысле противоположен эйлерову циклу, который проходит через все ребра только один раз. До определенного момента оба цикла могут показаться похожими, но дальнейшее изучение покажет, что цикл Гамильтона намного сложнее.

Формально путь Гамильтона и цикл Гамильтона (гамильтонов цикл) описываются следующим образом.

ОПРЕДЕЛЕНИЕ 14.68. Пусть G — граф. ***Гамильтонов путь*** — это простой путь, который проходит через каждую вершину графа G. ***Гамильтонов цикл*** — это простой цикл, который проходит через каждую вершину графа G.

Убедимся, что граф для игры Гамильтона действительно имеет гамильтонов цикл, оправдывающий свое название. Один из таких циклов изображен на рис. 14.70.

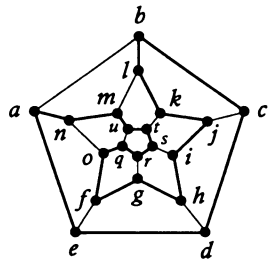


Рис. 14.70

ПРИМЕР 14.69. Граф на рис. 14.71 имеет гамильтонов цикл, изображенный на рис. 14.72.

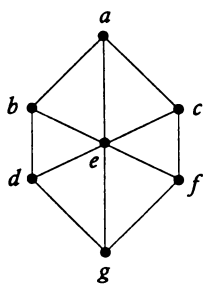


Рис. 14.71

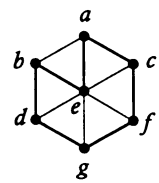


Рис. 14.72

□

ПРИМЕР 14.70. Гиперкуб порядка n при $n \geq 3$ имеет гамильтонов цикл. Его описывет код Грея для n (см. раздел 6.7). Таким образом, гамильтонов цикл для

гиперкуба порядка 4 имеет вид

$$
\begin{array}{cccc}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 \\
1 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 \\
0 & 1 & 1 & 1 \\
1 & 1 & 1 & 1
\end{array}
$$

□

ПРИМЕР 14.71. Полный граф K_n при $n \geq 3$ имеет гамильтонов цикл. Пусть v_1, v_2, $v_3, \ldots, v_n$ — вершины графа K_n. Поскольку между двумя любыми вершинами имеется ребро, то всегда существует ребро из v_i в v_{i+1} и, наконец, существует ребро из последней вершины v_n обратно в v_1. □

Для следующего примера необходима теорема, приведенная ниже.

ТЕОРЕМА 14.72. Для любой вершины из цикла Гамильтона существует ровно два ребра из этого цикла, инцидентные данной вершине.

ДОКАЗАТЕЛЬСТВО. По ходу цикла для каждой вершины V имеется ребро к циклу и ребро из цикла. Если бы существовало еще одно ребро цикла, инцидентное вершине V, то цикл вернулся бы в вершину V, и она опять появилась бы в цикле, что противоречит определению гамильтонова цикла. Следовательно, существует ровно два ребра, которые инцидентны вершине V из цикла Гамильтона. ■

Заметим, что граф, у которого есть гамильтонов цикл, называют гамильтоновым графом. И как следствие приведенной выше теоремы, получаем, что любой граф, содержащий вершину степени 1, не может быть гамильтоновым.

ПРИМЕР 14.73. Граф Петерсена, изображенный на рис. 14.73, имеет гамильтонов путь, но не имеет гамильтонова цикла. Путь показан на рис. 14.74.

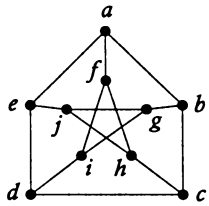


Рис. 14.73

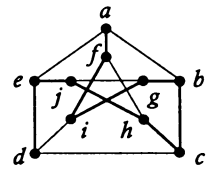


Рис. 14.74

Доказательство, что граф Петерсена не имеет гамильтонова цикла, потребует некоторого количества проб и ошибок. Попытаемся построить гамильтонов цикл. Всякий раз мы будем попадать в тупик. Для начала припомним, что звезда в центре должна соединяться с внешним пятиугольником, поэтому, не нарушая общности, можно предположить, что $\{a, f\}$ — ребро в цикле. Из вершины f цикл должен идти к вершине i или вершине h, но, в силу симметрии, это не имеет значения, поэтому предположим, что $\{f, i\}$ — ребро из цикла. Согласно предыдущей теореме ребро $\{f, h\}$ не может входить в цикл, поскольку тогда будут три ребра, инцидентных вершине f. Значит, ребра $\{j, h\}$ и $\{h, c\}$ должны входить в цикл, поскольку это единственно возможные ребра, инцидентные вершине h. Если ребро $\{i, d\}$ принадлежит циклу и ребро $\{f, i\}$ входит в цикл, получаем, что ребро $\{i, g\}$ не может принадлежать циклу, поскольку согласно предыдущей теореме только два ребра могут быть инцидентными вершине i. Поэтому ребра $\{j, g\}$ и $\{b, g\}$ должны находится в цикле, поскольку только два ребра могут быть инцидентными вершине g. Поскольку ребра $\{j, h\}$ и $\{j, g\}$ принадлежат циклу, ребро $\{j, e\}$ не может находиться в цикле, поскольку тогда будут три ребра, инцидентных вершине j. Поэтому ребра $\{d, e\}$ и $\{e, a\}$ должны входить в цикл, поэтому имеется два ребра, инцидентных вершине e. В результате имеем части цикла, изображенные на рис. 14.75.

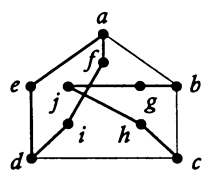


Рис. 14.75

Поскольку ребра $\{i, d\}$ и $\{d, e\}$ находятся в цикле, то ребро $\{d, c\}$ не может входить в цикл, т.к. в противном случае было бы три ребра, инцидентных вершине d, поэтому ребро $\{c, b\}$ должно быть в цикле, чтобы было два ребра, инцидентных вершине c. Поскольку ребра $\{b, g\}$ и $\{c, b\}$ находятся в цикле, то ребро $\{a, b\}$ не может входить в цикл, и в результате получаем граф, изображенный на рис. 14.76, в котором больше нет ребер, чтобы добавить в цикл. Таким образом, этот путь не может породить гамильтонов цикл.

Если вернуться к рассуждениям, согласно которым ребра $\{a, f\}$, $\{f, i\}$, $\{j, h\}$ и $\{h, c\}$ входят в цикл, а ребро $\{f, h\}$ не входит в цикл, поскольку $\{i, d\}$ не может быть в цикле, ребро $\{i, g\}$ должно быть в цикле, поэтому имеется два ребра, инцидентных вершине i. Ребра $\{e, d\}$ и $\{d, c\}$ должны быть в цикле, так что имеется

два ребра, инцидентных вершине d. Поскольку ребра $\{h, c\}$ и $\{d, c\}$ принадлежат циклу, ребро $\{b, c\}$ не может входить в цикл, поскольку тогда будут три ребра, инцидентные вершине c. Поэтому ребра $\{a, b\}$ и $\{g, b\}$ должны быть в цикле, чтобы было два ребра, инцидентных вершине b. Учитывая, что ребра $\{i, g\}$ и $\{g, b\}$ должны быть в цикле, ребро $\{j, g\}$ не может входить в цикл, поскольку тогда будут три ребра, инцидентные вершине g. На этом этапе имеем граф, изображенный на рис. 14.77.

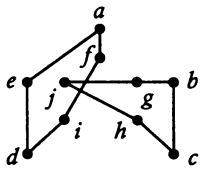


Рис. 14.76

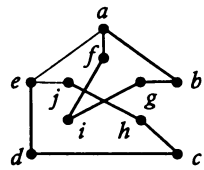


Рис. 14.77

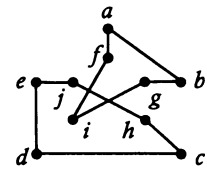


Рис. 14.78

Ребро $\{e, j\}$ должно входить в цикл, чтобы было два ребра, инцидентных вершине j. Поскольку ребра $\{e, d\}$ и $\{e, j\}$ должны входить в цикл, то ребро $\{e, a\}$ не может входить в цикл. В результате получаем граф, изображенный на рис. 14.78, в котором не осталось свободных ребер и который не является гамильтоновым циклом. Поскольку исчерпаны все возможности построения, приходим к выводу, что граф Петерсона не имеет цикла Гамильтона. □

Рассуждения, приведенные выше, убеждают в следующем: чтобы граф имел гамильтонов цикл, степень каждой вершины должна быть не меньше двух. Очевидно также, что граф должен быть связным, чтобы иметь гамильтонов цикл. Приведенная ниже теорема содержит дополнительную информацию по этому вопросу. Доказательство теоремы предоставляется читателю.

ТЕОРЕМА 14.74. Если граф G имеет разрезающее ребро, то он не может иметь гамильтонов цикл. Если компоненты графа, полученные путем удаления разрезающего ребра, имеют гамильтонов цикл, то граф G имеет гамильтонов путь.

Ранее нами были найдены весьма изящные критерии существования у графа эйлерова цикла. К сожалению, никому не удалось установить необходимые и достаточные условия существования у графа гамильтонова цикла. Тем не менее, в следующей теореме приводятся некоторые условия существования у графа гамильтонова цикла. Очевидно, чем больше ребер при фиксированном числе вершин имеет граф, тем выше степень вершин и более вероятно, что имеется цикл, проходящий через все вершины, без повторения вершин. На интуитивном уровне понятно, что если одна вершина имеет более низкую степень, то это компенсируется более высокой степенью другой вершины, что отображено в приведенной ниже теореме.

ТЕОРЕМА 14.75. Если $G(V, E)$ — связный граф с n вершинами, где $n \geq 3$, и для каждой пары различных несмежных вершин $u, v \in V$, $\deg(u) + \deg(v) \geq n$, тогда граф G имеет гамильтонов цикл.

ДОКАЗАТЕЛЬСТВО. Пусть $v_1v_2v_3\dots v_m$ — максимальный простой путь в графе G. Пусть $a = \deg(v_1)$ и $b = \deg(v_m)$. Сначала покажем, что перестановка вершин, входящих в путь, предоставляет возможность формирования простого цикла. Если вершины v_1 и v_m — смежные, то $v_1v_2v_3\dots v_mv_1$ — цикл, и искомый цикл найден. Если v_1 и v_m смежными не являются, то $a + b \geq n$. Мы хотим показать, что в данном случае существуют вершины v_i и v_{i+1}, входящие в путь при условии, что вершина v_1 является смежной с v_{i+1}, а вершина v_m является смежной к v_i. Если это сделано, то начинаем путь с $v_1v_2v_3\dots v_iv_{i+1}\dots v_m$. Удаляем ребро между v_i и v_{i+1}. Затем начинаем новый путь с вершины v_{i+1}, продолжаем его до вершины v_1, поскольку вершины v_{i+1} и v_1 являются смежными. Продолжаем далее и, поскольку вершины v_i и v_m являются смежными, получаем путь $v_{i+1}v_1v_2v_3\dots v_iv_m$, как показано на рис. 14.79.

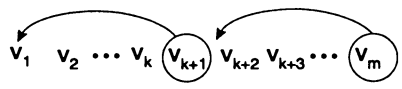


Рис. 14.79

Движемся в обратном направлении вдоль пути $v_mv_{m-1}v_{m-2}\dots v_{i+2}v_{i+1}$ и получаем искомый цикл $v_{i+1}v_1v_2v_3\dots v_iv_mv_{m-1}v_{m-2}\dots v_{i+2}v_{i+1}$.

Теперь требуется показать, что существуют вершины v_i и v_{i+1}, входящие в путь такие, что вершина v_1 — смежная с вершиной v_{i+1}, а вершина v_m — смежная с вершиной v_i. Предположим обратное. В таком случае вершина v_1 не является смежной ни к одной из вершин, которые не входят в путь, поскольку если существует вершина w, не входящая в путь и смежная с вершиной v_1, то $wv_1v_2v_3\dots v_m$ — простой путь, а это противоречит предположению о том, что $v_1v_2v_3\dots v_m$ — максимальный простой путь в графе G, и тогда существует a вершин, входящих в путь $v_1v_2v_3\dots v_m$, смежных с вершиной v_1. Аналогично, существует b вершин, входящих в путь $v_1v_2v_3\dots v_m$, смежных с вершиной v_m. Если включить вершину v_1, то путь $v_1v_2v_3\dots v_m$ будет содержать $a + 1$ вершин, совпадающих с вершиной v_1 или смежных с ней. Существует также b вершин, входящих в путь $v_1v_2v_3\dots v_m$ и смежных с вершиной v_m. Если для каждой вершины v_i, входящей в путь, смежной с вершиной v_m, вершина v_{i+1} не является смежной с вершиной v_1, то путь содержит $a + 1$ вершин, совпадающих с вершиной v_1 или смежных с ней, и b вершин, входящих в путь, которые не совпадают с вершиной v_1 и не смежные с ней. Таким образом, путь $v_1v_2v_3\dots v_m$ содержит $a + b + +1$ различных вершин, что невозможно. Следовательно, сделанное предположение было неверным, и существуют вершины v_i и v_{i+1}, входящие в путь и такие, что вершина v_1 — смежная с вершиной v_{i+1}, а вершина v_m — смежная с вершиной v_i. Таким образом, искомый цикл получен. Предположим для простоты, что вершины переобозначены, так что цикл имеет вид $v_1v_2v_3\dots v_mv_1$.

Покажем теперь, что этот цикл содержит все вершины множества V. Если это не выполняется и вершина v' не совпадает ни с одной из вершин v_i, то, поскольку граф G связный, существует путь из вершины v' в одну из вершин v_i, и существует вершина w, которая не входит в путь $v_1v_2v_3\dots v_m$ и является смежной с одной из вершин v_j. Но тогда

$$wv_jv_{j+1}v_{j+2}\dots v_mv_1v_2v_3\dots v_{j-1}$$

простой путь, который длиннее, чем путь $v_1v_2v_3\ldots v_m$, что является противоречием. Следовательно, цикл $v_1v_2v_3\ldots v_mv_1$ является гамильтоновым циклом. ■

Из теоремы непосредственно вытекает следствие, которое получено раньше и является более известным, чем сама теорема.

СЛЕДСТВИЕ 14.76. Если $G(V, E)$ — связный граф, имеющий n вершин, где $n \geq 3$, и если для каждой вершины $v \in V$ выполняется $\deg(v) \geq \frac{n}{2}$, то граф G имеет гамильтонов цикл.

При доказательстве теоремы 14.75 был использован только тот факт, что сумма степеней вершин v_1 и v_m максимального пути $v_1v_2v_3\ldots v_m$ больше, чем число вершин. Поэтому получаем такое развитие теоремы 14.75.

ТЕОРЕМА 14.77. Пусть $G(V, E)$ — связный граф с $n \geq 3$ вершинами и пусть u и v — несмежные вершины графа G такие, что $\deg(u) + \deg(v) \geq n$. Отсюда граф G^e, состоящий из графа G с присоединенным ребром $e = \{u, v\}$, имеет гамильтонов цикл тогда и только тогда, когда граф G имеет гамильтонов цикл.

ДОКАЗАТЕЛЬСТВО. Если граф G имеет гамильтонов цикл, то и граф G^e имеет гамильтонов цикл, т.к. для него используется тот же самый цикл. Обратно, предположим, что граф G^e имеет гамильтонов цикл. Если ребро e не входит в цикл, то вполне очевидно, что граф G имеет гамильтонов цикл. Если ребро e не принадлежит циклу, то гамильтонов цикл в графе G^e записывается в виде $uv_1v_2v_3\cdots v_mvu$. Но в этом случае $uv_1v_2v_3\cdots v_mv$ — гамильтонов путь в графе G и, следовательно, является максимальным путем, и, кроме того, $\deg(u)+\deg(v) \geq n$. Согласно доказательству предыдущей теоремы вершины этого пути можно переставить так, что они образуют цикл в графе G. Поскольку этот цикл содержит все вершины графа G, то это гамильтонов цикл. ■

Пусть имеется граф G с n вершинами. Добавляем ребра к несмежным вершинам u и v графа G, для которых

$$\deg(u) + \deg(v) \geq n,$$

до тех пор, пока это возможно. По завершении процесса полученный граф назовем замыканием графа G.

ОПРЕДЕЛЕНИЕ 14.78. Пусть G — граф с n вершинами. ***Замыканием*** графа G, обозначаемым $\mathrm{cl}(G)$, называется граф, полученный из графа G рекурсивным добавлением ребер к несмежным вершинам u и v графа G, для которых $\deg(u)+\deg(v) \geq n$ до тех пор, пока это возможно.

Таким образом, $\mathrm{cl}(G)$ обладает свойством, что если в графе $\mathrm{cl}(G)$ есть две несмежные вершины u и v, для которых $\deg(u)+\deg(v) \geq n$, то между этими вершинами существует ребро. Необходимо показать, что $\mathrm{cl}(G)$ определено корректно. Это означает, что если получить граф G' из графа G путем рекурсивного добавления ребер к несмежным вершинам u и v графа G, для которых $\deg(u)+\deg(v) \geq n$, пока это возможно, и граф G'' из графа G путем рекурсивного добавления ребер

к несмежным вершинам u и v графа G, для которых $\deg(u)+\deg(v) \geq n$, пока это возможно, но другим рекурсивным способом, то $G' = G''$. Для доказательства данного утверждения предположим, что $e'_1, e'_2, e'_3, \ldots, e'_l$ — ребра, рекурсивно добавленные к графу G для построения графа G', а $e''_1, e''_2, e''_3, \ldots, e''_m$ — ребра, рекурсивно добавленные к графу G для построения G'' Если ребра не совпадают, то существует, например, ребро в графе G', которое не принадлежит графу G''. Пусть e'_j — первое ребро, рекурсивно добавленное для получения графа G', которое не принадлежит графу G'', и пусть $e'_j = \{u, v\}$. Тогда с ребрами $e'_1, e'_2, e'_3, \ldots, e'_{j-1}$, добавленными к графу G, $\deg(u)+\deg(v) \geq n$. Но поскольку e'_j — первое ребро, рекурсивно добавленное для получения графа G', которое не входит в граф G'', то ребра $e'_1, e'_2, e'_3, \ldots, e'_{j-1}$ входят в граф G''. Следовательно, $\deg(u)+\deg(v) \geq n$ в графе G'', но между вершинами u и v не существует ребро, что и приводит к противоречию. Следовательно, $G' = G''$, поэтому замыкание $\mathrm{cl}(G)$ определено однозначно.

Отметим, что для произвольного графа G $\mathrm{cl}(\mathrm{cl}(G)) = \mathrm{cl}(G)$.

ПРИМЕР 14.79. Если G — граф, изображенный на рис. 14.80, то $\mathrm{cl}(G)$ — граф, изображенный на рис. 14.81.

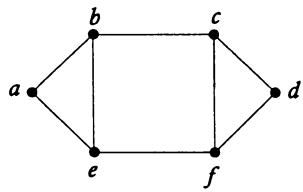


Рис. 14.80

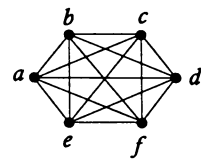


Рис. 14.81 □

ПРИМЕР 14.80. Если G — граф, изображенный на рис. 14.82, то $\mathrm{cl}(G)$ — граф, изображенный на рис. 14.83.

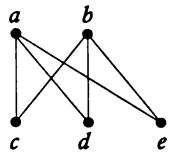


Рис. 14.82

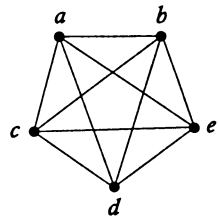


Рис. 14.83 □

ПРИМЕР 14.81. Если G — полный двудольный граф $K_{m,m}$ при $m \geq 1$, то $\mathrm{cl}(G)$ — полный граф K_{m+m}. □

Предлагаем читателю доказать приведенную ниже теорему, используя метод индукции и теорему 14.77.

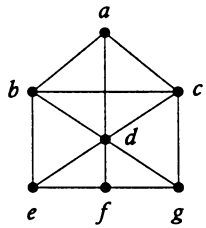
ТЕОРЕМА 14.82. Граф G имеет гамильтонов цикл тогда и только тогда, когда граф $\mathrm{cl}(G)$ имеет гамильтонов цикл.

На основе теоремы и примера, приведенных выше, получаем, что при $m \geq 1$ полный двудольный граф $K_{m,m}$ имеет гамильтонов цикл.

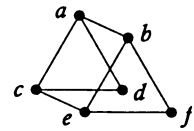
■ УПРАЖНЕНИЯ

1. Найдите гамильтонов цикл, если он существует, для каждого из приведенных ниже графов.

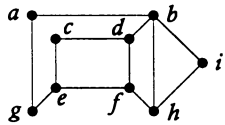
а)



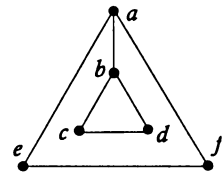
б)



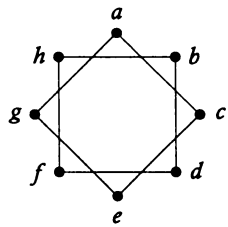
в)



г)

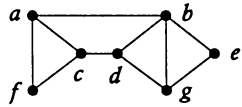


д)

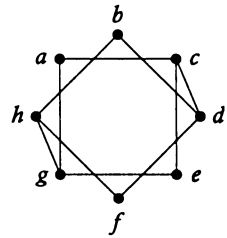


2. Найдите гамильтонов цикл, если он существует, для каждого из приведенных ниже графов.

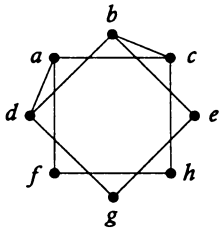
а)



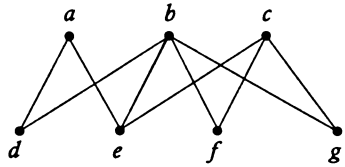
б)



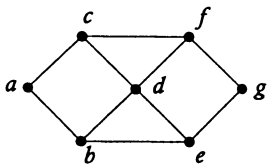
в)



г)

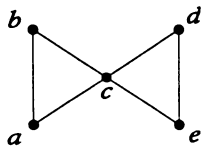


д)

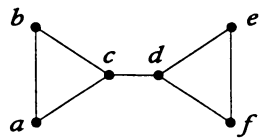


3. Найдите гамильтонов путь, если он существует, для каждого из приведенных ниже графов.

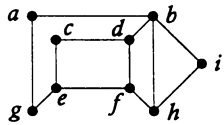
а)



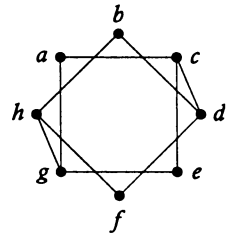
б)



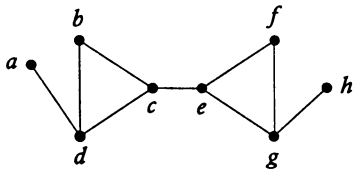
в)



г)

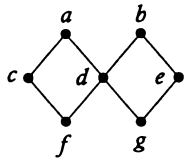


д)

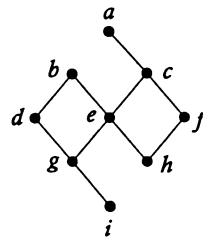


4. Найдите гамильтонов путь, если он существует, для каждого из приведенных ниже графов.

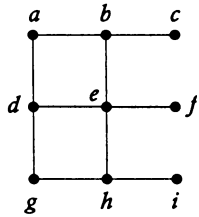
а)



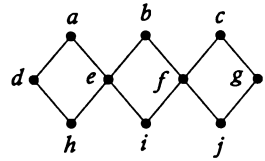
б)



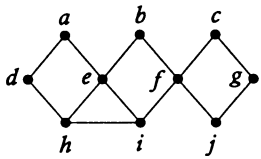
в)



г)



д)



5. Докажите теорему 14.74. Если граф G имеет разрезающее ребро, то граф G не может иметь гамильтонов цикл. Если компоненты графа, полученные при удалении разрезающего ребра, имеют гамильтоновы циклы, то граф G имеет гамильтонов путь.

6. Нарисуйте граф с шестью вершинами, который имеет гамильтонов цикл, но не имеет эйлерова цикла.

7. Нарисуйте граф с шестью вершинами, который имеет эйлеров цикл, но не имеет гамильтонова цикла.

8. Используя теорему 14.77 и метод индукции, докажите теорему 14.82. Граф G имеет гамильтонов цикл тогда и только тогда, когда его замыкание $\mathrm{cl}(G)$ имеет эйлеров цикл.

14.5. ВЗВЕШЕННЫЕ ГРАФЫ И АЛГОРИТМЫ ПОИСКА КРАТЧАЙШЕГО ПУТИ

До сих пор при рассмотрении графов нас интересовали вершины и ребра, по которым можно перемещаться. Теперь нас интересует не только перемещение из точки A в точку B, но и то, как это сделать наилучшим способом. Первый вопрос состоит, конечно, в том, что означает "наилучшим способом". Это может быть самый дешевый путь, самый безопасный путь, кратчайший путь или тот, который требует минимум энергии, или путь, выбранный в соответствии с каким-то иным критерием. Для определения наилучшего пути присвоим каждому ребру вес или меру. Если пытаться найти кратчайшее расстояние между двумя городами, то их необходимо представить в виде вершин, а вес, присвоенный ребрам, — это расстояние между городами. Если пытаться найти самый дешевый способ перелета из одного города в другой, то вес ребер между вершинами, представляющими города, будет стоимостью перелета из города в город. Если прямого перелета между городами нет, то не будет ребра между соответствующими вершинами. Хотя вес или мера, присвоенные ребрам, могут иметь различные значения, для упрощения будем рассматривать вес ребра как расстояние, а наилучший путь из точки A в точку B как кратчайший путь между точками A и B. Это всего лишь вопросы терминологии, и они ни в коей мере не ограничивают использование теории или общность результатов. Тем не менее, одно важное ограничение будет введено. Предположим, что вес или мера, названные теперь расстоянием и приписываемые ребрам между двумя различными точками, положительные. Существуют алгоритмы, которые с некоторыми ограничениями допускают отрицательные значения веса ребер. Например, алгоритм Флойда-Уоршолла, приведенный в данном разделе, можно использовать для ребер с отрицательным весом.

В оставшейся части раздела будем использовать символ ∞. Для упрощения рассмотрения предположим, что все целые числа меньше ∞, так что $\min(a, \infty) = a$ для каждого неотрицательного целого числа a и $\min(\infty, \infty) = \infty$. Примем также, что $a + \infty = \infty + \infty = \infty$. Это — для удобства обозначений.

Для улучшения системы обозначений используется следующее определение.

ОПРЕДЕЛЕНИЕ 14.83. Пусть $A = (a_1, a_2, a_3, \ldots, a_n)$ и $B = (b_1, b_2, b_3, \ldots, b_n)$ — матрицы-строки, где каждое из a_i и b_i — неотрицательные целые числа или ∞. Тогда

$$A \wedge B = (\min(a_1, b_1), \min(a_2, b_2), \min(a_3, b_3), \ldots, \min(a_n, b_n))$$

ОПРЕДЕЛЕНИЕ 14.84. Пусть c — число, а $A = (a_1, a_2, a_3, \ldots, a_n)$ — матрица-строка. Тогда

$$c + A = c + (a_1, a_2, a_3, \ldots, a_n) = (c + a_1, c + a_2, c + a_3, \ldots, c + a_n).$$

Первый алгоритм, который будет описан, называется алгоритмом Дейкстры. Существует несколько версий этого алгоритма. Первым будет приведен оригинальный алгоритм, а затем его улучшенная версия, которая является более эф-

фективной. Кроме того, алгоритм обладает дополнительным свойством, которое позволяет определить не только длину кратчайшего пути, но и сам путь. Это достигается с помощью указателя, который для каждой вершины из кратчайшего пути указывет предыдущую вершину пути. Таким образом, если найдена длина кратчайшего пути между A и B, то двигаясь вдоль кратчайшего пути в обратном направлении от B к A, можно найти сам путь. Далее будет продемонстрирована улучшенная версия алгоритма Дейкстры вместе с возможностью находить путь.

Прежде, чем начать, сформулируем теорему, которая может показаться интуитивно очевидной. Доказательство предоставляется читателю.

ТЕОРЕМА 14.85. Пусть $a = v_1$ и $b = v_n$. Если $v_1, v_2, \ldots, v_i, v_{i+1}, \ldots, v_j, v_{j+1}, \ldots, v_n$ есть кратчайший путь между a и b, то $v_i, v_{i+1}, \ldots, v_j$, часть этого пути между вершинами v_i и v_j, является кратчайшим путем между v_i и v_j.

Начнем с формулировки первого алгоритма Дейкстры, а затем дадим пример его использования. Согласно алгоритму отыскивается кратчайшее расстояние от вершины v_1 к вершине v_n. Начинаем с вершины v_1 и находим расстояние от v_1 до каждой из смежных с ней вершин. Выбираем вершину, расстояние от которой до вершины v_1 наименьшее, пусть это будет вершина v_i. Далее находим расстояние от вершины v_1 до каждой вершины, смежной к v_i вдоль пути, проходящего через вершину v_i. Если это расстояние меньше, чем текущее расстояние, присвоенное каждой из вершин, то заменяем им текущее расстояние. Снова выбираем вершину, ближайшую к v_1, но не совпадающую с выбранной ранее, и процесс повторяется.

Алгоритм Дейкстры(1)

Для данного взвешенного графа алгоритм дает кратчайшее расстояние от вершины v_1 к вершине v_n. Каждой вершине поставим в соответствие упорядоченную пару $(\infty, 0)$. Первая координата вершины $v_i(m, v_r)$ будет означать присвоенное расстояние от вершины v_1 к вершине v_i, а вторая координата — предыдущую вершину пути от v_1 к v_i.

(1) Начать в вершине $v_1(\infty, 0)$, заменить ее на $v_1(0, 0)$ и сделать постоянной. Остальные вершины на этот момент оставить временными.

(2) Когда вершина $v_k(m, v_r)$ станет постоянной, для каждой вершины v_j, смежной к v_k, прибавить величину m к расстоянию от вершины v_k к вершине v_j. Если это значение меньше, чем текущее расстояние, присвоенное вершине v_j, заменить текущее расстояние этой суммой и заменить вторую координату на v_k.

(3) Найти минимум из расстояний, приписанных временным вершинам. Первую из вершин с таким расстоянием делаем постоянной.

(4) Если v_n — не постоянная вершина, то возвращаемся к пункту (2).

(5) Если v_n — постоянная вершина, то расстояние, присвоенное вершине v_n, является кратчайшим расстоянием от v_1 к v_n.

(6) Для нахождения пути начать в вершине v_n, найти предшествующую ей вершину пути (вторая координата). Для каждой вершины пути v_j находить предшествующую ей вершину пути, пока не будет достигнута вершина v_1. Перестановка вершин в обратном порядке даст кратчайший путь.

ПРИМЕР 14.86. Пусть граф, изображенный на рис. 14.84, — взвешенный граф, в котором отыскивается кратчайшее расстояние от вершины A к вершине F. Поставив в соответствие каждой вершине упорядоченную пару $(\infty, 0)$, рассматриваемый граф приводим к виду, показанному на рис. 14.85.

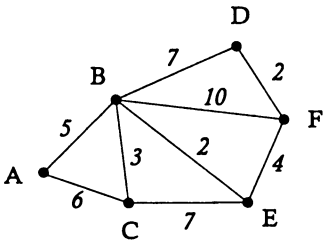


Рис. 14.84

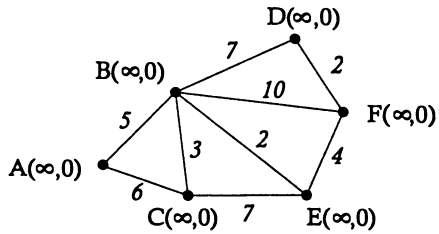


Рис. 14.85

Приступим к построению путей от вершины A к другим вершинам. Первая компонента упорядоченной пары покажет длину кратчайшего пути к вершине в момент достидения, а вторая компонента укажет на предыдущую вершину кратчайшего пути. Первая компонента будет содержать ∞, а вторая — 0 до тех пор, пока путь не найден. Вершина, которая стала постоянной, будет выделена жирным шрифтом. Выполнив шаг 1 алгоритма, получаем граф, изображенный на рис. 14.86.

Поскольку вершины B и C — смежные с вершиной A, выполняем шаг 2 и упорядоченной паре для вершины B присваиваем значение $(5, A)$, а упорядоченной паре для вершины C присваиваем значение $(6, A)$. (Фактичестки, изменения вносятся, тогда и только тогда, когда новые расстояния меньше старых, но поскольку старые расстояния до вершин B и C равны ∞, в данном случае это не имеет значения.) Выполнив шаг 3, выбираем наименьшее из временных присвоенных значений. В данном случае это расстояние до вершины A, равное 5, и вершину $B(5, A)$ делаем постоянной. Таким образом, получаем рис. 14.87.

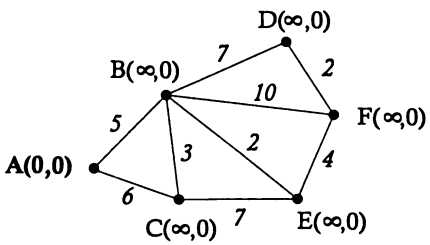


Рис. 14.86

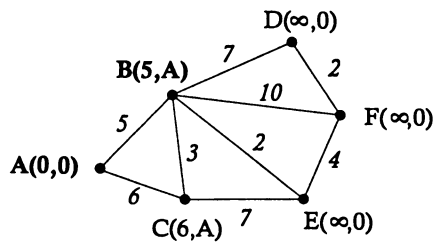


Рис. 14.87

Возвращаясь к шагу 2, рассмотрим временные вершины C, D, E и F, смежные с вершиной B. В каждом случае прибавляем расстояние от вершины A к вершине B к расстоянию от вершины B к данным вершинам. Таким образом, для вершины C это будет $5 + 3 = 8$. Для вершины D имеем $5 + 7 = 12$. Для вершины E имеем $5 + 2 = 7$. Для вершины F получаем $5 + 10 = 15$. Поскольку новое

расстояние до вершины C не меньше, чем уже присвоенное, упорядоченную пару $C(6, A)$ оставляем без изменения. Новые расстояния до вершин D, E и F меньше уже присвоенных, поэтому им задаем значения, которые получены для пути из вершины B, т.е. меняем их на $D(12, B)$, $E(7, B)$ и $F(15, B)$. Выполнив шаг 3, находим наименьшее из расстояний, присвоенных временным вершинам, поэтому берем $\min\{6, 12, 15, 7\} = 6$, и поскольку вершина C имеет это расстояние, делаем вершину $C(6, A)$ постоянной. Таким образом, получаем рис. 14.88.

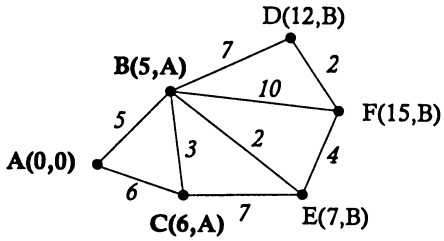


Рис. 14.88 □

Теперь берем новую постоянную вершину C. Выполнение шага 2 не приводит к изменениям. Выполнив шаг 3, делаем вершину $E(7, B)$ постоянной. Получаем в результате рис. 14.89.

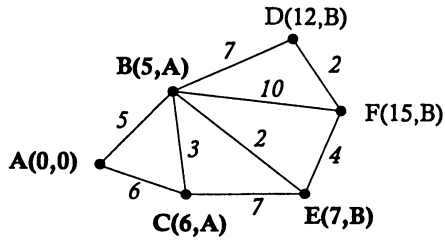


Рис. 14.89

Берем новую постоянную вершину E и, используя шаг 2, меняем $F(15, B)$ на $F(11, E)$. Выполнив шаг 3, делаем вершину $F(11, E)$ постоянной. Таким образом, получаем рис. 14.90.

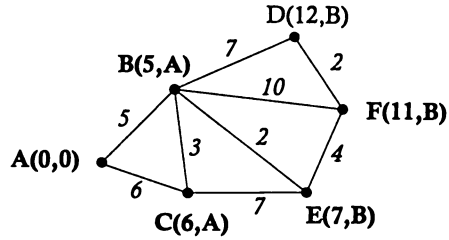


Рис. 14.90

Вершина F стала постоянной, поэтому процесс завершен и 11 — это кратчайшее расстояние от вершины A к вершине F. Если бы совокупность вершин, смежных с постоянной вершиной, была исчерпана до того, как мы достигли вершину F, то задача не имела бы решения, поскольку не было бы пути от вершины A к вершине F. Для нахождения кратчайшего пути заметим, что вершине F предшествует вершина E, вершине E предшествует вершина B, а вершине B предшествует вершина A. Поэтому кратчайшим путем является $ABEF$. □

Сформулируем теперь второй алгоритм Дейкстры, в котором для нахождения кратчайшего расстояния между двумя вершинами использованы матрицы. Пусть n — количество вершин в графе. Матрица W — матрица размерности $n \times n$, в которой значение $W(i,j)$ равно расстоянию между вершинами v_i и v_j, если они смежные, или полагается равным ∞ в противном случае. Обозначим i-ю строку матрицы W через $W(i)$. Используем вспомогательные матрицы P, T, Sum и $pred$ размерности $1 \times n$ и переменную V. Матрица-строка $pred(i)$ содержит индекс вершины, которая предшествует вершине v_i в кратчайшем пути к v_i. Матрица-строка $P(i)$ хранит постоянное расстояние от вершины v_1 к вершине v_i, если вершина v_i выбрана как постоянная вершина. Матрица-строка $T(i)$ содержит значения временного расстояния от вершины v_1 к вершине v_i. Переменная V отслеживает последнюю постоянную вершину. Матрица-строка Sum используется просто для временного хранения значений $P(V) + W(V)$.

Алгоритм Дейкстры(2)

(1) Установить $P(1) = 0$, $T(1) = \infty$ и $W_{j1} = \infty$ при $1 \le j \le 5$. Положить $V = 1$.

(2) Добавить $P(V)$ к каждому элементу $W(V)$. Более точно, пусть $Sum = P(V) + W(V)$. Для $Sum(j) < T(j)$ положить $pred(j) = V$. Далее, положить $T = T \wedge Sum$, и для наименьшего i такого, что $T(i) = \min\{T(j) : 1 \le j \le n\}$ положить $P(i) = T(i)$, $T(i) = \infty$ и $W_{ji} = \infty$ для $1 \le j \le n$. Положить $V = i$.

(3) Если $i \ne n$, вернуться к шагу 2.

(4) Если $i = n$, то $P(i)$ — кратчайшее расстояние от v_1 к v_n.

(5) Для нахождения пути использовать $pred(n)$ и найти вершину, которая предшествует n. Для каждой вершины пути v_j использовать $pred(j)$ и находить предшествующую ей вершину, пока не будет достигнута вершина v_1. Перестановка вершин в обратном порядке даст кратчайший путь.

ПРИМЕР 14.87. Определим кратчайший путь от вершины v_1 к вершине v_5 для графа, изображенного на рис. 14.91.

Продемонстрируем алгоритм Дейкстры с использованием матриц и покажем соответствующий граф. Упорядоченные пары, связанные с вершинами, введем аналогично предыдущему примеру, после чего получим рис. 14.92.

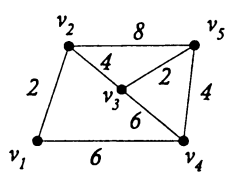


Рис. 14.91

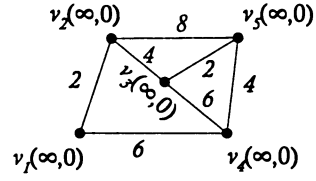


Рис. 14.92

Вместо того, чтобы использовать предыдущую вершину v_i в качестве второй координаты для вершины, принадлежащей пути, для представления вершины v_i

используем только индекс i. Таким образом, вместо записи $v_j(m, v_i)$, где m — присвоенное расстояние от вершины v_1 к v_j, а v_i — предыдущая вершина пути, используем запись $v_j(m, i)$.

Первоначально установим $pred(i) = 0$ для $1 \le i \le 5$, т.к. в данный момент ни одна из вершин не имеет предшественника. Положим $P(1) = 0$, $T(1) = \infty$ и $W_{j1} = \infty$ для $1 \le j \le 5$. Это делает вершину v_1 постоянной вершиной и приводит к ситуации, изображенной на рис. 14.93.

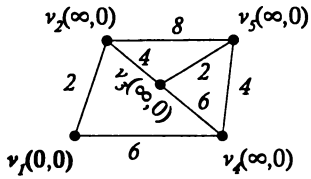


Рис. 14.93

Прибавляем $P(1) = 0$ к каждому элементу матрицы-строки

$$W(1) = (W_{11}, W_{12}, W_{13}, W_{14}, W_{15}) = (\infty, 2, \infty, 6, \infty).$$

Точнее, полагаем $Sum = P(1) + W(1)$. Для $Sum(j) < T(j)$, полагаем $pred(j) = 1$. Затем полагаем $T = T \wedge Sum$, так что $T = (\infty, 2, \infty, 6, \infty)$. Это не что иное, как присваивание расстояния вершинам, смежным с вершиной v_1. Для наименьшего i такого, что $T(i) = \min\{T(j) : 1 \le j \le 5\}$, полагаем $P(i) = T(i)$, $T(i) = \infty$ и $W_{ji} = \infty$ при $1 \le j \le 5$. Этим определяется выбор вершины v_2 как второй постоянной вершины. Полагаем $V = 2$. (Переменная V отслеживает последнюю постоянную вершину.) Таким образом, $T = (\infty, \infty, \infty, 6, \infty)$, $P = (0, 2, \infty, \infty, \infty)$ и $pred = (0, 1, 0, 1, 0)$. Матрица W имеет вид

$$W = \begin{bmatrix} \infty & \infty & \infty & 6 & \infty \\ \infty & \infty & 4 & \infty & 8 \\ \infty & \infty & 0 & 6 & 2 \\ \infty & \infty & 6 & 0 & 4 \\ \infty & \infty & 2 & 4 & 0 \end{bmatrix}.$$

Получаем граф, изображенный на рис. 14.94.

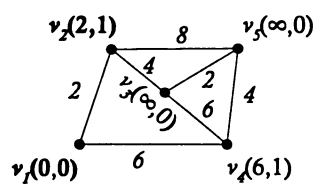


Рис. 14.94

Теперь прибавляем $P(2) = 2$ к каждому элементу матрицы-строки

$$W(2) = (W_{21}, W_{22}, W_{23}, W_{24}, W_{25}) = (\infty, \infty, 4, \infty, 8)$$

Точнее, пусть $Sum = P(2) + W(2)$, что дает расстояние от вершины v_1 к каждой вершине, смежной с вершиной v_2, вдоль пути, проходящего через v_2. Для

$Sum(j) < T(j)$ полагаем $pred(j) = 2$. Далее, пусть $T = T \wedge Sum$, так что $T = (\infty, \infty, 6, 6, 10)$, и для наименьшего i такого, что $T(i) = \min\{T(j) : 1 \le j \le 5\}$, полагаем $P(i) = T(i)$, $T(i) = \infty$ и $W_{ji} = \infty$ при $1 \le j \le 5$. Таким образом, v_3 выбрана как постоянная вершина. Полагаем $V = 3$. В результате $T = (\infty, \infty, \infty, 6, 10)$, $P = (0, 2, 6, \infty, \infty)$ и $pred = (0, 1, 2, 1, 2)$. Матрица W имеет вид

$$W = \begin{bmatrix} \infty & \infty & \infty & 6 & \infty \\ \infty & \infty & \infty & \infty & 8 \\ \infty & \infty & \infty & 6 & 2 \\ \infty & \infty & \infty & 0 & 4 \\ \infty & \infty & \infty & 4 & 0 \end{bmatrix}.$$

Получаем граф, изображенный на рис. 14.95.

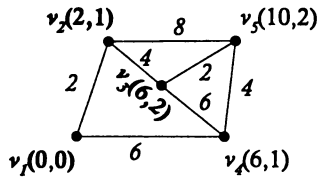


Рис. 14.95

Теперь прибавляем $P(3) = 6$ к каждому элементу матрицы-строки

$$W(3) = (W_{31}, W_{32}, W_{33}, W_{34}, W_{35}) = (\infty, \infty, \infty, 6, 2).$$

Точнее, полагаем $Sum = P(3) + W(3)$, что дает расстояние от вершины v_1 к каждой вершине, смежной с вершиной v_3, вдоль пути, проходящего через вершину v_3. Для $Sum(j) < T(j)$ полагаем $pred(j) = 3$. Далее полагаем $T = T \wedge Sum$, так что $T = (\infty, \infty, \infty, 6, 8)$, и для наименьшего i такого, что $T(i) = \min\{T(j) : 1 \le j \le 5\}$, полагаем $P(i) = T(i)$, $T(i) = \infty$ и $W_{ji} = \infty$ при $1 \le j \le 5$. Это определяет выбор вершины v_4 как постоянной вершины. В результате $T = (\infty, \infty, \infty, \infty, 10)$, $P = (0, 2, 6, 6, \infty)$, $pred = (0, 1, 2, 1, 3)$, а

$$W = \begin{bmatrix} \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & 8 \\ \infty & \infty & \infty & \infty & 2 \\ \infty & \infty & \infty & \infty & 4 \\ \infty & \infty & \infty & \infty & 0 \end{bmatrix}.$$

Имеем граф, изображенный на рис. 14.96.

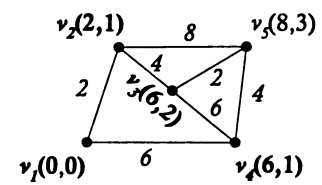


Рис. 14.96

Наконец, прибавляем $P(4) = 6$ к каждому элементу матрицы-строки

$$W(4) = (W_{41}, W_{42}, W_{43}, W_{44}, W_{45}) = (\infty, \infty, \infty, \infty, 4).$$

Точнее, полагаем $Sum = P(4) + W(4)$. Это дает расстояние от вершины v_1 к каждой вершине, смежной с вершиной v_4, вдоль пути, проходящего через вершину v_4. Для $Sum(j) < T(j)$ полагаем $pred(j) = 4$. Далее полагаем $T = T \wedge Sum$, так что $T = (\infty, \infty, \infty, \infty, 8)$, и для наименьшего i такого, что $T(i) = \min\{T(j) : 1 \leq j \leq 5\}$, полагаем $P(i) = T(i)$, $T(i) = \infty$ и $W_{ji} = \infty$ для $1 \leq j \leq 5$. Это определяет выбор вершины v_5 как последней постоянной вершины. В результате имеем $T = (\infty, \infty, \infty, \infty, \infty)$, $P = (0, 2, 6, 6, 8)$, $pred = (0, 1, 2, 1, 3)$, а

$$W = \begin{bmatrix} \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty \end{bmatrix}.$$

Таким обраазом, мы получили граф, изображенный на рис. 14.97, и завершили алгоритм. Длина пути равна 8. Отслеживая путь, как в первом примере, находим, что $v_1v_2v_3v_5$ — кратчайший путь.

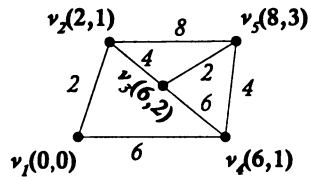


Рис. 14.97 □

Теперь рассмотрим алгоритм Флойда-Уоршолла. Этот алгоритм обладает преимуществом большей эффективности по времени и объему требуемой памяти. Используя алгоритм, также можно найти кратчайшее расстояние между двумя вершинами. Данный алгоритм очень напоминает алгоритм Уоршолла, который использовался для нахождения всех путей в графе путем определения матрицы $\hat{A}$, где $\hat{A}_{ij} = 1$ тогда и только тогда, когда существует путь из вершины v_i в вершину v_j. В предыдущем случае мы отыскивали все пути длины 1. Далее — множество всех 2-путей, в которых серединной вершиной была вершина v_1, и множество комбинировались затем со всеми путями длины 1. Затем отыскивали все пути длины 3 или менее, проходящие через вершину v_1 и/или v_2 и/или v_3 (если такие существовали), и продолжали процесс до тех пор, пока не находили все пути длины n или менее, проходящие через любую из n вершин.

На этот раз, вместо нахождения путей из вершины v_i к вершине v_j, на каждом этапе находим длину пути и сравниваем ее с найденной ранее длиной кратчайшего пути из вершины v_i в v_j (если он существует) и заменяем предыдущее значение в i-ой строке и j-ом столбце новой длиной, если найден более короткий пути из v_i к v_j. Очевидно, что на этот момент имеется две длины. Одна длина определена суммированием длин всех ребер, входящих в путь. Другая — нормальная длина пути, равная числу ребер пути. Надеемся, что удастся избежать путаницы, если говоря о длине ребра, иметь в виду его вес, как это было при рассмотрении алгоритмов Дейкстры.

Ниже приведены два алгоритма вычисления $\hat{A}$. Первый алгоритм более удобен при вычислениях вручную. Второй — для реализации на компьютере. Он приводится в конце раздела.

I-й алгоритм Флойда-Уоршолла

(1) Посмотрим на столбец 1 матрицы A. Если в j-ой строке столбца имеется положительное целое число, добавить его к строке 1, чтобы сформировать A_j^1. Положить A_j равной j-ой строке матрицы A и заменить j-ую строку матрицы A на $A_j^1 \wedge A_j$. Назвать эту матрицу $A^{(1)}$.

(2) Рассмотреть столбец 2 матрицы $A^{(1)}$, построенной на шаге 1. Если в j-ой строке столбца имеется положительное целое число, добавить его к строке 2, чтобы сформировать A_j^2. Положить A_j равной j-ой строке матрицы A и заменить j-ую строку матрицы $A^{(1)}$ на $A_j^2 \wedge A_j$. Назвать эту матрицу $A^{(2)}$.

(3) Рассмотреть столбец 3 матрицы $A^{(2)}$, построенной на шаге 2. Если в j-ой строке столбца имеется положительное целое число, добавить его к строке 3, чтобы сформировать A_j^3. Положить A_j равной j-ой строке матрицы A и заменить j-ую строку матрицы $A^{(2)}$ на $A_j^3 \wedge A_j$. Назвать эту матрицу $A^{(3)}$.

(4) Продолжать процесс, рассматривая следующий, например, i-ый столбец матрицы, построенной на предыдущем шаге. Если в j-ой строке этого столбца имеется положительное целое число, добавить его к соответствующей строке, чтобы сформировать A_j^i. Положить A_j равной j-ой строке матрицы A и заменить j-ю строку матрицы $A^{(i-1)}$ на $A_j^i \wedge A_j$. Назвать эту матрицу $A^{(i)}$.

(5) Продолжать, пока все столбцы не будут проверены.

ПРИМЕР 14.88. Задан граф, изображенный на рис. 14.98.

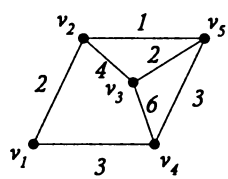


Рис. 14.98

Пусть A — матрица, где A_{ij} — вес ребра $\{v_i, v_j\}$, если ребро существует и ∞ в противном случае. В рассматриваемом случае

$$A = \begin{bmatrix} 0 & 2 & \infty & 3 & \infty \\ 2 & 0 & 4 & \infty & 1 \\ \infty & 4 & 0 & 6 & 2 \\ 3 & \infty & 6 & 0 & 3 \\ \infty & 1 & 2 & 3 & 0 \end{bmatrix}.$$

Найдем вес всех путей длины 2, или 2-путей, в которых средней вершиной является вершина v_1. Начинаем с первого столбца. Находим первую строку в столбце 1, в которой имеется положительное целое число. В данном случае это

число 2 в строке 2. Таким образом, существует ребро из вершины v_2 к вершине v_1 весом 2. Если в строке 1 на позиции $(1, j)$ имеется целое число m, то существует ребро из вершины v_1 к вершине v_j весом m. Тогда $2 + m$ — вес 2-пути из вершины v_2 в вершину v_j. Таким образом, если прибавить 2 к каждому элементу в строке 1, то получим строку, которую следует рассматривать как матрицу-строку, обозначая ее $A_2^1 = (2, 4, \infty, 5, \infty)$, так что $A_2^1(j)$ — вес пути длины 2, или 2-пути из вершины v_2 к вершине v_j. Если рассматривать $A_2 = (2, 0, 4, \infty, 1)$, строку 2 матрицы A как матрицу-строку, тогда $A_2(j)$ — вес пути длины 1 из вершины v_2 к вершине v_j. Поскольку нас интересует кратчайший путь для каждого j, заменяем строку 2 матрицы A на $A_2^1 \wedge A_2 = (2, 4, \infty, 5, \infty) \wedge (2, 0, 4, \infty, 1) = (2, 0, 4, 5, 1)$. Аналогичным образом, в строке 4 первого столбца находится число 3. Таким образом, существует ребро из вершины v_4 к вершине v_1 весом 3. Если в строке 1 на позиции $(1, j)$ имеется целое число m, то существует ребро из вершины v_4 к вершине v_j весом m. Тогда $3 + m$ — вес 2-пути из вершины v_4 к вершине v_j. Следовательно, если прибавить 3 к каждому элементу в строке 1, то получим строку, которую следует рассматривать как матрицу-строку, обозначая ее $A_4^1 = (3, 5, \infty, 6, \infty)$, так что $A_4^1(j)$ — вес 2-пути из v_4 к v_j. Если рассматривать $A_4 = (3, \infty, 6, 0, 3)$, строку 4 матрицы A как матрицу-строку, тогда $A_4(j)$ — вес пути длины 1, или 1-пути из вершины v_4 к вершине v_j. Поскольку нас интересует кратчайший путь для каждого j, заменяем строку 4 матрицы A на $A_4^1 \wedge A_4 = (3, 5, \infty, 6, \infty) \wedge (3, \infty, 6, 0, 3) = (3, 5, 6, 0, 3)$. И т.к. в других строках столбца 1 больше нет положительных целых чисел, первый этап завершен, в результате которого имеем

$$A^{(1)} = \begin{bmatrix} 0 & 2 & \infty & 3 & \infty \\ 2 & 0 & 4 & 5 & 1 \\ \infty & 4 & 0 & 6 & 2 \\ 3 & 5 & 6 & 0 & 3 \\ \infty & 1 & 2 & 3 & 0 \end{bmatrix}.$$

Теперь используем второй столбец матрицы $A^{(1)}$. Если в строке i столбца 2 имеется число k, то существует ребро из вершины v_i к вершине v_2 весом k. Если в строке j на позиции $(2, j)$ имеется целое число m, то существует ребро из вершины v_2 к вершине v_j весом m. Тогда $k + m$ — вес пути из v_i к v_j. Следовательно, если прибавить число k к каждому элементу в строке 2, то получим строку, которую следует рассматривать как матрицу-строку, обозначая ее A_i^2, так что $A_i^2(j)$ — длина 2-пути из вершины v_i к вершине v_j. Если рассматривать A_i, строку i матрицы $A^{(1)}$ как матрицу-строку, тогда $A_i(j)$ — длина пути из вершины v_i к вершине v_j. Поскольку нас интересует кратчайший путь для всех j, заменяем строку i матрицы $A^{(1)}$ на $A_i^2 \wedge A_i$. Во втором столбце имеется число 2 в строке 1, число 4 — в строке 3, число 5 — в строке 4 и число 1 — в строке 5. Используя этот процесс для каждой из указанных строк, получаем

$$A^{(2)} = \begin{bmatrix} 0 & 2 & 6 & 3 & 3 \\ 2 & 0 & 4 & 5 & 1 \\ 6 & 4 & 0 & 6 & 2 \\ 3 & 5 & 6 & 0 & 3 \\ 3 & 1 & 2 & 3 & 0 \end{bmatrix}.$$

Рассмотрим теперь третий столбец матрицы $A^{(2)}$. Для каждой позиции $(3, j)$ матрицы $A^{(2)}$, на которой имеется положительное целое число, прибавляем это число к третьей строке матрицы $A^{(2)}$, получая матрицу-строку A_j^3, после чего положим A_j равной j-ой строке матрицы $A^{(2)}$. Далее заменяем j-ю строку матрицы $A^{(2)}$ на $A_j^3 \wedge A_j$ и получаем

$$A^{(3)} = \begin{bmatrix} 0 & 2 & 6 & 3 & 3 \\ 2 & 0 & 4 & 5 & 1 \\ 6 & 4 & 0 & 6 & 2 \\ 3 & 5 & 6 & 0 & 3 \\ 3 & 1 & 2 & 3 & 0 \end{bmatrix}.$$

Теперь рассмотрим четвертый столбец матрицы $A^{(3)}$. Для каждого положительного элемента на позиции $(4, j)$ матрицы $A^{(3)}$ добавляем это значение к четвертой строке матрицы $A^{(3)}$, чтобы получить A_j^4 и положить A_j, равной j-ой строке матрицы $A^{(3)}$. Далее заменяем j-ю строку матрицы $A^{(3)}$ на $A_j^4 \wedge A_j$ и получаем

$$A^{(4)} = \begin{bmatrix} 0 & 2 & 6 & 3 & 3 \\ 2 & 0 & 4 & 5 & 1 \\ 6 & 4 & 0 & 6 & 2 \\ 3 & 5 & 6 & 0 & 3 \\ 3 & 1 & 2 & 3 & 0 \end{bmatrix}.$$

Наконец, рассмотрим пятый столбец матрицы $A^{(4)}$. Для каждого положительного элемента на позиции $(5, j)$ матрицы $A^{(4)}$ прибавляем это значение к пятой строке матрицы $A^{(4)}$, получая A_j^5, и полагаем A_j, равной j-ой строке матрицы $A^{(4)}$. Далее заменяем j-ю строку матрицы $A^{(4)}$ на $A_j^5 \wedge A_j$ и получаем

$$A^{(5)} = \begin{bmatrix} 0 & 2 & 5 & 3 & 3 \\ 2 & 0 & 3 & 4 & 1 \\ 5 & 3 & 0 & 5 & 2 \\ 3 & 4 & 5 & 0 & 3 \\ 3 & 1 & 2 & 3 & 0 \end{bmatrix}.$$

□

Второй метод использует то же самый процесс, записанный в псевдокоде, включая соглашения относительно ∞, принятые выше. Единственная разница состоит в том, что в отличие от метода, изложенного выше, в данной ситуации допускается, что A_{ik} может не быть положительным целым числом.

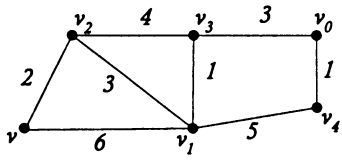
II-й алгоритм Флойда-Уоршолла

Цикл по k от 1 до n:
    Цикл по i от 1 до n:
        Цикл по j от 1 до n:
            $A_{ij} = A_{ij} \wedge (A_{ik} + A_{kj})$.

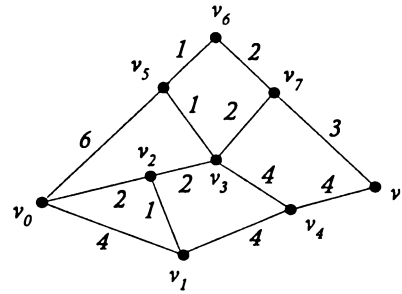
■ УПРАЖНЕНИЯ

1. Используйте алгоритм Дейкстры для нахождения в приведенных ниже графах кратчайшего расстояния между вершинами v_0 и v.

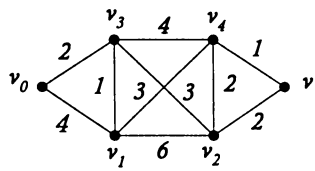
а)



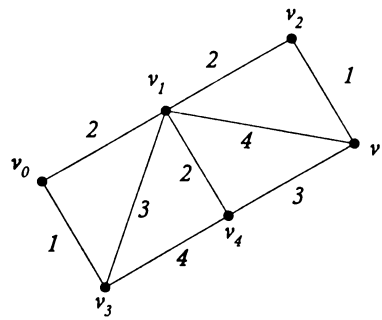
б)



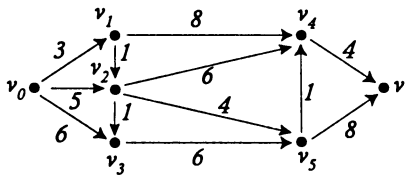
в)



г)



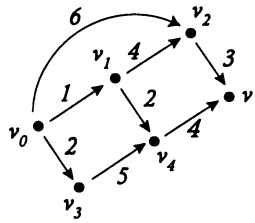
д)



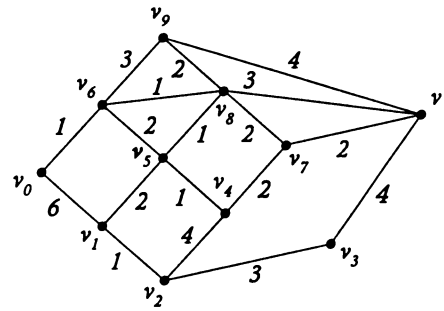
2. Используйте алгоритм Флойда-Уоршолла для нахождения кратчайшего расстояния между вершинами каждого графа из упражнения (1).

3. Используйте алгоритм Дейкстры для нахождения в приведенных ниже графах кратчайшего расстояния между вершинами v_0 и v.

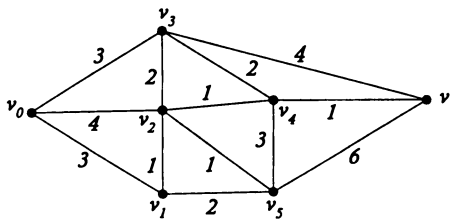
а)



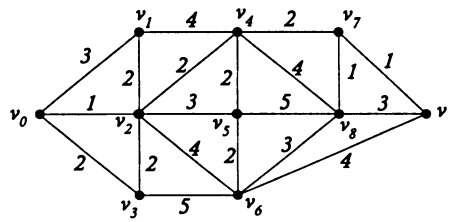
б)



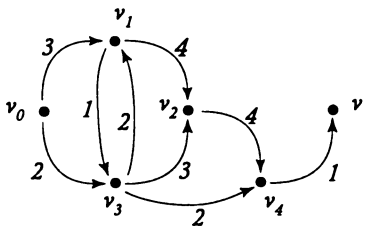
в)



г)



д)



4. Используйте алгоритм Флойда-Уоршолла для нахождения кратчайшего расстояния между вершинами каждого графа из упражнения 3.

ГЛАВА

15

ДЕРЕВЬЯ

15.1. СВОЙСТВА ДЕРЕВЬЕВ

В разделе 6.3 дерево определяется как связный граф без циклов. Следующая теорема обобщает свойства деревьев, приведенные в разделе 6.3.

ТЕОРЕМА 15.1. Следующие утверждения эквивалентны.

а) Граф G — дерево.
б) Граф G — связный и $v = e+1$, где v — количество вершин, а e — количество ребер графа G.
в) Для каждой пары различных вершин a и b существует единственный путь из a в b.
г) Граф G — ацикличный (не имеет циклов) и $v = e + 1$.

ДОКАЗАТЕЛЬСТВО. Эквивалентность пунктов (а) и (б) доказана в теоремах 6.37 и 6.38 раздела 6.3. Эквивалентность пунктов (а) и (в) доказана в теоремах 6.40 и 6.41. Пункт (г) есть прямое следствие пунктов (а) и (б). Для доказательства того, что утверждение (а) следует из утверждения (г), предположим, что граф G — ацикличный. В этом случае все компоненты $G_1, G_2, G_3, \ldots, G_m$ графа G — деревья. Для каждого G_i имеем $v_i = e_i + 1$, где v_i — количество вершин, а e_i — количество ребер графа G_i. Поэтому $\sum_{i=1}^{m} v_i = \sum_{i=1}^{m} e_i + m$. Но $v = \sum_{i=1}^{m} v_i$ и $e = \sum_{i=1}^{m} e_i$, так что $v = e + m$. Следовательно, $m = 1$, поэтому есть только одна компонента, и граф G — связный. Таким образом, граф G — дерево. ■

В главе 6 определено, что ***ориентированное*** T — это ориентированный граф без петель, соотнесенный граф которого является деревом, так что если существует путь из вершины a в вершину b, то он единственный. Ориентированное дерево T является ***корневым ориентированным деревом***, если существует единственная вершина v_0 такая, что $\text{indeg}(v_0) = 0$, и существует путь из вершины v_0 в каждую другую вершину дерева T. Как и в случае деревьев, представим несколько эквивалентных описаний корневых ориентированных деревьев. В приведенном ниже доказательстве, если p и q обозначают пути, то pq обозначает путь p, за которым следует путь q.

ТЕОРЕМА 15.2. Для ориентированного графа G следующие утверждения эквивалентны.

а) G — корневое ориентированное дерево.

б) G имеет единственный такой элемент v_0, что для любой вершины a графа G существует единственный ориентированный путь из v_0 в a.

в) Соотнесенный граф графа G связен, и G содержит единственный элемент v' такой, что $\text{indeg}(v') = 0$, и для любой другой вершины a графа G имеем, $\text{indeg}(a) = 1$.

г) Соотнесенный граф графа G связен, и G содержит единственный элемент v_0 такой, что для любой вершины a графа G существует единственный путь из v_0 в a.

ДОКАЗАТЕЛЬСТВО. (а) $\rightarrow$ (б) Поскольку G — корневое ориентированное дерево, то существует ориентированный путь из корня v_0 в любую заданную вершину a, и в силу того, что G — ориентированное дерево, этот путь единственный. Корень только один, и если вершина v_0' обладает тем же свойством, тогда существует путь p из v_0 в v_0' и путь p' из v_0' в v_0. Но в таком случае путь $pp'p$ — еще один путь из v_0 в v_0', и между v_0 и v_0' существует более чем один путь.

(б) $\rightarrow$ (в) Пусть v' — тот единственный элемент v_0, о котором идет речь в пункте (б). Тогда $\text{indeg}(v') = 0$, но если $\text{indeg}(v') \neq 0$, для некоторой вершины a существует ребро (a, v'). Отсюда существует также путь p из v' в a, и $p(a, v')p$ также является путем из p в a, поэтому между p и a существует более одного пути. Вершина v' — единственная вершина, степень входа которой равна 0. Если существует другая вершина v, у которой $\text{indeg}(v) = 0$, тогда между v' и v не существовало бы пути, что противоречит пункту (б). Каждая другая вершина имеет степень входа 1, потому что если вершина v имеет степень входа больше чем 1, тогда для двух различных вершин a и b существует ориентированное ребро из a в v и ориентированное ребро из b в v. Но существует путь p из v' в a и путь q из v' в b, так что $p(a, v)$ и $q(b, v)$ — различные пути из v' в v, что противоречит пункту (б). Соотнесенный граф графа G связный, поскольку для заданных вершин a и b существует путь из v' в a и из v' в b, и, следовательно, в соотнесенном графе существует путь из a в v' и путь из v' в b, т.е. существует путь из a в b.

(в) $\rightarrow$ (г) Очевидно, $v_0 = v'$. Согласно пункту (в) для любой вершины графа a существует путь из v_0 в a. Предположим, что для некоторой вершины a существует два пути из v_0 в a. Пусть b — первая вершина, начиная с которой оба пути из b в a совпадают. Ею может быть и сама вершина a. В этом случае $\text{indeg}(b) > 1$, что противоречит пункту (в).

(г) $\rightarrow$ (а) Предположим, что соотнесенный граф G' содержит два неориентированных пути из вершины v^* в вершину a. Если оба пути ориентированы из v^* в a, это противоречит гипотезе. Если один путь есть ориентированный путь p из a в v^*, а другой путь — ориентированный путь q из v^* в a, то qpq — также ориентированный путь из v^* в a, что противоречит условию единственности пути. Остается только предположить, что один из связных путей из v^* в a не является ориентированным путем. В этом случае в одном из путей для некоторой вершины c существует ребро (b, c) и ребро (d, c). Но поскольку существует ориентированный путь p' из v^* в b и ориентированный путь p'' из v^* в d, существуют

различные пути $p'(b,c)$ и $p''(d,c)$ из v^* в c, что опять является противоречием. Поэтому неориентированный путь из вершины v^* в вершину a единственный. Неориентированный путь из любой вершины a в любую вершину b должен быть единственным, так как если бы существовали два пути из a в b, то существовали бы два различных пути из v^* в a и в b или, другими словами, два различных пути из v^* в b, что неверно. Поэтому G — ориентированное дерево и, поскольку v^* удовлетворяет определению корня, G — корневое ориентированное дерево. ■

В оставшейся части книги рассматриваются только корневые ориентированные деревья; следовательно, под всеми ориентированными деревьями следует понимать корневые ориентированные деревья.

Некоторые из приведенных ниже определений для удобства изложения повторены. Обращаем внимание читателя, что во многих учебниках эти определения сформулированы по-другому.

ОПРЕДЕЛЕНИЕ 15.3. В ориентированном дереве ***уровень*** вершины v — это длина пути от корня дерева до вершины v. ***Высота*** ориентированного дерева — это длина самого длинного пути от корня до листа.

ОПРЕДЕЛЕНИЕ 15.4. ***m-арным ориентированным деревом*** называется такое ориентированное дерево, в котором $\text{outdeg}(v) \leq m$ для каждой его вершины v. Таким образом, родитель имеет не более m сыновей. ***Полным m-арным ориентированным деревом*** называется такое ориентированное дерево, в котором $\text{outdeg}(v) = m$ для каждой его вершины v, не являющейся листом, и каждый лист находится на одном и том же уровне. Таким образом, каждый родитель имеет в точности m сыновей.

ОПРЕДЕЛЕНИЕ 15.5. m-арное ориентированное деревом высоты h называется ***сбалансированным*** (***полным***, или ***почти полным***), если уровень каждого листа равен h или $h-1$.

ТЕОРЕМА 15.6. Если полное m-арное ориентированное дерево имеет n вершин и i внутренних вершин, то $n = mi + 1$. Решая относительно i, получаем $i = \dfrac{n-1}{m}$.

ДОКАЗАТЕЛЬСТВО. Каждая вершина, за исключением корня, является сыном внутренней вершины. Поскольку имеется i внутренних вершин и каждая внутренняя вершина имеет m сыновей, всего имеется im сыновей. Если учесть корень, то общее число вершин равно $mi + 1$. ■

ТЕОРЕМА 15.7. Если полное m-арное ориентированное дерево имеет n вершин, i внутренних вершин и l листьев, то $l = (m-1)i + 1$. Решая относительно i, получаем $i = \dfrac{l-1}{m-1}$.

ДОКАЗАТЕЛЬСТВО. По предыдущей теореме число вершин равно $n = mi + 1$. Вычитая i внутренних вершин из $n = mi + 1$, получаем $l = n - i = mi + 1 - i = (m-1)i + 1$ листьев. ■

ТЕОРЕМА 15.8. Полное m-арное ориентированное дерево высоты h имеет $\frac{m^{h+1}-1}{m-1}$ вершин и m^h листьев. В частности, полное бинарное ориентированное дерево высоты h имеет $2^{h+1}-1$ вершин и 2^h листьев.

ДОКАЗАТЕЛЬСТВО. Поскольку число вершин на каждом уровне в m раз больше, чем на предыдущем, это приводит к геометрической прогрессии, поэтому m-арное ориентированное дерево высоты h имеет

$$1 + m + m^2 + m^3 + \cdots + m^{h-1} + m^h = \frac{m^{h+1}-1}{m-1}$$

вершин. Поскольку листьями являются только те вершины, которые находятся на уровне h, их насчитывается m^h. ■

Доказательство следующей теоремы предоставляется читателю.

ТЕОРЕМА 15.9.

а) Если полное m-арное дерево высоты h имеет l листьев, то $h = \log_m(l)$.
б) Если m-арное дерево высоты h имеет l листьев, то $h \geq \log_m(l)$.
в) Если полное бинарное дерево высоты h имеет v вершин, то $h = \log_2(v+1)-1$.
г) Если бинарное дерево высоты h имеет v вершин, то $h \geq \log_2(v+1)-1$.

Напомним, что функция f из графа $G(V,E)$ в граф $G'(V',E')$ называется ***гомоморфизмом*** из G в G', обозначается $f : G \to G'$, если она обладает следующими свойствами.

ОПРЕДЕЛЕНИЕ 15.10.

а) Если $e \in E$, то $f(e) \in E'$ $(f(E) \subseteq E')$.
б) Если $v \in V$, то $f(v) \in V'$ $(f(V) \subseteq V')$.
в) Если вершины u и v инцидентны ребру e в G, то $f(u)$ и $f(v)$ инцидентны ребру $f(e)$ в G'.

Напомним также, что гомоморфизм $f : G \to G'$ является ***изоморфизмом***, если $f : V \to V'$ и $f : E \to E'$ являются взаимно однозначными соответствиями. Если $f : G \to G'$ изоморфизм, то говорят, что G и G' ***изоморфны***.

ОПРЕДЕЛЕНИЕ 15.11. Два корневых бинарных дерева $T(E,V)$ и $T'(E',V')$ ***изоморфны***, если существует изоморфизм f из T в T' такой, что

а) v_i — левый сын вершины v_j тогда и только тогда, когда $f(v_i)$ — левый сын вершины $f(v_j)$.
б) v_i — правый сын вершины v_j тогда и только тогда, когда $f(v_i)$ — правый сын вершины $f(v_j)$.
в) f отображает корень r дерева T в корень r' дерева T'.

Любителям заниматься подсчетами предлагаем следующую теорему.

ТЕОРЕМА 15.12. Число неизоморфных корневых бинарных деревьев с n вершинами равно числу Каталана C_n.

ДОКАЗАТЕЛЬСТВО. Пусть I_n — число изоморфных корневых бинарных деревьев с n вершинами. Если $n = 0$, имеем пустое дерево и полагаем $I_0 = 1$. Если $n = 1$, имеем одну вершину и, очевидно, только одно дерево, поэтому $I_1 = 1$. Если $n = 2$, имеем два корневых бинарных дерева, изображенных на рис. 15.1. Если $n = 3$, имеем пять корневых бинарных дерева, изображенных на рис. 15.2.

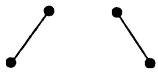


Рис. 15.1

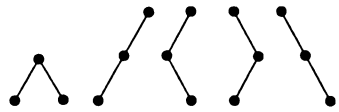


Рис. 15.2

Предположим, что $n > 3$, и выберем такое k, что $1 \le k \le n$. Пусть T_n обозначает дерево с n вершинами и пусть $T_{n,k}$ — дерево с n вершинами, определенное следующим образом. Поскольку одна вершина является корнем, предположим, что имеются правое поддерево T_{k-1} с $k-1$ вершиной и левое поддерево T_{n-k} с $n-k$ вершинами, как показано на рис. 15.3.

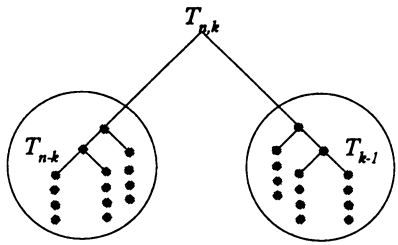


Рис. 15.3

Тогда по определению число способов, которыми можно построить дерево T_{k-1}, равно I_{k-1}, а число способов, которыми можно построить дерево T_{n-k}, равно I_{n-k}. Таким образом, число способов построения $T_{n,k}$ равно $I_{k-1} \cdot I_{n-k}$. Суммируя по k, получаем число всевозможных способов построения дерева T_n. Итак,

$$I_n = \sum_{k=1}^{n} I_{k-1} \cdot I_{n-k},$$

что совпадает с определением числа Каталана C_n. Поэтому

$$I_n = C_n = \frac{1}{n+1} \cdot \frac{(2n)!}{n! \cdot n!}.$$

■

■ УПРАЖНЕНИЯ

1. Найдите неизоморфные корневые бинарные деревья с n вершинами: когда

а) $n = 2$; **б)** $n = 3$; **в)** $n = 4$.

2. Сколько существует неизоморфных корневых бинарных деревьев с n вершинами:

а) $n = 5$; **б)** $n = 6$; **в)** $n = 8$;

г) $n = 10$; **д)** $n = 20$.

3. Сколько в полном m-арном дереве высоты h имеется листьев, вершин и внутренних вершин при условии:

а) $m = 2$ и $h = 5$;
б) $m = 3$ и $h = 4$;
в) $m = 2$ и $h = 8$;
г) $m = 4$ и $h = 3$;
д) $m = 1$ и $h = 10$.

4. Для дерева, изображенного на рис. 15.4, определите:

(i) высоту корневого дерева;
(ii) уровень вершины e;
(iii) уровень вершины g;
(iv) уровень вершины a;
(v) какая вершина является родителем i;
(vi) какие вершины являются сыновьями вершины b

а) если корнем выбрана вершина d;
б) если корнем выбрана вершина f;
в) если корнем выбрана вершина c;
г) если корнем выбрана вершина j;
д) если корнем выбрана вершина b.

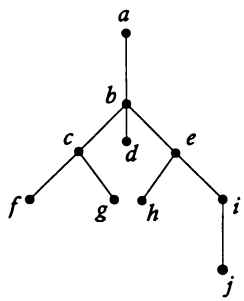


Рис. 15.4

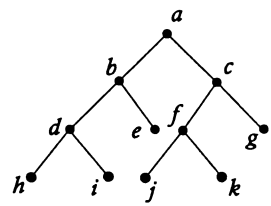


Рис. 15.5

5. Для дерева, изображенного на рис. 15.5, определите:

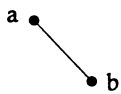
(i) высоту корневого дерева;
(ii) уровень вершины e;
(iii) уровень вершины g;
(iv) уровень вершины a;
(v) какая вершина является родителем i;
(vi) какие вершины являются сыновьями вершины b

а) если корнем выбрана вершина d;
б) если корнем выбрана вершина f;

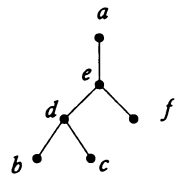
в) если корнем выбрана вершина c;
г) если корнем выбрана вершина j;
д) если корнем выбрана вершина b.

6. Какие из приведенных ниже деревьев являются сбалансированными? Какие из них являются полными?

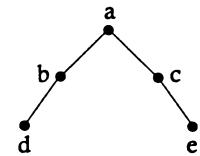
а)



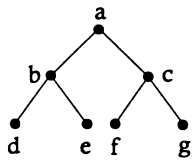
б)



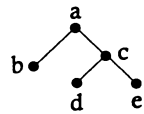
в)



г)

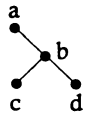


д)

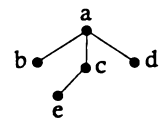


7. Какие из приведенных ниже деревьев являются сбалансированными? Какие из них являются полными?

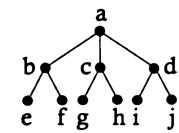
а)



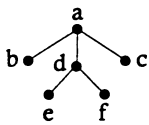
б)



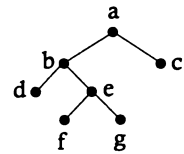
в)



г)



д)



8. При условии, что полное бинарное дерево имеет 32 листа, определите:
а) высоту этого дерева;
б) количество вершин.

9. При условии, что полное бинарное дерево имеет 128 листьев, определите:
а) высоту этого дерева;
б) количество вершин.

10. Докажите, что
а) для полного m-арного дерева высоты h с l листьями $h = \log_m(l)$;
б) для полного m-арного дерева высоты h с l листьями $h \geq \log_m(l)$;
в) для полного бинарного дерева высоты h с числом вершин, равным v, $h = \log_2(v+1) - 1$.

11. Докажите теорему 15.9

а) для полного m-арного дерева высоты h с l листьями $h = \log_m(l)$.

б) для полного m-арного дерева высоты h с l листьями $h \geq \log_m(l)$.

в) для полного бинарного дерева высоты h с числом вершин, равным v, $h = \log_2(v+1) - 1$.

г) для бинарного дерева высоты h, имеющего v вершин, $h \geq \log_2(v+1) - 1$.

12. Докажите, что дерево с двумя или более вершинами является двудольным графом.

13. Докажите, что связный граф является деревом тогда и только тогда, когда добавление нового ребра между вершинами графа всегда приводит к возникновению цикла.

14. Докажите или опровергните, что связный граф является деревом тогда и только тогда, когда каждое его ребро является разрезающим ребром.

15. Докажите или опровергните, что связный граф является деревом тогда и только тогда, когда каждая его вершина является разрезающей вершиной.

16. Какие вершины в корневом дереве являются разрезающими?

17. Пусть G — связный граф. Сформулируйте необходимые и достаточные условия того, что удаление из графа G любого ребра превращает его в дерево.

18. Пусть $T(V, E)$ — корневое дерево. Определим отношение $\leq$ на V как $a \leq b$, если $a = b$ или a — потомок b. Покажите, что отношение $\leq$ есть частичный порядок. Определим произведение $a \cdot b = \text{нвг}(a, b)$, так что $(V, \cdot)$ — верхняя полурешетка. Опишите $a \cdot b$ в терминах путей из вершин a и b в корень.

19. Пусть $T(V, E)$ — корневое дерево. На V определено умножение, как в предыдущем упражнении. Для заданной вершины a из V определим $\hat{a} : V \to V$ как $\hat{a}(v) = av$. Докажите, что $\hat{a}$ — гомоморфизм полугрупп.

15.2. БИНАРНЫЕ ДЕРЕВЬЯ ПОИСКА

Бинарное корневое дерево, которое назовем просто бинарным деревом, обеспечивает прекрасный метод организации данных, при котором любые конкретные данные можно легко найти или установить их отсутствие. Очевидно, что самый неэффективный способ поиска — последовательный просмотр всех данных, так как если необходимые данные отсутствуют, то для установления этого факта нужно просмотреть весь список. Бинарное дерево поиска позволяет избежать этого. Единственным требованием является введение на данных некоторого линейного порядка. Этим порядком может быть, к примеру, алфавитный или числовой порядок. Линейный порядок может быть на теге, указателе, файле или некотором другом ключе, определяющем данные. Но нас будет интересовать только наличие некоторого линейного порядка. Для простоты примем расположение имен в алфавитном порядке. Бинарное дерево поиска — это прежде всего бинарное дерево, в каждом узле которого находится имя (или другой ключ). Рассмотрим построение бинарного дерева поиска. В процессе построения станет понятно, как использовать это дерево. Предположим, что необходимо запомнить имена: Петерсон, Джонсон, Смит, Вейл, Спенсер, Рассел, Бауэр, Мартин, Уилсон и Лайман.

Начинаем с имени Петерсон и помещаем его в корень дерева. Поскольку следующее имя, Джонсон, в алфавитном порядке стоит перед именем Петерсон, делаем Джонсона левым сыном Петерсона, как показано на рис. 15.6. Следующее имя, Смит, в алфавитном порядке стоит после имени Петерсон, поэтому имя Смит становится правым сыном имени Петерсон, как показано на рис. 15.7.

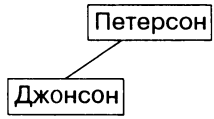


Рис. 15.6

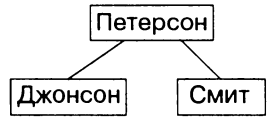


Рис. 15.7

Теперь рассмотрим имя Вейл. В алфавитном порядке имя Вейл стоит перед (или меньше, чем) именем Петерсон, спускаемся к левому сыну, Джонсону, и поскольку в алфавитном порядке имя Вейл расположено перед именем Джонсон, делаем Вейла левым сыном Джонсона, как показано на рис. 15.8.

Далее рассматриваем имя Спенсер. Поскольку в алфавитном порядке имя Спенсер идет после имени Петерсон, спускаемся к правому сыну, Смиту. Но так как в алфавитном порядке имя Спенсер расположено после имени Смита, делаем Спенсера правым сыном Смита, как показано на рис. 15.9.

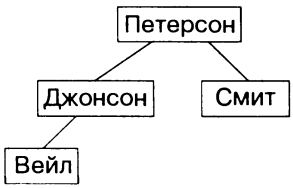


Рис. 15.8

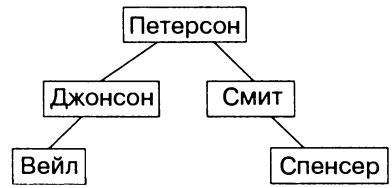


Рис. 15.9

Следующее имя Рассел. Поскольку по алфавиту имя Рассел стоит после имени Петерсон, спускаемся к правому сыну, Смиту. Имя Рассел по алфавиту расположено перед именем Смит; делаем Рассела левым сыном Смита, как показано на рис. 15.10.

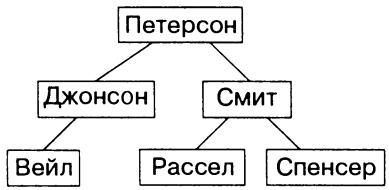


Рис. 15.10

Надеемся, что теперь схема начинает проясняться. Предположим, что есть имя, которое необходимо поместить в дерево. Если в вершине дерева уже есть имя и вставляемое имя в алфавитном порядке стоит после этого имени, то следует спуститься к левому сыну, а если оно стоит до этого имени, то нужно спуститься

к правому сыну. Если в вершине нет имени, то имя, которое нужно вставить, помещается в данную вершину.

Используя эту процедуру, доведем до конца рассмотрение нашего списка. Следующее имя — Бауэр. По алфавиту оно стоит перед именем Петерсон, поэтому спускаемся к левому сыну, Джонсону. По алфавиту имя Бауэр стоит перед именем Джонсон, поэтому спускаемся к левому сыну, Вейлу. В алфавитном порядке имя Бауэр располагается перед именем Вейл, поэтому спускаемся к левому сыну. Поскольку там нет имени, делаем Бауэр левым сыном, как показано на рис. 15.11.

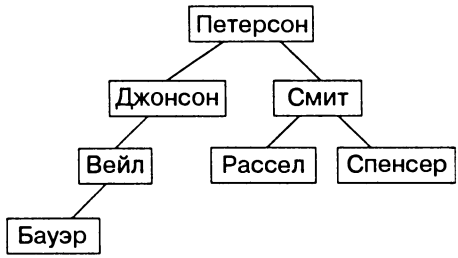


Рис. 15.11

Следующее имя, которое необходимо сохранить, — Мартин. Это имя по алфавиту стоит перед именем Петерсон, поэтому спускаемся к левому сыну, Джонсону. Но имя Мартин в алфавите следует после имени Джонсон, отсюда спускаемся к правому сыну и, поскольку там нет имени, делаем Мартина правым сыном, как показано на рис. 15.12.

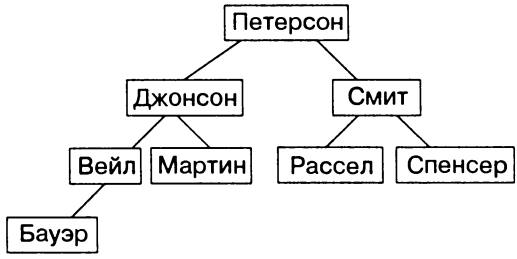


Рис. 15.12

Далее имя, которое необходимо сохранить, — Уилсон. Поскольку в алфавитном порядке оно стоит после имени Петерсон, спускаемся к правому сыну, Смиту. Но имя Уилсон расположено после имени Смит, значит, спускаемся к правому сыну, Спенсеру. Поскольку имя Уилсон стоит по алфавиту после имени Спенсер, и у Спенсера нет правого сына, делаем Уилсона правым сыном Спенсера, как показано на рис. 15.13.

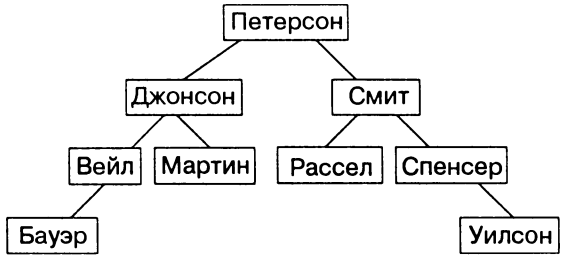


Рис. 15.13

Последним по счету, но не по значению, помещаем на свое место имя Лайман. Имя Лайман по алфавиту стоит перед именем Петерсон, спускаемся к левому сыну, Джонсону. Поскольку по алфавиту имя Лайман стоит после имени Джонсон, спускаемся к его правому сыну, Мартину. Так как имя Лайман по алфавиту стоит перед имением Мартин, спускаемся к его левому сыну. Но там имени нет, поэтому делаем Лаймана левым сыном Мартина.

При достижении левого (или правого) сына, который еще не определен, вероятно, точнее было бы сказать: "Если левый (правый) сын не существует, то создаем вершину и сохраняем имя", вместо: "Если там нет имени, то сохраняем имя". Но в данной ситуации внимание сосредотачивается, в основном, на концепции, а не на технических деталях. Теперь приведем алгоритм вставки имени в дерево поиска, который реально создает дерево поиска, за исключением размещения имени в корне дерева. Используя < и >, следует иметь в виду, что сказанное применимо для любого упорядочения.

Вставка(элемент)

(1) Начинаем с корня.
(2) Если элемент < объекта в вершине, переходим к левому сыну.
(3) Если элемент > объекта в вершине, переходим к правому сыну.
(4) Повторяем шаги 2 и 3, пока не достигнем вершины, которая не определена.
(5) Если достигнутая вершина не определена, то определяем вершину и вставляем элемент.

Поскольку метод создания дерева поиска описан, легко понять, как производить поиск элемента в дереве. Используется, в основном, тот же самый подход, кроме проверки, является ли данное имя больше или меньше имени в вершине, проверяется также несовпадение этого имени с именем в вершине. Если последнее выполняется, процесс поиска завершен. В противном случае повторяются предыдущие действия. Если достигается вершина, которая не определена, это означает, что данное имя не хранится в дереве. Предположим, например, что в рассмотренном выше дереве разыскивается имя Дженкинс. Поскольку оно "меньше, чем имя Петерсон", идем к левому сыну, Джонсону. Но так как имя Дженкинс "меньше, чем имя Джонсон", идем к левому сыну, Вейлу. Поскольку данное имя "больше, чем Вейл" (на самом деле нет имени больше, чем Вейл), идем к правому сыну. Но он не определен, поэтому данное имя отсутствует в списке. Далее приведен алгоритм поиска имени в дереве поиска.

Поиск(элемент)

(1) Начинаем с корня.
(2) Если элемент < объекта в вершине, идем к левому сыну.
(3) Если элемент > объекта в вершине, идем к правому сыну.
(4) Если элемент = объекту в вершине, то имя найдено; выполняем соответствующие действия и выходим.
(5) Повторяем шаги 2, 3 и 4, пока не достигнем вершины, которая не определена.

(6) Если достигнутая вершина не определена и в дереве нет имени, то выполняем соответствующие действия и выходим.

Вполне очевидно, что длина пути поиска элемента не может превышать высоту дерева. Поэтому, если дерево полное, то по теореме 15.9 в самом неблагоприятном варианте длина пути имеет порядок $O(\ln(n))$, где n — число элементов в дереве. Следовательно, если дерево сбалансированное, самый неблагоприятный случай имеет порядок $O(\ln(n))$.

В заключение рассмотрим, как удалять вершину из дерева. Предлагается один из методов. При изложении будут опущены все технические детали, касающиеся его фактической реализации.

Метод резюмирован в приведенном ниже алгоритме.

Удаление(элемент)

(1) Если вершина v_0 не имеет сыновей, просто удаляем ее.

(2) Если вершина v_0 имеет одного сына, удаляем v_0 и заменяем ее сыном.

(3) Если v_0 имеет двух сыновей, находим правого сына v_1 вершины v_0, а затем находим левого сына вершины v_1 (если он существует). Продолжаем выбирать левых сыновей каждой найденной вершины, пока не найдется такая вершина v, у которой не будет левого сына. Заменим v_0 на v и сделаем правого сына вершины v левым сыном родителя вершины v.

ПРИМЕР 15.13. Рассмотрим дерево, изображенное на рис. 15.14. Если удалить вершину x, то получится дерево, изображенное на рис. 15.15. Если же удалить вершину k, то получится дерево, изображенное на рис. 15.16.

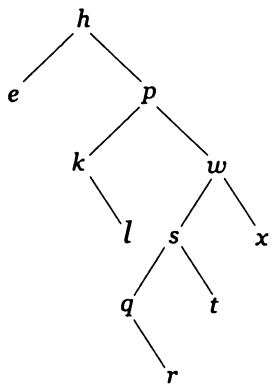


Рис. 15.14

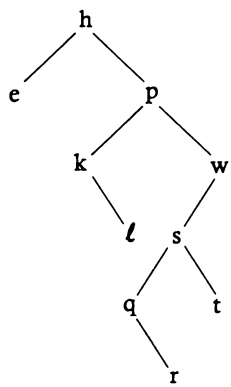


Рис. 15.15

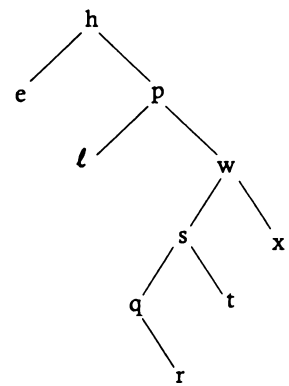


Рис. 15.16

Если в исходном дереве удаляется вершина h, нужно спуститься к правому сыну p, а затем к левому сыну k. Поскольку k не имеет левого сына, это искомый элемент для замены. Итак, меняем h на k и делаем l левым сыном p. В результате получаем дерево, изображенное на рис. 15.17.

Если вместо этого необходимо удалить вершину p, следует идти вправо к вершине w, затем влево к s, затем влево к q. Поскольку q не имеет левого сына,

это искомая вершина для замены. Поэтому заменяем p на q и делаем r левым сыном вершины s. В результате получаем дерево, изображенное на рис. 15.18.

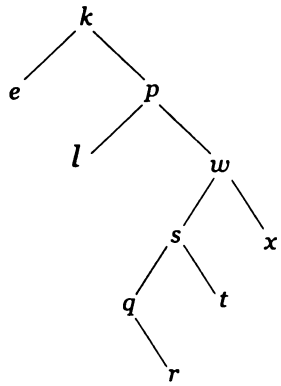


Рис. 15.17

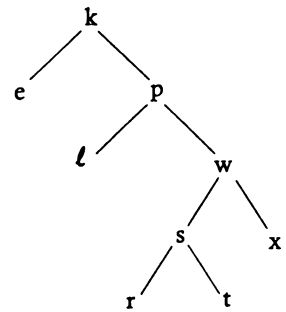


Рис. 15.18

□

■ УПРАЖНЕНИЯ

1. Задано дерево (рис. 15.19).

а) Вставьте букву j.

б) Вставьте букву u.

в) Вставьте букву m.

2. Задано дерево из предыдущего упражнения.

а) Сколько требуется сравнений для определения того, что i отсутствует?

б) Сколько требуется сравнений для определения того, что u отсутствует?

в) Сколько требуется сравнений для определения того, что f отсутствует?

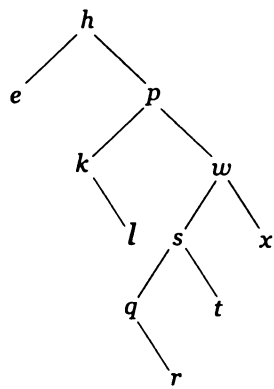


Рис. 15.19

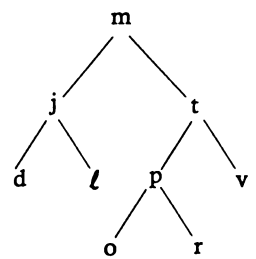


Рис. 15.20

3. Задано дерево (рис. 15.20).

а) Вставьте букву e.

б) Вставьте букву n.

в) Вставьте букву u.

4. Задано дерево из предыдущего упражнения.

а) Сколько требуется сравнений для определения того, что n отсутствует?

б) Сколько требуется сравнений для определения того, что u отсутствует?

в) Сколько требуется сравнений для определения того, что b отсутствует?

5. Опишите бинарное дерево с пятью вершинами, которое дает наихудший вариант для поиска?

6. Постройте бинарное дерево для следующих имен, хранящихся в алфавитном порядке, если данные вводились в таком порядке: Говард, Джеймс, Аарон, Спенсер, Джексон, Мэдисон, Винес, Кальхаун, Барнс, Макдав и Петерсон.

7. Постройте бинарное дерево для имен президентов Соединенных Штатов, хранящихся в алфавитном порядке, если имена вводились в том порядке, в котором президенты находились на своем посту первый срок.

8. Постройте бинарное дерево для приведенных ниже чисел, хранящихся в обычном порядке, если они вводились в следующем порядке: 25, 35, 15, 25, 48, 36, 22, 44, 18, 30, 42, 11, 39, 32.

9. Постройте бинарное дерево для приведенных ниже букв, хранящихся в алфавитном порядке, если они вводились в следующем порядке: $m, a, r, i, g, o, l, d, s, t, u, l, i, p, s$.

10. Постройте бинарное дерево для приведенных ниже марок автомобилей, хранящихся в алфавитном порядке, если данные поступили в следующем порядке: Бэнсон, Крайслер, Шевроле, Додж, Мэркьюри, Нэш и Олдсмобил.

11. Постройте бинарное дерево высоты 4 для букв английского алфавита.

12. Задано дерево (рис. 15.21).

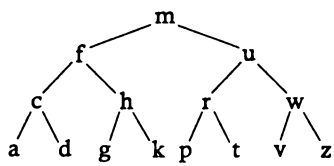


Рис. 15.21

а) Удалите a.

б) Удалите f.

в) Удалите m.

г) Удалите u.

д) Удалите r.

13. Задано дерево (рис. 15.22).

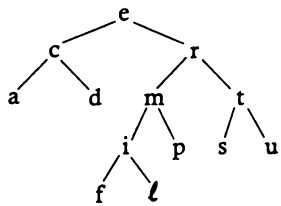


Рис. 15.22

а) Удалите a.
б) Удалите e.
в) Удалите r.
г) Удалите m.
д) Удалите t.

14. Напишите программу для алгоритма Вставка (элемент) в псевдокодах или на заданном языке.
15. Напишите программу для алгоритма Поиск (элемент) в псевдокодах или на заданном языке.
16. Напишите программу для алгоритма Удаление (элемент) в псевдокодах или на заданном языке.

15.3. ВЗВЕШЕННЫЕ ДЕРЕВЬЯ

В этом разделе рассматриваются две проблемы, имеющие одно и то же решение. В компьютере все буквы и другие символы хранятся в виде строк из единиц и нулей. Например, если символы хранятся как ASCII-символы, они состоят из строк длины 7 плюс восьмой символ, который используется для контроля четности. Пусть A — множество всех символов, которые необходимо сохранить. Всегда есть возможность сохранить A в виде строк некоторой фиксированной длины n, состоящих из единиц и нулей. Это дает существенное преимущество при "декодировании" строки обратно в символы из A. Первые n байтов (единиц и нулей в строке) образуют первый символ из A, вторые n байтов — второй символ и т.д.

Однако, если данных достаточно много, всегда желательно провести их компактификацию, т.е. использовать для их хранения как можно меньше места. Очевидный способ компактно скомпоновать данные — изменять длину строк, используемых для их хранения, так чтобы более короткие строки хранили часто используемые символы, а более длинные строки — редко используемые символы. И здесь возникает проблема: если строки, представляющие разные символы, имеют разную длину, то как узнать, где заканчивается строка одного символа и начинается строка другого? Например, если b представлено как 101, e как 11, a как 10, q как 1011 и w как 110, то строка 1011110 представляет bea или qw?

Определим ***однозначно декодируемый код*** для языка как такое множество, что каждая строка в языке может быть задана однозначно как конкатенация элементов C. В этом случае строки из единиц и нулей, представляющие элементы из A, будут нашим кодом. Если эти строки образуют однозначно декодируемый код, то, по крайней мере, разделяя строки на элементы, представляющие A, мы знаем, что представление однозначно и, следовательно, декодированные слова правильные. Хотелось бы это как-то улучшить. Определим код C как ***префиксный код***, если он обладает тем свойством, что никакой элемент кода не может быть начальной строкой другого элемента кода. Таким образом, прочитав строку из единиц и нулей, представляющую символ из A, мы знаем, что это искомый символ, а не начало другого символа. Поэтому следующие 1 или 0 должны начинать строку для следующего символа, который декодируется. Например, если

строки 111, 1011, 1001, 110 и 01 представляют в пятибуквенном алфавите соответственно буквы a, e, i, o и u и в качестве первых трех знаков строки имеются цифры 110, это означает представление буквы o, поскольку никакая другая буква не начинается с этих цифр. Аналогично, если следующие три цифры 011, это означает, что 01 представляет букву u, а 1 — начало следующей буквы.

Рассмотрим бинарное дерево с листьями t_1, t_2, t_3, t_4, t_5 и t_6, изображенное на рис. 15.23. Если рассматривать путь от корня дерева к любому из листьев, то каждое ребро вдоль пути должно вести к правому или левому сыну предыдущей вершины. Если ребро ведет к левому сыну, обозначим его через 0, а если к правому сыну, то пометим его 1. Поскольку путь к листу v_1 есть путь к левому сыну, а затем к другому левому сыну, то путь к v_1 обозначаем через 00. Аналогично пути к v_2, v_3, v_4, v_5 и v_6 обозначаются через 010, 011, 10, 110 и 111. Строку, соответствующую данному листу, назовем ***путевым кодом***. Заметим, что строки, соответствующие этим листьям, образуют префиксный код. Строка, соответствующая каждому листу, определяется однозначно, так как к каждому листу ведет единственный путь. Обобщая вышеизложенное, приходим к следующей теореме. Доказательство оставляем читателю.

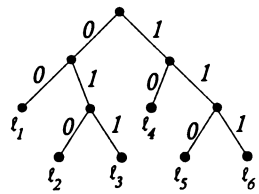


Рис. 15.23

ТЕОРЕМА 15.14. В любом бинарном дереве путевые коды для листьев дерева являются префиксным кодом.

Предположим, что имеется бинарное дерево и для каждого символа в алфавите A, который требуется представить с помощью двоичной строки, на дереве имеется лист. Присвоим каждой букве алфавита A путевой код соответствующего листа на бинарном дереве. Таким образом, если буквы a, e, i, o, u и w связаны соответственно с листьями v_1, v_2, v_3, v_4, v_5 и v_6 в упомянутом выше дереве, то буквам a, e, i, o, u и w соответствуют строки 00, 010, 011, 10, 110 и 111. Первоначальная формулировка задачи содержала требование, чтобы часто встречающиеся символы имели более короткие строки. В рамках изложенного выше метода это означает, что часто встречающимся символам должен соответствовать более короткий путь от корня или, что равносильно, они должны находиться на более низком уровне. Для того чтобы оценить, насколько близки мы от цели поставить в соответствие часто встречающимся буквам более короткие строки, т.е. минимизировать коды, необходим некоторый критерий. Предположим, что каждая буква a_i в алфавите символов A появляется с относительной частотой f_i и представляющая ее двоичная строка имеет длину l_i. Тогда величина

$$w = l_1 f_1 + l_2 f_2 + l_3 f_3 + l_4 f_4 + \cdots + l_n f_n = \sum_{i=1}^{n} l_i f_i$$

называется ***весом кода***. Если листья взвешенного дерева представляют знаки кода, вес листа равен частоте появления буквы, а расстояние от корня до вершины равно длине двоичного кода для символа в листе. Определим ***вес дерева*** как

величину

$$w = l_1 f_1 + l_2 f_2 + l_3 f_3 + l_4 f_4 + \cdots + l_n f_n = \sum_{i=1}^{n} l_i f_i.$$

В более общем виде ***вес дерева*** определяется как

$$w = l_1 w_1 + l_2 w_2 + l_3 w_3 + l_4 w + \cdots + l_n w_n = \sum_{i=1}^{n} l_i w_i,$$

где w_i — вес, приписанный вершине v_i, и l_i — длина пути от корня до вершины. Данное понятие не следует путать с понятием взвешенного графа, включающего деревья, где вес приписывается ребрам. Эти вопросы мы обсудим в разделе 15.6.

Чем меньше вес дерева, тем в большей степени достигнута поставленная цель. Таким образом, чтобы найти наилучший код для минимизации данных, необходимо найти код с минимальным весом. Процесс построения такого дерева называется ***алгоритмом Хаффмана***. Код, приписываемый символам согласно их путевому коду, называется ***кодом Хаффмана***. Прежде чем излагать алгоритм, следует разобраться в механизме его работы. Предположим, имеются следующие буквы и соответствующие им частоты:

символ	частота
a	7
e	10
w	3
k	5
m	8
#	2
%	1



Рис. 15.24

Для начала разместим частоты в возрастающем порядке: $(1, 2, 3, 5, 7, 8, 10)$. Затем формируем дерево, используя две буквы с наименьшими частотами в качестве вершин и сумму частот в качестве их родителя. Помещаем 0 на ребре к левому сыну и 1 на ребре к правому сыну, после чего получаем дерево, изображенное на рис. 15.24.

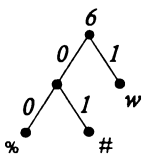


Рис. 15.25

В списке частот заменяем значения двух наименьших частот их суммой, после чего имеем $(3, 3, 5, 7, 8, 10)$. Если бы полученная сумма оказалась больше одной или нескольких частот в списке, его надо было бы отсортировать так, чтобы частоты снова располагались в возрастающем порядке. Опять выбираем две вершины или деревья в соответствии с наименьшими числами в списке частот и формируем дерево с этими вершинами или деревьями в качестве сыновей и их суммой в качестве родителя. В данном случае это дерево со значением 3 и буква w со значением 3 в качестве сыновей и их суммой 6 в качестве родителя.

Помещаем 0 на ребре к левому сыну и 1 на ребре к правому сыну, после чего получаем граф (рис. 15.25).

Теперь заменим в списке частот значения двух наименьших частот их суммой, 6, и расположим значения в возрастающем порядке. В результате получим список $(5, 6, 7, 8, 10)$. Во всех случаях значения вершин или деревьев, использованных в качестве сыновей, удаляются из списка частот и из строящегося дерева. Мы формируем дерево, выбирая две вершины или два дерева с наименьшими числами в списке частот и используя их как сыновей, а их сумму в качестве частоты родителя. В данном случае это вершина k со значением 5 и дерево со значением 6. Помещаем 0 на ребро к левому сыну и 1 на ребро к правому сыну, после чего получаем дерево, изображенное на рис. 15.26.

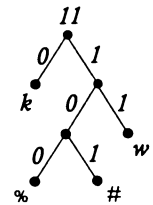


Рис. 15.26

В списке частот опять заменяем значения двух наименьших частот их суммой, 11, и располагаем значения частот в возрастающем порядке, после чего получаем список $(7, 8, 10, 11)$. Продолжая процесс, получаем следующие деревья и списки частот: рис. 15.27 $(10, 11, 15)$ и рис. 15.28 $(15, 21)$.

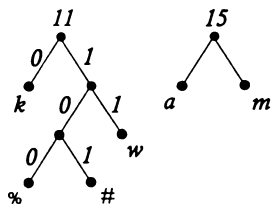


Рис. 15.27

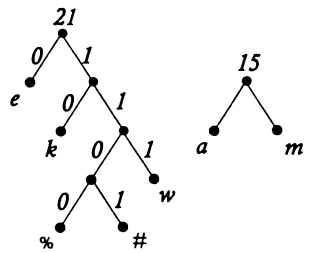


Рис. 15.28

Наконец, соединяем два оставшихся дерева, поместив 0 на ребро к левому сыну, 1 на ребро к правому сыну, в результате получаем дерево, изображенное на рис. 15.29.

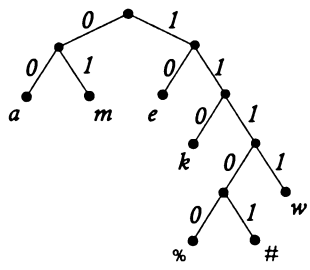


Рис. 15.29

Помещать сумму в корень нет необходимости. Заметим, что числами кода Хаффмана для a, e, w, k, m, # и % являются 00, 10, 1111, 01, 110, 11101 и 11100 соответственно. Ниже представлен алгоритм Хаффмана построения дерева для символов $s_1, s_2, \ldots s_n$, имеющих соответствующие частоты $f_1, f_2, \ldots, f_n$.

Алгоритм Хаффмана $((s_1, f_1), (s_2, f_2), \ldots, (s_n, f_n))$

(1) Расположить частоты в возрастающем порядке в таблице частот.
(2) Если f_i и f_j — две наименьшие частоты, то сформировать бинарное дерево с s_i и s_j в качестве сыновей, где s_i — левый, а $f_i + f_j$ — частота родителя. Поместить 0 на ребро к левому сыну, а 1 — на ребро к правому сыну.
(3) Удалить f_i и f_j из таблицы частот и заменить суммой $f_i + f_j$.
(4) Снова расположить частоты в возрастающем порядке.
(5) Удалить две наименьшие частоты из таблицы частот и сформировать дерево, используя в качестве сыновей символы или деревья, соответствующие этим частотам, когда меньший символ или дерево — левый, а сумма их частот — частота родителя. Если какой-либо из сыновей — дерево, удалить метку его частоты из построенного дерева.
(6) Заменить значение двух наименьших частот их суммой.
(7) Повторять шаги 4–7, пока в таблице частот не останется только один элемент.
(8) Удалить частоту из корня дерева и из таблицы частот.
(9) Сформировать код Хаффмана для элемента, составляя строку из единиц и нулей на ребрах пути из корня к заданному элементу.

Рассмотрим вторую из упомянутых выше задач с таким же решением. Предположим, имеется n элементов данных и известна частота использования каждого из них. Требуется сохранить их как листья бинарного дерева так, чтобы наиболее часто используемые элементы были более доступны при обратном поиске. Но это означает, что наиболее часто используемым элементам должны соответствовать самые короткие пути от корня. Это как раз то, что делалось до сих пор, поскольку наиболее часто используемые символы имеют самые короткие двоичные строки и, следовательно, самые короткие пути от корня.

ПРИМЕР 15.15. Предположим, имеются следующие буквы и их частоты:

символ	частота
A	15
B	25
C	10
D	30
E	15
F	5

Построим такое бинарное дерево, чтобы наиболее часто используемые элементы были возможно ближе расположенны к корню. Более точно, требуется найти такое дерево, что **вес дерева**

$$w = l_1 f_1 + l_2 f_2 + l_3 f_3 + l_4 f_4 + \cdots + l_n f_n = \sum_{i=1}^{n} l_i f_i$$

минимален, где l_i и f_i — уровень и частота заданного элемента. Сначала упорядочим частоты, приведенные в списке частот (5, 10, 13, 17, 25, 30). Воспользовавшись приведенным выше алгоритмом, получаем следующие деревья и списки частот: рис. 15.30, (13, 15, 17, 25, 30); рис. 15.31 (17, 25, 28, 30); рис. 15.32 (28, 30, 42); рис. 15.33 (42, 58).

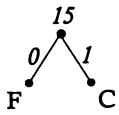


Рис. 15.30

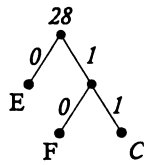


Рис. 15.31

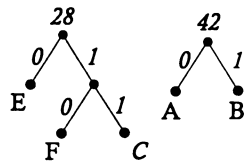


Рис. 15.32

В заключение объединяем два дерева и формируем искомое дерево, которое изображено на рис. 15.34.

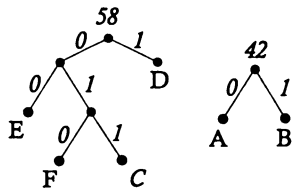


Рис. 15.33

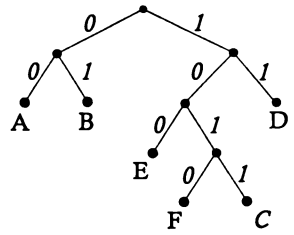


Рис. 15.34 □

Теперь необходимо показать, что дерево Хаффмана — это дерево с минимальным весом. Начнем с доказательства серии лемм.

ЛЕММА 15.16. Для заданного множества из n символов и их частот существует бинарное дерево минимального веса с символами в качестве листьев.

ДОКАЗАТЕЛЬСТВО. Существует только конечное число бинарных деревьев с n листьями. Для каждого из них подсчитаем

$$w = l_1 f_1 + l_2 f_2 + l_3 f_3 + l_4 f_4 + \cdots + l_n f_n = \sum_{i=1}^{n} l_i f_i$$

и выберем дерево, которое обеспечивает наименьшее значение. ■

ЛЕММА 15.17. В дереве с минимальным весом на максимальном уровне листья присутствуют в парах, т.е. везде, где есть левый сын, имеется и правый сын и наоборот.

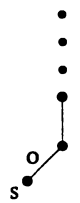


Рис. 15.35

ДОКАЗАТЕЛЬСТВО. Допустим, что на максимальном уровне имеется единственный сын; пусть это будет левый сын, помеченный символом s. Правого сына нет, поэтому имеем дерево, изображенное на рис. 15.35.

Если мы теперь удалим левого сына и обозначим родителя меткой s, то в результате получим дерево с меньшим весом, а это противоречит предположению о том, что дерево имеет минимальный вес. ■

ЛЕММА 15.18. В дереве с минимальным весом два символа с минимальными частотами расположены на максимальном уровне.

ДОКАЗАТЕЛЬСТВО. Если теорема не верна, то имеется символ a на уровне l_i с частотой f_i и символ b на уровне l_j с частотой f_j, где $l_i > l_j$ и $f_i > f_j$. В общий вес дерева эти два символа вносят величину $l_i f_i + l_j f_j$. Если эти символы поменять местами, то их вклад в общий вес дерева составит $l_i f_j + l_j f_i$. Но

$$\begin{aligned} l_i f_i + l_j f_j - (l_i f_j + l_j f_i) &= l_i f_i + l_j f_j - l_i f_j - l_j f_i = \\ &= (l_i - l_j)(f_i - f_j) > \\ &> 0, \end{aligned}$$

поскольку $l_i - l_j > 0$ и $f_i - f_j > 0$. Таким образом, $l_i f_i + l_j f_j > l_i f_j + l_j f_i$, поэтому перестановка символов приводит к уменьшению веса дерева, что противоречит предположению о том, что дерево имеет минимальный вес. ■

ЛЕММА 15.19. Существует дерево с минимальным весом, в котором два символа с наименьшими частотами имеют одного родителя.

ДОКАЗАТЕЛЬСТВО. По лемме 15.17 каждый лист на максимальном уровне имеет брата. По лемме 15.18 оба символа с минимальным весом, к примеру, a и b, находятся на одном уровне. Если они не братья, то поскольку находятся на одном уровне, замена a на b не повлияет на вес дерева. ■

Теперь все готово для формулировки основной теоремы.

ТЕОРЕМА 15.20. Для заданного множества символов с соответствующими частотами дерево Хаффмана является деревом с минимальным весом.

ДОКАЗАТЕЛЬСТВО. Докажем теорему, используя индукцию по числу символов. Для двух символов, скажем, a и b, дерево Хаффмана изображено на рис. 15.36 и действительно является деревом с минимальным весом.

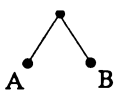


Рис. 15.36

Таким образом, теорема верна для $n = 2$. Предположим, что утверждение теоремы справедливо для дерева с n символами. Пусть T_{n+1} — дерево минимального веса, содержащее $n+1$ символ. Предположим, что a и b с частотами f_a и f_b являются символами с самыми низкими частотами. Тогда по лемме 15.18 они могут быть братьями на самом высоком уровне $l+1$. Теперь удалим a и b из дерева и присвоим их родителю символ c с частотой $f_a + f_b$. Назовем это дерево T_n. Поскольку в дереве только n символов, то по индуктивному предположению существует дерево

Хаффмана с минимальным весом с этими символами. Назовем дерево T'_n. Поскольку T'_n имеет минимальный вес, то вес T'_n меньше или равен весу дерева T_n. Теперь заменим c с частотой $f_a + f_b$ на родителя с сыновьями a и b, образуя дерево T'_{n+1}. Другими словам, для перехода от дерева T_{n+1} к дереву T_n используется обратный процесс. Но это совпадает с построением дерева Хаффмана с $n+1$ символом, поскольку для построения дерева Хаффмана с $n+1$ символом следовало бы начать с a и b и образовать дерево с родителем, имеющим частоту $f_a + f_b$, так что если рассматривать c как представителя этого дерева, то построение дерева Хаффмана T'_n является продолжением дерева Хаффмана T'_{n+1}.

Теперь имеются четыре дерева: деревья Хаффмана T'_n и T'_{n+1}, где T'_n — дерево с минимальным весом, и деревья T_n и T_{n+1}, где T_{n+1} — дерево с минимальным весом. Мы перешли от дерева T'_n к дереву T'_{n+1} и от дерева T_n к дереву T_{n+1}, заменяя c с частотой $f_a + f_b$ на вершины a и b с частотами f_a и f_b и их родителя. Теперь рассмотрим, каким образом это влияет на вес дерева. При переходе от дерева T_n к дереву T_{n+1} вершина c с частотой $f_a + f_b$, длиной l, общим весом

$$l(f_a + f_b)$$

заменяется вершиной a с частотой f_a, длиной $l+1$ и вершиной b с частотой f_b, длиной $l+1$, общий вес которых равен

$$f_a(l+1) + f_b(l+1) = l(f_a + f_b) + f_a + f_b.$$

Таким образом, увеличение веса дерева составило $f_a + f_b$. Аналогично при переходе от дерева T'_n к дереву T'_{n+1} увеличение веса дерева также равно $f_a + f_b$. Поскольку вес дерева T'_n меньше или равен весу дерева T_n и увеличение веса при переходе от дерева T'_{n+1} и T_{n+1} одно и то же, то вес дерева T'_{n+1} меньше или равен весу T_{n+1}. Таким образом, дерево T'_{n+1} — дерево минимального веса. ■

■ УПРАЖНЕНИЯ

1. Заданы следующие символы и их коды:

символ	код
c	111
d	101
e	01
i	11001
s	0010
u	1101

а) закодируйте слово *discuss*;
б) закодируйте слово *suede*;
в) декодируйте слово 10101111100110101101;
г) декодируйте слово 11001001000101101010010.

2. Заданы следующие символы и их коды:

символ	код
a	111
c	101
e	01
i	11011
g	0010
h	1001
m	0001
s	10001
t	11001

а) закодируйте слово *mathematics*;
б) закодируйте слово *mammals*;
в) декодируйте слово 11100101111000111001;
г) декодируйте слово 10010111011001010011100 1.

3. Заданы следующие символы с их частотами и три множества кодов.

символ	частота	код 1	код 2	код 3
c	7	111	010	111
d	8	101	11	101
e	10	01	001	001
i	2	11001	101	10001
s	9	0010	011	011
u	4	1101	000	100011

Найдите вес каждого кода и выберите наиболее подходящий для компактного хранения данных.

4. Заданы следующие символы с их частотами и три множества кодов.

символ	частота	код 1	код 2	код 3
a	8	111	1111	111
c	4	101	1110	011
e	13	01	1100	101011
i	9	11011	1001	11011
g	5	0010	1011	001
h	2	1001	1000	000
m	3	0001	1010	100
s	7	10001	0011	10111
t	11	11001	0111	10110

Найдите вес каждого кода и выберите наиболее подходящий для компактного хранения данных.

5. Заданы следующие символы с их частотами:

символ	частота
a	8
b	6
c	3
r	4

а) постройте дерево Хаффмана;
б) определите код Хаффмана;
в) найдите вес кода;
г) закодируйте слово *crab*;
д) закодируйте слово *bar*;
е) декодируйте 1100111;
ж) декодируйте 10011110.

6. Заданы следующие символы с их частотами:

символ	частота
a	8
b	5
e	10
k	1
r	3

а) постройте дерево Хаффмана;
б) определите код Хаффмана;
в) найдите вес кода;
г) закодируйте слово *break*;
д) закодируйте слово *barber*;
е) декодируйте 11011011000;
ж) декодируйте 11110110111101101.

7. Заданы следующие символы с их частотами:

символ	частота
a	8
d	7
e	12
k	1
r	4
s	6

а) постройте дерево Хаффмана;
б) определите код Хаффмана;
в) найдите вес кода;
г) закодируйте слово *eraser*;
д) закодируйте слово *darker*.
е) декодируйте 10010110001100;
ж) декодируйте 0011011001111001.

8. Заданы следующие символы с их частотами:

символ	частота
c	7
e	12
g	3
h	2
o	9
r	4
s	5
t	8
u	1

а) постройте дерево Хаффмана;
б) определите код Хаффмана;
в) найдите вес кода;
г) закодируйте слово *ghost*;
д) закодируйте слово *choose*;
е) декодируйте 100010010000111101;
ж) декодируйте 10001001110111101;
з) декодируйте 11100010000100110001.

9. Заданы следующие символы с их частотами:

символ	частота	символ	частота
a	8	h	2
c	4	m	3
e	13	s	7
i	9	t	11
g	5		

а) постройте дерево Хаффмана;
б) определите код Хаффмана;
в) найдите вес кода.

10. Заданы следующие символы с их частотами:

символ	частота	символ	частота
a	19	m	7
d	8	n	9
e	20	o	12
i	11	r	11
g	3	s	16
h	1	t	14
l	5		

а) постройте дерево Хаффмана;
б) определите код Хаффмана.
в) найдите вес кода;
г) закодируйте слова *moonlightandroses*;
д) закодируйте слово *roosters*.

11. Заданы следующие символы с их частотами:

символ	частота	символ	частота
a	20	n	7
d	9	o	13
e	25	r	15
f	4	s	11
g	8	t	5
h	2	u	14
l	6	w	1
m	10		

а) постройте дерево Хаффмана;
б) определите код Хаффмана;

в) найдите вес кода;
г) закодируйте слова *eastofthesunandwestofthemoon*;
д) закодируйте слово *sweater*.

12. Заданы следующие символы с их частотами:

символ	частота	символ	частота
a	25	n	12
b	10	o	20
e	30	r	9
g	7	s	11
h	6	t	15
i	18	u	3
l	8	y	13

а) постройте дерево Хаффмана;
б) определите код Хаффмана;
в) найдите вес кода;
г) закодируйте слова *eastofthesunandwestofthemoon*;
д) закодируйте слова *lightsout*.

13. Докажите теорему 15.14: В любом бинарном дереве путевые коды для листьев дерева являются префиксным кодом.

15.4. ОБХОД БИНАРНЫХ ДЕРЕВЬЕВ

В данном разделе рассматриваются три способа обхода бинарного дерева, при которых каждая вершина посещается в точности один раз. Внимание будет сосредоточено на деревьях, изображающих арифметичесткие выражения, хотя такие обходы имеют также другое использование. Сначала представим арифметические выражения с помощью деревьев, а затем рассмотрим способы их обхода.

Деревья, изображающие арифметические выражения, обладают тем свойством, что все их внутренние вершины являются бинарными операторами. Рассмотрим выражение $A+B$. Соответствующее ему дерево изображено на рис. 15.37.

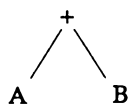


Рис. 15.37

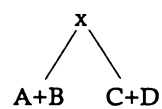


Рис. 15.38

Воспользуемся нисходящим принципом обработки, который продемонстрируем на примере. Рассмотрим выражение $(A + B) \times (C + D)$. Оно представляет собой произведение $A + B$ и $C + D$, поэтому сначала формируем дерево, изображенное на рис. 15.38. Далее, поскольку выражение $A + B$ — это сумма A и B, а $C + D$ — сумма C и D, то мы формируем дерево, изображенное на рис. 15.39.

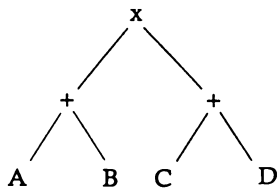


Рис. 15.39

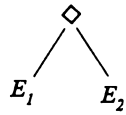


Рис. 15.40

Каждое арифметическое выражение, содержащее бинарный оператор, по сути является бинарной операцией над двумя арифметическими выражениями (которые могут и не содержать арифметические операции). Начинаем с корня и, если выражение в вершине имеет вид $E_1 \diamond E_2$, где $\diamond$ — бинарная операция над выражениями E_1 и E_2, заменяем вершину деревом, изображенным на рис. 15.40.

Повторяем тот же процесс для вершин E_1 и E_2. Продолжаем действовать таким образом, пока не останется вершин, содержащих бинарные операции. Теперь покажем три способа обхода дерева, которое задает арифметическое выражение. В каждом случае используем команду **обработать** (n), где n — узел. Оставим этот термин пока неопределенным, поскольку его значение будет зависеть от того, чего мы хотим достичь при прохождении вершин. Для наших целей достаточно рассмотреть обычную команду печати символа. Начнем с **обхода дерева в центрированном порядке**. При таком обходе мы обращаем процесс, использованный при создании дерева. Если задано дерево, изображенное на рис. 15.40, где $\diamond$ — бинарная операция над выражениями E_1 и E_2, обрабатываем (или печатаем) это как $E_1 \diamond E_2$. В результате получается выражение в **инфиксной записи**, как описано в разделе 5.5.

Приведенный ниже алгоритм описывает процедуру обхода дерева в центрированном порядке (ОЦП). Обозначим левого сына вершины v аббревиатурой лс(v), а правого сына вершины v — аббревиатурой пс(v).

Алгоритм обхода дерева в центрированном порядке — ОЦП (*корень*)

(1) Если лс(*корень*) существует, то ОЦП(лс(*корень*)).
(2) Обработать(*корень*).
(3) Если пс(*корень*) существует, то ОЦП(пс(*корень*)).

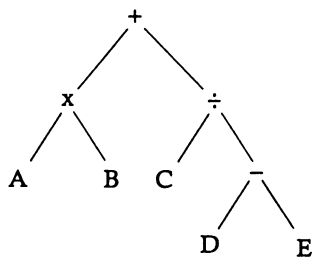


Рис. 15.41

Рассмотрим дерево, изображенное на рис. 15.41. Начнем с ОЦП(+); первая инструкция вызывает ОЦП(×). Первая инструкция команды в ОЦП(×) вызывает

ОЦП(A). Поскольку A не имеет левого сына, то переходим к шагу 2 команды ОЦП(A) и обрабатывем (печатаем) A. Переходим к шагу 3 в ОЦП(A) и, поскольку ОЦП(A) не имеет правого сына, обработка ОЦП(A) закончена, и мы возвращаемся к шагу 2 в ОЦП($\times$). Здесь обрабатывается (печатается) $\times$, а на шаге 3 вызывается ОЦП(B). Заметим, что результатом вызова ОЦП(A) была обработка (печать) листа A и возврат к родителю. Поэтому впредь, если имеется ОЦП(L), где L — лист, то три шага соответствующего алгоритма будем сокращенно описывать как **обработать лист L**. Используя это обозначение, печатаем B и возвращаемся к ОЦП($\times$). Теперь выполнена команда "обработать" (напечатать)

$$A \times B.$$

Команда ОЦП($\times$) выполнена, поэтому возвращаемся к ОЦП($+$). (К настоящему моменту общая концепция уже вполне ясна, в дальнейшем будем говорить просто "обработать", понимая под этим "печатать". Находясь на шаге 2 в ОЦП($+$), обрабатываем $+$ и переходим к шагу 3 в ОЦП($+$). На этом этапе вызываем команду ОЦП($\div$). На первом шаге ОЦП($\div$) вызываем ОЦП(C), обрабатываем лист C и возвращаемся на шаг 2 в ОЦП($\div$). Теперь обрабатываем $\div$, после чего уже обработано

$$A \times B + C \div .$$

Возвращаемся на шаг 3 в ОЦП($\div$) и вызываем ОЦП($-$). На шаге 1 в ОЦП($-$) вызываем ОЦП(D) и обрабатывем лист D. Затем возвращаемся на шаг 2 в ОЦП($-$) и вызываем ОЦП(E). Обрабатываем лист E, после чего возвращаемся к команде ОЦП($-$). Однако, команда ОЦП($-$) уже выполнена, поэтому возвращаемся к ОЦП($\div$). Но эта команда тоже закончена, поэтому возвращаемся к ОЦП($+$). Наконец, в завершение всего процесса выполняем команду ОЦП($+$). Итак, получено выражение

$$A \times B + C \div D - E.$$

Это выражение, согласно изложенному в разделе 5.5, имеет **инфиксную запись**. К сожалению, в отсутствие скобок выражение в том виде, в каком получено, не имеет смысла. Если вспомнить содержание раздела 5.5, то именно такая ситуация была основанием для перехода к **префиксной (польской)** или **постфиксной (польской инверсной)** записи выражений. К счастью, два следующих способа обхода бинарных деревьев порождают именно такие формы записи выражений. Следует отметить, что инфиксный способ обхода в применении к бинарному дереву поиска дает обработку элементов в алфавитном порядке.

Теперь рассмотрим **обход в прямом порядке**. Такой обход осуществим, например, для дерева, изображенного на рис. 15.42, где $\diamond$ — бинарная операция над выражениями E_1 и E_2. Эту операцию будем обрабатывать (или печатать) как $\diamond E_1 E_2$. В результате получим выражение в **префиксной форме**. Такая форма записи широко используется в логике и хороша тем, что в ней не используются скобки.

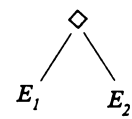


Рис. 15.42

Приведенный ниже алгоритм описывает процедуру обхода в прямом порядке (ОПП). Как и раньше, будем сокращенно обозначать левого сына вершины v аббревиатурой лс(v), а правого сына вершины v — аббревиатурой пс(v).

Алгоритм обхода дерева в прямом порядке — ОПП(*корень*)

(1) Обработать(*корень*).
(2) Если лс(*корень*) существует, то ОПП(лс(*корень*)).
(3) Если пс(*корень*) существует, то ОПП(пс(*корень*)).

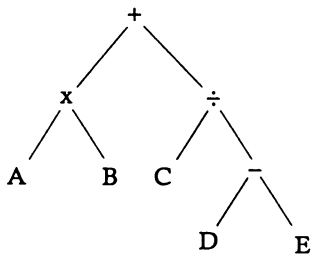


Рис. 15.43

Снова рассмотрим дерево (рис. 15.43). Начнем с ОПП(+); первая инструкция обрабатывает +. Шаг 2 вызывает ОПП(×). Первая инструкция ОПП(×) обрабатывает ×. Затем вторая инструкция вызывает (*A*). На шаге ОПП(*A*) обрабатываем *A*. На этом этапе уже обработано

$$+ \times A.$$

Поскольку *A* не имеет левого сына, переходим к шагу 3 в ОПП(*A*); так как нет и правого сына, возвращаемся к ОПП(×). Опять замечаем, что, как и в случае ОЦП(*A*), результатом вызова ОПП(*A*) является обработка (печать) листа *A* и возврат к родителю. Поэтому впредь, если имеется ОПП(*L*), где *L* — лист, то три шага соответствующего алгоритма будем сокращенно описывать как **обработать лист L**. Находясь на шаге 3 в ОПП(×), вызываем ОПП(*B*). Обрабатываем лист *B* и возвращаемся к ОПП(×). Это завершает ОПП(×), поэтому возвращаемся к ОПП(+). К этому моменту обработано (уже напечатано)

$$+ \times AB.$$

Затем на шаге 3 в ОПП(+) вызываем ОПП(÷). Первый шаг ОПП(÷) обрабатывает ÷. Второй шаг вызывает ОПП(*C*). Обрабатываем лист *C* и возвращаемся на шаг 3 в ОПП(÷). Эта инструкция вызывает ОПП(−). Первый шаг в ОПП(−) обрабатывает −. Шаг 2 вызывает ОПП(*D*). Обрабатываем лист *D* и возвращаемся на шаг 3 в ОПП(−). На данном этапе уже обработано

$$+ \times AB \div C - D.$$

Шаг 3 в ОПП(−) вызывает ОПП(*E*). Обрабатываем лист *E* и возвращаемся к ОПП(−). Но ОПП(−) завершено, поэтому возвращаемся к ОПП(÷). Однако, ОПП(÷) также завершено, поэтому возвращаемся к ОПП(+). Наконец, ОПП(+) тоже завершен, и мы получаем результат

$$+ \times AB \div C - DE,$$

который, согласно изложенному в разделе 5.5, выражен **префиксной**, или **польской записью**.

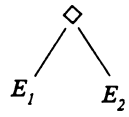


Рис. 15.44

И, наконец, рассмотрим **обход в обратном порядке**. Этот обход осуществим для дерева, изображеного на рис. 15.44, где ◇ — бинарная операция над выражениями E_1 и E_2, которую будем обрабатывать (или печатать как $E_1E_2\diamond$. В результате получим выражение в ***постфиксной***, или ***польской инверсной записи***, как описано в разделе 5.5. Всякий, кто пользовался калькулятором, знаком с постфиксной, или польской инверсной записью. Она обладает тем преимуществом, что позволяет без скобок, чем и объясняется ее популярность.

Приведенный ниже алгоритм описывает процедуру обхода в обратном порядке (ООП). Как и раньше, будем сокращенно обозначать левого сына вершины v аббревиатурой лс(v), а правого сына вершины v — аббревиатурой пс(v).

Алгоритм обхода дерева в обратном порядке — ООП(*корень*)

(1) Если лс(*корень*) существует, то ООП(лс(*корень*)).
(2) Если пс(*корень*) существует, то ООП(пс(*корень*)).
(3) Обработать(*корень*).

Снова рассмотрим дерево, изображенное на рисунке 15.45.

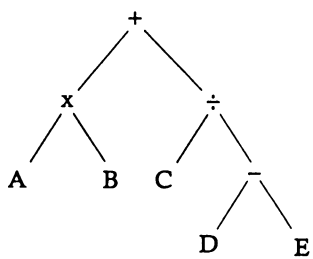


Рис. 15.45

Начинаем с ООП(+); первая инструкция вызывает ООП(×). Первая инструкция в ООП(×) вызывает ООП(A). Поскольку A не имеет левого сына, переходим к шагу 2 в ОПП(A). Но так как нет и правого сына, переходим к шагу 3 в ОПП(A), на котором обрабатываем A и возвращаемся к ООП(×). Опять замечаем, что, как и в случае ОЦП(A) и ОПП(A), результатом вызова команды ООП(A) является обработка (печать) листа A и возврат к родителю. Поэтому впредь, если имеется команда ООП(L), где L — лист, то три шага соответствующего алгоритма будем сокращенно описывать как **обработать лист L**. Находясь на шаге 2 команды ООП(×), вызываем ООП(B). Обрабатываем лист B и возвращаемся на шаг 3 в (×). Обрабатываем × и, поскольку это завершает ООП(×), возвращаемся к ООП(+). На этом этапе имеем обработанным

$$AB\times .$$

Теперь мы на шаге 2 в ООП(+) вызываем ООП(÷). Первый шаг ООП(÷) вызывает ООП(C). Обрабатываем лист C и возвращаемся на шаг 2 в ООП(÷). Эта

инструкция вызывает ООП(−). Певый шаг команды ООП(−) вызывает ООП(D). Обрабатываем лист D и возвращаемся на шаг 2 в ООП(−). Второй шаг в ООП(−) вызывает ООП(E). Обрабатываем E и возвращаемся на шаг 3 в ООП(−). Обрабатываем − и возвращаемся на шаг 3 в ООП(÷). На этом этапе результатом обработки является

$$AB \times CDE - .$$

На шаге 3 в ООП(÷) обрабатываем ÷ и возвращаемся на шаг 3 в (+). Обрабатываем + и завершаем процесс. В качестве результата обработки имеем

$$AB \times CDE - \div +,$$

т.е. выражение, преобразованное в **постфиксную запись**.

Умение обходить дерево дает возможность легко устанавливать изоморфность двух деревьев. Проходя два дерева одновременно, можно проверять наличие вершины у одного дерева при условии наличия ее у другого. При одновременном прохождении двух деревьев обработка вершин будет означать проверку обоих деревьев на наличие соответствующих вершин. В данном случае для прохождения деревьев представляется разумным использовать префиксный обход, так как это позволяет не интересоваться наличием сыновей, если один из родителей отсутствует. При обработке вершины V будем присваивать переменной *наличие*(V) значение N, если вершина V отсутствует, и значение Y, если вершина имеется. Последующая проверка соответствующих вершин покажет, что если для одной вершины значение переменной *наличие*(V) равно Y, а для другой вершины оно равно N, то два дерева не изоморфны.

Введем переменную, названную *Изо*, которая будет принимать значения T и F. Изначально ей присваивается значение T, но если на каком-то этапе выявляется, что деревья не изоморфны, то этой переменной присваивается значение F. Итак, имеем следующий алгоритм для двух деревьев T_1 и T_2 с корнями r_1 и r_2 соответственно.

Пусть *Изо* $= T$.

Алгоритм проверки изоморфности бинарных деревьев — ИБД(r_1, r_2)

(1) Обработать r_1 и r_2.

(2) Если *наличие*(r_1) $= Y$ и *наличие*(r_2) $= N$ или *наличие*(r_1) $= N$ и *наличие*(r_2) $= Y$, то *Изо* $= F$.

(3) Если *Изо* $= F$, то завершить ИБД(r_1, r_2).

(4) Если *наличие*(r_1) $= Y$ и *наличие*(r_2) $= Y$, то ИБД(лс(r_1), лс(r_2)).

(5) Если *наличие*(r_1) $= Y$ и *наличие*(r_2) $= Y$, то ИБД(пс(r_1), пс(r_2)).

Испытаем этот алгоритм на деревьях T_1 и T_2, изображенных на рис. 15.46.

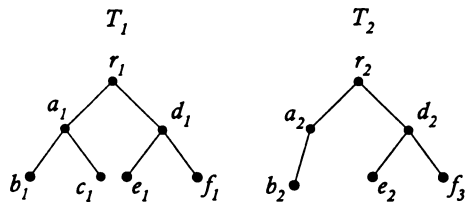


Рис. 15.46

Обрабатываем r_1 и r_2. Поскольку *наличие*$(r_1) = Y$ и *наличие*$(r_2) = Y$, переходим к шагу 3 ИБД(r_1, r_2). Здесь вызываем ИБД(лс(r_1), лс(r_2)), т.е. обрабатываем a_1 и a_2. Поскольку *наличие*$(a_1) = Y$ и *наличие*$(a_2) = Y$, переходим к шагу 3 алгоритма ИБД(a_1, a_2). Здесь вызываем ИБД(лс(a_1), лс(a_2)), т.е. обрабатываем b_1 и b_2. Поскольку *наличие*$(b_1) = Y$ и *наличие*$(b_2) = Y$, переходим к шагу 3 алгоритма ИБД (a_1, a_2). Здесь вызываем ИБД(лс(b_1), лс(b_2)), т.е. обрабатываем лс(b_1) и лс(b_2). А так как *наличие*(лс(b_1)) $= N$ и *наличие*(лс(b_2)) $= N$, возвращаемся к шагу 4 алгортма ИБД(a_1, a_2), который вызывает ИБД(пс(a_1), пс(a_2)). Обрабатываем c_1 и пс(a_2). Но *наличие*$(c_1) = Y$ и *наличие*(пс(a_2)) $= N$, поэтому *Изо* $= F$, и процесс прекращается. Таким образом, два рассмотренных дерева не изоморфны.

■ УПРАЖНЕНИЯ

1. Для дерева, изображенного на рис. 15.47,
- **а)** приведите инфиксную запись вершин дерева;
- **б)** приведите префиксную запись вершин дерева;
- **в)** приведите постфиксную запись вершин дерева.

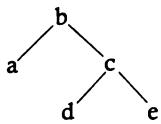


Рис. 15.47

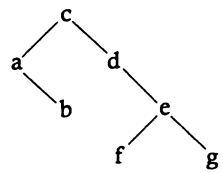


Рис. 15.48

2. Для дерева, изображенного на рис. 15.48,
- **а)** приведите инфиксную запись вершин дерева;
- **б)** приведите префиксную запись вершин дерева;
- **в)** приведите постфиксную запись вершин дерева.

3. Для дерева, изображенного на рис. 15.49,
- **а)** приведите инфиксную запись вершин дерева;
- **б)** приведите префиксную запись вершин дерева;
- **в)** приведите постфиксную запись вершин дерева.

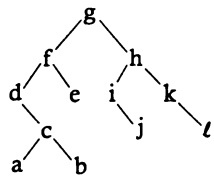


Рис. 15.49

4. Для дерева, изображенного на рис. 15.50,
- **а)** приведите инфиксную запись вершин дерева;

б) приведите префиксную запись вершин дерева;
в) приведите постфиксную запись вершин дерева.

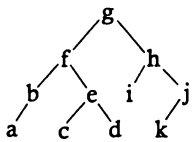


Рис. 15.50

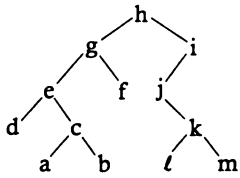


Рис. 15.51

5. Для дерева, изображенного на рис. 15.51,
а) приведите инфиксную запись вершин дерева;
б) приведите префиксную запись вершин дерева;
в) приведите постфиксную запись вершин дерева.
6. Для дерева, изображенного на рис. 15.52,
а) приведите инфиксную запись вершин дерева;
б) приведите префиксную запись вершин дерева;
в) приведите постфиксную запись вершин дерева.

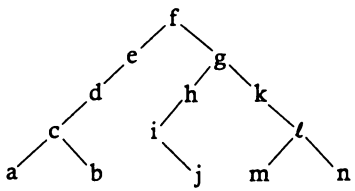


Рис. 15.52

7. Дана инфиксная запись $(a + b) \times c$:
а) найдите бинарное дерево, задающее это выражение;
б) укажите префиксную запись этого выражения;
в) укажите постфиксную запись этого выражения.
8. Дана инфиксная запись $(a + b) \times (c + d)$:
а) найдите бинарное дерево, задающее это выражение;
б) укажите префиксную запись этого выражения;
в) укажите постфиксную запись этого выражения.
9. Дана инфиксная запись $((a + b) \times c) \div (c + d)$:
а) найдите бинарное дерево, задающее это выражение;
б) укажите префиксную запись этого выражения.
в) укажите постфиксную запись этого выражения.
10. Дана инфиксная запись $(a \div b) + (c \times (d + e))$:
а) найдите бинарное дерево, задающее это выражение;
б) укажите префиксную запись этого выражения;
в) укажите постфиксную запись этого выражения.

11. Дана инфиксная запись $(a \times (b + (d - (e + f)))) \div ((g + h) \times (i + j))$:
- **а)** найдите бинарное дерево, задающее это выражение;
- **б)** укажите префиксную запись этого выражения;
- **в)** укажите постфиксную запись этого выражения.

12. Дана инфиксная запись $(a + b) \times ((((e - f) + g) \div h) - i)$:
- **а)** найдите бинарное дерево, задающее это выражение;
- **б)** укажите префиксную запись этого выражения;
- **в)** укажите постфиксную запись этого выражения.

13. Дана инфиксная запись $(((a + b) - c) \times d) \div (e - (f \times (g + h)))$:
- **а)** найдите бинарное дерево, задающее это выражение;
- **б)** укажите префиксную запись этого выражения;
- **в)** укажите постфиксную запись этого выражения.

14. Найдите такое бинарное дерево, обход которого в центрированном порядке дает *acbde*, а обход в обратном порядке дает *abced*.

15. Найдите такое бинарное дерево, обход которого в центрированном порядке дает *acbde*, а обход в обратном порядке дает *acedb*.

16. Найдите такое бинарное дерево, обход которого в центрированном порядке дает *acbdefg*, а обход в обратном порядке дает *dcbfeg*.

17. Найдите такое бинарное дерево, обход которого в центрированном порядке дает *acbdefg*, а обход в обратном порядке дает *edcabfg*.

18. Найдите дерево, содержащее не менее трех вершин, инфиксная запись которого совпадает с префиксной.

19. Найдите дерево, содержащее не менее трех вершин, инфиксная запись которого совпадает с постфиксной.

20. Постройте три неизоморфных бинарных дерева, которые при обходе дают постфиксное выражение *abcd*.

21. Постройте три неизоморфных бинарных дерева, которые при обходе дают префиксое выражение *abcd*.

22. Докажите, что задание для дерева инфиксного и постфиксного выражений однозначно определяет его структуру.

23. Если все арифметические операции бинарные, то какое из приведенных ниже выражений истинно для дерева, задающего данное арифметическое выражение?
- **а)** Бинарное дерево сбалансировано.
- **б)** Бинарное дерево полное.
- **в)** Лист может быть арифметическим выражением.
- **г)** Внутренняя вершина должна быть арифметическим выражением.

24. Напишите алгоритм, который правильным образом расставляет скобки в инфиксном арифметическом выражении, полученном при обходе дерева в центрированном порядке.

25. Сформулируйте необходимые и достаточные условия, при выполнении которых выражение было бы постфиксным арифметическим выражением.

26. Арифметическая операция возведения в степень является бинарной или унарной? Почему? Приведите пример унарной арифметической операции.

27. Опишите, где и каким образом алгоритм ИБД покажет, что деревья, изображенные на рис. 15.53, не изоморфны.

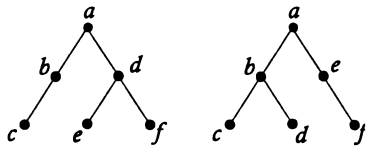


Рис. 15.53

15.5. ОСТОВНЫЕ ДЕРЕВЬЯ

В разделе 6.3 ***остовное дерево*** было определено как дерево T, которое является подграфом графа G и таким, что каждая вершина в G является вершиной в T; показано также, что каждый связный граф имеет остовное дерево.

В данном разделе рассматриваются два метода построения остовного дерева. Первый — метод поиска в ширину, а второй — метод поиска в глубину. Начнем с метода поиска остовного дерева в ширину. Согласно методу произвольную вершину v_0 графа G выбираем в качестве корня дерева T. Для каждой вершины v графа G, смежной с вершиной v_0, в дерево T добавляется вершина v и ребро $\{v, v_0\}$. Это вершины уровня 1. Затем берем каждую вершину v_i уровня 1 и для каждой вершины v_j графа G, смежной с вершиной v_i из тех, что еще не выбраны, добавляем в дерево T вершину v_j и ребро $\{v_i, v_j\}$. Вершины, добавленные на этом этапе, — это вершины уровня 2. Продолжаем процесс, пока в графе G не останется вершин, которые можно было бы добавить в дерево. По построению T является деревом. Если расстояние от v_0 до вершины v графа G равно n, то эта вершина будет добавлена в дерево на уровне n. Следовательно, T несомненно является остовным деревом.

При построении алгоритма требуется наверняка знать, что некоторые вершина и ребро уже были добавлены в дерево, с тем чтобы, с одной стороны, описывать дерево, а с другой стороны, — исключать возможность повторного использования вершин, так как в противном случае мы не получим дерево. Требуется также знать уровень, на котором добавляется вершина, поскольку для определения листьев на следующем уровне эффективнее использовать только вершины, добавленные на последнем уровне. При определении новых вершин для добавления в дерево необходимо также иметь возможность выбора вершин по одной. Существует несколько способов решения такого рода задач, в том числе решения с использованием массивов и очередей, поэтому мы не будем подробно останавливаться на том, как определить "следующую" вершину. Можно считать, что вершины графа G предварительно упорядочены. Пусть V и E обозначают множество вершин и множество ребер графа G, и пусть V^T и E^T обозначают множество вершин и множество ребер дерева T. При добавлении в дерево вершины v мы будем обозначать через $L(v)$ уровень, на котором она добавляется.

Алгоритм поиска остовного дерева в ширину — ПОДШ(G)

(1) Выбрать произвольный элемент v_0 графа G. Пусть $v_0 \in V^T$ и $L(v_0) = 0$.
(2) Для всех $v \in V - V^T$ таких, что v смежна с v_0, положить $v \in V^T$, $\{v_0, v\} \in E^T$ и $L(v) = 1$.
(3) Пусть $i = 1$.
(4) Выбрать $v_j \in V^T$ такое, что $L(v_j) = i$.
(5) Выбрать $v \in V - V^T$. Если v смежна с v_j, положить $v \in V^T$, $\{v_0, v\} \in E^T$ и $L(v) = i + 1$.
(6) Продолжать шаг 5, пока все элементы множества $V - V^T$ не будут рассмотрены. (Обратите внимание, что $V - V^T$ постоянно меняется.)
(7) Повторять шаги 4, 5, и 6, пока все v_j такие, что $L(v_j) = i$, не будут выбраны.
(8) Положить $i = i + 1$.
(9) Повторять шаги 4–8 до $V = V^T$.

ПРИМЕР 15.21. Рассмотрим граф, изображенный на рис. 15.54. Предположим, что вершина v_0 выбрана в качестве первой вершины. Тогда $L(v_0) = 0$ и $v_0 \in V^T$. Поскольку v_1 является смежной с v_0, положим $v_1 \in V^T$, $\{v_0, v_1\} \in E^T$ и $L(v_1) = 1$. Вершина v_2 также смежна с v_0, поэтому положим $v_2 \in V^T$, $\{v_0, v_2\} \in E^T$ и $L(v_2) = 1$. Наконец, вершина v_3 тоже смежна с v_0, поэтому полагаем $v_3 \in V^T$, $\{v_0, v_3\} \in E^T$ и $L(v_3) = 1$. На данный момент имеем дерево, изображенное на рис. 15.55.

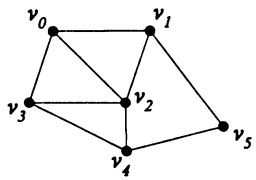


Рис. 15.54

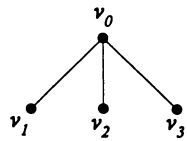


Рис. 15.55

Теперь рассматриваем вершины v такие, что $L(v) = 1$. Начинаем с v_1 и находим неиспользованные вершины, смежные с v_1. Поскольку вершина v_5 смежна с v_1, полагаем $v_5 \in V^T$, $\{v_1, v_5\} \in E^T$ и $L(v_5) = 2$. Больше нет неиспользованных и смежных с v_1 вершин, поэтому переходим к v_2. Поскольку v_4 смежна с v_2, полагаем $v_4 \in V^T$, $\{v_2, v_4\} \in E^T$ и $L(v_4) = 2$. Теперь все вершины использованы, процесс завершен. В результате имеем дерево, изображенное на рис. 15.56.

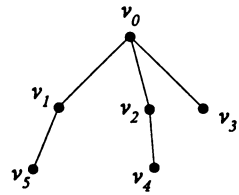


Рис. 15.56

□

ПРИМЕР 15.22. Рассмотрим граф, изображенный на рис. 15.57.

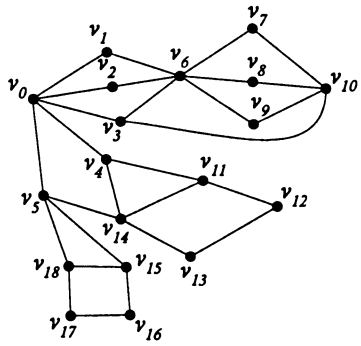


Рис. 15.57

Как и раньше, выберем v_0 как начальную вершину. Тогда $L(v_0) = 0$ и $v_0 \in V^T$. Поскольку v_1, v_2, v_3, v_4, v_5 являются смежными с v_0, положим $v_i \in V^T$, $\{v_0, v_i\} \in E^T$ и $L(v_i) = 1$, где $1 \leq i \leq 5$. После этого получаем дерево, изображенное на рис. 15.58.

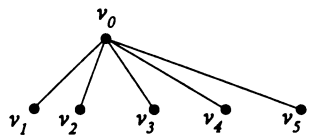


Рис. 15.58

Теперь начинаем на уровне 1 с вершины v_1. На этом уровне находятся: вершина v_6, смежная с v_1, вершина v_{10}, смежная с v_3, вершины v_{11} и v_{14}, смежные с v_4, и v_{15}, v_{18}, смежные с v_5. Поэтому $v_6, v_{10}, v_{11}, v_{14}, v_{15}, v_{18} \in V^T$, $L(v_6) = L(v_{10}) = L(v_{11}) = L(v_{14}) = L(v_{15}) = L(v_{18}) = 2$ и $\{v_1, v_6\}, \{v_3, v_{10}\}, \{v_4, v_{11}\}, \{v_4, v_{14}\}, \{v_5, v_{15}\}, \{v_5, v_{18}\} \in E^T$ В результате получаем дерево, изображенное на рис. 15.59.

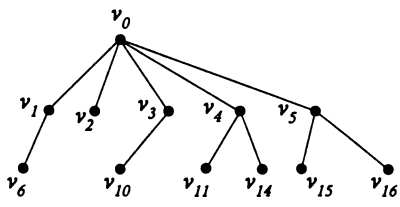


Рис. 15.59

Теперь начинаем на уровне 2. На этом уровне: вершины v_7, v_8, v_9, смежные с v_6, v_{12}, смежная с v_{11}, v_{13}, смежная с v_{14}, v_{16}, смежная с v_{15}, и v_{17}, смежная с v_{18}. Поэтому $v_7, v_8, v_9, v_{12}, v_{13}, v_{16}, v_{17} \in V^T$, $\{v_6, v_7\}, \{v_6, v_8\}, \{v_6, v_9\}, \{v_{11}, v_{12}\}$, $\{v_{14}, v_{13}\}, \{v_{16}, v_{13}\}, \{v_{17}, v_{18}\} \in E^T$, и $L(v_7) = L(v_8) = L(v_9) = L(v_{12}) = L(v_{13}) = L(v_{16}) = L(v_{17}) = 3$. В результате получаем дерево, изображенное на рис. 15.60.

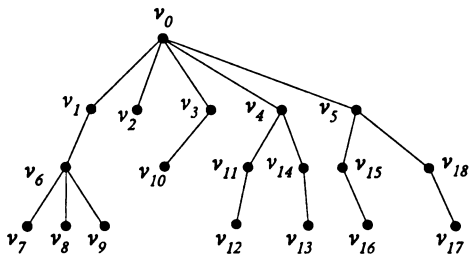


Рис. 15.60 □

При поиске в ширину первым делом отыскиваются все вершины, смежные с заданной вершиной, а потом осуществляется переход на следующий уровень. При поиске в ширину усилия направлены на построение для дерева как можно более длинного пути. Когда путь заходит в тупик и формирует лист, необходимо возвращаться к родителю листа и пытаться сформировать другой путь. Возврат к родителю происходит только после попытки построить все возможные пути от сына этого родителя. В образах создаваемого дерева это означает, что мы пытаемся достичь самый глубокий, или самый большой уровень, какой только возможен. Алгоритм начинается у заданной вершины графа, которую будем считать корнем. Выбираем вершину, смежную с корнем, и формируем ребро дерева. Затем выбираем вершину, смежную с ранее выбранной вершиной, и формируем новое ребро. По ходу процесса необходимо помечать использованные вершины, с тем чтобы каждая вершина использовалась один раз. Если, находясь в вершине v, мы выбираем другую вершину w и обнаруживаем, что вершина w уже была добавлена в дерево, то ребро $\{v, w\}$ между этими вершинами не может быть добавлено в дерево. Такое ребро назовем ***обратным ребром***, и будем хранить его след для дальнейшего использования. Прежде, чем объявить ребро обратным, нужно тщательно проверить, не является ли вершина w родителем вершины v, поскольку в этом случае ребро $\{v, w\}$ уже присутствует в дереве. По существу, при выборе w как возможной новой вершины следует избегать ситуации, когда вершина w является родителем v. Если $\{v, w\}$ — обратное ребро, то мы остаемся в v и выбираем новую смежную вершину, если это возможно. По определению, любое ребро графа, в конечном счете, — либо ребро дерева, либо обратное ребро.

В рассматриваемом алгоритме множество ребер дерева назовем РЕБРА ДЕРЕВА, а множество обратных ребер — ОБРАТНЫЕ РЕБРА. В самом начале предположим, что для "новых" вершин использованы метки "*нов*", а после добавления вершины в дерево эти метки меняются на "*исп*" — "использованные".

Алгоритм поиска остовного дерева в глубину — ПОДГ(G)

(1) Пометить каждую вершину графа G символом n.

(2) Выбрать произвольный элемент v_0 графа G и назвать его корнем дерева.

(3) Заменить метку вершины v_0 с "*нов*" на "*исп*" и положить $v = v_0$

(4) Пока имеются невыбранные вершины, смежные с v, выполнять следующие действия:

а) Выбрать вершину w, смежную с v.

б) Если w имеет метку "*нов*", добавить (v, w) в РЕБРА ДЕРЕВА, поменять метку w на "*исп*", положить $w = v$ и повторить шаг 4.

в) Если w имеет метку "*исп*" и не является родителем v, добавить (v, w) в ОБРАТНЫЕ РЕБРА и повторить шаг 4.

(5) Если $v \neq a$, положить $v = (v)$ и повторить шаг 4.

Заметим, что пока для вершины v имеется смежная с ней неиспользованная вершина w, мы продлеваем путь от v к w. Только в том случае, когда двигаться дальше нельзя, переходим к шагу 5, возвращаемся к родителю вершины v.

ПРИМЕР 15.23. Рассмотрим граф, изображенный на рис. 15.61.

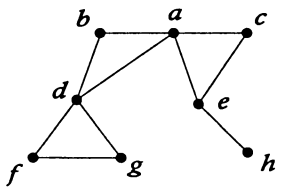


Рис. 15.61

Произвольным образом выбираем вершину a в качестве корня. Меняем метку a с "*нов*" на "*исп*". Поскольку вершина b смежна с a и имеет метку "*нов*", добавляем ребро $\{a, b\}$ в РЕБРА ДЕРЕВА и меняем метку вершины b на "*исп*", как показано на рис. 15.62.

От вершины b переходим к вершине d, поскольку она является смежной с b. Вершина d имеет метку "*нов*", поэтому добавляем ребро $\{b, d\}$ в РЕБРА ДЕРЕВА и меняем метку вершины d на "*исп*", как показано на рис. 15.63.



Рис. 15.62

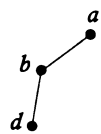


Рис. 15.63

Теперь выбираем вершину, смежную с вершиной d. Можно выбрать a, f или g. Выбор определит форму дерева, поэтому поиск в глубину не продуцирует дерево единственным образом. Предположим, что следующей вершиной выбираем g. Вершина g имеет метку "*нов*", поэтому добавляем ребро $\{d, g\}$ в РЕБРА ДЕРЕВА и меняем метку вершины g на "*исп*", как показано на рис. 15.64.

Из вершины g выбираем вершину f, так как она является смежной с g. Вершина f имеет метку "*нов*", поэтому добавляем ребро $\{g, f\}$ в РЕБРА ДЕРЕВА и меняем метку вершины f на "*исп*", как показано на рис. 15.65.

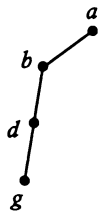


Рис. 15.64

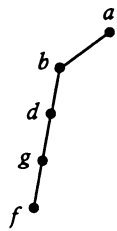


Рис. 15.65

Из вершины f выбираем вершину d, так как она является смежной с вершиной f. Однако вершина d уже имеет метку "*исп*", поэтому добавляем ребро $\{d, f\}$ в ОБРАТНЫЕ РЕБРА, как показано на рис. 15.66.

Поскольку больше нет ребер для проверки на смежность с вершиной f, кроме родителя, возвращаемся к вершине g. (Далее не будем упоминать родителя как возможную смежную вершину.) Но иных ребер для проверки на смежность с вершиной g тоже нет, поэтому возвращаемся к вершине d. Здесь единственной вершиной для проверки является a, но вершина a уже имеет метку "*исп*", поэтому ребро $\{d, f\}$ добавляем в ОБРАТНЫЕ РЕБРА и возвращаемся в вершину b, как показано на рис. 15.67.

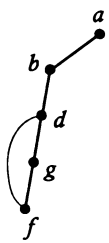


Рис. 15.66

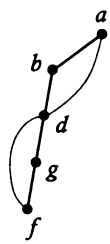


Рис. 15.67

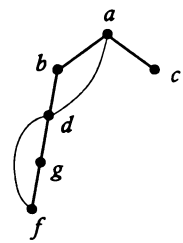


Рис. 15.68

Ребер для проверки на смежность с вершиной b больше нет; возвращаемся к вершине a. Заметим, что если бы была возможность достичь каждую вершину, построив путь из b, то пока дерево не построено полностью, нам не надо было бы возвращаться в вершину a. Вернувшись в вершину a, можем выбирать c или e. Предположим, что выбрана вершина c. А так как вершина c имеет метку "*нов*", добавляем ребро $\{a, c\}$ в РЕБРА ДЕРЕВА и меняем метку вершины c на "*исп*", как показано на рис. 15.68.

Вершина e смежна с вершиной c и имеет метку "*нов*"; добавляем ребро $\{c, e\}$ в РЕБРА ДЕРЕВА и меняем метку e на "*исп*", как показано на рис. 15.70. Допустим теперь, что выбрана вершина a, потому что она является смежной с вершиной e. Но вершина a уже имеет метку "*исп*", поэтому добавляем ребро $\{e, a\}$ в ОБРАТНЫЕ РЕБРА и возвращаемся в вершину e, как показано на рис. 15.70.

Поскольку вершина h смежна с e, добавляем ребро $\{e, h\}$ в РЕБРА ДЕРЕВА и меняем метку вершины h на "*исп*", как показано на рис. 15.71.

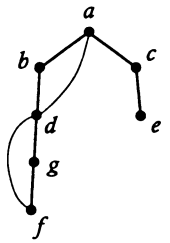


Рис. 15.69

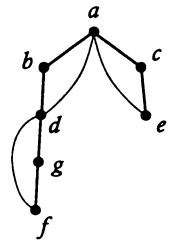


Рис. 15.70

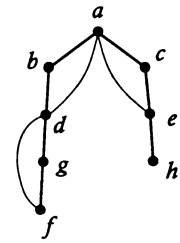


Рис. 15.71

Больше нет вершин для проверки из вершины h, поэтому возвращаемся в вершину e. Но больше нет вершин для проверки из вршины e, поэтому возвращаемся в вершину c. Поскольку нет больше вершин для проверки из вершины c, возвращаемся в вершину a. Других вершин для проверки из вершины a тоже нет, поэтому процесс завершен. □

ПРИМЕР 15.24. Найти остовное дерево графа, изображенного на рис. 15.72.

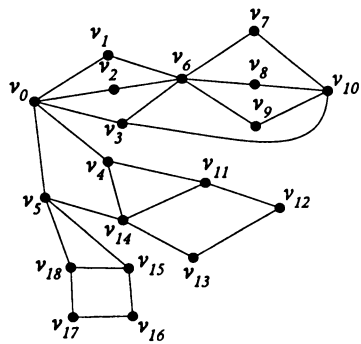


Рис. 15.72

Предположим, что все вершины графа имеют метку "*нов*". Пусть v_0 — первая выбранная вершина. Добавим ребра (v_0, v_5), (v_5, v_{18}), (v_{18}, v_{17}), (v_{17}, v_{16}) и (v_{16}, v_{15}) в РЕБРА ДЕРЕВА и заменим метки вершин $v_0, v_5, v_{18}, v_{17}, v_{16}$ и v_{15} на u. Из вершины v_{15} находим, что вершины v_5 и v_{18} — смежные с вершиной v_{15}, но они уже отмечены меткой "*исп*", поэтому ребра (v_{15}, v_{18}) и (v_{15}, v_5) добавляются в ОБРАТНЫЕ РЕБРА. Теперь последовательно возвращаемся в v_{16}, v_{17}, затем в v_{18} и, наконец, в v_5. Находясь в v_5, добавляем ребра (v_5, v_{14}), (v_{14}, v_{13}), (v_{13}, v_{12}), (v_{12}, v_{11}) и (v_{11}, v_4) в РЕБРА ДЕРЕВА и меняем метки вершин $v_{14}, v_{13}, v_{12}, v_{11}$ и v_4 на u. Вершины v_0 и v_{14} — смежные с вершиной v_4, но они уже имеют метку "*исп*", поэтому ребра (v_4, v_0) и (v_4, v_{14}) добавляются в ОБРАТНЫЕ РЕБРА. Возвращаемся в вершину v_{11}, откуда находим, что вершина v_{14} является смежной с v_{11}. Но v_{14} уже отмечена меткой "*исп*", поэтому добавляем ребро (v_{11}, v_{14}) в ОБРАТНЫЕ РЕБРА. Осуществляем возврат, пока не достигаем вершины v_0. На этот момент имеем дерево с обратными ребрами, изображенное на рис. 15.73.

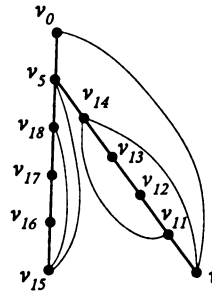


Рис. 15.73

Находясь в вершине v_0, выбираем вершину, смежную с v_0, например, v_3, добавляем ребро (v_0, v_3) в РЕБРА ДЕРЕВА и меняем метку вершины v_3 на "*исп*". Продолжаем процесс, добавляя $(v_3, v_6), (v_6, v_9), (v_9, v_{10})$ и (v_{10}, v_8) в РЕБРА ДЕРЕВА, и меняем метки вершин v_6, v_9, v_{10} и v_8 на "*исп*". Теперь мы находимся в вершине v_8. Вершина v_6 является смежной с v_8, но v_6 уже отмечена меткой "*исп*", поэтому добавляем ребро (v_8, v_6) в ОБРАТНЫЕ РЕБРА. Возвращаясь в v_{10}, находим, что вершина v_7 — смежная с v_{10}, поэтому добавляем (v_{10}, v_7) в РЕБРА ДЕРЕВА, меняя метку вершины v_7 на u. Вершина v_7 — смежная с вершиной v_6, но v_6 уже отмечена меткой "*исп*", поэтому добавляем ребро (v_7, v_6) в ОБРАТНЫЕ РЕБРА. Возвращаясь в вершину v_{10}, находим, что v_3 — смежная с v_{10}. Вершина v_3 уже отмечена меткой u, поэтому добавляем ребро (v_{10}, v_3) в ОБРАТНЫЕ РЕБРА. Осуществляем возврат в вершину v_6. Вершина v_1 — смежная с v_6, добавляем ребро (v_6, v_1) в РЕБРА ДЕРЕВА, меняя метку вершины v_1 на "*исп*". Находясь в вершине v_1, определяем, что вершина v_0 — смежная с v_1, и поскольку v_0 уже отмечена меткой "*исп*", добавляем ребро (v_1, v_0) в ОБРАТНЫЕ РЕБРА. Возвращаясь в v_6, находим, что вершина v_2 — смежная с v_6, поэтому добавляем ребро (v_6, v_2) в РЕБРА ДЕРЕВА и меняем метку вершины v_2 на "*исп*". Из вершины v_2 находим, что вершина v_0 — смежная с вершиной v_2 и отмечена меткой "*исп*", добавляем ребро (v_2, v_0) в ОБРАТНЫЕ РЕБРА. Построение дерева закончено, и на рис. 15.74 оно изображено вместе с обратными ребрами.

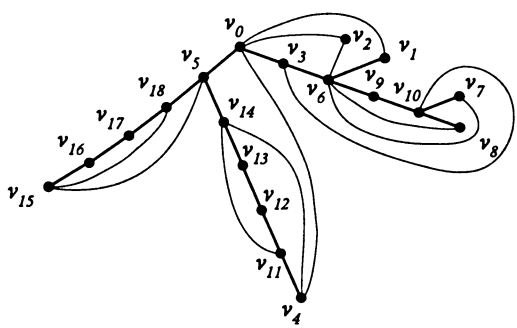


Рис. 15.74 □

Если граф не связный, то алгоритм ПОДГ будет применяться не ко всему графу, а только к его компоненте, содержащей вершину, выбранную в качестве корня. По этой причине алгоритм поиска остовного дерева в глубину можно использовать для проверки связности графа, для чего необходимо просто реализовывать алгоритм. Если какая-либо вершина не достижима, то рассматриваемый граф

не является связным. Алгоритм также может быть использован для построения остовного дерева каждой компоненты несвязного графа. Следует, как и прежде, выбрать вершину для корня первого дерева и построить остовное дерево. Затем выбрать одну из недостижимых вершин и построить второе дерево. Продолжать процесс, пока не останется вершин. Лес остовных деревьев, построенных таким образом, называется ***остовным лесом***.

Алгоритм построения остовного дерева в глубину можно также использовать для проверки графа на наличие точек сочленения. Для этого потребуется следующая теорема.

ТЕОРЕМА 15.25. Если T — глубинное остовное дерево графа $G(V, E)$ и $\{a, b\}$ — ребро графа $G(V, E)$, то либо a является потомком b, либо b — потомком a.

ДОКАЗАТЕЛЬСТВО. Если $\{a, b\}$ — ребро дерева T, то вывод очевиден. Если это не так, то одна из вершин, например, a, должна быть помещена в дерево первой. Но поскольку вершина b не была помещена в дерево на шаге 4 алгоритма ПОДГ, то поиск продолжается из вершины a до тех пор, пока не будет найдена вершина b. Поэтому вершина b является потомком a. ■

ТЕОРЕМА 15.26. Пусть T — глубинное остовное дерево графа $G(V, E)$. Вершина $a \in V$ является точкой сочленения графа $G(V, E)$ тогда и только тогда, когда вершина a (1) либо является корнем дерева T и имеет более одного сына, либо (2) вершина a не является корнем дерева T, и существует такой сын c, что между c, или одним из его потомков, и собственным предком вершины a не существует обратного ребра.

ДОКАЗАТЕЛЬСТВО. (1) Если a — корень дерева T и имеет только одного сына c, то между любой вершиной дерева T и вершиной c существует путь. Поэтому для двух вершин v и w, ни одна из которых не совпадает с a, существует путь из v в c в w, который не проходит через a, и, следовательно, a не является точкой сочленения. Если корень a имеет двух сыновей, c и c', и если поиск в глубину начинается от c, он может достичь c' только путем возврата в вершину a; поэтому не существует пути из c в c', который не проходит через a.

(2) Предположим, что существует сын c вершины a такой, что между c или одним из его потомков и собственным предком вершины a не существует обратное ребро. Пусть p — родитель a. Поскольку не существует обратного ребра между c или одним из его потомков и собственным предком вершины a, по предыдущей теореме каждое ребро из a или одного из ее потомков должно идти в a или одного из ее потомков. Таким образом, каждый путь, имеющий начало в вершине c, должен заканчиваться в c или в одном из ее потомков либо проходить через a. Поэтому единственный путь из c в p проходит через a, и a является точкой сочленения. Обратно, предположим, что для всех сыновей c вершины a существует обратное ребро из c в собственного предка вершины a. Пусть c и c' — сыновья вершины a, а b и b' — собственные предки вершины a, так что $e = \{d, b\}$ и $e = \{d', b'\}$ — обратные ребра, где d и d' — потомки вершин c и c' соответственно. Таким образом, существует путь $p_{c,b}$ из c в d в b и путь $p_{c',b'}$ из c' в d' в b'. Поскольку существует путь p из b в b', который не проходит через a, путь $p_{c,b}$, за которым следует путь p и путь, обратный к $p_{c',b'}$, является путем из c

в c', который не проходит через a. Пусть v и w — две вершины графа $G(V, E)$. Если v и w — потомки вершины a, то из вершин v и w в c и c' соответственно существуют пути, которые не проходят через a, где c и c' — сыновья вершины a. Поскольку существует также путь из c и c', который не проходит через a, то существует путь из v в c, далее в c' и далее в w, который не проходит через a. Если v — потомок вершины a, а w — нет, то поскольку существует путь из v к сыну c вершины a, путь $p_{c,b}$ из c в b, где b — собственный предок вершины a, и путь вдоль дерева из вершины b в w, и ни один из этих путей не содержит a, то существует путь из v в w, не содержащий a. Если ни v, ни w не являются потомками вершины a, то существует путь вдоль дерева из v в w, не содержащий a. Таким образом, всегда существует путь из v в w, не содержащий a, и поэтому a не является точкой сочленения. ■

Для нахождения точек сочленения графа G используем предыдущую теорему и попытаемся выяснить, имеется ли в дереве T сын c такой, что между c или одним из его потомков и собственным предком вершины a не существует обратного ребра. Прежде всего нам нужен такой способ упорядочения вершин, при котором если вершина a является собственным предком вершины b, то значение, приписываемое a, больше значения, приписываемого b. Это можно сделать путем присваивания вершинам порядкового номера их поступления в дерево. Тогда, если c является n-ой вершиной, попавшей в дерево, то ей будет присвоено значение счета n, которое мы обозначим через $ЗС(c)$. Используя это упорядочение, присвоим c наименьшее значение вершины a такой, что между a и c или одним из ее потомков существует обратное ребро. Обозначим это значение через $ОР(c)$ и назовем его $ОР$-значением вершины c. Как отмечалось раннее, при поиске в глубину, пока у вершины v имеется смежная с ней неиспользованная вершина w, мы проходим путь от v в w. И только в случае, когда нельзя продвигаться дальше, мы покидаем шаг 4. Желательно, чтобы новый алгоритм делал то же самое, но, кроме того, по мере добавления в дерево новой вершины w присваивал новое значение счета величине $ЗС(w)$. Временно положим $ОР(w) = ЗС(w)$. Будем определять $ОР$-значения для вершины w по восходящему принципу. Дойдя до листа l, сравним его $ОР$-значение со значением счетчика или $ЗС$-значения всех вершин, смежных с l, исключая его родителя, и присвоим $ОР(l)$ значение, минимальное из всех перечисленных. При возврате к родителю p сравним его текущее $ЗС$-значение с $ОР$-значением того сына c, от которого произошел возврат. Если $ОР(c) > ЗС(p)$, то никакой потомок вершины c не имеет обратного ребра к собственному предку вершины p, и согласно предыдущей теореме p является точкой сочленения.

Далее, положим $ОР(p)$ равным минимуму из его текущего значения и значения $ОР(c)$. Если новых ребер из p больше нет, положим $ОР(p)$ равным минимуму из его текущего значения и $ЗС$-значений всех вершин таких, что между p и этими вершинами существует обратное ребро. Продолжаем, пока не завершится поиск в глубину.

Заметим, что если достигнута точка сочленения a, и мы удалим те (еще неудаленные) ребра, вершиной которых является a и ее потомки, но оставим вершину a, то получим остовное дерево одной из компонент двусвязности графа G.

Формально процедура выражается следующим алгоритмом. Используем рекурсивное свойство, позволяющее алгоритму вызывать самого себя. Используем

счетчик, изначально установленный на 0. Как и ранее, предположим, что начальной меткой для каждой вершины будет метка "*нов*".

Алгоритм поиска точек сочленения — ПТС(v)

(1) Пометить v символом "*исп*".
(2) Присвоить i значение $i+1$.
(3) Положить $OP(v) = ЗС(v) = i$.
(4) Пока имеется невыбранная вершина, смежная с v, выполнять следующее:
  (а) Выбрать вершину w, которая является смежной с v.
  (б) Если w имеет метку "*нов*", то:
    (i) добавить (v, w) в РЕБРА ДЕРЕВА;
    (ii) положить $v = \text{parent}(w)$;
    (iii) вызвать ПТС(w);
    (iv) если $OP(w) \geq ЗС(v)$, то v — точка сочленения. Ребра, ниже v, которые охватывают компоненту, удаляются;
    (v) положить $OP(v) = \min(OP(v), OP(w))$.
  (г) Если w имеет метку "*исп*" и w не является $\text{parent}(w)$, то $OP(v) = \min(OP(v), ЗС(w))$.

ПРИМЕР 15.27. Рассмотрим граф, изображенный на рис. 15.75, и соответствующее дерево (рис. 15.76).

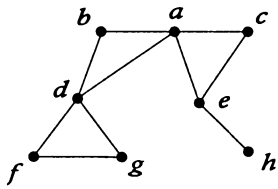


Рис. 15.75

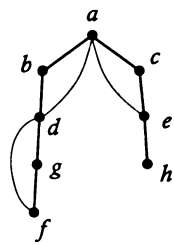


Рис. 15.76

Из пути $abdgf$ дерева определяем, что $OP(a) = ЗС(a) = 1$, $OP(b) = ЗС(b) = 2$, $OP(d) = ЗС(d) = 3$, $OP(g) = ЗС(g) = 4$ и $OP(f) = ЗС(f) = 5$. Из вершины f находим обратное ребро из f в d и присваиваем $OP(f) = ЗС(d) = 3$. Возвращаемся в вершину g, поскольку f — сын g, то $OP(g) = \min(OP(f), OP(g)) = \min(4, 3) = 3$. Поскольку $OP(g) = 3 \geq ЗС(d)$ и g — сын вершины d, d — точка сочленения, а дерево на рис. 15.77, сформированное ребрами $\{f, g\}$ и $\{d, g\}$, охватывает компоненту двусвязности (рис. 15.78).

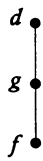


Рис. 15.77

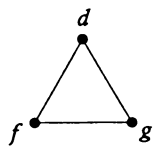


Рис. 15.78

Возвращаемся в вершину d, поскольку g — сын вершины d, полагаем $OP(d) = \min((OP(g), OP(d)) = \min(3,3) = 3$. Затем, так как существует обратное ребро из d в a, присваиваем $OP(d) = ЗС(a) = 1$. Возвращаемся в b. Поскольку d — сын вершины b, присваиваем $OP(b) = \min(OP(d), OP(b)) = \min(1,2) = 1$. Потому что $OP(b) \geq ЗС(a)$, вершина a — точка сочленения (a — корень с двумя сыновьями), и дерево (рис. 15.79), сформированное ребрами $\{d,b\}$ и $\{b,a\}$, охватывает компоненту двусвязности, изображенную на рис. 15.80.



Рис. 15.79

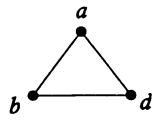


Рис. 15.80

Возвращаемся в вершину a; поскольку b — сын вершины a, сравниваем их OP-значения и находим, что $OP(a) = \min((OP(a), OP(b)) = \min(1,1) = 1$, поэтому $OP(a)$ остается неизменным. Из того, как сформирован путь $aceh$ дерева, находим $OP(c) = ЗС(c) = 6$, $OP(e) = ЗС(e) = 7$ и $OP(h) = ЗС(h) = 8$. А так как вершина h не имеет обратных ребер, $OP(h)$ остается неизменным. Из того, что $OP(h) \geq ЗС(e)$ и h — сын вершины e, следует, что e — точка сочленения и ребро $\{e,h\}$ — остовное дерево компоненты двусвязности, изображенной на рис. 15.81.

Возвращаемся в e; поскольку h — сын вершины e, имеем $OP(e) = \min(OP(e)$, $OP(h))\min(7,8) = 7$, поэтому $OP(e)$ остается неизменным. Но существует обратное ребро из e в a, значит $OP(e) = ЗС(a) = 1$. $OP(c) = \min(OP(c), OP(e)) = \min(1,6) = 1$. Поскольку $OP(e) \geq ЗС(a)$, то, как и ожидалось, вершина a — опять найдена как точка сочленения, дерево (рис. 15.82), сформированное ребрами $\{a,c\}$ и $\{c,e\}$, охватывает бикомпоненту, изображенную на рис. 15.83.



Рис. 15.81



Рис. 15.82

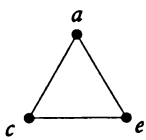


Рис. 15.83

Наконец, возвращаемся в вершину a, и процесс завершен. □

Заметим, что множество из n вершин ***размечено***, если каждой вершине приписано единственное целое число между 1 и n. Таким образом, имеем вершины $v_1, v_2, v_3, \ldots, v_n$. Если рассмотреть все графы, имеющие n размеченных вершин, то сколько остовных деревьев можно построить? Другими словами, сколько имеется остовных деревьев для n размеченных вершин? Ответ на этот вопрос дает формула Кэли. Это число остовных деревьев для K_n.

ТЕОРЕМА 15.28. (Формула Кэли для дерева) Число остовных деревьев для n размеченных вершин равно n^{n-2}.

ДОКАЗАТЕЛЬСТВО. Для доказательства утверждения покажем, что каждое остовное дерево однозначным образом можно представить в виде последовательности $a_1, a_2, a_3, \ldots, a_{n-2}$ длины $n-2$ такой, что для $1 \le k \le n-2$ элемент последовательности a_k может быть любым целым числом между 1 и n, и обратно, каждой последовательности $a_1, a_2, a_3, \ldots, a_{n-2}$ длины $n-2$ такой, что для $1 \le k \le n-2$ выполняется $1 \le a_k \le n$, соответствует единственное остовное дерево. Если это сделать, то вопрос сведется к тому, сколькими способами можно заполнить $n-2$ пробела, если имеется n способов заполнения каждого пробела, а таких способов, как известно, существует n^{n-2}. Вместо формального доказательства приведем алгоритм построения указанной последовательности по остовному дереву и алгоритм построения остовного дерева по соответствующей последовательности. Далее покажем на примере, что если начать с остовного дерева и использовать алгоритм построения последовательности, а затем применить к ней алгоритм построения остовного дерева, то получим исходное остовное дерево.

Начнем с алгоритма преобразования остовного дерева в последовательность.

Алгоритм перевода дерева в последовательность — ДвП(T) для $n \ge 3$

(1) Для выбора a_1 взять лист v_j с наименьшим значением j. Всегда существует единственное k такое, что $\{v_j, v_k\}$ является ребром дерева. Удалить это ребро и положить $a_1 = k$.

(2) Если выбрано a_{i-1}, то для выбора a_i из оставшегося дерева взять лист v_j с наименьшим значением j. Всегда существует единственное k такое, что $\{v_j, v_k\}$ является ребром дерева. Удалить это ребро и положить $a_i = k$.

(3) Продолжать, пока не будет выбрано a_{n-2}. ■

ПРИМЕР 15.29. Пусть T — дерево, изображенное на рис. 15.84. Вершина v_2 — лист с наименьшим индексом, поэтому удаляем ребро $\{v_2, v_1\}$ и полагаем $a_1 = 1$, как показано на рис. 15.85. Поскольку в оставшемся дереве вершина v_3 — лист с наименьшим индексом, удаляем ребро $\{v_3, v_4\}$ и полагаем $a_2 = 4$ (рис. 15.86).

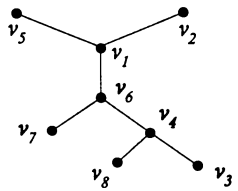


Рис. 15.84

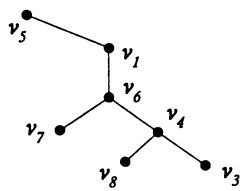


Рис. 15.85

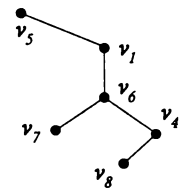


Рис. 15.86

Далее, v_5 — лист с наименьшим индексом, удаляем ребро $\{v_5, v_1\}$ и полагаем $a_3 = 1$, как показано на рис. 15.87. Поскольку в оставшемся дереве вершина v_1 — лист с наименьшим индексом, удаляем ребро $\{v_1, v_6\}$ и полагаем $a_4 = 6$, как показано на рис. 15.88. В оставшемся дереве вершина v_7 — лист с наименьшим индексом, поэтому удаляем ребро $\{v_7, v_6\}$ и полагаем $a_5 = 6$ (рис. 15.89).

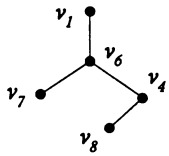


Рис. 15.87

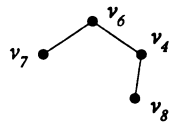


Рис. 15.88

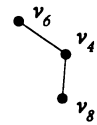


Рис. 15.89

Наконец, поскольку в оставшемся дереве v_6 — лист с наименьшим индексом, удаляем ребро $\{v_6, v_4\}$ и полагаем $a_6 = 4$. Осталось только ребро, изображенное на рис. 15.90. Поэтому искомая последовательность имеет вид

$$1, 4, 1, 6, 6, 4.$$

Отметим, что происходит число вхождений в последовательность индекса каждой вершины на единицу меньше ее степени. Это потому, что при удалении ребра вершина-лист не учитывается, а вторая его вершина учитывается как вершина в ребре, которое остается как лист, или как вершина в ребре, которое остается. □



Рис. 15.90

ПРИМЕР 15.30. Пусть T — дерево, изображенное на рис. 15.91. Поскольку v_2 — лист с наименьшим индексом, удаляем ребро $\{v_2, v_9\}$ и полагаем $a_1 = 9$, как показано на рис. 15.92. В оставшемся дереве v_3 — лист с наименьшим индексом, удаляем ребро $\{v_3, v_8\}$ и полагаем $a_2 = 8$, как показано на рис. 15.93.

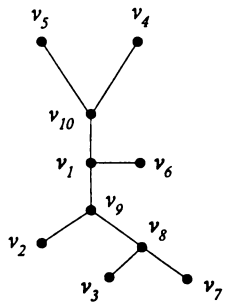


Рис. 15.91

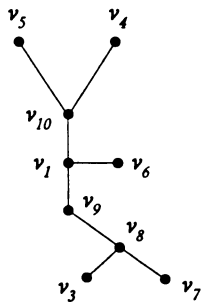


Рис. 15.92

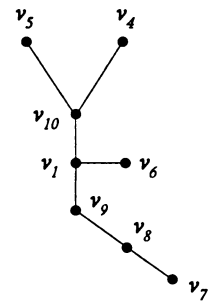


Рис. 15.93

Поскольку вершина v_4 стала листом с наименьшим индексом, удаляем ребро $\{v_4, v_{10}\}$ и полагаем $a_3 = 10$ (рис. 15.94). Теперь вершина v_5 стала листом с наименьшим индексом, удаляем ребро $\{v_5, v_{10}\}$ и полагаем $a_4 = 10$, как показано на рис. 15.95. Далее, поскольку вершина v_6 стала листом с наименьшим индексом, удаляем ребро $\{v_6, v_1\}$ и полагаем $a_5 = 1$ (рис. 15.96).

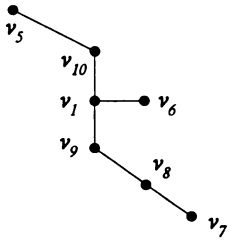


Рис. 15.94

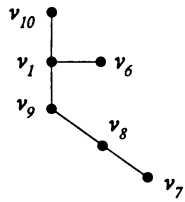


Рис. 15.95

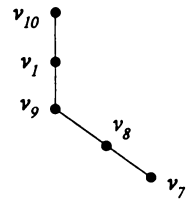


Рис. 15.96

Поскольку в оставшемся дереве вершина v_7 стала листом с наименьшим индексом, удаляем ребро $\{v_7, v_8\}$ и полагаем $a_6 = 8$, как показано на рис. 15.97. Теперь вершина v_8 стала листом с наименьшим индексом, поэтому удаляем ребро $\{v_8, v_9\}$ и полагаем $a_7 = 9$, как показано на рис. 15.98. Наконец, оставшаяся вершина v_9 стала листом с наименьшим индексом, и мы удаляем ребро $\{v_9, v_1\}$ и полагаем $a_8 = 1$. В результате осталось только ребро, изображенное на рис. 15.99.

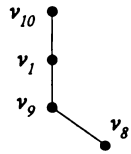


Рис. 15.97

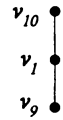


Рис. 15.98



Рис. 15.99

Таким образом, искомая последовательность имеет вид

$$9, 8, 10, 10, 1, 8, 9, 1.$$

Теперь нам нужен обратный алгоритм.

Алгоритм перевода последовательности в дерево — ПвД($a_1, a_2, \ldots, a_{n-2}$) для $n \geq 3$

(1) Для каждого v_i, если i появляется в последовательности n_i раз, положить $\deg(v_i) = n_i + 1$.

(2) Прочитать a_1 и сформировать ребро $\{v_{a_1}, v_j\}$, где j — наименьшая метка, такая что $\deg(v_j = 1)$.

(3) Уменьшить $\deg(v_{a_1})$ и $\deg(v_j)$ на 1.

(4) Предположим, что a_{i-1} прочитано, читать a_i и формировать ребро $\{v_{a_i}, v_j\}$, где j — наименьшая такая метка, что $\deg(v_j) = 1$.

(5) Уменьшить $\deg(v_{a_i})$ и $\deg(v_j)$ на 1.

(6) После того, как a_{n-2} прочитано, создать ребро между двумя оставшимися вершинами степени 1.

ПРИМЕР 15.31. Рассмотрим последовательность

$$1, 4, 1, 6, 6, 4$$

и воспользуемся ею для восстановления исходного дерева из примера 15.29.

Поскольку 1, 4 и 6 появляются в последовательности дважды, то

$$\deg(v_1) = \deg(v_4) = \deg(v_6) = 3.$$

А так как 2, 3 и 5 не встречаются вообще, то

$$\deg(v_2) = \deg(v_3) = \deg(v_5) = \deg(v_7) = \deg(v_8) = 1.$$

Запишем их с помощью восьмерки из чисел $(3,1,1,3,1,3,1,1)$, которую назовем восьмеркой степеней.

Теперь считываем $a_1 = 1$ и, поскольку v_2 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_1, v_2\}$ (рис. 15.100). Положим $\deg(v_1) = 2$ и $\deg(v_2) = 0$, поэтому восьмерка степеней теперь имеет вид $(2,0,1,3,1,3,1,1)$. Далее считываем $a_2 = 4$ и, потому что v_3 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_3, v_4\}$, после чего получаем граф, изображенный на рис. 15.101.

Полагаем $\deg(v_4) = 2$ и $\deg(v_3) = 0$, поэтому восьмерка степеней теперь имеет вид $(2,0,0,2,1,3,1,1)$. Второе ребро наклонено в другую сторону, потому что я смошенничал и посмотрел на дерево, которое пытаюсь воссоздать. Теперь считываем $a_3 = 1$ и, поскольку v_5 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_1, v_5\}$, после чего получаем граф, изображенный на рис. 15.102.

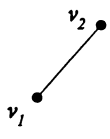


Рис. 15.100

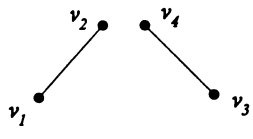


Рис. 15.101

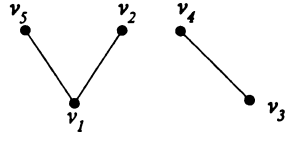


Рис. 15.102

Полагаем $\deg(v_1) = 1$ и $\deg(v_5) = 0$, поэтому восьмерка степеней теперь имеет вид $(1,0,0,2,0,3,1,1)$. Считываем $a_4 = 6$ и, поскольку v_1 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_1, v_6\}$, после чего получаем граф, изображенный на рис. 15.103.

Полагаем $\deg(v_1) = 0$ и $\deg(v_6) = 2$, поэтому восьмерка степеней теперь имеет вид $(0,0,0,2,0,2,1,1)$. Считываем $a_5 = 6$ и, потому что v_7 среди всех вершин степени 1 наименьшую степень, создаем ребро $\{v_7, v_6\}$, после чего получаем граф, изображенный на рис. 15.104.

Полагаем $\deg(v_7) = 0$ и $\deg(v_6) = 1$, поэтому восьмерка степеней теперь имеет вид $(0,0,0,2,0,1,0,1)$. Считываем $a_6 = 4$ и, поскольку v_6 среди всех вершин степени 1 наименьший индекс, создаем ребро $\{v_4, v_6\}$, после чего получаем граф изображенный на рис. 15.105.

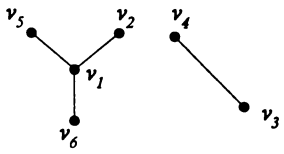


Рис. 15.103

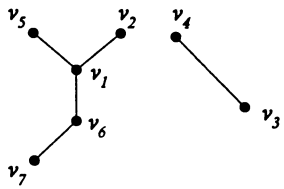


Рис. 15.104

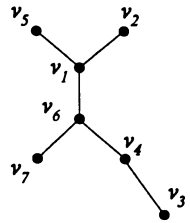


Рис. 15.105

Полагаем $\deg(v_4) = 1$ и $\deg(v_6) = 0$, поэтому восьмерка степеней теперь имеет вид $(0, 0, 0, 1, 0, 0, 0, 1)$. Наконец, поскольку все последовательности прочитаны и $\deg(v_4) = \deg(v_8) = 1$, формируем ребро $\{v_4, v_8\}$ и получаем искомое дерево из примера 15.29. □

ПРИМЕР 15.32. Теперь рассмотрим последовательность

$$9, 8, 10, 10, 1, 8, 9, 1$$

и воспользуемся ею для воссоздания исходного дерева из примера 15.30.

Поскольку 10, 9, 8 и 1 появляются в последовательности по два раза, то

$$\deg(v_{10}) = \deg(v_9) = \deg(v_8) = \deg(v_1) = 3.$$

А так как 2, 3 и 5 не встречаются в последовательности вообще, то

$$\deg(v_2) = \deg(v_3) = \deg(v_4) = \deg(v_5) = \deg(v_6) = \deg(v_7) = 1.$$

Запишем их с помощью десятки чисел $(3, 1, 1, 1, 1, 1, 1, 3, 3, 3)$, которую будем называть десяткой степеней.

Считываем $a_1 = 9$ и, поскольку v_2 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_9, v_2\}$, как показано на рис. 15.106.

Полагаем $\deg(v_9) = 2$ и $\deg(v_2) = 0$, поэтому десятка степеней имеет вид $(3, 0, 1, 1, 1, 1, 1, 3, 2, 3)$. Считываем $a_2 = 8$ и, потому что v_3 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_8, v_3\}$, как показано на рис. 15.107. Полагаем $\deg(v_8) = 2$ и $\deg(v_3) = 0$; десятка степеней теперь имеет вид $(3, 0, 0, 1, 1, 1, 1, 2, 2, 3)$. Считываем $a_3 = 10$ и, поскольку v_4 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_{10}, v_4\}$, как показано на рис. 15.108.

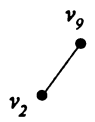


Рис. 15.106



Рис. 15.107

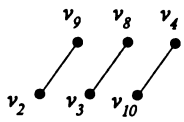


Рис. 15.108

Полагаем $\deg(v_{10}) = 2$ и $\deg(v_4) = 0$; десятка степеней теперь имеет вид $(3, 0, 0, 0, 1, 1, 1, 2, 2, 2)$. Считываем $a_4 = 10$ и, так как v_5 среди всех вершин

степени 1 имеет наименьший индекс, создаем ребро $\{v_{10}, v_5\}$, как показано на рис. 15.109. Полагаем $\deg(v_{10}) = 1$ и $\deg(v_5) = 0$, поэтому десятка степеней теперь имеет вид $(3,0,0,0,0,1,1,2,2,1)$. Считываем $a_5 = 1$ и поскольку v_6 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_1, v_6\}$ рис. 15.110.

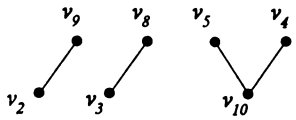


Рис. 15.109

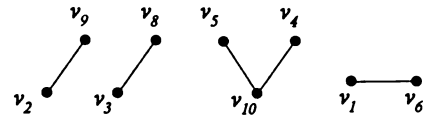


Рис. 15.110

Полагаем $\deg(v_1) = 2$ и $\deg(v_6) = 0$, поэтому десятка степеней имеет вид $(2,0,0,0,0,0,1,2,2,1)$. Считываем $a_6 = 8$ и, поскольку v_7 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_8, v_7\}$, как показано на рис. 15.111.

Полагаем $\deg(v_8) = 1$ и $\deg(v_7) = 0$; десятка степеней теперь имеет вид $(2,0,0,0,0,0,0,1,2,1)$. Считываем $a_7 = 9$ и, потому что v_8 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_8, v_9\}$, как показано на рис. 15.112.

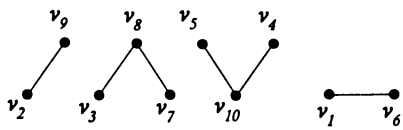


Рис. 15.111

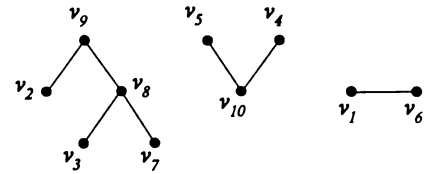


Рис. 15.112

Полагаем $\deg(v_9) = 1$ и $\deg(v_8) = 0$, поэтому десятка степеней теперь имеет вид $(2,0,0,0,0,0,0,0,1,1)$. Считываем $a_8 = 1$ и, поскольку v_9 среди всех вершин степени 1 имеет наименьший индекс, создаем ребро $\{v_1, v_9\}$, как показано на рис. 15.113.

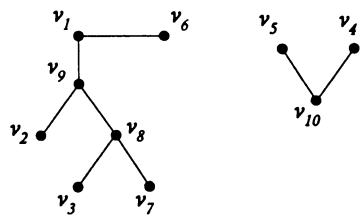


Рис. 15.113

Полагаем $\deg(v_1) = 1$ и $\deg(v_9) = 0$, поэтому десятка степеней теперь имеет вид $(1,0,0,0,0,0,0,0,0,1)$. Наконец, вся последовательность прочитана и $\deg(v_1) = \deg(v_{10}) = 1$; формируем ребро $\{v_1, v_{10}\}$ и получаем искомое дерево из упражнения 15.30. □

Предыдущая теорема дает количество различных остовных деревьев на n размеченных вершинах, но не рассматривает конкретный граф с этими вершинами. Когда рассматривается определенный граф, то количество различных остовных деревьев для этого графа уменьшается, так как используются только ребра графа. Данный раздел завершает метод определения количества остовных деревьев для графа с n размеченными вершинами.

ОПРЕДЕЛЕНИЕ 15.33. Пусть G — граф с n размеченными вершинами $v_1, v_2, v_3, \ldots, v_n$. ***Матрицей степеней*** графа G называется $n \times n$ матрица D, определенная следующим образом: $D_{ij} = 0$, если $i \neq j$, и D_{ii} равно степени вершины v_i.

Используя это определение, сформулируем следующую, весьма впечатляющую теорему, которую приведем без доказательства.

ТЕОРЕМА 15.34. (Матричная формула Кирхгофа) Пусть G — связный граф с помеченными вершинами. Пусть $K = D - A$, где A — матрица смежности графа G, а D — матрица степеней графа G. Число остовных деревьев графа G равно любому из алгебраических дополнений матрицы K.

ПРИМЕР 15.35. Найдем количество остовных деревьев графа, изображенного на рис. 15.114.

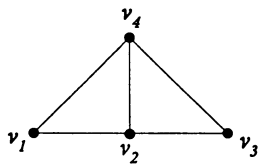


Рис. 15.114

Поскольку $\deg(v_1) = \deg(v_3) = 2$ и $\deg(v_2) = \deg(v_4) = 3$,

$$C = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Матрица A имеет вид

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix},$$

и

$$K = C - A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}.$$

Алгебраическое дополнение K_{11} есть

$$\det\begin{bmatrix} 3 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 3 \end{bmatrix} = 3\det\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} - (-1)\det\begin{bmatrix} -1 & -1 \\ -1 & 3 \end{bmatrix} +$$

$$+ (-1)\det\begin{bmatrix} -1 & 2 \\ -1 & -1 \end{bmatrix} = 3(5) - 4 - 3 = 8.$$

Значит, существует восемь остовных деревьев. □

■ УПРАЖНЕНИЯ

1. Задан граф (рис. 15.115), вершины которого упорядочены так, как помечены.
а) Найдите остовное дерево, удаляя ребро из каждого цикла.
б) Найдите остовное дерево поиском в ширину.
в) Найдите остовное дерево поиском в глубину.

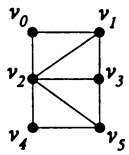


Рис. 15.115

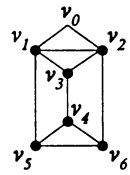


Рис. 15.116

2. Задан граф (рис. 15.116), вершины которого упорядочены так, как помечены.
а) Найдите остовное дерево, удаляя ребро из каждого цикла.
б) Найдите остовное дерево поиском в ширину.
в) Найдите остовное дерево поиском в глубину.

3. Задан граф (рис. 15.117), вершины которого упорядочены так, как помечены.
а) Найдите остовное дерево, удаляя ребро из каждого цикла.
б) Найдите остовное дерево поиском в ширину.
в) Найдите остовное дерево поиском в глубину.

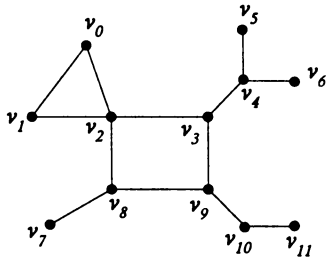


Рис. 15.117

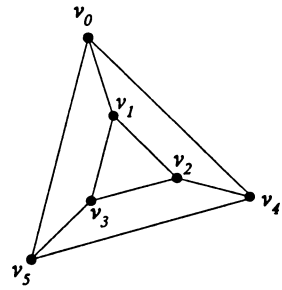


Рис. 15.118

4. Задан граф (рис. 15.118), вершины которого упорядочены так, как помечены.
   а) Найдите остовное дерево, удаляя ребро из каждого цикла.
   б) Найдите остовное дерево поиском в ширину.
   в) Найдите остовное дерево поиском в глубину.

5. Задан граф (рис. 15.119), вершины которого упорядочены так, как помечены.
   а) Найдите остовное дерево, удаляя ребро из каждого цикла.
   б) Найдите остовное дерево поиском в ширину.
   в) Найдите остовное дерево поиском в глубину.

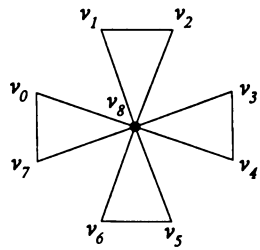


Рис. 15.119

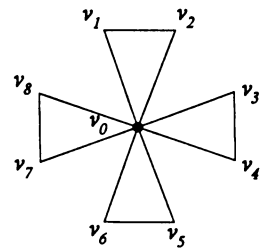


Рис. 15.120

6. Задан граф (рис. 15.120), вершины которого упорядочены так, как помечены.
   а) Найдите остовное дерево, удаляя ребро из каждого цикла.
   б) Найдите остовное дерево поиском в ширину.
   в) Найдите остовное дерево поиском в глубину.

7. Задан граф (рис. 15.121), вершины которого упорядочены так, как помечены.
   а) Найдите остовное дерево, удаляя ребро из каждого цикла.
   б) Найдите остовное дерево поиском в ширину.
   в) Найдите остовное дерево поиском в глубину.

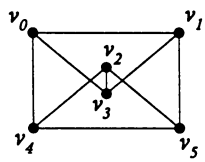


Рис. 15.121

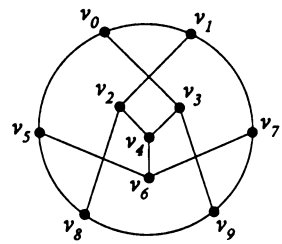


Рис. 15.122

8. Задан граф (рис. 15.122), вершины которого упорядочены так, как помечены.
   а) Найдите остовное дерево, удаляя ребро из каждого цикла.
   б) Найдите остовное дерево поиском в ширину.
   в) Найдите остовное дерево поиском в глубину.

9. Задан граф K_5.

а) Найдите остовное дерево, удаляя ребро из каждого цикла.

б) Найдите остовное дерево поиском в ширину.

в) Найдите остовное дерево поиском в глубину.

10. Задан граф $K_{3,4}$.

а) Найдите остовное дерево, удаляя ребро из каждого цикла.

б) Найдите остовное дерево поиском в ширину.

в) Найдите остовное дерево поиском в глубину.

11. Задан граф Петерсена.

а) Найдите остовное дерево, удаляя ребро из каждого цикла.

б) Найдите остовное дерево поиском в ширину.

в) Найдите остовное дерево поиском в глубину.

12. Используя алгоритм ПТС поиска точек сочленения, найдите точки сочленения и компоненты двусвязности для графа, изображенного на рис. 15.123.

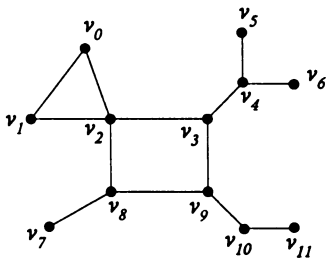


Рис. 15.123

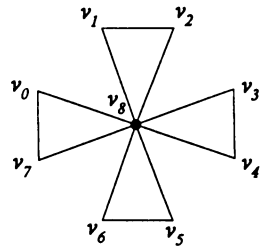
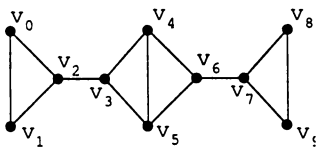


Рис. 15.124

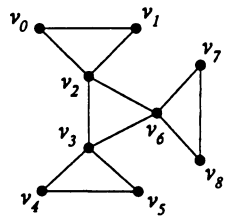
13. Используя алгоритм ПТС поиска точек сочленения, найдите точки сочленения и компоненты двусвязности для графа, изображенного на рис. 15.124.

14. Используя алгоритм ПТС поиска точек сочленения, найдите точки сочленения и компоненты двусвязности для приведенных ниже графов.

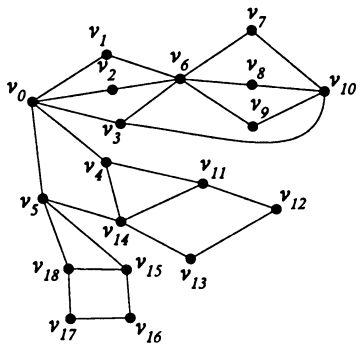
а)



б)



в)



15. Для графа G, представляющего собой простой цикл, опишите его остовные деревья, полученные поиском в ширину и в глубину. Если G имеет n вершин, то сколько у него остовных деревьев?

16. Если G — граф, представляющий собой цикл, но необязательно простой, изменится ли описание его остовных деревьев? Почему? Может ли граф G иметь точку сочленения?

17. Докажите, что ребро графа G принадлежит всем его остовным деревьям тогда и только тогда, когда это разрезающее ребро.

18. Найдите для графа G необходимые и достаточные условия равенства остовных деревьев, полученных поиском в глубину и в ширину?

19. Покажите, что остовное дерево, полученное поиском в глубину для K_n при $n \geq 3$, является простым путем, а остовное дерево, полученное поиском в ширину, таковым не является.

20. Найдите такое a, что если $n \geq a$, то K_n имеет два остовных дерева без общих ребер.

21. Докажите, что если E' — разрезающее множество графа G, то каждое остовное дерево содержит ребро из E'.

22. Сколько остовных деревьев имеет граф K_3? Сколько остовных деревьев имеет K_4? Сколько остовных деревьев имеет K_n?

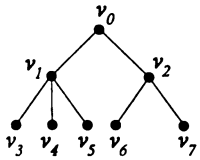
23. Сколько остовных деревьев имеет граф $K_{2,n}$?

24. Если G — граф с v вершин и e ребер, то сколько ребер можно удалить, не нарушив связности оставшегося графа?

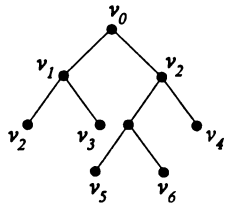
25. Пусть T и T' — различные остовные деревья графа G. Пусть e — ребро дерева T, которое не является ребром дерева T', и T_e — это дерево T, из которого удалено ребро e. Докажите, что в дереве T' имеется такое ребро e', добавив которое в T_e, получим граф, который будет опять остовным деревом графа G.

26. Используя алгоритм преобразования дерева в последовательность, определите соответствующие последовательности для каждого из приведенных ниже деревьев.

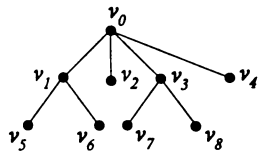
а)



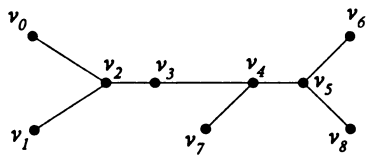
б)



в)

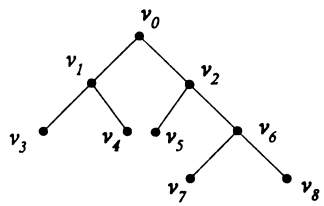


г)

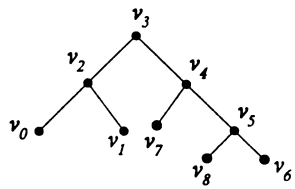


27. Используя алгоритм преобразования дерева в последовательность, определите последовательности, соответствующие каждому из следующих деревьев.

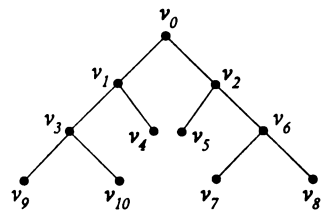
а)



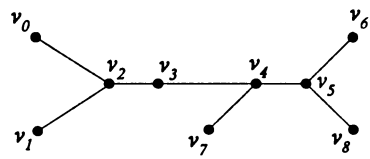
б)



в)



г)



28. Используя алгоритм перевода дерева в последовательность, постройте соответствующие деревья для приведенных ниже последовательностей.

а) 1, 2, 3, 3, 2, 3.

б) 1, 3, 3, 5, 5, 4, 4.

в) 1, 2, 2, 2, 5, 4, 5, 6.

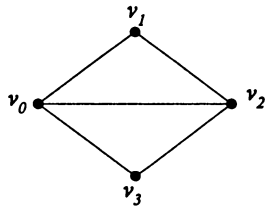
29. Используя алгоритм перевода дерева в последовательность, постройте соответствующие деревья для приведенных ниже последовательностей.

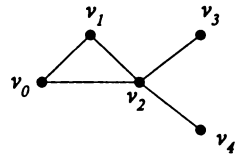
а) 1, 4, 2, 3, 2, 3.

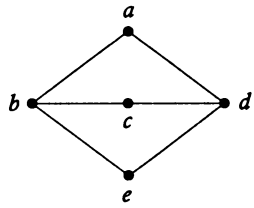
б) 1, 4, 3, 3, 4, 4, 4.

в) 1, 5, 2, 5, 3, 6, 5, 5.

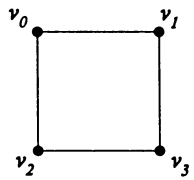
30. Используя матричную формулу Кирхгофа, определите число остовных деревьев для каждого из приведенных ниже графов.

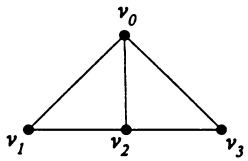
а) 

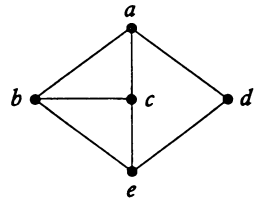
б) 

в) 

31. Используя матричную формулу Кирхгофа, определите число остовных деревьев для каждого из приведенных ниже графов.

а) 

б) 

в) 

32. Сколько остовных деревьев имеет граф $K_{n,n}$?

15.6. МИНИМАЛЬНЫЕ ОСТОВНЫЕ ДЕРЕВЬЯ

Взвешенный граф определен в предыдущей главе как граф, каждому ребру которого приписано положительное число, называемое весом. Этот вес может описывать расстояние, стоимость или другие данные. Например, если бы вершины представляли города, то вес мог бы быть расстоянием между городами или стоимостью перелета из одного города в другой. ***Вес*** остовного дерева взвешенного графа G равен сумме весов, приписанных ребрам остовного дерева. ***Минимальным остовным деревом*** называется такое остовное дерево графа G, что вес T меньше или равен весу любого другого остовного дерева графа G.

Мы рассмотрим два способа построения минимального остовного дерева взвешенного графа G. Первый называется алгоритмом Крускала. Идея метода состоит в том, чтобы формировать дерево $T(E, V)$, выбирая ребра с наименьшим весом так, чтобы не возникал цикл.

Алгоритм Крускала

(1) Выбрать в графе G ребро e минимального веса, не принадлежащее множеству E и такое, что его добавление в E не создает цикл в дереве T.

(2) Добавить это ребро во множество ребер E.

(3) Продолжать, пока имеются ребра, обладающие указанными свойствами.

ПРИМЕР 15.36. Рассмотрим взвешенный граф, изображенный на рис. 15.125.

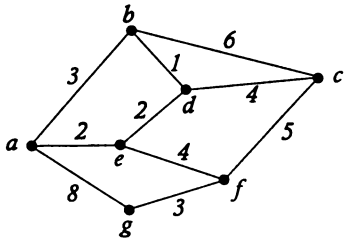


Рис. 15.125

Выберем ребро $\{b,d\}$ и добавим его в E. Далее — ребро $\{d,e\}$, так что $E = \{\{b,d\},\{d,e\}\}$. Затем добавим ребро $\{a,e\}$, после чего $E = \{\{b,d\},\{d,e\},\{a,e\}\}$. Ребро $\{a,b\}$ добавить нельзя, так как это приведет к образованию цикла, поэтому добавим ребро $\{f,g\}$, так что $E = \{\{b,d\},\{d,e\},\{a,e\},\{f,g\}\}$. Следующим можно добавить $\{e,f\}$, и $E = \{\{b,d\},\{d,e\},\{a,e\},\{f,g\},\{e,f\}\}$. Наконец, добавляем ребро $\{c,d\}$, после чего $E = \{\{b,d\},\{d,e\},\{a,e\},\{f,g\},\{e,f\}\},\{c,d\}$. Больше не осталось ребер, которые можно добавлять, поэтому искомое дерево выглядит, как изображенное на рис. 15.126.

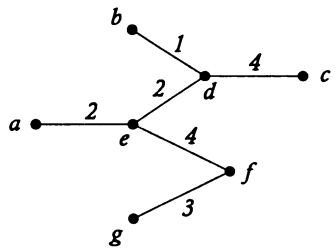


Рис. 15.126 □

ТЕОРЕМА 15.37. Алгоритм Крускала создает минимальное остовное дерево.

ДОКАЗАТЕЛЬСТВО. Алгоритм Крускала дает минимальное остовное дерево по построению. Пусть

$$e_1, e_2, e_3, e_4, \ldots, e_n$$

— ребра дерева T в том порядке, в котором их выбирает алгоритм Крускала. Пусть k — наибольшее целое число такое, что $e_1, e_2, e_3, e_4, \ldots, e_k$ — ребра минимального остовного дерева, и пусть T' является этим деревом. Если $k \neq n$, то присоединим ребро e_{k+1} к дереву T'. Теперь дерево T' с присоединенным ребром содержит цикл. Удалим из этого цикла e, где $e \neq e_i$, для $1 \leq i \leq k+1$, и обозначим полученное дерево T''. Вес ребра e не может быть меньше веса e_{k+1}, если бы это было так, то ребро уже было бы выбрано алгоритмом Крускала или создавало бы цикл, поэтому ребра $\{e_1, e_2, e_3, e_4, \ldots, e_k, e\}$ не могут составлять часть ребер дерева. Но это невозможно, поскольку данное множество ребер является частью дерева T'. Таким образом, поскольку дерево T'' создано из дерева T' заменой ребра e на ребро, которое не тяжелее, то вес дерева T'' меньше или равен весу дерева T'. Поэтому T'' — минимальное остовное дерево, которое как часть своего множества ребер содержит $\{e_1, e_2, e_3, e_4, \ldots, e_k, e_{k+1}\}$. Но это противоречит определению числа k. Поэтому $k = n$ и T — минимальное остовное дерево. ■

Второй и последний из представленных алгоритмов построения минимального остовного дерева называется алгоритмом Прима. Принципиальное отличие алгоритма состоит в том, что всегда имеется дерево, к которому ребра добавляют до тех пор, пока не получится остовное дерево. Используя алгоритм Прима, сначала выбираем вершину v_0 графа G, а затем выбираем ребро с наименьшим весом $e_1 = \{v_1, v_0\}$, инцидентное к вершине v_0, и формируем дерево T_1. Следующим выбираем ребро с наименьшим весом e_2 такое, что одна вершина ребра принадлежит дереву T_1, а вторая — нет, и добавляем его в дерево, после чего получаем дерево T_2 с двумя ребрами. Сформировав дерево T_k, формируем дерево T_{k+1}, выбирая ребро с минимальным весом e_{k+1} такое, что одна его вершина принадлежит дереву T_k, а другая — нет, и добавляем это ребро в дерево. Продолжаем процесс до тех пор, пока дерево не будет содержать все вершины графа G.

Теперь опишем алгоритм формально.

Алгоритм Прима нахождения минимального остовного дерева

(1) Выбрать вершину v_0 графа G и ребро с наименьшим весом e_1, для которого v_0 — одна из вершин, и сформировать дерево T_1.

(2) Для заданного дерева T_k с ребрами $e_1, e_2, e_3, \ldots, e_k$, если имеется вершина, не принадлежащая T_k, выбрать ребро с наименьшим весом, смежное с ребром дерева T_k и имеющее вершину вне дерева T_k. Добавить это ребро в дерево T_k, формируя дерево T_{k+1}.

(3) Продолжать, пока имеются вершины графа G, не принадлежащие дереву.

ПРИМЕР 15.38. Снова рассмотрим дерево (рис. 15.127).

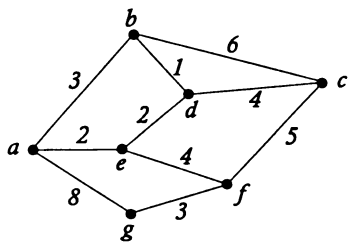


Рис. 15.127

Начав с вершины a, первым выбираем ребро $\{a, e\}$, поскольку оно имеет минимальный вес. Таким образом, дерево T_1 состоит из ребра $\{a, e\}$. Теперь выбираем ребро $\{e, d\}$, так как это ребро с минимальным весом из тех, что имеют только одну вершину в дереве T_1. Получаем дерево T_2, изображенное на рис. 15.128.

Далее выбираем ребро $\{d, b\}$, поскольку это ребро с минимальным весом из тех, что имеют только одну вершину в T_2. Получаем T_3 — дерево, изображенное на рис. 15.129. Теперь, так как имеется два ребра с одинаковым весом и одной вершиной в T_3, имеем выбор. Предположим, выбрано ребро $\{d, c\}$, поэтому дерево T_4 имеет следующий вид (рис. 15.130).



Рис. 15.128



Рис. 15.129

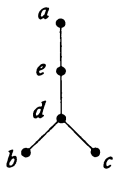


Рис. 15.130

Следующим выбираем ребро $\{e, f\}$. Это ребро имеет минимальный вес из тех, что имеют только одну вершину в дереве T_4; отсюда T_5 — дерево на рисунке 15.131.

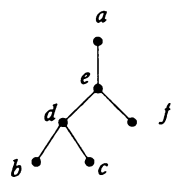


Рис. 15.131

Наконец, выбираем ребро $\{f, g\}$. Поскольку это ребро минимального веса из тех, что имеют одну вершину в T_4, а вершина g — последняя из оставшихся вершин, имеем дерево T_6, которое является искомым остовным деревом. □

Теперь докажем, что данный алгоритм в действительности создает минимальное остовное дерево.

ТЕОРЕМА 15.39. Для заданного взвешенного графа G дерево, создаваемое в соответствии с алгоритмом Прима нахождения минимального остовного дерева, является минимальным остовным деревом.

ДОКАЗАТЕЛЬСТВО. Из построения дерева следует, что оно остовное. Доказательство того, что оно является минимальным остовным деревом, аналогично доказательству подобного утверждения для алгоритма Крускала. Пусть

$$e_1, e_2, e_3, e_4, \ldots, e_n$$

— ребра дерева T в том порядке, в каком они выбраны согласно алгоритму Прима нахождения минимального остовного дерева, и k — наибольшее целое число такое, что $e_1, e_2, e_3, e_4, \ldots, e_k$ являются ребрами минимального остовного дерева, и T' является этим деревом. Пусть T_k — дерево, сформированное ребрами $e_1, e_2, e_3, e_4, \ldots, e_k$. Если $k \neq n$, присоединяем ребро e_{k+1} к дереву T'. Теперь T' с новым присоединеннным ребром содержит цикл. В цикле содержится другое ребро e, имеющее только одну вершину в дереве T_k. Удалим ребро e из дерева $T' \cup \{e_{k+1}\}$, формируя дерево T'' с ребром e_{k+1} вместо ребра e. Поскольку e_{k+1} — ребро с наименьшим весом среди тех, что имеют одну вершину в дереве T_k, вес e_{k+1} меньше или равен весу e. Поэтому вес T'' меньше или равен весу T', так что T'' также является минимальным остовным деревом с ребрами $e_1, e_2, e_3, e_4, \ldots, e_k, e_{k+1}$, но это противоречит определению k. Поэтому $k = n$, и алгоритм Прима нахождения минимального остовного дерева создает минимальное остовное дерево. ■

Алгоритм Прима можно также реализовать с помощью матриц. Для графа G с n вершинами требуется матрица размерности $(n+1)\times(n+1)$. Сначала определим $n \times n$ весовую матрицу W, где W_{ij} — вес ребра из v_i в v_j, если ребро из v_i в v_j существует, и 0, если такого ребра нет. После этого добавим $n+1$-ую строку и $n+1$-ый столбец, которые пока будут пустыми. Новую матрицу обозначим через W^*.

Формально изложим следующий алгоритм.

Матричный алгоритм Прима

(1) Создать весовую матрицу W.

(2) Добавить дополнительную строку и столбец, чтобы создать матрицу W^*.

(3) В строке 1 матрицы W^* поместить $*$ в последнем столбце. В столбце 1 заменить все числа на 0 и поместить U в последней строке.

(4) Выбрать наименьшее число, так что строка с этим числом имеет $*$ в столбце $n+1$, а столбец с этим числом не содержит U в строке $n+1$.

(5) Если число выбрано в строке i и столбце j, то поместить $*$ в последний столбец строки j, заменить оставшиеся числа в столбце j на 0, поместить U в строке $n+1$ столбца j и добавить ребро (v_i, v_j) в остовное дерево.

(6) Продолжать выполнение шагов 4 и 5, пока не останется чисел, которые можно выбирать.

ПРИМЕР 15.40. Предположим, что имеется граф, изображенный на рис. 15.132

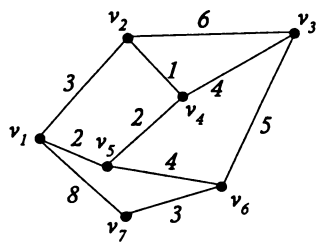


Рис. 15.132

поэтому матрицы W и W^* имеют, соответственно, вид

$$\begin{bmatrix} 0 & 3 & 0 & 0 & 2 & 0 & 8 \\ 3 & 0 & 6 & 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 4 & 0 & 5 & 0 \\ 0 & 1 & 4 & 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 & 4 & 0 \\ 0 & 0 & 5 & 0 & 4 & 0 & 3 \\ 8 & 0 & 0 & 0 & 0 & 3 & 0 \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} 0 & 3 & 0 & 0 & 2 & 0 & 8 \\ 3 & 0 & 6 & 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 4 & 0 & 5 & 0 \\ 0 & 1 & 4 & 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 & 4 & 0 \\ 0 & 0 & 5 & 0 & 4 & 0 & 3 \\ 8 & 0 & 0 & 0 & 0 & 3 & 0 \\ \end{bmatrix}.$$

Сначала выбираем строку 1, что означает выбор первой вершины v_1равной 0, так как не собираемся больше возвращаться к вершине v_1, и помещаем U в восьмой ($(n+1)$-ой) строке столбца 1, чтобы показать, что столбец 1 не будет больше использован. Помещаем $*$ в строке 1, тем самым указывая, что можно

выбирать значения из этой строки, поскольку она содержит значения ребер из v_1 в соседние вершины. Отсюда имеем

$$\begin{bmatrix} 0 & 3 & 0 & 0 & 2 & 0 & 8 & * \\ 0 & 0 & 6 & 1 & 0 & 0 & 0 & \\ 0 & 6 & 0 & 4 & 0 & 5 & 0 & \\ 0 & 1 & 4 & 0 & 2 & 0 & 0 & \\ 0 & 0 & 0 & 2 & 0 & 4 & 0 & \\ 0 & 0 & 5 & 0 & 4 & 0 & 3 & \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & \\ U & & & & & & & \end{bmatrix}.$$

Выбираем наименьшее ненулевое значение в этой строке (т.е. выбираем ребро с наименьшим весом, смежное с v_1). В данном случае это значение равно 2 и находится в столбце 5. Таким образом, в качестве первого ребра выбрано ребро (v_1, v_5). Все другие значения в столбце 5 полагаем равными 0 и в восьмой строке этого столбца устанавливаем значение, равное U, поэтому данную строку уже не будем использовать и не будем больше возвращаться к вершине v_5. Помещаем $*$ в восьмом столбце строки 5, показывая, что элементы в строке 5 можно теперь использовать так же, как элементы первой строки, поскольку эти строки представляют ребра из v_1 и v_5 в смежные вершины. Имеем

$$\begin{bmatrix} 0 & 3 & 0 & 0 & \underline{2} & 0 & 8 & * \\ 0 & 0 & 6 & 1 & 0 & 0 & 0 & \\ 0 & 6 & 0 & 4 & 0 & 5 & 0 & \\ 0 & 1 & 4 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 2 & 0 & 4 & 0 & * \\ 0 & 0 & 5 & 0 & 0 & 0 & 3 & \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & \\ U & & & & U & & & \end{bmatrix}.$$

Далее выбираем наименьшее ненулевое значение в строке 1 или в строке 5. В данном случае это значение 2 в столбце 4 строки 5. Таким образом, в качестве следующего ребра выбрано (v_5, v_4). Остальные значения в данном столбце полагаем равными 0 и в восьмой строке данного столбца устанавливаем значение, равное U, поэтому строку не будем больше использовать и не будем больше возвращаться к вершине v_4. Помещаем $*$ в восьмом столбце строки 4, показывая, что элементы в этой строке можно теперь использовать так же, как и элементы строки 1 и строки 5, поскольку эти строки представляют ребра из v_1, v_4, и v_5 в смежные вершины. Имеем

$$\begin{bmatrix} 0 & 3 & 0 & 0 & 2 & 0 & 8 & * \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & \\ 0 & 6 & 0 & 0 & 0 & 5 & 0 & \\ 0 & 1 & 4 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & \underline{2} & 0 & 4 & 0 & * \\ 0 & 0 & 5 & 0 & 0 & 0 & 3 & \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & \\ U & & & U & U & & & \end{bmatrix}.$$

Теперь выбираем наименьшее ненулевое значение из 1-ой, 4-ой или 5-ой строки. В данном случае это значение, равное 1 в столбце 2 строки 4. Таким образом, в качесте следующего ребра выбрано (v_4,v_2). Остальные значения в столбце полагаем равными 0 и в восьмой строке этого столбца устанавливаем значение, равное U. Помещаем $*$ в восьмом столбце строки 2, показывая, что эти элементы можно теперь использовать, как и элементы строк 1, 4 и 5, поскольку эти строки представляют ребра из v_1,v_2,v_4, и v_5 в смежные вершины. Имеем

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 0 & 8 & * \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & \\ 0 & \underline{1} & 4 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 2 & 0 & 4 & 0 & * \\ 0 & 0 & 5 & 0 & 0 & 0 & 3 & \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & \\ U & U & & U & U & & & \end{bmatrix}.$$

Используя ту же процедуру, выбираем значение 4 в строке 4 столбца 3 и значение 4 в строке 5 столбца 6, что дает ребра (v_4, v_3) и (v_5, v_6), соответственно, и матрицы

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 0 & 8 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & * \\ 0 & 1 & \underline{4} & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 2 & 0 & 4 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & \\ U & U & U & U & U & & & \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 0 & 8 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 1 & 4 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 2 & 0 & \underline{4} & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ U & U & U & U & U & U & & \end{bmatrix}.$$

Наконец, выбираем значение 3 в строке 6 столбца 7, что дает ребро (v_6, v_7) и матрицу

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 2 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 1 & 4 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 2 & 0 & 4 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{3} & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ U & U & U & U & U & U & & \end{bmatrix},$$

а также остовное дерево, изображенное на рис. 15.133.

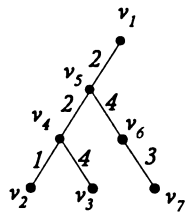


Рис. 15.133 □

■ УПРАЖНЕНИЯ

1. Заданы взвешенные графы, изображенные на рис. 15.134 и 15.135, найдите минимальные остовные деревья, воспользовавшись:

а) алгоритмом Крускала;

б) алгоритмом Прима.

в) матричным алгоритмом Прима;

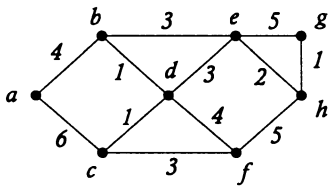


Рис. 15.134

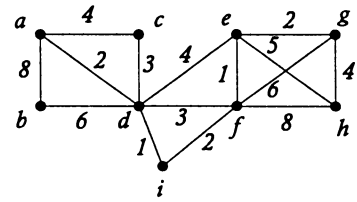


Рис. 15.135

2. Заданы взвешенные графы, изображенные на рис. 15.136 и 15.137, найдите минимальные остовные деревья, воспользовавшись;

а) алгоритмом Крускала;

б) алгоритмом Прима;

в) матричным алгоритмом Прима.

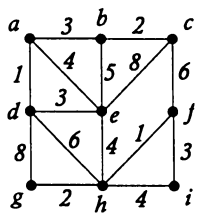


Рис. 15.136

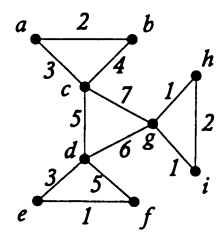


Рис. 15.137

3. Заданы взвешенные графы, изображенные на рис. 15.138 и 15.139, найдите минимальные остовные деревья, воспользовавшись:

а) алгоритмом Крускала;

б) алгоритмом Прима;

в) матричным алгоритмом Прима.

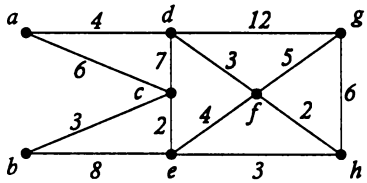


Рис. 15.138

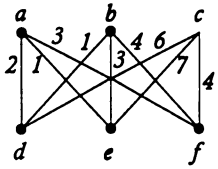


Рис. 15.139

4. Заданы взвешенные графы, изображенные на рис. 15.140 и 15.141, найдите минимальные остовные деревья, воспользовавшись:

а) алгоритмом Крускала;

б) алгоритмом Прима;

в) алгоритмом Прима с матрицами.

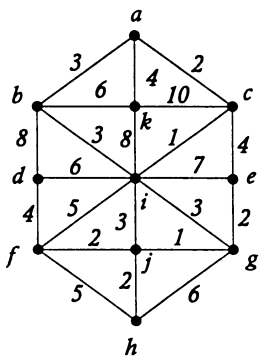


Рис. 15.140

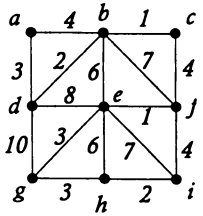


Рис. 15.141

ГЛАВА

16

СЕТИ

16.1. СЕТИ И ПОТОКИ

Сеть можно представить как систему, транспортирующую некий продукт из одной точки в другую. Этим продуктом могут быть люди, электроэнергия, природный газ, нефть и многое другое. Примером может служить система нефтепровода, где нефть течет из одной точки к другим точкам системы. Наша терминология будет соответствовать данной концепции. Используя такое представление, рассмотрим сеть как ориентированный граф, ребра которого — трубы между точками системы, а они, в свою очередь, представлены вершинами графа. Каждому ребру $e = (v_i, v_j)$ соответствует положительное целое число $c(e)$, называемое пропускной способностью e. Если между двумя вершинами не существует ребра, то пропускную способность полагаем равной нулю. В примере с нефтяными сетями пропускная способность связана с количеством нефти, проходящей через трубу (ребро).

Перед тем как определить сеть, ограничим определение ориентированного графа. Наличие петель у графа недопустимо, поскольку рассматриваемые задачи связаны только с транспортировкой продукта между различными точками. Будем считать, что если существует ребро из v_i в v_j, то нет ребра из v_j в v_i. Таким образом, рассматривается поток продукта только в одну строну. Необходимое требование: ориентированный граф должен быть связным, т.к. если имеется путь из a в z, нас будет интересовать только компонента, содержащая a и z. Если между a и z не существует пути, то и определять нечего. Ориентированный граф, удовлетворяющий этим условиям, часто называется простым связным ориентированным графом. Назовем такой граф ориентированным графом, имея в виду, что он удовлетворяет упомянутым выше ограничениям. Рассмотрим также особую вершину a, называемую источником, и особую вершину z, называвемую стоком. Степень входа вершины a равна 0, так что в источник ничто не втекает. Степень выхода вершины z равна 0, так что из стока ничто не вытекает. Таким образом, продукт перевозится из a и имеет местом назначения z. Более точно определим сеть следующим образом.

ОПРЕДЕЛЕНИЕ 16.1. ***Сеть*** — это ориентированный граф (G, V, E) вместе с весовой функцией $C : E \to N$ и выделенными вершинами a, z, такими что

(i) $\text{indeg}(a) = 0$;

(ii) $\text{outdeg}(z) = 0$.

Например, граф, изображенный на рис. 16.1, является примером сети, где числа на каждом ребре обозначают пропускную способность.

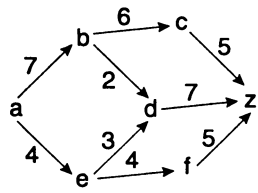


Рис. 16.1

Для этой сети, которая представлена как нефтепровод, введем понятие потока (количество нефти, проходящее через трубы нефтепровода). Таким образом, для каждого ребра e имеется значение $f(e)$, которое является потоком через конкретное ребро или трубу. Очевидно, величина потока не может превысить пропускную способность трубы. Потребуем также, чтобы поток, входящий в вершину, был равен потоку, выходящему из вершины, за исключением вершин a и z.

Это называется ***сохранением потока***. Пусть $\text{in}(v)$ — множество ребер, для которых v — конечная вершина, и $\text{out}(v)$ — множество ребер, для которых v — начальная вершина. Таким образом, $\text{out}(v)$ — множество ребер, выходящих из вершины v, и $\text{in}(v)$ — множество ребер, входящих в вершину v. Следовательно, имеем следующее определение.

ОПРЕДЕЛЕНИЕ 16.2. ***Поток*** в сети — это функция $f : E \to N \cup \{0\}$ такая, что

(i) для всех $e \in E$, $0 \le f(e) \le c(e)$;

(ii) для всех $v \in V$ таких, что $v \ne a, z$, $\sum_{e \in \text{in}(v)} f(e) = \sum_{e \in \text{out}(v)} f(e)$.

Допустим, имеется фиксированный поток. Пусть $\textit{поток}(a) = \sum_{e \in \text{out}(a)} f(e)$, т.е. $\textit{поток}(a)$ — поток, вытекающий из источника a, и $\textit{поток}(z) = \sum_{e \in \text{in}(z)} f(e)$, т.е. $\textit{поток}(z)$ — поток, втекающий в сток z. Пример потока в сети, где первый элемент упорядоченной пары на каждом ребре — пропускная способность, а второй — поток, демонстрирует рис. 16.2.

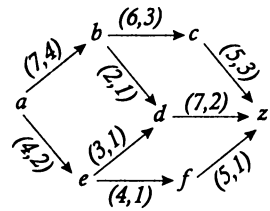


Рис. 16.2

Поток через каждое ребро меньше, чем его пропускная способность. Обратите внимание, что для вершины b поток, втекающий в b, равен 4 и поток, вытекающий из b, равен 4. Таким образом, оба потока совпадают, и в вершине b имеет место сохранение потока. Это верно также и для всех других вершин, за исключением вершин a и z. Воспользуемся принципом сохранения потока для доказательства того, что интуитивно кажется очевидным. А именно, поток из a равен потоку в z, т.е. $поток(a) = поток(z)$, что и утверждает приведенная ниже теорема. Рассмотрим сначала конкретную сеть.

Заметим, что для потока на рис. 16.2 выполняется равенство $поток(a) = поток(z) = 6$. Пусть S — множество вершин $\{b, c, d\}$, а $T = V - S$ — множество вершин $\{a, f, z\}$. Если просуммировать потоки в S, то согласно приведенной ниже таблице получим, что входящий поток равен 9 и выходящий поток равен 9. Таким образом, разность потоков равна 0.

вершина	входящий поток	ребро	выходящий поток	ребро
b	4	(a, b)	3	(b, c)
			1	(b, d)
c	3	(b, c)	3	(c, z)
d	1	(e, d)	2	(d, z)
	1	(b, d)		
сумма	9		9	

Если во множество S добавить вершину a, получим, что разность 15 (выходящий поток) и 9 (входящий поток) равна $6 = поток(a)$.

вершина	входящий поток	ребро	выходящий поток	ребро
b	4	(a, b)	3	(b, c)
			1	(b, d)
c	3	(b, c)	3	(c, z)
d	1	(e, d)	2	(d, z)
	1	(b, d)		
a	0		4	(a, b)
			2	(a, e)
сумма	9		15	

Заметим, что ребра (a, b), (b, c) и (b, d) присутствуют в колонке для входящего и выходящего потока, потому что обе вершины каждого ребра принадлежат S. При вычитании эти ребра сокращаются, поэтому, удалив их из обеих частей, получаем

вершина	входящий поток	ребро	выходящий поток	ребро
b				
c			3	(c,z)
d	1	(e,d)	2	(d,z)
a	0			
			2	(a,e)
сумма	1		7	

Снова разница входящего и выходящего потока равна 6. Но поскольку ребра, обе вершины которых принадлежат S, удалены, то выходящий поток — это сумма потоков ребер, которые идут из S в T, а входящий поток — это сумма потоков ребер, которые идут из T в S. Аналогичные рассуждения будут использованы в приведенной ниже теореме.

ТЕОРЕМА 16.3. Для любого фиксированного потока f

$$поток(a) = \sum_{e \in \text{out}(a)} f(e) = \sum_{e \in \text{in}(z)} f(e) = поток(z).$$

ДОКАЗАТЕЛЬСТВО. Пусть S — подмножество множества V, содержащее a, но не содержащее z, а $T = V - S$. Для $v \neq a, z$ из условия сохранения потока имеем $\sum_{e \in \text{in}(v)} f(e) = \sum_{e \in \text{out}(v)} f(e)$. Следовательно, суммирование по вершинам, принадлежащим $S - \{a\}$, дает

$$\sum_{v \in S-\{a\}} \sum_{e \in \text{in}(v)} f(e) = \sum_{v \in S-\{a\}} \sum_{e \in \text{out}(v)} f(e) \qquad \square$$

или

$$\sum_{v \in S-\{a\}} \sum_{e \in \text{out}(v)} f(e) - \sum_{v \in S-\{a\}} \sum_{e \in \text{in}(v)} f(e) = 0.$$

Поэтому, если включить a, получим

$$\sum_{v \in S} \sum_{e \in \text{out}(v)} f(e) - \sum_{v \in S} \sum_{e \in \text{in}(v)} f(e) = \sum_{e \in \text{out}(a)} f(e).$$

Пусть (S,T) — множество всех ребер из S в T, т.е. $(S,T) = \{e$: начальная вершина e принадлежит S, а конечная вершина e принадлежит $T\}$. Аналогично, пусть (T,S) — множество всех ребер из T в S. Рассмотрим опять

$$\sum_{v \in S} \sum_{e \in \text{out}(v)} f(e) - \sum_{v \in S} \sum_{e \in \text{in}(v)} f(e).$$

В случае, когда начальная и конечная вершины ребра e принадлежат S, e появляется в обеих суммах и, следовательно, взаимно сокращается. Если исключить такие ребра, то слева получим только ребра, идущие из S в T, а справа — только ребра, идущие из T в S. Таким образом, имеем

$$\sum_{v \in S} \sum_{e \in \text{out}(v)} f(e) - \sum_{v \in S} \sum_{e \in \text{in}(v)} f(e) = \sum_{e \in (S,T)} f(e) - \sum_{e \in (T,S)} f(e).$$

Следовательно,

$$\sum_{e \in (S,T)} f(e) - \sum_{e \in (T,S)} f(e) = \sum_{e \in \text{out}(a)} f(e),$$

т.е. для любого заданного множества вершин S, содержащего a, но не содержащего z, разность выходящего из S и входящего в S потоков равна $\textit{поток}(a)$, т.е. потоку, выходящему из источника a. Теперь пусть $S = V - \{z\}$, так что $T = \{z\}$. Тогда $\sum_{e \in (S,T)} f(e)$ есть просто поток, вливающийся в z, так что

$$\sum_{e \in (S,T)} f(e) = \sum_{e \in \text{in}(z)} f(e) = \textit{поток}(z)$$

и $\sum_{e \in (T,S)} f(e) = 0$, поскольку это поток, вытекающий из z. Следовательно,

$$\sum_{e \in (S,T)} f(e) - \sum_{e \in (T,S)} f(e) = \sum_{e \in \text{in}(z)} f(e) = \textit{поток}(z)$$

и $\textit{поток}(a) = \textit{поток}(z)$. ■

При доказательстве теоремы попутно доказано такое следствие.

СЛЕДСТВИЕ 16.4. Пусть S — подмножество множества V, содержащее a, но не содержащее z, и $T = V - S$. Тогда

$$\sum_{e \in (S,T)} f(e) - \sum_{e \in (T,S)} f(e) = \textit{поток}(a) = \textit{поток}(z).$$

ОПРЕДЕЛЕНИЕ 16.5. ***Величина потока*** f, обозначаемая как $val(f)$, равна $\textit{поток}(a) = \textit{поток}(z)$.

ОПРЕДЕЛЕНИЕ 16.6. Пусть S — подмножество множества V и $T = V - S$. Тогда $\{e : e \in (S,T)\}$ называется ***сечением***. Если $a \in S$ и $z \in T$, то сечение называется ***$a - z$ сечением***.

ОПРЕДЕЛЕНИЕ 16.7. Величина $C(S,T) = \sum_{e \in (S,T)} c(e)$ называется ***пропускной способностью сечения***.

ОПРЕДЕЛЕНИЕ 16.8. Поток $\hat{f}$ в сети называется ***максимальным потоком***, если $val(\hat{f}) \geq val(f)$ для любого возможного потока f в сети.

ОПРЕДЕЛЕНИЕ 16.9. $a - z$ сечение (S,T) называется ***минимальным сечением***, если $C(S,T)$ не больше пропускной способности любого другого $a - z$ сечения.

ТЕОРЕМА 16.10. Пусть S — подмножество множества V, содержащее a, но не содержащее z, и $T = V - S$. Тогда

$$val(f) \leq C(S,T).$$

ДОКАЗАТЕЛЬСТВО.

$$val(f) = \sum_{e \in (S,T)} f(e) - \sum_{e \in (T,S)} f(e) \leq \sum_{e \in (S,T)} f(e) \leq \sum_{e \in (S,T)} c(e) = C(S,T). \blacksquare$$

СЛЕДСТВИЕ 16.11. Если $val(f) = C(S,T)$ для некоторого потока f, а $a - z$ сечения (S,T), то f — максимальный поток, а C — минимальное сечение.

СЛЕДСТВИЕ 16.12. Для некоторого потока f и $a - z$ сечения (S,T) равенство $val(f) = C(S,T)$ имеет место тогда и только тогда, когда $f(e) = c(e)$ для всех $e \in (S,T)$ и $f(e) = 0$ для всех $e \in (T,S)$.

Опишем способы определения максимального потока сети. Рассмотрим сеть, изображенную на на рис. 16.3.

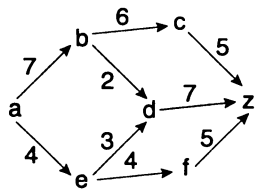


Рис. 16.3

Максимальный поток (не обязательно единственный) легко найти способом, проиллюстрированным на рис. 16.4.

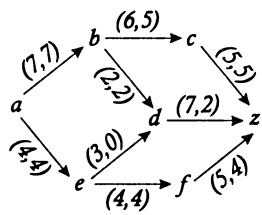


Рис. 16.4

В данном случае известно, что поток максимальный, потому что величина потока, выходящего из a, не может превысить сумму пропускных способностей ребер, выходящих из a.

Рассмотрим сеть, изображенную на рис. 16.5.

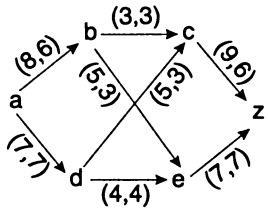


Рис. 16.5

Кажется, что у этой сети максимальный поток, потому что не существует ориентированного пути, по которому можно увеличить поток. Заметим, однако, что сеть на рис. 16.6 имеет больший поток, и, фактически, он максимальный.

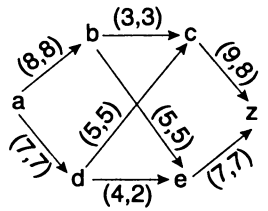


Рис. 16.6

Следует обратить внимание, что если поток из источника равен сумме пропускных способностей ребер, выходящих из источника, или поток, втекающий в сток, равен сумме пропускных способностей ребер, входящих в сток, то поток максимальный. Однако, поток может быть максимальным и без удовлетворения какого-либо из этих признаков. Например, сеть на рис. 16.7 имеет максимальный поток.

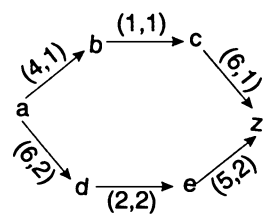


Рис. 16.7

Итак, как же находить максимальные в смысле потока сети? Для этого сформируем пути из a в z, не обращая внимания на направление ребер. Такие пути назовем ***цепями***. Рассмотрим снова сеть с потоком, изображенную на рис. 16.5. Одним из таких путей является

$$a \overset{(8,6)}{\rightarrow} b \overset{(3,3)}{\rightarrow} c \overset{(9,6)}{\rightarrow} z.$$

Очевидно, что нет возможности увеличить поток по этому пути, поскольку ребро из b в c наполнено до его пропускной способности. То же самое верно и для цепи

$$a \overset{(8,6)}{\rightarrow} b \overset{(5,3)}{\rightarrow} e \overset{(7,7)}{\rightarrow} z.$$

Однако, при рассмотрении цепи

$$a \overset{(8,6)}{\rightarrow} b \overset{(5,3)}{\rightarrow} e \overset{(4,4)}{\leftarrow} d \overset{(5,3)}{\rightarrow} c \overset{(9,6)}{\rightarrow} z$$

видим: если увеличить поток на 2 для стрелок, имеющих то же самое направление, что и цепь, и уменьшить поток на 2 для стрелок, имеющих противоположное направление, то получим

$$a \overset{(8,8)}{\rightarrow} b \overset{(5,5)}{\rightarrow} e \overset{(4,2)}{\leftarrow} d \overset{(5,5)}{\rightarrow} c \overset{(9,8)}{\rightarrow} z,$$

что увеличивает поток на 2. Первый вопрос, который следовало бы задать: "Почему выбираем 2?". Вполне очевидно, что поток желательно увеличить как можно больше. Но он не может превысить пропускную способность ни одного из заданных ребер. Именно по этой причине мы ограничены величиной 2. Кроме того, если ребро имеет направление, противоположное цепи, то нельзя уменьшить поток через это ребро более, чем на текущую величину потока через него, иначе получим отрицательный поток. Поэтому, если бы не было других ограничений, то

$$e \overset{(4,4)}{\leftarrow} d$$

ограничило бы изменение потока до 4. Второй вопрос: "Как это влияет на сохранение потока?". Убедимся, что условие сохранения потока выполняется. Рассмотрим, например, изменение

$$a \overset{(8,6)}{\rightarrow} b \overset{(5,3)}{\rightarrow} e$$

на

$$a \overset{(8,8)}{\rightarrow} b \overset{(5,5)}{\rightarrow} e.$$

Поток, выходящий из вершины b, увеличен на ту же самую величину, что и поток, входящий в вершину b. Поэтому чистый поток через b остается прежним. При изменении

$$b \overset{(5,3)}{\rightarrow} e \overset{(4,4)}{\leftarrow} d$$

на

$$b \overset{(5,5)}{\rightarrow} e \overset{(4,2)}{\leftarrow} d$$

поток из вершины b в вершину e увеличен на ту же величину, на которую уменьшен поток из вершины d в вершину e, поэтому чистый поток в вершине e остался тем же. Наконец, рассмотрим изменение

$$e \overset{(4,4)}{\leftarrow} d \overset{(5,3)}{\rightarrow} c$$

на

$$e \overset{(4,2)}{\leftarrow} d \overset{(5,5)}{\rightarrow} c$$

Поток из d в e уменьшен на ту же величину, на которую увеличен поток из d в c. Поэтому чистый поток и вершины d остался неизменным.

Процесс увеличения потока в сети чрезвычайно прост. Формируем цепь из a в z. Если возможно, то увеличиваем поток, определяя наибольшую величину, которую можно добавить к каждому из ребер, имеющих то же самое направление, что и цепь, чтобы поток не превысил пропускную способность, и которую можно вычесть из каждого ребра, имеющего противоположное направление, не

создавая отрицательный поток. Поскольку последнее ребро, входящее в z, имеет то же самое направление, что и цепь, то поток в z возрастает. Аналогично, ребро, выходящее из a, имеет то же самое направление, что и цепь, поэтому, как и ожидалось, поток из a возрастает. Так и должно было быть, потому что $поток(a) = поток(z)$. Потому что пропускная способность — конечная величина, а общая величина потока увеличивается, неизбежно наступает момент, когда наращивать поток далее нельзя. Когда это происходит, получаем максимальный поток. До сих пор рассматривался способ увеличения уже существующего потока. А с какого потока начинать? Ответ прост. Следует начинать с нулевого потока, т.е. когда поток через каждое ребро равен 0, а затем наращивать его. Теперь перейдем к изложению систематического алгоритма для нахождения максимального потока. Каждой вершине поставим в соответствие упорядоченную пару. Первый элемент пары — предшествующая вершина в рассматриваемой цепи предназначена для того, чтобы можно было найти обратный путь. Второй элемент пары — резерв, т.е. величина, на которую можно увеличить поток на каждом ребре вдоль пути, если ориентация ребра совпадает с направлением цепи (такую ориентацию назовем правильной), или уменьшить поток, если ребро ориентировано неправильно. Проще говоря, резерв заданной вершины равен максимальной величине, на которую поток вдоль цепи до этой вершины можно увеличить без нарушения самого потока. Рассмотрим также множество S, содержащее все вершины, не задействованные при построении цепи в z. Если S окажется пустым, до того как будет достигнута вершина z, то больше не будет вершин, которые можно испытать на пути к z, поэтому нет иного пути в z и нет более возможности увеличить поток, значит поток максимален.

Алгоритм (Форда-Фалкерсона) нахождения максимального потока

(1) Установить предшественника каждой вершины и резерв равными символу $-$ (непомечено). Вершина помечена, когда ее резерв не обозначен символом $-$. Установить резерв вершины a равным ∞, с тем чтобы она не ограничивала резерв других вершин. Положить $S = \{a\}$.

(2) Если S — пустое множество, то поток максимизирован. Если S не является пустым, то выбираем любой элемент из S и удаляем его. Полагаем этот элемент равным v.

(3) Если w не помечена, (v, w) является ребром и $f((v, w)) < c((v, w))$. Положить резерв вершины w равным минимуму величины $c((v, w)) - f((v, w))$ и резерва вершины v. Установить предшественника вершины w в v. Если $w \neq z$, добавить w в S.

(4) Если w не помечена, (w, v) является ребром и $f((v, w)) > 0$. Положить резерв вершины w равным минимуму из $f((v, w))$ и резерва вершины v. Установить предшественника вершины w в v. Добавить w во множество S.

(5) Если z помечена, то, используя функцию предшествования, вернуться в вершину a и для каждого ребра цепи добавить резерв вершины z к потоку каждого правильно ориентированного ребра, и вычесть резерв z из потока каждого неправильно ориентированного ребра. Вернуться к шагу 1.

(6) Вернуться к шагу 2.

Требуется еще доказать, что алгоритм действительно определяет максимальный поток, но сначала посмотрим, как он работает.

ПРИМЕР 16.13. Найдем максимальный поток для сети, изображенной на рис. 16.8.

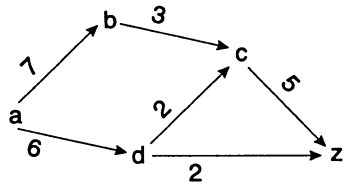


Рис. 16.8

На шаге 1 для каждой вершины устанавливаем предшественника и резерв равными символу $-$, для вершины a устанавливаем резерв, равный ∞, и полагаем $S = \{a\}$. После чего получаем сеть, изображенную на рис. 16.9.

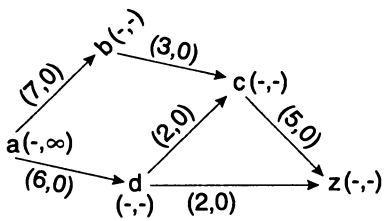


Рис. 16.9

На шаге 2 проверяем, не пусто ли множество S, и выбираем из него вершину a. На шаге 3 полагаем резерв вершины b равным $\min(7, \infty) = 7$ и устанавливаем вершину a в качестве предшественника вершины b. Помещаем вершину b во множество S. Устанавливаем резерв вершины d равным $\min(6, \infty) = 6$ и устанавливаем вершину a в качестве предшественника вершины d. Помещаем вершину d во множество S. Шаги 4 и 5 не применяем, поэтому возвращаемся к шагу 2.

Проверяем, не является ли множество S пустым, и выбираем из него вершину, например, b. Теперь множество S содержит только вершину d. На шаге 3 полагаем резерв вершины c равным $\min(3, 7) = 3$ и устанавливаем вершину b в качестве предшественника вершины c. Помещаем вершину c во множество S, так что $S = \{c, d\}$. Опять шаги 4 и 5 не применяем и возвращаемся к шагу 2.

Проверяем, не является ли множество S пустым, и выбираем из S вершину, например, c. Множество S опять содержит только вершину d. На шаге 3 полагаем резерв вершины z равным $\min(3, 5) = 3$ и устанавливаем вершину c в качестве предшественника вершины z. Не помещаем z во множество S. На шаге 4 следует пометить вершину d, но она уже помечена.

На шаге 5 видим, что вершина z уже помечена, и, используя функцию предшествования, находим цепь

$$a \overset{(7,0)}{\rightarrow} b \overset{(3,0)}{\rightarrow} c \overset{(5,0)}{\rightarrow} z.$$

Добавляя 3 (резерв вершины z) к потоку каждого ребра, получаем

$$a \overset{(7,3)}{\rightarrow} b \overset{(3,3)}{\rightarrow} c \overset{(5,3)}{\rightarrow} z,$$

что дает сеть, изображенную на рис. 16.10.

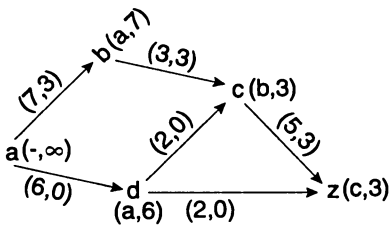


Рис. 16.10

Теперь возвращаемся к шагу 1, где переустанавливаем метки и предшественников и полагаем $S = \{a\}$. На шаге 2 выбираем вершину a из множества S. На шаге 3 устанавливаем резерв вершины b равным $\min(\infty, 7-3) = 4$ и устанавливаем a в качестве предшественника вершины b. Помещаем вершину b во множество S. Полагаем также резерв вершины d равным 6 и устанавливаем вершину a в качестве предшественника вершины d. Помещаем вершину d во множество S. Шаги 4 и 5 не применяем, поэтому возвращаемся к шагу 2. Выбираем вершину b из множества S. Поскольку поток из b в c равен пропускной способности из b в c, то вершину c не помечаем. Выбираем вершину d из множества S и продолжаем процесс, после чего получаем цепь

$$a \overset{(6,0)}{\rightarrow} d \overset{(2,0)}{\rightarrow} z$$

и, добавляя 2 (резерв вершины z) к потоку каждого ребра, получаем цепь

$$a \overset{(6,2)}{\rightarrow} d \overset{(2,2)}{\rightarrow} z.$$

Текущая ситуация изображена на рис. 16.11.

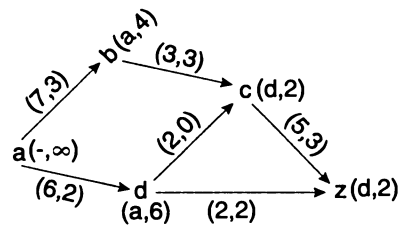


Рис. 16.11

Снова возвращаемся к шагу 1 и повторяем процесс, пока не получим цепь

$$a \overset{(6,2)}{\rightarrow} d \overset{(2,0)}{\rightarrow} c \overset{(5,3)}{\rightarrow} z.$$

Добавляя 2 (резерв вершины z) к потоку каждого ребра, получаем

$$a \overset{(6,4)}{\rightarrow} d \overset{(2,2)}{\rightarrow} c \overset{(5,5)}{\rightarrow} z.$$

Ситуация изображена на рис. 16.12.

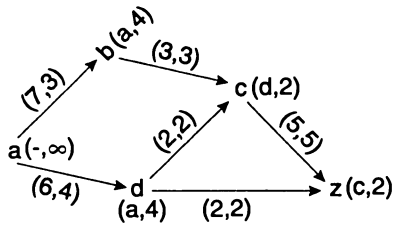


Рис. 16.12

Возвращаемся к шагу 1, переустанавливаем метки предшественника и полагаем $S = \{a\}$. На шаге 2 выбираем вершину a из множества S. Как и ранее, на шаге 3 полагаем резерв вершины b равным 4 и устанавливаем вершину a в качестве предшественника вершины b. Помещаем вершину b во множество S. Устанавливаем также резерв вершины d равным $6 - 4 = 2$ и помещаем вершину d во множество S. Возвращаемся к шагу 2. Выбираем вершину b из множества S. Поскольку поток из вершины b в вершину c равен пропускной способности из b в c, пометить c нельзя, поэтому, в конце концов, возвращаемся к шагу 2. Выбираем вершину d из множества S. Поскольку поток из вершины d в вершину c и вершину z равен их пропускным способностям, нельзя помечать c и z. Опять возвращаемся к шагу 2, но множество S пусто. Процесс закончен. Окончательный поток изображен на рис.16.13.

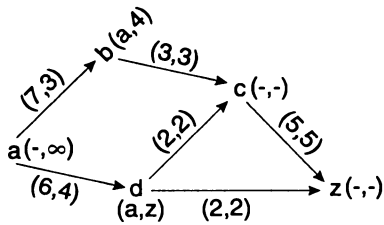


Рис. 16.13 □

ПРИМЕР 16.14. Найдите максимальный поток для сети, изображенной на рис. 16.14.

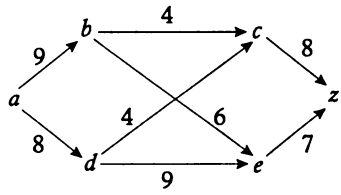


Рис. 16.14

Первые несколько проходов ничем не примечательны, поэтому опишем их вкратце. На первом проходе помечаем $a(-,\infty)$, $b(a,9)$, $d(a,8)$, $c(b,4)$, $e(b,4)$ и $z(c,4)$. Это дает $a - z$ путь

$$a \overset{(9,0)}{\rightarrow} b \overset{(4,0)}{\rightarrow} c \overset{(8,0)}{\rightarrow} z,$$

и поскольку резерв вершины z равен 4, добавляем 4 к каждому потоку, получая

$$a \overset{(9,4)}{\rightarrow} b \overset{(4,4)}{\rightarrow} c \overset{(8,4)}{\rightarrow} z,$$

что изображено на рис. 16.15.

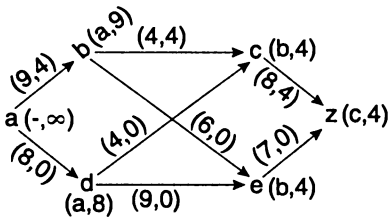


Рис. 16.15

На втором проходе помечаем $a(-,\infty)$, $b(a,5)$, $d(a,8)$, $c(d,4)$, $e(d,8)$ и $z(e,7)$. Это дает $a-z$ путь

$$a \overset{(8,0)}{\rightarrow} d \overset{(9,0)}{\rightarrow} e \overset{(7,0)}{\rightarrow} z,$$

и поскольку резерв вершины z равен 7, добавляем 7 к каждому потоку, получая

$$a \overset{(8,7)}{\rightarrow} d \overset{(9,7)}{\rightarrow} e \overset{(7,7)}{\rightarrow} z,$$

что изображено на рис. 16.16.

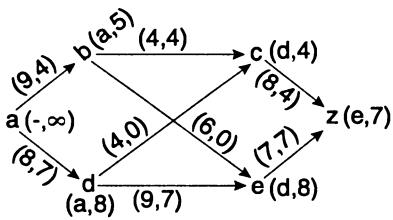


Рис. 16.16

На третьем проходе помечаем $a(-,\infty)$, $b(a,5)$, $d(a,1)$, $c(d,1)$, $e(d,1)$ и $z(c,1)$. Это дает $a-z$ путь

$$a \overset{(8,7)}{\rightarrow} d \overset{(4,0)}{\rightarrow} c \overset{(8,4)}{\rightarrow} z,$$

и поскольку резерв z равен 1, добавляем 1 к каждому потоку, получая

$$a \overset{(8,8)}{\rightarrow} d \overset{(4,1)}{\rightarrow} c \overset{(8,5)}{\rightarrow} z,$$

что изображено на рис. 16.17.

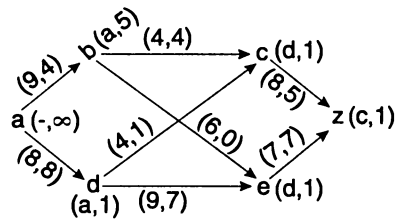


Рис. 16.17

Теперь начинается кое-что поинтересней, поэтому дальнейшее описание будет более подробным. Пометим $b(a,5)$, но d пометить нельзя; помещаем b в S. При выборе вершины b из множества S нельзя пометить вершину c, но можно

пометить $e(b,5)$; помещаем вершину c во множество S. Выбирая e из S, нельзя пометить вершину z, но т.к. d еще не помечена, ее можно пометить. Поскольку ребро (d,e) ориентировано неправильно, полагаем резерв вершины d равным $\min(7,5) = 5$. Устанавливаем e в качестве предшественника вершины d и помещаем вершину d во множество S. Выбирая вершину d из множества S, имеем единственную возможность — пометить вершину $c(d,3)$ и поместить c в S. Выбираем вершину c из множества S и помечаем $z(c,3)$. Это дает $a - z$ путь

$$a \overset{(9,4)}{\rightarrow} b \overset{(6,0)}{\rightarrow} e \overset{(9,7)}{\leftarrow} d \overset{(4,1)}{\rightarrow} c \overset{(8,5)}{\rightarrow} z,$$

и поскольку резерв z равен 3, прибавляем 3 к каждому потоку, исключая ребро (d,e). Это ребро ориентировано неправильно; вычитаем 3 из его потока, получая

$$a \overset{(9,7)}{\rightarrow} b \overset{(6,3)}{\rightarrow} e \overset{(9,4)}{\leftarrow} d \overset{(4,4)}{\rightarrow} c \overset{(8,8)}{\rightarrow} z,$$

что изображено на рис. 16.18.

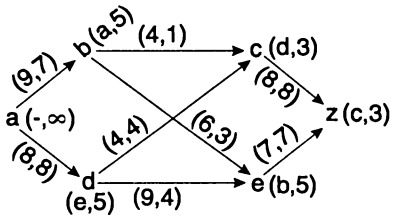


Рис. 16.18

Начинаем последний проход. Сначала $S = \{a\}$. Удаляем вершину a из множества S, помечаем вершину $b(a,2)$ и помещаем b в S. Удаляем вершину b из множества S, помечаем $e(b,2)$, $c(b,2)$ и помещаем вершины c, e во множество S. Удаляем e из S, помечаем $d(e,2)$, поскольку резерв d равен минимуму из потока ребра (d,e) и резерва вершины d. Помещаем d в S. Затем из S удаляем d, но ни один из оставшихся шагов нельзя выполнить, поэтому возвращаемся к шагу 2. Следующим из S удаляем c, но мы не можем выполнить ни один из оставшихся шагов и возвращаемся к шагу 2. Множество S пусто, поэтому алгоритм завершен. Результат изображен на рис. 16.19.

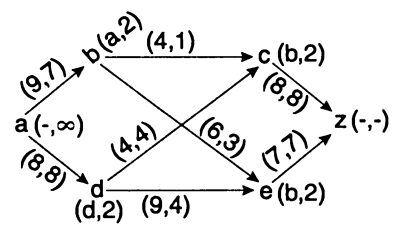


Рис. 16.19 □

ТЕОРЕМА 16.15. Алгоритм максимального потока строит максимальный поток для сети.

ДОКАЗАТЕЛЬСТВО. Пусть S — множество всех вершин, помеченных во время последнего прохода алгоритма, и $T = V - S$. Множество S не пусто, поскольку $a \in S$. Если e — ребро из S в T, например, $e = (s,t)$, так что t не помечено, то $f(e) = c(e)$, так как иначе t могла быть помечена. Если e — ребро из T в S, скажем, $e = (t,s)$, так что t не помечена, то $f(e) = 0$, так как в противном случае t могла быть помечена. Согласно следствиям 16.11 и 16.12 поток f — максимальный, и (S,T) — минимальное сечение. ■

Следствие 16.12 утверждает, что если $val(f) = C(S,T)$ для некоторого потока f и $a - z$ сечения (S,T), то f — максимальный поток, а C — минимальное сечение. В изложенном выше алгоритме показано, как находить минимальное сечение (S,T), так что f — максимальный поток, $val(f) = C(S,T)$. Соединяя все вместе, приходим к теореме.

ТЕОРЕМА 16.16. Поток f на заданной сети N является максимальным тогда и только тогда, когда существует сечение (S,T) такое, что $val(f) = C(S,T)$.

Заметим, что в примере 16.13 множество S состоит из всех вершин, которые можно пометить при последнем проходе, поэтому $S = \{a,b,d\}$ и $T = \{c,z\}$.

В примере 16.14 $S = \{a,b,c,d,e\}$ и $T = \{z\}$.

■ УПРАЖНЕНИЯ

1. Для сети, изображенной на рис. 16.20,
   - **а)** проверить сохранение потока в вершинах b, c и d;
   - **б)** найти $val(f)$, величину потока;
   - **в)** найти значение $C(S,T)$, где $S = \{a,b,c,d\}$;
   - **г)** найти значение $C(S,T)$, где $S = \{a,d,e\}$;
   - **д)** найти значение $C(S,T)$, где $S = \{a,b,d\}$.

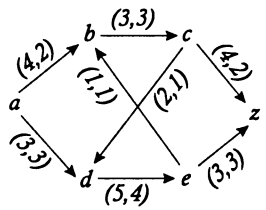


Рис. 16.20

2. Для сети, изображенной на рис. 16.21,
   - **а)** проверить сохранение потока в вершинах b, d и e;
   - **б)** найти $val(f)$, значение потока;
   - **в)** найти значение $C(S,T)$, где $S = \{a,b,c,d\}$;
   - **г)** найти значение $C(S,T)$, где $S = \{a,b,d,e\}$;
   - **д)** найти значение $C(S,T)$, где $S = \{a,b,e\}$.

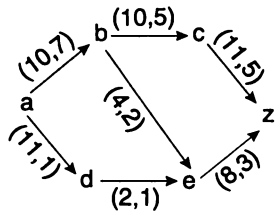


Рис. 16.21

3. Дополнить поток в сети, изображенной на рис. 16.22, так чтобы имело место сохранение потока.

4. Дополнить поток в сети, изображенной на рис. 16.23, так чтобы имело место сохранение потока.

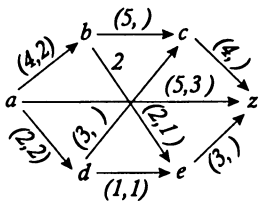


Рис. 16.22

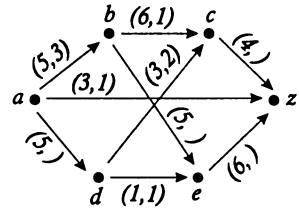


Рис. 16.23

5. Найти максимальный поток в сети, изображенной на рис. 16.24.

6. Найти максимальный поток в сети, изображенной на рис. 16.25.

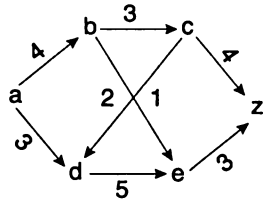


Рис. 16.24

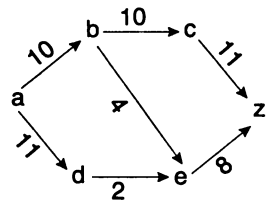


Рис. 16.25

7. Найти максимальный поток в сети, изображенной на рис. 16.26.

8. Найти максимальный поток в сети, изображенной на рис. 16.27.

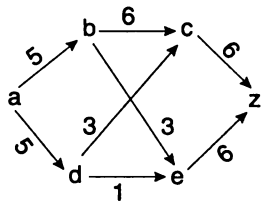


Рис. 16.26

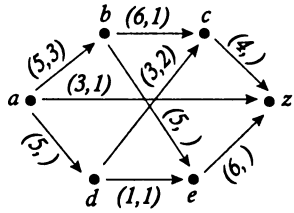


Рис. 16.27

9. Найти минимальное сечение для потока в сети из упражнения 5.
10. Найти минимальное сечение для потока в сети из упражнения 6.
11. Найти минимальное сечение — разрез для потока в сети из упражнения 7.
12. Найти минимальное сечение — разрез для потока в сети из упражнения 8.

16.2. ПАРОСОЧЕТАНИЕ

Теория потоков в сети имеет непосредственное применение в теории графов. Вспомним, что граф $G = (V, E)$ называется двудольным графом, если V можно выразить как несвязное объединение непустых множеств, например, $V = A \cup B$, так что каждое ребро имеет вид $\{a, b\}$, где $a \in A$ и $b \in B$. Таким образом, каждое ребро соединяет вершину из A с вершиной из B, и не существует двух вершин из одного и того же множества (A или B), которые соединены друг с другом. Подмножество M множества E называется ***паросочетанием***, если никакие два ребра из M не имеют общей вершины. Таким образом, никакие два ребра не являются инцидентными. Если $\{a, b\}$ — ребро в паросочетании, тогда как a, так и b имеет ***паросочетающие ребра***. При этом вершины a и b называются ***паросочетанными***. Например, для двудольного графа, изображенного на рис. 16.28, паросочетающие ребра, обозначенные на рисунке жирной линией, образуют паросочетание.

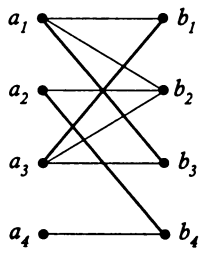


Рис. 16.28

Наглядным примером может служить соответствие между множеством работников и множеством заданий для них. Ребра графа показывают, какие задания может выполнить тот или иной работник. Еще один пример — паросочетание множества юношей и множества девушек. Паросочетающие ребра соединяют совместимые пары.(Будем считать, что эти множества образуют двудольный граф, и воздержимся от дальнейших комментариев.)

ОПРЕДЕЛЕНИЕ 16.17. Паросочетание M на двудольном графе $G = (V, E)$ называется ***максимальным***, если никакое другое паросочетание на G не содержит ребер больше, чем M.

ОПРЕДЕЛЕНИЕ 16.18. Паросочетание M на двудольном графе $G = (V, E)$, где $V = A \cup B$, называется ***полным***, если для каждого $a \in A$ существует $b \in B$ такое, что $\{a, b\} \in M$.

ПРИМЕР 16.19. Граф, изображенный на рис. 16.29, максимальный, но не полный.

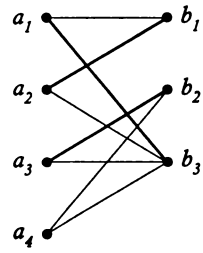


Рис. 16.29 □

Очевидно, нам нужен метод нахождения максимального паросочетания. Для упрощения задачи заменим рассматриваемый двудольный граф ориентированным графом, ребра которого имеют начало в вершинах множества A и заканчиваются в вершинах множества B. Добавим также две новые вершины a и z и ребра из вершины a в каждую вершину множества A, и ребра из каждой вершины множества B в вершину z. Граф, изображенный на рис. 16.30, превратится в граф, изображенный на рис. 16.31.

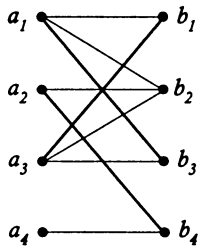


Рис. 16.30

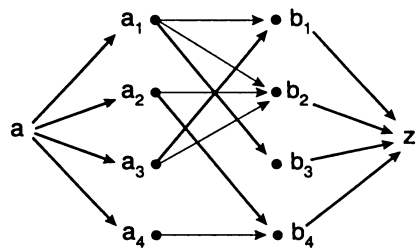


Рис. 16.31

Теперь двудольный граф превращен в сеть. Такая сеть называется ***сетью паросочетаний***. Для каждого $a_i \in A$ обозначим через e_i ребро из a в a_i. Положим пропускную способность $c(e_i) = 1$ для всех e_i. Аналогично, для каждого $b_i \in B$ обозначим через ε_i ребро из b_i в z и положим пропускную способность $c(\varepsilon_i) = 1$ для всех ε_i. Для ребра k_{ij} из a_i в b_j положим $c(k_{ij}) = |A| + 1$, так что пропускная способность больше, чем количество вершин в A. Причину объясним позже. Заметим, т.к. поток в каждую вершину a_i не может превысить 1, поток через каждое ребро k_{ij} из a_i в b_j должен быть 1 или 0 **независимо от пропускной способности** $c(k_{ij})$. Для ребра k_{ij} из a_i в b_j определим поток $f(k_{ij})$ равным 1, если имеется паросочетающее ребро между a_i и b_j, и равным 0 в противном случае. Величина потока, $val(f)$, в таком случае равна количеству ребер в паросочетании между A и B. Максимальный поток имеет место, когда имеется максимальное паросочетание. Таким образом, задача нахождения максимального паросочетания сведена к задаче построения максимального потока.

Воспользуемся алгоритмом нахождения максимального потока, описанным в предыдущем разделе. Начнем с ребра e_i, ведущего в вершину a_i, которое не

является паросочетающим, в противном случае $f(e_i) = c(e_i) = 1$. Затем движемся по ориентированному ребру k_{ij} к вершине b_j множества B. Если вершина b_j не паросочетается, создаем паросочетающее ребро между a_i и b_j, увеличивая поток. Если вершина b_j паросочетается, нужно вернуться по неправильно ориентированному ребру. Поэтому его поток должен быть равен 1, и мы продолжаем цепь назад во множество A по паросочетающему ребру. Снова, возвращаясь в B, выбираем ребро, которое не является паросочетающим. Продолжаем цепь, пока не достигнем вершины во множестве B, которая не паросочетается. В цепи необходимо добавить 1 к потоку правильно ориентированных ребер из A в B и вычесть 1 из потоков неправильно ориентированных ребер из B в A. В результате паросочетающими ребрами станут те, которые ими не были, а прежние паросочетающие ребра будут удалены. Это добавит одно паросочетающее ребро двудольному графу, и мощность паросочетания увеличится на 1. Таким образом, алгоритм нахождения максимального потока позволяет найти максимальное паросочетание.

Начать можно с пустого паросочетания, как того требует алгоритм нахождения максимального потока. В этом случае используем алгоритм построения паросочетающих ребер, а затем увеличим их число, чтобы найти максимальное паросочетание. В следующих примерах мы предполагаем, что изначально имеется некоторое паросочетание, которое будем увеличивать до максимального.

Мы видим, что при использовании этого метода нет необходимости заботиться о вершинах a и z. Мы просто начинаем с вершины из множества A, которая не паросочетается, и продолжаем движение назад и вперед, пока оно не закончится в вершине множества B, у которой нет паросочетающего ребра.

ПРИМЕР 16.20. Для графа, изображенного на рис. 16.32, найти максимальное паросочетание.

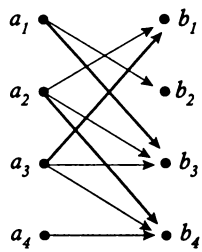


Рис. 16.32

Начинаем в вершине a_4, поскольку она не имеет паросочетающего ребра. По ориентированному ребру попадаем в вершину b_4, т.к. ребро (a_4, b_4) не принадлежит начальному паросочетанию. Продолжаем цепь назад в a_2, поскольку ребро (a_2, b_4) принадлежит текущему паросочетанию. Из a_2 цепь продолжаем в b_1, поскольку (a_2, b_1) — не паросочетающее ребро. Из b_1 цепь продолжаем в a_3, поскольку (a_3, b_1) — паросочетающее ребро. Из a_3 цепь продолжаем в вершину b_3, т.к. (a_3, b_3) не принадлежит паросочетанию. Из b_3 цепь продолжаем назад в a_1, поскольку (a_1, b_3) — паросочетающее ребро. Из a_1 цепь продолжаем в b_2, поскольку (a_1, b_2) не принадлежит паросочетанию. Из b_2 не выходит ни одно паросочетающее ребро, поэтому цепь построена и имеет вид: $a_4 \to b_4 \leftarrow a_2 \to b_1 \leftarrow a_3 \to b_3 \leftarrow a_1 \to b_2$. После этого удаляем из исходного

паросочетания ребра, имеющиеся в цепи, и добавляем ребра из цепи, которые в нем отсутствуют. Получаем граф, изображенный на рис. 16.33.

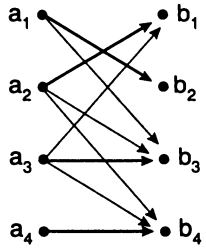


Рис. 16.33

Построенное паросочетание является максимальным и, к тому же, полным паросочетанием, потому что количество ребер в паросочетании равно количеству вершин в A. □

ПРИМЕР 16.21. Для графа, изображенного на рис. 16.34, найти максимальное паросочетание.

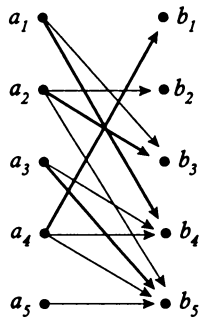


Рис. 16.34

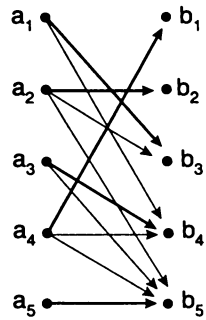


Рис. 16.35

Начинаем в вершине a_5, поскольку у нее нет просочетающего ребра. По ориентированному ребру попадаем в вершину b_5, т.к. (a_5, b_5) — не паросочетающее ребро. Затем продолжаем цепь назад в a_3, потому что (a_3, b_5) — паросочетающее ребро. Из a_3 продолжаем цепь в b_4, поскольку (a_3, b_4) — не паросочетающее ребро. Далее продолжаем цепь назад в a_1, т.к. (a_1, b_4) — паросочетающее ребро. Из a_1 продолжаем цепь в b_3, потому что (a_1, b_3) — не паросочетающее ребро. После этого возвращаем цепь в a_2, поскольку (a_2, b_3) — паросочетающее ребро. Из a_2 продолжаем цепь в b_2, (a_2, b_2) — не паросочетающее ребро. Из b_2 не выходит ни одно паросочетающее ребро, поэтому цепь построена и имеет вид:

$$a_5 \to b_5 \leftarrow a_3 \to b_4 \leftarrow a_1 \to b_3 \leftarrow a_2 \to b_2.$$

Удаляем из исходного паросочетания ребра, имеющиеся в цепи, и добавляем ребра из цепи, которые в нем отсутствуют. Получаем граф, изображенный на рис. 16.35.

Поскольку количество ребер в паросочетании равно количеству вершин в A, имеем максимальное паросочетание, которое является полным. □

ОПРЕДЕЛЕНИЕ 16.22. Для подмножества X множества A введем множество $R(X) = \{b : b \in B$ и b смежная с вершиной из $X\}$. $R(X)$ называется множеством значений X.

Это определение согласовано с определением множества значений, которое представлено в главе 2 при изучении отношений. Теперь можно сформулировать необходимые и достаточные условия полноты паросочетания.

ТЕОРЕМА 16.23. (Теорема Холла) Двудольный граф $G(E,V)$, в котором $V = A \cup B$, имеет полное паросочетание тогда и только тогда, когда для каждого подмножества $X \subseteq A$ выполняется $|X| \leq |R(X)|$.

ДОКАЗАТЕЛЬСТВО. Очевидно, что если граф $G(E,V)$ — полный, тогда необходимо $|X| \leq |R(X)|$ для каждого подмножества X множества A, т.к. в противном случае не для всех элементов из X и, следовательно, не для всех элементов из A существует паросочетающее ребро. В частности, каждый элемент из A смежен с элементом из B и $|A| \leq |B|$.

Обратно, предположим, что $|X| \leq |R(X)|$ для каждого подмножества X множества A. Согласно теореме 16.16 существует разрез (S,T) такой, что $val(f) = C(S,T)$. Известно, что $val(f)$ равно количеству вершин в максимальном паросочетании, поэтому $C(S,T)$ тоже равно количеству вершин в макисимальном паросочетании. Пусть $|A| = n$. Тогда $C(S,T) \leq n$, поскольку количество элементов в максимальном паросочетании не может превысить количество вершин в A. Вспомним, что $C(S,T)$ равно сумме пропускных способностей всех ребер между S и T. Известно также, что этот разрез не может содержать ребро k_{ij} между вершиной a_i из A и вершиной b_j из B, т.к. $c(k_{ij}) = n+1$. Поэтому, если $a_i \in S$, то $R(\{a\}) \subseteq S$. Отсюда следует, что $R(A \cap S) \subseteq S$. Между каждой вершиной из $R(A \cap S)$ и z существует ребро. Кроме того, каждое из этих ребер принадлежит разрезу (S,T), поскольку $z \in T$ и вносит 1 в $C(S,T)$. Множество A — несвязное объединение множеств $A \cap S$ и $A \cap T$. Следовательно, $n = |A| = |A \cap S| + |A \cap T|$. Между элементом a и каждым элементом множества $A \cap T$ существует ребро. Поскольку $a \in S$, каждое из этих ребер принадлежит разрезу (S,T) и вносит 1 в $C(S,T)$. Поскольку по предположению $|A \cap S| \leq |R(A \cap S)|$, отсюда следует, что

$$C(S,T) \geq |R(A \cap S)| + |A \cap T| \geq |A \cap S| + |A \cap T| = |A| = n.$$

Вспомним, однако, что $C(S,T)$ равно количеству ребер в максимальном паросочетании. Поэтому $C(S,T) \leq n$, откуда следует, что $C(S,T) = n$, т.е. максимальное паросочетание является полным. ■

Теперь представим другой алгоритм нахождения максимального паросочетания. На этот раз используем матрицы. Начнем с модифицированной матрицы инцидентности. Поскольку имеются ребра только из A в B, то необходима та часть матрицы A, в которой строка соответствует элементу из A, а столбец соответствует элементу из B. Строки именуем вершинами из A, столбцы — вершинами из B. Как и раньше, помещаем 1 в i-ую строку j-ого столбца, если существует ребро из a_i в b_j. Начнем с расстановки скобок вокруг каждой 1 в i-ой строке и j-ом столбце, если существует паросочетающее ребро из a_i в b_j. Проиллюстрируем этот процесс для паросочетания из примера 16.20. Помещаем символ # в

конец строки, соответствующей a_4, чтобы отметить, что в строке нет [1] и, следовательно, для a_4 нет паросочетающего ребра. Считаем строку, содержащую такой символ в конце, помеченной. Точно так же считаем помеченным каждый столбец, содержащий в конце указанный символ. На этом этапе матрица имеет следующий вид:

	b_1	b_2	b_3	b_4	
a_1	0	1	[1]	0	
a_2	1	0	1	[1]	
a_3	[1]	0	1	1	
a_4	0	0	0	1	#

Поскольку строка, содержащая a_4, помечена, то везде, где имеется 1 без скобок, помещаем a_4 под каждым столбцом, содержащим 1. Числа внизу и справа помогут найти обратный путь. Переход по 1 в столбце b_4 равносилен переходу из a_4 в b_4 по непаросочетающему ребру. В столбце, соответствующем b_4, выберем строку, содержащую [1], которая встречается в a_2, и поместим b_4 в правый столбец этой строки, что эквивалентно построению паросочетающего ребра из b_4 в a_2. На этом этапе имеем следующую матрицу:

	b_1	b_2	b_3	b_4	
a_1	0	1	[1]	0	
a_2	1	0	1	[1]	b_4
a_3	[1]	0	1	1	
a_4	0	0	0	1	#
				a_4	

Теперь ищем вхождение единиц в строке a_2 и помещаем a_2 внизу каждого столбца. В данном случае единицы имеются в столбцах b_1 и b_3. Имеем, таким образом, две возможности для непаросочетающего ребра. В столбцах, соответствующих b_1 и b_3, находим строки, содержащие [1]. Ими являются строки, соответствующие a_3 и a_1. Помещаем b_1 в правый столбец строки a_3 и b_3 — в правый столбец строки a_1. Получаем матрицу вида

	b_1	b_2	b_3	b_4	
a_1	0	1	[1]	0	b_3
a_2	1	0	1	[1]	b_4
a_3	[1]	0	1	1	b_1
a_4	0	0	0	1	#
	a_2		a_2	a_4	

В строке a_1 отыскиваем 1 в одном из столбцов. Таким столбцом является b_2. Под ним помещаем a_1. В строке a_3 отыскиваем 1 в столбцах. Ими являются столбцы b_3 и b_4, но они уже помечены. В столбце b_2 нет ни одной [1], поэтому вершина из B, не имеющая паросочетающего ребра, найдена. Следовательно,

искомая цепь построена. Теперь имеем матрицу.

	b_1	b_2	b_3	b_4	
a_1	0	1	[1]	0	b_3
a_2	1	0	1	[1]	b_4
a_3	[1]	0	1	1	b_1
a_4	0	0	0	1	#
	a_2	a_1	a_2	a_4	

Используя метки ребер внизу и в самом крайнем столбце таблицы, возвращаемся назад и восстанавливаем цепь

$$a_4 \to b_4 \leftarrow a_2 \to b_3 \leftarrow a_1 \to b_2.$$

Заметим, что цепь *не* совпадает с построенной в примере 16.20. Она короче. Изменение статуса ребер (паросочетающие — на непаросочетающие и наоборот), как в примере 16.20, дает паросочетание, изображенное на рис. 16.36.

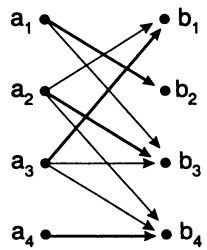


Рис. 16.36

Для получения новой матрицы каждую 1 по ходу движения меняем на [1], а каждую [1] — на 1. В результате получаем такую матрицу.

	b_1	b_2	b_3	b_4	
a_1	0	[1]	1	0	
a_2	1	0	[1]	1	
a_3	[1]	0	1	1	
a_4	0	0	0	[1]	

Используя приведенную выше процедуру для паросочетания из примера 16.21, начинаем с матрицы

	b_1	b_2	b_3	b_4	b_5	
a_1			1	[1]		
a_2			[1]		1	
a_3				1	[1]	
a_4	[1]				1	
a_5					1	

и в конце приходим к матрице

	b_1	b_2	b_3	b_4	b_5	
a_1			1	[1]		b_4
a_2		1	[1]		1	b_3
a_3				1	[1]	b_5
a_4	[1]			1	1	
a_5					1	#
	a	a_2	a_1	a_3	a_5	

Используя метки ребер внизу и в самом крайнем столбце таблицы, возвращаемся назад и восстанавливаем цепь

$$a_5 \to b_5 \leftarrow a_3 \to b_4 \leftarrow a_1 \to b_3 \leftarrow a_2 \to b_2,$$

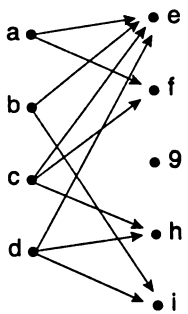
которая совпадает с полученной в примере 16.21. Матрица нового паросочетания имеет следующий вид.

	b_1	b_2	b_3	b_4	b_5	
a_1			[1]	1		b_4
a_2		[1]	1		1	b_3
a_3				[1]	1	b_5
a_4	[1]			1	1	
a_5					[1]	#
	a	a_2	a_1	a_3	a_5	

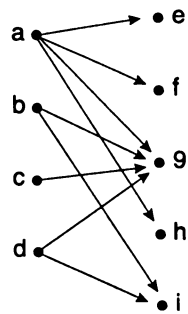
■ УПРАЖНЕНИЯ

1. Для приведенных ниже двудольных графов найдите максимальное паросочетание, используя методы, изложенные в этой главе. Является ли каждое паросочетание полным?

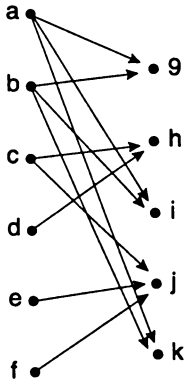
а)



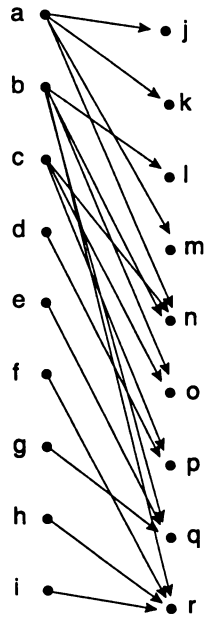
б)



в)

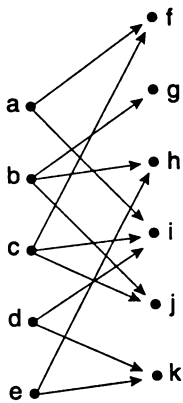


г)

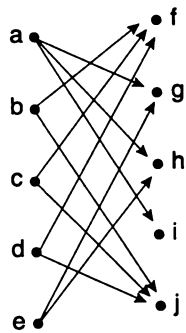


2. Выполните задание предыдущего упражнения для следующих графов.

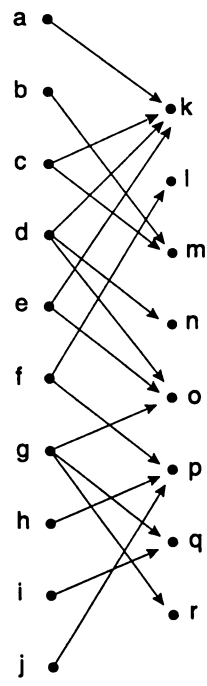
а)



б)



в)



3. Шесть человек ищут работу. Человек a готов выполнять работы A, D и F. Человек b готов выполнять работу C. Человек c готов выполнять работы B и D. Человек d готов выполнять работы A и E. Человек e готов выполнять работы B и D. Человек f — работу E. Найдите паросочетание, согласно которому наибольшее количество людей получит работу, которую они готовы выполнять. Все ли получат работу?
4. Шестеро девушек хотят выбрать партнера для танцев. Элис нравятся Чарльз и Эдвард; Бетти нравятся Альберт и Дэн; Клара предпочитает Чарльза и Эдварда; Донна любит танцевать с Дэном; Элизабет нравятся Барт и Эдвард. Подберите как можно большему числу девушек пару для танцев, которая была бы им по душе. Возможно ли каждой девушке подобрать парня, который ей нравится?
5. Эмми входит в состав комитетов a и b. Брюс — комитетов a и c. Чарльз — член комитета b. Дэнис — член комитетов b, c и e. Эллен входит в состав комитетов d и e. Председатель студенческого совета планирует пригласить на совещание пятерых членов совета, входящих в различные комитеты. Возможно ли это?
6. Шестеро друзей собираются на маскарад. Все хотят быть в разных костюмах. Анне нравится костюм клоуна; Бену нравятся костюмы приведения и Франкенштейна; Кортни — клоуна и костюмы из вестернов; Денис хотел бы нарядиться приведением, Дракулой или Франкенштейном; Эллен предпочитает костюм клоуна, костюм из вестернов или Шерлока Холмса; Френк — костюм Дракулы и клоуна. Есть ли возможность для всех пойти на маскарад в разных костюмах, причем в тех, которые им нравятся?
7. Составьте формальный алгоритм для нахождения максимального паросочетания M на двудольном графе $G = (V, E)$ без использования матриц.
8. Составьте формальный алгоритм для нахождения максимального паросочетания M на двудольном графе $G = (V, E)$ с использованием матриц.

16.3. СЕТИ ПЕТРИ

В этом разделе продолжается рассмотрение двудольных графов с ориентированными ребрами. Однако, в отличие от предыдущего раздела ориентация ребер допускается в обоих направлениях. Пусть $G(V, E)$ — двудольный ориентированный граф, в котором $V = P \cup T$. Множество P называется множеством ***позиций***, а множество T — множеством ***переходов***. E — множество ориентированных ребер между P и T. В рассмотрение включено также множество функций M таких, что каждая функция $\mu \in M$ ставит в соответствие неотрицательное целое число каждому элементу множества P. Функция μ называется ***разметкой*** графа G. Ориентированный граф, обладающий такими свойствами, называется ***сетью Петри***.

Сети Петри используются главным образом для моделирования параллельных процессов: для моделирования компонентов компьютера, параллельных вычислений, в робототехнике и даже для описания музыкальных структур (см. [41]). Вообще, сети Петри используются для нахождения дефектов в проекте системы,

хотя имеют и многие другие применения. Они обладают многими свойствами блок-схем и конечных автоматов, речь о которых пойдет дальше. Более подробно см. [85].

Каждая позиция в сети Петри обозначена кружком, а каждый переход — вертикальной линией. Если p — позиция, то $f(p)$ называется разметкой позиции p. Разметка множества G показана на графе с помощью больших черных точек, называемых ***метками***, помещенных в кружки, которые обозначают позиции. Если кружок позиции p пуст, это означает, что в p меток нет, поэтому $f(p) = 0$.

Рассмотрим граф, изображенный на рис. 16.37. В этой сети Петри p_1 и p_2 — позиции, а t_1 — переход. Позиция p_1 содержит одну метку, а p_2 меток не содержит.

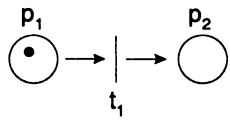


Рис. 16.37

Срабатывание перехода t — это удаление по одной метке из каждой позиции p_i, так что имеется ориентированное ребро из p_i в t, и добавление метки в каждую позицию p_j, так что имеется ориентированное ребро из t в p_j. Например, срабатывание t_1 в приведенном выше примере приводит к сети Петри, изображенной на рис. 16.38.

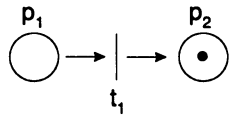


Рис. 16.38

В сети Петри, изображенной на рис. 16.39, срабатывание перехода t_1 даст в результате сеть Петри, изображенную на рис. 16.40.

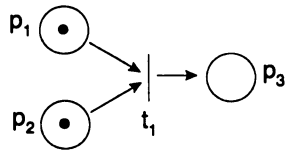


Рис. 16.39

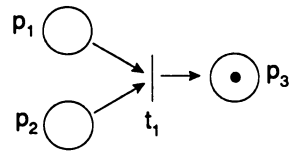


Рис. 16.40

В сети Петри, изображенной на рис. 16.41, срабатывание перехода t_1 даст в результате сеть Петри, изображенную на рис. 16.42.

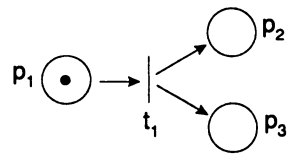


Рис. 16.41

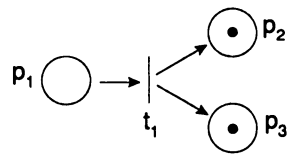


Рис. 16.42

В сети Петри, изображенной на рис. 16.43, срабатывание перехода t_1 даст в результате сеть Петри, изображенную на рис. 16.44, поскольку имеется ориентированное ребро из t_1 обратно в p_1.

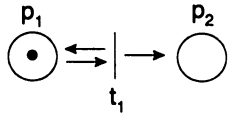


Рис. 16.43

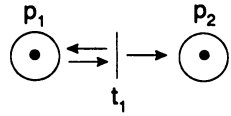


Рис. 16.44

Переход t может сработать, когда каждая позиция p_i такая, что имеется ориентированное ребро из p_i в t, содержит метку. Когда такая ситуация имеет место, говорят, что переход t ***разрешен***. Переход t_1 в сети Петри, изображенной на рис. 16.45, не может сработать, поскольку в p_2 нет метки.

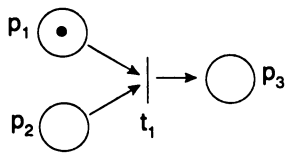


Рис. 16.45

В некоторых случаях необходимо, чтобы при срабатывании перехода удалялась или прибавлялась более чем одна метка. Если при срабатывании перехода t из позиции p необходимо удалить более чем одну метку, то ориентированное ребро из p в t помечается таким количеством меток, которое следует удалить из p. Если меток нет, то их количество по умолчанию предполагается равным 1. Если при срабатывании перехода t в позицию p необходимо добавить более чем одну метку, то ориентированое ребро из t в p помечается таким количеством меток, которое нужно добавить в p. Если меток нет, то количество опять предполагается равным 1. Если при срабатывании перехода t из позиции p следует удалить n меток, то p должна содержать, по крайней мере, n меток, чтобы переход t был разрешен. В сети Петри, изображенной на рис. 16.46, срабатывание перехода t_1 дает в результате сеть Петри, изображенную на рис. 16.47. Переход t_1 не может сработать снова, поскольку в p_1 нет достаточного количества меток.

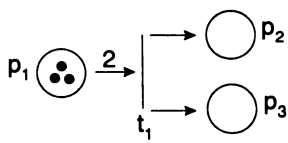


Рис. 16.46

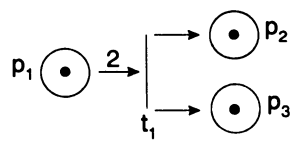


Рис. 16.47

Срабатывание перехода предполагается мгновенным и может произойти в любой момент, когда переход разрешен. Порядок срабатывания перехода ограничен только наличием достаточного количества меток, необходимых для разрешенности перехода. Вполне возможно, что в некоторый момент времени более

чем один переход будет разрешен. В этом случае любой из них может сработать первым. Например, если моделируются несколько процессоров, имеющих общие данные и периферию, то неизвестно, когда одному из процессоров понадобится принтер или данные. Сети Петри используются для проверки работоспособности системы даже в условиях такой неопределенности системы.

Состоянием s сети Петри называется количество меток в каждой позиции, определяемое функцией $\mu_i \in M$. Таким образом, состояние сети Петри — это то же самое, что и разметка сети Петри. Термин "состояние" используется при рассмотрении автоматов и других машин с конечным количеством состояний. Поскольку μ_i полностью определяют состояние, будем их отождествлять. Срабатывание перехода изменяет текущее состояние на состояние, определяемое другой функцией μ_j из M. Поэтому новое состояние сети Петри зависит от текущего состояния и сработавшего перехода. Таким образом, $\mu_j = \delta(\mu_i, t)$, где $\delta(\mu_i, t)$ — состояние, полученное из состояния μ_i. Если в некоторый момент времени не существует разрешенного перехода, то функция состояния более не определяется, и процесс, описывающий сеть Петри, завершается.

Рассмотрим сеть Петри, изображенную на рис. 16.48. При срабатывании перехода t_1 сеть Петри переходит в состояние, показанное на рис. 16.49. При срабатывании перехода t_4 имеем сеть Петри, изображенную на рис. 16.50.

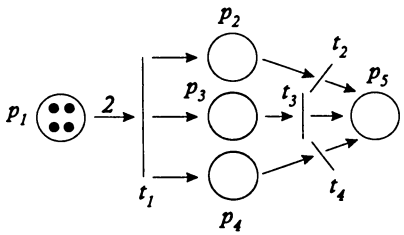


Рис. 16.48

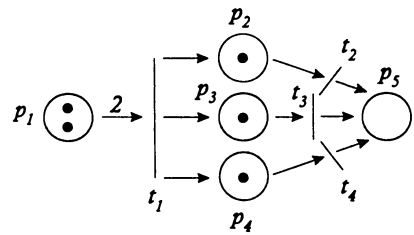


Рис. 16.49

Вполне очевидно, что в конце концов все метки переместятся в p_5, и процесс завершится. Процесс в сети Петри, изображенной на рис. 16.51, никогда не завершится, поскольку переход t_1 всегда разрешен.

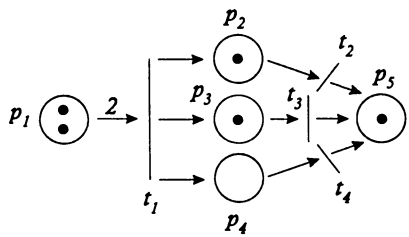


Рис. 16.50

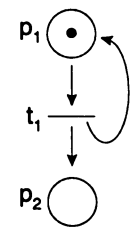


Рис. 16.51

Сеть Петри на рис. 16.52 выполняет вычисления $(a+b) \times (c+d)$. Каждый из переходов t_1 или t_2 может сработать первым; однако переход t_3 не будет разрешен, пока оба перехода t_1 и t_2 не сработают.

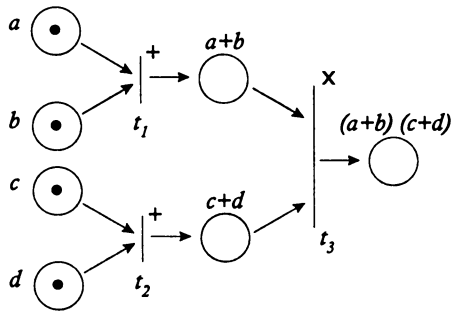


Рис. 16.52

Состояние μ_j называется ***непосредственно достижимым*** из состояния μ_i, если срабатывание какого-либо перехода t в то время, пока сеть находится в состоянии μ_i, приводит к состоянию μ_j. Более формально, состояние μ_j непосредственно достижимо из состояния μ_i, если существует переход t такой, что $\mu_j = \delta(\mu_i, t)$. Состояние μ_j ***достижимо*** из состояния μ_i, если, начиная из состояния μ_i, срабатывание последовательности переходов производит состояние μ_j. Состояние μ_{i_m} ***достижимо*** из состояния μ_{i_1}, если существуют состояния $\mu_{i_2}, \mu_{i_3}, \cdots, \mu_{i_{m-1}}$ и переходы $t_1, t_2, \ldots, t_{m-1}$ такие, что $\mu_{i_{k+1}} = \delta(\mu_{i_k}, t_k)$ для $k = 1, 2, \ldots, m-1$. В частности, состояние, изображенное на рис. 16.53, непосредственно достижимо из состояния, изображенного на рис. 16.54, поскольку оно производится срабатыванием перехода t_1.

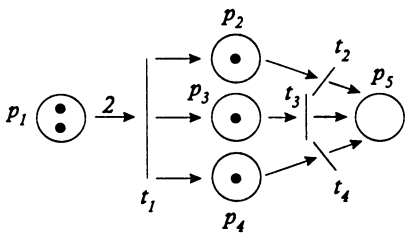


Рис. 16.53

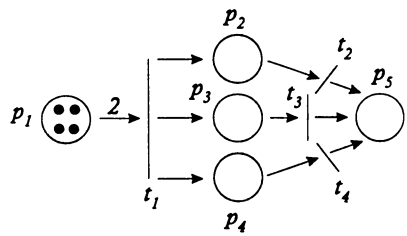


Рис. 16.54

Состояние сети Петри, изображенное на рис. 16.55, достижимо из состояния сети Петри, изображенного на рис. 16.56, поскольку оно производится срабатыванием перехода t_1, а затем срабатыванием перехода t_4.

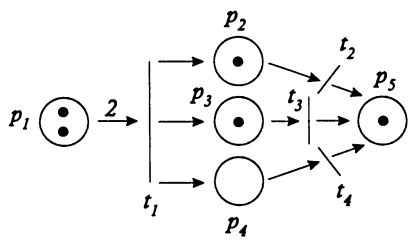


Рис. 16.55

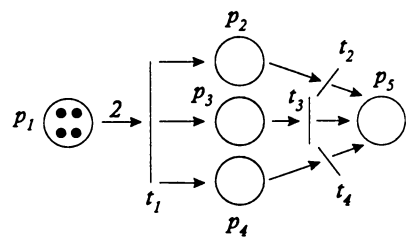


Рис. 16.56

Сеть Петри называется ***живой***, если для любого состояния μ и любого перехода t существует состояние μ', достижимое из состояния μ, так что переход t в состоянии μ' является разрешенным. Таким образом, каким бы ни было текущее состояние, существует такая последовательность срабатывания переходов с началом в текущем состоянии, что любой заданный переход может сработать. Очевидно, что сеть на рис. 16.57 является живой.

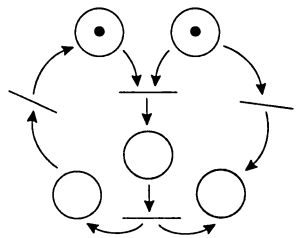


Рис. 16.57

Одно из определений тупиковой сети Петри — сеть не живая. Следовательно, существует такое состояние, что если сеть Петри находится в этом состоянии, то один или более переходов могут никогда не сработать. Назовем это состояние ***частичным тупиком***. Будем говорить, что сеть Петри ***находится в тупике***, если имеется состояние, в котором ни один из переходов не может сработать. Таким образом, существует такое состояние μ, когда $\delta(\mu, t)$ не определена для всех переходов t. Рассмотрим стандартный пример, когда два процессора используют совместно два ресурса, например, принтер и память, каждый из которых не может быть использован ими одновременно. Если один процессор имеет доступ к принтеру, а другой — к памяти, то оба не могут завершить задания. Они также не могут освободить используемый ресурс, поэтому система находится в тупике. Эта ситуация показана на рис. 16.58.

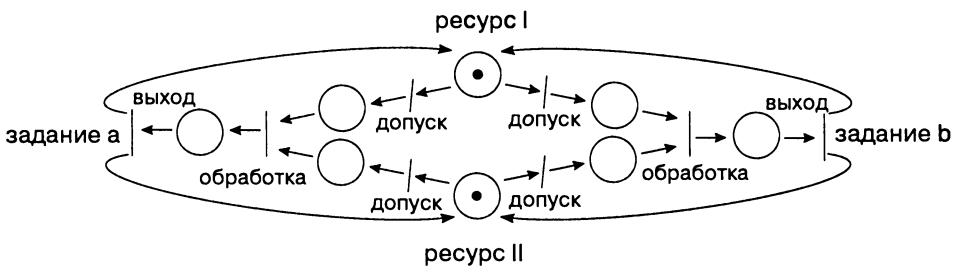


Рис. 16.58

Еще один пример, **проблему обедающих философов**, мы рассмотрим в упражнениях.

Другая, с которой приходится сталкиваться при использовании сети Петри для совместного использования файлов, — это проблема взаимного исключения. Предположим, что одновременный доступ двух людей к одним и тем же данным

нежелателен. Так, недопустимо, чтобы один человек читал данные, а другой в это же время их изменял. Эта проблема решается путем взаимного исключения. Пока один человек работает с данными, для другого доступ закрыт. Такая ситуация проиллюстрирована на рис. 16.59. Единственная метка в позиции *c* препятствует допуску другого субъекта до тех пор, пока метка не будет перемещена.

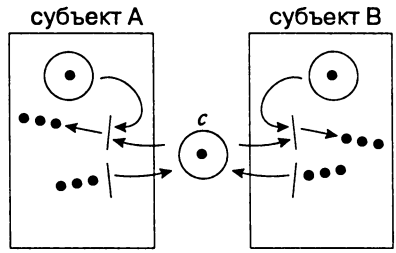


Рис. 16.59

Сеть Петри называется ***безопасной***, если каждая позиция содержит не более одной метки. Когда сеть Петри безопасна, в каждой позиции имеется либо одина метка, либо они отсутствуют вообще, поэтому помещение метки в позицию является бинарной операцией. Для большинства моделей контроля необходима безопасная сеть Петри, поскольку наличие метки может означать протекание процесса, а ее отсутствие — сигнал о его остановке.

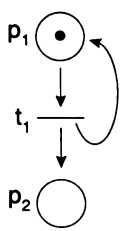


Рис. 16.60

Сеть Петри называется ***ограниченной***, если количество меток в каждой позиции не превышает некоторое целое число k. Ограниченная сеть Петри предоставляет возможность контроля проблемы переполнения. Очевидно, что безопасная цепь Петри — ограниченная. Сеть Петри, изображенная на рис. 16.60, ограниченной не является.

В некоторых сетях Петри на вместимость позиций накладываются ограничения, и по условию количество меток не может превышать эту вместимость.

Сеть Петри называется ***консервативной***, если общее число меток во всех позициях всегда постоянно. В этом случае количество меток на входе каждого перехода равно количеству меток на выходе перехода. Если метки описывают ресурсы, то консервативная сеть Петри гарантирует, что никакой ресурс не будет ни создан, ни утерян. Консервативная сеть Петри является очевидным образом ограниченной. Сеть Петри, изображенная на рис. 16.61, консервативна.

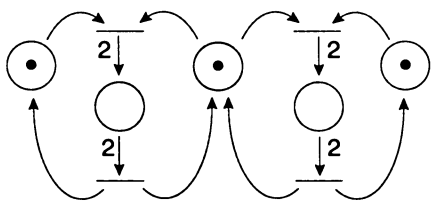
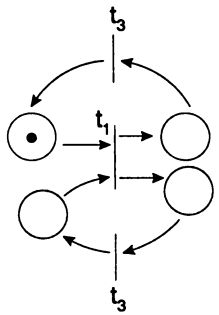


Рис. 16.61

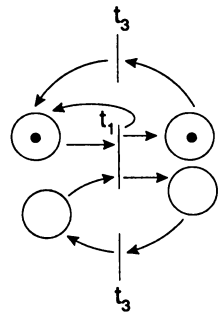
■ УПРАЖНЕНИЯ

1. Какие из приведенных ниже сетей Петри
   - **а)** являются безопасными?
   - **б)** являются консервативными?
   - **в)** являются ограниченными?
   - **г)** имеют достижимым каждое состояние?
   - **д)** являются живыми?
   - **е)** имеют все текущие переходы разрешенными?
   - **ж)** имеют переходы, которые могут быть разрешены в некотором состоянии, достижимом из текущего состояния?

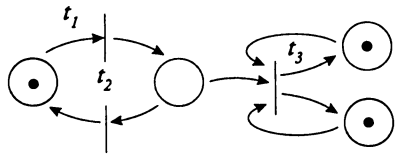
(i)



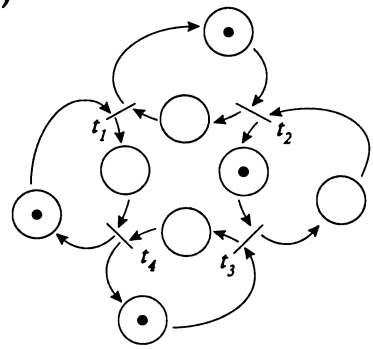
(ii)



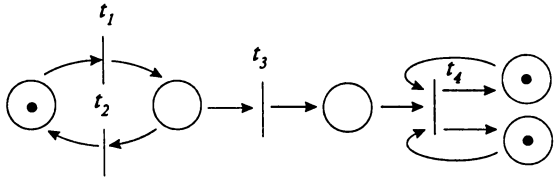
(iii)



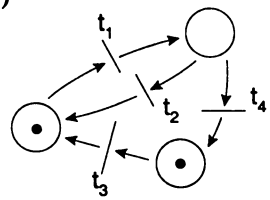
(iv)



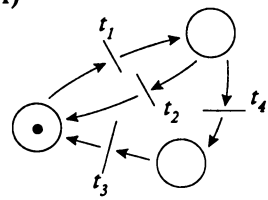
(v)



(vi)



(vii)



2. В сетях Петри (упражнение 1) найдите:
   а) все возможные состояния сети Петри;
   б) состояние после срабатывания перехода t_1;
   в) состояние после последовательного срабатывания переходов t_1 и t_2;
   г) любое тупиковое состояние;
   д) любое частично тупиковое состояние.
3. Постройте пример сети Петри, в которой два перехода разрешены, но срабатывание каждого из них блокирует другой.
4. Постройте пример сети Петри, которая является безопасной, но не является живой.
5. Постройте пример сети Петри с ограничением 4.
6. Постройте пример сети Петри, которая является безопасной, но не является ограниченной.
7. Постройте пример сети Петри, которая является консервативной, но не является безопасной.
8. Постройте сеть Петри, которая эквивалентна сети, изображенной на рис. 16.58, но не является тупиковой.
9. Может ли консервативная сеть Петри быть тупиковой?
10. Приведите пример сети Петри, которая имеет состояние, недостижимое из исходного состояния.
11. Постройте сеть Петри, которая частино тупиковая, но не тупиковая.
12. Проблема обедающих философов формулируется следующим образом: пять философов сидят за столом, чтобы есть и размышлять. В любой заданный промежуток времени они могут делать что-то одно. Философы едят китайское блюдо, используя палочки, поэтому для еды каждому необходимо иметь две палочки. Когда философы размышляют, палочки не нужны. На столе между двумя сидящими лежит одна палочка, т.е. всего пять палочек. Очевидно, что если философы одновременно возьмут правую палочку, то каждый будет лишен пищи. Постройте сеть Петри, моделирующую эту задачу, которая не была бы тупиковой, т.е. когда философ будет приниматься за еду, у него для этого будут две палочки по обе стороны от него.

ГЛАВА

17

ТЕОРИЯ ВЫЧИСЛЕНИЙ

17.1. РЕГУЛЯРНЫЕ ЯЗЫКИ

В примере 9.16 приведены следующие определения.

ОПРЕДЕЛЕНИЕ 17.1. Пусть A — конечное множество символов. Такое множество символов называется ***алфавитом***. Например, A может быть подмножеством английского алфавита или просто множеством $\{0,1\}$. ***Строка*** или ***слово*** символов из A имеет вид $a_1a_2a_3a_4\ldots a_n$, где $a_i \in A$.

Если $A = \{0,1\}$, то 1, 100, 101, 11011 и 0001 — строки символов из A. Определим также пустую строку — λ, не содержащую ни одного символа.

ОПРЕДЕЛЕНИЕ 17.2. Пусть A^* обозначает множество всех строк алфавита A. Определим бинарную операцию $\circ$, называемую ***конкатенацией*** на A^*, следующим образом: если $a_1a_2a_3a_4\ldots a_n$ и $b_1b_2b_3b_4\ldots b_m \in A^*$, то

$$a_1a_2a_3a_4\ldots a_n \circ b_1b_2b_3b_4\ldots b_m = a_1a_2a_3a_4\ldots a_nb_1b_2b_3b_4\ldots b_m.$$

Если $A = \{0,1\}$, то $11011 \circ 100010 = 11011100010$. В частности, если ω — строка из A^*, то $\lambda \circ \omega = \omega \circ \lambda = \omega$, и это свидетельствует о том, что если строке предшествует или следует за ней пустая строка, то в результате конкатенации получаем ту же самую строку.

ОПРЕДЕЛЕНИЕ 17.3. Пусть A^* обозначает множество всех строк алфавита A. Подмножество L множества A^* называется ***языком***.

Если A — множество букв английского алфавита, то L могло бы быть множеством слов английского языка. Если A — множество букв греческого алфавита, то L могло бы быть множеством слов греческого языка. Если A — множество символов языка программирования, то L — множество слов этого языка.

Если алфавит A={1, 0}, то следующие множества являются языками.

$L_1 = \{0, 1, 00, 11, 000, 111, \ldots\}$;
$L_2 = \{w : w \in A^*$ и содержит точно одну цифру $1\}$;
$L_3 = \{w : w \in A^*$ и содержит точно две цифры $1\}$;
$L_4 = \{w : w \in A^*$ и содержит не мене двух цифр $1\}$.

Если алфавит $A = \{a, b, c\}$, то следующие множества являются языками.

$L_1 = \{acb, aacbb, aaacbbb, \ldots\}$;
$L_2 = \{w : w = a^n b^n$ для $n \geq 1\}$;
$L_3 = \{w : w \in a^n b^m$ для $m, n \geq 1\}$;
$L_4 = \{w : w \in A^*$ и не содержит последовательных одинаковых букв$\}$.

ОПРЕДЕЛЕНИЕ 17.4. Пусть S — подмножество множества A^*. Тогда S^* — множество всех строк или слов, образованных конкатенацией слов из множества S, т.е. $S^* = \{w_1 w_2 \ldots w_n : w_i \in S\}$. Символ * называется звездой Клини в честь математика и логика Стивена Коула Клини.

Заметим, что данное определение согласовано с определением множества A^*.

ОПРЕДЕЛЕНИЕ 17.5. Пусть A — алфавит. ***Выражением*** над A называется строка символов, которая представляет собой подмножество множества A^*. ***Регулярное выражение*** состоит из символов *, $\vee$, λ, (,) и **a** для каждого элемента a из A. Класс ***регулярных выражений*** определяется следующими правилами.

i) Символ λ является регулярным выражением и для каждого $a \in A$ символ **a** — регулярное выражение.
ii) Если w_1 и w_2 — регулярные выражения, то $w_1 w_2$, $w_1 \vee w_2$, w_1^* и (w_1) — регулярные выражения.
iii) Все регулярные выражения образуются по правилам (i) и (ii).

Каждому выражению сопоставляется множество согласно соответствию $\pounds$, определенному следующим образом:

$$\pounds(\lambda) = \lambda$$
$$\pounds(\mathbf{a}) = \{a\} \text{ для всех } a \in A$$
$$\pounds(\vee) = \cup$$
$$\pounds(*) = *$$
$$\pounds(() = ($$
$$\pounds()) =)$$

Поэтому

$\mathcal{L}(\mathbf{aa}^*)$	$= \{a\} \circ \{a\}^*$	$= \{a, aa, aaa, aaaa, aaaa, \ldots\};$
$\mathcal{L}(\mathbf{a(b \vee c)d})$	$= \{a\} \circ \{\{b\} \cup \{c\}\} \circ \{d\}$	$= \{abd, acd\};$
$\mathcal{L}((\mathbf{a} \vee \mathbf{b})^*)$	$= \{a \cup b\}^*$	$= \{\lambda, a, b, ab, ba, abb, aba, \ldots\} =$
		$=$ все строки, состоящие из 0 или более символов a и b;
$\mathcal{L}(\mathbf{ab^*c})$	$= \{a\} \circ \{b\}^* \circ \{c\}$	$= \{ac, abc, abbc, abbbc, \ldots\};$
$\mathcal{L}(\mathbf{a} \vee \mathbf{b} \vee \mathbf{c})$	$= \{a\} \cup \{b\} \cup \{c\}$	$= \{a, b, c\}.$

Образ регулярного выражения является регулярным языком. Регулярные языки можно определить следующим образом.

ОПРЕДЕЛЕНИЕ 17.6. Класс R называется ***регулярным языком***, если обладает следующими свойствами.

i) Пустое слово, $\lambda \in R$, и если $a \in A$, то $\{a\} \in R$.
ii) Если s_1 и $s_2 \in R$, то $s_1 \cup s_2$, $s_1 \circ s_2$, $s_1^* \in R$.
iii) Только множества, образованные согласно (i) и (ii), принадлежат R.

Хотя следующие утверждения здесь и не доказываются, но пересечение регулярных множеств является регулярным множеством и дополнение регулярного множества является регулярным множеством.

ОПРЕДЕЛЕНИЕ 17.7. Пусть S и T — подмножества множества A^*. Если $S = T^*$, то S называется множеством, ***порожденным*** множеством T.

Если множество S порождено множеством T, то каждый элемент из множества S является конечной конкатенацией элементов множества T. Однако, элементы конкатенации, образующие данное слово, необязательно являются однозначно определимыми. Если $T = \{a, b\}$, то T^* состоит из пустого слова и всех возможных конечных строк из символов a и b. Если $\overline{T} = \{a, ab, b\}$, то $T^* = \overline{T}^*$, но хотя каждая строка в T^* может выражаться единственным образом конкатенацией элементов из T, это неверно для элементов множества $\overline{T}$, поскольку выражение **abab** можно записать как **(a)(b)(a)(b)**, а также как **(a)(b)(ab)** и как **(ab)(ab)**.

ОПРЕДЕЛЕНИЕ 17.8. Пусть S и C — подмножества множества A^*. Если $S = C^*$, то каждая строка в S может выражаться конкатенацией элементов множества C, и будем говорить, что C является ***кодом***. Код C называется ***однозначно декодируемым,*** если каждую строку из S можно однозначно выразить конкатенацией элементов из C.

Поэтому $\{ba, ab, c\}$, $\{ade, ddbee, dfc, ddd\}$ и $\{a, b, c, de\}$ являются однозначно декодируемыми кодами, хотя $\{a, ab, bc, c\}$, $\{ab, abc, cde, de\}$ и $\{a, bc, abc\}$ однозначно декодируемыми не являются.

ПРИМЕР 17.9. Функция на множестве S называется ***ретракционным отображением***, или ***ретракцией***, если $f \circ f = f$. Образ при отображении f называется ***ретрактом***. Пусть R — ретракт при ретракционном отображении $f : A^* \to A^*$. Том Хэд (Tom Head) показал, что ретракт порождается конечным кодом, обладающим таким свойством: каждое слово кода содержит элемент из A, который встречается в этом слове только один раз и не встречается ни в каком другом слове кода. Такие коды называются ***ключевыми кодами***. Следовательно, $\{a, b, c\}$, $\{a, be, ce\}$, $\{ade, ddbee, caa\}$ и $\{ad, bdde, cede\}$ являются ключевыми кодами, порождающими ретракты, поскольку каждый из элементов a, b и c встречаются точно в одном слове и при этом только один раз. Однако, $\{ba, ab, c\}$, $\{ab, abc, cde\}$ и $\{a, ab, bc, d\}$ тоже являются кодами, но не порождают ретракты. Пересечение ретрактов необязательно является ретрактом. На усмотрение читателя оставим доказательство, что пересечение множеств $\{ac, bc, d\}^*$ и $\{a, cb, cd\}^*$ порождено кодом, который можно описать выражением **ac(bc)*d**. □

ОПРЕДЕЛЕНИЕ 17.10. Пусть A — алфавит. Код $C \subseteq A^*$ называется ***префиксным кодом***, если для всех слов $u, v \in C$; если $u = vw$ для $w \in A^*$, то $u = v$ и $w = \lambda$. Это означает, что одно слово кода не может быть начальной строкой другого слова кода. Код $C \subseteq A^*$ называется ***суффиксным кодом***, если для всех слов $u, v \in C$; если $u = wv$ для $w \in A^*$, то $u = v$ и $w = \lambda$. Это означает, что одно слово кода не может быть конечной строкой другого слова кода.

Множество $\{a, ab, ac\}$ является кодом, но не является префиксным кодом, так как a является начальной строкой как для ab, так и для ac. Однако, данное множество — суффиксный код. Множество $\{a, ba, ca\}$ является префиксным кодом, но суффиксным не является, поскольку a — конечная строка для ba и для ca. Множество $\{ad, ab, ac\}$ — одновременно префиксный и суффиксный код.

■ УПРАЖНЕНИЯ

1. Найдите регулярные множества, соответствующие приведенным ниже выражениям. Если множество бесконечное, перечислите десять его элементов.
   а) **a(b∨c∨d)a**;
   б) **a*b*c**;
   в) **(a∨b)(c∨d)**;
   г) **(ab*λ)∨(cd)***;
   д) **a(bc)*d**.
2. Найдите регулярные множества, соответствующие приведенным ниже выражениям. Если множество бесконечное, перечислите десять его элементов.
   а) **bc(bc)***;
   б) **(a∨b* ∨ λ)(c∨d*)**;

в) $(a \vee bc \vee d)^*$;

г) $(a \vee b)(c \vee d)b$;

д) $a^*(b \vee c \vee)d^*$.

3. Найдите регулярные выражения, соответствующие приведенным ниже регулярным множествам.

а) $\{ab, ac, ad\}$;

б) $\{ab, ac, bb, bc\}$;

в) $\{a, ab, abb, abbb, abbbb, \ldots\}$;

г) $\{ab, abab, ababab, abababab, ababababab, \ldots\}$;

д) $\{ab, abb, aab, aabb\}$.

4. Найдите регулярные выражения, соответствующие приведенным ниже регулярным множествам.

а) $\{ab, acb, adb\}$;

б) $\{ab, abb, abbb, abbbb, \ldots\}$;

в) $\{ad, ae, af, bd, be, bf, cd, ce, cf\}$;

г) $\{abcd, abcbcd, abcbcbcd, abcbcbcbcd, \ldots\}$;

д) $\{abcd, abef, cdcd, cdef\}$.

5. Пусть $A = \{a, b, c\}$.

а) Постройте регулярное выражение для множества всех элементов множества A^*, содержащих точно два символа b.

б) Постройте регулярное выражение для множества всех элементов множества A^*, содержащих точно два символа b и два символа c.

в) Постройте регулярное выражение для множества всех элементов множества A^*, содержащих точно два или более символов b.

г) Постройте регулярное выражение для множества всех элементов множества A^*, которые начинаются и оканчиваются символом a и содержат, по крайней мере, один символ b и один символ c.

д) Постройте регулярное выражение для множества всех элементов множества A^*, которые состоят из одного или более символов a, за которым следует один или более символов b, а затем один или более символов c.

6. Пусть $A = \{a, b\}$.

а) Постройте регулярное выражение для множества всех элементов множества A^*, содержащих в точности два символа b или в точности два символа a.

б) Постройте регулярное выражение для множества всех элементов множества A^*, содержащих четное число символов b.

в) Постройте регулярное выражение для множества всех элементов множества A^*, которые начинаются и оканчиваются символом a и содержат не менее одного символа b.

г) Постройте регулярное выражение для множества всех элементов множества A^* таких, что количество символов a в каждой строке кратно 3, а количество символов b кратно 5.

д) Постройте регулярное выражение для множества всех элементов множества A^* таких, что длина каждой строки кратна 3.

7. Какое из приведенных ниже выражений является однозначно декодируемым кодом?

а) $\{ab, ba, a, b\}$;

б) $\{ab, acb, accb, acccb, \ldots\}$;

в) $\{a, b, c, bd\}$;

г) $\{ab, ba, a\}$;

д) $\{a, ab, ac, ad\}$.

8. Какие из приведенных ниже выражений описывают однозначно декодируемые коды?

а) $\mathbf{ab^*}$;

б) $\mathbf{ab^* \vee baaa}$;

в) $\mathbf{ab^*c \vee baaac}$;

г) $\mathbf{(a \vee b)(b \vee a)}$;

д) $\mathbf{(a \vee b \vee \lambda)(b \vee a \vee \lambda)}$.

9. Какое из перечисленных ниже выражений является однозначно декодируемым кодом? Какое является суффиксным кодом?

а) $\{ab, ba, \}$;

б) $\{ab, abc, bc\}$;

в) $\{a, b, c, bd\}$;

г) $\{aba, ba, c\}$;

д) $\{ab, acb, accb, acccb\}$.

10. Какие из приведенных ниже выражений описывают префиксные коды? Какие из них описывают суффиксные коды?

а) $\mathbf{ab^*}$;

б) $\mathbf{ab^*c}$;

в) $\mathbf{a^*bc^*}$;

г) $\mathbf{(a \vee b)(b \vee a)}$;

д) $\mathbf{a^*b}$.

11. Какие из приведенных ниже кодов являются однозначно декодируемыми? Обоснуйте свой ответ.

а) 1110, 111, 1011, 11, 0101;

б) 10, 110, 1110, 0101;

в) 010, 101, 0101, 1010, 10101;

г) 0101, 1001, 011,11, 10011, 00101;

д) 101, 1001, 10001, 11011, 11001, 00111.

12. Какие из приведенных ниже кодов являются однозначно декодируемыми? Обоснуйте свой ответ.

а) 110, 1110, 11110, 111110, 10;

б) 1011, 1010, 1100, 0101, 0011, 1100, 1100, 1111, 0000;

в) 101, 010, 11, 10101, 1011, 1010;

г) 110, 101, 1001, 11101, 0101, 0110, 0111, 10001;
д) 11011, 101, 010100, 11, 1011, 1110.

13. Какие из приведенных ниже кодов являются префиксными кодами?
а) 101, 1101, 0101, 001, 1001, 11101, 0110, 01001;
б) 1101, 0011, 1010, 1110, 1010, 0101, 1100, 1111;
в) 101, 1100, 01010, 10101, 0011, 0110;
г) 1011, 10101, 10110, 0101, 01100, 111000;
д) 1101, 10001, 1001, 01110, 10110, 110110.

14. Какие коды в упражнении 1 являются префиксными кодами?
15. Какие коды в упражнении 2 являются префиксными кодами?
16. а) Какие коды в упражнении 1 являются суффиксными кодами?
б) Какие коды в упражнении 2 являются суффиксными кодами?
в) Какие коды в упражнении 3 являются суффиксными кодами?
17. Код называется ***инфиксным кодом***, если никакая строка кода не является подстрокой другой строки кода. Всегда ли суффиксный и префиксный коды являются инфиксными кодами? Докажите или постройте контрпример.
18. Всегда ли инфиксный код является префиксным и суффиксным кодом одновременно? Докажите или постройте контрпример.
19. Какие коды в упражнении 1 являются инфиксными кодами?
20. Какие коды в упражнении 2 являются инфиксными кодами?
21. Какие коды в упражнении 3 являются инфиксными кодами?
22. Существует ли префиксный код, состоящий из слов длины $1, 2, 3, 3$?
23. Существует ли префиксный код, состоящий из слов длины 1, 3, 3, 3, 4, 5, 5, 5?
24. Существует ли инфиксный код, состоящий из слов длины $1, 3, 3, 3$?
25. Докажите, что ключевой код является префиксным и суффиксным кодом одновременно.

17.2. АВТОМАТЫ

По сути, ***автомат*** представляет собой устройство, распознающее, воспринимающее или допускающее определенные элементы множества A^*, где A — конечный алфавит. Различные автоматы (совокупность автоматов) распознают или допускают различные элементы множества A^*. Подмножество элементов множества A^*, допускаемое автоматом M,— это язык, который называется ***языком*** над алфавитом A, ***допускаемым*** автоматом M, и обозначается $M(L)$.

Для заданного конечного алфавита автомат состоит из множества S — множества состояний и функции $F : A \times S \to S$, называемой ***функцией переходов***. Множество S содержит элемент s_0 и одно или более финальных состояний. Автомат M обозначается через (A, S, s_0, T, F), где A — алфавит, S — множество состояний, s_0 — начальное или исходное состояние, T — множество финальных состояний и F — функция перехода. Обратите внимание, что входом для F является буква из A и состояние, принадлежащее множеству S. Выходом является

состояние из S (возможно, то же самое состояние). Если автомат находится в состоянии s и "читает" букву a, то (a, s) является входом для F, и $F(a, s)$ — следующее состояние. Для автомата в качестве "начала" можно рассматривать начальное состояние s_0. Если автомат "читает" букву a алфавита A, то он "переходит" в состояние $s = F(a, s_0)$. Если теперь автомат "читает" букву из A, например, b, то он переходит в состояние $s' = F(a, s)$. Следовательно, по мере "прочтения" букв алфавита автомат переходит из одного состояния в другое. Пусть M — автомат с алфавитом $A = \{a, b\}$, множеством состояний $S = \{s_0, s_1, s_2\}$, а функция F задана таблицей

F	s_0	s_1	s_2
a	s_1	s_1	s_2
b	s_2	s_2	s_1

Предположим, что M "читает" букву a, за которой следуют буквы b и a. Поскольку автомат начинает функционировать в состоянии s_0, и буква, которую он читает, это a, то $F(a, s_0) = s_1$, поэтому теперь автомат находится в состоянии s_1. Следующей буквой для чтения является b и $F(b, s_1) = s_2$. Наконец, читается последняя буква a, и так как $F(a, s_2) = s_2$, автомат остается в состоянии s_2. Функцию перехода F можно задать также совокупностью правил.

Если автомат находится в состоянии s_0 и читает a , то он переходит в состояние s_1.
Если автомат находится в состоянии s_1 и читает a , то он переходит в состояние s_1.
Если автомат находится в состоянии s_2 и читает a , то он переходит в состояние s_2.
Если автомат находится в состоянии s_0 и читает b, то он переходит в состояние s_2.
Если автомат находится в состоянии s_1 и читает b, то он переходит в состояние s_2.
Если автомат находится в состоянии s_2 и читает b, то он переходит в состояние s_1.

Такой автомат лучше всего изображать графически с помощью ***диаграммы состояний***, являющейся ориентированным графом, у которого состояния представлены вершинами, а ориентированное ребро из вершины s в вершину s' помечено буквой из алфавита A, например, буквой a, если $F(a, s) = s'$. Направленная стрелка, помеченная буквой a, будет называться ***a-стрелкой***. Поэтому приведенный выше автомат можно представить графически, как показано на рис. 17.1.

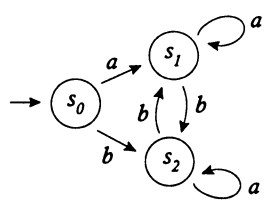


Рис. 17.1

(Отметим, что хотя диаграмма состояний на рис. 17.1 и является ориентированным графом, тем не менее, исходя из условий задания некоторых автоматов, соответсвующие диаграммы могут не быть ориентированными графами, поскольку из одного состояния в другое может вести не одно ориентированное ребро.

Более точным названием таких диаграмм является ***мультиграфы***, или ***помеченные орграфы***.

Слово или строку $a_0a_1a_2\ldots a_n$ из A^* автомат "читает" следующим образом: сначала читается a_0, затем a_1 и так до тех пор, пока не будет прочитано a_n. Как отмечалось выше, некоторые состояния определены как финальные состояния. Автомат ***допускает*** или ***распознает*** строку, $a_0a_1a_2\ldots a_n$, если после прочтения a_n он останавливается в финальном состоянии. Если s — начальное состояние, то соответствующая вершина обозначается схемой, изображенной на рис. 17.2. Если s — финальное состояние, то соответствующая вершина обозначается, как изображена на рис. 17.3.

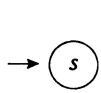


Рис. 17.2



Рис. 17.3

Например, для автомата, диаграмма состояний которого изображена на рис. 17.4, состояние s_0 — начальное, а состояние s_3 — финальное.

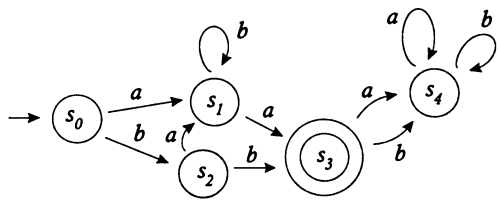


Рис. 17.4

Автомат допускает слово baa, поскольку после прочтения b он переходит в состояние s_2; после прочтения a — в состояние s_1; после прочтения второго a, он переходит в состояние s_3, которое является финальным состоянием. Можно убедиться, что автомат допускает также слова $abbba$ и bb, но не распознает слова bbb, $abab$ или abb. Как отмечалось выше, множество слов, допускаемых автоматом, называется языком, допускаемым автоматом.

ПРИМЕР 17.11. Рассмотрим автомат с диаграммой состояний, изображенной на рис. 17.5.

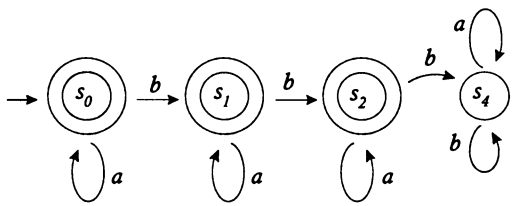


Рис. 17.5

Очевидно, что автомат допускает слово bb. Для слова a в каждом состоянии имеется петля, поэтому при чтении a состояние не меняется, что дает возможность до прочтения следующего слова b читать любое необходимое количество

слов a. Таким образом, автомат читает $aababaaa$, $baabab$, $baaab$, $babaaa$, $aabaabaa$ и может фактически прочесть любое слово языка, заданного регулярным выражением $\mathbf{a^*ba^*ba^*}$. Поскольку $A = \{a, b\}$, то такой язык можно описать также как множество всех слов, содержащих точно два b. Состояние s_4 является примером ***состояния зацикливания***. Попав в состояние зацикливания, автомат никогда не сможет из него выйти. □

ПРИМЕР 17.12. Рассмотрим автомат с диаграммой состояния, изображенной на рис. 17.6 (упрощенный вид — на рис. 17.7).

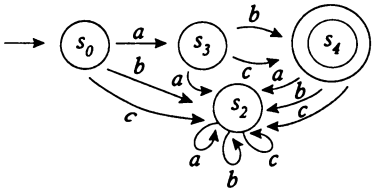


Рис. 17.6

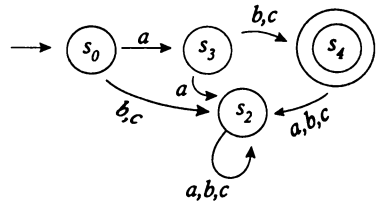


Рис. 17.7

Автомат допускает, очевидно, только слова ab и ac. Этот язык можно задать с помощью регулярного выражения $\mathbf{a(b \vee c)}$. □

ПРИМЕР 17.13. Рассмотрим автомат с диаграммой состояний, изображенной на рис. 17.8.

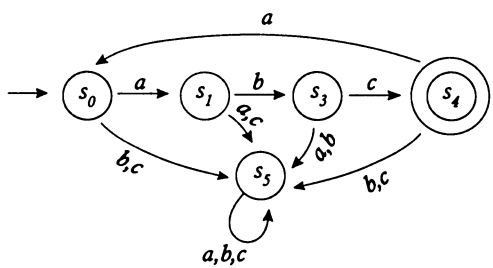


Рис. 17.8

Каждое слово следует начинать с ab и заканчивать на c. Однако петлю $aabc$ можно повторять требуемое количество раз, поскольку она начинается и заканчивается в s_3. Поэтому регулярным выражением для этого автомата будет $\mathbf{abc(aabc)^*}$. □

ПРИМЕР 17.14. Рассмотрим автомат с диаграммой состояний, изображенной на рис. 17.9.

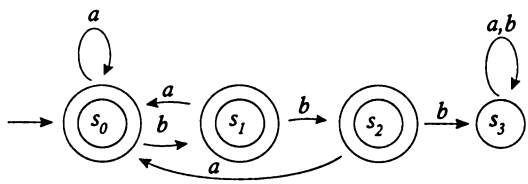


Рис. 17.9

Автомат при чтении трех последовательных b переходит в состояние s_3, которое является состоянием зацикливания. Это единственный способ попасть в состояние s_3, и все другие состояния являются финальными состояниями. Таким образом, язык, который допускает автомат, состоит из всх слов, не включающих три последовательных b. □

Рассмотренные автоматы часто называются ***детерминированными автоматами,*** так как в любом состоянии и для любого значения из алфавита, которое подается в автомат для чтения, существует одно и только одно состояние. Иными словами, такая ситуация обусловлена тем, что $F : A \times S \to S$ является функцией. Зачастую удобно ослабить правила таким образом, чтобы F была не функцией, а отношением. Если опять рассмотреть F как множество правил для $a \in A$ и $s \in S$, то могут существовать правила, предоставляющие возможность перехода из данного состояния в некоторые другие состояния, или такие правила не могут существовать. В последнем случае автомат, как говорят, "зависает" и не может далее функционировать. Автоматы, для которых F не обязательно является функцией, называются ***недетерминированными автоматами***. Рассмотрим, например, приведенную ниже диаграмму состояний, изображенную на рис. 17.10.

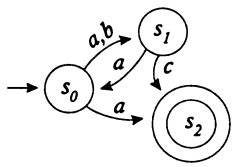


Рис. 17.10

Если в состоянии s_0 прочитывется слово, символ a, то автомат переходит в состояние s_1 или s_2. Однако, если автомат находится в состоянии s_1 и прочитывается слово, символ b, то автомат "зависает", поскольку нет b-стрелки, выходящей из состояния s_1. Отметим, что в соответствии с приведенным выше определением множество детерминированных автоматов принаждлежит множеству недетерминированных автоматов.

ПРИМЕР 17.15. Легко убедиться, что автомат с диаграммой состояний, изображенной на рис. 17.11, допускает язык, задаваемый регулярным выражением **ab*c**.

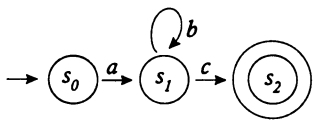


Рис. 17.11 □

ПРИМЕР 17.16. Автомат с диаграммой состояний, изображенной на рис. 17.12, допускает язык, задаваемый регулярным выражением **a∨b**.

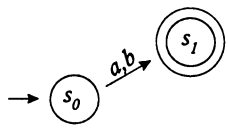


Рис. 17.12 □

ПРИМЕР 17.17. Автомат с диаграммой состояний, изображенной на рис. 17.13, допускает язык, задаваемый регулярным выражением **aa*bb***.

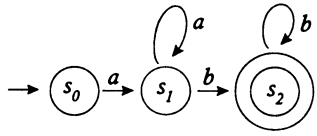


Рис. 17.13 □

ПРИМЕР 17.18. Автомат с диаграммой состояний, изображенной на рис. 17.14, допускает язык, состоящий их строк, содержащих не менее двух символов a.

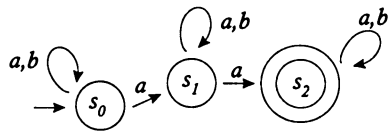


Рис. 17.14 □

Преимущество недетерминированных автоматов состоит в том, что нет необходимости включать в них несущественные ребра. Поэтому в некоторых дальнейших построениях будет намного проще использовать недетерминированные автоматы. Любой язык, который допускается детерминированным автоматом, очевидно, допускается и недетерминированным автоматом, так как множество недетерминированных автоматов включает в себя множество детерминированных автоматов. Следующая теорема показывает, что любой язык, который допускается недетерминированным автоматом, допускается и детерминированным автоматом.

ТЕОРЕМА 17.19. Для каждого недетерминированного автомата существует эквивалентный детерминированный автомат, который допускает тот же язык.

Вместо формального доказательства продемонстрируем построение детерминированного автомата, который допускает язык, допускаемый недетерминированным автоматом. Пусть S — множество состояний недетерминированного автомата. Используем элементы $\mathcal{P}(S)$, т.е. элементы подмножества множества S в качестве состояний для создаваемого детерминированного автомата. Рассмотрим недетерминированный автомат N, изображенный на рис. 17.15.

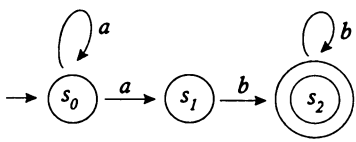


Рис. 17.15

Пусть $\{s_0\}$ — начальное состояние. Построим a-стрелку из $\{s_0\}$ во множество всех таких состояний, что существует a-стрелка из s_0 в это состояние. В данном случае имеется a-стрелка из s_0 в s_0 и a-стрелка из s_0 в s_1, т.е. построена a-стрелка из $\{s_0\}$ в $\{s_0, s_1\}$. Не существует b-стрелки из s_0 в любое другое состояние, поэтому строим b-стрелку из $\{s_0\}$ в пустое множество $\varnothing$. Теперь рассмотрим состояние $\{s_0, s_1\}$. Строим a-стрелку из $\{s_0, s_1\}$ во множество всех таких состояний, что существует a-стрелка либо из s_0, либо из s_1 в это состояние. В данном случае строим a-стрелку из $\{s_0, s_1\}$ в самое себя. Строим b-стрелку из

$\{s_0, s_1\}$ во множество всех таких состояний, что существует b-стрелка из s_0 или из s_1 в это состояние. В данном случае строим b-стрелку из $\{s_0, s_1\}$ в $\{s_2\}$. Поскольку не существует ни a-стрелки, ни b-стрелки из любого состояния в пустом множестве ни в какое другое состояние, строим a-стрелку и b-стрелку из пустого множества в себя. Теперь рассмотрим $\{s_2\}$. Поскольку не существует a-стрелки из s_2 ни в какое другое состояние, строим a-стрелку из $\{s_2\}$ в пустое множество. Учитывая, что только b-стрелка ведет из s_2 в себя, строим b-стрелку из $\{s_2\}$ в себя. Финальное состояние (состояния) состоит из всех множеств, содержащих элемент финального множества автомата N. В данном случае $\{s_2\}$ — единственное финальное состояние. Итак, построена диаграмма состояний, изображенная на рис. 17.16, которая является диаграммой состояний детерминированного автомата. Этот автомат допускает тот же язык, что и автомат N, а именно, язык, заданный выражением **a*abb***.

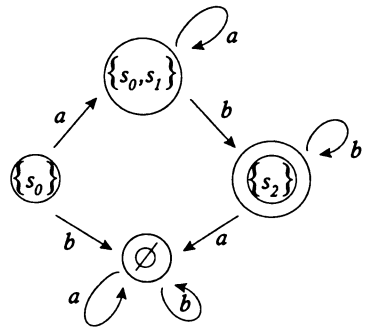


Рис. 17.16

В общем случае имеется следующая процедура для построения детерминированного автоматата из недетерминированного.

1. Начать из состояния $\{s_0\}$, где s_0 — начальное состояние недетерминированного автомата.
2. Для каждого $a_i \in A$ построить a_i-стрелку из $\{s_0\}$ во множество всех состояний таким образом, что имеется a_i-стрелка из s_0 в это состояние.
3. Для каждого вновь построенного состояния S_j и для каждого $a_i \in A$ построить a_i-стрелку из S_j во множество всех таких состояний, что существует a_i-стрелка из элемента множества S_j в это состояние.
4. Продолжать процесс, пока есть возможность создавать новые состояния.
5. Преобразовать все множества состояний S_j, содержащие элемент финального множества недетерминированного автомата, в финальное состояние.

ПРИМЕР 17.20. Построим детерминированный автомат, соответствующий недетерминированному автомату, изображенному на рис. 17.17.

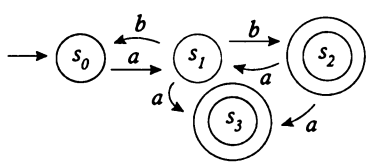


Рис. 17.17

Поскольку имеется a-стрелка из s_0 в s_1, строим a-стрелку из $\{s_0\}$ в $\{s_1\}$. Так как не существует b-стрелки из s_0 в любое иное состояние, строим b-стрелку из s_0 в пустое множество $\varnothing$. Имеется a-стрелка из s_1 в s_3, поэтому строим a-стрелку из $\{s_1\}$ в $\{s_3\}$. Поскольку имеется b-стрелка из s_1 в s_0 и из s_1 в s_2, строим b-стрелку из $\{s_1\}$ в $\{s_0, s_2\}$. Стрелок из s_3 в другое состояние не существует, поэтому строим a-стрелку и b-стрелку из s_3 в пустое множество. Имеются a-стрелки из s_0 в s_1, из s_2 в s_1 и из s_2 в s_3, поэтому строим a-стрелку из $\{s_0, s_2\}$ в $\{s_1, s_3\}$. Не существует b-стрелок ни из состояния s_0, ни из состояния s_2, поэтому строим b-стрелку из $\{s_0, s_2\}$ в пустое множество. Так как имеется a-стрелка из s_1 в s_3, строим a-стрелку из $\{s_1, s_3\}$ в $\{s_3\}$. Имеются b-стрелки из s_1 в s_0 и из s_1 в s_2, поэтому строим b-стрелку из $\{s_1, s_3\}$ в $\{s_0, s_2\}$. Этим завершается построение детерминированного автомата, изображенного на рис. 17.18.

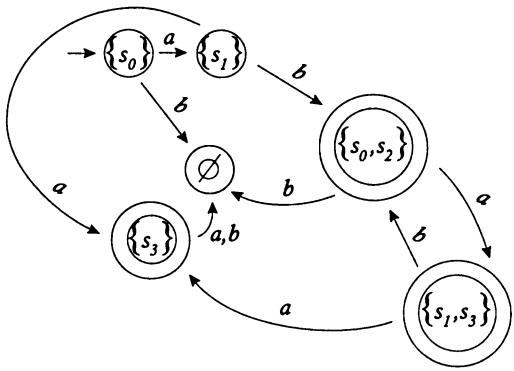


Рис. 17.18 □

Проведенное нами исследование регулярных языков и автоматов подытоживает следующая теорема, которая приводится без доказательства.

ТЕОРЕМА 17.21. (Клини) Язык является регулярным тогда и только тогда, когда он допускается автоматом.

■ УПРАЖНЕНИЯ

1. Какие из перечисленных ниже слов допускаются автоматом, изображенным на рис. 17.19?
   а) *abba*
   б) *aabbb*
   в) *babab*
   г) *aaabbb*
   д) *bbaab*

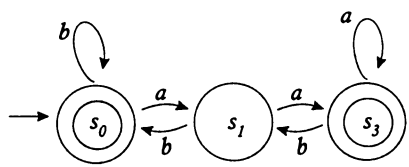


Рис. 17.19

2. Какие из перечисленных ниже слов допускаются автоматом, изображенным на рис. 17.20?

а) *aaabb*

б) *abbbabbb*

в) *bababa*

г) *aaabab*

д) *bbbabab*

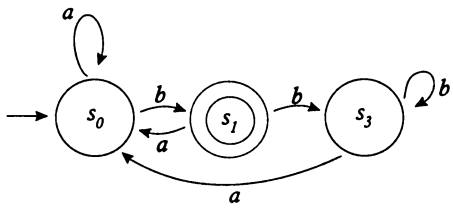


Рис. 17.20

3. Составьте выражение для языка, который допускается автоматом, изображенным на рис. 17.21.

4. Составьте выражение для языка, который допускается автоматом, изображенным на рис. 17.22.

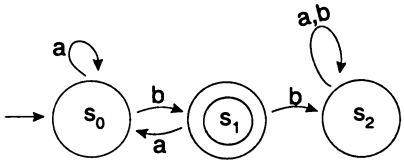


Рис. 17.21

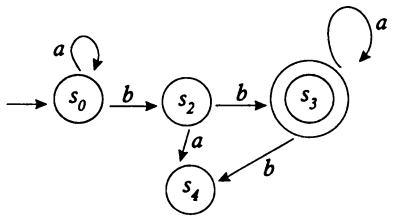


Рис. 17.22

5. Составьте выражение для языка, который допускается автоматом, изображенным на рис. 17.23.

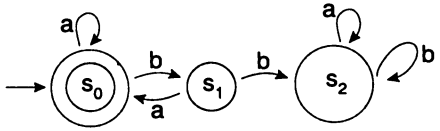


Рис. 17.23

6. Составьте выражение для языка, который допускается автоматом, изображенным на рис. 17.24.

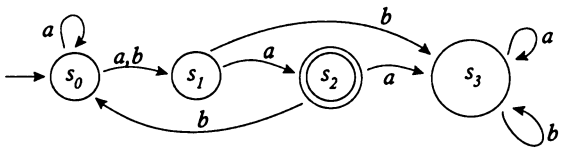


Рис. 17.24

7. Найдите детерминированный автомат, который допускает язык, заданный выражением **aa*bb*cc***.

8. Найдите детерминированный автомат, который допускает язык, заданный выражением **(a*ba*ba*b)***.
9. Найдите детерминированный автомат, который допускает язык, заданный выражением **(a*(ba)*bb*a)***.
10. Найдите детерминированный автомат, который допускает язык, заданный выражением **(a*b)∨(b*a)***.
11. Найдите недетерминированный автомат, который допускает язык, заданный выражением **aa*bb*cc***.
12. Найдите недетерминированный автомат, который допускает язык, заданный выражением **(a*b)∨(c*b)∨(ac)***.
13. Найдите недетерминированный автомат, который допускает язык, заданный выражением **(a∨b)*(aa∨bb).(a∨b)***
14. Найдите недетерминированный автомат, который допускает язык, заданный выражением **((aa*b)∨bb*a)ac***
15. Найдите детерминированный автомат, который допускает тот же язык, что и недетерминированный автомат, изображенный на рис. 17.25.

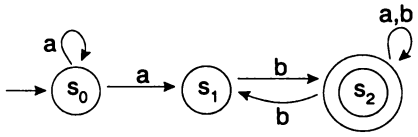


Рис. 17.25

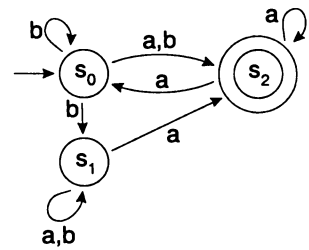


Рис. 17.26

16. Найдите детерминированный автомат, который допускает тот же язык, что и недетерминированный автомат, изображенный на рис. 17.26.
17. Найдите детерминированный автомат, который допускает тот же язык, что и недетерминированный автомат, изображенный на рис. 17.27.

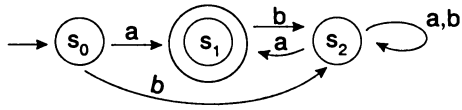


Рис. 17.27

18. Найдите детерминированный автомат, который допускает тот же язык, что и недетерминированный автомат, изображенный на рис. 17.28.

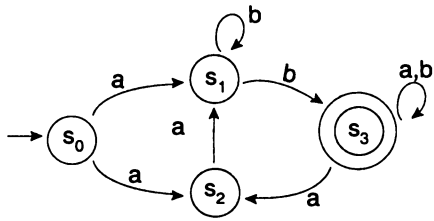


Рис. 17.28

17.3. ГРАММАТИКИ

Грамматика — это совокупность правил, которые используются для построения языка. Правила позволяют заменять символы или строки символов другими символами или строками символов. Допустим, имеется начальное слово *add*, которое можно заменить на $A + B$, а A и B можно заменить любым неотрицательным целым числом, меньшим 10. Пользуясь этим правилом, A можно заменить на 3 и B — на 2, получив $3 + 2$. Если A заменить на 5, а B — на 6, то получим $5 + 6$.

При вводе дополнительного правила о возможности замены A на $A + B$ появится возможность начать с замены *add* на $A + B$. Если после этого заменить A на $A+B$, получим $A+B+B$. Заметим следующее: поскольку A можно заменить как на $A + B$, так и на целое число, то при замене символов однозначности не требуется. Кроме того, и одно, и второе B в выражении $A+B+B$ не обязательно заменять одним и тем же значением. Если A заменить на 3, первое B заменить на 5, а второе B — на 7, то получим $3 + 5 + 7$. С другой стороны, можно снова было бы заменить A на $A + B$, получить $A + B + B + B$ и продолжать строить суммы произвольной длины.

Отметим, что в приведенных выше правилах символы *add*, A и B можно заменять другими символами, в то время как символ + и целые числа заменять нельзя. Символы, которые можно заменять другими символами, называются ***нетерминальными символами***, а символы, которые нельзя заменять другими символами, называются ***терминальными символами***. Правила, указывающие, как заменять символы, называются ***продукциями***. Обозначим продукцию (или правило), смысл которой состоит в том, что *add* можно заменить на $A + B$, через

$$add \rightarrow A + B.$$

Таким образом, продукция для первого из приведенных выше примеров имеет вид:

$$\begin{array}{ll} add \rightarrow A + B & \\ A \rightarrow 0 & B \rightarrow 0 \\ A \rightarrow 1 & B \rightarrow 1 \\ A \rightarrow 2 & B \rightarrow 2 \\ \vdots & \vdots \\ A \rightarrow 9 & B \rightarrow 9 \end{array}$$

Если добавить правило, состоящее в том, что A можно заменить на $A + B$, то продукции будут иметь вид:

$$\begin{array}{ll} add \rightarrow A + B & A \rightarrow A + B \\ A \rightarrow 0 & B \rightarrow 0 \\ A \rightarrow 1 & B \rightarrow 1 \\ A \rightarrow 2 & B \rightarrow 2 \\ \vdots & \vdots \\ A \rightarrow 9 & B \rightarrow 9 \end{array}$$

Формально грамматика определяется следующим образом.

ОПРЕДЕЛЕНИЕ 17.22. ***Формальная грамматика***, или ***грамматика непосредственных составляющих*** состоит из конечного множества ***нетерминальных символов*** N, конечного множества ***терминальных символов*** T, элемента $S \in N$, называемого ***начальным символом***, и конечного множества продукций P, которое является отношением в $(N \cup T)^*$, таким что каждый первый элемент упорядоченной пары из P содержит символ из N и, по крайней мере, один первый элемент некоторого правила является начальным элементом S. Если W и W' — элементы из $(N \cup T)^*$, $W = uvw$, $W' = uv'w$, а $v \to v'$ — продукция, то непосредственный вывод W' из W обозначается $W \Rightarrow W'$. Если

$$W_1 \Rightarrow W_2 \Rightarrow W_3 \Rightarrow \cdots \Rightarrow W_n$$

для $n \geq 0$, то говорят, что W_n **выводится** из W_1. Обозначается это следующим образом: $W_1 \Rightarrow_n^* W_n$. Формальная грамматика Γ обозначается четверкой (N, T, S, P). Множество всех строк элементов из T, которые можно породить с помощью множества продукций P, называется ***языком, порождаемым грамматикой*** Γ, и обозначается $\Gamma(L)$.

При порождении слова в грамматике Γ для вывода новых строк продолжаем использовать продукции до тех пор, пока не получим строку, содержащую только терминальные элементы.

Таким образом, для последнего из приведенных выше примеров имеем:

$$N = \{add, A, B\},$$

$$T = \{+, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\},$$

$$S = add$$

и

$$P = \{(add, A{+}B), (A, A{+}B), (A, 0), (A, 1), \ldots, (A, 9), (B, 0), (B, 1), \ldots, (B, 9)\},$$

где $(add, A + B)$ обозначено через $add \to A + B$, $(A, A + B)$ обозначено через $A \to A + B$ и т.д. Язык, порожденный грамматикой Γ, — это множество всех формальных выражений конечных сумм неотрицательных целых чисел, меньших, чем 10.

ПРИМЕР 17.23. В описанной выше грамматике выведем выражение $3 + 2 + 4$.

Начнем с продукции

$$add \to A + B,$$

чтобы вывести

$$A + B.$$

Затем применим продукцию

$$A \to A + B,$$

чтобы вывести

$$A + B + B.$$

Далее применим продукции

$$A \to 3 \qquad B \to 2 \qquad B \to 4,$$

чтобы вывести

$$+2+4. \qquad \square$$

ПРИМЕР 17.24. Пусть требуется построить грамматику, позволяющую вывести арифметические выражения для множества целых чисел $\{0,1,2,3,4,5,6,7,8,9\}$. Язык, порожденный этой грамматикой, есть множество всех конечных арифметических выражений для множества целых чисел $\{0,1,2,3,4,5,6,7,8,9\}$. Примерами служат выражения $3 \times (5+4)$ и $(4+5) \div (3\hat{}2)$, где символ $\hat{}$ обозначает показатель степени. Понятно, что желательно исключить такие выражения, как $3 + \times 6$ и $3 + \div 6 \times 4 - 5$. Пусть множество $N = \{S, A, B\}$ и $T = \{+, -, \times, \div, \hat{}, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, (,)\}$. Нам понадобятся следующие продукции:

$$\begin{array}{lcl} S \to (A+B) \\ S \to (A-B) \\ S \to (A \times B) \\ S \to (A \div B) \\ S \to (A\hat{}B) \\ A \to (A+B) \\ A \to (A-B) \\ A \to (A \times B) \\ A \to (A \div B) \\ A \to (A\hat{}B) \end{array} \qquad \begin{array}{lcl} B \to (A+B) \\ B \to (A-B) \\ B \to (A \times B) \\ B \to (A \div B) \\ B \to (A\hat{}B) \\ A \to 0 \\ \vdots \\ A \to 9 \\ B \to 0 \\ \vdots \\ B \to 9 \end{array}$$

Используем грамматику для вывода арифметического выражения $((2+3)\times(4+5))$. Начнем с продукции

$$S \to (A \times B).$$

Затем применим продукции

$$A \to (A+B)$$

и

$$B \to (A+B),$$

чтобы получить

$$((A+B) \times (A+B)).$$

Продукции

$$A \to 2 \qquad \text{и} \qquad B \to 3$$

дают

$$((2+3) \times (A+B)).$$

Наконец, применим продукции

$$A \to 4 \qquad \text{и} \qquad B \to 5,$$

чтобы вывести

$$((2+3)\times(4+5)).$$

Теперь применим грамматику для вывода арифметического выражения $((3\hat{}2)\div(5+7)$. Начнем с продукции

$$S\to(A\div B).$$

Затем применим продукции

$$A\to(A\hat{}B)\qquad\text{и}\qquad B\to(A+B),$$

чтобы получить

$$((A\hat{}B)\div(A+B)).$$

Продукции

$$A\to 3\qquad\text{и}\qquad B\to 2$$

дают

$$((3\hat{}2)\div(A+B)).$$

Наконец, применяя продукции

$$A\to 5\qquad\text{и}\qquad B\to 7,$$

выводим (

$$(3\hat{}2)\div(5+7)).$$

□

ПРИМЕР 17.25. Аналогичным образом формируются арифметические выражения в постфиксной форме. Пусть множество $N=\{S,A,B\}$ и

$$T=\{+,-,\times,/,\hat{},0,1,2,3,4,5,6,7,8,9\}.$$

Нам потребуются следующие продукции:

$$\begin{array}{cccc}
S\to AB+ & A\to AB+ & B\to AB+ & A\to 0\\
S\to AB- & A\to AB- & B\to AB- & \vdots\\
S\to AB\times & A\to AB\times & B\to AB\times & A\to 9\\
S\to AB\div & A\to AB\div & B\to AB\div & B\to 0\\
S\to AB\hat{} & A\to AB\hat{} & B\to AB\hat{} & \vdots\\
 & & & B\to 9
\end{array}$$

Рассмотрим выражение $32+47+\times$. Поскольку все рассматриваемые целые числа меньше, чем 10, то 32+ представляет собой символ целого числа 3, за которым следует символ целого числа 2 и символ +. Для построения этого выражения начнем с продукции

$$S\to AB\times\ .$$

Затем воспользуемся продукциями

$$A\to A+B\qquad\text{и}\qquad B\to A+B,$$

чтобы вывести

$$AB + AB + \times .$$

Продукции

$$A \to 2 \qquad \text{и} \qquad B \to 3$$

дают

$$23 + AB + \times .$$

Наконец, применяя продукцию

$$A \to 4 \qquad \text{и} \qquad B \to 7 ,$$

выводим

$$3 + 47 + \times . \qquad \square$$

ПРИМЕР 17.26. Грамматику можно также использовать для вывода истинных выражений. Такие выражения истинны в том смысле, что они грамматически правильны, хотя даже лишены смысла. Предположим, что нужно построить грамматику, которая наряду с другими выводит следующие утверждения:

Джо преследует Фреда.
Большой пес прыгает через старый забор.
Том исчезал медленно за горизонт.

Перед началом фактического построения грамматики выберем ее структуру. В каждом из рассматриваемых предложений имеется именная группа (им. гр.), глагольная группа (гл. гр.) и еще одна именная группа. Кроме того, два последних предложения содержат предлог (предл.). Поэтому, пусть первой продукцией будет

$$S \to \langle \text{им. гр.} \rangle \langle \text{гл. гр.} \rangle \langle \text{предл.} \rangle \langle \text{им. гр.} \rangle .$$

В нашем примере наиболее общей формой именной группы является прилагательное (прил.), за которым следует существительное (сущ.). Следующим правилом будет

$$\langle \text{им. гр.} \rangle \to \langle \text{прил.} \rangle \langle \text{сущ.} \rangle .$$

Наиболее общей глагольной формой является глагол (гл.), за которым следует наречие (нар.). Поэтому пусть следующей продукцией будет

$$\langle \text{гл. гр.} \rangle \to \langle \text{нар.} \rangle \langle \text{гл.} \rangle .$$

Заранее известно, что финальное множество имеет вид $T = \{$Джо, преследует, Фреда, большой, пес, прыгает, через, старый, забор, Том, исчезал, медленно, за, горизонт$\}$. Нетерминальное множество имеет вид $N = \{S$, $\langle$им. гр.$\rangle$, $\langle$гл. гр.$\rangle$, $\langle$прил.$\rangle$, $\langle$сущ.$\rangle$, $\langle$нар.$\rangle$, $\langle$гл.$\rangle$, $\langle$предл.$\rangle\}$.

Теперь понадобятся продукции, приписывающие значения для $\langle$прил.$\rangle$, $\langle$сущ.$\rangle$, $\langle$нар.$\rangle$, и $\langle$гл.$\rangle$. В некоторых из приведенных предложений не нужны $\langle$прил.$\rangle$, $\langle$предл.$\rangle$ и $\langle$нар.$\rangle$. Для решения этой задачи введем продукции

$$\langle \text{прил.} \rangle \to \lambda \qquad \langle \text{нар.} \rangle \to \lambda \qquad \langle \text{предл.} \rangle \to \lambda .$$

Поставив в соответствие этим символам пустое множество, мы можем вычеркивать их за ненадобностью. Оставшаяся часть продукций включает следующее:

⟨сущ.⟩ → Фред	⟨сущ. ⟩ → забор	⟨нар.⟩ → медленно
⟨прил.⟩ → большой	⟨сущ.⟩ → пес	⟨гл.⟩ → преследует
⟨прил.⟩ → старый	⟨сущ.⟩ → Том	⟨гл.⟩ → прыгает
⟨сущ.⟩ → Джо	⟨сущ.⟩ → горизонт	⟨предл.⟩ → за
⟨гл.⟩ → исчезает	⟨предл.⟩ → через	

Для вывода предложения "Джо преследует Фреда" начнем с

$$S \rightarrow \langle\text{им. гр.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle\,,$$

чтобы вывести

$$\langle\text{им. гр.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle\,.$$

Используя продукцию

$$\langle\text{им. гр.}\rangle \rightarrow \langle\text{прил.}\rangle\langle\text{сущ.}\rangle\,,$$

выводим

$$\langle\text{прил.}\rangle\langle\text{сущ.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle\,.$$

Применив

$$\langle\text{прил.}\rangle \rightarrow \lambda\,,$$

выводим

$$\langle\text{сущ.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle\,.$$

Повторяя процесс для второй именной группы ⟨им. гр.⟩, выводим

$$\langle\text{сущ.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{сущ.}\rangle\,.$$

Применив

$$\langle\text{гл. гр.}\rangle \rightarrow \langle\text{нар.}\rangle\langle\text{гл.}\rangle\,,$$

выводим

$$\langle\text{сущ.}\rangle\langle\text{нар.}\rangle\langle\text{гл.}\rangle\langle\text{предл.}\rangle\langle\text{сущ.}\rangle\,.$$

Применив

$$\langle\text{предл.}\rangle \rightarrow \lambda\,,$$

выводим

$$\langle\text{сущ.}\rangle\langle\text{гл.}\rangle\langle\text{сущ.}\rangle\,.$$

Воспользовавшись

$$\langle\text{сущ.}\rangle \rightarrow \text{Джо} \qquad \langle\text{сущ.}\rangle \rightarrow \text{Фред} \qquad \langle\text{гл.}\rangle \rightarrow \text{преследует}\,,$$

выводим

$$\text{Джо преследует Фреда}\,.$$

Выведем предложение: "Большой пес прыгает через старый забор". Как и в предыдущем случае, начнем с продукции

$$S \to \langle\text{им. гр.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle,$$

чтобы вывести

$$\langle\text{им. гр.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle.$$

Применив продукцию

$$\langle\text{им. гр.}\rangle \to \langle\text{прил.}\rangle\langle\text{сущ.}\rangle,$$

выводим

$$\langle\text{прил.}\rangle\langle\text{сущ.}\rangle\langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle.$$

Используя

$$\langle\text{прил.}\rangle \to \text{большой} \qquad \langle\text{сущ.}\rangle \to \text{пес},$$

выводим

$$\text{Большой пес } \langle\text{гл. гр.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle.$$

Применив

$$\langle\text{гл. гр.}\rangle \to \langle\text{нар.}\rangle\langle\text{гл.}\rangle,$$

выводим

$$\text{Большой пес}\langle\text{нар.}\rangle\langle\text{гл.}\rangle\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle.$$

Используя

$$\langle\text{нар.}\rangle \to \lambda \qquad \langle\text{гл.}\rangle \to \text{прыгает},$$

выводим

$$\text{Большой пес прыгает}\langle\text{предл.}\rangle\langle\text{им. гр.}\rangle.$$

Применив

$$\langle\text{предл.}\rangle \to \text{через},$$

выводим

$$\text{Большой пес прыгает через}\langle\text{им. гр.}\rangle.$$

Используя продукцию

$$\langle\text{им. гр.}\rangle \to \langle\text{прил.}\rangle\langle\text{сущ.}\rangle,$$

получаем

$$\text{Большой пес прыгает через}\langle\text{прил.}\rangle\langle\text{сущ.}\rangle.$$

Применив

$$\langle\text{прил.}\rangle \to \text{старый} \qquad \langle\text{сущ.}\rangle \to \text{забор},$$

выводим

$$\text{Большой пес прыгает через старый забор}.$$

Вывод последнего предложения предоставляем читателю. □

ОПРЕДЕЛЕНИЕ 17.27. На рис. 17.29 изображено ***соответствующее дерево*** для продукции $P \to w_1 w_2 w_3 \ldots w_n$.

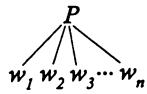


Рис. 17.29

Таким образом, соответствующим деревом для $S \to A + B$ является дерево, изображенное на рис. 17.30.



Рис. 17.30

ОПРЕДЕЛЕНИЕ 17.28. Если соответствующие деревья продукций, использованных для вывода данного выражения, связны, то они формируют дерево с корнем S, называемое ***деревом грамматического разбора*** или ***деревом вывода***. Листья дерева при чтении слева направо формируют выражение.

ПРИМЕР 17.29. В примере 17.23 продукции использовались для вывода $3+2+4$.

Для построения дерева начнем с первой использованной продукции

$$add \to A + B,$$

чтобы сформировать соответствующее дерево, изображеное на рис. 17.31.



Рис. 17.31



Рис. 17.32

Затем воспользуемся деревом, изображенным на рис. 17.32, которое соответствует продукции

$$A \to A + B,$$

чтобы сформировать дерево, изображенное на рис. 17.33.

Далее воспользуемся деревьями, соответствующими следующим продукциям

$$A \to 3, \qquad B \to 2, \qquad B \to 4,$$

чтобы сформировать дерево грамматического разбора, изображенное на рис. 17.34.

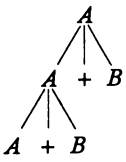


Рис. 17.33

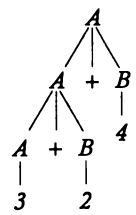


Рис. 17.34

□

ПРИМЕР 17.30. В примере 17.25 для вывода выражения $((2+3)\times(4+5))$ использовались продукции

$$\begin{array}{llll} S\to (A\times B), & A\to (A+B), & A\to 2, & B\to 3,\\ B\to (A+B), & A\to 4, & B\to 5. & \end{array}$$

Поэтому дерево грамматического разбора — это дерево, изображенное на рис. 17.35.

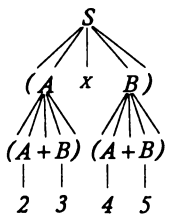


Рис. 17.35 □

ПРИМЕР 17.31. В примере 17.26 для вывода выражения "Большой пес прыгает через старый забор" использовались продукции

$S \to$ ⟨им. гр.⟩⟨гл. гр.⟩⟨предл.⟩⟨им. гр.⟩ ⟨им. гр.⟩ → ⟨прил.⟩⟨сущ.⟩
⟨им. гр.⟩ → ⟨прил.⟩⟨сущ.⟩ ⟨прил.⟩ →большой ⟨сущ.⟩ →пес
⟨гл. гр.⟩ → ⟨нар.⟩⟨гл.⟩ ⟨нар.⟩ → λ ⟨гл.⟩ →прыгает
⟨предл.⟩через
⟨прил.⟩ →старый ⟨сущ.⟩ →забор

Соответствующее дерево грамматического разбора изображенно на рис. 17.36.

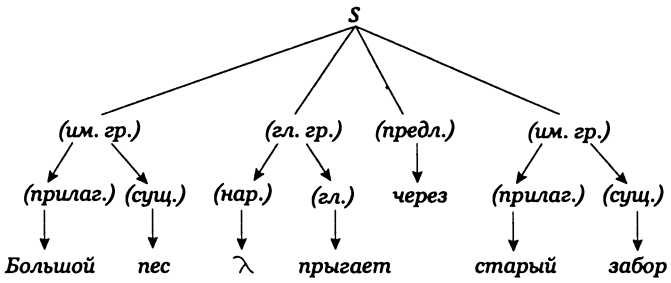


Рис. 17.36

Во всех грамматиках данного раздела продукции имели вид $A \to W$, где A — нетерминальный символ. Следовательно, эти продукции можно использовать всюду, где возникает A, невзирая на его расположение в выражении. Такие грамматики называются ***контекстно-свободными грамматиками***. Если грамматика содержит продукцию вида $aAb \to W$, где A — нетерминальный символ, и a либо b не являются пустым словом, то эту продукцию можно использовать только тогда, когда a находится слева от A, а b — справа. Поэтому такую продукцию нельзя применять непосредственно к A, и, следовательно, грамматика зависит от контекста, в котором находится A. Такая грамматика называется ***контекстно-зависимой***, или ***контекстной грамматикой***. □

В приведенных ниже примерах рассматриваются контекстно-свободные грамматики, порождающие более абстрактные языки.

ПРИМЕР 17.32. Пусть $\Gamma = (N, T, S, P)$ — грамматика, определенная с помощью $N = \{S, A, B\}$, $T = \{a, b\}$ и множества продукций P.

$$S \to AB, \qquad A \to a, \qquad B \to Bb, \qquad B \to \lambda.$$

Применив продукцию $S \to AB$, выводим AB. Далее с помощью продукций $A \to a$ и $B \to \lambda$ выводим a. Если теперь последовательно использовать продукции

$$S \to AB, \qquad A \to a, \qquad B \to Bb, \qquad B \to \lambda,$$

то получим ab. Можно породить также $abbb$, $abbbb$ и $abbbbb$. Фактически, для любых неотрицательных чисел n можно породить ab^n. Следовательно, язык, порожденный грамматикой Γ, задается выражением ab^*. □

ПРИМЕР 17.33. Пусть $\Gamma' = (N, T, S, P)$ — грамматика, определенная с помощью $N = \{S, A\}$, $T = \{a, b\}$ и множества продукций P.

$$S \to aAb, \qquad A \to aAb, \qquad A \to \lambda.$$

Применив продукции $S \to aAb$ и $A \to \lambda$, выводим ab. Применяя последовательно продукции

$$S \to aAb, \qquad A \to aAb, \qquad A \to \lambda,$$

получаем $aabb$, или a^2b^2. Воспользовавшись последовательно продукциями

$$S \to aAb \quad , \quad A \to aAb \quad , \quad A \to aAb \quad , \quad A \to ab,$$

выводим $aaabbb$, или a^3b^3. Легко показать, что язык, порожденный грамматикой Γ', — это $\{a^nb^n : n$ — положительное целое число $\}$. Отметим, что это не то же самое, что a^*b^*, поскольку последнее может включать также a^mb^n, где числа m и n не обязательно равны. □

ПРИМЕР 17.34. Пусть $\Gamma'' = (N, T, S, P)$ — грамматика, определенная с помощью $N = \{S, A, B\}$, $T = \{a, b\}$ и множества продукций P.

$$S \to ABABABA, \qquad A \to Aa, \qquad A \to \lambda, \qquad B \to b.$$

Можно показать, что язык, порожденный грамматикой Γ'', задается выражением $a^*ba^*ba^*ba^*$. Это язык, состоящий из всех слов, содержащих в точности три символа b. □

В примере 17.33 порожден язык $\{a^nb^n : n$ — положительное целое число$\}$. Интуитивно понятно, что это не регулярный язык, поскольку единственный способ построения бесконечного регулярного языка на основе конечного алфавита состоит в использовании звезды Клини $*$. В данном случае единственной возможностью является a^*b^*, но, как отмечалось ранее, эта возможность не реализуема, так как сюда включаются и слова a^mb^n, где числа m и n не обязательно равны.

Возникает вопрос, существует ли такая грамматика, которая порождает только регулярные языки. Ответ положительный.

ОПРЕДЕЛЕНИЕ 17.35. Контекстно-свободная грамматика $\Gamma = (N, T, S, P)$ называется ***регулярной грамматикой***, если каждая продукция $p \in P$ имеет вид $n \to w$, где строка w содержит не более одного нетерминального символа, который встречается только в конце строки.

Следовательно, w может иметь вид $aacA$, ab или bA, где a, b и c — терминальные символы, а символ A — нетерминальный. Однако, строка w не может иметь вид aAb, aAB или Aa. Продукцию $n \to abcA$ можно заменить продукцией

$$n \to aB, \qquad B \to bC, \qquad C \to cA\,.$$

Возможно также, что строка w не содержит ни одного терминального символа и один нетерминальный, поэтому имеем $B \to C$, но если за этим следует $C \to tC$, где t — терминальный символ, то можно скомбинировать две продукции и получить $B \to tC$. Поэтому нет необходимости вводить ограничение на то, чтобы каждая продукция имела одну из следующих форм:

$$A \to aB, \qquad B \to b, \qquad C \to \lambda\,.$$

где A, B и C — нетерминальные символы, а символы a и b — терминальные.

Приведенная ниже теорема объединяет понятия регулярных грамматик, регулярных языков и автоматов. Теорема сформулирована без доказательства.

ТЕОРЕМА 17.36. Язык является регулярным тогда и только тогда, когда он порожден регулярной грамматикой.

■ УПРАЖНЕНИЯ

1. Воспользовавшись грамматикой из примера 17.32, постройте дерево грамматического разбора для строки $abbb$.
2. Воспользовавшись грамматикой из примера 17.33, постройте дерево грамматического разбора для строки $aaabbb$.
3. Воспользовавшись грамматикой из примера 17.34, постройте дерево грамматического разбора для строки $babaab$.
4. Найдите язык, порожденный грамматикой $\Gamma = (N, T, S, P)$, определенной с помощью $N = \{S, A, B\}$, $T = \{a, b\}$ и заданного множества продукций P:

$$S \to AB, \qquad A \to aA, \qquad A \to \lambda, \qquad B \to Bb, \qquad B \to \lambda\,.$$

5. Найдите язык, порожденный грамматикой $\Gamma = (N, T, S, P)$, определенной с помощью $N = \{S, A, B\}$, $T = \{a, b\}$ и заданного множества продукций P:

$$S \to aB, \qquad B \to bA, \qquad A \to aB, \qquad B \to b\,.$$

6. Найдите язык, порожденный грамматикой $\Gamma = (N, T, S, P)$, определенной с помощью $N = \{S, A, B\}$, $T = \{a, b\}$ и заданного множества продукций P:

$$S \to aA, \quad B \to aA, \quad S \to bB, \quad A \to aB,$$
$$B \to bB, \quad A \to bA, \quad B \to b, \quad A \to a.$$

7. Найдите язык, порожденный грамматикой $\Gamma = (N, T, S, P)$, определенной с помощью $N = \{S, A, B\}$, $T = \{a, b\}$ и заданного множества продукций P:

$$S \to C, \quad A \to aB, \quad C \to bC, \quad B \to bB,$$
$$C \to aA, \quad B \to aA, \quad A \to bA, \quad B \to \lambda.$$

ГЛАВА

18

ТЕОРИЯ КОДОВ

18.1. ВВЕДЕНИЕ

Когда речь идет о кодах, как правило, имеются в виду секретные коды. Большинство читателей наверняка вспомнят, как в детстве придумывали свои коды, непонятные взрослым. Проблема передачи информации, которая должна быть понятна только тому, для кого она предназначена, весьма важна для военных целей, для обеспечения конфиденциальности переписки и в других случаях, когда информация может попасть в руки посторонних лиц. Параллельно с разработкой кодов идет работа по их взламыванию. Имеется в виду перехват и декодирование сообщений упомянутыми выше посторонними лицами. Это серьезная задача для военных и всех, кто заинтересован в защите целостности своих кодов и взламывании чужих. В наши дни взламывание кодов стало весьма популярным увлечением. Существует своего рода состязание между теми, кто создает коды, и теми, кто их взламывает. Одним из наиболее примечательных случаев взлома была разгадка секретного кода Германии в период Второй мировой войны (см. [54], [55]). Раздел теории кодов, посвященный этим вопросам, называется криптологией.

В данном разделе рассматриваются различные типы кодов. Как и в разделе 15.3, мы определим ***код*** как представление множества символов строками, состоящими из единиц и нулей. Это множество символов обычно включает буквы алфавита, цифры и, как правило, определенные контрольные символы. Коды представляются бинарными строками с целью использования их компьютерами для хранения и передачи данных. Желательно, чтобы коды обладали некоторыми свойствами. Мы увидим, однако, что некоторые из этих свойств несовместимы. Наиболее важное свойство кода состоит в том, что когда сообщение кодируется как двоичная строка, состоящая из конкатенации элементов кода, эта конкатенация однозначна. При декодировании сообщения не должно возникать проблем с тем, какую букву представляет элемент кода. Такой код назовем однозначно декодируемым кодом (см. определение 17.8). Существует несколько способов достижения этой цели. Один из них — кодирование всех символов двоичными строками одной длины. Такой код называется ***блоковым***. Например, если для кодирования каждого символа используется 8 бит, то известно, что каждые восемь бит пред-

ставляют один символ передаваемого сообщения. Блоковый код особенно полезен при ограничении длины кода для каждого посланного символа или буквы. Другим способом построения однозначно декодируемого кода является использование ***префиксного кода***, который рассматривался в разделе, посвященном взвешенным деревьям. Согласно определению 17.10, код C является префиксным, если обладает тем свойством, что элемент кода не может быть начальной строкой другого элемента кода. Таким образом, при чтении строки из единиц и нулей, изображающей символ A, всегда известен момент завершения строки. Префиксный код называют также ***мгновенным кодом***.

Разновидностью префиксного кода является ***кома-код***. При его использовании каждый символ кодируется строкой из единиц, в конце которой стоит нуль. Значит, множество строк кода имеет вид $\{0, 10, 110, 11110, 111110, \ldots\}$. Этот код имеет явный недостаток: элементы кода могут быть очень длинными и занимать большой объем памяти.

Часто необходимо сжимать данные, чтобы минимизировать объем памяти для их хранения или время для передачи данных. Что касается минимизации объема памяти, то наиболее эффективным является ***код Хаффмана***. Кодирование при помощи кода Хаффмана описано в разделе 15.3. Преимущество кода Хаффмана состоит также в том, что это мгновенный код.

Хорошо известным примером кода, минимизирующего время передачи, является ***код Морзе***. В кодах Хаффмана и Морзе буквы или символы, которые встречаются наиболее часто, имеют более короткий код. В коде Морзе буквы разделены "пробелами", а слова — тремя "пробелами". В данном случае "пробелы" — это единицы времени.

Код Морзе					
A	·−−−	J	·−−−	S	···
B	−···	K	−·−−−	T	−−−
C	−·−·	L	·−··	U	··−−−
D	−··	M	−−	V	···−
E	·	N	−·	W	·−−
F	··−·	O	−−−	X	−··−−−
G	−−·	P	·−−·	Y	−·−−
H	····	Q	−−·−−−	Z	−−··
I	··	R	·−·		

В процессе передачи данных могут возникать ошибки. Все, что может стать причиной ошибок, называется неопределенным термином "шум". Например, данные, полученные от такого отдаленного космического корабля, как "Voyager" или "Galileo", наверняка подвержены различного рода шумам. В некоторых случаях интерес представляет только определение наличия ошибки, что соответствует ситуации передачи данных еще раз. Коды, обладающие свойством определения наличия ошибок, называются ***кодами, обнаруживающими ошибки***.

В другом случае, когда данные не могут быть переданы еще раз (например, данные от удаленного космического корабля), требуется дополнительная инфор-

мация о данных с целью не только обнаружения, но и исправления ошибки. Коды, позволяющие исправлять ошибки, называются ***кодами, исправляющими ошибки***. Может показаться разумным всегда использовать коды с исправлением ошибок. Проблема использования кодов с исправлением ошибок и кодов с обнаружением ошибок состоит в том, что они должны включать в себя дополнительную информацию, поэтому они являются менее эффективными в отношении минимизации объема памяти. К сожалению, использование кодов с исправлением ошибок и кодов с обнаружением ошибок не дает абсолютной гарантии того, что ошибка будет исправлена или обнаружена. Проблема заключается в возможности многих ошибок. Несомненно, ошибку можно исправить или обнаружить, если она одна. Все, что мы можем сделать, — это уменьшить вероятность того, что ошибки останутся необнаруженными и неисправленными. Но проблема, опять-таки, в том, что чем в большей степени будет уменьшена эта вероятность, тем больше информации потребуется переслать, и тем менее эффективным будет наш код.

Прежде чем начать изучение кодов с исправлением ошибок и кодов с обнаружением ошибок, необходимо ввести еще один тип кодов. Существует классический пример, демонстрирующий применение такого кода. Предположим, имеется вращающийся диск, разделенный на секторы, и серия щеток или лазерных лучей, выделяющих информацию о скорости вращения диска. Если двоичные строки, записывающие нумерацию соседних секторов, существенно различаются в отдельных цифрах при переходе от одного сектора к следующему, тогда считывающее устройство воспримет это так, чтобы измененный сектор мог выработать число, совершенно отличное от числа, выработанного любым из секторов. В данном случае желательно пронумеровать секторы таким образом, чтобы двоичные строки, определяющие соседние секторы, различались только одной цифрой. Код, обладающий таким свойством, называется ***кодом Грея***. Его построению посвящен раздел 6.7.

В качестве первого метода обнаружения ошибок рассмотрим бит контроля четности. Продемонстрируем этот метод на примере кода ASCII. ASCII-код является блоковым кодом, который использует 7 битов, поэтому любой закодированный символ изображается строкой из семи символов 1 и 0. Восьмой бит добавляется таким образом, чтобы количество единиц было четным. Поэтому, если код переданной строки получен с единственной ошибкой, то количество единиц будет нечетным, и получатель информации поймет, что произошла ошибка. К сожалению, если произошло две ошибки, их нельзя будет обнаружить, поскольку количество единиц опять будет четным. Предположим, что вероятность ошибки при передаче равна 0.01 как для изменения 1 на 0, так и для изменения 0 на 1. Предположим далее, что вероятность ошибки не зависит от расположения ошибки и от того, что является ошибкой: изменение 1 на 0 или наоборот. Предположим также, что появление одной ошибки не влияет на вероятность появления другой.

Из биномиальной теоремы для вероятности (теорема 8.101) известно, что вероятность появления в точности одной ошибки равна $\binom{8}{1}(.01)(.99)^7$, что приблизительно равно 0.07. Вероятность же появления в точности двух ошибок равна $\binom{8}{2}(.01)^2(.99)^6$, что приблизительно равно 0.002 и является существенно меньшей величиной.

Код ASCII с битом контроля четности							
код	*символ*	*код*	*символ*	*код*	*символ*	*код*	*символ*
00000000	NUL	10100000	SP	11000000	@	01100000	‘
10000001	SOH	00100001	!	01000001	A	11100001	a
10000010	STX	00100010	”	01000010	B	11100010	b
00000011	ETX	10100011	#	11000011	C	01100011	c
10000100	EOT	00100100	$	01000100	D	11100100	d
00000101	ENQ	10100101	%	11000101	E	01100101	e
00000110	ACK	10100110	&	11000110	F	01100110	f
10000111	BEL	00100111	’	01000111	G	11100111	g
10001000	BS	00101000	(	01001000	H	11101000	h
00001001	HT	10101001	)	11001001	I	01101001	i
00001010	LF	10101010	*	11001010	J	01101010	j
10001011	VT	00101011	+	01001011	K	11101011	k
00001100	FF	10101100	’	11001100	L	01101100	l
10001101	CR	00101101	–	01001101	M	11101101	m
10001110	SO	00101110	.	01001110	N	11101110	n
00001111	SI	10101111	/	11001111	O	01101111	o
10010000	DLE	00110000	0	01010000	P	11110000	p
00010001	DC1	10110001	1	11010001	Q	01110001	q
00010010	DC2	10110010	2	11010010	R	01110010	r
10010011	DC3	00110011	3	01010011	S	11110011	s
00010100	DC4	10110100	4	11010100	T	01110100	t
10010101	NAK	00110101	5	01010101	U	11110101	u
10010110	SYN	00110110	6	01010110	V	11110110	v
00010111	ETB	10110111	7	11010111	W	01110111	w
00011000	CAN	10111000	8	11011000	X	01111000	x
10011001	EM	00111001	9	01011001	Y	11111001	y
10011010	SUB	00111010	:	01011010	Z	11111010	z
10011011	ESC	10111011	;	11011011	[	01111011	{
10011100	FS	00111100	<	01011100	\	11111100	|
00011101	GS	10111101	=	11011101	]	01111101	}
00011110	RS	10111110	>	11011110	^	01111110	~
10011111	US	00111111	?	01011111	_	11111111	DEL

Поскольку 3 ошибки будут обнаружены, вероятность не обнаружить более двух ошибок меньше, чем вероятность наличия четырех или более ошибок, поскольку любое нечетное количество ошибок будет обнаружено. Эта вероятность пренебрежимо мала по сравнению с вероятностью одной или менее ошибок.

Рассмотрим код, в котором для кодирования используется простое повторение кодируемой строки заданное число раз. Например, если при кодировании каждую строку кода нужно повторить один раз, то 10110 будет закодировано как 1011010110. Если при кодировании каждую строку кода нужно повторить дважды, то 10110 будет закодировано как 101101011010110. Если при кодировании каждая строка повторена один раз, то в результате получаем код с обнаружением

ошибок. Если произошла ошибка, то соответствующие позиции не будут совпадать. Например, если закодированная строка имеет вид 111111010110111011, то ошибки имеются в третьем и в последнем битах. Исправить ошибки мы не можем, поскольку не знаем, в какой из копий какая ошибка присутствует. Если кодируемая строка повторяется дважды, то лучше всего выявить ошибку. Если имеются три копии строки, то можно исправить код при наличии единственной ошибки. Если имеется отличие в битах в соответствующих позициях строк, то выбирается значение, которое встречается дважды. Например, если строка имеет длину 4 и нам передано 110110011101, то во второй позиции мы два раза получим 1 и один раз 0. Таким образом, предполагаем, что правильное значение равно 1, и правильным вариантом закодированной строки является 1101. Конечно, если ошибка появляется в одной и той же позиции строки более одного раза, возникает проблема. Если строка повторена так, что имеется n ее копий, то исправление ошибки дает правильный результат, если при повторении ошибка встречается в одной и той же позиции менее, чем $\lfloor \frac{n}{2} \rfloor$ раз.

18.2. ПОРОЖДАЮЩИЕ МАТРИЦЫ

С этого момента мы будем предполагать, что все строки кода имеют фиксированную длину n, и будем трактовать эти строки как векторы или $(1 \times n)$-матрицы. Следовательно, можно использовать сложение векторов. Мы, однако, определим сложение по модулю 2, так что $1 + 1 = 0$. Таким образом,

$$11110001 + 10100111 = 01010110\,.$$

Коды, представленные в данном разделе, называются ***линейными кодами***. Они известны также как ***групповые коды***. Как уже упоминалось, строками кода C будут двоичные строки длины n, которые мы будем рассматривать также как векторы или $(1 \times n)$-матрицы. Если B_n — множество всех двоичных строк длины n, то C — подмножество множества B_n. Предлагаем читателю доказать, что множество B_n вместе с указанной выше операцией сложения образует группу относительно сложения. Каждая строка из B_n совпадает со своим обратным элементом, так что в этом смысле сложение и вычитание — это одно и то же. Если в некоторых ситуациях покажется, что следует вычитать, а не складывать, то нужно помнить, что это не имеет значения. Код C называется ***линейным***, если C — подгруппа группы B_n. Если вспомнить линейную алгебру, то можно заметить, что в действительности C — линейное векторное пространство, и многие из свойств линейной алгебры могут оказаться здесь полезными. Однако, рассматривая код C, мы будем использовать здесь только те свойства, что C является группой и что элемент из C — вектор, поэтому может быть умножен на матрицу соответствующего размера. Будет использован также закон дистрибутивности для матриц, согласно которому

$$A(B + C) = AB + AC$$

для любых матриц A, B и C, произведение которых определено. Читатель может вспомнить, что если $u = (u_1, u_2, u_3, \ldots, u_n)$ и $v = (v_1, v_2, v_3, \ldots, v_n)$, то ***скалярное***

произведение векторов u и v, обозначаемое $u \bullet v$, равно

$$u_1v_1 + u_2v_2 + u_3v_3 + \cdots + u_nv_n\,.$$

Весом строки кода c, обозначаемым $wt(c)$, называется количество единиц в строке. Например, если $c = 1011010$, то $wt(c) = 4$.

Предположим, что имеется такая $(k \times n)$-матрица G, что ее первые k столбцов и строк образуют единичную матрицу I_k размера $(k \times k)$, все столбцы которой различны. Таким образом, матрица G имеет вид $[I_k|A_{n-k}]$. Например,

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

является такой матрицей. Матрица G называется ***порождающей матрицей***. Рассмотрим строки порождающей матрицы как векторы или строки кода. Обозначим это множество строк S. Например, для матрицы, приведенной выше,

$$S = \{100101, 010110, 001011\}.$$

Пусть C — код, образованный всеми векторами, которые являются конечными суммами строк из S. Предлагаем читателю доказать, что C — подгруппа B_n. В нашем примере получаем 110011, складывая первые две строки из S, поэтому 110011 будет принадлежать C. Говорят, что группа C ***порождена*** множеством S. Это S является также минимальным множеством, порождающим код C, поскольку никакие элементы из S не являются суммами других элементов из S. То, что группа C порождена множеством S, обозначим $C = S^*$. Код C вида $[I_k|A_{n-k}]$ (т.е. порожденный строками $[I_k|A_{n-k}]$) называется $[n, k]$-кодом.

ТЕОРЕМА 18.1. $[n, k]$-код C содержит 2^k строк.

ДОКАЗАТЕЛЬСТВО. Первые k битов строки C определяют элементы C. Позиции, в которых находятся единицы в первых k битах строки из C, показывают, какие строки из S были просуммированы. Например, если единицы находятся в первом и третьем битах строки из C, то эта строка получена в результате сложения первой и третьей строк кода G. Поскольку существует 2^k различных способов сформировать первые k битов, то в коде C имеется 2^k строк. ■

Если необходимо передать строки сообщения длины k, то мы кодируем их, умножая справа на матрицу G. Таким образом, если $w = w_1w_2w_3\cdots w_k$ или $(w_1, w_2, w_3, \ldots, w_k)$, то мы кодируем это строкой wG. В нашем примере закодируем 110 или $(1, 1, 0)$ как

$$(1,1,0)\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} = (1,1,0,0,1,1)$$

или 110011. Заметим, что строка сообщения совпадает с первыми тремя битами закодированной строки. В общем случае строка сообщения длины k будет

совпадать с первыми k битами закодированной строки, поскольку единичная матрица I_k, образующая первые k столбцов матрицы G, просто повторяет исходную строку. Поэтому мы можем осуществлять декодирование, взяв первые k битов закодированной строки. Заметим также, что выше, при умножении $(1,1,0)$ на G, мы получили

$$(1,1,0,0,1,1) = 1\cdot(1,0,0,1,0,1) + 1\cdot(0,1,0,1,1,0) + 0\cdot(0,0,1,0,1,1)\,,$$

так как для любого вектора v имеем $1\cdot v = v$ и $0\cdot v = (0,0,0,0,0,0)$. Таким образом, закодированная строка является суммой векторов из множества S и, следовательно, принадлежит C, поскольку C — группа. В общем случае, если $S = (s_1, s_2, s_3, \ldots, s_k)$ — множество строк порождающей матрицы и $w = (w_1, w_2, w_3, \ldots, w_k)$ — код сообщения, то закодированная строка имеет вид

$$w_1 s_1 + w_2 s_2 + w_3 s_3 + \ldots + w_k s_k\,.$$

Она представляет собой сумму строк из S, так как каждое w_i равно либо 1, либо 0 и, следовательно, принадлежит C, поскольку C — группа, порожденная множеством S.

Отметим, что G имеет вид $[I_k|A_{n-k}]$ и что I_k передает сообщение. А что делает A_{n-k}? Рассмотрим тот же пример. Если (w_1, w_2, w_3) — строка сообщения, то закодированная строка имеет вид

$$(w_1,w_2,w_3)\begin{bmatrix}1&0&0&1&0&1\\0&1&0&1&1&0\\0&0&1&0&1&1\end{bmatrix} = (w_1, w_2, w_3, w_1+w_2, w_2+w_3, w_1+w_3)\,.$$

Таким образом, четвертый бит закодированной строки должен быть равен w_1+w_2, пятый бит должен быть равен w_2+w_3 и шестой бит должен быть равен w_1+w_3. Следовательно, если закодированная строка имеет вид $(w_1, w_2, w_3, w_4, w_5, w_6)$, то $w_4 = w_1 + w_2$, $w_5 = w_2 + w_3$ и $w_6 = w_1 + w_3$. Если любая закодированная строка после передачи не удовлетворяет этим соотношениям, то понятно, что при передаче произошла ошибка. Например, если получена закодированная строка 101100, то, учитывая, что $w_1 = 1$, $w_2 = 0$ и $w_3 = 1$, должны иметь $w_4 = 1 = 1 + 0$, $w_5 = 0 = 0 + 1$ и $w_6 = 0 = 1 + 1$. Поскольку соотношение для w_5 не выполняется, становится понятным, что имеется ошибка. Таким образом, матрица A_{n-k} служит для контроля точности передачи данных так же, как ранее это делал бит контроля четности. В общем случае, если имеется закодированная строка $w_1 w_2 w_3 \ldots w_k \ldots w_n$ и

$$R = \begin{bmatrix}1&0&0&\cdots&0&A_{1,k+1}&A_{1,k+2}&\cdots&A_{1,n}\\0&1&0&\cdots&0&A_{2,k+1}&A_{2,k+2}&\cdots&A_{2,n}\\0&0&1&\ddots&\vdots&A_{3,k+1}&A_{3,k+2}&\cdots&A_{3,n}\\\vdots&\vdots&\vdots&\ddots&0&\vdots&\vdots&\ddots&\\0&0&0&0&1&A_{k,k+1}&A_{k,k+2}&&A_{k,n}\end{bmatrix},$$

то для $i > k$ имеем $w_i = w_1 A_{1,i} + w_2 A_{2,i} + w_3 A_{3,i} + \cdots + w_k A_{k,i}$, и закодированная строка должна удовлетворять всем этим $n-k$ соотношениям.

Следующая проблема — исправление ошибки, если известно, что произошла единственная ошибка. Изложенный ниже метод известен как использование **лидеров смежных классов**. Метод будет проиллюстрирован нашим примером, в котором

$$R = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Известно, что

$$S = \{100101, 010110, 001011\},$$

так что

$$C = \{000000, 100101, 010110, 001011, 110011, 011101, 101110, 111000\}.$$

Теперь, как это описано в разделе 9.4, сформируем для C смежные классы в B_n. Первым смежным классом является сам C. Для формирования следующего смежного класса выберем элемент из B_n, который имеет минимальный вес и не принадлежит C. Например, можно выбрать $b_1 = 100000$. Смежный класс

$$b_1 + C = \{100000, 000101, 110110, 101011, 010011, 111101, 001110, 011000\}.$$

Опять выбираем элемент из B_n, который имеет минимальный вес и не принадлежит ни одному из предыдущих смежных классов. Например, можно выбрать $b_2 = 010000$. Смежный класс

$$b_2 + C = \{010000, 110101, 000110, 011011, 100011, 001101, 111110, 101000\}.$$

Продолжая этот процесс, получаем следующую таблицу для B_n, разделенного на смежные классы, где элемент с минимальным весом выписан первым. Элементы первого столбца называются лидерами смежных классов.

000000	100101	010110	001011	110011	011101	101110	111000
100000	000101	110110	101011	010011	111101	001110	011000
010000	110101	000110	011011	100011	001101	111110	101000
001000	101101	011110	000011	111011	010101	100110	110000
000100	100001	010010	001111	110111	011001	101010	111100
000010	100111	010100	001001	110001	011111	101100	111010
000001	100100	010111	001010	110010	011100	101111	111001
100010	000111	110100	101001	010001	111111	001100	011010

В последнем смежном классе необходимо использовать строки веса 2. Можно выбрать одну из строк 100010, 010001 или 001100. Откровенно говоря, нам повезло, что мы нашли хотя бы одну. Процесс работает следующим образом. Получив закодированную строку, ищем ее в таблице. Предположим, например, что получена строка 110110. Смотрим на первый столбец строки, содержащей 110110, и получаем 100000, поэтому предполагаем, что ошибка возникла в первом бите,

поскольку 100000 была добавлена каждому элементу C при построении этого смежного класса. Смотрим на первую строку столбца, содержащего 110110, и получаем 010110, что, по нашему предположению, является правильной строкой кода C. Теперь предположим, что получена строка 001010. Смотрим на первый столбец строки, содержащей 001010, и получаем 000001, поэтому предполагаем, что ошибка возникла в шестом бите. Смотрим на первую строку столбца, содержащего 001010, и находим 001011, что, по нашему предположению, является правильной строкой кода C.

Теперь поищем более простой способ обнаружения ошибок. Пусть v и w — векторы (или строки) длины n. Вектор v называется ***ортогональным*** вектору w, если их скалярное произведение $v \bullet w = 0$. Пусть задан код C. ***Двойственным кодом*** к коду C, обозначаемым $C^{\perp}$, называется множество всех строк из B_n, ортогональных каждой строке из кода C. Предоставляем читателю доказать, что $C^{\perp}$ — подгруппа группы B_n. Если код C является $[n, k]$-кодом, так что он порожден k строками, то $C^{\perp}$ порожден $n - k$ строками. Поскольку доказательство этого утверждения требует знания линейной алгебры, то доказывать его не будем, а просто примем на веру. Следующую теорему, однако, мы докажем.

ТЕОРЕМА 18.2. Пусть C — групповой код, а $C^{\perp}$ — двойственный ему код. Строка t принадлежит коду $C^{\perp}$ тогда и только тогда, когда она ортогональна каждой строке из S, множества порождающих элементов кода C.

ДОКАЗАТЕЛЬСТВО. Очевидно, что если строка t принадлежит коду $C^{\perp}$, то она ортогональна каждой строке из множества S, поскольку она ортогональна каждой строке из кода C и $S \subseteq C$. Предположим, что $S = \{s_1, s_2, s_3, \ldots s_k\}$, и строка t ортогональна строке s_i для всех $s_i \in S$, так что $t \bullet s_1 = 0$ для всех $s_i \in S$. Каждый элемент из C имеет вид

$$w_1 s_1 + w_2 s_2 + w_3 s_3 + \ldots + w_k s_k\,,$$

где каждое w_1 равно 1 или 0. Согласно свойству линейности матриц имеем

$$\begin{aligned} t \bullet (w_1 s_1 + w_2 s_2 + \ldots + w_k s_k) &= w_1(t \bullet s_1) + w_2(t \bullet s_2) + \ldots + w_k(t \bullet s_k) = \\ &= 0 + 0 + 0 + \ldots + 0 = \\ &= 0\,. \end{aligned}$$

■

Вспомним, что в рассматриваемом примере

$$G = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [I_3 | A_3]\,.$$

Определим теперь

$$G^{\perp} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} = [A_3^t | I_3]\,,$$

где A_3^t — транспозиция матрицы A_3, полученная заменой строк матрицы A_3 на столбцы. Матрица $G^\perp$ называется ***матрицей контроля четности***.

Перебирая все возможные варианты, можно показать, что скалярное произведение любой строки из G с любой строкой из $G^\perp$ равно 0. Отсюда получаем, что

$$G^\perp r_i^t = 0,$$

где r_i^t — транспозиция i-ой строки матрицы G. По определению транспозиции r_i^t является i-ой строкой матрицы G, преобразованной в столбец. Мы превращаем строку в столбец, чтобы на него можно было умножить матрицу $G^\perp$:

$$\begin{aligned} G^\perp(w_1 r_1^t + w_2 r_2^t + w_3 r_3^t) &= w_1 G^\perp r_1^t + w_2 G^\perp r_2^t + w_3 G^\perp r_3^t = \\ &= 0 + 0 + 0 = \\ &= 0\,. \end{aligned}$$

Таким образом, если умножить матрицу $G^\perp$ на транспозицию любого элемента из C, то получим 0.

В общем случае, если

$$G = [I_k | A_{n-k}] = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & A_{1,k+1} & A_{1,k+2} & \cdots & A_{1,n} \\ 0 & 1 & 0 & \cdots & 0 & A_{2,k+1} & A_{2,k+2} & \cdots & A_{2,n} \\ 0 & 0 & 1 & \ddots & \vdots & A_{3,k+1} & A_{3,k+2} & \cdots & A_{3,n} \\ \vdots & \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & 0 & 1 & A_{k,k+1} & A_{k,k+2} & & A_{k,n} \end{bmatrix},$$

то

$$G^\perp = [A_{n-k}^t | I_{n-k}] = \begin{bmatrix} A_{1,k+1} & A_{2,k+1} & \cdots & A_{k,k+1} & 1 & 0 & 0 & \cdots & 0 \\ A_{1,k+2} & A_{2,k+2} & \cdots & A_{k,k+2} & 0 & 1 & 0 & \cdots & 0 \\ A_{1,k+3} & A_{2,k+3} & \cdots & A_{k,k+3} & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & & \vdots & \vdots & \vdots & \ddots & 0 \\ A_{1,n} & A_{2,n} & & A_{k,n} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Скалярное произведение i-ой строки матрицы G на j-ю строку матрицы $G^\perp$ равно

$$0 + 0 + \cdots 0 + A_{i,j} + 0 + \cdots 0 + A_{i,j} + \cdots + 0 + 0 = 0\,,$$

так что в общем случае

$$G^\perp r_i^t = 0,$$

где r_i^t — транспозиция i-ой строки матрицы R. Если умножить матрицу $G^\perp$ на транспозицию любого элемента из C, то, используя те же самые рассуждения, имеем в результате 0. Мы также получаем еще один замечательный результат. Если два элемента b_1 и b_2 из B_n принадлежат одному и тому же смежному классу, образованному в B_n с использованием группы C, как это было сделано выше, то

$$G^\perp b_1^t = G^\perp b_2^t\,.$$

Для доказательства воспользуемся тем фактом, что если b_1 и b_2 принадлежат одному смежному классу, то $b_1 = b_2 + c$ для некоторого $c \in C$. Следовательно, $b_1^t = b_2^t + c^t$ и

$$\begin{aligned} G^{\perp} b_1^t &= G^{\perp}(b_2^t + c^t) = \\ &= G^{\perp} b_2^t + G^{\perp} c^t = \\ &= G^{\perp} b_2^t + 0 = \\ &= G^{\perp} b_2^t \ , \end{aligned}$$

т.к.

$$G^{\perp} c^t = 0$$

для всех $c \in C$.

Поскольку $G^{\perp} b^t$ одно и то же для всех b из класса смежности, то можно выбрать любое b из класса смежности и определить это значение. Таким образом, в каждую строку приведенной выше таблицы можно добавить это общее значение образа $G^{\perp}$, т.к. элементы каждой строки определяют класс смежности. Мы выбираем лидера смежного класса, потому что он простейший, и помещаем значение его образа во второй столбец. Эти значения называются ***синдромами***. Уже известно, что первый синдром есть $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Находим, что

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

поэтому $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ — второй синдром, и

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

поэтому $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ — третий синдром.

Продолжая процесс, получаем следующую таблицу.

000000	0 0 0	100101	010110	001011	110011	011101	101110	111000
100000	1 0 1	000101	110110	101011	010011	111101	001110	011000
010000	1 1 0	110101	000110	011011	100011	001101	111110	101000
001000	0 1 1	101101	011110	000011	111011	010101	100110	110000
000100	1 0 0	100001	010010	001111	110111	011001	101010	111100
000010	0 1 0	100111	010100	001001	110001	011111	101100	111010
000001	0 0 1	100100	010111	001010	110010	011100	101111	111001
100010	1 1 1	000111	110100	101001	010001	111111	001100	011010

Заполнив таблицу, предположим, что получена переданная строка 101100. Умножение матрицы $G^{\perp}$ на транспозицию строки дает

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Отсюда следует, что 101100 находится в строке 6. Лидер смежного класса — 000010, элемент крайнего левого столбца строки, содержащей 101100. Элемент C, расположенный в верхней строке столбца, содержащего 101100, равен 101110. Согласно способу построения таблицы имеем $101100 = 101110 + 000010$, поэтому можем предположить, что переданная строка 101100 должна иметь вид 101110, и в пятом бите была ошибка.

Этот метод намного быстрее, поскольку требует только умножить матрицу $G^{\perp}$ на транспозицию строки-сообщения, найти строку, содержащую синдром,

и найти переданное сообщение. Лидер для этого сообщения — индикатор ошибки, а элемент кода C, расположенный в первой строке столбца, содержащего переданное сообщение, есть исправленное сообщение.

Заметим, однако, что процесс можно сделать еще быстрее. При этом потребуются только два первых столбца приведенной выше таблицы. Предположим, что получено сообщение 110000. Умножение матрицы $G^{\perp}$ на его транспозицию дает

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

поэтому синдром — $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, а лидер смежного класса — 001000. Поскольку лидер указывает, что ошибка появилась в третьем бите, то, добавляя 001000 к 110000, получаем 111000, исправленный код. Таким образом, этот метод прост. Умножить транспозицию сообщения на матрицу $G^{\perp}$, чтобы найти синдром. Для этого нужно найти лидера смежного класса и прибавить его к сообщению, чтобы получить исправленный код. Обратите внимание, что используются только первые два столбца таблицы. Однако, теперь возникла еще одна проблема. Если полученным сообщением будет строка 101001, то синдром будет $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, и в соответствующей строке будет три элемента с весом 2. Вспомним, что строка 100010 была выбрана произвольно. Любая из таких строк равновероятно может содержать ошибку, поэтому в данном случае совершенно безнадежно использовать синдромы для исправления ошибки. Кроме того, мы пытаемся в данном случае исправить строку с двумя ошибками вместо одной.

■ УПРАЖНЕНИЯ

1. Какие из приведенных ниже матриц в действительности являются порождающими? Если нет, то почему?

а) $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$; **б)** $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$;

в) $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$; **г)** $\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$.

2. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

а) найдите $G^{\perp}$;
б) закодируйте $111, 011, 101, 110$;
в) декодируйте $111011, 110100, 010010, 011110$;
г) используя $G^{\perp}$, определите, какие из строк пункта (c) являются правильными кодированными строками.

3. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

а) найдите $G^{\perp}$;
б) закодируйте $1111, 0101, 1001, 1010$;
в) декодируйте $1111110, 0111010, 0111101, 1011110$;
г) используя $G^{\perp}$, определите, какие строки из пункта (c) являются правильными закодированными строками.

4. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

а) найдите $G^{\perp}$;
б) постройте таблицу смежных классов вместе с лидерами;
в) используйте эту таблицу для исправления ошибок (если они имеются) в переданных строках 1111100, 1111000, 0110101 и 1011000.

5. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

а) найдите $G^{\perp}$;
б) найдите синдромы для смежных классов;
в) используйте эти синдромы для исправления ошибок (если они имеются) в переданных строках 111101, 111001, 110010, 101001.

6. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

а) найдите $G^{\perp}$;

б) найдите синдромы для смежных классов;

в) используйте эти синдромы для исправления ошибок (если они имеются) в переданных строках 111101, 111001, 110101, 101001.

7. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

а) найдите $G^{\perp}$;

б) постройте таблицу смежных классов вместе с лидерами;

в) используйте эту таблицу для исправления ошибок (если они имеются) в переданных строках 1110011, 1111000, 1101001, 1011101.

8. Объясните, что случится при использовании порождающей матрицы в упражнении 7, если будут две или более ошибок.

9. Используя порождающую матрицу из упражнения 5, найдите другие строки, отличные от строк матрицы $G^{\perp}$, которые принадлежат $C^{\perp}$, и, используя эти строки, постройте матрицу контроля четности.

10. Используя порождающую матрицу из упражнения 4, найдите другие строки, отличные от строк матрицы $G^{\perp}$, которые принадлежат $C^{\perp}$, и, используя эти строки, постройте матрицу контроля четности.

11. Докажите, что B_n, множество всех двоичных строк длины n, — группа относительно сложения, введенного в этой главе.

12. Пусть C — код, образованный с помощью всех векторов, которые являются конечными суммами строк из S. Докажите, что C — подгруппа B_n относительно сложения.

13. Докажите, что **двойственный код** для кода C, обозначаемый $C^{\perp}$, является подгруппой группы B_n относительно сложения.

18.3. КОДЫ ХЕММИНГА

В конце предыдущего раздела было показано, что существовали определенные трудности при попытке исправить код для некоторых строк, поскольку все лидеры имели вес 1. Эти трудности устраняются путем использования матрицы, называемой ***матрицей Хемминга***, в качестве порождающей матрицы. Перед тем как перейти к рассмотрению матрицы Хемминга G_H, посмотрим на матрицу контроля четности $G_H^{\perp}$. Пусть $G_H^{\perp}$ — матрица, в которой r таких строк, что столбцы состоят из всевозможных строк длины r, за исключением строки, состоящей только из нулей. Предположим, что $r \geq 3$. Существует $2^r - 1$ таких строк, поэтому $G_H^{\perp}$ есть $(r \times n)$-матрица, где $n = 2^r - 1$. Будем использовать столбцы с весом 1 как последние r столбцов, формирующих единичную матрицу, так что матрица $G_H^{\perp}$ имеет вид $[A^t | I_r]$, где A^t — $(r \times (n - r))$-матрица. Матрица Хемминга G_H

— это $((n-r)\times n)$-матрица вида $[I_{n-r}|A]$, где A — $((n-r)\times r)$-матрица. Код, образованный строками матрицы Хемминга, называется ***кодом Хемминга***.

Например, пусть $r=3$ и $G_H^{\perp}$ — матрица

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Тогда G_H — матрица

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Для изучения матриц Хемминга необходимо понятие расстояния и его связь с весом каждой из строк. Начнем с теоремы о весах строк.

ТЕОРЕМА 18.3. Для строк c и c' вес $wt(c+c')\le wt(c)+wt(c')$.

ДОКАЗАТЕЛЬСТВО. Пусть $c=c_1c_2c_3\ldots c_n$ и $c'=c'_1c'_2c'_3\ldots c'_n$. Если $c_i+c'_i=1$, то либо $c_i=1$, либо $c'_i=1$. Поэтому существованию каждой единицы в $c+c'$ соответствует существование единицы либо в c, либо в c'. ■

Расстояние Хемминга, или просто ***расстояние*** между двумя строками кода c и c', имеющими одинаковую длину, — это число соответствующих битов в строке, где одна строка имеет цифру 1, а другая имеет цифру 0. Будем обозначать функцию расстояния через $\delta(c,c')$. Например, если $c=101011$ и $c'=110010$, то $\delta(c,c')=3$, так как две строки отличаются во второй, третьей и шестой позициях. Очевидно, что чем больше расстояние между двумя строками кода, тем больше ошибок можно обнаружить. Поскольку δ названа функцией расстояния, мы должны показать, что она обладает основными свойствами функции расстояния.

ТЕОРЕМА 18.4. Функция расстояния Хемминга имеет следующие свойства:

а) для строк c и c' расстояние $\delta(c,c')=0$ тогда и только тогда, когда $c=c'$;
б) для строк c и c' расстояние $\delta(c,c')=\delta(c',c)$;
в) для строк c, c' и c'' выполняется соотношение $\delta(c,c'')\le\delta(c,c')+\delta(c',c'')$.

ДОКАЗАТЕЛЬСТВО. Пункты (а) и (б) следуют непосредственно из определения, и их доказательство оставляем читателю. Для доказательства части (в) заметим, что для строк c и c'' вес $wt(c+c'')=\delta(c,c'')$, а также, что если $c=c_1c_2c_3\ldots c_n$ и $c''=c''_1c''_2c''_3\ldots c''_n$, то $c_i+c''_i$ вносит 1 в вес $wt(c+c'')$ тогда и только тогда, когда $c_i=0$ и $c''_i=1$ или $c_i=1$ и $c''_i=0$. Но это верно в том и только том случае, если c_i и c''_i различны, что вносит 1 в $\delta(c,c'')$. Заметим также, что для любой строки c' строка $c'+c'$ состоит только из нулей. Назовем такую строку $\mathbf{0}$. По определению сложения $\mathbf{0}+c=c$ для каждой строки c. Следовательно,

$$\delta(c,c'')=wt(c+c'')=wt(c+\mathbf{0}+c'')=wt(c+c'+c'+c'')\le$$
$$\le wt(c+c')+wt(c'+c'')=\delta(c,c')+\delta(c',c'').$$ ■

Важно знать минимальное расстояние между двумя строками кода. Если C — код, то ***минимальное расстояние*** кода C, обозначаемое $D(C)$, равно наименьшему расстоянию между двумя строками из C. В приведенной ниже теореме сформулирован важный критерий для определения числа ошибок, которые могут быть исправлены или обнаружены с использованием кода.

ТЕОРЕМА 18.5. Для кода C,

а) если $D(C) = k + 1$, то использование кода позволяет обнаружить вплоть до k ошибок;

б) если $D(C) = 2k + 1$, то использование кода позволяет исправить вплоть до k ошибок.

ДОКАЗАТЕЛЬСТВО. (а) Если $D(C) = k+1$, то для данной строки $c \in C$. Отсюда следует, что c отличается от любой другой строки кода по крайней мере в $k + 1$ позициях. Поэтому, если переданная строка c имеет k или менее ошибок, то она не может быть другой строкой кода, и ошибка определена.

(б) Если строка c передана как c' с k или менее ошибками, то для любой строки $c \in C$ имеем $\delta(c, c') \leq k$. Если для некоторой строки c'' из C расстояние $\delta(c', c'') \leq k$, то $\delta(c, c') + \delta(c', c'') \leq 2k$. Но $\delta(c', c'') \leq \delta(c, c') + \delta(c', c'')$ и $\delta(c, c'') \geq 2k + 1$, что приводит к противоречию. Следовательно, c' можно исправить на c, единственную строку, расстояние которой от c' меньше, чем $k + 1$. ■

Возникает проблема определения $D(C)$, наименьшего расстояния между любыми двумя строками из кода C. Для решения задачи нам сначала потребуется следующая теорема.

ТЕОРЕМА 18.6. Минимальное расстояние $D(C)$ кода C равно $W(C) = \min\{wt(c) : c \in C \text{ и } c \neq \mathbf{0}\}$.

ДОКАЗАТЕЛЬСТВО. По определению $D(C)$, существует $w, w' \in C$ такое, что $\delta(w, w') = D(C)$. Но $\delta(c, c'') = wt(c + c'')$ и, поскольку $c + c'' \in C$, то $W(C) \leq wt(c + c'')$. Следовательно, $W(C) \leq D(C)$. Обратно, для $c \in C$ имеем $wt(c) = wt(c+\mathbf{0}) = \delta(c,\mathbf{0}) \geq D(C)$. Следовательно, $W(C) \geq D(C)$. Отсюда $W(C) = D(C)$. ■

Покажем, что для кода Хемминга C выполняется соотношение $W(C) \geq 3$. Прежде всего, не существует $c \in C$ с весом 1. Если бы такая строка c существовала, то она содержала бы нули во всех позициях, за исключением позиции j, в которой стояла бы единица; в силу ортогональности строки c к каждой строке матрицы $G_H^{\perp}$ имели бы, что j-ый столбец матрицы $G_H^{\perp}$ содержит только 0, что противоречит построению матрицы $G_H^{\perp}$. Не существует также $c \in C$ с весом 2. Если бы такие $c \in C$ существовали, то c была бы строкой, состоящей только из нулей, за исключением двух единиц, например, в позициях i и j. Из ортогональности строки c к каждой строке матрицы $G_H^{\perp}$ следует, что i-ый и j-ый столбцы каждой строки матрицы $G_H^{\perp}$ должны быть одновременно 1 или 0. Но тогда i-ый и j-ый столбцы матрицы $G_H^{\perp}$ должны совпадать, что опять же противоречит построению матрицы $G_H^{\perp}$. Следовательно, $W(C) \geq 3$, и код C можно использовать для исправления единственной ошибки.

Теперь необходимо показать, что в каждом смежном классе, за исключением самого C, имеется один элемент с весом 1. Доказательство этого факта значительно упростит проблему декодирования. По теореме 18.1, $[n, k]$-код C содержит 2^k строк. Поскольку G_H имеет вид $[I_{n-r}|A]$, то код Хемминга C является $[n, n-r]$-кодом, поэтому код C содержит 2^{n-r} элементов. Множество B_n содержит 2^n элементов. Следовательно, существует

$$\frac{2^n}{2^{n-r}} = 2^r$$

смежных классов, включая C. Строка из C имеет длину $2^r - 1$. Поэтому существует $2^r - 1$ строк с весом 1. Теперь нужно показать, что никакой смежный класс не содержит две строки с весом 1. Предположим, что строки s и s' принадлежат одному смежному классу, и каждая из них имеет вес 1. Тогда, согласно определению смежного класса, $s = s' + c$ для некоторой строки $c \in C$. Следовательно, $c = s + s'$, поэтому, по теореме 18.3,

$$wt(c) \leq wt(s) + wt(s') \leq 1 + 1 = 2.$$

Но $wt(c) \geq 3$, что противоречит полученному. Поэтому каждый смежный класс содержит ровно одну строку с весом 1, за исключением C, который содержит строку с весом 0.

Вернемся к нашему примеру, в котором матрица $G_H^{\perp}$ имеет вид

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix},$$

а матрица G_H имеет вид

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Учитывая, что каждый смежный класс содержит лидера с весом 1, рассмотрим строку 0010000. Если умножить ее на матрицу $G_H^{\perp}$, получим

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

поэтому синдром равен $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, что совпадает с третьим столбцом матрицы $G_H^{\perp}$. По сути, если 1 встречается в j-ой позиции строки с весом 1, то синдром после умножения матрицы $G_H^{\perp}$ на транспозицию строки становится j-ым столбцом

матрицы $G_H^{\perp}$. Поэтому всякий раз, когда мы получаем правильную строку переданного сообщения, умножаем на ее транспозицию матрицу $G_H^{\perp}$, то получаем синдром со всеми нулями. Если имеется единственная ошибка, то мы получим один из столбцов матрицы $G_H^{\perp}$, поскольку строка переданного сообщения должна быть в одном из смежных классов, поэтому существует лидер смежного класса с весом 1. Следовательно, если i-ый столбец матрицы $G_H^{\perp}$ является синдромом, это говорит о том, что лидер смежного класса имеет 1 в i-ом столбце, поэтому ошибка имеется в i-ом столбце или в i-ом бите строки. Например, предположим, получена строка 1110110 переданного сообщения. Умножение матрицы $G_H^{\perp}$ на транспозицию этой строки дает

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

что совпадает со второй строкой матрицы $G_H^{\perp}$. Следовательно, ошибка во втором бите, и переданным сообщением должна быть строка 1010110.

Оставшуюся часть этого раздела посвятим небольшому обзору других кодов. Первый из них — код Голе. Коды Хемминга были открыты независимо Хеммингом в 1950 г. и Голе в 1949 г. Не будем пытаться объяснить, почему они названы именно кодами Хемминга. Это длинная и запутанная история, о которой подробно можно прочитать у Томпсона [112]. Тем не менее, существует и код Голе. Он был описан в работе Голе в 1949 г. Речь идет о разработанной им модели $(23, 12, 7)$, где цифры означают, что используется порождающая матрица размерности $(23, 12)$ с минимальным расстоянием 7 между строками кода C. Этот ***код Голе*** имеет порождающую матрицу $G = [I_{11}|A]$, где

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Эта матрица имеет геометрическую интерпретацию с использованием пяти прямых в плоскости (см. [112]). Код легче изучить, используя расширенную поро-

ждающую матрицу $G = [I_{12}|A]$ с параметрами $(24, 12, 8)$, где

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(см. [44]). Матрица получена путем перестановок в порождающей матрице Голе и добавления бита четности. Симметричность матрицы упрощает ее изучение. Легко видеть, что $G^{\perp} = [A|I_{12}]$, поскольку $A = A^t$. Голе ввел еще несколько кодов, среди которых $(4096, 244, 8)$-код, использованный космическим кораблем "Voyager" для передачи изображений планет Юпитер, Уран и Нептун.

Для заданной строки s длины n положим $\bar{s} = s+\mathbf{1}$, где $\mathbf{1}$ — строка длины n, состоящая только из единиц. Для заданной строки S положим

$$Plot(S) = \{ss : s \in S\} \cup \{s\bar{s} : s \in S\}.$$

Для заданного множества

$$S = \{0000, 0011, 1100, 1111\}$$

положим: $S_1 = Plot(S)$, $S_2 = Plot(S_1)$, $S_3 = Plot(S_2)$ и $S_n = Plot(S_{n-1})$. Конструкция *Plot* названа в честь разработавшего ее М. Плоткина (см. [88]). Коды, порождаемые множествами S_1, S_2, S_3, ..., называются ***кодами Рида-Мюллера***. Множество S_3 — это $(64, 32, 16)$-код, который использовался для исправления ошибок на изображении, переданном космическим кораблем "Mariner 9". В частности, матрицу Рида-Мюллера часто называют матрицей Адамара, определение которой будет дано ниже.

Рассмотрим матрицы, определенные рекурсивно следующим образом:

$$A_1 = [0];$$
$$A_{2n} = \begin{bmatrix} A_n & A_n \\ A_n & \bar{A}_n \end{bmatrix},$$

где матрица $\bar{A}_n$ определена соотношением $\bar{A}_{ij} = A_{ij}$ для $1 \le i, j \le n$. Следовательно,

$$A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix};$$

$$A_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

Пусть H_n обозначает матрицу, полученную из матрицы A_n заменой 0 на 1, а 1 — на -1, для каждого элемента A_{ij} для $1 \leq i, j \leq n$. Таким образом,

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Матрица H_n обладает свойством: $H_n H_n^t = nI$, где I — единичная матрица. Матрицы, обладающие таким свойством, называются ***матрицами Адамара***. Порождаемый ими код называется ***кодом Адамара***.

■ УПРАЖНЕНИЯ

1. Пусть $G_H^{\perp}$ — матрица

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Найдите G_H.

2. Пусть $G_H^{\perp}$ — матрица

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Найдите матрицу G_H.

3. Найдите расстояние между строками 110010101 и 010101111.

4. Найдите расстояние между строками 110011001 и 111100001.

5. Найдите три строки, ортогональные к строке 110011001.

6. Найдите три строки, ортогональные к строке 010101111.

7. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

а) найдите матрицу $G_H^{\perp}$;

б) используя только матрицу $G_H^{\perp}$, исправьте ошибки (если они имеются) в переданных сообщениях 1110110, 1011000, 1101100 и 1111110.

8. Для заданной порождающей матрицы

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

а) найдите матрицу $G_H^{\perp}$;

б) используя только матрицу $G_H^{\perp}$, исправьте ошибки (если они имеются) в переданных сообщениях 1111110, 1011001, 1101100 и 1111101.

9. Для каждой матрицы Хемминга из упражнений 7 и 8, если строка 1001110 передана как 1011010, то каков исправленный код? Как это повлияло на исходное закодированное слово?

10. Какие из приведенных ниже матриц являются матрицами Хемминга?

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

11. Докажите теорему 18.4: Функция расстояния Хемминга обладает следующими свойствами:

а) для строк c и c' расстояние $\delta(c, c') = 0$ тогда и только тогда, когда $c = c'$;

б) для строк c и c' расстояние $\delta(c, c') = \delta(c', c)$.

12. Постройте код Рида-Мюллера, S_1.

13. Постройте матрицы Адамара H_2, H_4, H_8 и H_{16}.

ГЛАВА

19

ПЕРЕЧИСЛЕНИЕ ЦВЕТОВ

19.1. ТЕОРЕМА БЕРНСАЙДА

Вначале рассмотрим количество способов раскраски вершин правильного n-угольника с помощью m цветов. Поскольку случай $m = 1$ не представляет интереса, начнем с двух цветов, например, красного (R) и синего (B). Сначала рассмотрим квадрат, т.е. будем раскрашивать двумя цветами вершины квадрата. Назовем раскрашивание вершин просто ***раскраской***. Таким образом, если выкрасить верхнюю левую вершину красным цветом, а все остальные — синим, то это будет примером раскраски. Если не допускать движений квадрата, то имеем два способа выбора для каждой из вершин, поэтому существует $2^4 = 16$ способов раскрасить квадрат. Чтобы усложнить задачу, разрешим вращение квадрата и поворот вокруг осей, проходящих через противоположные вершины или середины противолежащих сторон. Первый класс движений назовем ***вращениями***, а второй класс — ***отражениями***. Все раскраски, которые можно получить из данной раскраски вершин квадрата путем вращения и отражения, будем считать эквивалентными. Таким образом, все раскраски, изображенные на рис. 19.1, эквивалентны, поскольку каждая может быть получена из другой вращением.

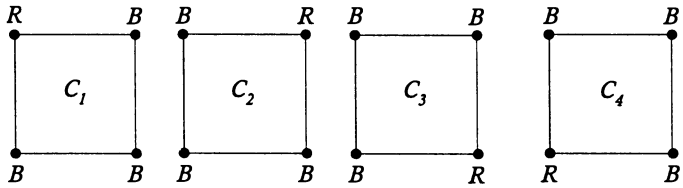


Рис. 19.1

Это также справедливо для раскрасок, изображенных на рис. 19.2.

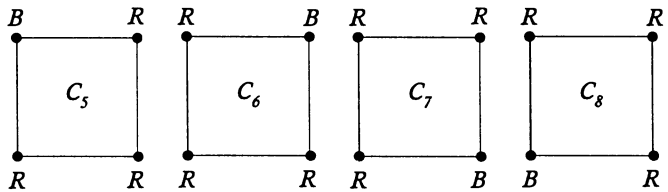


Рис. 19.2

Раскраска, изображенная на рис. 19.3, эквивалентна только самой себе, т.к. любое вращение или отражение дает в результате ту же раскраску. Аналогично, раскраска, изображенная на рис. 19.4, эквивалентна только самой себе.

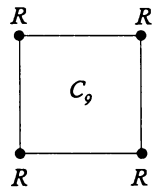


Рис. 19.3

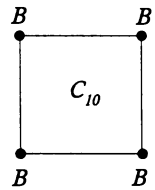


Рис. 19.4

Следующие две раскраски на рис. 19.5 эквивалентны, поскольку каждая получена из другой вращением на 90°, и никакие другие раскраски в результате вращения или отражения не могут быть получены.

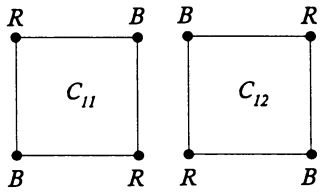


Рис. 19.5

Наконец, легко видеть, что все раскраски множества на рис. 19.6 эквивалентны, так как каждая раскраска может быть получена из другой вращением.

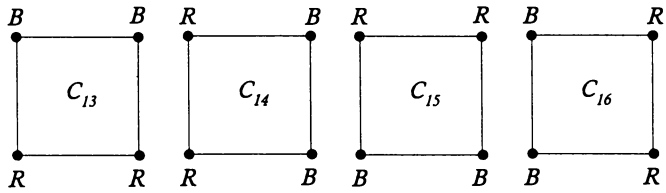


Рис. 19.6

Можно заметить, что в случае двухцветных раскрасок квадрата отражения фактически не нужны, т.к. всевозможные эквивалентные раскраски можно получить в результате одних лишь вращений.

Вспомним раздел 9.4, в котором вращения и отражения квадрата описаны подгруппой S_4 (группой перестановок четырех элементов). Эта группа называется ***октической группой***, или ***группой симметрий квадрата***.

Для квадрата, изображенного на рис. 19.7,

$$\delta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} \quad \text{и} \quad \delta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

являются отражениями, которые меняют местами вершины 1 и 3 и вершины 2 и 4, соответственно.

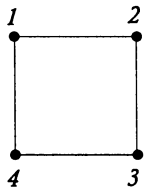


Рис. 19.7

Перестановки

$$\phi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \quad \text{и} \quad \phi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

отражают квадрат относительно горизонтальной оси, проходящей через центр квадрата, и относительно вертикальной оси, проходящей через центр квадрата. Перестановки ρ_1, ρ_2 и ρ_3 вращают квадрат по часовой стрелке соответственно на 90°, 180° и 270°.

$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}; \quad \rho_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}; \quad \rho_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

I — тождественная перестановка

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

Заметим, что группа симметрий правильного n-угольника всегда содержит $2n$ преобразований. Всегда имеется n вращений. Если число n нечетное, то имеется n отражений относительно осей, проходящих через каждую вершину и середину противоположной стороны. Если число n четное, то имеется $\frac{n}{2}$ отражений относительно осей, проходящих через противоположные вершины, и $\frac{n}{2}$ отражений относительно осей, проходящих через середины противоположных сторон многоугольника. В каждом случае получаем n вращений и n отражений.

Возвращаясь к двухцветным раскраскам квадрата, отметим, что для каждой перестановки σ можно образовать перестановку $\hat{\sigma}$ раскрасок, такую что $\hat{\sigma}(C_i) = C_j$, где C_j — раскраска, полученная в результате того, что квадрат с раскраской C_i преобразуется перестановкой σ. Например, $\hat{\rho}_1(C_1) = C_2$ и $\hat{\rho}_1(C_{11}) = C_{12}$. Для простоты будем писать вместо $\hat{\sigma}$ просто σ и использовать σ как для обозначения перестановок вершин, так и для обозначения перестановок цветов. Ранее мы говорили об эквивалентных раскрасках. Теперь, согласно такой эквивалентности раскрасок, введем отношение эквивалентности.

Определим отношение R на множестве раскрасок соотношением $C_i R C_j$, если в циклической группе существует перестановка σ такая, что $\sigma(C_i) = C_j$. Покажем, что R — отношение эквивалентности. Рефлексивность отношения R очевидна, поскольку $I(C_i) = C_i$ для любой раскраски C_i. Если $C_i R C_j$, то существует перестановка σ такая, что $\sigma(C_i) = C_j$. Но поскольку перестановка σ^{-1} принадлежит октической группе и $\sigma^{-1}(C_j) = C_i$, то $C_j R C_i$. Если $C_i R C_j$ и $C_j R C_k$, то существуют такие перестановки σ и σ', что $\sigma(C_i) = C_j$ и $\sigma(C_j) = C_k$. Следовательно, $\sigma'\sigma(C_i) = C_k$, и, поскольку $\sigma'\sigma$ также является перестановкой из

группы симметрий квадрата, имеем C_iRC_k. Следовательно, R — отношение эквивалентности, которое можно аналогичным образом получить для произвольного множества K раскрасок на множестве S и группы перестановок S. Это дает нам следующую теорему.

ТЕОРЕМА 19.1. Пусть K — множество раскрасок на множестве S, и G — группа перестановок S. Пусть отношение R на K определено следующим образом: C_iRC_j, если существует перестановка $\sigma \in G$ такая, что $\sigma(C_i) = C_j$. Отношение R есть отношение эквивалентности.

Возвращаясь к двухцветным раскраскам квадрата, отметим, что когда впервые упоминалась эквивалентность раскрасок, мы имели в виду, что одна раскраска может быть получена из другой посредством вращения и отражения. Говоря об эквивалентных раскрасках в упомянутой выше теореме, мы имеем в виду, что одна раскраска может быть получена из другой посредством перестановки. Поскольку каждую перестановку можно рассматривать как вращение или отражение, становится очевидным, что это, по сути, одна и та же эквивалентность.

Необходимо помнить, что наша цель — найти количество "различных" раскрасок, где раскраски, принадлежащие одному классу эквивалентности, считаются "одинаковыми". Следовательно, нас фактически интересует количество классов эквивалентности, поскольку оно совпадает с количеством "различных" раскрасок. Например, рассмотрим круглый стол, за которым сидит n человек. Предположим, что оба способа разместить людей являются одинаковыми, если у каждого из сидящих тот же сосед справа и тот же сосед слева. Следовательно, если стол вращать, то взаимное расположение людей за столом при этом не меняется. Мы не будем рассматривать отражения, т.к. в результате сосед слева и сосед справа поменяются местами, а вся еда окажется на полу. Существует $n!$ способов разместить n человек за столом. Поскольку два способа размещения людей эквивалентны, или одинаковы, если один способ получен из другого вращением и всего имеется n вращений, то количество различных способов разместить людей равно количеству классов эквивалентности. Но число классов эквивалентности равно $n!$ (число способов разместить n человек за столом), деленному на n (число таких размещений в каждом классе эквивалентности), и поэтому равно $\frac{n!}{n} = (n-1)!$. Такой метод подсчета оправдывает себя, потому что каждое вращение дает новый элемент в классе эквивалентности. На вопрос, куда подевались цвета, ответим, что каждого человека можно рассматривать как представителя своего собственного цвета, поэтому фактически рассматривается n-угольник, в котором все вершины окрашены в различные цвета.

Снова, возвращаясь к двухцветной раскраске квадрата, заметим, что не каждая перестановка дает другую раскраску. Например, $\rho_2(C_{11}) = C_{11} = \delta_1(C_{11}) = \delta_2(C_{11})$. На самом деле нам уже известно, что классы эквивалентности различны по размеру. Поэтому нужен иной подход.

Для фиксированной раскраски C пусть $G_C = \{\sigma : \sigma(C) = C\}$, так что G_C — множество всех перестановок в G, которые не меняют C. Множество G_C называется ***стабилизатором***, или ***стационарной подгруппой*** C.

ТЕОРЕМА 19.2. Для фиксированной раскраски C стабилизатор G_C есть подгруппа группы G, группы перестановок.

ДОКАЗАТЕЛЬСТВО. Для доказательства, что G_C — группа, покажем сначала, что если $\sigma_i, \sigma_j \in G_C$, то $\sigma_i \circ \sigma_j \in G_C$. Но $\sigma_i(C) = C$ и $\sigma_j(C) = C$, а значит,

$$\sigma_i \circ \sigma_j(C) = \sigma_i(\sigma_j(C)) = \sigma_i(C) = C$$

и $\sigma_i \circ \sigma_j \in G_C$. Очевидно, что тождественная перестановка I принадлежит G_C, поскольку $I(C_i) = C_i$ для каждой раскраски C_i. Если $\sigma \in G_C$, то $\sigma(C) = C$ и, следовательно, $\sigma^{-1}(C) = C$, т.к.

$$C = I(C) = \sigma^{-1} \circ \sigma(C) = \sigma^{-1}(\sigma(C)) = \sigma^{-1}(C)\,.$$

Следовательно, $\sigma^{-1} \in G_C$. Отсюда G_C — группа. ■

Пусть C и C' принадлежат одному классу эквивалентности. Количество перестановок в G, которые отображают C в C', равно $|G_C|$. Для доказательства положим $\sigma_0(C) = C'$ и $\sigma(C) = C$. Тогда

$$\sigma_0 \circ \sigma(C) = \sigma_0(\sigma(C)) = \sigma_0(C) = C'\,.$$

Таким образом, при фиксированном σ_0 для каждой перестановки σ такой, что $\sigma(C) = C$, имеется перестановка $\sigma_0 \circ \sigma$, которая переводит C в C'. Для каждой σ такой, что $\sigma(C) = C$, положим $f(\sigma) = \sigma_0 \circ \sigma$. Тогда f отображает каждую перестановку, переводящую C в саму себя, в перестановку, которая переводит C в C'. Более того, если σ_1 и σ_2 — две перестановки, которые переводят C в себя, и $\sigma_0 \circ \sigma_1 = \sigma_0 \circ \sigma_2$, то

$$\sigma_0^{-1} \circ (\sigma_0 \circ \sigma_1) = \sigma_0^{-1} \circ (\sigma_0 \circ \sigma_2)\,.$$

Поэтому

$$(\sigma_0^{-1} \circ \sigma_0) \circ \sigma_1 = (\sigma_0^{-1} \circ \sigma_0) \circ \sigma_2\,,$$

что дает

$$I \circ \sigma_1 = I \circ \sigma_2$$

и

$$\sigma_1 = \sigma_2\,.$$

Следовательно, функция f инъективна. Пусть σ переводит C в C'. Тогда

$$f(\sigma_0^{-1} \circ \sigma) = \sigma_0 \circ (\sigma_0^{-1} \circ \sigma) = (\sigma_0 \circ \sigma_0^{-1}) \circ \sigma = I \circ \sigma = \sigma\,,$$

так что f — сюръекция. Следовательно, f — взаимно однозначное соответствие, поэтому существует одинаковое количество перестановок, которые переводят C в C', и тех, которые отображают C в себя. Следовательно, если C_1 и C_2 принадлежат E, тому же классу эквивалентности, что и C, то количество перестановок, которые переводят C в C_1, равно количеству перестановок, которые переводят C в C_2, а именно, $|G_C|$. Это число называется ***кратностью*** класса эквивалентности E. Отсюда следует, что $|G|$, количество перестановок в G, равно $|G_C| \cdot |E|$, где $|E|$ — количество раскрасок в E, или

$$|G_C| \cdot |E| = |G|\,.$$

Последнее соотношение можно также переписать в виде

$$|G| = \sum_{C \in E} |G_C|,$$

и тогда, если N — количество классов эквивалентности,

$$N \cdot |G| = \sum_{E} \sum_{C \in E} |G_C| = \sum_{C \in K} |G_C|,$$

где K — множество раскрасок. Следовательно, количество различных раскрасок, т.е. число N классов эквивалентности, дает следующая формула

$$N = \frac{\sum\limits_{C \in K} |G_C|}{|G|} = \frac{1}{|G|} \sum_{C \in K} |G_C|.$$

Для каждой перестановки σ пусть $\varphi(\sigma)$ равно количеству раскрасок, которые σ не изменяют. Заметим, что если для каждой раскраски найти число перестановок, не изменяющих ее, и просуммировать эти числа по всем раскраскам, то получим тот же результат, как если бы мы находили количество раскрасок, которые не изменяет каждая перестановка. Поэтому, суммируя по перестановкам, получаем

$$\sum_{C \in K} |G_C| = \sum_{\sigma \in G} \varphi(\sigma)$$

и

$$N = \frac{1}{|G|} \sum_{C \in K} |G_C| = \frac{1}{|G|} \sum_{\sigma \in G} \varphi(\sigma).$$

Подведем итог в приведенной ниже теореме.

ТЕОРЕМА 19.3. (Лемма Бернсайда о подсчете) Если K — число раскрасок над множеством S, G — группа перестановок на множестве S и N — число классов эквивалентности (т.е. различных раскрасок), то

$$N = \frac{1}{|G|} \sum_{C \in K} |G_C| = \frac{1}{|G|} \sum_{\sigma \in G} \varphi(\sigma).$$

■ УПРАЖНЕНИЯ

1. Определите группу симметрий равностороннего треугольника.
2. Определите группу симметрий правильного пятиугольника.
3. Определите группу симметрий правильного восьмиугольника.
4. Определите группу симметрий тетраэдра.
5. Если стержень украшен пятью полосками, каждая из которых может быть одного из двух цветов, то сколькими способами можно раскрасить стержень? (Напомним, что стержень можно перевернуть.)
6. Сколькими способами можно раскрасить равносторонний треугольник, используя два цвета?
7. Сколькими способами, используя два цвета, можно раскрасить прямоугольник, который не является квадратом?

19.2. ТЕОРЕМА ПОЙА

Сделав первый шаг, будем развивать дальше теорию подсчета числа раскрасок. С этой целью рассмотрим теорему Пойа, которая дает производящую функцию для нахождения числа раскрасок неориентированной фигуры с использованием заданного количества цветов. При этом нам будет особенно полезна последняя часть уравнения Бернсайда. Если удастся найти простой способ определения количества раскрасок, которые остаются неизменными при каждой перестановке, то мы сможем просуммировать эти числа и разделить сумму на число перестановок. В разделе 9.5 показано, что каждую перестановку можно представить как произведение непересекающихся циклов. Если вернуться к двухцветной раскраске квадрата, то

$$\phi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

можно выразить как

$$(14)(23).$$

В цикле раскраска будет оставаться неизменной, если цвет каждой вершины будет совпадать с цветом той вершины, в которую она переходит. Поэтому все вершины в цикле должны быть одного цвета. Если ϕ_1 должна оставить раскраску неизменной, то вершины 1 и 4 должны быть одного цвета и вершины 2 и 3 должны быть одного цвета. Следовательно, вершины 1 и 4 могут быть красными или синими, и вершины 2 и 3 могут быть красными или синими. Это значит, что есть четыре раскраски, которые остаются неизменными при перестановке ϕ_1. Это раскраски C_9, C_{10}, C_{14} и C_{16}. Поскольку для каждого цикла есть выбор из двух цветов, то в общем случае число раскрасок, которые не меняются при перестановке σ, равно $2^{Cy(\sigma)}$, где $Cy(\sigma)$ — количество циклов в перестановке σ, включая циклы длины 1. Тождественную перестановку I можно представить в виде $(1)(2)(3)(4)$. Она имеет четыре цикла, поэтому оставляет неизменными 2^4 раскрасок. В приведенной ниже таблице указана каждая перестановка, ее выражение через циклы, ее цикловая структура, количество циклов (N_1), раскраски, которые она не меняет (инвариантные раскраски), и количество таких раскрасок (N_2). Для записи цикловой структуры будем использовать обозначение c_m^n, где n — число циклов длины m в перестановке. Таким образом, цикловая структура для ϕ_1 имеет вид c_2^2, т.е. ϕ_1 выражается через два цикла длины 2.

Перестановка	Цикловая структура	N_1	Инвариантные раскраски	N_2
$\delta_1 = (13)(2)(4)$	$c_1^2 c_2$	3	$C_2, C_4, C_6, C_8, C_9, C_{10}, C_{11}C_{12}$	8
$\delta_2 = (1)(3)(24)$	$c_1^2 c_2$	3	$C_1, C_3, C_5, C_7, C_9, C_{10}, C_{11}C_{12}$	8
$\phi_1 = (14)(23)$	c_2^2	2	$C_9, C_{10}, C_{14}C_{16}$	4
$\phi_2 = (12)(34)$	c_2^2	2	$C_9, C_{10}, C_{13}C_{15}$	4
$\rho_1 = (1234)$	c_4	1	C_9, C_{10}	2
$\rho_2 = (13)(24)$	c_2^2	2	$C_9, C_{10}, C_{11}C_{12}$	4
$\rho_3 = (1432)$	c_4	1	C_9, C_{10}	2
$I = (1)(2)(3)(4)$	c_1^4	4	$C_1, C_2, C_3, C_4, \ldots, C_{16}$	16
Всего	$c_1^4 + 2c_1^2c_2 + 3c_2^2 + 2c_4$			48

Как отмечалось ранее, для определения количества раскрасок, инвариантных относительно перестановки σ, достаточно найти $2^{Cy(\sigma)}$, где $Cy(\sigma)$ — количество циклов в перестановке σ. Затем нужно просуммировать по всем перестановкам и найти $\sum_{\sigma\in G}\varphi(\sigma)$. Другой метод — взять цикловую форму и присвоить каждому циклу значение 2. Например, $\delta_2 = c_1^2c_2$ оставляет неизменными $2^2\cdot 2 = 8$ раскрасок. Затем нужно найти сумму $\sum_{\sigma\in G}\varphi(\sigma)$. Третий метод — сложить вместе все цикловые формы. Согласно таблице имеем

$$c_1^4 + 2c_1^2c_2 + 3c_2^2 + 2c_4\,.$$

Если теперь присвоить каждому циклу значение 2, то получим

$$2^4 + 2\cdot 2^2\cdot 2 + 3\cdot 2^2 + 2\cdot 2 = 48\,.$$

Итак, всего имеется 48 раскрасок, которые не меняются при перестановках. Если теперь это число разделить на 8, количество перестановок, то получим

$$\frac{1}{|G|}\sum_{\sigma\in G}\varphi(\sigma) = \frac{48}{8} = 6$$

различных раскрасок.

Далее предположим, что имеется три цвета. Используя аналогичные рассуждения и таблицу, можно вычислить $3^{Cy(\sigma)}$, где $Cy(\sigma)$ — количество циклов в перестановке σ, чтобы найти количество раскрасок, которые остаются неизменными при перестановке σ, а затем просуммировать эти числа и найти $\sum_{\sigma\in G}\varphi(\sigma)$. Затем снова можно разделить на 8 и получить количество различных раскрасок. Второй метод: учитывая, что уже имеется

$$c_1^4 + 2c_1^2c_2 + 3c_2^2 + 2c_4\,,$$

присвоим каждому циклу значение 3, так что сумма всех раскрасок, которые не меняются при перестановках, равна

$$3^4 + 2\cdot 3^2\cdot 3 + 3\cdot 3^2 + 2\cdot 3 = 168\,.$$

Разделив теперь на 8, получаем

$$\frac{1}{|G|}\sum_{\sigma\in G}\varphi(\sigma) = \frac{168}{8} = 21$$

различную раскраску. Пусть имеется m цветов, тогда

$$\frac{1}{|G|}\sum_{\sigma\in G}\varphi(\sigma) = \frac{m^4 + 2\cdot m^2\cdot m + 3\cdot m^2 + 2\cdot m}{8}\,.$$

Поскольку в общем случае нам нужна величина $\frac{1}{|G|}\sum_{\sigma\in G}\varphi(\sigma)$, которую можно получить из

$$\frac{1}{|G|}(c_1^4 + 2c_1^2c_2 + 3c_2^2 + 2c_4)\,,$$

то положим

$$P_G(c_1, c_2, c_3, c_4) = \frac{1}{|G|}(c_1^4 + 2c_1^2c_2 + 3c_2^2 + 2c_4)\,,$$

где c_i — цикл длины i, и назовем $P_G(c_1, c_2, c_3, c_4)$ ***цикловым индексом***.

ПРИМЕР 19.4. Ранее отмечалось, что для двухцветных раскрасок квадрата достаточно одних вращений. Предположим, что $G = \{I, \rho_1, \rho_2, \rho_3\}$. В таком случае получим таблицу.

Перестановки	Цикловая структура	N_1	Инвариантные раскраски	N_2
$\rho_1 = (1234)$	c_4	1	C_9, C_{10}	2
$\rho_2 = (13)(24)$	c_2^2	2	$C_9, C_{10}, C_{11}C_{12}$	4
$\rho_3 = (1432)$	c_4	1	C_9, C_{10}	2
$I = (1)(2)(3)(4)$	c_1^4	4	$C_1, C_2, C_3, C_4, \ldots, C_{16}$	16
Всего	$c_1^4 + c_2^2 + 2c_4$			24

(Здесь N_1 — количество циклов, N_2 — количество инвариантных раскрасок.)

Тогда

$$\frac{1}{|G|}\sum_{\sigma\in G}\varphi(\sigma) = \frac{24}{4} = 6,$$

и, как и ранее, получаем шесть различных раскрасок. Опять, если присвоить каждому циклу значение 2 в выражении

$$P_G(c_1, c_2, c_3, c_4) = \frac{1}{4}(c_1^4 + c_2^2 + 2c_4)\,,$$

то получим

$$\frac{1}{4}(2^4 + 2^2 + 2\cdot 2) = 6\,,$$

и снова получаем шесть различных раскрасок. □

Рассмотрим общий случай, когда группа перестановок G содержит $|G|$ элементов раскрашивание проводится с помощью m красок, и сумма циклов равна $a_1c_1^{e_1} + a_2c_1^{e_2} + a_3c_1^{e_3} + \cdots + a_nc_n^{e_n}$. Цикловой индекс в этом случае равен

$$P_G(c_1, c_2, c_3, \ldots, c_n) = \frac{1}{|G|}(a_1c_1^{e_1} + a_2c_1^{e_2} + a_3c_1^{e_3} + \cdots + a_nc_n^{e_n})\,,$$

а количество различных раскрасок равно $P_G(m, m, m, \ldots, m)$. Теперь перейдем от подсчета структур к рассмотрению структуры раскрасок, сохраняющихся при конкретных перестановках. Опять начнем с двухцветной раскраски квадрата. Рассмотрим цикловую стркутуру $c_1^2c_2$ перестановки $\delta_1 = (13)(2)(4)$. Сначала посмотрим на элемент структуры c_1^2, который представляет два цикла длины 1. Для двух циклов имеется четыре способа выбора цветов, при которых раскраска не будет меняться: вершины в обоих циклах могут быть окрашены только в красный или только в синий цвет, вершины в одном цикле могут быть окрашены в

красный, а в другом — в синий, и наоборот. Очевидно, существует только один способ раскраски вершин в обоих циклах красным и один способ раскрасить их в обоих циклах синим. Если вершины в одном цикле красные, а в другом — синие, то имеются две возможности выбора синих вершин; тогда в другом цикле вершины должны быть красные. Запишем эту структуру в форме $bb + 2br + rr$ (b — синий, r — красный), где знак $+$ можно читать как "или". Для удобства запишем эту структуру как $b^2 + 2br + r^2$. Заметим, что алгебраически это $(b+r)^2$. Другой цикл задается c_2, что представляет один цикл длины 2. Поскольку это единственный цикл длины 2, то в нем обе вершины должны быть окрашены в красный или обе — в синий. Запишем эту структуру как $bb + rr$, или $b^2 + r^2$. Поскольку структура раскраски двух вершин есть $(b + r)^2$, а для других двух вершин эта структура есть $b^2 + r^2$, то структура раскрасок, не изменяющихся при перестановке δ_1, может быть описана как $(b + r)^2(b^2 + r^2)$.

Допустим, что имеется цикл c_m^n. Это произведение n циклов длины m. Пусть R представляет раскраску цикла красным, а B представляет раскраску цикла синим. Таким образом, $R^j B^{n-j}$ представляет раскраску j циклов красным и $n-j$ циклов — синим. Количество способов раскраски j циклов красным и $n-j$ циклов синим равно количеству вариантов выбора j циклов из n циклов для окраски в красный цвет. Это равно $C(n, j) = \binom{n}{j}$. Поэтому будем использовать запись $\binom{n}{j} R^j B^{n-j}$ для обозначения раскрасок с j красными и $n - j$ синими циклами. Но это не что иное, как коэффициент при $R^j B^{n-j}$ в разложении $(R + B)^n$. Таким образом, можно представить раскраски цикла c_m^n как $(R+B)^n$. Поскольку каждая раскраска цикла состоит из раскраски m вершин, положим $R = r^m$ и $B = b^m$. Теперь структуры раскрасок для c_m^n можно описать выражением $(r^m + b^m)^n$. В результате для каждой перестановки получим следующее множество раскрасок, которое назовем ***перечнем*** раскрасок для перестановки.

Перестановка	Цикловая структура	Перечень инвариантных раскрасок
$\delta_1 = (13)(2)(4)$	$c_1^2 c_2$	$(r + b)^2(r^2 + b^2)$
$\delta_2 = (1)(3)(24)$	$c_1^2 c_2$	$(r + b)^2(r^2 + b^2)$
$\phi_1 = (14)(23)$	c_2^2	$(r^2 + b^2)^2$
$\phi_2 = (12)(34)$	c_2^2	$(r^2 + b^2)^2$
$\rho_1 = (1234)$	c_4	$(r^4 + b^4)$
$\rho_2 = (13)(24)$	c_2^2	$(r^2 + b^2)^2$
$\rho_3 = (1432)$	c_4	$(r^4 + b^4)$
$I = (1)(2)(3)(4)$	c_1^4	$(r + b)^4$
Всего	$c_1^4 + 2c_1^2 c_2 + 3c_2^2 + 2c_4$	$8r^4 + 8r^3 b + 16r^2 b^2 + 8rb^3 + 8b^4$

Последняя строка в столбце перечней раскрасок получается, если раскрыть скобки в выражениях для всех перечней и результаты сложить.

Ранее было показано, что для раскраски C количество перестановок в G, обозначенное $|G|$, равно $|G_C| \cdot |E|$, где $|E|$ — количество раскрасок в классе эквивалентности E, содержащем C, и $|G_C|$ — количество перестановок, которые не

меняют C. Поскольку каждая перестановка переводит раскраску в раскраску с той же цикловой структурой, то в сумму наших перечней раскраски, эквивалентные C и описываемые той же структурой, вошли $|G|$ раз. Таким образом, для определения перечня неэквивалентных структур раскрасок мы определяем общий перечень инвариантных раскрасок как $|G| = 8$ и для нашего перечня структур раскрасок получаем

$$\frac{1}{8}(8r^4 + 8r^3b + 16r^2b^2 + 8rb^3 + 8b^4) = r^4 + r^3b + 2r^2b^2 + rb^3 + b^4 .$$

Для получения полного перечня структур мы брали инвариантные раскраски из цикловой структуры каждой перестановки, а затем суммировали их. Можно было сначала сложить цикловые структуры, как в колонке 2 приведенной выше таблицы, а затем для получения такого же результата перевести циклы в раскраски. Таким образом, для суммы

$$c_1^4 + 2c_1^2c_2 + 3c_2^2 + 2c_4$$

получаем полный перечень структур

$$(r^4 + b^4) + 2(r + b)^2(r^2 + b^2)(r + b)^2 + 3(r^2 + b^2)^2 + 2(r^4 + b^4) ,$$

что после раскрытия скобок и приведения подобных членов равно

$$8r^4 + 8r^3b + 16r^2b^2 + 8rb^3 + 8b^4 .$$

Таким образом, если сумма циклов имеет вид $a_1c_1^{e_1} + a_2c_1^{e_2} + a_3c_1^{e_3} + \cdots + a_nc_n^{e_n}$, то перечень неэквивалентных структур раскрасок имеет вид

$$P_G(r + b, r^2 + b^2, \ldots, r^n + b^n) = \\ = \frac{1}{|G|}(a_1(r + b)^{e_1} + a_2(r^2 + b^2)^{e_2} + \cdots + a_n(r^n + b^n)^{e_n}) .$$

Если бы мы использовали три цвета: красный (r), белый (w) и синий (b), то перечень неэквивалентных структур раскрасок имел бы вид

$$P_G(r + w + b, r^2 + w^2 + b^2, \ldots, r^n + w^n + b^n) .$$

В общем случае имеет место следующая теорема.

ТЕОРЕМА 19.5. (Теорема Пойа о подсчете) Если задано множество S, цикловой индекс $P_G(c_1, c_2, \ldots, c_n)$ и цвета $k_1, k_2, k_3, \ldots, k_m$, то перечень неэквивалентных структур раскрасок имеет вид

$$P_G(\sum_{i=1}^{n} k_i, \sum_{i=1}^{n} k_i^2, \ldots, \sum_{i=1}^{n} k_i^m) .$$

ПРИМЕР 19.6. Снова рассмотрим двухцветные раскраски квадрата с использованием только вращений. Пусть $G = \{I, \rho_1, \rho_2, \rho_3\}$. В таком случае получаем таблицу

Перестановка	Цикловая структура	Перечень инвариантных раскрасок
$\rho_1 = (1234)$	c_4	$(r^4 + b^4)$
$\rho_2 = (13)(24)$	c_2^2	$(r^2 + b^2)^2$
$\rho_3 = (1432)$	c_4	$(r^4 + b^4)$
$I = (1)(2)(3)(4)$	c_1^4	$(r + b)^4$
Всего	$c_1^4 + c_2^2 + 2c_4$	$4r^4 + 4r^3b + 8r^2b^2 + 4rb^3 + 4b^4$

$$P_G(r + b, r^2 + b^2, r^3 + b^3, r^4 + b^4) = \frac{1}{4}((r + b)^4 + (r^2 + b^2)^2 + 2(r^4 + b^4)) = \\ = r^4 + r^3b + 2r^2b^2 + rb^3 + b^4,$$

что дает совпадение перечня неэквивалентных структур раскрасок с уже рассмотренным случаем, в котором была использована вся группа симметрий. □

ПРИМЕР 19.7. Пусть S — вершины треугольника, а G — группа симметрий треугольника. Найдем перечень неэквивалентных структур раскрасок. Пусть $\rho_1 = (123)$ и $\rho_2 = (132)$ — вращения вершин. Пусть $\delta_1 = (1)(23)$, $\delta_2 = (13)(2)$ и $\delta_3 = (12)(3)$ — отражения, а $I = (1)(2)(3)$ — тождественное преобразование. В таком случае имеем таблицу

Перестановка	Цикловая структура	Перечень инвариантных раскрасок
$\rho_1 = (123)$	c_3	$(r^3 + b^3)$
$\rho_2 = (132)$	c_3	$(r^3 + b^3)$
$\delta_1 = (1)(23)$	c_1c_2	$(r + b)(r^2 + b^2)$
$\delta_2 = (13)(2)$	c_1c_2	$(r + b)(r^2 + b^2)$
$\delta_2 = (12)(3)$	c_1c_2	$(r + b)(r^2 + b^2)$
$I = (1)(2)(3)(4)$	c_1^3	$(r + b)^3$
Всего	$c_1^3 + 3c_1c_2 + 2c_3$	$6r^3 + 6r^2b + 6rb^2 + 6b^3$

$$P_G(r + b, r^2 + b^2, r^3 + b^3) = \frac{1}{6}((r + b)^3 + 3(r^2 + b^2)(r + b) + (r^3 + b^3)) = \\ = r^3 + r^2b + rb^2 + b^3.$$ □

■ УПРАЖНЕНИЯ

В приведенных ниже заданиях допускается, что все вращения и отражения возможны.

1. На браслете пять бусинок. Если их нужно раскрасить двумя цветами, то сколько раскрасок существует? Сколько существует трехцветных раскрасок?
2. Сколькими способами можно раскрасить клетки шахматной доски 3×3?
3. Сколько существует способов раскраски равностороннего треугольника тремя цветами?

4. Сколько существует способов раскраски куба двумя цветами?
5. Сколько существует способов раскраски правильного пятиугольника двумя цветами?
6. Сколько существует способов раскраски правильного пятиугольника тремя цветами?
7. Сколько существует способов раскраски правильного восьмиугольника двумя цветами?
8. Найдите структурный перечень двухцветных раскрасок браслета с семью бусинками.
9. Найдите структурный перечень двухцветных раскрасок прямоугольника, который не является квадратом.
10. Найдите структурный перечень трехцветных раскрасок равностороннего треугольника.
11. Найдите структурный перечень двухцветных раскрасок пятиугольника.
12. Найдите структурный перечень двухцветных раскрасок тетраэдра.
13. Найдите структурный перечень двухцветных раскрасок шахматной доски размера 3×3.
14. Найдите структурный перечень двухцветных раскрасок восьмиугольника.
15. Десять кеглей выставлены в треугольник стандартным образом. Найдите структурный перечень двухцветных раскрасок этих кеглей без ограничений на движения. Найдите структурный перечень при условии, что разрешены только вращения.

ГЛАВА

20 КОЛЬЦА, ОБЛАСТИ ЦЕЛОСТНОСТИ И ПОЛЯ

20.1. КОЛЬЦА И ОБЛАСТИ ЦЕЛОСТНОСТИ

В разделах, посвященных решеткам и булевым алгебрам, рассматривались две бинарные операции и были описаны некоторые связи между ними. Здесь мы будем действовать по той же схеме, но обобщим понятие целого числа. Фактически речь пойдет об операциях $+$ и $\cdot$. Мы выделим определенные свойства целых чисел, и любое множество с указанными двумя операциями, удовлетворяющее этим свойствам, назовем кольцом. Определив кольца, на некоторое время снова вернемся к целым числам.

ОПРЕДЕЛЕНИЕ 20.1. ***Кольцом*** называется непустое множество R вместе с бинарными операциями, называемыми умножением и сложением, которые обозначаются соответственно $\cdot$ и $+$ и удовлетворяют следующим условиям.

1. Множество R замкнуто относительно сложения, т.е., если $x \in R$ и $y \in R$, то $x + y \in R$.
2. Сложение в R ассоциативно, т.е. $x + (y + z) = (x + y) + z$ для всех x, y и z из R.
3. Множество R содержит аддитивную единицу (нейтральный элемент относительно сложения, обозначаемый 0), так что $x + 0 = 0 + x = x$ для всех x из R.
4. Для каждого элемента x из R множество R содержит элемент $-x$, обратный x относительно сложения, так что $x + -x = -x + x = 0$.
5. Сложение в R коммутативно, т.е. $x + y = y + x$ для всех x и y из R.
6. Множество R замкнуто относительно умножения, т.е., если $x \in R$ и $y \in R$, то $x \cdot y \in R$.
7. Умножение в R ассоциативно, т.е. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ для всех x, y и z из R.

8. Для всех x, y и z из R выполняются следующие законы дистрибутивности:

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z)$$

и

$$(y + z) \cdot x = (y \cdot x) + (z \cdot x).$$

Если во множестве R существует элемент 1 (мультипликативная единица, или нейтральный элемент относительно умножения) такой, что $1 \cdot r = r \cdot 1 = r$ для всех r из R, то множество R называется ***кольцом с единицей***. Если $r \cdot r' = r' \cdot r$ для всех r и r' из R, то множество R называется ***коммутативным кольцом***.

Заметим, что кольцо R является группой относительно сложения и полугруппой относительно умножения. Примером колец относительно обычных операций сложения и умножения могут служить целые, действительные, рациональные и комплексные числа. Другими примерами колец являются множество $(n \times n)$-матриц для фиксированного целого числа n и множество полиномов — оба относительно обычных операций сложения и умножения. Среди перечисленных выше колец только матрицы не образуют коммутативное кольцо.

ОПРЕДЕЛЕНИЕ 20.2. ***Областью целостности*** называется коммутативное кольцо с единицей, не совпадающей с 0, так что из условия $ab = 0$ следует $a = 0$ или $b = 0$.

Множества целых, рациональных и действительных числа являются областями целостности. Множество (2×2)-матриц не является областью целостности, поскольку

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

т.е. произведение двух ненулевых матриц может быть нулевой матрицей.

Области целостности обладают свойством сокращения относительно умножения, которое означает, что если $ab = ac$ и $a \neq 0$, то $b = c$. Это очень важно, т.к. кольца не обязательно содержат взаимно обратные элементы, поэтому не всегда возможно умножить обе части уравнения $ab = ac$ на a^{-1} и получить $b = c$. Доказательство следующей теоремы оставляем читателю.

ТЕОРЕМА 20.3. Пусть R — коммутативное кольцо с единицей. Кольцо R является областью целостности тогда и только тогда, когда из $ab = ac$ следует, что $b = c$ для всех b, c и ненулевых элементов a из R.

Теперь проверим, является ли множество классов вычетов Z_n кольцом или областью целостности. Нетрудно убедиться, что Z_n всегда является кольцом. Ранее было показано, что Z_n удовлетворяет всем законам кольца, за исключением законов дистрибутивности. Выполнение законов дистрибутивности также легко показать. Рассмотрев, однако, Z_6, можно обнаружить, что $[3] \odot [2] = [0]$, и, следовательно, множество Z_6 не является областью целостности. На самом деле, если

число n не простое, например, $n = pq$, то $[p] \odot [q] = [0]$, поэтому Z_n не будет областью целостности.

Если n — простое число, то для заданного целого числа $a \not\equiv 0 \pmod{n}$, не равного нулю, из теоремы 10.6 следует, что сравнение $ax \equiv 1 \pmod{n}$ имеет решение, например, a'. Следовательно, имеем $[a] \odot [a'] = [1]$, и для каждого элемента существует обратный элемент. Поэтому, если $ab = 0$ и $a \neq 0$, то $b = a^{-1}ab = a^{-1}0 = 0$. Таким образом, если n — простое число, то множество Z_n — область целостности.

В случае, когда число n не является простым, рассмотрим подмножество $R = \{[x] : x$ — взаимно простое с $n\}$. Легко показать, что это множество образует группу относительно умножения. Произведение двух целых чисел, взаимно простых с n, является числом, взаимно простым с n. Единица, 1, есть число, взаимно простое с n; и если число b — взаимно простое с n, то сравнение $bx \equiv 1 \pmod{n}$ имеет единственное решение, поэтому для элемента $[b]$ существует обратный элемент. Однако, относительно сложения множество R не образует даже полугруппу, поскольку сумма двух чисел, взаимно простых с n, не обязательно является числом, взаимно простым с n.

ОПРЕДЕЛЕНИЕ 20.4. Пусть R и R' — кольца и пусть $f : R \to R'$ — функция из R в R'. Функция f называется ***гомоморфизмом колец*** тогда и только тогда, когда

$$f(a + b) = f(a) + f(b),$$
$$f(a \cdot b) = f(a) \cdot f(b)$$

для всех $a, b \in R$. Сложение и умножение в соответствующих кольцах одинаково определены. Если гомоморфизм колец $f : R \to R'$ — инъекция, то его называют ***мономорфизмом***. Если гомоморфизм колец $f : R \to R'$ — сюръекция, то его называют ***эпиморфизмом***. Гомоморфизм колец $f : R \to R'$ называют ***изоморфизмом***, если функция $f : R \to R'$ — биекция. Обычно при описании гомоморфизма из кольца R с единицей в кольцо R' с единицей требуется, чтобы мультипликативная единица кольца R отображалась на мультипликативную единицу кольца R'.

Например, рассмотрим кольцо целых чисел Z и кольцо

$$Z_7 = \{[0], [1], [2], [3], [4], [5], [6]\}$$

классов вычетов по модулю 7. Предположим, что функция $f : Z \to Z_7$ определена соотношением $f(a) = [a]$; другими словами, $f(a)$ — класс эквивалентности целых чисел, сравнимых с a по модулю 7. Если каждое из чисел a и b является целым, то, используя определение умножения и сложения классов эквивалентности Z_7, получим, что $f(ab) = [ab] = [a][b] = f(a)f(b)$ и $f(a+b) = [a+b] = [a]+[b] = f(a)+f(b)$. Таким образом, функция f — гомоморфизм. Ясно, что f представляет собой эпиморфизм, потому что $f(i) = [i]$ для $0 \leq i \leq 6$. Функция f не является инъекцией, поскольку $7 \equiv 0 \pmod{7}$ и, следовательно, $f(7) = [7] = [0] = f(0)$. Поэтому

функция f не является ни мономорфизмом, ни изоморфизмом. Те же самые рассуждения, несомненно, применимы к кольцам Z и Z_n для любого положительного целого числа $n > 1$.

Изоморфные кольца имеют одинаковую алгебраическую структуру и отличаются только именованием своих элементов.

ТЕОРЕМА 20.5. Для всех a из кольца R выполняется соотношение $a \cdot 0 = 0$.

ДОКАЗАТЕЛЬСТВО. Доказательство теоремы оставляем читателю. ■

ОПРЕДЕЛЕНИЕ 20.6. Подмножество R' кольца R называется ***подкольцом*** кольца R, если R' — это кольцо с той же самой операцией.

Например, целые числа образуют подкольцо кольца рациональных чисел. Рациональные числа образуют подкольцо кольца действительных чисел. Действительные числа образуют подкольцо кольца комплексных чисел. Множество $(n \times n)$-матриц с целочисленными элементами образуют подкольцо кольца $(n \times n)$-матриц с рациональными элементами. Можно легко проверить, что $\{[0], [2], [4]\}$ — это подкольцо кольца Z_6.

Дадим теперь определение поля, хотя само понятие будет использовано лишь в следующей главе.

ОПРЕДЕЛЕНИЕ 20.7. ***Полем*** называется коммутативное кольцо с единицей, не совпадающей с 0, каждый ненулевой элемент которого имеет обратный элемент относительно умножения.

Множества рациональных, действительных и комплексных чисел являются полями. Если рассмотреть только положительные действительные числа и положительные рациональные числа, то получим области целостности. Во многих случаях область целостности не является полем, но образует подкольцо поля. Ранее было показано, что в случае простого числа n каждый элемент множества классов вычетов Z_n имеет обратный элемент, поэтому Z_n — поле.

Приведенная ниже теорема устанавливает связь между областью целостности и полем.

ТЕОРЕМА 20.8. Поле является областью целостности. Конечная область целостности является полем.

ДОКАЗАТЕЛЬСТВО. Если F — поле и $ab = 0$ для $a, b \in F$, и если $a \neq 0$, то

$$a^{-1}ab = a^{-1}0 = 0\,.$$

Но $a^{-1}ab = b$, поэтому $b = 0$ и F — область целостности. Обратно, если D — конечная область целостности с элементами $0, 1, a_1, a_2, a_3, \ldots, a_n$ и $a = a_i$ для некоторого $0 \leq i \leq n$, то все элементы $a \cdot 0, a \cdot 1, a \cdot a_1, a \cdot a_2, a \cdot a_3, \ldots, a \cdot a_n$ являются различными. Если $a \cdot a_j = a \cdot a_k$, то $a \cdot (a_j - a_k) = 0$, поэтому $(a_j - a_k) = 0$ и $a_j = a_k$. Следовательно, $a \cdot 0, a \cdot 1, a \cdot a_1, a \cdot a_2, a \cdot a_3, \ldots, a \cdot a_n$ — в точности элементы области D, так что $a \cdot a_k = 1$ для некоторого k, где $0 \leq k \leq n$, поэтому a имеет обратный элемент. Таким образом, область целостности D является полем. ■

Заметим, что целые числа образуют область целостности, но не являются полем. Следовательно, не каждая область целостности является полем. Тем не менее, будет показано, что каждая область целостности вкладывается в поле. Это означает, что для заданной области целостности A существует поле F и мономорфизм из A в F.

Пусть A — область целостности. В частности, A может быть множеством целых чисел Z. Рассмотрим множество упорядоченных пар

$$P = \{(a, b) : (a, b) \in A \times A \text{ и } b \neq 0\}$$

и определим отношение $\sim$ на P следующим образом.

ОПРЕДЕЛЕНИЕ 20.9. Если пары (a, b) и (c, d) принадлежит P, то $(a, b) \sim (c, d)$ тогда и только тогда, когда $ad = bc$.

Например, если $A = Z$, то класс эквивалентности $[(2, 3)]$ содержит такие упорядоченные пары: $(2, 3), (4, 6), (6, 9), \ldots, (-2, -3), (-4, -6), \ldots$, которые соответствуют представлениям рациональных чисел в виде $2/3, 4/6, 6/9, \ldots, (-2)/(-3)$, $(-4)/(-6), \ldots$. Все они являются различными представлениями одного и того же рационального числа, $[(2, 3)]$.

ТЕОРЕМА 20.10. Отношение $\sim$ на множестве P есть отношение эквивалентности.

Доказательство этой теоремы тривиально и предоставляется читателю.

Обозначим через a/b класс эквивалентности, содержащий $(a, b) \in P$, т.е. $[(a, b)]$, а через F — множество классов эквивалентности на множестве P.

ОПРЕДЕЛЕНИЕ 20.11. Для заданных элементов a, b, c и d из области целостности A сложение на F определено соотношением $a/b + c/d = (ad + bc)/bd$, а умножение на F — соотношением $(a/b)(c/d) = ac/bd$.

Покажем, что операции сложения и умножения на F определены корректно, т.е. не зависят от выбора представителей классов эквивалентности.

ТЕОРЕМА 20.12.

а) Сложение в F определено корректно.
б) Умножение в F определено корректно.

ДОКАЗАТЕЛЬСТВО. Пусть $a/b = a'/b'$ и $c/d = c'/d'$. Имеем непосредственно, что $ab' = a'b$ и $cd' = c'd$.

а) По определению сложения имеем $a/b + c/d = (ad + bc)/bd$ и $a'/b' + c'/d' = (a'd' + b'c')/b'd'$. Необходимо показать, что $(ad + bc)/bd = (a'd' + b'c')/b'd'$ или, что равносильно, $b'd'(ad + bc) = bd(a'd' + b'c')$, но

$$\begin{aligned} b'd'(ad + bc) &= b'add' + cd'bb' = \\ &= a'bdd' + c'dbb' = \\ &= bd(a'd' + b'c'), \end{aligned}$$

где второе равенство выполняется, поскольку $ab' = a'b$ и $cd' = c'd$.

б) Из определения умножения следует, что $(a/b)(c/d) = ac/bd$ и $(a'/b')(c'/d') = a'c'/b'd'$. Необходимо показать, что $ac/bd = a'c'/b'd'$ или, что равносильно, $acb'd' = a'c'bd$. Но $acb'd' = ab'cd' = a'bc'd = a'c'bd$.

■

ТЕОРЕМА 20.13. Множество классов эквивалентности F является коммутативным кольцом с аддитивной единицей $0/1$ и мультипликативной единицей $1/1$.

Доказательство этой теоремы тривиально и предоставляется читателю.

ЛЕММА 20.14. Для a и b из A имеем $a/b = 0/1$ тогда и только тогда, когда $a = 0$.

ДОКАЗАТЕЛЬСТВО. Очевидно, что равенство $a/b = 0/1$ имеет место тогда и только тогда, когда $a = a(1) = b(0) = 0$. ■

ТЕОРЕМА 20.15. Коммутативное кольцо F является полем.

ДОКАЗАТЕЛЬСТВО. Если $a/b \neq 0$, то $a \neq 0$ и $b/a \in F$. Но

$$(a/b)(b/a) = ab/ab = 1/1 = 1\,.$$

Следовательно, для каждого ненулевого элемента из F существует обратный элемент, поэтому F — поле. ■

Пусть $f : A \to F$ определено соотношением $f(a) = a/1$. Имеем $f(ab) = ab/1 = (a/1)(b/1) = f(a)f(b)$ и $f(a+b) = (a+b)/1 = (a(1)+b(1))/1 = a/1+b/1 = f(a) + f(b)$. Кроме того, функция f — инъекция, потому что если $f(a) = f(b)$, то $a/1 = b/1$ и $a = a \cdot 1 = b \cdot 1 = b$. Следовательно, f — мономорфизм, и элементу a из A однозначно соответствует элемент $\{a/1 : a \in A\}$ из F. Таким образом, область целостности A можно рассматривать как подкольцо поля F.

ТЕОРЕМА 20.16. Отображение $f : A \to F$, определенное соотношением $f(a) = a/1$, является мономорфизмом, при этом говорят, что область целостности A ***вложена*** в поле F или что поле F содержит область целостности A.

Легко показать, что это поле F есть наименьшее поле, в которое область целостности A может быть вложена.

ОПРЕДЕЛЕНИЕ 20.17. Поле F называется ***полем частных*** области целостности A. Если A — множество целых чисел Z, то поле F — множество ***рациональных чисел***, обозначаемое, как правило, через Q.

В этом разделе развиваются алгебраические структуры целых чисел, и целые числа алгебраически определяются с точностью до изоморфизма. Отметим, что из такого рассмотрения не следует существование целых чисел, а только наличие присущих им алгебраических свойств, если сами эти числа существуют.

ОПРЕДЕЛЕНИЕ 20.18. Подмножество I кольца R называется ***идеалом*** кольца R, если

а) I — подкольцо кольца R;
б) если x принадлежит I и r принадлежит R, то $x \cdot r$ и $r \cdot x$ принадлежат I.

Например, в кольце целых чисел множество всех чисел, кратных фиксированному целому числу p, образует идеал.

ОПРЕДЕЛЕНИЕ 20.19. Пусть R — коммутативное кольцо. Идеал I кольца R называется ***главным идеалом***, порожденным элементом a, если I состоит из всех произведений a на элементы кольца R, т.е. $I = \langle a \rangle = \{ar : r \in R\}$.

ТЕОРЕМА 20.20. Каждый непустой идеал I кольца целых чисел является главным идеалом.

ДОКАЗАТЕЛЬСТВО. Для $I \neq \{0\}$ пусть p — наименьшее положительное целое число в идеале I и пусть m принадлежит I. Согласно алгоритму деления, $m = pq + r$, где $0 \leq r < p$. Поскольку $r = m - pq$, то r принадлежит I. Таким образом, поскольку $r < p$ и p — наименьшее положительное целое число в идеале I, то $r = 0$. Следовательно, каждое целое число в идеале I кратно числу p и $I = \langle p \rangle$. ■

Заметим, что при доказательстве этой теоремы используется принцип вполне упорядочения и алгоритм деления, каждый из которых выполняется для целых чисел.

ПРИМЕР 20.21. Рассмотрим кольцо Z целых чисел и два главных идеала, порожденных целыми числами 8 и 12:

$$\begin{aligned} \langle 8 \rangle &= \{8r : r \in Z\} = \\ &= \{\ldots, -24, -16, -8, 0, 8, 16, 24, \ldots\}; \\ \langle 12 \rangle &= \{12s : s \in Z\} = \\ &= \{\ldots, -24, -12, 0, 12, 24, \ldots\}. \end{aligned}$$

Пересечение множеств $\langle 8 \rangle \cap \langle 12 \rangle$ есть множество

$$\{\ldots, -48, -24, 0, 24, 48, \ldots\},$$

которое является главным идеалом, порожденным целым числом 24. Заметим, что 24 — наименьшее общее кратное чисел 8 и 12. Вообще,

$$\langle a \rangle \cap \langle b \rangle = \langle \text{НОК}(a, b) \rangle .$$

Результаты, имеющие отношение к этому примеру, сформулированы в приведенной ниже теореме. Доказательство предоставляется читателю. □

ТЕОРЕМА 20.22. Если s и t — ненулевые целые числа и $\langle s \rangle$ и $\langle t \rangle$ — соответствующие главные идеалы в кольце Z, то

а) если $\langle s \rangle \subseteq \langle t \rangle$, то $t \mid s$;
б) $\langle s \rangle \cap \langle t \rangle = \langle u \rangle$, где $u = \text{НОК}(s, t)$.

ПРИМЕР 20.23. Если $\langle a, b \rangle$ — наименьший идеал, содержащий a и b, то

$$\langle a, b \rangle = \langle \text{НОД}(a, b) \rangle .$$

Например,

$$\langle 8, 12\rangle = \{-16, -12, -8, -4, 0, 4, 8, 12, 16, \ldots\} = \langle \text{НОД}(8, 12)\rangle \ ,$$

поскольку $\text{НОД}(8, 12) = 4$ и

$$\langle 4\rangle = \{\ldots, -8, -4, 0, 4, 8, \ldots\} \, .$$

Пусть I — наименьший идеал, содержащий целые числа a и b. Тогда каждый элемент идеала I имеет вид $am + bn$, где m и n — целые числа. При этом говорят, что идеал I порожден элементами a и b. По теореме 20.22 идеал I порожден наименьшим положительным целым числом такого вида. Но это число, как упоминалось ранее, является наибольшим общим делителем чисел a и b. Эту идею можно обобщить на случай идеала, порожденного любым конечным множеством положительных целых чисел. И опять этот идеал будет порожден наибольшим общим делителем всех этих чисел. □

Доказательство следующей теоремы предоставляется читателю.

ТЕОРЕМА 20.24. Идеал I кольца R с единицей совпадает с R тогда и только тогда, когда 1, мультипликативная единица кольца R, принадлежит идеалу I.

ТЕОРЕМА 20.25. Поле не содержит собственных идеалов.

ОПРЕДЕЛЕНИЕ 20.26. Идеал I коммутативного кольца R называется ***простым идеалом***, если $ab \in I$ имеет следствием, что $a \in I$ или $b \in I$.

ТЕОРЕМА 20.27. В кольце целых чисел идеал $\langle a\rangle$ является простым идеалом тогда и только тогда, когда a — простое число.

ДОКАЗАТЕЛЬСТВО. Предположим, что число a не является простым, например, $a = rs$, где r и s — целые числа, каждое из которых больше 1. Тогда число a принадлежит идеалу $\langle a\rangle$, однако, числа r и s этому идеалу не принадлежат. Если число a — простое и $rs \in \langle a\rangle$, то a делит rs; следовательно, по теореме 3.43, a делит r или a делит s. Поэтому либо r принадлежит идеалу $\langle a\rangle$, либо s принадлежит $\langle a\rangle$. ■

ОПРЕДЕЛЕНИЕ 20.28. Область целостности D является ***областью главных идеалов***, если каждый идеал в области D является главным идеалом.

Было уже показано, что Z, область целостности целых чисел, является областью главных идеалов.

ОПРЕДЕЛЕНИЕ 20.29. Если A — коммутативное кольцо с единицей, то пусть A^* обозначает множество $\{a \in A :$ существует $b \in A$ такое, что $ab = 1\}$. Подмножество A^* является группой относительно умножения, которая называется ***группой делителей единицы*** кольца A. Каждый элемент множества A^* называется ***делителем единицы*** кольца A. В кольце с единицей элемент s называется ***неприводимым***, если он ненулевой, не равен единице и не может быть выражен как произведение двух элементов, не являющихся делителями единицы.

В области целостности Z имеем $ab = 1$ только тогда, когда $a = b = 1$ или $a = b = -1$, поэтому 1 и -1 — делители единицы области целостности Z. В поле каждый ненулевой элемент является делителем единицы, поскольку $a \cdot a^{-1} = 1$ для $a \neq 0$.

ПРИМЕР 20.30. Множество $Z_6 = \{[0], [1], [2], [3], [4], [5]\}$ — коммутативное кольцо с единицей [1] и нулем [0]. Таблица умножения в Z_6 имеет вид

$\odot$	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]
[2]	[0]	[2]	[4]	[0]	[2]	[4]
[3]	[0]	[3]	[0]	[3]	[0]	[3]
[4]	[0]	[4]	[2]	[0]	[4]	[2]
[5]	[0]	[5]	[4]	[3]	[2]	[1]

Таблица показывает, что $[3] \odot [2] = [0]$, но $[3] \mid [0]$ и $[2] \mid [0]$, так что [3] и [2] — ненулевые делители элемента [0]. Таким образом, Z_6 не является областью целостности. □

Как показывает предыдущий пример, в Z_6 свойство сокращения не удовлетворяется, поскольку $[3] \odot [1] = [3] = [15] = [3] \odot [5]$ и, хотя $[3] \neq [0]$, имеем $[1] \neq [5]$. Делители единицы Z_6 соответствуют целым числам, взаимно простым с 6; поэтому, как следует из приведенной выше таблицы умножения, делителями единицы кольца Z_6 являются [1] и [5].

В кольце $Z_5 = \{[0], [1], [2], [3], [4]\}$ с единицей [1] и нулем [0] каждый ненулувой элемент Z_5 представляет собой делитель единицы, поскольку является взаимно простым с простым числом 5. Следовательно, каждый элемент имеет обратный элемент, поэтому Z_5 — поле. Таблица результатов операций для Z_5 дана в примере 3.61.

■ УПРАЖНЕНИЯ

1. Докажите, что если F — поле, то уравнение $ax = b$ имеет в поле единственное решение.
2. Элемент a кольца R называется делителем нуля, если в кольце R существует такой элемент b, что $ab = 0$. Докажите или опровергните, что сумма двух элементов, каждый из которых не является делителем нуля, не будет делителем нуля.
3. Докажите или опровергните, что в кольце R произведение двух элементов, каждый из которых не является делителем нуля, не будет делителем нуля.
4. Что представляют собой делители нуля в Z_{10}?
5. Что представляют собой делители нуля в Z_7?
6. Докажите теорему 20.25. Поле не содержит собственных идеалов.
7. Докажите теорему 20.3. Пусть R — коммутативное кольцо с единицей. Кольцо R является областью целостности тогда и только тогда, когда из того, что $ab = ac$, следует, что $b = c$ для всех b, c и ненулевых a из R.

8. Пусть M — кольцо $(2\times)2$-матриц вида $\begin{bmatrix} a & b \\ c & 0 \end{bmatrix}$. Докажите, что множество всех матриц вида $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ является идеалом кольца M.

9. Докажите, что сумма идеалов I и J коммутативного кольца R, определенного соотношением $I + J = \{i + j \mid i \in I \text{ и } j \in J\}$, является идеалом кольца R.

10. Пусть $A = \langle 14 \rangle$ и $B = \langle 16 \rangle$. Найдите $A \cap B$ и $A + B$.

11. Опишите наименьший идеал целых чисел, содержащих числа 6, 9 и 12.

12. Докажите, что если I — идеал кольца R и $-1 \in I$, то $I = R$.

13. Пусть $f : R \to R'$ — гомоморфизм колец. Докажите, что $f(R)$ — кольцо.

14. Пусть $f : R \to R'$ — гомоморфизм колец. Докажите, что если R — поле, то $f(R)$ — поле.

15. Пусть $f : R \to R'$ — гомоморфизм колец. Если R — область целостности, то является ли областью целостности $f(R)$?

16. Пусть $f : R \to R'$ — гомоморфизм колец. Докажите, что $\{x \in R \mid f(x) = 0\}$ — идеал кольца R.

17. Докажите или опровергните, что делители нуля коммутативного кольца R образуют идеал кольца R.

18. Образует ли множество полиномов с целыми коэффициентами область целостности?

19. Элемент a кольца R называется идемпотентом, если $a^2 = a$. Докажите, что если R — область целостности, то идемпотентами будут только элементы 0 и 1.

20. Докажите теорему 20.24. Идеал I кольца R совпадает с R тогда и только тогда, когда 1, мультипликативная единица кольца R, принадлежит I.

21. Докажите, что для любых элементов a, b кольца R имеют место равенства $a(-b) = -(ab) = (-a)b$ и $(-a)(-b) = ab$.

20.2. ОБЛАСТИ ЦЕЛОСТНОСТИ

ОПРЕДЕЛЕНИЕ 20.31. Говорят, что область целостности D является ***гауссовым кольцом***, если выполнены следующие условия:

а) если элемент области D не нуль и не делитель единицы, то его можно представить в виде произведения конечного числа неприводимых элементов;

б) если элемент области D имеет разложения $p_1 \cdots p_r$ и $q_1 \cdots q_s$ в виде произведения неприводимых элементов, то $r = s$ и q_j можно перенумеровать, так что p_i и q_i для всех i будут отличаться делителем единицы, т.е. $p_i = a_i q_i$ для некоторого делителя единицы a_i.

Известно, что множество целых чисел является гауссовым кольцом. Простое число было определено как число, у которого нет нетривиальных множителей. Иными словами, простое число неприводимо. Альтернативное определение простого числа сводится к тому, что p — простое число тогда и только тогда, когда из того, что $p \mid ab$, следует, что $p \mid a$ или $p \mid b$. Для целых чисел эти определения эквивалентны, но в случае произвольных областей целостности это неверно. Например, пусть A — множество всех комплексных чисел вида

$$a + b\sqrt{5}i\,.$$

Легко показать, что A — область целостности и что

$$21 = 3 \cdot 7 = (1 + 2\sqrt{5}i)(1 - 2\sqrt{5}i)\,.$$

Все эти множители неприводимые, и, следовательно, A не является гауссовым кольцом. Однако, ни один из этих множителей не является простым согласно альтернативному определению простого числа.

ОПРЕДЕЛЕНИЕ 20.32. Коммутативное кольцо A с единицей называется ***упорядоченным кольцом*** тогда и только тогда, когда существует непустое подмножество A^+ кольца A, называемое подмножеством положительных элементов кольца A таких, что

а) если $a, b \in A^+$, то $a + b \in A^+$;

б) если $a, b \in A^+$, то $a \cdot b \in A^+$;

в) для заданного элемента $a \in A$ выполняется одно и только одно из перечисленных ниже альтернативных условий:

(i) $a \in A^+$;

(ii) $a = 0$;

(iii) $-a \in A^+$.

Говорят, что коммутативное кольцо с единицей, которое содержит такое множество A^+, удовлетворяет аксиоме трихотомии. Если $a \in A^+$, то будем говорить, что $a > 0$. Если $-a \in A^+$, то будем говорить, что $a < 0$.

Доказательство приведенной ниже теоремы предоставляется читателю.

ТЕОРЕМА 20.33. Каждое упорядоченное кольцо является областью целостности. Для любого заданного $a \neq 0$ имеем $a^2 > 0$. В частности, $1^2 > 0$ и, следовательно, $1 > 0$.

ОПРЕДЕЛЕНИЕ 20.34. Упорядоченная область целостности A называется ***вполне упорядоченной*** тогда и только тогда, когда любое непустое подмножество S множества A^+ имеет первый элемент, т.е. существует такой элемент $s \in S$, что если $t < s$, то $t \notin S$.

Было уже показано, что целые числа образуют вполне упорядоченную область целостности.

ТЕОРЕМА 20.35. Если A — вполне упорядоченная область целостности, то не существует такой элемент c области A, что $0 < c < 1$.

ДОКАЗАТЕЛЬСТВО. Пусть S — подмножество всех таких элементов c из множества A, что $0 < c < 1$. Если S не пусто, то существует наименьший элемент s из S. Но s^2 принадлежит A; и учитывая, что $s > 0$, получаем $s^2 > 0$. Поскольку из условий $s < 1$, $0 < s$ следует, что $s^2 < s$, то имеем $0 < s^2 < s$. Поэтому $s^2 \in S$ и $s^2 < s$, что противоречит тому факту, что s — наименьший элемент множества S. Следовательно, S — пустое множество. ■

ОПРЕДЕЛЕНИЕ 20.36. В упорядоченной области целостности пусть $\overline{n+1} = \bar{n} + 1$, где 1 — мультипликативная единица.

ТЕОРЕМА 20.37. В любой упорядоченной области целостности A для подмножества положительных элементов A^+ следующие утверждения эквивалентны:

1) Первый принцип индукции.
2) Принцип полного упорядочения.
3) Второй принцип индукции.

ДОКАЗАТЕЛЬСТВО. (1) → (2). Пусть $S = \{\bar{n} \in A^+$: если $\bar{n} \in T$, где T — подмножество множества A^+, то T имеет наименьший элемент $\}$. Необходимо показать, что $S = A^+$. Очевидно, $1 \in S$, так как если $1 \in T$, то 1 — наименьший элемент множества T. Предположим, что $k \in S$ и пусть $k + 1 \in T$. Требуется показать, что множество T содержит наименьший элемент. Если $1 \in T$, то T содержит наименьший элемент, что и требовалось доказать. Предположим, что $1 \notin T$. Пусть $T' = \{s : s + 1 \in T\}$. Поскольку $k \in T'$, то T' имеет наименьший элемент, например, r. По определению, $r+1 \in T$. Кроме того, $r+1$ — наименьший элемент множества T. В противном случае, если $u \in T$ и $u < r + 1$, то в силу того, что $u = v + 1$ для некоторого $v \in T'$, получаем, что $v < r$, что является противоречием. Следовательно, множество T имеет наименьший элемент, $k+1 \in S$ и $S = A^+$.

(2) → (3). Предположим, что S — подмножество множества A^+, обладающее следующими свойствами:

а) $1 \in S$;

б) если $k \in A^+$ и $s \in S$ для всех $s < k$, то $k \in S$.

Требуется показать, что $S = A^+$. Выберем подмножество S с перечисленными выше свойствами. Пусть $T = \{\overline{n} : \overline{n} \in A^+$ и $\overline{n} \notin S\}$. Теперь необходимо показать, что множество T пустое. Если T не пусто, то по принципу вполне упорядочения T содержит наименьший элемент, например, t, причем $t \neq 1$, потому что 1 принадлежит множеству S, а t ему не принадлежит. Следовательно, $t = r + 1$ для некоторого r. Множество всех таких $x \in A^+$, что $x < t$, не пусто, т.к. ему принадлежит r. Все эти элементы принадлежат S, поскольку t — наименьшее

положительное целое число, которое не принадлежит S. Следовательно, согласно пункту (б) имеем $t \in S$, что является противоречием, поэтому множество T пусто.

(3) $\rightarrow$ (1). Пусть S — подмножество A^+. Пусть $p(1)$ есть утверждение: "$\bar{n} \in S$ имеет следствием $(\overline{n+1}) \in S$", а $p(2)$ есть утверждение: "Если для всех $m < \bar{n}$ имеем $m \in S$, то $\bar{n} \in S$". Очевидно, что из $p(1)$ вытекает $p(2)$. Второй принцип индукции можно сформулировать так: "Если $1 \in S$ и $p(2)$, то $S = A^+$", а первый принцип индукции можно сформулировать таким образом: "Если $1 \in S$ и $p(1)$, то $S = A^+$". Поскольку $1 \in S$ и $p(1)$ имеет следствием $1 \in S$ и $p(2)$, то из утверждения: "$1 \in S$ и $p(2)$ имеет следствием $S = A^+$" вытекает утверждение "$0 \in S$ и $p(1)$ имеет следствием $S = A^+$)". (Это верно, потому что $(p \rightarrow q) \Longrightarrow ((q \rightarrow r) \rightarrow (p \rightarrow r))$.) ■

ТЕОРЕМА 20.38. Любые две вполне упорядоченные, упорядоченные области целостности являются изоморфными, поэтому они изоморфны Z.

ДОКАЗАТЕЛЬСТВО. Пусть A — вполне упорядоченная, упорядоченная область целостности и N — множество положительных целых чисел. Пусть I — мультипликативная единица области целостности A. Определим $f : N \rightarrow A^+$ соотношением $f(1) = I$, и если $f(k) = K$, то $f(k+1) = K + I$. Легко показать, что $f(N)$ удовлетворяет первому принципу индукции и, следовательно, $f(N) = A^+$. Необходимо показать, что $f(m+n) = f(m) + f(n)$ и $f(m \cdot n) = f(m)f(n)$.

Используя индукцию по n, покажем сначала, что $f(m+n) = f(m) + f(n)$. Для $n = 1$, $f(m+1) = f(m) + I$ по определению функции f, поэтому $f(m+1) = f(m) + f(1)$. Предположим, что $f(m+k) = f(m) + f(k)$. Тогда

$$\begin{aligned} f(m + (k+1)) &= f((m+k)+1) = \\ &= f(m+k) + I = \\ &= f(m) + f(k) + I = \\ &= f(m) + f(k+1) \end{aligned}$$

и для любых $m, n \in N$

$$f(m+n) = f(m) + f(n).$$

Индукцией по n покажем, что $f(m \cdot n) = f(m)f(n)$. При $n = 1$

$$f(m \cdot 1) = f(m) = f(m) \cdot I = f(m)f(1).$$

Предположим, что $f(m \cdot k) = f(m)f(k)$. Тогда

$$\begin{aligned} f(m \cdot (k+1)) &= f((m \cdot k) + m) = \\ &= f(m \cdot k) + f(m) = \\ &= f(m)f(k) + f(m) = \\ &= f(m)f(k) + f(m)I = \\ &= f(m)(f(k) + I) = \\ &= f(m)f(k+1) \end{aligned}$$

и $f(m \cdot n) = f(m)f(n)$ для всех $m, n \in N$. Поэтому функция f — изоморфизм. Этот изоморфизм легко расширить до изоморфизма из Z в A. ■

ОПРЕДЕЛЕНИЕ 20.39. Область целостности называется ***минимальной областью*** тогда и только тогда, когда она не содержит никакой подобласти, кроме самой себя.

Минимальную область можно найти, построив пересечение всех подобластей области целостности. Доказательство следующей теоремы оставляем читателю.

ТЕОРЕМА 20.40. Любые две упорядоченные минимальные области целостности изоморфны. Они изоморфны целым числам и, следовательно, вполне упорядочены.

■ УПРАЖНЕНИЯ

1. Докажите, что пересечение всех подобластей произвольной области целостности является областью целостности, и используйте этот факт для доказательства того, что каждая область целостности содержит минимальную область целостности.
2. Докажите, что любые две упорядоченные минимальные области целостности изоморфны. Докажите, что они изоморфны множеству целых чисел, поэтому вполне упорядочены.
3. Докажите, что в упорядоченной области целостности для любого ненулевого элемента a имеем $a^2 > 0$. В частности, $1 > 0$ (теорема 20.33).

20.3. ПОЛИНОМЫ

В этом разделе будут формально определены полиномиальные формы и изучены некоторые их алгебраические свойства. Полиномиальные функции $f(x)$ и $g(x)$ равны при условии, что $f(b) = g(b)$ для всех b из их общей области определения. Полиномиальные формы f и g равны при условии, что равны коэффициенты при соответствующих степенях x. Для полиномов над целыми числами эти два понятия совпадают; однако, в общем случае коммутативных колец с единицей это неверно. В конце раздела будет показана согласованность этих двух понятий равенства. Большинство из приведенных ниже рассуждений изложено в терминах коммутативных колец с единицей и областей целостности, однако, использоваться будут, как правило, кольца целых или рациональных чисел. Для простоты, в большей части последующего изложения можно представлять себе A как множество целых чисел.

В дальнейшем станет понятно, что фактически полиномы определяются как производящие функции.

ОПРЕДЕЛЕНИЕ 20.41. Пусть A — коммутативное кольцо с единицей и пусть S — множество всех последовательностей $(a_0, a_1, a_2, \ldots)$ элементов кольца A таких, что если $f \in S$, то существует целое число N_f, так что $a_j = 0$ для всех $j > N_f$. Если $f \in S$, то говорят, что f — ***полином***, или ***полиномиальная форма*** над кольцом A.

Например,

$$(1,0,0,0,0,\ldots) \quad \text{и} \quad (1,1,1,0,0,0,\ldots)$$

принадлежат S, но

$$(1,1,1,1,1,1,1,1,\ldots) \quad \text{и} \quad (1,0,1,0,1,0,1,0,1,\ldots)$$

множеству S не принадлежат. Если $A = Z$, то

$$(0,-5,2,0,0,\ldots) \quad \text{и} \quad (3,7,5,8,0,0,\ldots)$$

являются полиномами из S.

ОПРЕДЕЛЕНИЕ 20.42. Пусть A — коммутативное кольцо с единицей и пусть $f = (a_i)^* = (a_0, a_1, a_2, \ldots)$, и $g = (b_i)^* = (b_0, b_1, b_2, \ldots)$ принадлежат S, множеству полиномов над кольцом A. Определим ***сумму*** полиномов f и g как последовательность $f + g = (a_i + b_i)^* = (a_0 + b_0, a_1 + b_1, \ldots)$ и ***произведение*** полиномов f и g как последовательность $fg = (c_k)^*$, где $c_k = \sum_{i+j=k} a_i b_j$.

ТЕОРЕМА 20.43. Пусть $f, g \in S$, где S — множество полиномов над коммутативным кольцом A с единицей. Тогда $f + g \in S$ и $fg \in S$.

ДОКАЗАТЕЛЬСТВО. Случай суммы $f + g$ мы оставляем читателю. Для доказательства того, что $fg \in S$, заметим, что для любого $k > N_f + N_g + 1$, если $i + j = k$ и $i \leq N_f$, то $j > N_g + 1$, и если $j \leq N_g$, то $i > N_f + 1$. Следовательно, $c_k = 0$ для любого $k > N_f + N_g + 1$, поскольку в выражении для c_k либо $a_i = 0$, либо $b_j = 0$. ■

Доказательство следующей теоремы оставляем читателю.

ТЕОРЕМА 20.44. Если S — множество полиномов над коммутативным кольцом A с единицей, то S — также коммутативное кольцо с единицей. Его единицей является $(1,0,0,0,\ldots)$, а нулевым элементом — $(0,0,0,\ldots)$.

ОПРЕДЕЛЕНИЕ 20.45.

а) Пусть A — коммутативное кольцо с единицей, пусть $f \in S$ и $f = (a_i)^*$. Для $f \neq 0$ пусть $\deg(f)$ равно наибольшему целому числу k такому, что $a_k \neq 0$. Функция $\deg(f)$ называется ***степенью*** полинома f. Множество S называется ***кольцом полиномов*** над кольцом A. Множество S обозначается через $A[x]$. Произвольный элемент множества $A[x]$ называется ***полиномом*** над кольцом A. Любой полином степени 0 или равный нулю называется ***константой***.

б) Пусть $f = (a_0, a_1, a_2, \ldots)$ принадлежит $A[x]$. Члены последовательности a_i называются ***коэффициентами*** полинома f. Если $f \neq 0$ и $n = \deg(f)$, то a_n называется ***старшим коэффициентом*** полинома f. Если $a_n = 1$, то полином f называется ***нормированным***. Если $f \neq 0$ обладает таким свойством, что наибольший общий делитель всех его ненулевых коэффициентов равен единице, то f называется ***примитивным*** полиномом.

в) Два элемента f и g множества $A[x]$ ***равны*** ($f = g$), если равны их соответствующие коэффициенты; т.е. если $f = (a_0, a_1, a_2, \ldots)$ и $g = (b_0, b_1, b_2, \ldots)$, то $f = g$ тогда и только тогда, когда $a_i = b_i$ для любого неотрицательного целого числа i.

г) Полином f ***делит*** полином g в том случае, если существует такой полином h, что $fh = g$. В этом случае говорят, что f и h являются ***делителями*** полинома g.

Например, если

$$f = (0, 1, 1, 0, 1, 0, 0, \ldots),$$

то $\deg(f) = 4$. Если

$$g = (1, 0, 0, 0, \ldots),$$

то $\deg(g) = 0$ и g — константа. Тождественно нулевой полином

$$(0, 0, 0, \ldots)$$

не имеет степени. Предоставляем читателю доказать, что равенство есть отношение эквивалентности на множестве $A[x]$.

ТЕОРЕМА 20.46. Пусть A — коммутативное кольцо с единицей, пусть $A[x]$ — кольцо полиномов над кольцом A и пусть f и g принадлежат $A[x]$.

а) Если $f, g \neq 0$, то $\deg(f + g) \leq \max(\deg(f), \deg(g))$.

б) Либо $fg = 0$, либо $\deg(fg) \leq \deg(f) + \deg(g)$.

в) Если A — область целостности, то либо $fg = 0$, либо $\deg(fg) = \deg(f) + \deg(g)$.

г) Если A — область целостности, то $A[x]$ — также область целостности.

ДОКАЗАТЕЛЬСТВО.

а) Пусть $f = (a_0, a_1, a_2, \ldots)$ и $g = (b_0, b_1, b_2, \ldots)$, где $n = \deg(f)$ и $m = \deg(g)$. Поскольку $a_i = 0$ для $i > n$ и $b_j = 0$ для $j > m$, то если $k > \max(m, n)$, то $a_k + b_k = 0$ и $\deg(f + g) \leq \max(\deg(f), \deg(g))$.

б) Пусть $n = \deg(f)$, $m = \deg(g)$ и $fg = (c_0, c_1, c_2, \ldots)$, так что $c_k = \sum_{i+j=k} a_i b_j$. Произведение $a_i b_j = 0$, когда $i > n$, поскольку $a_i = 0$. Если $i \leq n$ и $i + j > m + n$, то $j > m + (n - i)$, поэтому $b_j = 0$ и, в любом случае, $c_k = 0$. Следовательно, $\deg(fg) \leq \deg(f) + \deg(g)$.

в) Согласно пункту (б) $\deg(fg) \leq \deg(f) + \deg(g)$. Однако, поскольку a_n и b_m не равны 0, то $c_{n+m} = a_n b_m$ не равно 0. Следовательно, $\deg(fg) \geq \deg(f) + \deg(g)$, и требуемое равенство доказано.

г) Пусть f и g — ненулевые элементы множества $A[x]$ степени n и m, соответственно. Поскольку A — область целостности, то $\deg(fg) = \deg(f) + \deg(g)$ и fg не равно 0.

■

ТЕОРЕМА 20.47. Существует мономорфизм из A в $A[x]$, кольцо полиномов над кольцом A, для которого образ кольца A является подкольцом кольца $A[x]$. Если A — область целостности, то каждый делитель единицы в $A[x]$ соответствует делителю единицы в A согласно мономорфизму.

ДОКАЗАТЕЛЬСТВО. Определим функцию $\phi : A \to A[x]$ следующим образом: если $a \in A$, то пусть $\phi(a) = (a, 0, 0, 0, \ldots)$. Если a и b принадлежат A, то

$$\begin{aligned}\phi(a+b) &= (a+b, 0, 0, 0, 0, \ldots) = \\ &= (a, 0, 0, 0, 0, \ldots) + (b, 0, 0, 0, 0, \ldots) = \\ &= \phi(a) + \phi(b)\,,\end{aligned}$$

$$\begin{aligned}\phi(a \cdot b) &= (a \cdot b, 0, 0, 0, \ldots) = \\ &= (a, 0, 0, 0, \ldots) \cdot (b, 0, 0, 0, \ldots) = \\ &= \phi(a) \cdot \phi(b)\,.\end{aligned}$$

Кроме того, $\phi(1) = (1, 0, 0, 0, \ldots) = 1$. Следовательно, ϕ — гомоморфизм из A в $A[x]$.

Если $(c, 0, 0, 0, \ldots) = \phi(c) = 0 = (0, 0, 0, 0, \ldots)$, то $c = 0$. Если $\phi(a) = \phi(b)$, то $\phi(a) - \phi(b) = 0$; т.е. $\phi(a - b) = 0$ и $a - b = 0$. Поэтому $a = b$, и функция ϕ — инъекция.

Напомним, что делитель единицы кольца — это элемент, который имеет обратный элемент; т.е. если b — делитель единицы, то существует такой элемент b^{-1}, что $b \cdot b^{-1} = 1$.

Предположим, что A — область целостности и f — делитель единицы в $A[x]$. Пусть g — соответствующий обратный элемент, поэтому $fg = 1 = (1, 0, 0, 0, \ldots)$. Но $0 = \deg(fg) = \deg(f) + \deg(g)$, откуда следует, что $\deg(f)$ и $\deg(g)$ — обе равны 0. Таким образом, $f = (a, 0, 0, 0, \ldots)$ и $g = (b, 0, 0, 0, \ldots)$ для некоторых $a, b \in A$. Поскольку

$$\phi(a \cdot b) = fg = (1, 0, 0, 0, \ldots) = \phi(1)$$

и, учитывая, что ϕ — инъекция, получаем, что $ab = 1$, поэтому a и b — делители единицы области целостности A.

Поскольку мономорфный образ $\phi(A)$ в кольце $A[x]$ является кольцом, изоморфным кольцу A, то A можно отождествить с $\phi(A)$ как подкольцом кольца $A[x]$ и считать A подкольцом кольца $A[x]$. ■

ОПРЕДЕЛЕНИЕ 20.48. Символ Кронекера — δ_{ij} — для целых чисел i и j определяется соотношениями:

$$\delta_{ij} = \begin{cases} 1, & \text{если } i = j; \\ 0, & \text{если } i \neq j. \end{cases}$$

Например, если $f = (a_0, a_1, a_2, \ldots)$, где $a_i = \delta_{i3}$, то

$$f = (0, 0, 0, 1, 0, 0, \ldots).$$

Если $g = \sum_{j=1}^{5} (\delta_{ij})^*$, то

$$g = (0, 1, 1, 1, 1, 1, 0, 0, \ldots).$$

ТЕОРЕМА 20.49. Если $x = (0, 1, 0, 0, 0, \ldots) = (c_i)^*$, где $c_i = \delta_{i1}$, то для каждого $k > 0$ имеем $x^k = (a_0, a_1, a_2, \ldots)$, где $a_i = \delta_{ik}$.

ДОКАЗАТЕЛЬСТВО. При $k = 1$ теорема, очевидно, справедлива. Предположим, что теорема справедлива для $k = n$. Поскольку $x^{n+1} = x \cdot x^n$, то простое умножение показывает, что i-ый член в последовательности x^{n+1} равен $\delta_{i,n+1}$. ■

Если $a \in A$ и $f = (a_0, a_1, a_2, \ldots)$, то $af = (aa_0, aa_1, aa_2, \ldots)$. В частности, если $a \in A$, то $ax^k = (0, 0, 0, \ldots, a, 0, 0, 0, \ldots)$, где a находится на $(k+1)$-ом месте.

Пусть $f \in A[x]$ имеет степень n. Тогда

$$\begin{aligned} f &= (a_0, a_1, a_2, \ldots, a_n, 0, 0, \cdots) = \\ &= (a_0, 0, 0, \ldots) + (0, a_1, 0, \ldots) + \cdots + (0, 0, \ldots, 0, a_n, 0, 0, \ldots) = \\ &= a_0(1, 0, 0, \ldots) + a_1(0, 1, 0, 0, \ldots) + \cdots + a_n(0, 0, \ldots, 0, 1, 0, 0, \ldots) = \\ &= a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n. \end{aligned}$$

ОПРЕДЕЛЕНИЕ 20.50. Пусть A — коммутативное кольцо с единицей и пусть $A[x]$ — множество полиномов над кольцом A. Символ x назовем ***переменной*** над кольцом A. С каждым полиномом

$$f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in A[x]$$

связана функция из A в A вида

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

или

$$f(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0,$$

которая называется ***полиномиальной функцией***. Пусть $A(x) = \{f(x) : f \in S\}$. Степень функции $f(x)$ совпадает со степенью соответствующего полинома $f \in A[x]$. Элементы f множества $A[x]$ будем называть ***полиномиальными формами***, чтобы иметь возможность отличать их от полиномиальных функций.

Определим функцию $\theta : A[x] \to A(x)$ соотношением $\theta(f) = f(x)$. Легко показать, что отображение θ, которое устанавливает соответствие между элементами множества $A[x]$ и соответствующими полиномиальными функциями из $A(x)$, представляет собой гомоморфизм. Поскольку отображение θ — сюръекция, оно является также эпиморфизмом. Таким образом, все свойства множества $A[x]$,

сохраненные гомоморфизмом, перенесены на коммутативное кольцо полиномиальных функций $A(x)$. Вскоре будет показано, что если A — бесконечная область целостности, то $\theta : A[x] \to A(x)$ — изоморфизм.

ОПРЕДЕЛЕНИЕ 20.51. Пусть A — коммутативное кольцо с единицей. Если

$$f(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0$$

и

$$g(x) = b_n x^n + \cdots + b_2 x^2 + b_1 x + b,$$

где некоторые из a_i и b_i из A могут быть 0, включая a_n или b_n, то

$$f(x) + g(x) = (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$

и

$$f(x)g(x) = c_m x^m + \cdots + c_2 x^2 + c_1 x + c_0,$$

где $c_k = \sum_{i+j=k} a_i b_j$.

Будем говорить, что $f(x) = g(x)$ тогда и только тогда, когда $f(b) = g(b)$ для всех $b \in A$. ***Решением*** уравнения $f(x) = 0$ называется элемент $a \in A$ такой, что $f(a) = 0$.

На основании вышесказанного мы можем сформулировать следующую теорему относительно полиномиальных функций над областью целостности A.

ТЕОРЕМА 20.52. Если $f(x)$ и $g(x)$ — полиномиальные функции над областью целостности A, степень $f(x)$ равна n, а степень $g(x)$ равна m, то

а) степень $f(x) + g(x)$ меньше или равна, чем $\max\{m, n\}$;
б) степень $f(x)g(x)$ равна $m + n$.

Согласно определению две полиномиальные формы f и g равны тогда и только тогда, когда равны их коэффициенты. Две полиномиальные функции $f(x)$ и $g(x)$ по определению равны тогда и только тогда, когда $f(b) = g(b)$ для всех $b \in A$. В следующем разделе для полинома f степени n будет показано, что уравнение $f(x) = 0$ имеет не более n решений. Сейчас без доказательства сформулируем теорему, утверждение которой используем для доказательства более слабого результата.

ТЕОРЕМА 20.53. Если f — полином степени n над бесконечной областью целостности и $f(x)$ — соответствующая полиномиальная функция, то уравнение $f(x) = 0$ имеет не более n решений.

СЛЕДСТВИЕ 20.54. Пусть $f(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0$ — полиномиальная функция над бесконечной областью целостности A. Если $f(a) = 0$ для всех $a \in A$, то $a_0 = a_1 = a_2 = \cdots = a_n = 0$.

ТЕОРЕМА 20.55. Пусть f и g — полиномы над бесконечной областью целостности A. Тогда $f = g$ тогда и только тогда, когда соответствующие полиномиальные функции $f(x)$ и $g(x)$ обладают таким свойством, что $f(b) = g(b)$ для всех $b \in A$.

ДОКАЗАТЕЛЬСТВО. Если f и g — полиномы над областью целостности A такие, что $f = g$, то очевидно, $f(b) = g(b)$ для всех $b \in A$. Пусть

$$f(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0$$

и

$$g(x) = b_n x^n + \cdots + b_2 x^2 + b_1 x + b$$

— полиномиальные функции степени n и m, соответственно, и $f(c) = g(c)$ для каждого $c \in A$. Пусть f и g — соответствующие полиномы. Тогда полиномиальная функция $h(x) = c_s x^s + \cdots + c_2 x^2 + c_1 x + c$, где $s = \max(n, m)$ и $c_i = a_i - b_i$, обладает свойством: $h(c) = f(c) - g(c) = 0$ для всех $c \in A$. Но, согласно предыдущему следствию, если $h(c) = 0$ для всех c из A, то $h(x)$ должен быть нулевым полиномом. Поэтому $c_i = a_i - b_i = 0$ или $a_i = b_i$ для $i \geq 1$ и $f = g$. ■

Из предыдущей теоремы непосредственно следует теорема, упоминавшаяся нами ранее.

ТЕОРЕМА 20.56. Пусть A — бесконечная область целостности, определим $\theta : A[x] \to A(x)$ соотношением $\theta(f) = f(x)$. Тогда функция θ является изоморфизмом.

Следующий пример демонстрирует, что для справделивости приведенных выше следствия и теоремы существенно, чтобы область целостности A была бесконечной.

ПРИМЕР 20.57. Множество $Z_5 = \{[0], [1], [2], [3], [4]\}$ классов вычетов по модулю 5 является полем, поэтому по теореме 20.8 представляет собой область целостности. По теореме Ферма, если для простого числа p имеем $a \not\equiv 0 \pmod{p}$, то $a^{p-1} \equiv 1 \pmod{p}$. Тогда $a^p \equiv a \pmod{p}$ для любого целого числа a, даже если $a \equiv 0 \pmod{p}$. Для $p = 5$ имеем $a^5 \equiv a \pmod{5}$ для всех $a \in Z$. Это сравнение равносильно $[a]^5 = [a]$ или $[a]^5 - [a] = [0]$ для любого целого числа a. Таким образом, если $g(x) = x^5 - x$, то $g(x) = [0]$ выполняется для любого x из Z_5. Однако, если h — полином, равный нулю, то $h(x) = [0]$ для всех x из Z_5. Полиномы g и h не равны, но $g(b) = h(b)$ для всех b из области целостности Z. Таким образом, над конечной областью целостности равенство полиномиальных форм не эквивалентно равенству полиномиальных функций. С другой стороны, поскольку множество целых чисел бесконечно, то понятия равенства для полиномиальных форм и полиномиальных функций над Z совпадают. □

■ УПРАЖНЕНИЯ

1. Докажите теорему 20.44.
2. Покажите, что равенство на множестве S в определении 20.45 есть отношение эквивалентности.
3. Докажите или опровергните: пусть f и g — полиномы над коммутативным кольцом A. Если $f, g \neq 0$, то

$$\deg(f + g) = \max(\deg(f), \deg(g)) .$$

4. Найдите такие два полинома из множества полиномов над областью целостности Z_{12} (т.е. коэффициенты полиномов принадлежат Z_{12}), что $\deg(fg) < \deg(f) + \deg(g)$.
5. Докажите или опровергните: для множества полиномиальных функций над областью целостности Z_6, если $f(x) = g(x)$ для всех x, то соответствующие полиномы равны.

20.4. АЛГЕБРЫ И ПОЛИНОМЫ

В этом разделе рассматриваются особые подкольца и подполя комплексных чисел. Они состоят из комплексных целых чисел, алгебраических целых чисел и алгебраических чисел. В данном разделе будут изучены также другие свойства полиномов, и для заданного поля F будут найдены такие расширения поля, что полиномы, которые не могут быть разложены на множители в поле F, могут быть разложены на множители в расширении этого поля. Напомним, что делитель единицы является делителем числа 1; т.е.элемент u — делитель единицы при условии, что существует такой элемент u', что $uu' = 1$. Напомним также, что каждый ненулевой элемент поля является делителем единицы.

В дальнейшем будет построена алгебраическая структура, которую мы затем применим к полиномам. Все рассматриваемые кольца будут предполагаться коммутативными.

ОПРЕДЕЛЕНИЕ 20.58. Поле F называется ***подполем*** поля F', если оно содержится в F' и имеет те же самые операции. Поле F' называется ***расширением поля*** F. В общем случае, поле F' называется расширением поля F, если существует изоморфизм $\eta : F \to E$, где E — подполе поля F'.

ОПРЕДЕЛЕНИЕ 20.59. Пусть A — область целостности. Ненулевой элемент a из A называется ***неприводимым***, если из того, что $a = bc$, следует, что b или c является делителем единицы; в противном случае элемент a называется ***приводимым***. Пусть $A[x]$ обозначает множество полиномов с коэффициентами из A. Полином f из $A[x]$ с ненулевой степенью называется ***неприводимым***, если из того, что $f = g \cdot h$, следует, что g или h — элемент области целостности A; в противном случае полином f называется ***приводимым***.

Таким образом, если полином f приводимый, то существуют полиномы ненулевой степени g и h такие, что $f = g \cdot h$.

Далее покажем, что определения неприводимости для областей целостности и полиномов совпадают, если A — поле или рассмотрение ограничено нормированными полиномами.

ОПРЕДЕЛЕНИЕ 20.60. Полином f из множества $A[x]$ называется ***простым***, если его степень больше 0, и всякий раз, когда $f \mid g \cdot h$, то или $f \mid g$, или $f \mid h$. Вообще, элемент p интегральной области является простым, если из того, что $p \mid ab$, следует, что $p \mid a$ или $p \mid b$.

Во множестве целых чисел простые числа и неприводимые числа совпадают. Однако для некоторых колец полиномов это неверно.

ТЕОРЕМА 20.61. В области целостности простой элемент всегда неприводим.

ДОКАЗАТЕЛЬСТВО. Предположим, что p — простой элемент и $p = ab$; тогда, по определению, $p \mid a$ или $p \mid b$. Предположим, что $p \mid a$; тогда $a = mp$ для некоторого m. Далее, $p = mpb = mbp$, поэтому $mb = 1$ и b — делитель единицы. Следовательно, элемент p — неприводимый. ■

ОПРЕДЕЛЕНИЕ 20.62. Пусть $A^{\emptyset}$ обозначает ненулевые элементы области целостности A. Тогда A называется ***евклидовым кольцом***, если существует такая функция $\phi : A^{\emptyset} \to N$, где N — множество положительных целых чисел, что

1. Если $a, b \in A^{\emptyset}$, то $\phi(a) \leq \phi(ab)$, или, что эквивалентно, если $a \mid c$, то $\phi(a) \leq \phi(c)$.
2. Если $a, b \in A^{\emptyset}$, то существуют $q, r \in A^{\emptyset}$ такие, что $a = bq + r$ и либо $r = 0$, либо $\phi(r) < \phi(b)$.

В евклидовом кольце A величина $\phi(a)$, где $a \in A^{\emptyset}$, называется ***нормой*** элемента a.

ТЕОРЕМА 20.63. Пусть F — поле. Тогда $F[x]$ — евклидово кольцо с нормой $\phi(f) = \deg(f)$ для $f \in F[x]$.

ДОКАЗАТЕЛЬСТВО. Пусть $A^{\emptyset} = F[x] - \{(0)^*\}$. Определим $\phi : A^{\emptyset} \to N$ соотношением $\phi(f) = deg(f)$, когда $f \neq 0$. Часть (а) следует из теоремы 20.46. Для доказательства части (б) пусть $f \in F[x]$ — полином степени n, а g — полином степени k. Пусть S — множество полиномов вида $f - q \cdot g$. Если окажется, что произвольный элемент множества S равен 0, то доказательство завершено. В противном случае выберем один элемент с наименьшей степенью, например, r, и пусть q — соответствующее частное. Степень этого элемента меньше, чем n, потому что если старший коэффициент полинома b равен v, то выберем полином степени $n - k$ со старшим коэффициентом, равным c, так что vc — старший коэффициент полинома a. В таком случае степень $f - q \cdot g$ меньше, чем n. Аналогично, если $\deg(r)$ не меньше, чем $\deg(b)$, то

$$r = q' \cdot f + r',$$

где $\deg(r')$ меньше, чем $\deg(r)$, поэтому

$$f - q \cdot g + r = q' \cdot g + r'$$

и

$$f - (q + q')g = r',$$

где $\deg(r') < \deg(r)$. Но это противоречит определению элемента r. Следовательно, $\deg(r) < \deg(b)$. ■

Следующая теорема является частным случаем теоремы 20.89, которая будет доказана далее.

ТЕОРЕМА 20.64. Пусть F — поле и $p \in F[x]$ — полином степени n. Тогда для произвольного $a \in F$ полином $x - a$ будет делителем полинома p тогда и только тогда, когда $p(a) = 0$.

ДОКАЗАТЕЛЬСТВО. Если $x-a$ — делитель полинома p, то, несомненно, $p(a) = 0$. Обратно, предположим, что $p(a) = 0$. По предыдущей теореме $p = (x - a) \cdot q + r$, где $r \in F$. Поэтому $p(a) = (a - a) \cdot q + r = r$ и $r = 0$. Следовательно, $(x - a)$ — делитель полинома p. ■

Используя две предыдущие теоремы и принцип индукции, приходим к следующему утверждению.

ТЕОРЕМА 20.65. Пусть F — поле и $p \in F[x]$ — полином степени n. Тогда уравнение $p(x) = 0$ имеет не более n решений.

Поскольку область целостности можно вложить в поле, то приходим к такому следствию.

СЛЕДСТВИЕ 20.66. Пусть A — область целостности и $p \in A[x]$ — полином степени n. Тогда уравнение $p(x) = 0$ имеет не более n решений.

ТЕОРЕМА 20.67. Любое евклидово кольцо A есть кольцо главных идеалов.

ДОКАЗАТЕЛЬСТВО. Пусть I — ненулевой идеал евклидового кольца A. Пусть d — ненулевой элемент идеала I с минимальным значением нормы и пусть $p \in I$. Однако, $p = dq+r$, где норма элемента r меньше, чем норма элемента d. Поскольку $r \in I$, то элемент d имеет минимальную норму, а норма r меньше, чем норма d, имеем $r = 0$ и $p = dq$. Следовательно, элемент d порождает I. ■

СЛЕДСТВИЕ 20.68. Пусть F — поле, тогда $F[x]$ — кольцо главных идеалов.

ОПРЕДЕЛЕНИЕ 20.69. Идеал I называется ***максимальным*** в кольце R, если для любого идеала J из того, что $I \subseteq J \subseteq R$, следует, что $I = J$ или $J = R$.

ТЕОРЕМА 20.70. Идеал I из кольца главных идеалов является максимальным, тогда и только тогда, когда p, порождающий элемент идеала I, является неприводимым над полем F.

ДОКАЗАТЕЛЬСТВО. Предположим, что идеал I — максимальный, и элемент p не является неприводимым, например, $p = rs$. Тогда элемент p принадлежит идеалу J, порожденному элементом r(или s), и $I \subset J$, что является противоречием.

Обратно, если элемент p приводимый и идеал I — не максимальный, например, $I \subset J$ и элемент r порождает J, то $p = rs$ для некоторого элемента s, что невозможно, поскольку элемент p — неприводимый. ■

ТЕОРЕМА 20.71. Пусть I — идеал, принадлежащий кольцу R. Определим отношение $\sim$ на R таким условием: $a \sim b$ тогда и только тогда, когда $a - b \in I$. В этом случае отношение $\sim$ есть отношение эквивалентности.

ДОКАЗАТЕЛЬСТВО. Отношение $\sim$ имеет следующие свойства:

а) рефлексивность: $a \sim a$, поскольку $0 \in I$;
б) симметричность: если $a \sim b$, то $a - b \in I$; следовательно, $b - a = -(a - b) \in I$, поскольку I — группа относительно сложения. Поэтому $b \sim a$;
в) транзитивность: пусть $a \sim b$ и $b \sim c$, так что $a - b \in I$ и $b - c \in I$. Поэтому

$$(a - b) + (b - c) = a - c \in I\,.$$

Таким образом, $a \sim c$. ■

ТЕОРЕМА 20.72. Если $a \sim b$ и $c \sim d$, то $a + c \sim b + d$ и $ac \sim bd$.

ДОКАЗАТЕЛЬСТВО. Если $a \sim b$ и $c \sim d$, то $a - b \in I$ и $c - d \in I$, тогда

$$(a - b) + (c - d) = (a + c) - (b + d) \in I$$

и $a + c \sim b + d$. Кроме того, $c(a - b) = ac - bc \in I$ и $b(c - d) = bc - ad \in I$, поскольку I — идеал. Следовательно, $(ac - bc) + (bc - bd) = ac - bd \in I$ и $ac \sim bd$. ■

Теперь можно определить сложение и умножение между классами отношения эквивалентности. Если $a + I$ обозначает класс эквивалентности, содержащий элемент a, и если $b + I$ — класс эквивалентности, содержащий элемент b, то определим

$$(a + I) + (b + I) = (a + b) + I$$

и

$$(a + I) \cdot (b + I) = a \cdot b + I\,.$$

Когда есть возможность умножать и складывать классы эквивалентности, то, как отмечалось ранее, отношение называется ***отношением конгруэнтности***. Кольцо классов эквивалентности обозначается через R/I. Заметим, что $0 + I$ — это нулевой элемент кольца, а $1 + I$ — мультипликативная единица. Класс эквивалентности $a + I$ называется ***левым смежным классом*** идеала I, точнее говоря, левым смежным классом, порожденным элементом a. Следовательно, различные левые смежные классы идеала I — это именно классы эквивалентности в R/I.

ТЕОРЕМА 20.73. Пусть R — коммутативное кольцо с единицей и I — идеал кольца R. В этом случае R/I — поле тогда и только тогда, когда I — максимальный идеал.

ДОКАЗАТЕЛЬСТВО. Предположим, что идеал I — максимальный и пусть $a + I \in R/I$, где $a \notin I$. Пусть $M = \{ra + b \mid r \in R \text{ и } b \in I\}$. Можно непосредственно показать, что M — идеал. Поскольку $a \notin I$, получаем, что $M \subset I$ и $M = R$. Следовательно, $1 \in M$ и существуют элемент b из R и элемент c из I такие, что $1 = ab + c$. Таким образом, $1 + I = ab + I = (a + I)(b + I)$ и элемент $a + I$ имеет обратный элемент. Следовательно, R/I — поле.

Обратно, предположим, что R/I — поле. Допустим, что существует такой идеал J, что $I \subset J \subseteq R$. Рассмотрим элемент $a + I$, где элемент a принадлежит J, но не принадлежит I. Тогда существует такой элемент a' из R, что $a' + I$ — элемент, обратный элементу $a + I$, откуда получаем, что $aa' + I = 1 + I$. Более того, поскольку J — идеал, то $aa' \in J$. Кроме того, $1 = aa' + c$ для некоторого c из I. Поскольку aa' и c принадлежат J, то 1 принадлежит J. Но если 1 принадлежит J, то $r1$ принадлежит J для всех r, и $J = R$. ■

ТЕОРЕМА 20.74. Пусть R — коммутативное кольцо с единицей. В этом случае R/I — область целостности тогда и только тогда, когда I — простой идеал.

ДОКАЗАТЕЛЬСТВО. Если R/I — область целостности и ij — элемент идеала I, то $ij + I = 0 + I$. Следовательно, $(i + I)(j + I) = 0 + I$. Но поскольку R/I — область целостности, то либо $i + I$, либо $j + I$ равняется $0 + I$. Следовательно, либо $i \in I$, либо $j \in I$, и I — простой идеал.

Обратно, если I — простой идеал и $(i + I)(j + I) = 0$, то $ij + I = 0$ и $ij \in I$. Но поскольку I — простой идеал, то $i \in I$ или $j \in I$; поэтому, $i + I = 0$ или $j + I = 0$, и I — область целостности. ■

ТЕОРЕМА 20.75. Любой максимальный идеал является простым идеалом.

ДОКАЗАТЕЛЬСТВО. Доказательство следует непосредственно из теорем 20.73 и 20.74. ■

ТЕОРЕМА 20.76. В кольце главных идеалов A каждый неприводимый элемент является простым.

ДОКАЗАТЕЛЬСТВО. Пусть p — неприводимый элемент кольца A и пусть $\langle p \rangle$ — главный идеал, порожденный элементом p. Поскольку элемент p неприводим, то по теореме 20.70 $\langle p \rangle$ — максимальный идеал и, вследствие теоремы 20.75, является простым идеалом. Поэтому, если $p \mid (ab)$, то $ab \in \langle p \rangle$. Таким образом, $a \in \langle p \rangle$ или $b \in \langle p \rangle$, так что $p \mid a$ или $p \mid b$, и элемент p — простой. ■

Комбинируя следствие 20.68 и теорему 20.76, получаем следующий результат.

СЛЕДСТВИЕ 20.77. Если F — поле, то каждый неприводимый элемент множества $F[x]$ является простым.

Следующая теорема является аналогом принципа единственности разложения целого числа на простые множители.

ТЕОРЕМА 20.78. Если F — поле, то каждый неконстантный полином из множества $F[x]$ можно единственным образом представить в виде произведения неприводимых полиномов в том смысле, что множители являются единственными с точностью до умножения на делитель единицы.

ДОКАЗАТЕЛЬСТВО. Воспользовавшись индукцией по степеням полинома, покажем, что любой полином степени n, не равный константе, можно единственным образом разложить на неприводимые элементы. Для $n = 1$ имеем полином вида $ax + b$, который уже является неприводимым, поскольку в противном случае $ax+b = fg$, но $\deg(ax+b) = \deg(f)+\deg(g) \geq 2$, что составляет очевидное противоречие. Предположим, что $k > 1$, и теорема справедлива для всех $n < k$, и пусть полином f имеет степень k. Если полином f — неприводимый, то доказательство завершено. Если нет, то f можно разложить на полиномы g и h, степень каждого из которых меньше, чем степень f. Следовательно, каждый из полиномов g и h можно разложить на неприводимые элементы, поэтому полином f также можно разложить на неприводимые элементы.

Для доказательства единственности разложения опять воспользуемся индукцией по числу множителей в разложении. Для $n = 1$ результат очевиден. Предположим, что единственность имеет место для $n = k$ множителей. Предположим, что f можно разложить на $k + 1$ неприводимых элементов $g_1, g_2, \ldots, g_{k+1}$ и на m неприводимых элементов $g'_1, g'_2, \ldots, g'_m$. Из $k+1$ неприводимых элементов выберем множитель g_1. Поскольку, в силу следствия 20.77, неприводимый элемент является простым и g_1 делит f, то g_1 должен делить один из m неприводимых элементов во втором разложении, например, g'_j. Но так как g'_j — неприводимый элемент, то g_1 и g'_j отличаются множителем — делителем единицы. Следовательно, делим оба разложения на g_1, оставляя первое разложение полинома f состоящим из k неприводимых множителей $g_2, \ldots, g_{k+1}$, а второе разложение — состоящим из $m - 1$ неприводимых множителей $g'_1, \ldots, g'_{j-1}, g_{j+1}, \ldots, g'_m$ и возможного множителя — делителя единицы. По индуктивному предположению разложение полинома f/g_1 единственно, и поэтому разложение полинома f единственно. ■

Напомним, что полином $f = a_0 + a_1x + \cdots + a_nx^n$ примитивный, если НОД$(a_0, a_1, \ldots, a_n)$ равен единице.

ТЕОРЕМА 20.79. (Гаусс) Пусть A — область целостности. Произведение двух примитивных полиномов из множества $A[x]$ является примитивным полиномом.

ДОКАЗАТЕЛЬСТВО. Пусть $f = a_0 + a_1x + \cdots + a_nx^n$ и $g = b_0 + b_1x + \cdots + b_mx^m$, и пусть p — неприводимый элемент из A. Поскольку полином f — примитивный, то существует такой коэффициент a_i, что p не делит a_i. Пусть a_r — первый такой коэффициент. Аналогичным образом, существует такой коэффициент b_j, что p не делит b_j. Пусть b_s — первый такой коэффициент. Пусть $h = c_0 + c_1x + \cdots + c_{m+n}x^{m+n}$ — произведение полиномов f и g. Тогда

$$c_{r+s} = a_0b_{r+s} + \cdots + a_{r-1}b_{s+1} + a_rb_s + a_{r+1}b_{s-1} + \cdots + a_{r+s}b_0.$$

Но p делит $a_0, \ldots, a_{r-1}$, поэтому p делит $a_0b_{r+s} + \cdots + a_{r-1}b_{s+1}$; при этом p делит $b_0, \ldots, b_{s-1}$, поэтому p делит $a_{r+1}b_{s-1} + \cdots + a_{r+s}b_0$. Но p не делит a_rb_s, поэтому p не делит h. Следовательно, полином h — примитивный. ■

Может показаться, что полином, неприводимый над полем целых чисел, может и не быть неприводимым, если допустить разложение на полиномы с рациональными коэффициентами. Приведенная ниже теорема, однако, показывет, что это не так.

ТЕОРЕМА 20.80. Пусть A — область единственности разложения на множители, т.е. область целостности, в которой разложение на простые множители единственно, и пусть F — его поле частных. Если $p \in A[x]$ и если многчлен p приводимый в $F[x]$, то он приводимый в $A[x]$. Если полином p — примитивный в $A[x]$, то он приводимый в $F[x]$ тогда и только тогда, когда он приводимый в $A[x]$.

ДОКАЗАТЕЛЬСТВО. Пусть полином $p \in A[x]$ — приводимый в $F[x]$, например, $p = fg$. Пусть u — наименьшее общее кратное знаменателей коэффициентов полиномов f и g. Тогда $u \cdot p = f'g'$, где f' и g' принадлежат множеству $A[x]$. Но $f' = af''$ и $g' = bg''$, где f'' и g'' примитивные полиномы, и $p = cp'$, где c — примитивный полином. Следовательно, $uc \cdot p' \cdot f''g''$. Но поскольку $f''g''$ — также примитивный полином, то легко показать, что два выражения для примитивного полинома отличаются множителем — делителем единицы. Итак, $ucv = ab$, где v — делитель единицы. Следовательно, $ucp' = ucvf''g''$. Поэтому

$$up = ucvf''g''$$

или

$$p = cvf''g'',$$

и мы получили разложение полинома p на множители в $A[x]$.

Очевидно, что если полином разлагается на множители в $A[x]$, то он разлагается на множители в $F[x]$, но если полином p не является примитивным в $A[x]$, то он приводимый в $A[x]$, потому что он является произведением целого числа и полинома. Однако, в поле $F[x]$ полином может быть неприводимым, поскольку целое число может быть делителем единицы. ■

ТЕОРЕМА 20.81. Область целостности A является областью единственности разложения тогда и только тогда, когда таковой является область $A[x]$.

ДОКАЗАТЕЛЬСТВО. Пусть F — поле частных области целостности A и p — полином из $A[x]$. Тогда полином p разложим на неприводимые множители в $F[x]$, следовательно, p разложим на неприводимые множители в $A[x]$, где каждый неприводимый множитель разложения в $F[x]$ равен соответствующему неприводимому множителю разложения в $A[x]$, умноженному на элемент поля F.

Пусть d равно наибольшему общему делителю коэффициентов полинома p, так что $p = dp'$, где p' — примитивный полином. Поскольку полином d разложим единственным образом, то необходимо только показать, что p' разложим единственным образом. Поэтому предположим, что полином p — примитивный. Поскольку, по теореме 20.78, поле $F[x]$ — область единственности разложения на множители, то пусть

$$p' = r_1 r_2 \cdots r_n = s_1 s_2 \cdots s_n .$$

Допустим, что после перегруппировки имеем, что r_i и s_i отличаются только ненулевым множителем из F. Например, $r_i = (a/b)s_i$, поэтому $br_i = as_i$, но поскольку r_i и s_i примитивные полиномы, то a и b отличаются множителем — делителем единицы из $A[x]$, поэтому разложение будет единственным.

Обратно, поскольку область целостности A — подмножество множества $A[x]$, то если $A[x]$ — область единственности разложения на множители, то каждый элемент из A должен раскладываться на множители единственным образом, поэтому A — область единственности разложения на множители. ■

ТЕОРЕМА 20.82. Пусть A — область единственности разложения на множители, тогда два полинома из $A[x]$ имеют наибольший общий делитель.

ДОКАЗАТЕЛЬСТВО. Пусть p и q — полиномы из $A[x]$. Поскольку $A[x]$ — область единственности разложения на множители, то разложим p и q на неприводимые множители. Тогда НОД(p, q) равен произведению множителей, которые присутствуют в разложениях как p, так и q, в самых низких степенях. ■

ТЕОРЕМА 20.83. Пусть A — кольцо главных идеалов и I — идеал, порожденный элементами p и q. Если d порождает I, то $d =$ НОД(p, q) и d можно записать в виде $up + vq$.

ДОКАЗАТЕЛЬСТВО. Вполне очевидно, что d делит p и q. Поскольку $d \in I$, то d можно записать в виде $up + vq$. Если c делит p и c делит q, то c делит $up + vq$ и c делит d. Таким образом, $d =$ НОД(p, q). ■

Начиная с области целостности A, такой как целые числа, несомненно, можно найти поле частных области целостности A, например, поле F, и сформировать множество $F[x]$. Поскольку $A[x]$ — область целостности, то можно сформировать поле частных области целостности $A[x]$, порождая, по существу, элементы вида p/q, где p и q принадлежат $A[x]$. Поскольку $F[x]$ — также область целостности, то можно сформировать поле частных области целостности $F[x]$, которое состоит из элементов вида p/q, где p и q принадлежат$F[x]$. Это поле называется ***полем рациональных функций*** над полем F.

ОПРЕДЕЛЕНИЕ 20.84. Элемент $c \in C$, множеству комплексных чисел, называется ***алгебраическим целым числом***, если он является нулем некоторого нормированного полинома p из $Z[x]$, т.е. $p(c) = 0$.

ОПРЕДЕЛЕНИЕ 20.85. Элемент a поля E — расширения поля F — называется ***алгебраическим*** над полем F, если $f(a) = 0$ для некоторого ненулевого полинома $f \in F[x]$.

ОПРЕДЕЛЕНИЕ 20.86. Элемент $c \in C$, алгебраический над полем рациональных чисел Q, называется ***алгебраическим числом***; т.е. $c \in C$ — алгебраическое число, если существует такой полином $p \in Q[x]$, что $p(c) = 0$.

ОПРЕДЕЛЕНИЕ 20.87. Пусть A — область целостности. Элемент $a \in A$ называется ***корнем*** полинома p, если $p(a) = 0$

Далее в абстрактной форме покажем, как по заданному полю и неприводимому полиному над этим полем можно найти расширение поля, содержащее корень неприводимого полинома. Затем сравним это расширение поля с таким, которое было бы сформировано расширением поля корнем неприводимого полинома в комплексных числах.

ТЕОРЕМА 20.88. Пусть F — поле и p — неприводимый полином над полем F. Тогда F изоморфно подполю поля $F[x]/\langle p\rangle$, и существует такой элемент f поля $F[x]/\langle p\rangle$, что $p(f) = 0$.

ДОКАЗАТЕЛЬСТВО. Согласно теоремам 20.70 и 20.73, множество $F[x]/\langle p\rangle$ представляет собой поле. Поле F можно отождествить с подполем F' поля $F[x]/\langle p\rangle$ следующим образом: пусть $g : F \to F'$ определено соотношением $g(a) = a + \langle p\rangle$. Функция g является мономорфизмом (т.е. инъекцией), поскольку если $g(a) = g(b)$, то $a + \langle p\rangle = b + \langle p\rangle$, и p делит $a - b$. Поскольку p — полином, и полиномы a и b принадлежат F, то $a - b = 0$ и $a = b$.

Пусть $f = x + \langle p\rangle$ и $p = a_0 + a_1x + \cdots + a_nx^n$. Тогда

$$\begin{aligned} p(f) &= a_0 + a_1(x + \langle p\rangle) + \cdots + a_n(x + \langle p\rangle)^n = \\ &= a_0 + a_1x + \cdots + a_nx^n + \langle p\rangle = \\ &= p + \langle p\rangle = \\ &= 0 + \langle p\rangle = \\ &= 0 \end{aligned}$$

в $F[x]/\langle p\rangle$. ■

ТЕОРЕМА 20.89. Пусть p — неприводимый полином из $F[x]$ и $p(a) = 0$ в расширении поля F. Если $g(a) = 0$ для полинома g из $F[x]$, то p делит g.

ДОКАЗАТЕЛЬСТВО. Можно показать, что $I = \{g : g \in F[x] \text{ и } g(a) = 0\}$ — идеал. Вполне очевидно, что этот идеал порожден элементом p, так как p — неприводимый элемент. Следовательно, если $g \in I$, т.е. $g(a) = 0$, то полином g должен быть кратным полиному p. ■

ТЕОРЕМА 20.90. Пусть F — поле и $p \in F[x]$ — неприводимый полином. Если K — расширение поля F, содержащее корень a полинома p, т.е. $p(a) = 0$, то $F[a]$ изоморфно $F[x]/\langle p\rangle$.

ДОКАЗАТЕЛЬСТВО. Определим $h : F[x]/\langle p(x)\rangle \to F[a]$ соотношением $h(f + \langle p\rangle) = f(a)$. Вполне очевидно, что h — гомоморфизм. Если $h(f + \langle p\rangle) = h(g + \langle p\rangle)$, то $f(a) = g(a)$ и $(f - g)(a) = 0$. Следовательно, по теореме 20.89 полином p делит $f - g$ и $f + \langle p\rangle = g + \langle p\rangle$, поэтому функция h — инъекция. Пусть $f \in F[a]$, тогда $h(f + \langle p\rangle) = f(a)$, поэтому h — сюръекция и, следовательно, h — изоморфизм. ■

Пусть F — поле действительных чисел и $p = x^2+1$, и пусть K — расширение поля F, содержащее элемент i — корень полинома $p(x)$, т.е. $i^2 + 1 = 0$. По теореме 20.89 поле $F[x]/\langle p\rangle$ изоморфно полую $F[i]$. Поскольку $p(x)$ имеет степень 2, необходимо рассматривать только $f \in F[i]$ степени 1. Следовательно, каждый элемент поля $F[i]$ будет иметь вид $a+bi$, где a и b — действительные числа. Итак, мы создали комплексные числа.

Подкольцо, состоящее из элементов поля $F[i]$, где a и b — целые числа, называется множеством ***комплексных целых чисел***, или множеством ***гауссовых чисел***. Его можно также рассматривать как наименьшее подкольцо комплексных чисел, содержащее целые числа и число i. Очевидно, что число 5 не является неприводимым в этом кольце, поскольку $5 = (1-2i)(1+2i)$. Пусть норма элемента

n_G на множестве комплексных целых чисел определена соотношением $n_G(a+bi) = a^2 + b^2$. Тогда

$$\begin{aligned} n_G((a+bi)(c+di)) &= n_G((ac-bd)+(ad+bc)i) = \\ &= (ac-bd)^2 + (ad+bc)^2 = \\ &= a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2c^2 = \\ &= (a^2+b^2)(c^2+d^2) = \\ &= n_G(a+bi)n_G(c+di)\,, \end{aligned}$$

так что n_G — мультипликативный гомоморфизм из множества Гауссовых чисел во множество целых чисел. Поскольку $n_G(1-2i) = n_G(1+2i) = 5$, то $1-2i$ и $1+2i$ — неприводимые элементы в кольце комплексных целых чисел. Произведение неприводимых комплексных целых чисел $2+i$ и $2-i$ есть также число 5. Следовательно, единственность разложения на неприводимые множители не имеет места. Обратите внимание, что неприводимые множители в данном примере не являются простыми.

Еще один пример. Пусть F — поле рациональных чисел, и $p(x) = x^2 - 2$. Тогда $F[x]/\langle p\rangle = F[\sqrt{2}]$. Опять достаточно рассмотреть полиномы степени 1, и $F[x]/\langle p\rangle$ — поле, состоящее из рациональных чисел вида $a + b\sqrt{2}$, где a и b — рациональные числа. Мы можем рассмотреть подкольцо, состоящее только из элементов множества $F[x]/\langle p\rangle$ вида $a + b\sqrt{2}$, где a и b — целые числа. Но $(3+\sqrt{2})(3-\sqrt{2}) = 7$, поэтому число 7 не является неприводимым в этом множестве.

Мы предположили, что, начиная с рациональных чисел, все алгебраические целые числа являются подмножеством множества комплексных чисел. В рамках такого предположения докажем следующую теорему.

ТЕОРЕМА 20.91. *Множество A всех алгебраических чисел над полем рациональных чисел Q является подполем поля комплексных чисел.*

ДОКАЗАТЕЛЬСТВО. Пусть a и b — ненулевые алгебраические числа. Если число a не является рациональным, то существует такой неприводимый полином p степени 2 или выше, что a — корень полинома p. Пусть $F = Q[x]/\langle p\rangle$. Если число b не принадлежит F, то в F существует такой неприводимый полином, скажем, q, степени выше 2, что b — корень полинома q. Пусть $G = F[x]/\langle q\rangle$. Тогда G — поле, которое содержит элементы a и b, поэтому оно содержит $a+b$, $a-b$, ab и a/b. Поскольку поле G содержится в A, то все эти элементы принадлежат A, и поэтому A — поле. ■

■ УПРАЖНЕНИЯ

1. Докажите, что множество Гауссовых чисел образует подкольцо множества комплексных чисел.
2. Докажите или опровергните, что множество Гауссовых чисел является областью целостности.
3. Докажите, что $\sin(x)$ нельзя выразить полиномом.

4. Докажите или опровергните, что произведение двух простых полиномов будет простым полиномом.
5. Докажите или опровергните, что сумма двух простых полиномов будет простым полиномом.
6. Найдите порождающий элемент наименьшего главного идеала, содержащего полиномы $x^2 - 4$ и $x^2 + 4x + 4$, где полиномы рассматриваются над кольцом целых чисел.
7. Докажите или опровергните, что любое комплексное целое число является алгебраическим числом.
8. Найдите идеал целых чисел, который не является максимальным идеалом.
9. Опишите наименьшее подполе поля действительных чисел, содержащее корни полинома $x^2 - 5 = 0$.
10. Опишите наименьшее подполе поля действительных чисел, содержащее корни полинома $x^2 + 5 = 0$.
11. Опишите наименьшее подполе поля действительных чисел, содержащее корни полинома $x^3 - 1 = 0$.
12. Опишите наименьшее подполе поля действительных чисел, содержащее корни полинома $x^4 - 1 = 0$.

ГЛАВА

21

ХАРАКТЕРЫ ГРУПП И ПОЛУГРУПП

21.1. КОМПЛЕКСНЫЕ ЧИСЛА

В предыдущей главе комплексные числа рассматривались как корни многочленов. Теперь мы введем их формально и рассмотрим с различных точек зрения.

ОПРЕДЕЛЕНИЕ 21.1. Пусть $\mathbb{R}$ — множество действительных чисел и $\mathbb{C}$ — множество вида

$$\{(r,s) : r,s \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R},$$

так что каждый элемент множества $\mathbb{C}$ — упорядоченная пара действительных чисел. Определим сложение на $\mathbb{C}$ соотношением

$$(a,b) + (c,d) = (a+c, b+d),$$

а умножение определим соотношением

$$(a,b) \cdot (c,d) = (ac - bd, ad + bc).$$

Множество $\mathbb{C}$ называется множеством ***комплексных чисел***.

Легко показать, что $\mathbb{C}$ — поле. Отображение $f : \mathbb{R} \to \mathbb{C}$, определенное соотношением $f(a) = (a,0)$, является мономорфизмом и изоморфизмом в подполе $\{(a,0) : a \in \mathbb{R}\}$. Поэтому будем считать, что поле действительных чисел $\mathbb{R}$ содержится в поле комплексных чисел, и называть поле $\mathbb{R}$ ***подполем*** поля $\mathbb{C}$. Далее, $(0,1)\cdot(0,1) = (-1,0) = -1$. Если определить $i = (0,1)$, то $i^2 = -1$. Следовательно, элемент (a,b) можно записать в виде $(a,0) + (b,0) \cdot (0,1) = a + bi$. Таким образом, определения операций сложения и умножения принимают вид

$$(a+bi) + (c+di) = (a+c) + (b+d)i$$

и

$$(a+bi) \cdot (c+di) = (ac - bd) + (ad + bc)i.$$

Пусть $z = a + bi = (a, b)$ — комплексное число. Тогда a называют его ***действительной частью*** и записывают $\mathrm{Re}(z) = \mathrm{Re}(a + bi) = a$, а b называют его ***мнимой частью*** и записывают $\mathrm{Im}(z) = \mathrm{Im}(a + bi) = b$. Поскольку упорядоченная пара действительных чисел обычно изображается в виде точки на плоскости с введенной на ней декартовой системой координат, то комплексное число $a + bi$ также можно изобразить точкой плоскости, как и пару чисел (a, b). В этом случае горизонтальня ось называется действительной осью, а вертикальная ось называется мнимой осью. Таким образом, действительная часть числа $a + bi$ (т.е. $\mathrm{Re}(a + bi) = a$) измеряется вдоль горизонтальной оси, а мнимая часть числа $a + bi$ (т.е. $\mathrm{Im}(a + bi) = b$) измеряется вдоль вертикальной оси. Когда плоскость с декартовой системой координат используется таким образом, ее называют ***комплексной плоскостью***.

ОПРЕДЕЛЕНИЕ 21.2. Пусть $\mathbb{C}$ — множество комплексных чисел. Для $c = a + bi$ из $\mathbb{C}$ величина $|c| = \sqrt{a^2 + b^2}$ называется ***абсолютным значением***, или ***длиной*** числа c.

В комплексной плоскости величина $|c|$ равна расстоянию от точки $c = a + bi$ до начала координат. Легко показать, что для комплексных чисел c и d имеем $|c|\,|d| = |cd|$ и $|c + d| \leq |c| + |d|$. Пусть r — комплексное число, которое является корнем n-ой степени из 1, т.е. $r^n = 1$, тогда $|r|^n = |r^n| = |1| = 1$. Поэтому число r имеет длину 1 и находится на ***единичной окружности*** с центром в начале координат, т.е. окружности, состоящей из всех точек комплексной плоскости, расстояние от которых до начала координат равно 1. Следующая теорема, формулируемая нами без доказательства, устанавливает важное свойство комплексных чисел.

ОСНОВНАЯ ТЕОРЕМА АЛГЕБРЫ. (Гаусс) Для любого многочлена p положительной степени с коэффициентами из множества действительных чисел $\mathbb{R}$ или множества комплексных чисел $\mathbb{C}$ существует такой элемент b из $\mathbb{C}$, что $p(b) = 0$.

Поэтому, в частности, каждый многочлен над Z имеет, по крайней мере, один корень во множестве комплексных чисел.

21.2. ХАРАКТЕРЫ ГРУПП

ОПРЕДЕЛЕНИЕ 21.3. ***Характер группы*** коммутативной группы G — это гомоморфизм из G во множество ненулевых комплексных чисел с операцией умножения.

ОПРЕДЕЛЕНИЕ 21.4. Числом, ***сопряженным*** с данным комплексным числом $c = a + bi$, называется число $a - bi$, которое обозначается через $\overline{c}$.

В оставшейся части этого раздела ограничимся рассмотрением характеров коммутативных конечных групп. Пусть χ — характер группы. Таким образом, существует такой элемент $g \in G$, что $\chi(g) \neq 0$. Поскольку группа G конечная, то существует такое n, что $g^n = 1$. Следовательно, $\chi(g^n) = (\chi(g))^n = 1$, $\chi(g)$ является корнем n-ой степени из 1, и элемент g отображен на единичную окружность. Таким образом, характер группы отображает группу G в единичную окружность. Доказательство приведенной ниже теоремы предоставляем читателю.

ТЕОРЕМА 21.5. Если $\chi(g) = a + bi$, то $\chi(g^{-1})$ является сопряженным к $\chi(g)$, т.е. $\chi(g^{-1}) = a - bi$.

ОПРЕДЕЛЕНИЕ 21.6. ***Первообразным корнем n-ой степени*** r комплексного числа a называется такой корень n-ой степени из a, что $r^k \neq a$ для $0 < k < n$.

ОПРЕДЕЛЕНИЕ 21.7. Характер группы χ_1, определенный соотношением $\chi_1(g) = 1$ для всех g из группы G, называется ***главным характером*** группы G. ***Группа характеров*** группы G состоит из множества характеров группы G вместе с операцией $(\chi \cdot \chi')(g) = \chi(g)\chi'(g)$ для характеров групп χ и χ'.

ОПРЕДЕЛЕНИЕ 21.8. Пусть $G_1, G_2, \ldots, G_m$ — подгруппы группы G. Группа G называется ***прямой суммой*** подгрупп $G_1, G_2, \ldots, G_m$, обозначается $G_1 \oplus G_2 \oplus \cdots \oplus G_m$, если каждый элемент группы G может быть единственным образом записан в виде $a_1 a_2 \cdots a_m$, где $a_k \in G_k$.

ТЕОРЕМА 21.9. Любая конечная абелева группа есть прямое произведение циклических подгрупп. Таким образом, если $G = G_1 \oplus G_2 \oplus G_3 \oplus \cdots \oplus G_m$ для циклических групп G_i с порождающим элементом g_i порядка $k(i)$, тогда каждый элемент $g \in G$ можно единственным образом записать в виде:

$$g = g_1^{j(1)} g_2^{j(2)} g_3^{j(3)} \cdots g_m^{j(m)},$$

где $0 \leq j(i) < k(i)$.

ДОКАЗАТЕЛЬСТВО. Покажем, что существует целое число m такое, что

$$G = G_1 \oplus G_2 \oplus G_3 \oplus \cdots \oplus G_m,$$

и каждый элемент группы G может быть единственным образом записан в виде $g_1^{j(1)} g_2^{j(2)} g_3^{j(3)} \cdots g_m^{j(m)}$, где $0 \leq j(i) < k(i)$. Пусть $g_1 \in G$ имеет порядок $k(1)$, и G_1 — подгруппа, порожденная элементом g_1. Если $G_1 = G$, то $m = 1$, и доказательство завершено. Если нет, то выберем элемент g_2 из $G - G_1$, и пусть G_2 — группа, порожденная элементом g_2. Тогда $g_2^n = 1$ для некоторого n, поскольку группа G конечная. Следовательно, некоторая степень элемента g_2 принадлежит G_1. Пусть $k(2)$ — наименьшая степень элемента g_2 такая, что $g_2^{k(2)}$ принадлежит G_1. Если $G_1 \cup G_2 = G_2$, то переобозначим $G_1 = G_2$ и начнем снова. Если $G_1 \cup G_2 \neq G_2$, то в этом случае легко показать, что каждый элемент из

$G_1 \cup G_2$ может быть единственным образом записан в виде $g_1^j g_2^s$, где $0 \le j < k(1)$, $0 \le s < k(2)$. Если выбрать элемент g_3 из $G - (G_1 \oplus G_2)$ и продолжить процесс, то получим $G_1 G_2 G_3 \cdots G_r G_{r+1} \cdots$. А поскольку группа G конечная, то существует число m такое, что $G = G = G_1 \oplus G_2 \oplus G_3 \oplus \cdots \oplus G_m$, и каждый элемент группы G может быть единственным образом записан в виде $g_1^{j(1)} g_2^{j(2)} g_3^{j(3)} \cdots g_m^{j(m)}$, где $0 \le j(i) < k(i)$. ■

ТЕОРЕМА 21.10. Группа характеров группы G изоморфна группе G.

ДОКАЗАТЕЛЬСТВО. Пусть $G = G_1 \oplus G_2 \oplus \cdots \oplus G_m$, где элемент g_i порождает группу G_i, и пусть a_i — первообразный корень n_i-ой степени из единицы, где G_i имеет порядок n_i. Для каждого элемента $g \in G$ определим χ_g соотношением

$$\chi_g(b) = a_1^{j(1)\cdot h(1)} a_2^{j(2)\cdot h(2)} \cdots a_m^{j(m)\cdot h(m)},$$

где элементы g и b группы G имеют представления через прямые суммы

$$g = g_1^{j(1)} g_2^{j(2)} g_3^{j(3)} \cdots g_m^{j(m)};$$
$$b = g_1^{h(1)} g_2^{h(2)} g_3^{h(3)} \cdots g_m^{h(m)},$$

где $0 \le j(i) < n_i$ и $0 \le h(i) < n_i$. Можно показать, что для любого $g \in G$ отображение χ_g является гомоморфизмом из G в корень n-ой степени из единицы. Далее, если $g \ne g'$, то $\chi_g \ne \chi_{g'}$, поэтому существуют $\prod_{i=1}^{m} n_i = n$ характеров группы, когда группа G содержит n элементов. Определим отображение η из G в группу характеров группы G следующим образом:

$$\eta(g) = \eta\left(g_1^{j(1)} g_2^{j(2)} g_3^{j(3)} \cdots g_m^{j(m)}\right) = (\chi_{g_1})^{j(1)} (\chi_{g_2})^{j(2)} \cdots (\chi_{g_m})^{j(m)}.$$

Легко показать, что отображение η является изоморфизмом. ■

Кроме того, если A — матрица, где $A_{ij} = \chi_{g_i}(g_j)$, то

$$A = \begin{bmatrix} a_1 & 1 & 1 & \cdots & 1 \\ 1 & a_2 & 1 & \cdots & 1 \\ 1 & 1 & a_3 & \cdots & 1 \\ & & \vdots & & \\ 1 & 1 & 1 & \cdots & a_k \end{bmatrix}.$$

ПРИМЕР 21.11. Рассмотрим множество $G_5 = \{[1], [2], [3], [4]\}$, которое состоит из ненулевых элементов множества Z_5. Поскольку число 5 простое, то классы приведенных вычетов как раз являются ненулевыми классами в Z_5, а ненулевые классы, соответствующие приведенным вычетам, образуют группу относительно умножения. Таким образом, G_5 — группа относительно умножения. Таблица умножения для G_5 имеет вид

$\odot$	[1]	[2]	[3]	[4]
[1]	[1]	[2]	[3]	[4]
[2]	[2]	[4]	[1]	[3]
[3]	[3]	[1]	[4]	[2]
[4]	[4]	[3]	[2]	[1]

Чтобы сделать более понятным приведенное ниже построение характеров групп, будем использовать обозначения, введенные при доказательстве теоремы 21.10.

Комплексное число

$$e^{2\pi ti/n} = \cos\left(\frac{2\pi t}{n}\right) + i\sin\frac{2\pi t}{n}$$

находится на единичной окружности, потому что

$$\left|e^{2\pi ti/n}\right| = \sqrt{\cos^2\left(\frac{2\pi t}{n}\right) + \sin^2\left(\frac{2\pi t}{n}\right)} = 1\,.$$

Для любого целого числа t число $e^{2\pi ti/n}$ является корнем n-ой степени из единицы, поскольку

$$\left(e^{2\pi ti/n}\right)^n = e^{2\pi ti} = \cos(2\pi t) + i\sin(2\pi t) = 1\,.$$

Фактически, множество

$$\left\{e^{2\pi ti/n} : t = 0, 1, 2, \ldots, (n-1)\right\}$$

является совокупностью n различных комплексных корней n-ой степени из единицы.

Для $n = 4$, порядка группы G_5, приведенная выше формула дает четыре корня четвертой степени из единицы.

t	$e^{(2\pi ti)/4}$
0	1
1	i
2	-1
3	$-i$

Подсчитывая $1^4 = i^4 = (-1)^4 = (-i)^4 = 1$, проверяем, что $1, i, -1$ и $-i$ являются корнями четвертой степени из единицы.

Если положить $g_1 = [3]$, то

$$\begin{aligned} g_1^2 &= [3]^2 = [4], \\ g_1^3 &= [3]^3 = [2], \\ g_1^4 &= [3]^4 = [1] = g_1^0. \end{aligned}$$

Следовательно, $k(1) = 4$ и $G_1 = G$, так что $m = 1$ в обозначениях доказательства теоремы 21.10. Для $g_1 = [3]$ выберем корень четвертой степени $r_1 = i$. Поскольку $r_1^2 = -1$, $r_1^3 = -i$ и $r_1^4 = 1$, то корень $r_1 = i$ является первообразным. С помощью этого первообразного корня можно определить характер для каждого элемента g группы G_5. Следующая таблица дает определение всех характеров группы G_5.

$j(1)$	$g = g_1^{j(1)}$	$\chi_g(b) = \chi_g(g_1^{h(1)})$
0	[1]	$r_1^{0\cdot h(1)} = i^0$
3	[2]	$r_1^{3\cdot h(1)} = i^{3h(1)}$
1	[3]	$r_1^{1\cdot h(1)} = i^{h(1)}$
2	[4]	$r_1^{2\cdot h(1)} = i^{2h(1)}$

Следовательно, четыре элемента группы характеров для G_5 дает следующая таблица, в которой $b = g_1^{h(1)}$:

	b	=	[1]	[2]	[3]	[4]
	$h(1)$	=	0	3	1	2
g	$\chi_g(b)$					
[1]	$\chi_{[1]}(b)$	=	1	1	1	1
[2]	$\chi_{[2]}(b)$	=	1	i	$-i$	-1
[3]	$\chi_{[3]}(b)$	=	1	$-i$	i	-1
[4]	$\chi_{[4]}(b)$	=	1	-1	-1	1

Очевидно, $\chi_{[1]} = 1$ — это главный характер.

Для иллюстрации гомоморфной природы характеров групп рассмотрим $\chi_{[2]}$:

$$\chi_{[2]}([3][4]) = \chi_{[2]}([2]) = i;$$
$$\chi_{[2]}([3]) = -i;$$
$$\chi_{[2]}([4]) = -1;$$
$$\chi_{[2]}([3])\chi_{[2]}([4]) = (-i)(-1) = i.$$

Остальные 15 произведений для $\chi_{[2]}$ можно проверить аналогичным образом.

Используя таблицу определений характеров групп, приведенную выше, легко построить таблицу умножения для группы характеров.

$\cdot$	$\chi_{[1]}$	$\chi_{[2]}$	$\chi_{[3]}$	$\chi_{[4]}$
$\chi_{[1]}$	$\chi_{[1]}$	$\chi_{[2]}$	$\chi_{[3]}$	$\chi_{[4]}$
$\chi_{[2]}$	$\chi_{[2]}$	$\chi_{[4]}$	$\chi_{[1]}$	$\chi_{[3]}$
$\chi_{[3]}$	$\chi_{[3]}$	$\chi_{[1]}$	$\chi_{[4]}$	$\chi_{[2]}$
$\chi_{[4]}$	$\chi_{[4]}$	$\chi_{[3]}$	$\chi_{[2]}$	$\chi_{[1]}$

□

■ УПРАЖНЕНИЯ

1. Докажите теорему 21.5.
2. Докажите, что если b — элемент коммутативной группы G и $b \neq e$, единичному элементу группы G, то существует такой характер группы χ, что $\chi(b) \neq 1$.

3. Пусть g и g' — элементы конечной коммутативной группы G такие, что $g \neq g'$. Докажите, что $\chi_g \neq \chi_{g'}$.
4. Завершите доказательство теоремы 21.10.

21.3. ХАРАКТЕРЫ ПОЛУГРУПП

ОПРЕДЕЛЕНИЕ 21.12. Пусть S — конечная коммутативная полугруппа. ***Характером полугруппы*** на S называется гоморфизм из S в множество комплексных чисел с операцией умножения. Как и для случая групп, в данном разделе будем предполагать, что все рассматриваемые полугруппы являются коммутативными и конечными.

Пусть $s \in S$, где S — конечная коммутативная полугруппа. Если $|\chi(s)| > 1$, то

$$\begin{aligned} |\chi(s^2)| &> |\chi(s)|, \\ |\chi(s^3)| &> |\chi(s^2)|, \\ &\vdots \\ |\chi(s^{k+1})| &> |\chi(s^k)|, \end{aligned}$$

что невозможно, поскольку полугруппа S конечна. Точно так же, если $0\,|\chi(s)| < 1$, то

$$\begin{aligned} 0 < |\chi(s^2)| &< |\chi(s)|, \\ 0 < |\chi(s^3)| &< |\chi(s^2)|, \\ &\vdots \\ 0 < |\chi(s^{k+1})| < |\chi(s^k)| &< \cdots < |\chi(s^2)| < |\chi(s)| < 1, \end{aligned}$$

что опять-таки невозможно. Следовательно, либо $|\chi(s)| = 0$, либо $|\chi(s)| = 1$ для каждого элемента s из полугруппы S.

ОПРЕДЕЛЕНИЕ 21.13. Подполугруппа I коммутативной полугруппы S называется ***идеалом*** полугруппы S, если из того, что $i \in I$ и $s \in S$, следует, что $s \cdot i \in I$. Идеал I называется ***простым идеалом***, если $S - I$ — полугруппа.

Доказательство следующей теоремы оставляем читателю.

ТЕОРЕМА 21.14. Если S — конечная коммутативная полугруппа и χ — характер полугруппы S, то $I = \{s : |\chi(s)| = 0\}$ является идеалом. Обратно, если I — любой идеал в S и существует такой характер χ, что $\chi(s) \neq 0$ для некоторого s из $S - I$, то существует такой ненулевой характер χ_I, что $\chi_I(i) = 0$ тогда и только тогда, когда i принадлежит I.

Теперь рассмотрим множество Z_n классов вычетов по модулю n. Напомним, что множество элементов Z_n, состоящее из классов эквивалентности, содержащих целые числа, взаимно простые с n, называется множеством приведенных классов вычетов. Множество приведенных классов вычетов образует группу относительно умножения в классе. Обозначим эту группу через G_n. Множество Z_n является полугруппой относительно умножения в классе. Множество $Z_n - G_n$ является идеалом полугруппы, поскольку произведение любого целого числа и целого числа, не взаимно простого с n, не является взаимно простым с n. Следовательно, можно сформировать характеры Z_n, которые являются характерами группы G_n, и отобразить все элементы группы $Z_n - G_n$ в 0. Эти характеры называются ***характерами Дирихле***. Известно, что количество характеров Дирихле по модулю n равно количеству характеров группы G_n. Поскольку множество характеров группы G_n изоморфно G_n и существуют $\phi(n)$ элементов в G_n, то существует $\phi(n)$ характеров Дирихле по модулю n. Итак, мы приходим к следующей теореме.

ТЕОРЕМА 21.15. Существует $\phi(n)$ различных характеров Дирихле по модулю положительного целого числа n.

ЛЕММА 21.16. Пусть G — конечная группа, g' — произвольный элемент группы G и $G' = \{x : x = g'g$ для g из $G\}$. Тогда $G = G'$.

ДОКАЗАТЕЛЬСТВО. Предположим, что $G' = \{x : x = g'g$ для g из $G\}$. Вполне очевидно, что $G' \subseteq G$. Пусть g_1 и g_2 — элементы группы G. Тогда $g'g_1$ и $g'g_2$ принадлежат G'. Если $g'g_1 = g'g_2$, то $g_1 = (g')^{-1}g'g_1 = g_2$. Следовательно, в G' имеется столько же элементов, сколько и в G; и поскольку группа G конечная, то $G = G'$. ■

ТЕОРЕМА 21.17. Пусть G — конечная коммутативная группа порядка n. Если χ — характер группы, то

$$\sum_{g \in G} \chi(g) = \begin{cases} n, & \text{если } \chi = \chi_1; \\ 0, & \text{если } \chi \neq \chi_1. \end{cases}$$

ДОКАЗАТЕЛЬСТВО. Поскольку $\chi_1(g) = 1$ для всех элементов g из G, то первая часть утверждения очевидна. Если $\chi \neq \chi_1$ и g' — такой элемент группы G, что $\chi(g') \neq 1$, то

$$\begin{aligned} \sum_{g \in G} \chi(g) &= \sum_{g \in G} \chi(g'g) = && \text{по лемме 21.16} \\ &= \sum_{g \in G} \chi(g')\chi(g) = \\ &= \chi(g') \sum_{g \in G} \chi(g), \end{aligned}$$

так что

$$[1 - \chi(g')] \sum_{g \in G} \chi(g) = 0,$$

и, поскольку $1 - \chi(g') \neq 0$, то $\sum_{g \in G} \chi(g) = 0$. ■

ТЕОРЕМА 21.18. Если G — конечная коммутативная группа порядка n и $g \in G$, то

$$\sum_{\chi} \chi(g) = \begin{cases} n, & \text{если } g = e; \\ 0, & \text{если } g \neq e, \end{cases}$$

где e — единичный элемент группы, и суммирование осуществляется по всем характерам группы χ.

ДОКАЗАТЕЛЬСТВО. Если $g = e$, то $\chi(e) = 1$ для всех χ, так как каждый характер χ является гомоморфизмом. Таким образом, сумма равна n. Предположим, что $g \neq e$ и χ_a — такой характер, что $\chi_a(g) \neq 1$. Такой характер существует, потому что $g = g_1^{j(1)} g_2^{j(2)} g_3^{j(3)} \cdots g_m^{j(m)}$, где $0 \leq j(i) < k(i)$, как показано в доказательстве теоремы 21.10. Далее, поскольку $g \neq e$, то $j(s) \neq 0$ для некоторого s. Теперь можем положить $\chi_a(g)$ равным $\chi_{g_s}(g)$ по той причине, что $\chi_{g_s}(g) = r_s^{j_s} \neq 1$, где r_s — первообразный корень $k(s)$-ой степени из единицы. Тогда

$$\begin{aligned} \sum_{\chi} \chi(g) &= \sum_{\chi} (\chi_a \chi)(g) = && \text{согласно лемме 21.16} \\ &= \sum_{\chi} \chi_a(g)\chi(g) = \\ &= \chi_a(g) \sum_{\chi} \chi(g), \end{aligned}$$

и

$$[1 - \chi_a(g)] \sum_{\chi} \chi(g) = 0.$$

Поскольку $\chi_a(g) \neq 1$, получаем $\sum_{\chi} \chi(g) = 0$. ■

Следующая теорема формулирует два полезных свойства ортогональности характеров групп. Доказательство теорем оставляем читателю.

ТЕОРЕМА 21.19. Пусть G — конечная коммутативная группа порядка n, пусть $g_i, g_j \in G$, пусть χ_a и χ_b — характеры группы G. Тогда

а) $\sum_{\chi} \chi(g_i^{-1})\chi(g_j) = \begin{cases} n, & \text{если } g_i = g_j; \\ 0, & \text{если } g_i \neq g_j. \end{cases}$

б) $\sum_{g \in G} \chi_a(g)\chi_b^{-1}(g) = \begin{cases} n, & \text{если } \chi_a = \chi_b; \\ 0, & \text{если } \chi_a \neq \chi_b. \end{cases}$

ТЕОРЕМА 21.20. Предположим, что группа G имеет порядок n и σ — (n)-матрица, в которой $\sigma_{ij} = \chi_i(g_j)$, где $\chi_1, \chi_2, \chi_3, \ldots, \chi_n$ — характеры группы G. Если σ' — матрица, в которойh $\sigma'_{ij} = \overline{\sigma_{ji}}$, то

$$\sigma'\sigma = \begin{bmatrix} n & 0 & 0 & \cdots & 0 \\ 0 & n & 0 & \cdots & 0 \\ 0 & 0 & n & \cdots & 0 \\ & & \vdots & & \\ 0 & 0 & 0 & \cdots & n \end{bmatrix} = \sigma\sigma'.$$

ДОКАЗАТЕЛЬСТВО. Пусть $C = [C_{ij}] = \sigma'\sigma$. Тогда

$$C_{ij} = \sum_{k=1}^{n} \sigma'_{ik}\sigma_{kj} = \sum_{k=1}^{n} \overline{\sigma'_{ki}}\sigma_{kj} =$$
$$= \sum_{k=1}^{n} \overline{\chi_k(g_i)}\chi_k(g_j) =$$
$$= \sum_{k=1}^{n} \chi_k(g_i^{-1})\chi_k(g_j),$$

поэтому первое утверждение теоремы следует из теоремы 21.19. Доказательство второго утверждения предоставляем читателю. ■

ТЕОРЕМА 21.21. Всего существует $\phi(m)$ различных характеров Дирихле по модулю положительного целого числа m. Каждый характер Дирихле является вполне мультипликативным и обладает тем свойством, что $\chi([a+m]) = \chi([a])$, где $[a]$ — класс вычетов по модулю m, содержащий a, или, что эквивалентно, $\chi(a+m) = \chi(a)$.

Замечание: поскольку сложение и умножение в Z_m не зависят от выбора представителя класса из Z_m, то, как правило, вместо $\chi([a+m])$ используется запись $\chi(a+m)$. Таким образом, область определения функции χ расширена до Z. Именно в этом обобщенном смысле мы говорим о том, что каждый характер является вполне мультипликативным.

ДОКАЗАТЕЛЬСТВО. Поскольку каждый характер Дирихле является единственным расширением характера группы из группы приведенных классов вычетов по модулю m, то первая часть утверждения теоремы следует из того факта, что количество характеров группы совпадает с количеством элементов в группе. Поскольку каждое "целое" в образе характера Дирихле на самом деле представляет класс сравнимости по модулю m, то $[a] = [a+m]$ и $\chi([a+m]) = \chi([a])$.

Характер χ является вполне мультипликативным, так как если элементы a и b — взаимно простые с m, то $[a]$ и $[b]$ принадлежат группе приведенных вычетов, и χ — гомоморфизм на эту группу. Если a или b не является взаимно простым с m, то таковым не является и их произведение, поэтому оба $\chi([a])\chi([b])$ и $\chi([ab])$ равны 0 или, что эквивалентно, оба $\chi(a)\chi(b)$ и $\chi(ab)$ равны 0. ■

■ УПРАЖНЕНИЯ

1. Докажите теорему 21.14.
2. Докажите теорему 21.19.
3. Завершите доказательство теоремы 21.20.

ГЛАВА

22

ПРИЛОЖЕНИЯ ТЕОРИИ ЧИСЕЛ

22.1. ПРИЛОЖЕНИЕ: ПОИСК ПО ОБРАЗУ

В компьютерных приложениях часто возникает задача поиска по образу. Простой пример такой задачи можно найти в компьютерных программах обработки текстов, когда при написании или редактировании документа в нем требуется определять местоположение конкретного слова или фразы, например, слова "effect". Если документ содержит несколько разделов, то визуальный поиск отнимает много времени, это весьма скучное занятие, и возможны ошибки. Более сложным примером является ситуация, когда объект представлен двумерным прямоугольным изображением или картинкой, которая в рамках компьютерного приложения имеет вид двумерного массива нулевых или единичных битов, указывающих на белый или черный цвет. Такое цифровое изображение может иметь размер 1024×1024 битов и быть выведено на экран компьютера. В этом случае задача может состоять в поиске некоторой специальной формы или конкретного изображения белого или черного цвета, например, креста — ✠.

Метод прямого поиска предполагает сравнение образца с каждым возможным претендентом. В примере с поиском слова "effect" мы бы сравнивали шесть букв слова "effect" с каждой последовательностью шести символов из текста документа. Эта операция потребовала бы в худшем случае сравнения шести букв с каждым шестисимвольным словом документа. Для коротких образов и коротких текстов этот метод может быть достаточно эффективен. Однако, как правило, тексты бывают весьма объемными, а образы — достаточно длинными, что делает метод прямого поиска непрактичным. Эффективным будем называть такой метод, который для решения задачи требует относительно "короткого" времени или относительно "небольшого" числа сравнений или вычислений. Более конкретно, нам нужен алгоритм поиска, использующий минимум памяти, для которого число требуемых сравнений/вычислений линейно зависит от длины образа и текста поиска, и который, что может быть важно для некоторых приложений, функционирует в реальном времени (считывает текст поиска по одному символу, не зная всего текста заранее, и требует для обработки каждого символа текста одно и то же время). Желательно также, чтобы помимо эффективности алгоритм был

применим ко многим типам поисковых контекстов. Ниже мы приводим достаточно общий алгоритм поиска по образцу, разработанный Ричардом М. Карпом и Майклом О. Рабином [56], в основе которого лежит выбор случайных простых чисел. Этот алгоритм обладает описанными выше свойствами и является примером использования вероятностных методов.

Все числа и символы (цифры и буквы алфавита) представлены в компьютере как последовательности нулей и единиц. Например, представлением буквы "e" на основе ASCII 8-битовой кодировки является последовательность битов

$$[0\,1\,1\,0\,0\,1\,0\,1],$$

которая трактуется как двоичное целое число

$$0\cdot 2^7+1\cdot 2^6+1\cdot 2^5+0\cdot 2^4+0\cdot 2^3+1\cdot 2^2+0\cdot 2^1+1\cdot 2^0.$$

Слово "effect" можно представить в виде

$$[[01100101][01100110][01100110][01100101][01100011][01110110].$$

Рассматривая эту последовательность как битовую строку, мы можем трактовать ее как двоичное целое число

$$0\cdot 2^{47}+1\cdot 2^{46}+1\cdot 2^{45}+\cdots+1\cdot 2^1+0\cdot 2^0.$$

Строка битов [01100101] для буквы "e" может быть представлена в шестнадцатеричной системе счисления как 65_{16} или в восьмеричной — как 145_8. При практической реализиции алгоритма можно выбирать более естественную систему счисления, но при описании алгоритма Карпа и Рабина мы будем рассматривать "образ" и "текст" как битовые строки. Пусть X — образ и Y — текст, тогда

$$X=[x_1\,x_2\,x_3\,\cdots\,x_n], \qquad \text{где } x_i=0 \text{ или } 1;$$
$$Y=[y_1\,y_2\,y_3\,\cdots\,y_m], \qquad \text{где } y_i=0 \text{ или } 1.$$

Образ имеет длину n битов и текст имеет длину m битов, где $m\geq n$. Пусть $Y(i)$ — подстрока из n последовательных битов строки Y, которая начинается с позиции i, $1\leq i\leq m-n+1$, так что

$$Y(i)=[y_i\,y_{i+1}\,y_{i+2}\,\cdots\,y_{i+n-1}].$$

Очевидно, для всякого i такого, что $X=Y(i)$, имеется соответствие между образом и частью текста. Заметим, что X и $Y(i)$ можно рассматривать как элементы множества $\{0,1\}^n=\{0,1\}\times\{0,1\}\times\cdots\times\{0,1\}$ — декартовой n-степени множества $\{0,1\}$, или множества n-элементных наборов с элементами из $\{0,1\}$. Пусть $N(X)$ и $N(Y(i))$ — целые двоичные числа, которые имеют представления:

$$N(X)=x_1 2^{n-1}+x_2 2^{n-2}+\cdots+x_{n-1}2^1+x_n;$$
$$N(Y(i))=y_i 2^{n-1}+y_{i+1}2^{n-2}+\cdots+y_{i+n-2}2^1+y_{i+n-1},$$

при этом соответствие имеет место тогда и только тогда, когда $N(X)=N(Y(i))$.

Сравнение $N(X)$ и $N(Y(i))$ потребует не меньше усилий, чем сравнение X и $Y(i)$. Поэтому необходимо "сжать" информацию, содержащуюся в $N(X)$ и $N(Y(i))$, с тем, чтобы это сравнение требовало меньше работы. Одним из способов осуществления такого сжатия является отображение целых чисел $N(X)$ и $N(Y(i))$ на меньшие целые числа, используя модульную арифметику. Пусть p — простое число из интервала $1 \le p \le M$, где M — некоторое достаточно большое положительное целое число, которое будет определено ниже. Пусть N_p — отображение множества битовых строк длины n на множество первичных вычетов по модулю p, $\{0, 1, 2, \ldots, (p-1)\}$, определенное соотношением

$$N_p(W) = [[N(W)]]_p,$$

так что $N_p(X) = [[N(X)]]_p$ и $N_p(Y(i)) = [[N(Y(i))]]_p$. В типичных приложениях величина n может быть порядка 200, а p может находиться в диапазоне от 2^{16} до 2^{64}. Большинство компьютеров могут эффективно сравнивать 16-битовые целые числа с 64-битовыми. Функция N_p вырабатывает "доверенное" число $N_p(W)$ для $N(W)$, которое будет использоваться в сравнениях вместо W или $N(W)$. Такую функцию часто называют ***идентифицирующей функцией***, а $N_p(W)$ называют ***идентификатором*** для W. Заметим, что в этом случае идентифицирующая функция N_p не является однозначной, поэтому можно иметь две битовые строки W и V длины n с одним и тем же идентификатором, т.е. $N_p(W) = N_p(V)$. Если $N_p(W) = N_p(V)$, но $W \neq V$, то в этом случае будем говорить, что имеет место коллизия — ошибочное соответствие. В дальнейшем мы решим эту проблему. Если $N(W)$ сводить по модулю p до $N_p(W)$, начиная с $N(W)$ для каждого новой строки W, то это преобразование потребует значительного количества времени. К счастью, имеется возможность находить последовательность $N_p(Y(1))$, $N_p(Y(2))$, $N_p(Y(3))$, ... без лишних вычислений. Сведение $N(W)$ к $N_p(W)$ может быть выполнено постепенно. Заметим, что

$$\begin{aligned} N(W) &= w_1 2^{n-1} + w_2 2^{n-2} + \cdots + w_{n-1} 2^1 + w_n = \\ &= (\cdots((w_1 \cdot 2 + w_2) \cdot 2 + w_3) \cdot 2 + \cdots + w_{n-1}) \cdot 2 + w_n. \end{aligned}$$

Порядок выполнения операций в последнем выражении определен однозначно. Согласно теореме 3.54, если $a \equiv b \pmod p$, то b можно подставить вместо a в любом выражении, содержащем умножение, сложение и вычитание, не меняя при этом вычет выражения по модулю p. Будем говорить, что операция $a \odot b$ для целых чисел a и b выполнена по модулю p, если после выполнения операции результат сведен по модулю p до наименьшего неотрицательного вычета; т.е. подставляем первичный вычет $[[a \odot b]]_p$ вместо $a \odot b$. Для сокращения записи иногда будем писать $a \odot_p b$ вместо $[[a \odot b]]_p$, чтобы подчеркнуть, что сведение по модулю p должно быть выполнено после выполнения операции. Например, если подставить $w_1 \cdot_p 2$ вместо $w_1 \cdot 2$, а затем подставить $(w_1 \cdot_p 2) +_p w_2$ вместо $(w1 \cdot_p 2) + w_2$, то в результате получим

$$N_p(W) = (\cdots((w_1 \cdot_p 2 +_p w_2) \cdot_p 2 +_p w_3) \cdot_p 2 +_p \cdots +_p w_{n-1}) \cdot_p 2 +_p w_n.$$

Таким образом, приведение $N(W)$ к его первичному вычету по модулю p, $N_p(W)$, может выполняться по одной операции, начиная с самых внутренних круглых скобок. Остается еще необходимость сравнивать X и $Y(i)$ для каждого i, сравнивая

для каждого i "доверенные" целые числа $N_p(X)$ и $N_p(Y(i))$; таким образом, хотя $N_p(Y(i))$ может быть вычислено эффективно для каждого i, алгоритм поиска может оказаться в целом неэффективным. Однако, это тот случай, когда $N_p(Y(i+1))$ можно получить из $N_p(Y(i))$ и y_{i+1}, выполнив несколько арифметических операций, без полного пересчета $N_p(Y(i+1))$ из рабочего файла. Поскольку

$$N(Y(i+1)) = (N(Y(i)) - 2^{n-1} \cdot y_i) \cdot 2 + y_{i+n},$$

то, используя только операции по модулю p, имеем

$$N_p(Y(i+1)) = (N_p(Y(i)) -_p [[2^{n-1}]]_p \cdot_p y_i) \cdot_p 2 +_p y_{i+n}.$$

Такая формула, как описанная выше, позволяющая легко вычислять "следующее" значение по текущему, называется ***корректирующей формулой***. Таким образом, можно сказать, что $N_p(Y(i))$ легко корректируется. Основная идея алгоритма Карпа и Рабина состоит в выборе случайным образом простого числа p из соответствующего множества $1 \le p \le M$ и вычислении идентификатора $N_p(X)$ битовой строки образа X. Далее, необходимо сравнивать $N_p(X)$ с идентификатором $N_p(Y(i))$ каждой последовательности $Y(i)$ из n последовательных битов в текстовой строке битов Y для $i = 1, 2, 3, \ldots, m - n + 1$. Если случается так, что $N_p(X) = N_p(Y(k))$ для некоторого k, то имеем доказательство того, что $X = Y(k)$, и процесс завершен. Обычно, если алгоритм применяется с одним простым числом p, необходимо проверять, выполняется ли соотношение $X = Y(k)$, поскольку функция N_p не является однозначной. Если $X \neq Y(k)$, необходимо продолжить алгоритм для $i = k + 1$; вскоре мы увидим, что использование нескольких простых чисел и нескольких идентифицирующих функций иллюстрирует сущность вероятностной методологии.

АЛГОРИТМ ПОИСКА ПО ОБРАЗУ (Карп и Рабин) Предположим, что битовой строкой X длины n задан образ; битовой строкой Y длины m, где $m \ge n$, задан текст; задано целое положительное число M, которое будет зависеть от m и n. Пусть $S = \{p : p$ простое число и $1 \le p \le M\}$. Для простого числа p из множества S пусть $N_p : \{0,1\}^n \to \{0, 1, 2, \ldots, (p-1)\}$ — идентифицирующая функция. Выберем случайным образом k простых чисел p_1, p_2, ..., p_k из множества $\{1, 2, \ldots, M\}$, выбирая случайным образом целые числа из совокупности $\{1, 2, 3, \ldots, M\}$, проверяя, являются ли они простыми, и останавливая процесс поиска, когда k простых чисел найдены. Процедура поиска состоит в следующем. Для каждого i, $1 \le i \le m-n+1 = t$, сравнить $N_{p(j)}(X)$ с $N_{p(j)}(Y(i))$ для каждого простого числа $p_j = p(j)$, где $1 \le j \le k$, пока не найдется такое i, что

$$N_{p(j)}(X) = N_{p(j)}(Y(i)) \text{ для всех } j,\ 1 \le j \le k,$$

и тогда объявляется, что "соответствие" имеет место, или пока i не достигнет значения $i = t$, что будет означать, что "соответствие" не найдено.

Отметим, что для полного соответствия необходимо, чтобы все k идентифицирующих функций давали соответствие. Другим важным моментом является выбор простых чисел $p_1, p_2, \ldots, p_k$ случайным образом из достаточно большого множества. При этом наличие ошибочного соответствия все еще возможно, а

именно, когда $N_{p(j)}(X) = N_{p(j)}(Y(i))$ для всех j, но $X \neq Y(i)$. Суть, однако, заключается в том, что вероятность или возможность ошибочного соответствия может быть сделана сколь угодно малой. Эта последняя особенность воплощает в себе то, что понимают под вероятностным алгоритмом.

ПРИМЕР 22.1. Чтобы лучше понять идею приведенного выше алгоритма, рассмотрим в качестве примера небольшую задачу, предназначенную для иллюстрации применения метода. Предположим, что

$$X = [1001\,1011] \text{ где } n = 8$$

и

$$Y = [1011011001101100 \cdots 0], \text{ где } m = 300.$$

Пусть $t = m - n + 1 = 293$ и $M = nt^2 = 686792$. Предположим, что $k = 2$ и случайным образом получены простые числа $p_1 = 47$ и $p_2 = 31$. Вычисляя первичные вычеты, получим сначала идентификаторы строки X

$$N_{p(1)}(X) = [[N(X)]]_{p(1)} = [[155_{10}]]_{47} = 14;$$
$$N_{p(2)}(X) = [[N(X)]]_{p(2)} = [[155_{10}]]_{31} = 0.$$

Начиная с $i = 1$, вычислим идентификаторы (строки) $Y(i)$, взамно простые с простыми числами 47 и 31, и проверим на соответствие. Соответствия с каждым из двух идентификаторов отмечены звездочкой (*).

i	$Y(i)$	$N(Y(i))$	$N_{47}(Y(i))$	$N_{31}(Y(i))$	Соответствует?
1	[1011 0110]	182	41	27	Нет
2	[0110 1100]	108	14 *	15	Нет
3	[1101 1001]	217	29	0 *	Нет
4	[1011 0011]	179	38	24	Нет
5	[0110 0110]	102	8	9	Нет
6	[1100 1101]	205	17	19	Нет
7	[1001 1011]	155	14 *	0 *	Да

Первое соответствие наблюдается на позиции 7 в битовой строке Y. Следует отметить, что в этом примере мы провели прямое вычисление вычетов, поскольку числа достаточно малы. Компьютерная реализация потребовала бы использования модулярных операций и корректирующей формулы, описанных выше. □

Следуя Карпу и Рабину, получим теперь предельное значение вероятности ошибочного соответствия при использовании этого алгоритма. Для начала нам потребуются два понятия: вероятности и теоретико-числовой функции.

Пусть T — непустое множество "возможностей", и любая из этих возможностей может быть выбрана. Пусть подмножество A множества T содержит интересующие нас возможности. Выберем один элемент из T. Вероятность выбора элемента из A определена соотношением

$$\text{PROB}(A) = \frac{\text{Количество элементов в } A}{\text{Количество элементов в } T}.$$

В такой ситуации будем говорить, что элемент выбран случайным образом из множества T (каждый элемент множества T имеет такую же возможность быть выбранным, как и любой другой) и событие A имеет место, если выбранный элемент принадлежит множеству A. Формула, приведенная выше, дает вероятность события A. Например, предположим, что T — множество граней игрального кубика, так что $T = \{1, 2, 3, 4, 5, 6\}$, и мы бросаем кубик один раз (выбираем сторону случайным образом). Пусть A — событие (множество возможных исходов), состоящее в получении числа точек, выраженного простым числом. $A = \{2, 3, 5\}$. Тогда $\mathrm{PROB}(A) = 3/6 = 0.5$. Очевидно, вероятность любого события лежит между 0 и 1 включительно.

Пусть $\pi(w)$ количество простых чисел, которые меньше или равны w. Таким образом, $\pi(1) = 0$, $\pi(2) = 1$, $\pi(5) = 3$ и $\pi(20) = 8$.

Сформулируем два результата, полученные Дж. Россером и Л. Шонфилдом в работе [103].

ТЕОРЕМА 22.2. Если $w \geq 29$, то

$$p_1 p \cdots p_{\pi(w)} > 2^w,$$

где $p_1, p_2, p_3, \ldots, p_{\pi(w)}$ — простые числа, которые меньше или равны w.

ТЕОРЕМА 22.3. Если $w \geq 17$, то

$$\frac{w}{\log(w)} \leq \pi(w) \leq \frac{1.25506w}{\log(w)}.$$

ТЕОРЕМА 22.4. Если $w \geq 29$ и $b \leq 2^w$, то b имеет менее, чем $\pi(w)$ простых делителей.

ДОКАЗАТЕЛЬСТВО. Допустим, что $w \geq 29$ и $b \leq 2^w$, но b имеет не менее $\pi(w)$ простых делителей. Пусть $d_1 < d_2 < \cdots < d_r$ — простые делители числа b, так что $r \geq \pi(w)$. Пусть $p_1 < p_2 < \cdots < p_k$ — первые k простых чисел для любого k. Тогда

$2^w \geq b$	дано
$b \geq d_1 d_2 d_3 \cdots d_r$,	поскольку $d_1 d_2 \cdots d_r \mid b$
$d_1 d_2 \cdots d_r \geq p_1 p_2 \cdots p_r$,	поскольку $d_i \geq p_i$ для каждого i
$p_1 p_2 \cdots p_r \geq p_1 p_2 \cdots p_{\pi(w)}$,	поскольку $r \geq \pi(w)$
$p_1 p_2 \cdots p_{\pi(w)} > 2^w$	по теореме 22.2.

Но $2^w > 2^w$ является противоречием. Таким образом, b должно иметь менее, чем $\pi(w)$ простых делителей. ■

Как упоминалось выше, рассмотренный алгоритм поиска имеет возможность завершения с ошибочным соответствием. От этого алгоритма мало пользы, если ошибочное соответствие будет наблюдаться часто. Приведенная ниже теорема утверждает, что вероятность получения ошибочного соответствия в результате применения алгоритма не превышает некоторое определенное значение.

ТЕОРЕМА 22.5. Если алгоритм применен к задаче поиска для X и Y с характеристиками n, m, t, M, k и S, то вероятность ошибочного соответствия меньше или равна

$$t\left(\frac{\pi(n)}{\pi(M)}\right)^k \quad \text{при } n \geq 29.$$

ДОКАЗАТЕЛЬСТВО. Допустим, что $n \geq 29$, $1 \leq r \leq t$ и $X \neq Y(r)$. Поскольку $0 \leq N(X) < 2^n$ и $0 \leq N(Y(r)) < 2^n$, то $|N(X) - N(Y(r))| < 2^n$. По теореме 22.4 количество простых чисел, которые делят $|N(X) - N(Y(r))|$, меньше, чем $\pi(n)$; следовательно, для любого j, $1 \leq j \leq k$, учитывая, что простое число p_j выбрано случайным образом из множества S, в котором имеется $\pi(M)$ элементов, имеем

$$\text{PROB}\,(p_j \text{ делит } |N(X) - N(Y(r))|) \leq \frac{\pi(n)}{\pi(M)}.$$

Поскольку k простых чисел $p_1, p_2, \ldots, p_k$ выбраны случайно и независимо, то по свойству вероятности пересечения k таких событий имеем

$$\text{PROB}\,(p_j \text{ делит } |N(X) - N(Y(r))| \text{ для всех } j) \leq \left(\frac{\pi(n)}{\pi(M)}\right)^k.$$

Поскольку имеется t взаимоисключающих событий для r, то по свойству вероятности все простые числа $p_1, p_2, \ldots, p_k$ делят $|N(X) - N(Y(r))|$ для некоторого r, и сама вероятность меньше или равна $t \cdot \left(\frac{\pi(n)}{\pi(M)}\right)^k$. ■

ТЕОРЕМА 22.6. При выполнении предположений теоремы 22.5, если $M = nt^2$, где $n \geq 29$, то вероятность ошибочного соответствия меньше или равна

$$(1.25506)^k\, t^{-(2k-1)}\, (1 + 0.6\, \log(t))^k.$$

ДОКАЗАТЕЛЬСТВО. Из теоремы 22.5 следует, что

$$\text{PROB}\,(\text{ошибочного соответствия}) \leq t \cdot \left(\frac{\pi(n)}{\pi(M)}\right)^k.$$

Из теоремы 22.3 следует, что

$$\pi(n) \leq c\frac{n}{\log(n)}, \text{ где } c = 1.25506, \text{ и } \frac{nt^2}{\log(nt^2)} \leq \pi(nt^2).$$

Таким образом,

$$\text{PROB}\,(\text{ошибочного соответствия}) \leq t \cdot \left[\frac{c \cdot n}{\log(n)} \cdot \frac{\log(nt^2)}{nt^2}\right]^k =$$

$$= t \cdot \left[\frac{c}{t^2} \cdot \left(1 + 2 \cdot \frac{\log(t)}{\log(n)}\right)\right]^k.$$

Но для $n \geq 29$ имеем $\log(n) \geq \log(29)$. Следовательно, теорема верна, поскольку непосредственное вычисление показывает, что $\frac{2}{\log(n)} \leq \frac{2}{\log(29)} \leq 0.6$. ■

Для иллюстрации теоремы 22.6 предположим, что имеется задача сопоставления строк с образом X длины $n0$, которая соответствует 25 в 8-битовой ASCII-кодировке, с текстовой строкой Y длины 32000, соответствующей 4000 в кодировке ASCII. Тогда $t = 31801$ и $M = nt^2 < 200(32000)^2 = 2^{11} \cdot 10^8 \approx 2^{11} \cdot 2^{26.6} < 2^{38}$, так что компьютерные представления простых чисел p_i и соответствующих вычетов в двоичной системе счисления будут иметь менее 38 битов. Если $k = 4$, то вероятность ошибочного соответствия не будет превосходить величину $2 \cdot 10^{-28}$.

Алгоритм Карпа и Рабина является алгоритмом, выполняющимся в реальном времени и требующим одинакового времени вычисления для каждого символа текста. Он также требует постоянного количества памяти, зависящего только от n, k и t. Можно доказать, что алгоритм имеет малую вероятность ошибки. Более того, метод можно леко обобщить для применения в многомерных задачах. Помимо N_p этот метод может использовать и другие "идентифицирующие" функции. Заметим, что Карпу и Рабину принадлежат другие подобные, но отличные от вероятностных алгоритмы поиска по образцу, которые можно найти в ссылках, приведенных ранее. Информацию о некоторых эффективных невероятностных алгоритмах поиска для сравнения можно найти в работах (1) *Fast Pattern Matching in Strings*, авторы Д. Э. Кнут, Дж. Х. Моррис и В. Р. Пратт [65]; (2) *A Fast String Searching Algorithm*, авторы Р. С. Боер и Дж. С. Мур [12]; (3) *Pattern Matching Alternatives: Theory vs. Practice*, автор Джек Пардум [93] и (4) *A Very Fast Substring Search Algorithm*, автор Даниэль М. Сандей [111].

■ УПРАЖНЕНИЯ

1. Докажите, что если W и V — битовые строки длины n, то
$$|N(W) - N(V)| < 2^n.$$
2. В теореме 22.2 было показано, что произведение всех простых чисел, которые меньше или равны w, превышает величину 2^w, если $w \geq 29$. Определите, для какого положительного целого числа $w < 29$ также выполняется утверждение этой теоремы.
3. В теореме 22.3 были показаны два неравенства, которые имеют место для каждого положительного числа $w \geq 17$. Используя непосредственные вычисления, покажите, какое положительное целое число $w < 17$ также удовлетворяет этим неравенствам.
4. Для образа X длины $n = 0$ и текста длины $m\,000$, используя теорему 22.6, определите верхнюю границу вероятности ошибочного соответствия для $k = 1, 2, 3$ и 4.
5. Предположим, что задан образ длины n битов, количество простых чисел k и такое число M, что все идентифицирующие простые числа и их вычеты имеют не более, чем 32 бита. Кроме того, будет использована описанная в данном разделе корректировка. Определите количество 32-битовых компьютерных слов информационной памяти, необходимых для реализации алгоритма Карпа и Рабина.

22.2. ПРИЛОЖЕНИЕ: ФУНКЦИИ ХЕШИРОВАНИЯ

В компьютерных приложениях часто возникает ситуация, когда имеется n наборов информации $I_1, I_2, \ldots, I_n$, где каждый набор идентифицирован ключом, например, $k_1, k_2, \ldots, k_n$, соответственно. При этом желательно иметь возможность быстро находить информацию I_j, если известен ключ k_j. Ключи можно рассматривать как целые числа, хотя они могут иметь вид конечных последовательностей символов алфавита. Просмотр списков ключей может занимать достаточно много времени; более того, если поиск с использованием ключей необходимо проводить неоднократно, то вычислительные затраты могут быть значительными. Компилятор — это компьютерная программа, преобразующая другую компьютерную программу, написанную на языке программирования высокого уровня (напрмер, ФОРТРАН или Паскаль), в эквивалентную программу на машинном языке. В течение процесса трансляции компилятор должен генерировать таблицы и неоднократно обращаться к ячейкам таблиц для получения информации. Важным моментом здесь является то, что поисковая таблица создается один раз, а используется многократно. Поскольку таких таблиц может быть достаточно много, необходимо также использовать как можно меньше памяти для хранения.

Одно из решений проблемы "размещения" состоит в использовании ***функции хеширования*** $h : K \to \{0, 1, \ldots, m\}$, где $K = \{k_1, k_2, \ldots, k_n\}$ и $m + 1 \geq n$. Таким образом, мы обеспечиваем $m + 1$ компьютерных ячеек для хранения n наборов информации. Итак, чтобы найти I_j, нужно вычислить $h(k_j)$ и перейти в ячейку $h(k_j)$, где хранится I_j. Вообще, найти идеальную функцию хеширования, т.е. такую что $h(k_r) \neq h(k_s)$, когда $k_r \neq k_s$ и $m + 1 = n$, достаточно трудно. Наиболее широко используемые функции хеширования отображают множество K во множество, содержащее более, чем n элементов, поэтому не являются инъективными. Для функции h, которая не является взаимно однозначным отображением, должно быть выработано соглашение поведения при возникновении коллизии, т.е. в ситуации, когда k_r и k_s — разные ключи, но $h(k_r) = h(k_s)$. Дональд Э. Кнут в томе 3 книги *Искусство программирования* [64] рассматривал различные способы выбора функций хеширования и методы разрешения коллизий. Яэшке в книге *Reciprocal Hashing: A Method for Generating Minimal Perfect Hashing Functions* [51] описывает особый тип минимальной функции хеширования, т.е. такой функции хеширования, которая является инъективной и использует минимальный объем памяти (***минимальная совершенная функция хеширования***) . Для функции хеширования, которую ввел Яэшке, необходима возможность отображения ключей $\{k_1, k_2, \ldots, k_n\}$, которые могут не быть попарно взаимно простыми, во множество попарно взаимно простых целых чисел $\{f(k_1), f(k_2), \ldots, f(k_n)\}$. Он доказал следующую теорему.

ТЕОРЕМА 22.7. Пусть $K = \{k_1, k_2, \ldots, k_n\}$ — множество n различных положительных целых чисел. Пусть существуют целые числа D и E такие, что если функция $f(x)$ определена соотношением $f(x) = Dx + E$, то элементы набора положительных целых чисел $\{f(k_1), f(k_2), \ldots, f(k_n)\}$ являются попарно взаимно простыми. Таким образом, целые числа $Dk_1 + E$, $Dk_2 + E$, ... и $Dk_n + E$ являются попарно взаимно простыми.

ДОКАЗАТЕЛЬСТВО. Доказательство теоремы является слишком длинным и при-

водиться не будет (см. [51]). Тем не менее, далее рассмотрим специальный случай, вполне адекватный нашим целям. ■

Основной результат, который получил Яэшке, сформулирован в приведенной ниже теореме. В ней $\lfloor w \rfloor$ — набольшее положительное целое число, которое меньше или равно w.

ТЕОРЕМА 22.8. Если множество $K = \{k_1, k_2, \ldots, k_n\}$ содержит различные целые положительные ключи, то в нем существуют целые числа C, D и E такие, что

$$h(x) = \left[\left[\left\lfloor \frac{C}{Dx+E} \right\rfloor\right]\right]_n \equiv \left\lfloor \frac{C}{Dx+E} \right\rfloor \pmod{n},$$

где $x \in K$ является минимальной совершенной функцией хеширования.

ДОКАЗАТЕЛЬСТВО. Пусть $k_1 < k_2 < \cdots < k_n$ и пусть целые числа D и E заданы в соответствии с теоремой 22.7, так что $f(k_j) = Dk_j + E$ и НОД$(f(k_i), f(k_j)) = 1$ для $i \neq j$. Поскольку D и E могут быть выбраны таким образом, что $f(k_j) > n$ для каждого j, имеется возможность выбрать n целых чисел $a_1, a_2, \ldots, a_n$ таких, что $a_i \not\equiv a_j \pmod{n}$ и таких, что

$$(i-1) \cdot (Dk_i + E) \leq a_i < i \cdot (Dk_i + E).$$

Тогда, согласно обобщенной китайской теореме об остатках, применимой к модулям, не являющимся взаимно простыми (теорема 10.14), существует такое число C, что

$$\begin{aligned} C &\equiv a_1 \pmod{n(Dk_1 + E)}, \\ C &\equiv a_2 \pmod{n(Dk_2 + E)}, \\ &\vdots \\ C &\equiv a_n \pmod{n(Dk_n + E)}. \end{aligned}$$

Поэтому, существуют целые числа q_i такие, что $C = q_i[n(Dk_i + E)] + a_i$ для всех i. Следовательно,

$$\left\lfloor \frac{C}{Dk_i + E} \right\rfloor = q_i n + (i-1) \equiv i - 1 \pmod{n}.$$

Из этого следует, что функция h, определенная в теореме, обладает тем свойством, что $h(k_i) \neq h(k_j)$, когда $i \neq j$, поэтому функция h является инъективной. Более того, множеством значений функции h является $\{0, 1, \ldots, (n-1)\}$. ■

Яэшке привел алгоритмы вычисления чисел C, D и E; однако, эти методы не являются конструктивными. С другой строны, Чанг и Ши в работе *Pairwise Relatively Prime Generating Polynomials and Their Applications* [19] дают способ вычисления требуемых многочленов вида $f(x) = Dx + 1$, где $E = 1$. Доказательство приведенной ниже леммы предоставляется читателю.

ЛЕММА 22.9. Если a и b положительные целые числа, $a > b$ и d кратно $a - b$, то $d(a-b)$ и $da + 1$ — взаимно простые числа.

ТЕОРЕМА 22.10. Пусть $K = \{k_1, k_2, \ldots, k_n\}$ — множество, содержащее n различных положительных целых чисел, для которых $k_i < k_{i+1}$. Если $\{t_1, t_2, \ldots, t_s\} = \{k_i - k_j : 1 \le j < i \le n\}$ — множество, содержащее $s = n(n-1)/2$ разностей, то $D = w \cdot \text{НОК}(t_1, t_2, \ldots, t_s)$, где w — произвольное положительное целое число, обладает тем свойством, что $Dk_1 + 1, Dk_2 + 1, \ldots, Dk_n + 1$ являются попарно взаимно простыми числами.

ДОКАЗАТЕЛЬСТВО. Поскольку $\text{НОД}(a, b) = \text{НОД}(a - b, a)$, где a и b — целые числа, для которых $\text{НОД}(a, b)$ и $\text{НОД}(a - b, b)$ определены, получаем, что $\text{НОД}(Dk_i + 1, Dk_j + 1) = \text{НОД}(D(k_i - k_j), Dk_i + 1)$ для $i > j$, где D задано согласно теореме. По определению, D кратно $(k_i - k_j)$, следовательно, из леммы 22.9 следует, что $\text{НОД}(Dk_i + 1, Dk_j + 1) = 1$ для $i > j$. ■

Заметим, что для доказательства теоремы требуется только, чтобы D было кратно всем $t_1, t_2, \ldots, t_s$. Естественно, можно было бы выбирать D как можно меньшим ($w = 1$); однако, для доказательства теоремы 22.8 необходимо, чтобы D было таким, чтобы $f(k_i) = Dk_i + 1 > n$. Такое D может быть всегда получено за счет выбора w достаточно большим.

ПРИМЕР 22.11. Предположим, что $K = \{3, 6, 7, 12\}$. Тогда

$$\begin{aligned}\{k_i - k_j : 1 \le j < i \le 4\} &= \{6-3, 7-3, 12-3, 7-6, 12-6, 12-7\} = \\ &= \{3, 4, 9, 1, 6, 5\},\end{aligned}$$

так что $D = \text{НОК}(3, 4, 9, 1, 6, 5) = 180$. Таким образом, $f(x) = 180x + 1$ и

i	k_i	$f(k_i) = Dk_i + 1$
1	3	$180 \cdot 3 + 1 = 541$
2	6	$180 \cdot 6 + 1 = 1081$
3	7	$180 \cdot 7 + 1 = 1261$
4	12	$180 \cdot 12 + 1 = 2161$

Проверка показывает, что любые два из чисел 541, 1081, 1261 и 2161 являются взаимно простыми. □

Для реализации алгоритма обратного хеширования, предложенного Яэшке, необходимо иметь возможность вычислять целое число C. К счастью, Чанг и Ши в работе *A Fast Algorithm for Constructing Reciprocal Hashing Functions* [18] представили алгоритм вычисления целого числа C.

В оставшейся части этого раздела мы будем использовать как функцию целой части $\lfloor w \rfloor$, описанную ранее, так и функцию округления $\lceil w \rceil$, равную наименьшему положительному целому числу, большему или равному w. Таким образом, $\lceil 5.23 \rceil = 6$ и $\lceil 8 \rceil = 8$.

ТЕОРЕМА 22.12. Если

а) $m_1, m_2, \ldots, m_n$ — попарно взаимно простые целые числа;
б) $m_i > n$ для всех i, $1 \le i \le n$;

в) $M = \prod_{j=1}^{n} m_j$;

г) $M_i = n \prod_{j \neq i} m_j = \dfrac{nM}{m_i}$ для всех i;

д) b_i является таковым, что $M_i b_i \equiv n \pmod{nm_i}$, и

е) $N_i = \left\lceil \dfrac{(i-1)m_i}{n} \right\rceil$ для каждого i,

тогда

1. $\left\lfloor \dfrac{\sum_{j=1}^{n} M_j b_j N_j}{m_i} \right\rfloor \equiv (i-1) \pmod{n}$ для всех i,

2. $C = \left[\left[\sum_{j=1}^{n} M_j b_j N_j\right]\right]_{nM}$ — наименьшее положительное целое число такое, что $\left\lfloor \dfrac{C}{m_i} \right\rfloor \equiv (i-1) \pmod{n}$ для всех i.

ДОКАЗАТЕЛЬСТВО. Пусть $a_i = N_i \cdot n$, тогда a_i — наименьшее число, кратное n, так что $(i-1) \cdot m_i \leq a_i < i \cdot m_i$ для всех i и $a_i \not\equiv a_j \pmod{n}$, когда $i \neq j$. Сумма $\sum_{j=1}^{n} M_j b_j N_j$ является решением n сравнений

$$\begin{aligned} x &\equiv a_1 \pmod{nm_1}; \\ x &\equiv a_2 \pmod{nm_2}; \\ &\vdots \\ x &\equiv a_n \pmod{nm_n}, \end{aligned}$$

потому что, если $j \neq i$, то M_j содержит множитель nm_i. Поэтому $M_j b_j N_j \equiv 0 \pmod{nm_i}$ и

$$\sum_{j=1}^{n} M_j b_j N_j \equiv M_i b_i N_i \pmod{nm_i} \equiv$$
$$\equiv nN_i = a_i \pmod{nm_i}.$$

Поэтому,

$$W_i = \frac{\sum_{j=1}^{n} M_j b_j N_j}{m_i} = \frac{\sum_{j \neq i}^{n} M_j b_j N_j}{m_i} + \frac{M_i b_i N_i}{m_i} = nJ_i + \frac{M_i b_i N_i}{m_i}$$

для соответствующего целого числа J_i. Поскольку $M_i b_i \equiv n \pmod{nm_i}$, то существует целое число t_i такое, что $M_i b_i = t_i(nm_i) + n$, поэтому

$$W_i = nJ_i + t_i nN_i + \frac{nN_i}{m_i}.$$

Таким образом, учитывая $(i-1)\cdot m_i \le a_i < i\cdot m_i$, имеем

$$\lfloor W_i \rfloor \equiv \left\lfloor \frac{nN_i}{m_i} \right\rfloor \equiv \left\lfloor \frac{a_i}{m_i} \right\rfloor \equiv (i-1) \pmod n,$$

что доказывает часть 1.

Пусть $C = \left[\left[\sum_{j=1}^{n} M_j b_j N_j\right]\right]_{nM}$. Нам нужно показать, что $\left\lfloor \frac{C}{m_i} \right\rfloor \equiv (i-1) \pmod n$ для $1 \le i \le n$. Согласно алгоритму деления, существует целое число J такое, что $\sum_{j=1}^{n} M_j b_j N_j = J(nM) + C$. Для $1 \le i \le n$, опять согласно алгоритму деления, существуют целые числа q_i и r_i, где $0 \le r_i < m_i$ такие, что

$$\sum_{j=1}^{n} M_j b_j N_j = q_i m_i + r_i,$$

поэтому

$$\left\lfloor \frac{\sum_{j=1}^{n} M_j b_j N_j}{m_i} \right\rfloor = q_i.$$

Согласно результату части 1 существует целое число t_i такое, что

$$\begin{aligned} q_i &= (i-1) + t_i n; \\ q_i m_i &= (i-1)m_i + t_i n m_i; \\ q_i m_i + r_i &= (i-1)m_i + t_i n m_i + r_i; \\ J(nM) + C &= (i-1)m_i + t_i n m_i + r_i, \end{aligned}$$

поэтому

$$\frac{C}{m_i} = (i-1) + \left(t_i - \frac{JM}{m_i}\right) n + \frac{r_i}{m_i}$$

и

$$\left\lfloor \frac{C}{m_i} \right\rfloor = (i-1) + \left(t_i - \frac{JM}{m_i}\right) n,$$

так что

$$\left\lfloor \frac{C}{m_i} \right\rfloor \equiv (i-1) \pmod n.$$

Читателю остается доказать, что если C' — любое другое положительное целое число такое, что $\left\lfloor \frac{C'}{m_i} \right\rfloor \equiv (i-1) \pmod n$ для всех i, $1 \le i \le n$, то $C' > C$. ■

Объединяя последние несколько результатов, получаем следующую теорему.

ТЕОРЕМА 22.13. Пусть $K = \{k_1, k_2, \ldots, k_n\}$ — множество, содержащее n различных положительных целых чисел; пусть D и E — целые числа и пусть $f(x) = Dx + E$ — функция, описанная в теоремах 22.8 22.10. Предположим также, что число C задано согласно части 2 теоремы 22.12, где $m_i = f(k_i)$. Тогда

$$h(x) = \left[\left[\left\lfloor \frac{C}{Dx+E} \right\rfloor\right]\right]_n$$

является минимальной совершенной функцией хеширования.

ДОКАЗАТЕЛЬСТВО. Если $m_i = f(k_i)$, то из теоремы 22.12 следует, что

$$h(k_i) = \left[\left[\left\lfloor \frac{C}{f(k_i)} \right\rfloor\right]\right] = (i-1) \text{ для } 1 \le i \le n.$$

Следовательно, h является биекцией на $\{0, 1, \ldots, (n-1)\}$. ■

ПРИМЕР 22.14. Вернемся к примеру 22.11, в котором $K = \{k_1, k_2, k_3, k_4\} = \{3, 6, 7, 12\}$ и $f(x) = 180x + 1$. Было показано, что $\{f(k_1), f(k_2), f(k_3), f(k_4)\} = \{541, 1081, 1261, 2161\}$. Пусть $m_i = f(k_i)$. Числа m_i являются попарно взаимно простыми и $m_i > n = 4$. Используя теорему 22.12, получаем следующую таблицу

i	m_i	M_i	b_i	N_i	$M_i b_i N_i$	$h(k_i)$
1	541	11782990804	235	0	0	0
2	1081	5896945444	12	273	19318393274544	1
3	1261	5055192724	172	631	548650176721168	2
4	2161	2949837124	1303	1621	6230536829339212	3

Например, чтобы найти b_3, решаем сравнение $5055192724 \cdot b_3 \equiv 4 \pmod{5044}$, что эквивалентно решению сравнения $1263798181 \cdot b_3 \equiv 1 \pmod{1261}$ или решению сравнения $22 \cdot b_3 \equiv 1 \pmod{1261}$. Следовательно, требуется найти b_3 и y такие, что $22b_3 + 1261y = 1$. Используя алгоритм Евклида и обратную подстановку, приходим к результату $22 \cdot 172 + (-3) \cdot 1261 = 1$. Таким образом,

$$C = \left[\left[\sum_{j=1}^{n} M_j b_j N_j\right]\right]_{nM} = [[6798505399334924]]_{nM}\, 183904723300.$$

Далее используем C, f и h для хеширования $k_3 = 7$.

$$h(k_3) = h(7) = \left[\left[\left\lfloor \frac{3183904723300}{1261} \right\rfloor\right]\right]_4 = [[2524904618]]_4 = 2. \quad \square$$

Обратные функции хеширования, введенные Яэшке, порождают большие целые числа. Вычислительные затраты, необходимые для реализации этого, должны быть соизмеримы с ресурсами, необходимыми для использования неминимальной и несовершенной функции хеширования: (1) максимально возможные затраты памяти, (2) возможность разрешения коллизий и (3) использование алгоритмов поиска для нахождения элементов данных в таблице.

■ УПРАЖНЕНИЯ

1. Докажите, что если a и b — целые числа, для которых определены НОД(a, b) и НОД$(a - b, a)$, то НОД$(a, b) =$ НОД$(a - b, a)$.
2. Для каждого из приведенных ниже множеств найдите многочлен $f(x) = Dx+1$, который биективно отображает это множество на множество попарно взаимно простых чисел:
   а) $\{12, 15, 18, 24\}$;
   б) $\{5, 15, 20, 25\}$.
3. Пусть $a_1, a_2, \ldots, a_n$ являются n различными положительными целыми числами. Докажите или опровергните утверждение: НОД$(a_1, a_2, \ldots, a_n) = 1$ тогда и только тогда, когда НОД$(a_i, a_j) = 1$ для всех $i \neq j$.
4. В доказательстве теоремы 22.12 покажите, что $a_i = N_i n$ является наименьшим числом, кратным числу n, так что $(i - 1) \cdot m_i \leq a_i < i \cdot m_i$.
5. Проверьте числа в таблице примера 22.14.
6. Для
   а) $K = \{12, 15, 18, 24\}$ из упражнения 2(а);
   б) $K = \{5, 15, 20, 25\}$ из упражнения 2(б).
   вычислите, используя теорему 22.12, целое число C и, соответственно, минимальную совершенную обратную функцию хеширования из теоремы 22.13. Проверьте тот факт, что $h(k_i) = i - 1$ при любом i.

22.3. ПРИЛОЖЕНИЕ: КРИПТОГРАФИЯ

Во всем мире правительства, компании и обычные люди хотят отправлять сообщения таким образом, чтобы прочитать их мог только адресат. Военные посылают приказы, банки передают информацию о движении капиталов, люди делают покупки, используя кредитные карточки. Сообщения, посланные по телефонным линиям, по радио, через спутник или через компьютерные сети могут быть перехвачены посторонними, возможно, враждебными лицами. Чтобы не допустить возможности получения информации или ее изменения посторонними лицами, сообщения изменены таким образом, что оригинал скрыт. Процесс засекречивания информации называется шифрованием или кодированием, а процесс рассекречивания — дешифровкой или декодированием. Исходное сообщение называется открытым или незашифрованным текстом, а засекреченное сообщение — зашифрованным текстом (см. рис. 22.1).

Конечно, и отправитель, и получатель сообщения должны согласовать процедуру шифрования и дешифровки. Эта процедура обычно состоит из общего метода и "ключа". Ключ обеспечивает специфическую информацию, что вместе с общим методом позволяет отправителю легко зашифровать, а получателю — легко дешифровать сообщение. Однако, в отсутствие ключа любой, перехвативший сообщение, не может за приемлемое время, используя самые мощные вычислительные ресурсы, получить исходный текст сообщения, даже зная общий метод шифрования. Например, восстановление сообщения при использовании самого быстродействующего компьютера потребовало бы, скажем, миллион лет или



Рис. 22.1

исследования всех возможных сообщений на основе полного перебора, используя метод проб и ошибок. Попытка расшифровать открытый текст по зашифрованному тексту без использования общего метода или ключа называется криптоанализом. Рис. 22.1, на котором E — функция шифрования, а D — функция дешифровки, иллюстрирует механизм поиска. Таким образом, $D(E(M)) = M$. Криптография — это наука, которая изучает методы шифрования и криптоанализа этих методов.

Шифрование и дешифровка обычно выполняются компьютерами, поэтому сообщение M сначала преобразуется в целое число или в последовательность целых чисел с тем, чтобы появилась возможность преобразовывать эти числа, используя компьютерную арифметику. Таким образом, будем считать, что M — целое число. Сообщение в виде целого числа M обычно делят на блоки $M_1, M_2, M_3, \ldots, M_k$, так что каждый блок представляет собой целое число M_i, где $0 \le M_i < n$, и число n выбрано заранее. В таком случае шифрование выполняется блок за блоком как $E(M_i) = C_i$, и дешифровка проводится аналогично относительно $D(C_i) = M_i$. Например, можно следующим образом использовать 8-битовые компьютерные коды ASCII для каждого символа алфавита:

$$\begin{array}{ll} M & 01001101_2 = 77_{10}, \\ O & 01001111_2 = 79_{10}, \\ N & 01001110_2 = 78_{10}, \\ E & 01001010_2 = 69_{10}, \\ Y & 01011001_2 = 89_{10}. \end{array}$$

Тогда M_i могло бы соответствовать группам из двух букв, что дает

$$\text{“}MO\text{”} \rightarrow M_1 = 0100110101001111_2 = 19791_{10},$$

так что $0 \le M_i < n = 2^{16} = 65536$. Конечно, полное сообщение можно было бы рассматривать как единую строку двоичных цифр, которые затем можно группировать в последовательность блоков любого удобного размера.

Р. Л. Райвест, А. Шамир и Л. Адлеман в работе *Method for Obtaining Digital Signatures and Public Key Cryptosystems* [98] представили общий метод

шифрования, известный по фамилиям авторов как RSA-метод. Метод обладает многими хорошими характеристиками.

Пусть p и q — два достаточно больших числа, каждое из которых имеет, например, 100 разрядов, и пусть $n = p \cdot q$. Для шифрования и дешифровки будут использованы два целых числа e и d, которые связаны с числами p, q и n. Оба числа e и d будут определены ниже. Для шифровки сообщения $M = M_i$ вычислим

$$C = E(M) = [[M^e]]_n,$$

а для дешифровки вычислим

$$M = D(C) = \left[\left[C^d\right]\right]_n.$$

Поскольку произведения M^e и C^d приведены по модулю n, то зашифрованный текст и блоки исходного текста — это целые числа в диапазоне от 0 до $(n-1)$. Ключ для шифрования — пара целых чисел (n, e), и ключ для дешифровки — пара целых чисел (n, d).

Сначала выберем в качестве d "большое" целое число, взаимно простое с произведением $(p-1) \cdot (q-1)$. Затем определим целое число e, $0 \leq e < n$, т.е. единственное решение (теорема 3.65) сравнения

$$d \cdot x \equiv 1 \pmod{(p-1)(q-1)}.$$

Следовательно, учитывая соотношение $\phi(n) = \phi(p)\phi(q) = (p-1)(q-1)$, имеем

$$d \cdot e \equiv 1 \pmod{\phi(n)}.$$

ТЕОРЕМА 22.15. Если

а) $n = p \cdot q$, где p и q — различные простые числа;
б) НОД$(d, \phi(n)) = 1$;
в) $d \cdot e \equiv 1 \pmod{\phi(n)}$;
г) для $0 \leq J < n$, определим функции

$$E(J) = [[J^e]]_n \text{ и } D(J) = \left[\left[J^d\right]\right]_n,$$

тогда для $0 \leq J < n$

$$D(E(J)) = J \text{ и } E(D(J)) = J.$$

ДОКАЗАТЕЛЬСТВО. Пусть $0 \leq J < n$. Из соотношения $A \equiv B \pmod{n}$ следует, что $[[A]]_n = [[B]]_n$. Поэтому

$$\begin{aligned} D(E(J)) = D([[J^e]]_n) &= \left[\left[[[J^e]]_n^d\right]\right]_n = \\ &= \left[\left[(J^e)^d\right]\right]_n = \left[\left[J^{ed}\right]\right]_n = \\ &= \left[\left[J^1\right]\right]_n = J, \end{aligned}$$

потому что по теореме 10.31 соотношения $d \cdot e \equiv 1 \pmod{\phi(n)}$ и $r \equiv s \pmod{\phi(n)}$ имеют место тогда и только тогда, когда $x^r \equiv x^s \pmod{n}$. Результат $E(D(J)) = J$ может быть доказан аналогично. ■

ПРИМЕР 22.16. Предположим, что мы решили использовать 8-битовый код ASCII для символов алфавита с размером блока, равным одному символу или букве. Нам потребуется $n \geq 2^8 = 256$. Пусть $p = 41$ и $q = 73$, так что $n = p \cdot q = 2993$, $\phi(n) = 40 \cdot 72 = 2880 = 2^6 \cdot 3^2 \cdot 5$. Пусть $d = 217 = 31 \cdot 7$, так что НОД$(217, 2880) = 1$. Используя алгоритм Евклида, в качестве решения сравнения $217x \equiv 1 \pmod{2880}$ получаем $e = 1513$. В таком случае, слово "$MONEY$", зашифрованное с использованием RSA-метода с ключом $(n, e) = (2993, 1513)$, имеет вид, приведенный в следующей таблице

i	**Блок**	M_i	$E(M_i) = [[M_i^e]]_n = C_i$
1	"M"	77	1683
2	"O"	79	79
3	"N"	78	2560
4	"E"	69	872
5	"Y"	89	2571

Для эффективного вычисления M_i^e по модулю n был использован алгоритм из раздела 3.6. Можно показать также, что $D(C_i) = M_i$, используя в каждом случае тот же самый алгоритм с ключом $(n, d) = (2993, 217)$. Сравнительно небольшая величина числа n в этом примере не позволяет полностью раскрыть возможности шифровальной функции E, поскольку малое n можно легко разложить на множители. Кроме того, размер блока, как правило, выбирается большим, чем один символ, иначе мы бы имели шифр простой замены, который легко поддается криптоанализу. □

RSA-метод используют следующим образом. Оба простых числа p и q, которые дают $n = pq$, а также целое число d, взаимно простое с $\phi(n)$, сохраняются в тайне. Вместе с тем, целые числа n и e становятся общедоступными; известен также и общий метод. Поэтому всякий, знающий ключ шифрования (n, e), может зашифровать сообщение M, получив в результате зашифрованное сообщение C, прочесть которое может только тот, кто обладает ключом (n, d). Каждый может иметь собственный набор ключей криптосистемы RSA. Ключ общего пользования каждого человека, (n, e), может быть внесен в список, например, в регистр, как это сейчас сделано для телефонных номеров. Если некая особа, например, Алиса, хочет послать Бобу тайное послание M_A, то Алисе необходимо найти его ключ общего пользования, (n_B, e_B) в общедоступном регистре. Затем Алисе нужно зашифровать сообщение M_A следующим образом:

$$E_B(M_A) = [[M_A^{e_B}]]_{n_B} = C_A.$$

Зашифрованное сообщение C_A следует передать Бобу любым удобным способом. Всякий, помимо Боба, перехвативший сообщение C_A, не сможет за приемлемое время восстановить сообщение M_A. Получив зашифрованное сообщение C_A, Боб сможет восстановить оригинал, воспользовавшись своим личным ключом (n_B, d_B) следующим образом:

$$D_B(C_A) = \left[\left[C_A^{d_B}\right]\right]_{n_B} = M_A.$$

Особенность RSA, отличающая этот метод от обычных криптосистем, заключается в том, что часть ключа является общеизвестной. Такие методы называются криптосистемами с ключами общего пользования. Они были введены У. Диффи и М. Э. Хеллмэн в работе *New Directions in Cryptography* [28].

Согласно теореме 22.15 $D(E(J)) = J$ и $E(D(J)) = J$, поэтому криптосистема RSA обладает очень полезной характерной особенностью, на которую указали Диффи и Хеллмэн. Ключ Боба общего пользования (n_B, e_B) доступен для каждого, поэтому каждый может послать ему собщение; поэтому, если Алиса посылает Бобу сообщение, то как он может быть уверен, что сообщение пришло от Алисы, а не от кого-то другого? Алиса может набрать подписанное сообщение S_A, содержащее ее имя и другую идентифицирующую информацию. Далее, Алиса использует собственный ***секретный дешифрирующий ключ*** (n_A, d_A)

$$D_A(S_A) = \left[\left[S_A^{d_A}\right]\right]_{n_A} = T_A$$

и включает T_A вместе или как часть основного зашифрованного текста C_A сообщения M_A. Когда Боб получает сообщение Алисы, он выделяет T_A и применяет ***ключ шифрования общего пользования*** Алисы, (n_A, e_A), что дает

$$E_A(T_A) = E_A(D_A(S_A)) = S_A.$$

Учитывая, что, по-видимому, только Алиса могла создать сообщение T_A, которое преобразуется в S_A при помощи ключа общего пользования Алисы (n_A, e_A), соответствующее сообщение M_A является подлинным. Инструмент установления подлинности для такого применения не должен зависеть от сообщения и подписывающего лица. Простой способ реализации такого определения подлинности сообщения Алисы для Боба, состоящего из блоков $M_1, M_2, \ldots, M_k$, мог бы состоять в том, что Алиса создаст подпись S, состоящую из блоков $S_1, S_2, \ldots, S_k$, которая преобразуется в $T_i = D_A(S_i)$ путем использования дешифрующего ключа Алисы. Можно даже положить $S_i = M_i$, после чего $T_i = D_A(M_i)$. Комбинация суперблоков $T_1M_1, T_2M_2, \ldots, T_kM_k$ шифруется ключом Боба общего пользования. Может возникнуть необходимость переписать целые числа $T_1M_1, T_2M_2, \ldots, T_kM_k$ в виде $J_1, J_2, \ldots, J_m$, где $0 \leq J_i < n_B$, чтобы они удовлетворяли модулю Боба, n_B. Таким образом, имеем

$$U_i = E_B(J_i).$$

Получив последовательность $U_1, U_2, \ldots, U_k$, Боб воспользуется своим секретным ключом и получит

$$J_i = D_B(U_i),$$

что может быть разблокировано и снова даст $T_1M_1, T_2M_2, \ldots, T_kM_k$. Затем Боб сверит подпись с

$$E_A(T_i) = M_i$$

в предположении, что было использовано $S_i = M_i$. Поскольку T_i, преобразуемое в $M_i = E_A(T_i)$, могло быть создано только Алисой, сообщение является подлинным.

Безопасность криптосистемы RSA зависит от того, насколько сложно найти разложение на множители $n = p \cdot q$, получая посредством этого p и q. Поскольку

e общеизвестно, то d можно определить, решая, при условии, что известны p и q, сравнение $ex \equiv 1 \pmod{(p-1)(q-1)}$. Безопасность метода зависит также от сложности определения индексов. Если M известно, то $C = [[M^e]]_n$ можно вычислить, используя общедоступную информацию. Криптоаналитику потребуется решить сравнение $M \equiv C^d \pmod{n}$, если $d = \operatorname{ind}_C M$ по модулю n, что также является непростым делом. Кроме того, p и q — случайным образом выбранные большие простые числа, что делает разложение на множители числа n трудновыполнимым. Более того, d выбирается случайным образом так, чтобы оно было взаимно простым с $\phi(n)$ и достаточно большим, чтобы сделать быстрый полный перебор при малых d неэффективным.

Традиционные методы шифрования требуют, чтобы обе стороны, и только они, владели одинаковым ключом. Примером является 56-битовый ключ, используемый в стандарте кодировке данных DES — The United States Data Encryption Standard (см. работу Мейера и Матиаса *Cryptography: A New Dimension in Computer Data Security* [71]). Для передачи сообщений большого объема использование DES может быть предпочтительнее, чем использование метода RSA, т.к. DES требует меньших вычислительных затрат. Вопрос состоит в том, как стороны могут сообщать ключ друг другу? Диффи и Хелмэн [28] предложили следующий метод обмена ключами.

Предположим, что p — простое число и g — первообразный корень простого числа p. Числа p и g могут быть общеизвестны. Предположим, что Алиса и Боб хотят пользоваться одним ключом. Алиса выбирает случайное целое число a такое, что $1 \leq a \leq p-2$, и Боб выбирает случайное целое число b такое, что $1 \leq b \leq p-2$. Числа a и b хранятся в секрете и известны только Алисе и Бобу, соответственно. Алиса вычисляет

$$A = [[g^a]]_p$$

и передает A Бобу любым удобным способом. Со своей стороны, Боб вычисляет

$$B = [[g^b]]_p$$

и передает B любым способом Алисе. Алиса вычисляет

$$K_1 = [[B^a]]_p,$$

и Боб вычисляет

$$K_2 = [[A^b]]_p.$$

Но $K_1 = K_2$, потому что

$$[[A^b]]_p = [[(g^a)^b]]_p = [[(g^b)^a]]_p = [[B^a]]_p.$$

Таким образом, Боб и Алиса имеют теперь один и тот же ключ, $K = K_1 = K_2$.

Поскольку B, A, p и g могут быть известны, криптоаналитик может взломать шифр, решая сравнение $A \equiv g^a \pmod{p}$ для a или решая $B \equiv g^b \pmod{p}$ для b, т.е. нужно иметь возможность вычислить $\operatorname{ind}_g A$ или $\operatorname{ind}_g B$ по модулю p, что в каждом случае связано с большими вычислительными трудностями. Сложность

этого вычисления, согласно обсуждениям, проведенным в разделе 3.14, имеет порядок $\sqrt{p}$, поэтому p должно быть большим простым числом, а $p-1$ должно иметь в разложении хотя бы один большой простой множитель.

Общие сведения о криптографии можно найти в работе Дэвида Кана *The Codebreakers: The Story of Secret Writing* [54]. Вопросы криптографии с ключами общего пользования рассмотрены в работе М. Э. Хеллмэна *The Mathematics of Public-Key Cryptography* [43]. RSA-метод был запатентован Райвестом, Шамиром и Адлеманом в 1983 году; срок патента истек осенью 2000 года.

■ УПРАЖНЕНИЯ

1. В примере 22.16 проверьте, что $D(J) = \left[\left[J^{217}\right]\right]_{2993}$ дешифрует сообщение

$$1683, 79, 2560, 872, 2571.$$

2. Создайте RSA-криптосистему, используя $n = 83 \cdot 107$, путем нахождения ключа общего пользования (n, e) и личного ключа (n, d).

а) Используйте эти ключи, чтобы зашифровать сообщения "$HELP$".

б) Пусть RSA-ключи из примера 22.16 принадлежат Алисе, а ключи для $n = 83 \cdot 107$ принадлежат Бобу. Используйте метод подписей, описанный в этом разделе, для пересылки Бобу подписанного сообщения "$MONEY$" от Алисы.

в) Используя личный ключ Боба и ключ Алисы общего пользования, дешифруйте сообщение из части (б) и проверьте, что оно пришло от Алисы.

3. Покажите, что в RSA-криптосистеме с ключом общего пользования (n, e), если $\phi(n)$ может быть найдено, то криптосистема может быть взломана посредством вычисления d.

4. Рассмотрите RSA-криптосистему, когда $n = p \cdot q$ и $\phi(n)$ — известны, но простые числа p и q — неизвестны.

а) Покажите, что $p + q$ можно выразить через n и $\phi(n)$.

б) Покажите, что если $q > p$, то $q - p = \sqrt{(p+q)^2 - 4n}$.

в) Определите p и q.

г) Выполните пример 22.16 в предположении, что $n = 2993$ и $\phi(n) = 2880$. Используйте метод, приведенный выше, для нахождения p и q.

Литература

[1] Jiří Adámek, *Foundations of Coding*, Wiley, New York, 1991.

[2] L. M. Adleman, "A Subexponential Algorithm for the Discrete Logarithm Problem with Applications to Cryptography," *Proc. 20th IEEE Found. Comput. Sci. Symp.* (1979), 55-60.

[3] W. R. Alford, A. Granville, and C. Pomerance, "There Are Infinitely Many Carmichael Numbers," *Ann. Math.*, Volume 140 (1994), 703-722.

[4] Ian Anderson, *A First Course in Combinatorial Mathematics*, Oxford University Press, New York, 1995.

[5] James A. Anderson and James M. Bell, *Number Theory With Applications*, Prentice-Hall, Upper Saddle River, NJ, 1997.

[6] E. T. Bell, *The Development of Mathematics*, 2nd ed., McGraw-Hill, New York, 1945.

[7] E. T. Bell, *Men of Mathematics*, Simon & Schuster, New York, 1965.

[8] C. Berge, *Principles of Combinatorics*, Academic Press, New York, 1971.

[9] Jean Berstel and Dominique Perrin, *Theory of Codes*, Academic Press, New York, 1985.

[10] Norman Biggs, *Algebraic Graph Theory*, Cambridge University Press, Cambridge, 1993.

[11] G. D. Birkhoff, A Determinant Formula for the Number of Ways of Coloring a Map. *Ann. of Math.*, 14 (1912) 42-46).

[12] R. S. Boyer and J. S. Moore, "A Fast String Searching Algorithm," *Commun. ACM*, Volume 20, No. 10 (1977), 762-772.

[13] R. P. Brent, "Irregularities in the Distribution of Primes and Twin Primes," *Math. Comput.*, Volume 29 (1975), 43-56.

[14] R. P. Brent and J. M. Pollard, "Factorization of the Eighth Fermat Number," *Math. of Comput.*, Volume 36, No. 154 (1981), 627-630.

[15] Richard A. Brualdi, *Introductory Combinatorics*, Prentice-Hall, Upper Saddle River NJ, 1999.

[16] F. Cajori, *A History of Mathematics*, 2nd ed., Macmillan, New York, 1961.

[17] R. D. Carmichael, "On Composite Numbers P Which Satisfy the Fermat Congruence $a^{P-1} \equiv 1 \bmod P$," *Amer. Math. Monthly*, Volume 19, No. 2 (1912), 22-27.

[18] C. C. Chang and J. C. Shieh, "A Fast Algorithm for Constructing Reciprocal Hashing Functions," *Proc. Internat. Symp. New Directions Comput.* (1985), 232-236.

[19] C. C. Chang and J. C. Shieh, "Pairwise Relatively Prime Generating Polynomials and Their Applications," *Proc. Internat. Workshop Discrete Algorithms Complexity*, November (1989), 137-139.

[20] Nan-xian Chen, "Modified Möbius Inversion Formula and Its Application to Physics," *Phys. Rev. Lett.*, Volume 64, No. 11 (1990), 1193-1195.

[21] L. W. Cohen and G. Ehrlich, *The Structure of the Real Number System*, Van Nostrand Reinhold, New York, 1963.

[22] H. Cohn, *Advanced Number Theory*, Dover, New York, 1980.

[23] D. Coppersmith, "Cryptography," *IBM J. Res. Develop.*, Volume 31, No. 2 (1987), 244-248.

[24] R. Crandall, J. Doenias, C. Norrie, and J. Young, "The Twenty-Second Fermat Number Is Composite," *Math. Comput.*, Volume 64 (1995), 863-868.

[25] T. Dantzig, *Number: The Language of Science*, 4th ed., Macmillan, New York, 1954.

[26] H. Davenport, *The Higher ABithmetic*, 6th ed., Cambridge University Press, New York, 1992.

[27] L. E. Dickson, *Studies in the Theory of Numbers*, Chelsea, New York, 1957.

[28] W. Diffie and M. E. Hellman, "New Directions in Cryptography," *IEEE Trans. Inform. Theory*, Volume IT-22, No. 6 (1976), 644-654.

[29] H. W. Eves, *In Mathematical Circles, Quadrants I and II*, Prindle, Weber & Schmidt, Boston, 1969.

[30] William Feller, *An Introduction to Probability Theory and its Applications*, Wiley, New York, 1950.

[31] G. S. Fishman and L. R. Moore, "A Statistical Evaluation of Multiplicative Congruential Random Number Generators with Modulus $2^{31} - 1$," *J. Amer. Statist. Assoc.*, Volume 77 (1982), 129-136.

[32] G. S. Fishman and L. R. Moore III, "An Exhaustive Analysis of Multiplicative Congruential Random Generators with Modulus $2^{31} - 1$," *SIAM J. Sci. Statist. Comput.*, Volume 7, No. 1 (1986), 24-45.

[33] M. Gardner, "Through the Looking Glass," by Lewis Carroll, from *The Annotated Alice*, Bramhall House, New York, 1960.

[34] C. W. Gear, *Computer Organization and Programming*, 2nd ed., McGraw-Hill, New York, 1974.

[35] Samuel Goldberg, *Introduction to Difference Equations*, Dover, New York, 1986.

[36] Edgar G. Goodaire and Michael M. Paramenter, *Discrete Mathematics with Graph Theory*, Prentice-Hall, Upper Saddle River NJ, 1998.

[37] Ronald Gould, *Graph Theory*, Benjamin/Cummings, Menlo Park, 1988.

[38] Ronald R. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics*, Addison Wesley, New York, 1994.

[39] E. Grosswald, *Representations of Integers as Sums of Squares*, Springer-Verlag, New York, 1985.

[40] Richard Hamming, *Coding and Information Theory*, Prentice-Hall, Upper Saddle River NJ, 1986.

[41] Goffredo Haus and Sametti ScoreSnyth: A System for the Synthesis of Music Scores Based on Petri Nets and a Music Algebra *Readings in Computer-Generated Music* Denis Baggi ed. pp 53-77.IEEE Computer Press Society Los Alamitos, California, 1992

[42] P. J. Heawood, Map-Color Theorem, *Quart. J. Math.*, 23(1890), 332-339)

[43] M. E. Hellman, "The Mathematics of Public-Key Cryptography," *Sci. Amer.*, Volume 241, August (1979), 146-157.

[44] Raymond Hill, *A First Course in Coding Theory,* Oxford University Press, 1993).

[45] F. S. Hillier and G. Lieberman, *Operations Research*, 2nd ed., Holden-Day, San Francisco, 1974.

[46] Paul G. Hoel, *Introduction to Mathematical Statistics*, Wiley, New York.

[47] Robert V. Hogg and Allen T. Craig, *Introduction to Mathematical Statistics*, Macmillan, New York, 1965.

[48] Ellis Horowitz and Sartaj Sahni, *Fundamentals of Data Structures in Pascal*, Computer Science Press, Rockville MD, 1984.

[49] John M. Howie, *Automata and Languages*, Oxford University Press, New York, 1991.

[50] A. E. Ingham, *The Distribution of Prime Numbers*, Cambridge University Press, New York, 1992.

[51] G. Jaeschke, "Reciprocal Hashing: A Method for Generating Minimal Perfect Hashing Functions," *Commun. ACM*, Volume 24, No. 12 (1981), 829-833.

[52] Richard Johnsonbaugh, Discrete Mathematics, Prentice-Hall, Upper Saddle River NJ, 1997.

[53] Charles Jordan, *Calculus of Finite Differences*, Chelsea, New York, 1979.

[54] D. Kahn, *The Codebreakers: The Story of Secret Writing*, Macmillan, New York, 1968.

[55] David Kahn, *Kahn on Codes*, MacMillan, New York, 1983).

[56] R. M. Karp and M. O. Rabin, "Efficient Randomized Pattern-Matching Algorithms," *IBM J. Res. Develop.*, Volume 31, No. 2, March (1987), 249-260.

[57] Steven G. Kellison, *Fundamentals of Numerical Analysis*, Richard D. Irwin, Homewood IL, 1975.

[58] John G. Kemeny and J. Laurie Snell, *Mathematical Models in the Social Sciences*, Blaisdell, Waltham, 1962.

[59] John G. Kemeny, J. Laurie Snell, and Gerald L. Thompson, Introduction to Finite Mathematics, Prentice-Hall, Upper Saddle River NJ, 1966.

[60] Alfred Bay Kempe, On the Geographical Problem of the Four Colors *Amer. J. Math.*, 2(1879),193-200.

[61] A. Ya. Khinchin, *Continued Fractions*, The University of Chicago Press, Chicago, 1964.

[62] H. A. Klein, *The Science of Measurement: A Historical Survey*, Dover, New York, 1988.

[63] M. Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, New York, 1972.

[64] D. E. Knuth, *The Art of Computer Programming*, 2nd ed., Volumes 2 and 3, Addison-Wesley, Reading, Mass., 1981 and 1973.

[65] D. E. Knuth, J. H. Morris, and V. R. Pratt, "Fast Pattern Matching in Strings," *SIAM J. Comput.*, Volume 6, No. 2, June (1977), 323-350.

[66] P. L'Ecuyer, "Efficient and Portable Combined Random Number Generators," *Commun. ACM*, Volume 31 (1988), 742-774.

[67] A. K. Lenstra and H. W. Lenstra, Jr., editors, *The Development of the Number Field Sieve*, Springer-Verlag, New York 1993.

[68] H.Levy, and F. Lessman, *Finite Difference Equations*, Dover, New York, 1992.

[69] Harry Lewis and Christos H. Papadimitriou, *Elements of the Theory of Computation*, Prentice Hall, Upper Saddle River NJ, 1981

[70] G. Marsaglia, "Random Numbers Fall Mainly in the Planes," *Proc. Nat. Acad. Sci. U.S.A.*, Volume 61 (1968), 25-28.

[71] C. H. Meyer and S. M. Matyas, *Cryptography: A New Dimension in Computer Data Security*, Wiley, New York, 1982.

[72] R. P. Millane, "A Product Form of the Möbius Transform," *Phys. Lett. A*, Volume 162 (1992), 213-214.

[73] R. P. Millane, "Möbius Transform Pairs," *J. Math. Phys.*, Volume 34 (1993), 875-877.

[74] L. M. Milne-Thompson, *The Calculus of Finite Differences*, Chelsea, New York, 1981.

[75] O. Neugebauer, *The Exact Sciences in Antiquity*, 2nd ed., Dover, New York, 1969.

[76] O. Neugebauer and A. Sachs, *Mathematical Cuneiform Texts*, American Oriental Society, New Haven, Conn. 1945.

[77] J. R. Newman, *The World of Mathematics*, Simon & Schuster, New York, 1956.

[78] B. W. Ninham et al., "Möbius, Mellin, and Mathematical Physics," *Physica A*, Volume 186 (1992), 441-481.

[79] J. M. H. Olmsted, *The Real Number System*, Appleton-Century-Crofts, New York, 1962.

[80] G. Ottewell, *The Astronomical Companion,* Astronomical Workshop, Furman University, Greenville, S.C., 1979.

[81] R. E. A. C. Paley, "On Orthogonal Matrices," *J. of Math. Phys.*, Volume 12 (1933), 311-320.

[82] B. K. Parady, J. R. Smith, and S. Zarantonello, "Largest Known Twin Primes," *Math. Comput.*, Volume 55 (1990), 381-382.

[83] S. K. Park and K. W. Miller, "Random Number Generators: Good Ones Are Hard to Find," *Commun. ACM*, Volume 31 (1988), 1192-1201.

[84] Emanuel Parzen, *Modern Probability Theory and its Application*, Wiley, New York, 1960.

[85] James L. Peterson, *Petri Net Theory and the Modeling of Systems,* Prentice Hall, Upper Saddlle River NJ, 1981.

[86] R. L. Plackett and J. P. Burman, "The Design of Optimum Multifactorial Experiments," *Biometrika*, Volume 33 (1946), 305-325.

[87] Vera Pless, *Introduction to the Theory of Error-Correcting Codes*, Wiley, New York, 1998.

[88] M. Plotkin, Binary codes with specified minimal distance, *IEEE Trans. Info. Theory,* 6, 445-450).

[89] S. Pohlig and M. Hellman, "An Improved Algorithm for Computing Logarithms over $GF(p)$ and Its Cryptographic Significance," *IEEE Trans. Inform. Theory*, Vol. IT-24, No. 1 (1978), 106-110.

[90] J. M. Pollard, "Theorems on Factorization and Primality Testing," *Proc. Cambridge Philos. Soc.*, Volume 76 (1974), 521-528.

[91] J. M. Pollard. "A Monte Carlo Method for Factorization," *Bit*, (1975), 331-334.

[92] C. Pomerance, *Lecture Notes on Primality Testing and Factoring*, The Mathematical Association of America, Washington, D.C., 1984.

[93] J. Purdum, "Pattern Matching Alternatives: Theory vs. Practice," *Comput. Language*, November (1987), 33-44.

[94] M. Rabin, "Probabilistic Algorithm for Testing Primality," *J. Number Theory*, Volume 12 (1980), 128-138.

[95] F. Reid and K. Honeycutt, "A Digital Clock for Sidereal Time," *Sky & Telescope*, July (1976), 59-63.

[96] S. Y. Ren and J. D. Dow, "Generalized Möbius Transforms for Inverse Problems," *Phys. Lett. A*, Volume 154 (1991), 215-216.

[97] P. Ribenboim, *The Little Book of Big Primes*, Springer-Verlag, New York, 1991.

[98] R. L. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public Key Cryptosystems," *Commun. ACM*, Volume 21, No. 2 (1978), 120-126.

[99] Fred S. Roberts, *Applied Combinatorics*, Prentice-Hall, Upper Saddle River NJ, 1984.

[100] Steven Roman, *Introduction to Coding and Information Theory*, Springer-Verlag, New York, 1991.

[101] Gian-Carlos Rota, *Studies in Combinatorics*, Mathematical Association of America, 1978.

[102] Joseph J. Rotman, *An Introduction to the Theory of Groups*, Springer Verlag, New York, 1995.

[103] J. B. Rosser and L. Schoenfeld, "Approximate Formulas for Some Functions of Prime Numbers," *Illinois J. Math.*, Volume 6 (1962), 64-94.

[104] A. Schimmel, *The Mystery of Numbers*, Oxford University Press, New York, 1993.

[105] M. R. Schroeder, *Number Theory in Science and Communication*, 2nd ed., Springer-Verlag, Berlin, 1986.

[106] Robert Sedgewick, *Algorithms*, Addison Wesley, New york, 1988.

[107] R. W. Sinnott, *Sky & Telescope*, January (1989), 80-82.

[108] D. E. Smith, *History of Mathematics*, Dover, New York, 1958.

[109] D. E. Smith, *A Source Book in Mathematics*, Dover, New York, 1959.

[110] R. Solovay and V. Strassen, "A Fast Monte-Carlo Test for Primality," *SIAM J. Comput.*, Volume 6, No. 1 (1977), 84-85. "Erratum: A Fast Monte-Carlo Test for Primality," *SIAM J. Comput.*, Volume 7, No. 1 (1978), 118.

[111] D. M. Sunday, "A Very Fast Substring Search Algorithm," *Commun. ACM*, Volume 33, No. 8 (1990), 132-142.

[112] Thomas M. Thompson, *From Error-Correcting Codes through Sphere Packings to Simple Groups*, Mathematical Mathematical Association, 1983.

[113] J. A. Todd, "A Combinatorial Problem," *J. Math. Phys.*, Volume 12 (1933), 321-333.

[114] Alan Tucker, *Applied Combinatorics*, Wiley, New York, 1984.

[115] Lekh R. Vermani, *Elements of Algebraic Coding Theory*, Chapman and Hill, New York, 1996.

[116] B. L. van der Waerden, *Science Awakening*, Vol. I, 3rd ed., Oxford University Press, New York, 1971.

[117] Douglas B. West, *Introduction to Graph Theory*, Prentice-Hall, Upper Saddle River NJ, 1996.

[118] A. E. Western and J. C. Miller, *Tables of Indices and Primitive Roots*, Royal Society Mathematical Tables, Vol. 9, Cambridge University Press, Cambridge, 1968.

Ответы к упражнениям

Раздел 1.1.

1. а) нет; **б)** да, истинное; **в)** да, ложное; **г)** нет; **д)** да, ложное.

3. а) $\sim r \wedge \sim q$;
б) $(\sim r \wedge p) \vee q$;
в) $\sim (r \wedge q)$;
г) $(\sim p \wedge q) \vee (p \wedge \sim q)$.

5. а)

Случай	q	r	$\sim$	r	$\wedge$	$\sim$	q
1	*T*	*T*	*F*	*T*	**F**	*F*	*T*
2	*T*	*F*	*T*	*F*	**F**	*F*	*T*
3	*F*	*T*	*F*	*T*	**F**	*T*	*F*
4	*F*	*F*	*T*	*F*	**T**	*T*	*F*
					*		

б)

Случай	p	q	r	$(\sim$	r	$\wedge$	$p)$	$\vee$	q
1	*T*	*T*	*T*	*F*	*T*	*F*	*T*	**T**	*T*
2	*T*	*T*	*F*	*T*	*F*	*T*	*T*	**T**	*T*
3	*T*	*F*	*T*	*F*	*T*	*F*	*T*	**F**	*F*
4	*T*	*F*	*F*	*T*	*F*	*T*	*T*	**T**	*F*
5	*F*	*T*	*T*	*F*	*T*	*F*	*F*	**T**	*T*
6	*F*	*T*	*F*	*T*	*F*	*F*	*F*	**T**	*T*
7	*F*	*F*	*T*	*F*	*T*	*F*	*F*	**F**	*F*
8	*F*	*F*	*F*	*T*	*F*	*F*	*F*	**F**	*F*
								*	

в)

Случай	q	r	$\sim$	$(r$	$\wedge$	$q)$
1	*T*	*T*	**F**	*T*	*T*	*T*
2	*T*	*F*	**T**	*F*	*F*	*T*
3	*F*	*T*	**T**	*T*	*F*	*F*
4	*F*	*F*	**T**	*F*	*F*	*F*
			*			

г)

Случай	p	q	$(\sim$	p	$\wedge$	$q)$	$\vee$	$(p$	$\wedge$	$\sim$	$q)$
1	*T*	*T*	*F*	*T*	*F*	*T*	**F**	*T*	*F*	*F*	*T*
2	*T*	*F*	*F*	*T*	*F*	*F*	**T**	*T*	*T*	*T*	*F*
3	*F*	*T*	*T*	*F*	*T*	*T*	**T**	*F*	*F*	*F*	*T*
4	*F*	*F*	*T*	*F*	*F*	*F*	**F**	*F*	*F*	*T*	*F*
							*				

7. а) я играю в шахматы, и игра в шахматы требует времени;
б) эта игра не очень трудна, или я не буду играть в шахматы;
в) эта игра очень трудна, или игра в шахматы требует времени, и я играю в шахматы;
г) эта игра очень трудна, и я играю в шахматы, и игра в шахматы требует времени.

9. а)

Случай	q	r	$(q \wedge r)$
1	*T*	*T*	**T** *T*
2	*T*	*F*	**F** *F*
3	*F*	*T*	**F** *T*
4	*F*	*F*	**F** *F*
			*

б)

Случай	p	q	$\sim$	p	$\vee$	$\sim$	q
1	*T*	*T*	*F*	*T*	**F**	*F*	*T*
2	*T*	*F*	*F*	*T*	**T**	*T*	*F*
3	*F*	*T*	*T*	*F*	**T**	*F*	*T*
4	*F*	*F*	*T*	*F*	**T**	*T*	*F*
					*		

в)

Случай	p	q	r	$(p$	$\vee$	$r)$	$\wedge$	q
1	*T*	*T*	*T*	*T*	*T*	*T*	**T**	*T*
2	*T*	*T*	*F*	*T*	*T*	*F*	**T**	*T*
3	*T*	*F*	*T*	*T*	*T*	*T*	**F**	*F*
4	*T*	*F*	*F*	*T*	*T*	*F*	**F**	*F*
5	*F*	*T*	*T*	*F*	*T*	*T*	**T**	*T*
6	*F*	*T*	*F*	*F*	*F*	*F*	**F**	*T*
7	*F*	*F*	*T*	*F*	*T*	*T*	**F**	*F*
8	*F*	*F*	*F*	*F*	*F*	*F*	**F**	*F*
							*	

г)

Случай	p	q	r	$(p$	$\wedge$	$q)$	$\wedge$	r
1	*T*	*T*	*T*	*T*	*T*	*T*	**T**	*T*
2	*T*	*T*	*F*	*T*	*T*	*T*	**F**	*F*
3	*T*	*F*	*T*	*T*	*F*	*F*	**F**	*T*
4	*T*	*F*	*F*	*T*	*F*	*F*	**F**	*F*
5	*F*	*T*	*T*	*F*	*F*	*T*	**F**	*T*
6	*F*	*T*	*F*	*F*	*F*	*T*	**F**	*F*
7	*F*	*F*	*T*	*F*	*F*	*F*	**F**	*T*
8	*F*	*F*	*F*	*F*	*F*	*F*	**F**	*F*
							*	

11. а)

Случай	p	q	r	$p \wedge (q \vee \sim r)$
1	T	T	T	$T\ \mathbf{T}\ T\ T\ F\ T$
2	T	T	F	$T\ \mathbf{T}\ T\ T\ T\ F$
3	T	F	T	$T\ \mathbf{F}\ F\ F\ F\ T$
4	T	F	F	$T\ \mathbf{T}\ F\ T\ T\ F$
5	F	T	T	$F\ \mathbf{F}\ T\ T\ F\ T$
6	F	T	F	$F\ \mathbf{F}\ T\ T\ T\ F$
7	F	F	T	$F\ \mathbf{F}\ F\ F\ F\ T$
8	F	F	F	$F\ \mathbf{F}\ F\ T\ T\ F$
				$\quad\ *$

б)

Случай	p	q	r	$(q \wedge \sim r) \vee (\sim p \wedge r)$
1	T	T	T	$T\ F\ F\ T\ \mathbf{F}\ F\ T\ F\ T$
2	T	T	F	$T\ T\ T\ F\ \mathbf{T}\ F\ T\ F\ F$
3	T	F	T	$F\ F\ F\ T\ \mathbf{F}\ F\ T\ F\ T$
4	T	F	F	$F\ F\ T\ F\ \mathbf{F}\ F\ T\ F\ F$
5	F	T	T	$T\ F\ F\ T\ \mathbf{T}\ T\ F\ T\ T$
6	F	T	F	$T\ T\ T\ F\ \mathbf{T}\ T\ F\ F\ F$
7	F	F		$F\ F\ F\ T\ \mathbf{T}\ T\ F\ T\ T$
8	F	F	F	$F\ F\ T\ F\ \mathbf{F}\ T\ F\ F\ F$
				$\quad\ *$

в)

Случай	p	q	r	$\sim (p \wedge r) \vee (\sim q \wedge r)$
1	T	T	T	$F\ T\ T\ T\ \mathbf{F}\ F\ T\ F\ T$
2	T	T	F	$T\ T\ F\ F\ \mathbf{T}\ F\ T\ F\ F$
3	T	F	T	$F\ T\ T\ T\ \mathbf{T}\ T\ F\ T\ T$
4	T	F	F	$T\ T\ F\ F\ \mathbf{T}\ T\ F\ F\ F$
5	F	T	T	$T\ F\ F\ T\ \mathbf{T}\ F\ T\ F\ T$
6	F	T	F	$T\ F\ F\ F\ \mathbf{T}\ F\ T\ F\ F$
7	F	F	T	$T\ F\ F\ T\ \mathbf{T}\ T\ F\ T\ T$
8	F	F	F	$T\ F\ F\ F\ \mathbf{T}\ T\ F\ F\ F$
				$\quad\ *$

г)

Случай	p	q	r	$\sim (\sim p \vee (q \wedge \sim r)$
1	T	T	T	$\mathbf{T}\ F\ T\ F\ T\ F\ F\ T$
2	T	T	F	$\mathbf{F}\ F\ T\ T\ T\ T\ T\ F$
3	T	F	T	$\mathbf{T}\ F\ T\ F\ F\ F\ F\ T$
4	T	F	F	$\mathbf{T}\ F\ T\ F\ F\ F\ T\ F$
5	F	T	T	$\mathbf{F}\ T\ F\ T\ T\ F\ F\ T$
6	F	T	F	$\mathbf{F}\ T\ F\ T\ T\ T\ T\ F$
7	F	F	T	$\mathbf{F}\ T\ F\ T\ F\ F\ F\ T$
8	F	F	F	$\mathbf{F}\ T\ F\ T\ F\ F\ T\ F$
				$\quad\ *$

д)

Случай	p	q	r	$(p \wedge r) \vee (p \wedge \sim q)$
1	T	T	T	$T\ T\ T\ \mathbf{T}\ T\ F\ F\ T$
2	T	T	F	$T\ F\ F\ \mathbf{F}\ T\ F\ F\ T$
3	T	F	T	$T\ T\ T\ \mathbf{T}\ T\ T\ T\ F$
4	T	F	F	$T\ F\ F\ \mathbf{T}\ T\ T\ T\ F$
5	F	T	T	$F\ F\ T\ \mathbf{F}\ F\ F\ F\ T$
6	F	T	F	$F\ F\ F\ \mathbf{F}\ F\ F\ F\ T$
7	F	F	T	$F\ F\ T\ \mathbf{F}\ F\ F\ T\ F$
8	F	F	F	$F\ F\ F\ \mathbf{F}\ F\ F\ T\ F$
				$\quad\ *$

Раздел 1.2.

1. а) $r \to (p \wedge q)$;
б) $\sim p \to \sim q$;
в) $(r \to q) \wedge (\sim r \to \sim p)$;
г) $\sim (r \vee p) \to \sim q$.

3. а) если ему нравятся фиолетовые галстуки и он популярен, то у него странные друзья;
б) если он популярен, то у него нет странных друзей;
в) если ему нравятся фиолетовые галстуки, то он популярен, или у него странные друзья;
г) если ему нравятся фиолетовые галстуки, то он не популярен, и если он популярен, то у него странные друзья.

5. а)

Случай	p	q	r	$(p \to q) \to r$
1	T	T	T	$T\ T\ T\ \mathbf{T}\ T$
2	T	T	F	$T\ T\ T\ \mathbf{F}\ F$
3	T	F	T	$T\ F\ F\ \mathbf{T}\ T$
4	T	F	F	$T\ F\ F\ \mathbf{T}\ F$
5	F	T	T	$F\ T\ T\ \mathbf{T}\ T$
6	F	T	F	$F\ T\ T\ \mathbf{F}\ F$
7	F	F	T	$F\ T\ F\ \mathbf{T}\ T$
8	F	F	F	$F\ T\ F\ \mathbf{F}\ F$
				$\quad\ *$

б)

Случай	p	q	r	$p \to (q \to r)$
1	T	T	T	$T\ \mathbf{T}\ T\ T\ T$
2	T	T	F	$T\ \mathbf{F}\ T\ F\ F$
3	T	F	T	$T\ \mathbf{T}\ F\ T\ T$
4	T	F	F	$T\ \mathbf{T}\ F\ T\ F$
5	F	T	T	$F\ \mathbf{T}\ T\ T\ T$
6	F	T	F	$F\ \mathbf{T}\ T\ F\ F$
7	F	F	T	$F\ \mathbf{T}\ F\ T\ T$
8	F	F	F	$F\ \mathbf{T}\ F\ T\ F$
				$\quad\ *$

в)

Случай	p	q	r	$(q \to (p \wedge r))$	$\leftrightarrow$	$((q \to p) \wedge (q \to r))$
1	T	T	T	$T\ T\ T\ T$	**T**	$T\ T\ T\ T\ T$
2	T	T	F	$T\ F\ T\ F$	**T**	$T\ T\ F\ T\ F$
3	T	F	T	$F\ T\ T\ T$	**T**	$F\ T\ T\ T\ T$
4	T	F	F	$F\ T\ F\ F$	**T**	$F\ T\ T\ T\ F$
5	F	T	T	$T\ F\ F\ F$	**T**	$T\ F\ F\ F\ T$
6	F	T	F	$T\ F\ F\ F$	**T**	$T\ F\ F\ F\ F$
7	F	F	T	$F\ T\ F\ F$	**T**	$F\ T\ F\ T\ F$
8	F	F	F	$F\ T\ F\ F$	**T**	$F\ T\ F\ T\ F$
					*	

г)

Случай	p	q	r	$((p \to q) \vee r)$	$\to$	$(\sim p \vee \sim q)$
1	T	T	T	$T\ T\ T\ T$	**F**	$F\ T\ F\ F\ T$
2	T	T	F	$T\ T\ T\ F$	**F**	$F\ T\ F\ F\ T$
3	T	F	T	$T\ F\ T\ T$	**T**	$F\ T\ T\ T\ F$
4	T	F	F	$T\ F\ F\ F$	**T**	$F\ T\ T\ T\ F$
5	F	T	T	$F\ T\ T\ T$	**T**	$T\ F\ T\ F\ T$
6	F	T	F	$F\ T\ T\ F$	**T**	$T\ F\ T\ F\ T$
7	F	F	T	$F\ T\ F\ T$	**T**	$T\ F\ T\ T\ F$
8	F	F	F	$F\ T\ F\ F$	**T**	$T\ F\ T\ T\ F$
					*	

7. а)

Случай	p	q	$(p \to q)$	$\leftrightarrow$	$(\sim q \to \sim p)$
1	T	T	$T\ T\ T$	**T**	$F\ T\ T\ F\ T$
2	T	F	$T\ F\ F$	**T**	$T\ F\ F\ F\ T$
3	F	T	$F\ T\ T$	**T**	$F\ T\ T\ T\ F$
4	F	F	$F\ T\ F$	**T**	$T\ F\ T\ T\ F$
				*	

б)

Случай	p	q	r	$(p \wedge \sim (q \vee \sim r))$	$\leftrightarrow$	$(p \to q)$
1	T	T	T	$T\ F\ F\ T\ T\ F\ T$	**F**	$T\ T\ T$
2	T	T	F	$T\ F\ F\ T\ T\ T\ F$	**F**	$T\ T\ T$
3	T	F	T	$T\ T\ T\ F\ F\ F\ T$	**F**	$T\ F\ F$
4	T	F	F	$T\ F\ F\ F\ T\ T\ F$	**T**	$T\ F\ F$
5	F	T	T	$F\ F\ F\ T\ T\ F\ T$	**F**	$F\ T\ T$
6	F	T	F	$F\ F\ F\ T\ T\ T\ F$	**F**	$F\ T\ T$
7	F	F	T	$F\ F\ T\ F\ F\ F\ T$	**F**	$F\ T\ F$
8	F	F	F	$F\ F\ F\ F\ T\ T\ F$	**F**	$F\ T\ F$
					*	

в)

Случай	p	q	r	$(p \vee r) \to q$
1	T	T	T	$T\ T\ T\ T\ T$
2	T	T	F	$T\ T\ F\ T\ T$
3	T	F	T	$T\ T\ T\ F\ F$
4	T	F	F	$T\ T\ F\ F\ F$
5	F	T	T	$F\ T\ T\ T\ T$
6	F	T	F	$F\ F\ F\ T\ T$
7	F	F	T	$F\ T\ T\ F\ F$
8	F	F	F	$F\ F\ F\ T\ F$
				*

г)

Случай	p	q	r	$((p \to q) \wedge (q \to \sim r)) \to (r \to p)$
1	T	T	T	$T\ T\ T\ T\ T\ F\ F\ T\ \mathbf{T}\ T\ T$
2	T	T	F	$T\ T\ T\ T\ T\ T\ T\ F\ \mathbf{T}\ F\ T$
3	T	F	T	$T\ F\ F\ F\ F\ T\ F\ T\ \mathbf{T}\ T\ T$
4	T	F	F	$T\ F\ F\ F\ F\ T\ T\ F\ \mathbf{T}\ F\ T$
5	F	T	T	$F\ T\ T\ F\ T\ F\ F\ T\ \mathbf{T}\ T\ F$
6	F	T	F	$F\ T\ T\ T\ T\ T\ T\ F\ \mathbf{T}\ F\ F$
7	F	F	T	$F\ T\ F\ T\ F\ T\ F\ T\ \mathbf{F}\ T\ F$
8	F	F	F	$F\ T\ F\ T\ F\ T\ T\ F\ \mathbf{T}\ F\ T$
				$\qquad\qquad\qquad

г)

Случай	p	q	$p \to q$	$\sim p$	$\vee$	q
1	T	T	**T**	F	**T**	T
2	T	F	**F**	F	**F**	F
3	F	T	**T**	T	**T**	T
4	F	F	**T**	T	**T**	F
			*		*	

3.

Случай	p	q	$p \to q$	q	$\to$	p
1	T	T	**T**	T	**T**	T
2	T	F	**F**	F	**T**	T
3	F	T	**T**	T	**F**	F
4	F	F	**T**	F	**T**	F
			*		*	

Таблицы истинности не совпадают, поэтому высказывания не эквивалентны.

$p \to q \equiv \sim p \vee q \equiv$ теорема 1.3(з)

$\equiv q \vee \sim p \equiv$ закон коммутативности

$\equiv \sim\sim q \vee \sim p \equiv$ двойное отрицание

$\equiv \sim q \to \sim p$ теорема 1.3(з)

Следовательно, $\sim p \to \sim q \equiv q \to p$, и инверсия высказывания эквивалентна конверсии.

5. а) если он кентавр, то у него шесть ног;
б) если он преуспевающий политик, то он избран;
в) если он имеет деньги, то он популярен.

7. а) если я хороший гражданин, то я голосую;
б) если я не голосую, то я не являюсь хорошим гражданином;
в) если я не являюсь хорошим гражданином, то я не голосую.

Раздел 1.4.

1. а)

Случай	p	q	r	$p \to q$	$q \to r$	$p \to r$	
1	T	T	T	**T**	**T**	**T**	*
2	T	T	F	T	F	F	
3	T	F	T	F	T	T	
4	T	F	F	F	T	F	
5	F	T	T	**T**	**T**	**T**	*
6	F	T	F	T	F	T	
7	F	F	T	**T**	**T**	**T**	*
8	F	F	F	**T**	**T**	**T**	*

б)

Случай	p	q	r	s	p	$p \to (r \vee s)$	$r \to q$	$s \to q$	q	
1	T	T	T	T	T	**T**	**T**	**T**	**T**	*
2	T	T	T	F	T	**T**	**T**	**T**	**T**	*
3	T	T	F	T	T	**T**	**T**	**T**	**T**	*
4	T	T	F	F	T	F	T	T	T	
5	T	F	T	T	T	T	F	F	F	
6	T	F	T	F	T	T	F	T	F	
7	T	F	F	T	T	T	T	F	F	
8	T	F	F	F	T	F	T	T	F	
9	F	T	T	T	F	**T**	**T**	**T**	**T**	*
10	F	T	T	F	F	**T**	**T**	**T**	**T**	*
11	F	T	F	T	F	**T**	**T**	**T**	**T**	*
12	F	T	F	F	F	**T**	**T**	**T**	**T**	*
13	F	F	T	T	F	T	F	F	F	
14	F	F	T	F	F	T	F	T	F	
15	F	F	F	T	F	T	T	F	F	
16	F	F	F	F	F	T	T	T	F	

в)

Случай	r	w	$\sim w \rightarrow (r \wedge \sim r)$	w	
1	T	T	T	T	*
2	T	F	F	F	
3	F	T	T	T	*
4	F	F	F	F	

3. а) неправильное; **б)** правильное; **в)** правильное; **г)** правильное.

5. а) правильное; **б)** неправильное; **в)** неправильное; **г)** правильное.

7. а)

1.	$\sim (s \wedge t)$	дано
2.	$\sim w \rightarrow t$	дано
3.	$\sim s \vee \sim t$	1 закон де Моргана
4.	$s \rightarrow \sim t$	3 и теорема 1.3(и)
5.	$\sim t \rightarrow w$	2 и контрапозиция
6.	$s \rightarrow w$	силлогизм;

б)

1.	$p \rightarrow q$	дано
2.	$\sim r \rightarrow \sim q$	дано
3.	$\sim r$	дано
4.	$\sim q$	2, 3 и modus ponens
5.	$\sim q \rightarrow \sim p$	1 и контрапозиция
6.	$\sim p$	4, 5 и modus ponens;

в)

1.	$\sim (\sim p \vee q)$	дано
2.	$\sim z \rightarrow \sim s$	дано
3.	$(p \wedge \sim q) \rightarrow s$	дано
4.	$\sim z \vee r$	дано
5.	$\sim\sim p \wedge \sim q$	1 и закон де Моргана
6.	$p \wedge \sim q$	5 и двойное отрицание
7.	s	3, 6 и modus ponens
8.	$s \rightarrow z$	2 и контрапозиция
9.	z	7, 8 и modus ponens
10.	$z \rightarrow r$	4 и теорема 1.3(и)
11.	r	9, 10 и modus ponens.

г)

1.	$\sim x \rightarrow \sim w$	дано
2.	$(x \vee \sim w) \rightarrow z$	дано
3.	$\sim p \rightarrow \sim z$	дано
4.	$p \rightarrow (\sim r \vee \sim s)$	дано
5.	$x \vee \sim w$	1 и теорема 1.3(h)
6.	z	2, 5 и modus ponens
7.	$z \rightarrow p$	3 и контрапозиция
8.	p	6, 7 и modus ponens
9.	$\sim r \vee \sim s$	4, 8 и modus ponens.

9. а) нет; **б)** да; **в)** нет; **г)** да; **д)** нет.

Раздел 1.5.

1. а)

Случай	p	q	p	$\downarrow$	q	$\sim p$
1	T	T	T	**F**	T	**F**
2	T	F	T	**F**	T	**F**
3	F	T	F	**T**	F	**T**
4	F	F	F	**T**	F	**T**
				*		*

б)

Случай	p	q	$(p$	$\downarrow$	$p)$	$\downarrow$	$(q$	$\downarrow$	$q)$	$p \wedge q$
1	T	T	T	F	T	**T**	T	F	T	**T**
2	T	F	T	F	T	**F**	F	T	F	**F**
3	F	T	F	T	F	**F**	T	F	T	**F**
4	F	F	F	T	F	**F**	F	T	F	**F**
						*				*

в)

Случай	p	q	$(p$	$\downarrow$	$q)$	$\downarrow$	$(p$	$\downarrow$	$q)$	$p \vee q$
1	T	T	T	F	T	**T**	T	F	T	**T**
2	T	F	T	F	F	**T**	T	F	F	**T**
3	F	T	F	F	T	**T**	F	F	T	**T**
4	F	F	F	T	F	**F**	F	T	F	**F**
						*				*

3. а) $(p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge r) \vee (p \wedge \sim q \wedge \sim r) \vee (\sim p \wedge \sim q \wedge \sim r)$;

б) $(\sim p \wedge q \wedge r) \vee (\sim p \wedge q \wedge \sim r) \vee (\sim p \wedge \sim q \wedge \sim r)$;

в) $(p \wedge q \wedge r) \vee (p \wedge \sim q \wedge r) \vee (p \wedge \sim q \wedge \sim r) \vee (\sim p \wedge \sim q \wedge r)$.

Раздел 1.6.

1. а) $r \vee (p \wedge \sim q)$;

б) $(r \wedge s) \vee (p \wedge \sim q \wedge \sim s) \vee (p \wedge \sim r \wedge \sim s)$;

в) $(p \wedge \sim r) \vee (\sim r \wedge s) \vee (\sim q \wedge r \wedge \sim s)$.

3. а)

	q	q	$\sim q$	$\sim q$	
p	×	×			
$\sim p$	×	×	×		
	r	$\sim r$	$\sim r$	r	

$q \vee (\sim p \wedge \sim r)$;

б)

	q	q	$\sim q$	$\sim q$	
p		×	×		s
p		×	×		$\sim s$
$\sim p$	×	×		×	$\sim s$
$\sim p$		×	×		s
	r	$\sim r$	$\sim r$	r	

$(p \wedge \sim r) \vee (q \wedge \sim r) \vee (\sim p \wedge \sim r \wedge s) \vee (\sim p \wedge \sim s \wedge r)$;

в)

	q	q	$\sim q$	$\sim q$	
p	×				s
p			×	×	$\sim s$
$\sim p$			×	×	$\sim s$
$\sim p$	×				s
	r	$\sim r$	$\sim r$	r	

$(\sim q \wedge \sim s) \vee (q \wedge r \wedge s)$.

Раздел 1.7.

1. а) $pq(q + r')$;

б) $(q + r')(p' + r)(q + r' + s)$;

в) $(pq'r)' + q'r + pq'r'$;

г) $(p' + qr')'$;

д) $pr + pq'(p + r + s')$.

3. а) $(p'q')' + q'r$;

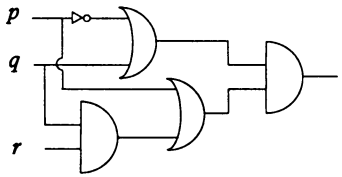
б) $(pr)' + qr'$;

в) $(pq + r) + (q + r)'$;

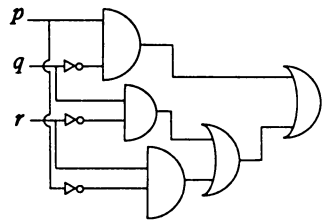
г) $((p + q + r)(p' + q + r')')'$;

д) $(p + q)((qr')' + (p' + r))$.

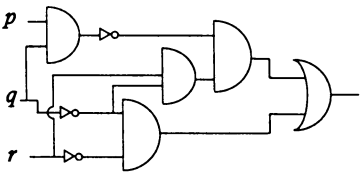
5. а)



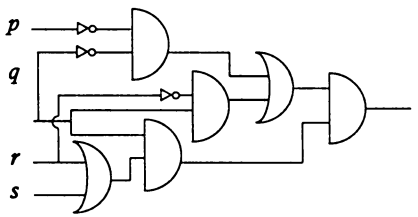
б)



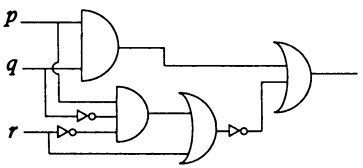
в)



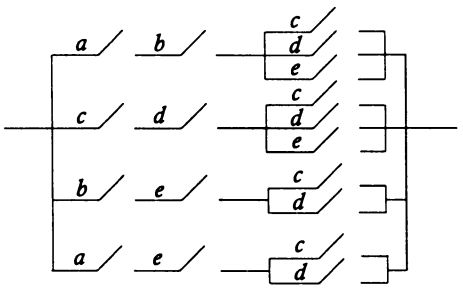
г)



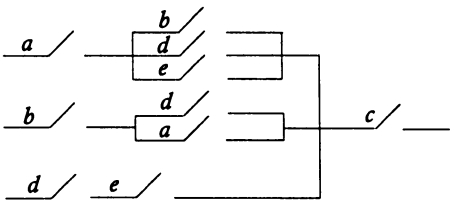
д)



7.



9.



Раздел 2.1.

1. $\{-9, -8, -7, -6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

3. $\{a, e, i, o, u\}$.

5. $\{x : 0 < x \leq 24$ и x кратно $3\}$.

7. $\{x : x$ — название американского штата, которое начинается с буквы "О"$\}$.

9. $\varnothing, \{a\}$.

11. $\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}$.

13. $\varnothing$.

15. **а)** истинно; **б)** ложно; **в)** ложно; **г)** истинно; **д)** ложно.

17. **а)** 2; **б)** 1; **в)** 4; **г)** 6; **д)** 3.

Раздел 2.2.

1. **а)** $\{1, 2, 3, 4, 5, 6, 7, 8, 10\}$; **б)** $\{4, 5, 6, 7\}$; **в)** $\{2, 4, 5, 6, 7\}$;
г) $\{2, 4, 5, 6, 7, 8, 10\}$; **д)** $\{1, 2, 3, 8, 9, 10\}$; **е)** $\varnothing$;
ж) $\{1, 2, 3, 8, 9, 10\}$; **з)** $\{1, 2, 3\}$.

3. **а)** $\{(1, a), (2, a), (3, a), (1, b), (2, b), (3, b)\}$;
б) $\{(a, a), (a, b), (b, a), (b, b)\}$;
в) $\varnothing$.

5. $\{\varnothing\}$.

7. $\{\varnothing, \{\varnothing\}\}$.

9. **а)** T; **б)** T; **в)** F; **г)** T; **д)** T.

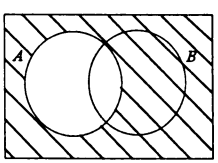
11.
$$
\begin{aligned}
x \in (A \cup B)' &\Leftrightarrow x \notin (A \cup B) && \text{определение дополнения} \\
&\Leftrightarrow \sim (x \in (A \cup B)) && \text{определение } \notin \\
&\Leftrightarrow \sim ((x \in A) \vee (x \in B)) && \text{определение объединения} \\
&\Leftrightarrow \sim (x \in A) \wedge \sim (x \in B) && \text{закон де Моргана;} \\
&\Leftrightarrow (x \notin A) \wedge (x \notin B) && \text{определение } \notin \\
&\Leftrightarrow x \notin A \cap B && \text{определение пересечения.}
\end{aligned}
$$

13. **а)** $\{\varnothing\}$; **б)** $\{\varnothing, \{\varnothing\}\}$; **в)** $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$.

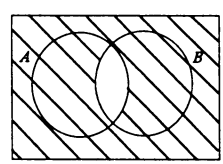
15. N.

Раздел 2.3.

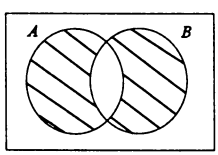
1. **а)**



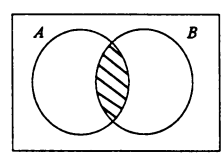
б)



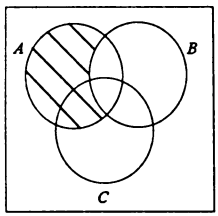
в)



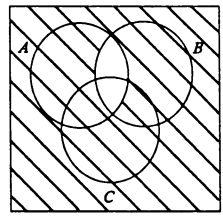
г)



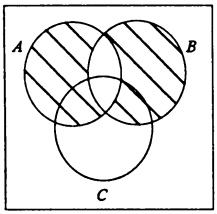
3. а)



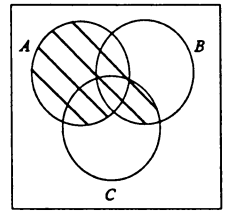
б)



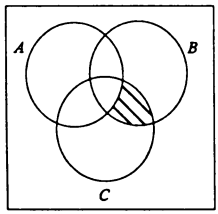
в)



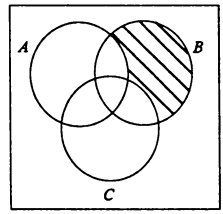
г)



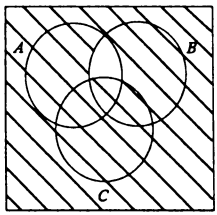
д)



е)



ж)



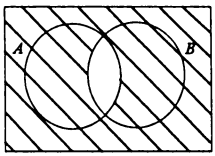
5. а) A';

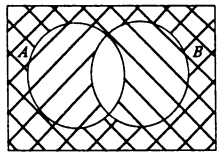
б) $((A \cup B) - C) \cup (A \cap B)$;

в) $((A \cap B) \cup (A \cap C) \cup (B \cap C))'$;

г) $((A \cap B) \cup (A \cap C) \cup (B \cap C)) - (A \cap B \cap C)$.

7.





Раздел 2.4.

1. $x \cdot x = x \cdot x + 0 =$ закон тождества
$= x \cdot x + x \cdot x' =$ закон дополнения
$= x \cdot (x + x') =$ закон дистрибутивности
$= x \cdot 1 =$ закон дополнения
$= x$ закон тождества

3. $x \cdot (x + y) = (x + 0) \cdot (x + y) =$ закон тождества
$= x + (0 \cdot y) =$ закон дистрибутивности
$= x + 0 =$ свойства констант
$= x$ закон тождества.

5. $1 + 0 = 1$ закон тождества
$1 \cdot 0 = 0$ закон тождества.
Следовательно, 0 удовлетворяет закону дополнения для 1 и, согласно закону единственности дополнения, является единственным элементом, который обладает таким свойством.

7. **а)** $xy' + (z'y)'$; **б)** $((0 \cdot x) + y)(x + y' + z)$;
в) $((x + y) \cdot 0) + (1 \cdot x) + z$.

9. Если $x + y = x + z$ и $x' + y = x' + z$, то $y = z$.

11. $x + y' = 1$ тогда и только тогда, когда $x + y = x$.

13. Объединение и пересечение двух конечных множеств является конечным множеством. Объединение и пересечение двух коконечных множеств коконечно. Объединение конечного множества и коконечного множество коконечно. Пересечение конечного множества и коконечного множества конечно. Поэтому свойство объединения и пересечения определено для всех рассматриваемых множеств. Для множеств было показано выполнение законов коммутативности, ассоциативности и дистрибутивности. Пустое множество конечно, а множество $\boldsymbol{U}$ коконечно. Поэтому они относятся к рассматриваемым множествам. Дополнение конечного множества коконечно, а дополнение коконечного множества конечно. Следовательно, дополнение определено для всех рассматриваемых множеств. Справедливость законов тождества и законов дополнения для множеств была доказана ранее.

Раздел 2.5.

1. **а)** область значений $= \{1, 2, 4, 5\}$; область определения $= \{a, c, d\}$;
б) область значений $= \{2, 4, 6, \ldots\}$; область определения $= \{1, 2, 3, \ldots\}$.
в) область значений $= R$; область определения $= \{x : x \in R \text{ и } x \geq 0\}$.

3. **а)** $R^{-1} = \{(7,1), (6,4), (6,5), (8,2)\}$,
$S^{-1} = \{(10,6), (11,6), (10,7), (13,8)\}$;
б) $S \circ R = \{(1,10), (4,10), (4,11), (5,10), (5,11), (2,13)\}$;
в) $S \circ S^{-1} = \{(10,10), (10,11), (11,10), (11,11), (13,13)\}$,
$S^{-1} \circ S = \{(6,6), (6,7), (7,6), (7,7), (8,8)\}$;
г) $R^{-1} \circ S^{-1} = \{(10,4), (10,5), (11,4), (11,5), (10,1), (13,2)\}$;
д) $T \circ (S \circ R) = \{(1, \triangle), (4, \triangle), (5, \triangle), (2, *), (2, \bigcirc)\}$;
е) $T \circ S = \{(6, \triangle), (7, \triangle), (8, *)(8, \bigcirc)\}$;
ж) $(T \circ S) \circ R = \{(1, \triangle), (4, \triangle), (5, \triangle), (2, *), (2, \bigcirc)\}$.

5. **а)** $\{(x, y) : y = 9x^2 + 5\}$;
б) $\{(x, y) : y = 3x^2 + 15\}$;
в) $\{(x, y) : y = \pm\sqrt{x - 5}$;
г) $\{(x, y) : y = \frac{1}{3}x$;

д) $\{y : y \geq 5\}$.

7. а) $\{(a,a),(a,b),(b,a),(b,c),(c,b),(b,d),(d,b),(c,e),(e,c),(e,d),(d,e)\}$;

б) $\{(a,a),(a,b),(b,a),(b,c),(c,b),(b,d),(d,b),(c,e),(e,c),(e,d),(d,e),(b,b),(c,c),(d,d),(e,e)\}$;

в) $A \times A$;

г) $\{(a,a),(a,b),(b,c),(b,d),(c,e),(e,d),(c,b),(a,c),(a,d),(b,e),(b,b),(c,d),(c,c)\}$.

9. а) $U \cap V = \{(a,b),(b,c),(b,b),(e,e),(b,a),(c,b),(d,d),(a,c),(c,a)\}$;

б) $S \cup T = \{(a,a),(a,b),(b,c),(b,d),(c,e),(e,d),(c,a),(b,a),(e,e),(d,e),(c,b)\}$;

в) $U - T = \{(a,a),(b,b),(c,c),(d,d),(a,c),(c,a)\}$;

г) $U \Delta S = \{(b,b),(e,e),(b,a),(c,b),(c,c),(d,d),(a,c),(b,d),(c,e),(e,d)\}$

11. Пусть R и S — симметричные отношения и $(a,b) \in R \cap S$. Поскольку R симметрично, то $(b,a) \in R$. Поскольку S симметрично, то $(b,a) \in S$. Следовательно, $(b,a) \in R \cap S$ и $R \cap S$ симметрично.

13. а) $\{(a,a),(b,b),(c,c),(d,d),(e,e),(a,b),(b,a),(c,b),(b,c)\}$;

б) $\{(a,a),(b,b),(c,c),(d,d)\}$;

в) $\{(a,a),(b,b),(c,c),(d,d),(e,e),(a,b)\}$.

15. а) истинно; **б)** истинно;

в) ложно. Пусть $R = \{(b,c),(c,b)\}$ и $S = \{(b,a),(a,b)\}$, тогда $(a,c) \in R \circ S$, но $(c,a) \notin R \circ S$.

г) истинно; **д)** истинно.

17. а) истинно;

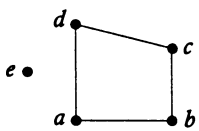
б) ложно. Пусть $R = \{(b,c)\}$ и $S = \{(a,b)\}$, тогда R и S транзитивны, но $R \cup S = \{(a,b),(b,c)\}$ не транзитивно;

в) ложно. Пусть $R = \{(c,a)(b,b)\}$ и $S = \{(a,b),(b,c),(a,c)\}$, тогда R и S транзитивны, но $R \circ S = \{(a,b),(b,a),(a,a)\}$ не транзитивно;

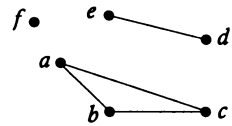
г) ложно. Пусть $R = \{(a,b),(b,c),(a,c)\}$ и $S = \{(a,c)\}$, тогда R и S транзитивны, но $R - S = \{(a,b),(b,c)\}$ не транзитивно;

д) $R = \{(a,b),(b,c),(a,c)\}$ и $S = \{(a,c)\}$, тогда R и S транзитивны, но $R \Delta S = \{(a,b),(b,c)\}$ не транзитивно.

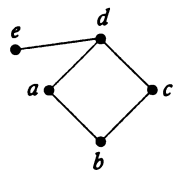
19. а)



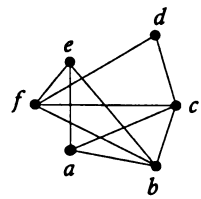
б)



в)



г)



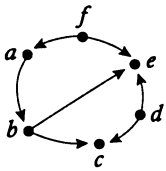
21. а) $V = \{a, b, c, d\}$;
$E = \{\{a, b\}, \{a, d\}, \{b, d\}, \{c, d\}, \{a, c\}\}$;
$R = \{(a, b), (b, a), (a, d), (d, a), (b, d), (d, b), (c, d), (d, c), (a, c), (c, a)\}$.

б) $V = \{a, b, c, d\}$;
$E = \{\{a, b\}, \{a, d\}, \{b, d\}, \{c, d\}, \{a, c\}, \{b, c\}\}$;
$R = \{(a, b), (b, a), (a, d), (d, a), (b, d), (d, b), (c, d), (d, c), (a, c), (c, a), (b, c), (c, b)\}$.

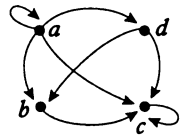
в) $V = \{a, b, c, d, e, f\}$;
$E = \{\{a, c\}, \{a, d\}, \{a, e\}, \{a, f\}, \{b, c\}, \{b, d\}, \{b, e\}, \{b, f\}\}$;
$R = \{(a, c), (c, a), (a, d), (d, a), (a, e), (e, a), (a, f), (f, a), (b, c), (c, b), (b, d);$
$(d, b), (b, e), (e, b), (b, f), (f, b)\}$.

г) $V = \{a, b, c, d, e, \}$;
$E = \{\{a, b\}, \{c, d\}, \{d, e\}; \{c, e\},$
$R = \{(a, b), (b, a), (c, d), (d, c), (d, e), (e, d), (c, e), (e, c)\}$.

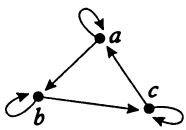
23. а)



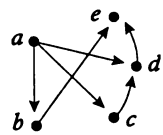
б)



в)



г)



25. а) $V = \{a, b, c, d, e, \}$;
$E = \{(a, b), (b, a), (a, e), (b, e), (c, d), (d, c), (d, e), (c, e)\}$.

б) $V = \{a, b, c, d\}$;
$E = \{(d, b), (d, a), (c, d), (b, c)\}$.

в) $V = \{a, b, c, d, e, \}$;
$E = \{(a, b), (b, c), (c, d), (d, e), (e, a), (b, d), (b, e), (c, e), (d, a)\}$.

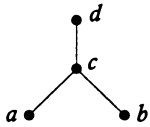
г) $V = \{a, b, c, d, e, \}$;
$E = \{(a, b), (b, c), (c, d), (d, e), (e, a), (b, d), (b, e), (c, e), (e, d), (a, c)\}$.

27. Покажите, что $(T \circ S) \circ R \subseteq T \circ (S \circ R)$. Пусть $(a, d) \in (T \circ S) \circ R$, тогда существует $b \in B$ такое, что $(a, b) \in R$ и $(b, d) \in T \circ S$. Поскольку $(b, d) \in T \circ S$, то существует $c \in C$ такое, что $(b, c) \in S$ и $(c, d) \in T$. Поскольку $(a, b) \in R$ и $(b, c) \in S$, то $(a, c) \in S \circ R$. Поскольку $(b, d) \in T \circ S$ и $(a, b) \in R$, то $(a, d) \in (T \circ S) \circ R$. Таким образом, $T \circ (S \circ R) \subseteq (T \circ S) \circ R$.

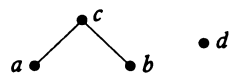
Раздел 2.6.

1. (а), (г).

3. а)



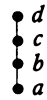
б)



в)



г)



5. а) $\{(a,a),(b,b),(c,c),(d,d),(e,e),(d,a),(d,c),(d,b),(d,e),(a,b),(c,b)\}$;
б) $\{(a,a),(b,b),(c,c),(d,d),(e,e),(d,b),(d,a),(d,e),(c,e),(c,a),(c,b),(e,a),(e,b)\}$;
в) $\{(a,a),(b,b),(c,c),(d,d),(e,e),(f,f),(a,b),(a,c),(f,b),(f,c),(f,d),(e,d),;(e,c)\}$
г) $\{(a,a),(b,b),(c,c),(d,d),(e,e),(e,a),(e,b),(e,c),(e,d),(d,a),(d,b),(d,c)\}$;
д) $\{(a,a),(b,b),(c,c),(d,d),(e,e),(c,a),(c,b),(e,a),(e,b)\}$.

7. Никакие два элемента не сравнимы.

Раздел 2.7.

1. а) нет; **б)** нет.
в) да, $[p]=\{q : q$ имеет такую же таблицу истинности, как и $p\}$;
г) да, $[(a,b]=\{(c,d) : ad=bc\}$;
д) да, $[n]=\{n,-n\}$;
е) да, $[n]=\{n\}$.

3. а) да, $[B]=\{C : C$ содержит такое же количество элементов, как и $B\}$;
б) да, $[1]=A$;
в) да, $[1]=\{1,3,5,7,9\}, [2]=\{2,4,6,8,10\}$;
г) нет;
д) Да, $[n]=\{m : m$ параллельно $n\}$;
е) Да, $[p]=\{q : p$ написано на том же языке программирования, что и $q\}$.

5. а) для всех $a \in A$ имеем $f(a)=f(a)$, поэтому R рефлексивно. Если $f(a)=f(b)$, то $f(b)=f(a)$, поэтому R симметрично. Если $f(a)=f(b)$ и $f(b)=f(c)$, то $f(a)=f(c)$, поэтому R транзитивно. Следовательно, R — отношение эквивалентности;
б) $[1]=\{1,-1\}$; $[2]=\{2,-2\}$; $[0]=\{0\}$; $[3]=\{3\}$; $[-\frac{1}{2}]=\{\frac{1}{2};-\frac{1}{2}\}$; $[4]=\{4\}$.

7. а) нет;
б) $R=\{(1,1),(1,2),(2,1),(2,2),(3,3),(4,4),(5,5),(3,4)(4,3),(3,5),(5,3),(4,5),(5,4),(6,6),(6,7),(7,6),(7,7)\}$;
в) нет; **г)** нет; **д)** $A \times A$.

9. Да. Покажем, что R рефлексивно. Имеем, что $(a,b)R(a,b)$ тогда и только тогда, когда $ab=ab$. Покажем, что R симметрично. Допустим, что $(a,b)R(c,d)$. Но $(a,b)R(c,d)$ тогда и только тогда, когда $ad=bc$. Также $(c,d)R(a,b)$ тогда и только тогда, когда $cb=ad$. Следовательно, $(c,d)R(a,b)$. Покажем, что R транзитивно. Пусть $(a,b)R(c,d)$ и $(c,d)R(e,f)$. Следовательно, $ad=bc$ и $cf=ed$. Из первого равенства $adf=bcf$. Подставляя $cf=ed$ в это равенство, получаем $adf=bed$. Разделив на d, получаем $af=be$ и $(a,b)R(e,f)$.

Раздел 3.1.

1. а) $3^2+4^2 \geq 5^2$; **б)** $2=\frac{8-4}{8+4}$; **в)** $|3-1| \leq 7$; **г)** $24=4!$.
д) Джон любит Сью больше, чем Мэри.

3. а) существуют x и y такие, что $x^2 + y^2 \geq 25^2$;

б) существует x такое, что $7 = \frac{x-4}{x+4}$;

в) для каждого r существует x такое, что $|x - 1| \leq r$;

г) существует n такое, что для всех y выполняется $y = n!$;

д) Джон любит Сью больше, чем кого-либо.

5. а) предметная область для x — множество всех улиц. Область для y — множество всех праздников. $R(x, y)$: x имеет свой y.
$\forall x \exists y R(x, y)$;

б) предметная область для x — множество всех машин. Область для y — множество всех людей. $S(x, y)$: x умнее, чем y.
$\exists x \forall y S(x, y)$;

в) предметная область — люди. $T(x, y)$: x играет в теннис лучше, чем y.
$\forall x T(x,$Фрэд$)$;

г) предметная область для x — множество всех действий. Область для y — множество всех противодействий. $U(x, y)$: x имеет равное и противоположно направленное y.
$\forall x \exists y U(x, y)$;

д) предметная область — игроки в гольф. $V(x, y)$: x будет в конце концов обыгран более сильным y.
$\forall x \exists y V(x, y)$.

7. а) истинно; **б)** истинно; **в)** истинно; **г)** истинно; **д)** ложно.

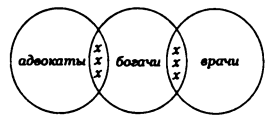
9.

$\forall x P(x) \vee \forall x Q(x)$	дано
$P(a) \vee \forall x Q(x)$ для произвольного a	универсальная конкретизация
$P(a) \vee Q(b)$ для произвольного b	универсальная конкретизация
$P(a) \vee Q(a)$	поскольку b произвольно, выберем $b = a$
$\forall x (P(x) \vee Q(x))$	универсальное обобщение.

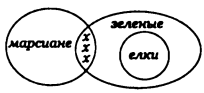
11. а)



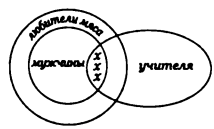
б)



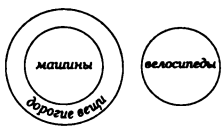
12. а)



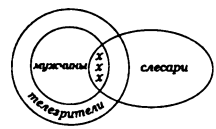
б)



14. а)



б)



Раздел 3.2.

1. а) $b+a=c+a$ дано

$(b+a)+(-a)=(c+a)+(-a)$ I2

$b+(a+(-a))=(c+a)+(-a)$ закон ассоциативности

$b+0=c+0$ I7 аддитивно обратный элемент

$b=a$ I6 аддитивно нейтральный элемент;

б) $a\cdot 0=a\cdot(0+0)=$ I6 аддитивно нейтральный элемент

$=a\cdot(0)+a\cdot(0)$ закон дистрибутивности.

Таким образом,

$a\cdot 0-a\cdot 0=(a\cdot(0)+a\cdot(0))-a\cdot 0=$ аддитивно обратный элемент

$a\cdot 0-a\cdot 0=a\cdot(0)+(a\cdot(0))-a\cdot 0)=$ закон ассоциативности

$0=a\cdot(0)+0=$ аддитивно обратный элемент

$=a\cdot(0)$ аддитивно нейтральный элемент;

в) поскольку $-a$ является обратным к a, имеем $a+(-a)=(-a)+a$. Но это удовлетворяет требованию, при котором a является обратным к $-a$. Следовательно, $a=-(-a)$;

г) $a\cdot(-b)+a\cdot b=a\cdot(-b+b)=$ согласно закону дистрибутивности

$=a\cdot 0=$ по определению обратного элемента

$=0$ согласно пункту (б).

Аналогично, $a\cdot b+a\cdot(-b)=0$, поэтому, по определению, $a\cdot(-b)$ является аддитивно обратным элементом для $a\cdot b$, поэтому $a\cdot(-b)=-ab$;

д) $a\cdot(-b)+(-a)\cdot(-b)=(-b)\cdot a+(-b)\cdot(-a)=$ закон коммутативности

$=(-b)\cdot(a+(-a))=$ закон дистрибутивности

$=(-b)\cdot 0=$ определение обратного элемента

$=0$ согласно пункту (б).

Аналогично, $(-a)\cdot(-b)+a\cdot(-b)=0$, поэтому $(-a)\cdot(-b)$ — обратный элемент для $a\cdot(-b)$. Но $a\cdot b$ — элемент, обратный для $a\cdot(-b)$. Следовательно, $a\cdot b=(-a)\cdot(-b)$.

3. Учитывая, что если $a>b\geq 0$ и $c>d\geq 0$, то $ac>bd$. Требуется рассмотреть только случай, когда $a=c$ или $c=d$. Если $a=b$ и $c\geq d$, то $ac\geq ad$. Следовательно, $ac\geq bd$. Случай $c=d$ аналогичен.

5. Поскольку $(a-b)^2\geq 0$, имеем

$$(a-b)^2=(a-b)(a-b)=a(a-b)+(-b)(a-b)\geq 0,$$

согласно закону дистрибутивности. Используя совместно закон дистрибутивности, пункты (д) и (е) задачи 1 и закон коммутативности, получаем $a(a-b)+(-b)(a-b)=a^2-ab-ab+b^2$. Следовательно, $a^2-2ab+b^2\geq 0$. Прибавляя $2ab$ к обеим частям неравенства, получаем $a^2+b^2\geq 2ab$.

Раздел 3.3.

1. Для $n=1$ имеем $1=\frac{1(3-1)}{2}$, поэтому утверждение истинно. Предположим, что утверждение истинно для $n=k$, так что $1+4+7+\cdots+(3k-2)=\frac{k(3k-1)}{2}$. Теперь необходимо доказать, что утверждение истинно для $n=k+1$, так что

$$1+4+7+\cdots+(3k-2)+(3(k+1)-2)=\frac{(k+1)(3(k+1)-1)}{2}=\frac{(k+1)(3k+2)}{2}.$$

Если прибавить $(3(k+1)-2) = 3k+1$ к обеим частям утверждения, записанного для $n = k$, имеем

$$1+4+7+\cdots+(3k-2)+(3(k+1)-2) =$$

$$= \frac{k(3k-1)}{2} + 3k + 1 = \frac{3k^2+5k+2}{2} = \frac{(k+1)(3k+2)}{2},$$

и справедливость утверждения для $n = k+1$ доказана.

3. Для $n = 1$ имеем $1 = 2-1$, поэтому утверждение истинно для $n = 1$. Предположим, что оно истинно для $n = k$, поэтому $1+2+2^2+\cdots+2^{k-1} = 2^k-1$. Мы хотим доказать истинность утверждения для $n = k+1$, т.е. $1+2+2^2+\cdots+2^{k-1}+2^k = 2^{k+1}-1$. Прибавляя 2^k к обеим частям утверждения, записанного для $n = k$, имеем $1+2+2^2+\cdots+2^{k-1}+2^k = 2^k-1+2^k$. Но $2^k-1+2^k = 2\cdot 2^k-1 = 2^{k+1}-1$. Таким образом, истинность утверждения для $n = k+1$ доказана.

5. Для $n = 1$ имеем $1 = 1^2$, поэтому утверждение истинно. Предположим, что оно истинно для $n = k$, поэтому имеем $1+3+5+\cdots+(2n-1) = n^2$. Мы хотим доказать истинность утверждения для $n = k+1$, т.е. равенство $1+3+5+\cdots+(2n-1)+(2n+1) = (n+1)^2$. Прибавляя $2n+1$ к обеим частям равенства, записанного для $n = k$, имеем $1+3+5+\cdots+(2n-1)+(2n+1) = n^2+(2n+1) = (n+1)^2$, что доказывает истинность утверждения для $n = k+1$.

7. Для $n = 1$ имеем $1(1+1) = \dfrac{1(1+1)(1+2)}{3}$ или $2 = 2$, поэтому утверждение истинно. Предположим, что оно истинно для $n = k$, поэтому имеем $1\cdot 2+2\cdot 3+\cdots+k(k+1) = \dfrac{k(k+1)(k+2)}{3}$. Мы хотим доказать, что утверждение истинно для $n = k+1$, т.е. равенство

$$1\cdot 2+2\cdot 3+\cdots+k(k+1)+(k+1)(k+2) = \frac{(k+1)(k+2)(k+3)}{3}.$$

Прибавляя $(k+1)(k+2)$ к обеим частям равенства, записанного для $n = k$, имеем

$$1\cdot 2+2\cdot 3+\cdots+k(k+1)+(k+1)(k+2) = \frac{k(k+1)(k+2)}{3}+(k+1)(k+2),$$

но

$$\frac{k(k+1)(k+2)}{3}+(k+1)(k+2) =$$

$$= \frac{k(k+1)(k+2)+3(k+1)(k+2)}{3} = \frac{(k+3)(k+1)(k+2)}{3},$$

что доказывает истинность утверждения для $n = k+1$.

9. Пусть $n = 1$. Тогда $ab = ab$, что и требовалось. Предположим истинность утверждения для $n = k$, согласно которому $(ab)^k = a^k b^k$.

Мы хотим доказать, что $(ab)^{k+1} = a^{k+1}b^{k+1}$. Но

$$\begin{aligned}
(ab)^{k+1} &= (ab)^k(ab) = && \text{согласно определению, введенному в предыдущей задаче}\\
&= a^k b^k(ab) = && \text{согласно индуктивному предположению}\\
&= a^k a^k b^k b = && \text{согласно коммутативности умножения}\\
&= a^{k+1}b^{k+1} && \text{согласно определению, введенному в предыдущей задаче.}
\end{aligned}$$

11. В этой задаче используются результаты предыдущей задачи. Пусть $n = 5$, тогда $2^5 > 5^2$, что и требовалось. Предположим, что утверждение истинно для $n = k$, поэтому $2^k > k^2$. Мы хотим доказать, что $2^{k+1} > (k+1)^2$. Но по индуктивному предположению $2^k > k^2$ и согласно результату предыдущей задачи $2^k > 2k + 1$. Поэтому, складывая, получаем $2^k + 2^k > k^2 + 2k + 1$ или $2^{k+1} > (k+1)^2$.

13. Для $n = 8$ имеем $8 = 3(1) + 5(1)$. Предположим истинность утверждения для $n = k$, поэтому $k = 3r + 5s$. Если $r \geq 3$, то $k + 1 = 3(r-3) + 5(s+2)$. Если $r < 3$, то $s \geq 1$, поэтому $k+1 = 3(r+2) + 5(s-1)$. Так или иначе, мы выразили $k+1$ в надлежащем виде.

15. Для $n = 1$ имеем $a^m = a^m$, поэтому равенство истинно для $n = 1$. Предположим истинность для $n = k$, поэтому $a^{mk} = (a^m)^k$. Используя затем индуктивное предположение, имеем

$$a^{m(k+1)} = a^{mk+m} = a^{mk} \cdot a^m = (a^m)^k \cdot a^m = (a^m)^{k+1}.$$

17. Для $n = 2$ имеем $1 - \frac{1}{4} = \frac{2+1}{4}$ или $\frac{3}{4} = \frac{3}{4}$, поэтому утверждение истинно. Предположим истинность для $n = k$, поэтому

$$(1 - \frac{1}{4})(1 - \frac{1}{9})(1 - \frac{1}{16}) \cdots \left(1 - \frac{1}{k^2}\right) = \frac{k+1}{2k}.$$

Мы хотим показать, что

$$(1 - \frac{1}{4})(1 - \frac{1}{9})(1 - \frac{1}{16}) \cdots (1 - \frac{1}{k^2}) \left(1 - \frac{1}{(k+1)^2}\right) = \frac{k+2}{2k+2}.$$

Умножая на $(1 - \frac{1}{(k+1)^2})$ обе части равенства, записанного для $n = k$, имеем

$$(1 - \frac{1}{4})(1 - \frac{1}{9})(1 - \frac{1}{16}) \cdots (1 - \frac{1}{k^2}) \left(1 - \frac{1}{(k+1)^2}\right) = \frac{k+1}{2k} \left(1 - \frac{1}{(k+1)^2}\right).$$

Но

$$\frac{k+1}{2k} \left(1 - \frac{1}{(k+1)^2}\right) = \frac{k+1}{2k} \left(\frac{(k+1)^2 - 1}{(k+1)^2}\right) = \frac{k^2 + 2k}{2(k+1)} = \frac{k+2}{2k+2},$$

и утверждение истинно для $n = k + 1$.

19. Некорректный переход от $k = 1$ к $k + 1 = 2$.

21. **а)** для $n = 1$ имеем $A_1 \cap A = A_1 \cap A$. Предположим истинность утверждения для $n = k$, так что $(\bigcap_{i=1}^{k} A_i) \cap A = \bigcap_{i=1}^{k} (A_i \cap A)$. Мы хотим показать, что $(\bigcap_{i=1}^{k+1} A_i) \cap A = \bigcap_{i=1}^{k+1} (A_i \cap A)$. Но

$$(\bigcap_{i=1}^{k+1} A_i) \cap A = ((\bigcap_{i=1}^{k} A_i) \cap A_{k+1}) \cap A = \quad \text{определение из предыдущей задачи}$$

$$= ((\bigcap_{i=1}^{k} A_i) \cap A_{k+1}) \cap A \cap A = \quad \text{поскольку } A \cap A = A$$

$$= ((\bigcap_{i=1}^{k} A_i) \cap A) \cap A_{k+1} \cap A = \quad \text{закон коммутативности}$$

$$= (\bigcap_{i=1}^{k} (A_i \cap A)) \cap (A_{k+1} \cap A) = \quad \text{индуктивное предположение}$$

$$= (\bigcap_{i=1}^{k+1} (A_i \cap A)) \quad \text{определение из предыдущей задачи;}$$

б) для $n = 1$ имеем $A_1' = A_1'$, поэтому равенство справедливо. Предположим истинность утверждения для $n = k$, поэтому имеем $\left(\bigcap_{i=1}^{k} A_i\right)' = \bigcup_{i=1}^{k} A_i'$. Мы хотим доказать, что $\left(\bigcap_{i=1}^{k+1} A_i\right)' = \bigcup_{i=1}^{k+1} A_i'$. По закону де Моргана $\left(\bigcap_{i=1}^{k+1} A_i\right)' = \left(\bigcap_{i=1}^{k} A_i \cap A_{k+1}\right)' = \left(\bigcap_{i=1}^{k} A_i\right)' \cup A_{k+1}'$. Воспользовавшись индуктивным предположением, имеем $\left(\bigcap_{i=1}^{k} A_i\right)' \cup A_{k+1}' = \bigcup_{i=1}^{k} A_i' \cup A_{k+1}' = \bigcup_{i=1}^{k+1} A_i'$. Следовательно, утверждение истинно для $n = k+1$.

23. Предположим, что T удовлетворяет условиям (а) и (б) второго принципа индукции. Пусть T' — множество всех целых чисел, для которых утверждение T не является истинным. Если T' не пусто, тогда, согласно принципу полного упорядочения, T' содержит наименьшее положительное целое число, например, n. Но поскольку n — наименьший элемент в T' для всех m таких, что $m < n$, утверждение T истинно. Это противоречит (б), поэтому T' пусто и T истинно для всех n.

25. Предположим, что T не пусто. Тогда, согласно принципу вполне упорядочения, это множество имеет наименьший элемент, например, n. Но по условию 2 существует $m < n$, которое принадлежит T. Это является противоречием, поэтому T пусто.

Раздел 3.4.

1. **а)** $54, 27, 18, 9, 6, 3, 2, 1$; **б)** $63, 21, 9, 7, 3, 1$;
в) $72, 36, 18, 9, 24, 12, 6, 3, 8, 4, 2, 1$; **г)** $73, 1$;
д) $74, 37, 2, 1$.

3. **а)** НОД$(54, 27) = 27$, НОК$(54, 27) = 54$; НОД$(54, 27) \cdot$ НОК$(54, 27) = 1458$;
б) НОД и НОК не определены;
в) НОД$(6, 15) = 3$, НОК$(6, 15) = 30$; НОД$(6, 15) \cdot$ НОК$(6, 15) = 90$;
г) НОД$(12, 16) = 4$, НОК$(12, 16) = 48$; НОД$(12, 16) \cdot$ НОК$(12, 16) = 192$;
д) НОД$(33, 1) = 1$, НОК$(33, 1) = 33$; НОД$(33, 1) \cdot$ НОК$(33, 1) = 33$.

5. В качестве НОК$(a, 0)$ можно было бы взять 0. Число 0 является кратным для a и для 0. Если некоторое число делит a и 0, то оно делит 0.

7. НОД(a, b) делит a и b, поэтому он делит $a - b$. Следовательно, НОД$(a, b) |$ НОД$(a - b, b)$. Обратно, если любое целое число делит $a - b$ и b, то оно делит a. Поэтому НОД$(a - b, b) |$ НОД(a, b). Следовательно, НОД$(a - b, b) =$ НОД(a, b).

Раздел 3.5.

1. **а)** $7 \cdot 2^3 \cdot 13$;
б) $1599 = 1600 - 1 = (40 - 1)(40 + 1) = 3 \cdot 13 \cdot 41$;
в) $3 \cdot 23 \cdot 71$; **г)** 131; **д)** 523.

3. 5 и 7, 11 и 13, 17 и 19.

5. Если одно из простых чисел не равно 2, то сумма $a^2 + b^2$ никогда не будет простым числом, поскольку она делится на 2.

7. $479001603 = 12! + 3$ и $479001603 = 12! + 7$.

9. Пусть p, $p + 2$ и $p + 4$ — три последовательных нечетных числа. Число p может быть записано как $3m$, $3m + 1$ или $3m + 2$. Если оно имеет вид $3m$, то оно делится на 3. Если оно имеет вид $3m + 1$, то $p + 2$ делится на 3. Если оно имеет вид $3m + 2$, то $p + 4$ делится на 3. Поскольку одно из трех чисел должно делиться на 3, эти числа не могут быть все простыми.

11. Если $0 \le b(i) \le a(i)$, то $p_i^{b(i)} | p_i^{a(i)}$ для всех i, поэтому $b|a$. Если $b|a$, то $p_i^{b(i)} | a$. Учитывая единственность разложения целого числа на простые множители, можем утверждать, что $p_i^{b(i)}$ является частью единственного разложения числа a, поэтому $p_i^{b(i)} | p_i^{a(i)}$.

Раздел 3.6.

1. **а)** 1; **б)** 3; **в)** 2; **г)** 14.

3. **а)** 0; **б)** 1; **в)** 1; **г)** 1.

5. Если a — нечетное целое число, тогда $a = 2k + 1$ для некоторого k.
Следовательно,

$$a^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 4(k^2 + k) + 1,$$

так что

$$a^2 - 1 = 4(k^2 + k) = 4k(k+1),$$

и поскольку либо k, либо $k+1$ — четно, то $a^2 - 1$ делится на 8, и $a^2 \equiv 1 \pmod 8$.

7.

·	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

9.

·	[0]	[1]	[2]	[3]	[4]	[5]	[6]
[0]	[0]	[0]	[0]	[0]	[P]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]	[6]
[2]	[0]	[2]	[4]	[6]	[1]	[3]	[5]
[3]	[0]	[3]	[6]	[2]	[5]	[1]	[4]
[4]	[0]	[4]	[1]	[5]	[2]	[6]	[3]
[5]	[0]	[5]	[3]	[1]	[6]	[4]	[2]
[6]	[0]	[6]	[5]	[4]	[3]	[2]	[1]

11. **а)** [3]; **б)** [2]; **в)** [1]; **г)** [3]; **д)** [0], [2].

13. Если $a \equiv b \pmod{mn}$, то $a - b$ равно kmn для некоторого k. Учитывая, что $a - b$ равно $(kn)m$, имеем $a \equiv b \pmod m$. Аналогично, $a \equiv b \pmod n$.

15. Уже известно, что существует одно и только одно r, где $0 \leq r < n$ такое, что $a \equiv r \pmod n$. Мы должны показать, что r является также взаимно простым с n. Поскольку $a = nq + r$ для некоторого q, и если r не является взаимно простым с n, то существует целое число m такое, что $n = cm$ и $r = dm$ для целых чисел c и d. Следовательно,

$$a = cmq + dm = (cq + d)m,$$

так что m также является множителем a, что противоречит предположению о том, что a и n — взаимно простые числа.

Раздел 4.1.

1. **а)** область определения f — множество R; область значений f — $\{y : y \geq 4\}$;
б) область определения f — $\{x : x \geq 2\}$; область значений f — $\{y : y \geq 0\}$;
в) область определения f — $\{x : x > 2\}$; область значений f — $\{y : y > 0\}$;
г) область определения f — множество R; область значений f — $\{y : 0 \leq \frac{1}{4}\}$;
д) область определения f — множество $R - \{2, -2\}$; область значений f — $\{y : y \leq -\frac{1}{4}\} \cup \{y : y > 0\}$;
е) область определения f — множество R; область значений f — $\{y : y \geq 0\}$.

3. **а)** $f(g(x)) = (x+3)^2 + 1$, $g(f(x)) = x^2 + 1 + 3 = x^2 + 4$;
б) $f(g(x)) = \sqrt{(x^2+3)^2 + 2}$, $g(f(x)) = x^2 + 5$;

в) $f(g(x)) = \dfrac{1}{2x+3}$, $g(f(x)) = \dfrac{2}{x} + 3$.

5. **а)** не удовлетворяет ни одному из условий;

б) не удовлетворяет ни одному из условий;

в) инъективна, сюръективна, имеет обратную функцию;

г) не удовлетворяет ни одному из условий;

д) сюръективна.

7. $f^{-1} = f^{-1}I_B = f^{-1}(fg) = (f^{-1}f)g = I_A\, g = g$.

9. **а)** предположим, что f не инъективна, тогда существуют $a, a' \in A$ такие, что $f(a) = f(a') = b$ для некоторого $b \in B$. Следовательно, $f^{-1}(f\{a\}) = f^{-1}\{b\} = \{a, a'\}$, поэтому неверно, что $f^{-1}(f(W) = W$ для всех $W \subseteq A$. Предположим, что существует $W \subseteq A$ такое, что $f^{-1}(f(W) \neq W$. Пусть $a \in f^{-1}(f(W)$ и $a \notin W$. Но $f(a) \in f(W)$ по определению f^{-1}, например, $f(a) = b$. Поскольку $b \in f(W)$, то существует $a' \in W$ такое, что $f(a') = b$, и f не является инъекцией.

б) если f не является сюръекцией, то существует $b \in B$ такое, что оно не принадлежит $f(A)$. Следовательно, $ff^{-1}(A \cup \{b\}) \neq A \cup \{b\}$. С другой стороны, предположим, что f — сюръекция. По определению функции f^{-1} соотношение $ff^{-1}(W) \subseteq W$ выполняется всегда. Пусть $w \in W$. Поскольку f — сюръекция, то существует $a \in A$ такое, что $f(a) = w$, и при этом $a \in f^{-1}(W)$. Следовательно, $f(a) \in ff^{-1}(W)$, поэтому $w \in ff^{-1}(W)$ и $W \subseteq ff^{-1}(W)$. Таким образом, $W = ff^{-1}(W)$.

11. Пусть $a \in A$ и $f(a) = a'$. $I(f(a)) = I(a') = a' = f(a) = f(I(a))$. Если f имеет обратную функцию, то $f^{-1}f(a) = f^{-1}(a') = a = I(a)$. Пусть $a \in A$ и $a'' = f^{-1}(a)$, тогда $ff^{-1}(a) = f(a'') = a = I(a)$.

Раздел 4.3.

1. **а)** $f \circ g = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$;
$g \circ f = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$;
$f^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$;
$g^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$;

б) $f \circ g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$;
$g \circ f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$;
$f^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$;
$g^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$;

в) $f \circ g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$;
$g \circ f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$;
$f^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$;
$g^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$;

г) $f \circ g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$;
$g \circ f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$;
$f^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$;
$g^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$.

3. **а)** 14; **б)** 630630; **в)** -9; **г)** -2;
д) -2; **е)** -3; **ж)** 5.

5. **а)** $6, 11, 18, 2738$; **б)** $2, 3, 2, \frac{5}{6}, \frac{1}{4}$;
в) $-4, -2, -1, -1, 0$; **г)** $1, -\frac{1}{2}\frac{1}{3}, -\frac{1}{4}, \frac{1}{5}$.

7. **а)** $A(n) = 2n^2$; **б)** $A(n) = \frac{1}{2}(-\frac{1}{3})^{n-1}$;
в) $A(n) = (-1)^{n+1}\left\lfloor \frac{n+1}{2} \right\rfloor$; **г)** $A(n) = n^2 + 2$.

Раздел 4.3.

1. а) $\begin{bmatrix} 1 & -4 & -3 & 12 \\ 39 & 5 & 21 & 8 \\ -10 & 5 & 0 & -20 \end{bmatrix}$; **б)** $\begin{bmatrix} -3 & 5 & 15 \\ 11 & 23 & 41 \\ 7 & 13 & 67 \end{bmatrix}$;

в) $\begin{bmatrix} -2 & 3 & -5 & 0 \\ -6 & 9 & -15 & 0 \\ -8 & 12 & -20 & 0 \\ -10 & 15 & -25 & 0 \end{bmatrix}$; **г)** $\begin{bmatrix} -13 \end{bmatrix}$;

д) $\begin{bmatrix} 10 & -2 \\ -12 & 14 \end{bmatrix}$; **е)** $\begin{bmatrix} -25 & -29 \\ 66 & -1 \end{bmatrix}$.

3. а) $\begin{bmatrix} -16 & 15 \\ 18 & 42 \end{bmatrix}$; **б)** $\begin{bmatrix} -5 & 13 \\ 69 & 41 \end{bmatrix}$; **в)** $\begin{bmatrix} -10 & 24 \\ 33 & 31 \end{bmatrix}$; **г)** $\begin{bmatrix} -26 & 60 \\ -2 & 47 \end{bmatrix}$;

д) $\begin{bmatrix} -16 & 18 \\ 15 & 42 \end{bmatrix}$; **е)** $\begin{bmatrix} -16 & 18 \\ 15 & 42 \end{bmatrix}$; **ж)** $\begin{bmatrix} -7 & 1 \\ 9 & -4 \end{bmatrix}$; **з)** $\begin{bmatrix} 7 & -1 \\ -9 & 4 \end{bmatrix}$;

и) $\begin{bmatrix} -15 & 10 \\ 35 & 25 \end{bmatrix}$; **к)** $\begin{bmatrix} 17 & -4 \\ -25 & 3 \end{bmatrix}$.

5. Пусть $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ и $C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Тогда $AB = AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, но $B \neq C$.

7. Предположим, что A и B — матрицы, и их произведение AB определено. Пусть $C = AB$, так что $C_{ij} = \sum A_{ik}B_{kj}$. При этом, в матрице $(AB)^t$ элемент $C^t_{ji} = \sum A_{ik}B_{kj}$ или $C^t_{ij} = \sum A_{jk}B_{ki}$. Пусть $\hat{C} = B^tA^t$, тогда $\hat{C}_{ij} = \sum B_{ki}A_{jk} = \sum A_{jk}B_{ki}$. Следовательно, $B^tA^t = (AB)^t$.

9. Пусть R и S — отношения с матричными представлениями M и N, соответственно. Пусть $A = M \wedge N$. Если $(a_i, a_j), \in R \cap S$, то $(a_i, a_j), \in R$ и $(a_i, a_j), \in S$. Поэтому $M_{ij} = N_{ij}, = 1$, так что $A_{ij} = M_{ij} \wedge N_{ij} = 1$. Следовательно, если $A_{ij} = M_{ij} \wedge N_{ij} = 1$, то $M_{ij} = N_{ij} = 1$, так что $(a_i, a_j), \in R$ и $(a_i, a_j), \in S$. Следовательно, $(a_i, a_j), \in R \cap S$. Поэтому $M_{ij} \wedge N_{ij}$ является матричным представлением для $R \cap S$.

11. $R = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$; $S = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$; $T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$.

а) $R^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$;

$S^{-1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$;

б) $S \circ R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$;

в) $S \circ S^{-1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$;

г) $S^{-1} \circ S = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$;

д) $R^{-1} \circ S^{-1} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$; **е)** $T \circ (S \circ R) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$;

ж) $T \circ S = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$; **з)** $(T \circ S) \circ R = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.

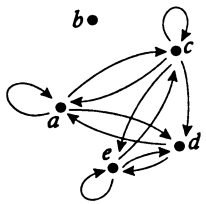
13. $S = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$; $T = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$;

$U = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$; $V = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

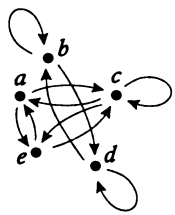
а) U, V; **б)** U; **в)** U; **г)** S.

15. а) $\{(1,a),(1,c),(2,b),(3,c),(4,a),(4,b)\}$;
б) $\{(1,b),(1,c),(2,a),(2,b),(3,a),(3,c),(4,a),(4,c)\}$;
в) $\{(1,a),(1,c),(2,b),(3,a),(3,c),(4,b)\}$.

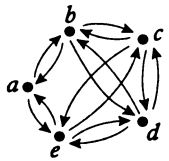
17. а)



б)



в)



19. а) $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$; **б)** $\begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$; **в)** $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

Раздел 4.4.

1. а) 7; **б)** 11; **в)** счетно бесконечное.

3.

1	2	3	4	5	...
↓	↓	↓	↓	↓	
−1	−2	−3	−4	−5	...

$a(n) = -n$

5.

1	2	3	4	5	...	11	12	...
↓	↓	↓	↓	↓		↓	↓	
−10	−9	−8	−7	−6	...	0	1	...

$a(n) = n - 11$

Раздел 4.5.

1. $\{2, 4\}$.

3. $\emptyset$.

5. Согласно теореме 4.51 множество положительных рациональных чисел счетно бесконечное. Поскольку существует взаимно однозначное соответствие между положительными рациональными числами и отрицательными рациональными числами, то множество отрицательных рациональных чисел счетно. Поэтому по теореме 4.52 их объединение счетно. Поскольку множество $\{0\}$ счетно, то его объединение с ненулевыми рациональными числами счетно. Поэтому множество рациональных чисел счетно.

7. Используя индукцию, покажем сначала, что если A счетно, то $A \times A \times A \times \cdots \times A$, декартово произведение n копий множества A, счетно. Существует взаимно однозначное соответствие между многочленами пятой степени с целыми коэффициентами и $Z \times Z \times Z \times Z \times Z \times Z$, где $a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$ соответствует $(a_5, a_4, a_3, a_2, a_1, a_0)$. Следовательно, множество многочленов пятой степени с целыми коэффициентами счетно.

Раздел 5.1.

1. Процедура **Транспонирование матрицы** (A, m, n):

```
Цикл по i от 1 до m:
    Цикл по j от 1 до n:
        T_ji = A_ij;
    Конец цикла;
Конец цикла;
Конец процедуры.
```

3. а) $\begin{bmatrix} -2 & 1 & 2 \\ 4 & 1 & -4 \\ 3 & 7 & 1 \end{bmatrix}$; б) $\begin{bmatrix} 29 & 23 & -17 \\ 23 & 51 & 5 \\ -17 & 5 & 11 \end{bmatrix}$; в) $\begin{bmatrix} 9 & -15 & 3 \\ -15 & 33 & 15 \\ 3 & 15 & 59 \end{bmatrix}$.

5. а) $\begin{bmatrix} -31 & 2 \\ -14 & -12 \end{bmatrix}$; б) $\begin{bmatrix} 3 & -8 & 5 \\ 11 & -16 & -15 \\ 6 & 0 & -30 \end{bmatrix}$; в) $\begin{bmatrix} 1 & 2 \\ 0 & -4 \\ -5 & 0 \end{bmatrix}$;

г) $\begin{bmatrix} -1 & 3 & 6 \\ 2 & 4 & 0 \end{bmatrix}$; д) $\begin{bmatrix} 3 & 11 & 6 \\ -8 & -16 & 0 \\ 5 & -15 & -30 \end{bmatrix}$; е) $\begin{bmatrix} 3 & 11 & 6 \\ -8 & -16 & 0 \\ 5 & -15 & -30 \end{bmatrix}$.

7. Процедура **Наименьшее целое число** $(a_1, a_2, a_3, \ldots, a_n)$ (находит наименьшее значение a_k):

```
Положить k = 1;
Цикл по i от 2 до n;
    Если a_i ≤ a_k, то положить k = i;
Конец цикла;
Конец процедуры.
```

9. Процедура **Проверка симметричности матрицы** (A, n):

```
Положить T = 1;
Цикл по i от 1 до n:
    Цикл по j от i до n:
        Если A_ij ≠ A_ji, то положить T = 0;
    Конец цикла;
Конец цикла;
Конец процедуры.
```

11. Процедура **Среднее арифметическое** $(a_1, a_2, a_3, \ldots, a_n)$:
Положить $S = 0$;
Цикл по i от 1 до n:
Заменить S на $S + a_i$;
Конец цикла;
Положить $A = \dfrac{S}{n}$;
Конец процедуры.

13. Следующая процедура вычитает b из a, где b имеет размерность m, a имеет размерность n, $m \le n$ и $b \le a$. Предположим, что $b = b_1 b_2 b_3 \cdots b_n$ и $a = a_1 a_2 a_3 \cdots a_n$, где $b_i = 0$ для $i < n - m$. Кроме того, предположим, что у нас имеется процедура **Сложение** (a, b), построенная в предыдущей задаче, которая складывает числа a и b. Пусть $b^{(k)} = b_1 b_2 b_3 \cdots b_k$.
Процедура **Вычитание (a,b,m,n)**:
Цикл по i от 1 до n:
Условие: если $b_{n-i+1} \le a_{n-i+1}$, то положить $c_{n-i+1} = a_{n-i+1} - b_{n-i+1}$;
Условие: если $b_{n-i+1} > a_{n-i+1}$, то
положить $c_{n-i+1} = 10 + a_{n-i+1} - b_{n-i+1}$;
заменить $b^{(n-i)}$ на **Сложение** $(b^{(n-i)}, 1)$;
Конец цикла;
Конец процедуры.

15. Следующая процедура делит a на b, определяя частное q и остаток r.
Процедура **Деление (a,b)**:
Если $b > a$, положить $q = 0$ и $r = a$;
Положить $i = 1$;
Условный цикл: до тех пор, пока $r > b$,
положить $r = a - (i \cdot b)$;
положить $q = i$;
положить $i = i + 1$;
Конец условного цикла;
Конец процедуры.

Раздел 5.2.

1. **а)** $f(1) = 9$ $f(2) = 27$ $f(3) = 81$ $f(4) = 243$;
б) $f(1) = 5$ $f(2) = 16$ $f(3) = 41$ $f(4) = 94$;
в) $f(1) = 4$ $f(2) = 16$ $f(3) = 256$ $f(4) = 65536$;
г) $f(1) = 1$ $f(2) = 4$ $f(3) = 36$ $f(4) = 576$;
д) $f(1) = 2$ $f(2) = 4$ $f(3) = 16$ $f(4) = 65536$.

3. **а)** $f(2) = 5$ $f(3) = 7$ $f(4) = 9$ $f(5) = 11$;
б) $f(2) = 1$ $f(3) = 0$ $f(4) = -1$ $f(5) = 1$;
в) $f(2) = 7$ $f(3) = 56$ $f(4) = 3145$ $f(5) = 9890994$;
г) $f(2) = 2$ $f(3) = 0$ $f(4) = -48$ $f(5) = -5760$;
д) $f(2) = -1$ $f(3) = -1$ $f(4) = 1$ $f(5) = -1$.

5. **а)** $f(n) = 2^n$; **б)** $f(n) = n + 1$; **в)** $f(n) = \dfrac{1}{n!}$;
г) $f(n) = 1 + 2n$; **д)** $f(n) = 5^n$.

7. $a_0 = 7 - 2^{0+1} = 5$. Требуется показать, что $7 - 2^{k+1} = 2(7 - 2^k) - 7$. Подставляя в рекурсивную функцию, имеем

$$2(7 - 2^k) - 7 = 7 - 2 \cdot 2^k = 7 - 2^{k+1}.$$

9. $a_0 = -1 - 2^{0+1} = -3$. $a_1 = -1 - 2^{1+1} = -5$. Подставляя в рекурсивную функцию, имеем

$$-1 - 2^{k+1} = 6(-1 - 2^k) - 8(-1 - 2^{k-1}) - 3.$$

Но

$$6(-1-2^k)-8(-1-2^{k-1})-3 = -1-6\cdot 2^k+8\cdot 2^{k-1} = -1-3\cdot 2^{k+1}+2\cdot 2^{k+1} = -1-2^{k+1}.$$

11. $a_1 = 1^2 + 1 + 1 = 3$. Подставляя в рекурсивную функцию, имеем

$$(k-1)^2 + k - 1 + 1 + 2k = k^2 - 2k + 1 + 3k = k^2 + k + 1.$$

13. $a_0 = -2(-1)^0 + 2\cdot 3^0 = 0$, $a_1 = -2(-1)^1 + 2\cdot 3^1 = 8$. Подставляя в рекурсивную функцию, имеем

$$\begin{aligned}2(-2(-1)^{k-1} + 2\cdot 3^{k-1}) + 3(-2(-1)^{k-2} + 2\cdot 3^{k-2}) &= \\ = -4(-1)^{k-1} - 6(-1)^{k-2} + 4\cdot 3^{k-1} + 6\cdot 3^{k-2} &= \\ = 4(-1)^k - 6(-1)^k + 4\cdot 3^{k-1} + 2\cdot 3^{k-1} &= -2(-1)^k + 2\cdot 3^k.\end{aligned}$$

15. $a_0 = 3(-2)^0 + 2\cdot 3^0 - 3\cdot 2^0 = 2$, $a_1 = 3(-2)^1 + 2\cdot 3^1 - 3\cdot 2^1 = -6$. Подставляя в рекурсивную функцию, имеем

$$\begin{aligned}3(-2)^k + 2\cdot 3^k - 3\cdot 2^k = 3(-2)^{k-1} + 2\cdot 3^{k-1} - 3\cdot 2^{k-1} + \\ + 6(3(-2)^{k-2} + 2\cdot 3^{k-2} - 3\cdot 2^{k-2}) + 3\cdot 2^k = \\ = (3-9)(-2)^{k-1} + (2+4)3^{k-1} + (-3-9)2^{k-1} + 3\cdot 2^k = \\ = 3(-2)^k + 2\cdot 3^k - 6\cdot 2^k + 3\cdot 2^k = 3(-2)^k + 2\cdot 3^k - 3\cdot 2^k.\end{aligned}$$

17. $\approx 5.8 \times 10^{11}$ лет.

Раздел 5.3.

1. $2^8 - 1 = 255$.

3. **а)** (а) 5 (б) 4; **б)** (а) 14 (б) 8; **в)** (а) 18 (б) 9.

5. (б), (в), (г).

7. Для $n = 1$ имеем $1! = 1 = 1^1$, поэтому утверждение истинно. Предположим, что утверждение истинно для $n = k$, т.е. $k! \le k^k \le (k+1)^k$ поскольку разложение $(k+1)^k$ включает k^k как один из множителей. Умножая обе части неравенства $k! \le (k+1)^k$ на $(k+1)$, получаем $(k+1)! \le (k+1)^{k+1}$.

9. Если $f(n) = O(g(n))$, то существуют k_1, m_1 такие, что $f(n) \le k_1 g(n)$ для $n > m_1$. Если $g(n) = O(h(n))$, то существуют k_2, m_2 такие, что $g(n) \le k_2 h(n)$ для $n > m_2$. Полагая $M = \max(m_1, m_2)$, получаем $f(n) \le k_1 g(n) \le k_1 k_2 h(n)$ для $n > M$, поэтому $f(n) = O(h(n))$.

11. Пусть $P(m)$ является утверждением "Если r и m — положительные целые числа, $r \le m$ и $n > 1$, то $n^r \le n^m$". Утверждение несомненно истинно, если $m = 1$. Предположим, что $P(k)$ истинно, так что если $r \le k$, то $n^r \le n^k$. Нужно доказать, что $P(k)$ истинно, так что если $r \le k+1$, то $n^r \le n^{k+1}$. Поскольку $1 < n$, то $n^k \cdot 1 \le n^k \cdot n$. Следовательно, $n^k \le n^{k+1}$. Предположим, что $r \le k+1$. Если $r < k+1$, то $r \le k$, поэтому $n^r \le n^k \le n^{k+1}$. Если $r = k+1$, то $n^r \le n^{k+1}$. Таким образом, $P(k+1)$ истинно.

Раздел 5.4.

1. $0, 1, 2, 3, 4, 5, 7, 8, 9, 11$.

3. $a, b, c, d, f, j, m, n, p, r, s, t, u, x, y, z$.

5. **а)** $9, 8, 7, 6, 5, 4, 3, 2, 1$ (для n объектов количество сравнений всегда одинаково);
б) $9, 8, 7, 6, 5, 4, 3, 2, 1$ (для n объектов количество сравнений всегда одинаково);
в) $1, 2, 3, 4, 5, 6, 7, 8, 9$;
г) $5, 3, 1, 2, 4, 3, 7$;
д) $1, 2, 3, 4, 5, 6, 7, 8, 9$.

7. $2D(\frac{n}{2})(\log_2(\frac{n}{2}) + A) + Dn = Dn(\log_2 n - \log_2 2 + A) + Dn =$

$$= Dn(\log_2 n - 1 + A) + Dn = Dn(\log_2 n + A).$$

9.
$$\frac{Q_{2^m}}{2^m} = \frac{cQ_{2^{m-1}}}{2 \cdot 2^{m-1}} + D =$$ рекурсивная формула, деленная на $n = 2^m$

$$= \left(\frac{c}{2}\right)^2 \frac{Q_{2^{m-2}}}{2^{m-2}} + \frac{c}{2}D + D =$$ повторное применение рекурсивной формулы

$$= \left(\frac{c}{2}\right)^3 \frac{Q_{2^{m-3}}}{2^{m-3}} + \left(\left(\frac{c}{2}\right)^2 + \frac{c}{2} + 1\right) D =$$ повторное применение рекурсивной формулы

$$\vdots$$

$$= \left(\frac{c}{2}\right)^m Q_1 + \left(\left(\frac{c}{2}\right)^{m-1} + \cdots + \left(\frac{c}{2}\right)^2 + \frac{c}{2} + 1\right) D =$$ повторное применение рекурсивной формулы

$$= \frac{c^{\log_2 n}}{2^{\log_2 n}} Q_1 + \left(\left(\frac{c}{2}\right)^{m-1} + \cdots + \left(\frac{c}{2}\right)^2 + \frac{c}{2} + 1\right) D =$$ $m = \log_2 n$, поскольку $n = 2^m$

$$= \frac{c^{\log_2 n}}{n} Q_1 + \left(\frac{1 - \left(\frac{c}{2}\right)^m}{1 - \frac{c}{2}}\right) D =$$ сумма геометрической прогрессии

$$= \frac{n^{\log_2 c}}{n} Q_1 + \left(\frac{1 - \left(\frac{c}{2}\right)^m}{1 - \frac{c}{2}}\right) D =$$ поскольку $c^{\log_2 n} = n^{\log_2 c}$ (возьмите $\log_2$ от обеих частей)

$$= \frac{n^{\log_2 c}}{n}\left(A + \frac{2D}{2 - c}\right) + \left(\frac{1 - \frac{n^{\log_2 c}}{n}}{1 - \frac{c}{2}}\right) D =$$ подставьте значение Q_1

$$= \frac{n^{\log_2 c}}{n} A + \frac{2D}{2 - c}\frac{n^{\log_2 c}}{n} + \frac{2D}{2 - c} - \frac{2D}{2 - c}\frac{n^{\log_2 c}}{n} = \frac{n^{\log_2 c}}{n} A + \frac{2D}{2 - c}.$$

Следовательно, $\frac{Q_n}{n} = \frac{n^{\log_2 c}}{n} A + \frac{2D}{2 - c}$ и $Q_n = n^{\log_2 c} A + \frac{2D}{2 - c} n$.

Раздел 5.5.

1. **а)** $a2\hat{}b2\hat{}+$; **б)** $a3\hat{}3a2\hat{} \times +7+$;
в) $a2\hat{}b + cd2\hat{}+$; **г)** $a2\hat{}b3\hat{} + 4\hat{}$;
д) $a2\hat{}bc + \times$.

3. **а)** $(a + b) \div (c - d)$, **б)** $(a + b\hat{}2) \times (c + d\hat{}2)$;
в) $3 \times a \times b + 4 \times c \times d\hat{}2$; **г)** $(a + b)\hat{}2 \div (c + d)$;
д) $(a\hat{}2 + b) \times (c + d\hat{}2)$.

5. **а)** $+\hat{}a2\hat{}b2$; **б)** $+ + \hat{}a3 \times 3\hat{}a27$;
в) $\times + \hat{}a2b + c\hat{}d2$; **г)** $\hat{} + \hat{}a2\hat{}b34$;
д) $\times\hat{}a2 + bc$.

7. **а)** $(a + b) \div (c - d)$ **б)** $(a\hat{}2 + d\hat{}2) \div (c\hat{}2 + d\hat{}2)$;
в) $3ab + 4cd + 5ad$; **г)** $(a\hat{}2 + b\hat{}2 + c\hat{}2)\hat{}3$;
д) $(a\hat{}2 + b)(c + d\hat{}2)$.

9.
1. Добавлять скобки в выражение, пока оно не станет полноскобочным.
2. Начиная с самых внутренних скобок, удалить пару скобок и поместить находящийся внутри скобок оператор на место соответствовавшей ему левой скобки. При наличии более чем одной самой внутренней пары скобок начинать с правой пары.
3. Перемещаясь к следующей паре скобок, снова удалить их и поместить оператор, находившийся внутри скобок, на место соответствовавшей ему левой скобки.
4. Продолжать шаг (3), пока все скобки не будут удалены.

Раздел 5.6.

1. 1, 10, 11, 100, 101, 110, 111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111
10000, 10001, 10010, 10011, 10100, 10101, 1011010111.

3. **а)** 101101; **б)** 1010110; **в)** 11110011;
г) 11000010; **д)** 100111000.

5. **а)** 11; **б)** 26; **в)** 36; **г)** 59; **д)** 53.

7. **а)** $3F$; **б)** 74; **в)** 178; **г)** $2DE$; **д)** $147E$.

9. **а)** 74; **б)** 755; **в)** 1274; **г)** 16893; **д)** 53731.

11. **а)** 10111.111; **б)** 101011.11; **в)** 10111010.001;
г) $1011110.\overline{1100}$; **д)** $110100.01\overline{0011}$.

13. **а)** 14.75; **б)** 10.625; **в)** 4.5625;
г) 3.4375; **д)** 13.3125.

15. **а)** $24.A$; **б)** $A1.C8$; **в)** 250.28;
г) $754.8BD7$; **д)** $147E.59$.

17. **а)** 92.2265625; **б)** 3858.6796875;
в) 1199.767578125; **г)** 12971.75415039;
д) 5070.821289062.

19. **а)** $5.A$; **б)** 37.54; **в)** $1B6.6F8$;
г) $3AE.36C$; **д)** $3D4.228$.

21. **а)** 101100.01001011; **б)** 111110010.101000111100;
в) 101011111000.010010001100; **г)** 1010000100101110.000100011100;
д) 1100000101001010.001011010100.

Раздел 5.7.

1. **а)** 10001010; **б)** 01001011;
в) 11001001; **г)** 00010101.

3. **а)** 01001001; **б)** 10011011;
в) 11011011; **г)** 11110010.

5. **а)** 00100110; **б)** 01111110;
в) 10111111; **г)** 11000101.

7. **а)** 01000111; **б)** 01011101;
в) 11101000; **г)** 00000110.

9. **а)** $8DC6$; **б)** $54E0$; **в)** $05EE$; **г)** $DC44$.

11. **а)** $010E$; **б)** $FC83$; **в)** $FFEF$; **г)** $D970$.

13. **а)** $C5C0$; **б)** $F6C9$; **в)** $EF7C$; **г)** $35b3$.

15. **а)** $04AA$; **б)** $15BD$; **в)** $D4B1$; **г)** $FD31$.

17. -2147483648 to 2147483647.

Раздел 5.8.

1. **а)** -1; **б)** 22; **в)** 15; **г)** -234.

3. -1.

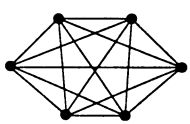
5. Если A — перестановочная матрица, то в каждой строке и в каждом столбце этой матрицы только один элемент равен 1, а все остальные равны 0. Пусть $B = AA^t$. Пусть A_i — строка матрицы A и пусть в i-ой строке $A_{ik} = 1$. Если A_j^t — j-ый столбец в транспонированной матрице A^t, то $B_{ij} = 1$ тогда и только тогда, когда $A_{kj}^t = 1$. В противном случае $B_{ij} = 0$. Но $A_{kj}^t = A_{jk}$, поэтому $B_{ij} = 1$ тогда и только тогда, когда $A_{jk} = 1$. Но поскольку имеется только одна 1 в каждой строке и столбце, то это имеет место, только если $i = j$. Следовательно, $B_{ij} = 1$ тогда и только тогда, когда $i = j$, поэтому B — единичная матрица.

Раздел 6.1.

1. **а)** путь, простой путь, длина 5; **б)** путь, не простой путь, длина 6;
в) путь, не простой путь, длина 7; **г)** путь, не простой путь, длина 8.

3. **а)** не цикл; **б)** цикл, не простой, длина 8;
в) не цикл; **г)** цикл, простой, длина 6.

5. **а)**



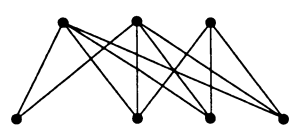
б)



в)



г)



7. Компонента представляет собой наименьший подграф такой, что если точка, изображающая палочку, принадлежит графу и палочка касается другой палочки, то вершина, изображающая вторую палочку, и ребро между ними принадлежат подграфу. Иными словами, это граф, представляющий совокупность палочек, лежащих одна поверх другой.

Раздел 6.2.

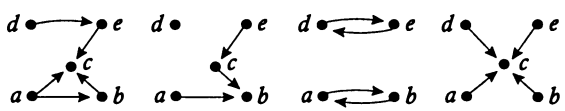
1. **а)** $V = \{a, b, c, d, e\}$. Вершина a имеет степень входа 2 и степень выхода 1, вершина b — степень входа 1 и степень выхода 1, вершина c — степень входа 2 и степень выхода 1, вершины d и e — степень входа 0 и степень выхода 1, вершины d и e — источники;
б) $V = \{a, b, c, d, e\}$. Вершины a, b, d и c имеют степень входа 0 и степень выхода 1.

Вершина e имеет степень входа 4 и степень выхода 0. Вершины a,b,d и c — источники. Вершина e — сток;

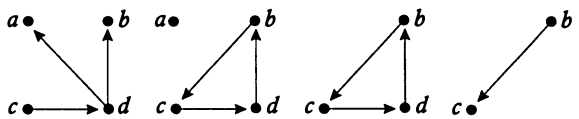
в) $V = \{a, b, c, d, e, f\}$. Вершина a имеет степень входа 0 и степень выхода 2. Вершины b, c и d имеют степень входа 1 и степень выхода 2. Вершина e имеет степень входа 1 и степень выхода 1. Вершина f имеет степень входа 5 и степень выхода 0. Вершина a — источник, а вершина f — сток.

г) $V = \{a, b, c, d, e\}$. Вершина a имеет степень входа 2 и степень выхода 1. Вершины b, d и e имеют степень входа 1 и степень выхода 1. Вершина c имеет степень входа 1 и степень выхода 2.

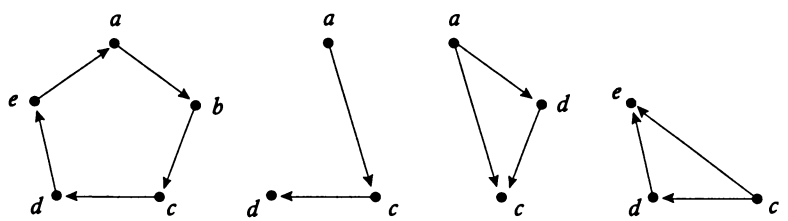
3. а)



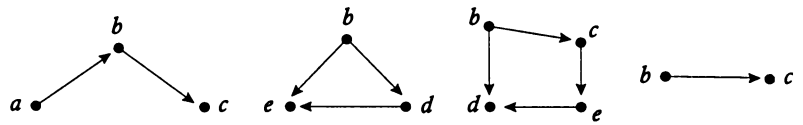
б)



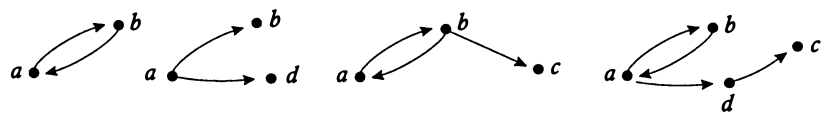
в)



г)

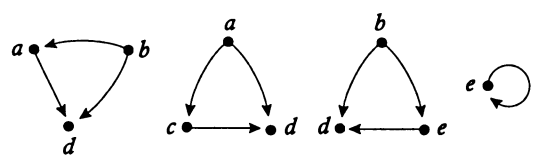


5. а)

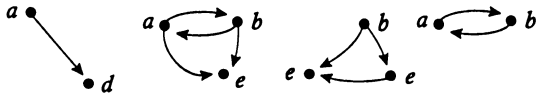


б) совпадают с построенными в (а)

в)



г)



7. Ориентированным циклом называется ориентированный путь, который начинается и заканчивается в одной и той же вершине.

9. Орграфы (а), (б), (г) — связные, граф (б) — сильно связный.

11. Орграфы (а), (в), (г) — связные, граф (в) — сильно связный.

Раздел 6.3.

1. (а); (б) и (д).

3. (а); (б) и (в).

5. Поскольку каждая компонента C_i — дерево, то $v_i = e_i + 1$. Следовательно, для каждого целого дерева имеем $\sum_{i=1}^{m} v_i = \sum_{i=1}^{m}(e_i + 1)$, поэтому $v = e + m$.

7. **а)** v_6, v_7, v_8, v_9; **б)** v_0, v_3, v_6; **в)** v_1;
г) 2; **д)** v_6, v_7; **е)** 3;
ж) $v_2, v_4, v_5, v_7, v_8, v_9$; **з)** нет, v_0 имеет степень выхода 3.

9. Воспользуемся индукцией по количеству вершин. Если if $v = 1$, то $e = 0$. Таким образом, теорема верна, если $v = 1$. Предположим истинность утверждения теоремы для $v = k$. Пусть T — дерево с $k + 1$ вершинами. Удалим лист и инцидентное ему ребро. Теперь дерево имеет k вершин. Следовательно, по индуктивному предположению, для этого дерева $v = e + 1$. Если возвратить на место лист и ребро, то количество вершин и количество ребер увеличится на один. Поэтому для дерева с $k + 1$ вершинами по-прежнему будем иметь $v = e + 1$.

Раздел 6.4.

1. а)

фронтальная	задняя	левая	правая
Ж	с	к	з
с	к	з	Ж
к	Ж	с	к
з	з	Ж	с

б)

фронтальная	задняя	левая	правая
с	Ж	з	с
к	з	с	Ж
з	с	Ж	к
Ж	к	к	з

в)

фронтальная	задняя	левая	правая
з	с	к	Ж
Ж	з	з	с
с	к	Ж	з
к	Ж	с	к

г)

фронтальная	задняя	левая	правая
з	Ж	к	Ж
к	к	з	с
Ж	с	Ж	з
с	з	с	к

д) решение отсутствует.

Раздел 6.5.

1. (а).

3. (в).

5. (а), (в).

7. (б).

9. Предположим, что граф G связный и ровно две его вершины имеют нечетную степень.

Пусть a и b — вершины нечетной степени. Если между a и b ребро отсутствует, добавим его. Теперь каждая вершина имеет четную степень, поэтому новый граф имеет эйлеров цикл. В этом цикле следует проходить и ребро $\{a, b\}$. Например, проходя цикл, мы движемся от a к b. Если начать с этого ребра и пройти цикл, то легко заметить, что если ребро удалить, начать путь в вершине b и пройти его, следуя циклу, то получим эйлеров путь из b в a. Если между a и b имеется ребро, удалим его. Новый граф имеет эйлеров цикл, если он по-прежнему связный. Пусть эйлеров цикл начинается и заканчивается в вершине a. Если мы пройдем этот цикл, а затем проследуем вдоль удаленного ребра от a к b, то получим эйлеров путь от a к b. Если новый граф перестал быть связным, то он имеет эйлеров цикл для компоненты, содержащей вершину a, который начинается и заканчивается в вершине a, а также для компоненты, содержащей вершину b, который начинается и заканчивается в вершине b. Пройдем эйлеров цикл от a к a, удаленное ребро от a к b, затем пройдем эйлеров цикл от b к b. В результате получим эйлеров путь от a к b.

Предположим, что граф G имеет эйлеров путь. Пусть, например, он начинается в вершине a и заканчивается в вершине b. После первого ребра пути, выходящего из a, для каждого ребра пути, который ведет в a, должно существовать ребро, выходящее из a. Поэтому вершина a должна иметь нечетную степень. Аналогично, вершина b должна иметь нечетную степень. В любых других вершинах, для любого ребра пути, который ведет в эту вершину, должно существовать ребро, которое выходит из этой вершины. Поэтому вершина имеет четную степень.

11. Если граф сильно связный, то для любой его вершины v, несомненно, любая другая вершина достижима из вершины v, и v достижима из любой другой его вершины. Обратно, пусть имеется вершина v, обладающая указанным свойством. Пусть a и b — произвольные вершины. Поскольку существует путь из a в v и путь из v в b, то существует путь из a в b, поэтому рассматриваемый граф сильно связный.

Раздел 6.6.

1. а)

	e_1	e_2	e_3	e_4	e_5	e_6
v_1	1	0	0	1	1	0
v_2	1	1	0	0	0	0
v_3	0	1	1	0	0	0
v_4	0	0	0	0	1	1
v_5	0	0	1	1	0	1

б)

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	e_{12}
v_1	1	1	0	0	0	0	1	0	0	0	0	0
v_2	1	0	0	1	1	0	0	0	0	0	0	0
v_3	0	0	0	0	0	0	1	0	1	1	0	0
v_4	0	0	0	0	1	0	0	0	0	1	1	0
v_5	0	1	1	0	0	0	0	1	0	0	0	0
v_6	0	0	1	1	0	1	0	0	0	0	0	0
v_7	0	0	0	0	0	0	0	1	1	0	0	1
v_8	0	0	0	0	0	1	0	0	0	0	1	1

в)

	e_1	e_2	e_3	e_4	e_5	e_6
v_1	1	0	0	1	0	0
v_2	1	1	1	0	0	0
v_3	0	1	0	0	1	1
v_4	0	0	1	1	1	0
v_5	0	0	0	0	0	1

г)

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}
v_1	0	0	1	1	0	0	0	1	1	0
v_2	1	1	0	0	0	0	0	0	1	1
v_3	0	0	1	0	1	1	0	0	0	1
v_4	0	1	0	1	0	1	1	0	0	0
v_5	1	0	0	0	1	0	1	1	0	0

3. а)

	v_1	v_2	v_3	v_4	v_5
v_1	0	1	0	1	1
v_2	1	0	1	0	0
v_3	0	1	0	0	1
v_4	1	0	0	0	1
v_5	1	0	1	1	0

б)

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
v_1	0	1	1	0	1	0	0	0
v_2	1	0	0	1	0	1	0	0
v_3	1	0	0	1	0	0	1	0
v_4	0	1	1	0	0	0	0	1
v_5	1	0	0	0	0	1	1	0
v_6	0	1	0	0	1	0	0	1
v_7	0	0	1	0	1	0	0	1
v_8	0	0	0	1	0	1	1	0

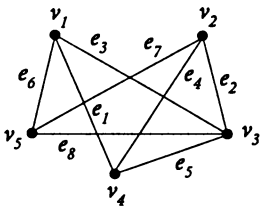
в)

	v_1	v_2	v_3	v_4	v_5
v_1	0	1	0	1	0
v_2	1	0	1	1	0
v_3	0	1	0	1	1
v_4	1	1	1	0	0
v_5	0	0	1	0	0

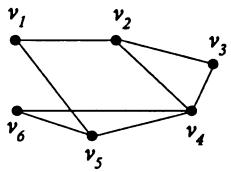
г)

	v_1	v_2	v_3	v_4	v_5
v_1	0	1	1	1	1
v_2	1	0	1	1	1
v_3	1	1	0	1	1
v_4	1	1	1	0	1
v_5	1	1	1	1	0

5.



7.



9. $A =$

	v_1	v_2	v_3	v_4	v_5
v_1	0	1	1	0	0
v_2	1	0	1	0	0
v_3	1	1	0	0	0
v_4	0	0	0	0	1
v_5	0	0	0	1	0

$A^2 =$

	v_1	v_2	v_3	v_4	v_5
v_1	1	1	1	0	0
v_2	1	1	1	0	0
v_3	1	1	1	0	0
v_4	0	0	0	1	0
v_5	0	0	0	0	1

2-реберные из v_1 в v_1, v_2, v_3, из v_2 в v_1, v_2, v_3, из v_3 в v_1, v_2, v_3, из v_4 в v_4, из v_5 в v_5.

11. а) $\begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$; **б)** $\begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$; **в)** $\begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$.

13. $R = \{(v_1, v_1), (v_1, v_2), (v_1, v_3), (v_2, v_1), (v_2, v_2), (v_2, v_3), (v_3, v_1), (v_3, v_2), (v_3, v_3), (v_4, v_4), (v_5, v_5), (v_4, v_5), (v_5, v_4)\}$.

15. $A \times A$.

17. $R = \{(v_1, v_1), (v_1, v_2), (v_1, v_3), (v_2, v_1), (v_2, v_2), (v_2, v_3), (v_3, v_1), (v_3, v_2), (v_3, v_3), (v_4, v_4), (v_5, v_5), (v_4, v_5), (v_5, v_4)\}$.

19. $A \times A$.

21. $R = \{(v_1, v_3), (v_1, v_2), (v_4, v_4), (v_4, v_5), (v_5, v_4), (v_5, v_5)\}$.

23. Ребрам соответствуют одинаковые столбцы.

25. По определению матрицы смежности, теорема верна, когда $n = 1$. Предположим, что она верна для $n = k$. Следовательно, k-путь из вершины v_i в вершину v_j существует тогда и

только тогда, когда $A_{ij}^{\odot k} = 1$. Теперь,

$$A_{ij}^{\odot k+1} = \bigvee_{m=1}^{n} (A_{im}^{\odot k} \wedge A_{mj}) = 1$$

тогда и только тогда, когда существует m, так что $A_{im}^{\odot k} = 1$ и $A_{mj} = 1$. Но, по индуктивному предположению, это имеет место тогда и только тогда, когда существует k-путь из i в m и 1-путь из m в j. Последнее утверждение истинно тогда и только тогда, когда существует $k+1$-путь из i в j.

27. Путь из вершины v_i в вершину v_j существует тогда и только тогда, когда существует k-путь из v_i в v_j для некоторого $1 \leq k \leq m$, т.е. тогда и только тогда, когда $A_{ij}^{\odot k} = 1$ для некоторого $1 \leq k \leq m$, или тогда и только тогда, когда $\hat{A}_{ij} = 1$.

29. Если дерево $T(V, E)$ не является ориентированным, то транзитивным замыканием является $V \times V$. Если $T(V, E)$ — ориентированное дерево, то aRb, если a — предок b.

Раздел 6.7.

1. **а)** 1; **б)** 2; **в)** 2; **г)** 4.

3.

1	1	1	1	1	0	0	1	1	1
1	1	1	1	0	0	1	1	1	0
1	1	1	0	0	0	0	1	0	0
1	1	1	0	1	0	0	1	0	1
1	1	0	0	1	0	0	0	0	1
1	1	0	0	0	0	0	0	0	0
1	1	0	1	0	0	0	0	1	0
1	1	0	1	1	0	0	0	1	1
1	0	0	1	1	0	1	0	1	1
1	0	0	1	0	0	1	0	1	0
1	0	0	0	0	0	1	0	0	0
1	0	0	0	1	0	1	0	0	1
1	0	1	0	1	0	1	1	0	1
1	0	1	0	0	0	1	1	0	0
1	0	1	1	0	0	1	1	1	0
1	0	1	1	1	0	1	1	1	1

5.

	111	110	100	101	001
11	11111	11110	11100	11101	11001
10	10111	10110	10100	10101	10001
00	00111	00110	00100	00101	00001
01	01111	01110	01100	01101	01001

7.

	111	110	100	101	001	000
111	111111	111110	111100	111101	111001	111000
110	110111	110110	110100	110101	110001	110000
100	100111	100110	100100	100101	100001	100000
101	101111	101110	101100	101101	101001	101000
001	001111	001110	001100	001101	001001	001000

9. Поскольку $m = 2^i$, первые m цифр вершин в каждом столбце образуют код Грея для m. Следовательно, первые m цифр первой вершины и последней вершины в столбце отличаются ровно одной цифрой. Последние n цифр вершин в каждом столбце совпадают. Следовательно, первая и последняя вершины в столбце отличаются ровно одной цифрой, поэтому они являются смежными. Поскольку $n = 2^j$, последние n цифр вершин в каждой

строке образуют код Грея для n. Таким образом, последние n цифр первой вершины и последней вершины в строке отличаются ровно одной цифрой. Первые m цифр вершин в каждой строке совпадают. Вполне очевидно, что первая и последняя вершины в строке отличаются ровно одной цифрой, поэтому они являются смежными.

Раздел 7.1.

1. 1, 2, 3, **4**, 5, **6**, 7, **8**, **9**, **10**, 11, **12**, 13, **14**, **15**, **16**, 17, **18**, 19, **20**, **21**, **22**, 23, **24**, **25**, **26**, **27**, **28**, 29, **30**, 31, **32**, **33**, **34**, **35**, **36**, 37, **38**, **39**, **40**, 41, **42**, 43, **44**, **45**, **46**, 47, **48**, **49**, **50**, **51**, **52**, 53, **54**, **55**, **56**, **57**, **58**, 59, **60**, 61, **62**, **63**, **64**, **65**, **66**, 67, **68**, **69**, **70**, 71, **72**, 73, **74**, **75**, **76**, **77**, **78**, 79, **80**, **81**, **82**, 83, **84**, **85**, **86**, **87**, **88**, 89, **90**, **91**, **92**, **93**, **94**, **95**, **96**, 97, **98**, **99**, **100**, 101, **102**, 103, **104**, **105**, **106**, 107, **108**, 109, **110**, **111**, **112**, 113, **114**, **115**, **116**, **117**, **118**, **119**, **120**, **121**, **122**, **123**, **124**, **125**, **126**, 127, **128**, **129**, **130**, 131, **132**, **133**, **134**, **135**, **136**, 137, **138**, 139, **140**, **141**, **142**, **143**, **144**, **145**, **146**, **147**, **148**, 149, **150**, 151, **152**, **153**, **154**, **155**, **156**, 157, **158**, **159**, **160**, **161**, **162**, 163, **164**, **165**, **166**, 167, **168**, **169**, **170**, **171**, **172**, 173, **174**, **175**, **176**, **177**, **178**, 179, **180**, 181, **182**, **183**, **184**, **185**, **186**, **187**, **188**, **189**, **190**, 191, **192**, 193, **194**, **195**, **196**, 197, **198**, 199, **200**.

3. $37 \cdot 13$.

5. $37 \cdot 47$.

7. $47 \cdot 83$.

Раздел 7.2.

1. $32^2 = 45^2 - 1001$, поэтому 1001 не является простым числом.

3. $70^2 - 7^2 = 4851$, поэтому 4851 не является простым числом.

5. $90^2 - 7^2 = 8051$, поэтому 8051 не является простым числом.

7. $90^2 - 13^2 = 7931$, поэтому 7931 не является простым числом.

Раздел 7.3.

1. **а)** $75 = 8 \times 9 + 3$; **б)** $102 = 5 \times 20 + 3$; **в)** $81 = 9 \times 9 + 0$; **г)** $16 = 25 \times 0 + 16$.

3. **а)** 75; **б)** 54; **в)** 11799; **г)** 271377; **д)** 179196.

5. Если $ax + by = \text{НОД}(a, b)$, то $\frac{a}{\text{НОД}(a,b)}x + \frac{b}{\text{НОД}(a,b)}y = 1$. Любой делитель чисел x и y должен также делить 1. Поэтому числа x и y — взаимно простые.

7. Процедура **Алгоритм деления** (a, b):
Положить $q = 0$ и $r = a$;
Условный цикл: до тех пор, пока $r \geq b$,
Положить $q = q + 1$ и $r = r - b$;
Конец условного цикла;
Конец процедуры.

9. Допустим, что существуют такие r и s, что $rs = a$ и $\text{НОД}(r, s) = b$. Тогда $r = ub$ и $s = vb$, поскольку $\text{НОД}(r, s)$ делит как r, так и s. Таким образом, $a = ubvb = uvb^2$ и $b^2 \mid a$. Обратно, если $b^2 \mid a$, то $ub^2 = a$ для некоторого целого u. Положим $r = ub$ и $s = b$. Тогда $rs = a$ и $\text{НОД}(r, s) = b$.

11. $\text{НОД}(a, b) = ax + by = a(x - b) + b(x + a)$.

13. $(n \cdot \text{НОД}(a, b)) \mid na$ и $(n \cdot \text{НОД}(a, b)) \mid nb$. Таким образом, $(n \cdot \text{НОД}(a, b)) \mid \text{НОД}(na, nb)$. $n \cdot \text{НОД}(a, b) = n(ua + vb)$ при некоторых целых a и b. $\text{НОД}(na, nb) \mid nua$ и $\text{НОД}(na, nb) \mid nvb$. Таким образом, $\text{НОД}(na, nb) \mid n(ua + vb)$, так что $\text{НОД}(na, nb) \mid n \cdot \text{НОД}(a, b)$ и $\text{НОД}(na, nb) = n \cdot \text{НОД}(a, b)$.

Раздел 7.4.

1. **а)** $[3; 2, 1, 3]$, $[3; 2, 1, 2, 1]$; **б)** $[0; 20, 1, 8, 1, 1, 2]$, $[0; 20, 1, 8, 1, 1, 1]$;
в) $[-1; 1, 6, 1, 3, 1, 5, 1, 1, 1, 2]$, $[-1; 1, 6, 1, 3, 1, 5, 1, 1, 1, 1, 1]$;

г) $[0;3,2,1,3]$, $[0;3,2,1,2,1]$;

3. а) $[2;4,4,4,4,4,2+\sqrt{5}]$;

в) $[3;7,15,1,292,1,1.736\ldots]$;

д) $[-1;1,7,1,4]$, $[-1;1,7,1,3,1]$.

б) $[1;2,2,2,2,2,1+\sqrt{2}]$;

г) $[1;1,1,1,1,1,\frac{1+\sqrt{5}}{2}]$.

Раздел 7.5.

1. Доказательство проведем индукцией. Поскольку $q_0 = 1$, то $q_0 > 0$. Предположим, что $q_m > 0$ для всех $m < k$. Но $q_k = q_{k-1}t_k + q_{k-2}$. По индуктивному предположению, q_{k-1} и q_{k-2} больше, чем 0. По определению конечной цепной дроби, t_k положительно. Поэтому q_k положительно.

3. а)

k	p_k	q_k	$x-(p_k/q_k)$
0	2	1	$0.236\cdots$
1	9	4	$-0.0139\cdots$
2	38	17	$7.73\cdots\times10^{-4}$
3	161	72	$-4.31\cdots\times10^{-5}$
4	682	305	$2.40\cdots\times10^{-6}$
5	2899	1292	$-1.33\cdots\times10^{-7}$

б)

k	p_k	q_k	$x-(p_k/q_k)$
0	1	1	$0.414\cdots$
1	3	2	$-8.57\cdots\times10^{-2}$
2	7	5	$1.42\cdots\times10^{-2}$
3	17	12	$-2.45\cdots\times10^{-3}$
4	41	29	$4.20\cdots\times10^{-4}$
5	99	70	$-7.21\cdots\times10^{-5}$

в)

k	p_k	q_k	$x-(p_k/q_k)$
0	3	1	$0.141\cdots$
1	22	7	$-1.26\cdots\times10^{-3}$
2	333	106	$8.32\cdots\times10^{-5}$
3	355	113	$-2.66\cdots\times10^{-7}$
4	103993	33102	$5.77\cdots\times10^{-10}$
5	104348	33215	$-3.31\cdots\times10^{-10}$

г)

k	p_k	q_k	$x-(p_k/q_k)$
0	1	1	$0.618\cdots$
1	2	1	$-0.381\cdots$
2	3	2	$0.118\cdots$
3	5	3	$-0.0486\cdots$
4	8	5	$0.0180\cdots$
5	13	8	$-0.00696\cdots$

5. По теореме 7.11,

$$\frac{q_k}{q_{k-1}} = \frac{q_{k-1}t_k + q_{k-2}}{q_{k-1}} = t_k + \frac{q_{k-2}}{q_{k-1}} = t_k + \frac{1}{t_{k-1} + \dfrac{q_{k-3}}{q_{k-2}}} = t_k + \frac{1}{t_{k-1} + \dfrac{1}{t_{k-2}\cdots}}.$$

7. $p_{-2} = 0$, $q_{-2} - 1$.

Раздел 8.1.

1. 139, 233, 46.

3. а) $53 \times 52 \times 51 \times 50 = 7027800$;

б) $8 \times 52 \times 51 \times 50 = 1060800$;

в) $45 \times 52 \times 51 \times 50 = 5967000$;

г) $20 \times 33 \times 32 \times 31 + 33 \times 32 \times 31 \times 30 = 1636800$.

5. $10^4 - 10 \times 9 \times 8 \times 7 = 4960$.

7. $2^{30} = 1073741824$.

9. $2 \times 26^2 + 2 \times 26^3 = 36504$.

11. а) 16; **б)** 1; **в)** 15.

13. 10^9.

15. 3^6.

17. $699 - 87 = 612$.

19. $26^2 + 26^3 = 18252$.

21. 254.

23. $2^7 - 8 = 120$.

25. $10^4 - 9^4 = 3439$.

27. 26×36^6.

Раздел 8.2.

1. $120 + 110 - 80 = 150$, 50.

3. $57 + 36 - 5 = 88$.

5. $40 + 26 - 13 = 53$.

7. $100 - 14 = 86$.

9. $2 \times 8! = 80640$.

11. **а)** 11111; **б)** 10000; **в)** 20000; **г)** 79,999.

13. $400 + 286 + 182 - 57 - 36 - 26 + 5 = 754$.

15. $333 + 286 + 250 - 47 - 83 - 35 + 11 = 715$.

17. **а)** 60; **б)** 120; **в)** 65; **г)** 200; **д)** 40.

19. 38212

Раздел 8.3.

1. **а)** 6720; **б)** 6652800; **в)** 792; **г)** 91; **д)** 91.

3. 210, 70, 120, 60.

5. $18 \cdot 17 \cdot 16 = 4896$.

7. **а)** $10! = 3628800$; **б)** $2^5 \cdot 5! = 3840$.

9. **а)** 3360; **б)** 8820; **в)** 30240; **г)** 36960; **д)** 50400.

11. 2880.

13. 387897600.

15. $\binom{52}{13}$.

17. $4 \cdot \binom{13}{7} \cdot \binom{39}{6} = 22394644272$.

19. 235237860.

21. 58656, включая все "фулл хаус", или 54912, исключая все "фулл хаус".

23. 252.

25. 210, 968.

27. 30045015.

29. 167960.

31. $a^8 + 8a^7b + 28a^6b^2 + 56a^5b^3 + 70a^4b^4 + 56a^3b^5 + 28a^2b^6 + 8ab^7 + b^8$.

33. -8646663168.

35. -262440.

Раздел 8.4.

1. 21453.

3. 21534.

5. $\{a, c, e\}$.

7. $\{2, 3, 4\}$.

9. $\{2, 3\}$.

11. $\{1, 5, 6\}$.

13. $\{1, 3, 4, 6\}$.

15. $xyz, xzy, yxz, yzx, zxy, zyx$.

17. $\{a, b, c, d\}, \{a, b, c, e\}, \{a, b, d, e\}, \{a, c, d, e\}, \{b, c, d, e\}$.

19. $abcdefghijklmnopqrstuvwxyz$, $zyxwvutsrqponmlkjihgfedcba$.

21. $\{abcde\}, \{vwxyz\}$.

23. Цикл по i от 1 до $C(n, r)$:
Использовать процедуру **Сочетание (n,r)** для образования i-го r-сочетания из n элементов;
Использовать процедуру **Обобщенная перестановка**, чтобы упорядочить все перестановки из r элементов, образованных процедурой **Сочетание (n,r)**;
Конец цикла.

Раздел 8.5.

1. а) $\frac{3}{36}$; б) $\frac{6}{36}$; в) $\frac{5}{36}$; г) $\frac{4}{36}$; д) $\frac{6}{36}$.

3. а) $\frac{12}{36}$; б) $\frac{16}{36}$; в) $\frac{30}{36}$; г) $\frac{22}{36}$.

5. а) $\frac{100}{301}$; б) $\frac{14}{301}$; в) $\frac{129}{301}$; г) $\frac{172}{301}$.

7. $\frac{15120}{59049}$.

9. а) $\frac{53}{250}$; б) $\frac{105}{250}$; в) $\frac{97}{250}$; г) $\frac{23}{250}$.

11. а) $\frac{1098240}{2598960} \approx .4226$; б) $\frac{54912}{2598960}$; в) $\frac{36}{2598960}$;
г) $\frac{4}{2598960}$; д) $\frac{5148}{2598960}$; е) $\frac{9216}{2598960}$.

13. $\frac{345}{567}$.

15. $\frac{4464}{9000}$.

17. $\frac{3168}{9000}$, $\frac{5146}{9000}$.

Раздел 8.6.

1. $\dfrac{8!}{4!2!}$.

3. $\dfrac{6!}{3!}$.

5. $\dfrac{26!}{8!6!9!3!}$.

7. $\dfrac{35!}{7!7!7!7!}$.

9. $\dfrac{24!}{(4!)^6}$.

11. $\dfrac{12!}{3!6!3!}2^6 3^3 = 31933440$.

13. $\dfrac{18!}{4!5!6!3!}2^4 4^5 5^3$.

15. $\dfrac{11!}{2!4!2!3!}4^2 2^4 5^3 = 2217600000$.

Раздел 8.7.

1. $\dfrac{16!}{4!12!} = 1820$.

3. $\dfrac{12!}{9!3!} = 220$.

5. $\dfrac{14!}{4!10!} = 1001$.

7. $\dfrac{12!}{2!10!} = 66$.

9. $\dfrac{15!}{3!12!} - 40 = 415.$

11. $\dfrac{16!}{3!13!} = 560.$

Раздел 8.8.

1. 21.

3. Если у одного человека 29 взаимных друзей, то каждый имеет 29 взаимных друзей. Если это не так, то количество взаимных друзей, которое люди могут иметь, изменяется от 0 до 28, поэтому существует 29 возможностей. Но в компании имеется 30 человек, поэтому, по крайней мере, 2 человека должны иметь одинаковое количество взаимных друзей.

5. 51.

7. 25.

9. Пусть $a_1, a_2, a_3, \ldots, a_n$ — последовательность из n положительных целых чисел. Пусть $s_1 = a_1$, $s_2 = a_1 + a_2$, $s_3 = a_1 + a_2 + a_3$, ..., $s_n = a_1 + a_2 + a_3 + \cdots + a_n$. Каждое s_i эквивалентно r_i по модулю n, где $0 \leq r_i \leq n-1$. Если какое-либо из r_i равно 0, то s_i делится на n, и доказательство завершено. Если нет, то существует $n-1$ значений для r_i и n штук r_i. Следовательно, два из r_i, например, r_j и r_k, должны иметь одно и то же значение. Предположим, что $j < k$, тогда $a_{j+1} + a_{j+2} + \cdots + a_k$ делится на n.

11. $26^2 + 1$.

13. 11.

15. $\left\lceil \dfrac{170}{14} \right\rceil = 13.$

17. Рассмотрим пары чисел $\{1,2\}, \{3,4\}, \{5,6\}, \ldots \{2n-1, 2n\}$. Имеется n пар чисел, поэтому если у нас $n+1$ чисел, то два должны принадлежать одной и той же паре. Эти два числа отличаются на 1.

19. Поскольку $0 + 1 + 2 + 3 + \cdots + n - 1 = n(n-1)/2$, существует только $n-1$ различных целых чисел, сумма которых меньше, чем $n(n-1)/2$. Поэтому m является суммой не более $n-1$ различных целых чисел. Имеется только $n-1$ различных целых чисел и n ящиков, поэтому два ящика должны содержать целые числа, которые не будут различными. Эти ящики содержат одинаковое количество шаров.

Раздел 8.9.

1. $\dfrac{1}{4}$.

3. $\dfrac{6}{18}$.

5. $\dfrac{1}{2}$, $\dfrac{1}{2}$.

7. $\dfrac{1}{7}$.

9. $\dfrac{15}{19}$.

11. $\dfrac{19}{6}$.

13. $\dfrac{483}{56}$.

15. $\dfrac{252}{36}$.

17. -13 центов.

19. 3456.

21. $\dfrac{35}{2^7}$.

23. $\binom{10}{2}(.1)^2(.9)^8 \approx .1937$.

25. $\dfrac{\binom{13}{2}\binom{39}{8}}{\binom{52}{10}} \approx .3$.

27. $\dfrac{36}{38}$.

29. $\dfrac{36}{38}$.

31. 3, 2.1.

Раздел 8.10.

1. $\frac{250}{535} \approx .47$.

3. $\frac{5}{22} \approx .23$.

5. $\frac{4}{13} \approx .31$.

7. 125.

Раздел 8.11.

1. **а)** .25; **б)** .232; **в)** .4989.

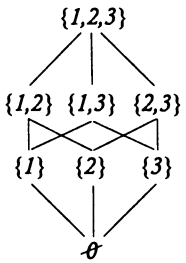
3. **а)** .26; **б)** .28; **в)** .585.

5. **а)** .365ь **б)** .6365.

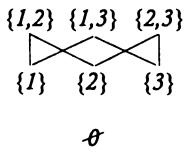
7. **а)** .181632; **б)** $1000 \cdot .67032 \approx 670$; **в)** $1000 \cdot .797704 \approx 798$.

Раздел 9.1.

1.



3.



5. **а)** наибольший элемент f;
б) наименьший элемент a;
в) максимальный элемент f, минимальный элемент a;
г) и то, и другое.

7. **а)** наибольший элемент g;
б) наименьший элемент не существует;
в) максимальный элемент g, минимальные элементы a, b, c, d;

г) верхней.

9. а) наибольший элемент i;

б) наименьший элемент a;

в) максимальный элемент i, минимальный элемент a;

г) и то, и другое.

11. а) докажем $a \vee (b \vee c) = (a \vee b) \vee c$

$$
\begin{aligned}
b \le (a \vee b) \le (a \vee b) \vee c \quad & \text{определение нвг} \\
c \le (a \vee b) \vee c \quad & \text{определение нвг} \\
\therefore (b \vee c) \le (a \vee b) \vee c \quad & \text{определение нвг} \\
a \le (a \vee b) \le (a \vee b) \vee c \quad & \text{определение нвг} \\
\therefore a \vee (b \vee c) \le (a \vee b) \vee c \quad & \text{определение нвг}
\end{aligned}
$$

Точно так же, $(a \vee b) \vee c \le a \vee (b \vee c)$ и $(a \vee b) \vee c = a \vee (b \vee c)$.
Докажем $a = a \vee a$. $a \le a \vee a$ по определению верхней грани. Поскольку $a \le a$ по определению наименьшей верхней грани, то $a \vee a \le a$. Таким образом, $a = a \vee a$.
Докажем $a \vee b = b \vee a$. $a \le a \vee b$ по определению наименьшей верхней грани. Точно так же, $b \le b \vee a$. Следовательно, по определению наименьшей верхней грани, $a \vee b \le b \vee a$. Точно так же, $b \vee a \le a \vee b$, и $a \vee b = b \vee a$;

б) докажем $a \wedge (b \wedge c) = (a \wedge b) \wedge c$.

$$
\begin{aligned}
b \ge (a \wedge b) \ge (a \wedge b) \wedge c \quad & \text{определение ннг} \\
c \ge (a \wedge b) \wedge c \quad & \text{определение ннг} \\
\therefore (b \wedge c) \ge (a \wedge b) \wedge c \quad & \text{определение ннг} \\
a \ge (a \wedge b) \ge (a \wedge b) \wedge c \quad & \text{определение ннг} \\
\therefore a \wedge (b \wedge c) \ge (a \wedge b) \wedge c \quad & \text{определение ннг}
\end{aligned}
$$

Точно так же, $(a \wedge b) \wedge c \ge a \wedge (b \wedge c)$ и $(a \wedge b) \wedge c = a \wedge (b \wedge c)$.
Докажем $a = a \wedge a$. $a \ge a \wedge a$ по определению наибольшей нижней грани. Поскольку $a \ge a$ по определению наибольшей нижней грани, то $a \wedge a \ge a$. Таким образом, $a = a \wedge a$.
Докажем $a \wedge b = b \wedge a$. $a \ge a \wedge b$ по определению наибольшей нижней грани. Точно так же, $b \ge b \wedge a$. Таким образом, по определению наибольшей нижней грани, $a \wedge b \ge b \wedge a$. Точно так же, $b \wedge a \ge a \wedge b$, и $a \wedge b = b \wedge a$.

13. Доказательство проведем индукцией по числу элементов в полурешетке, имеющей наименьшую верхнюю грань. Вполне очевидно, что единственный элемент имеет наименьшую верхнюю грань. По определению полурешетки, это утверждение истинно для любых двух элементов. Предположим, что любые k элементов имеют наименьшую верхнюю грань. Если полурешетка содержит только k элементов, доказательство завершено. В противном случае, пусть $a_1, a_2, a_3, \ldots a_k, a_{k+1}$ и σ_k — наименьшая верхняя грань для $a_1, a_2, a_3, \ldots a_k$. Тогда $\sigma_k \vee a_{k+1}$ — наименьшая верхняя грань для $a_1, a_2, a_3, \ldots a_k, a_{k+1}$. Вполне очевидно, $a_i \le \sigma_k \vee a_{k+1}$ для $1 \le i \le k+1$. Предположим, что $a_i \le b$ для $1 \le i \le k+1$. Поскольку $a_i \le b$ для $1 \le i \le k$, то $\sigma_k \le b$. Поэтому, учитывая, что $a_{k+1} \le b$ и $\sigma_k \le b$, имеем $\sigma_k \vee a_{k+1} \le b$. Следовательно, $\sigma_k \vee a_{k+1}$ — наименьшая верхняя грань для $a_1, a_2, a_3, \ldots a_k, a_{k+1}$.

15. а) если $A \le B$ и $B \le C$, то $A_{ij} \le B_{ij}$ для всех $1 \le i, j \le n$ и $B_{ij} \le C_{ij}$ для всех $1 \le i, j \le n$. Следовательно, $A_{ij} \le C_{ij}$ для всех $1 \le i, j \le n$ и $A \le C$;

б) поскольку $A_{ij}, B_{ij} \le A_{ij} \vee B_{ij}$ для всех $1 \le i, j \le n$, то $A \le A \vee B$ и $B \le A \vee B$. Следовательно, $A \vee B$ — верхняя грань для A и B. Предположим, что D — верхняя

грань для A и B. Тогда $A_{ij}, B_{ij} \leq D_{ij}$ для всех $1 \leq i, j \leq n$. Если $A_{ij} \vee B_{ij} = 1$, то $A_{ij} = 1$ или $B_{ij} = 1$, и $D_{ij} = 1$. Следовательно, $A_{ij} \vee B_{ij} \leq D_{ij}$ для всех $1 \leq i, j \leq n$, и $A \vee B \leq D$;

в) поскольку $A_{ij} \wedge B_{ij} \leq A_{ij}, B_{ij}$ для всех $1 \leq i, j \leq n$, то $A \wedge B$ — нижняя грань для A и B. Предположим, что D — нижняя грань для A и B. Если $A_{ij} \wedge B_{ij} = 1$, то $A_{ij} = 0$ или $B_{ij} = 0$, и $D_{ij} = 0$. Следовательно, $D_{ij} \leq A_{ij} \wedge B_{ij}$ для всех $1 \leq i, j \leq n$, и $D \leq A \wedge B$;

г) поскольку любые два элемента имеют как наименьшую верхнюю границу, так и наибольшую нижнюю границу, то S является как верхней, так и нижней полурешеткой;

д) наибольший элемент — матрица U, где $U_{ij} = 1$ для всех $1 \leq i, j \leq n$. Наименьший элемент — матрица L, где $L_{ij} = 0$ для всех $1 \leq i, j \leq n$.

Раздел 9.2.

1. $\{\{a\}, \{a, b, c\}\}$, $\{\emptyset, \{a\}, \{a, b\}, \{a, b, c\}\}$; $\{\{a\}, \{b\}, \{a, b\}, \{a, b, c\}\}$.

3. Требуется только показать, что произведение двух целых чисел вида $4k + 1$ имеет такую же форму. Пусть $a = 4m + 1$ и $b = 4n + 1$ — целые числа такого вида. Произведение

$$\begin{aligned} ab &= (4m + 1)(4n + 1) = \\ &= 16mn + 4m + 4 + 1 = \\ &= 4(4mn + m + n) + 1 \end{aligned}$$

также имеет вид $4k + 1$, где $k = 4mn + m + n$.

5. Поскольку $[a] \oplus [b] = [a + b]$ и $[a] \odot [b] = [a + b]$ для всех положительных целых чисел a и b, то имеет место замыкание.

Поскольку

$$\begin{aligned} ([a] \oplus [b]) \oplus [c] &= [a + b] \oplus [c] = \\ &= [(a + b) + c] \\ &= [a + (b + c)] = \\ &= [a] \oplus [b + c] = \\ &= ([a] \oplus ([b]) \oplus [c]), \end{aligned}$$

то сложение ассоциативно.

Поскольку

$$\begin{aligned} ([a] \odot [b]) \odot [c] &= [ab] \odot [c] = \\ &= [(ab)c] = \\ &= [a(bc)] = \\ &= [a] \odot [bc] = \\ &= ([a] \odot ([b]) \odot \oplus [c]), \end{aligned}$$

то умножение ассоциативно.

7. Пусть a и b — конечные произведения элементов из $\{a_1, a_2, a_3, \ldots, a_k\}$. Тогда ab несомненно является конечным произведением элементов из $\{a_1, a_2, a_3, \ldots, a_k\}$. Поскольку A^* — подмножество множества S, то имеет место ассоциативность. Следовательно, A^* — подполугруппа полугруппы S. Любая подгруппа A, содержащая все конечные произведения элементов из A, следовательно, должна содержать A^*.

9. **а)** в главе 2 было показано, что композиция ассоциативна, поэтому требуется только показать, что если f и g биективны, то $f \circ g$ также является биекцией. Если $(f \circ g)(x) =$

$(f \circ g)(y)$, то $f(g(x)) = f(g(y))$ и $g(x) = g(y)$, поскольку f биективна. Но тогда $x = y$, так как g биективна. Следовательно, $f \circ g$ — инъективна. Пусть $z \in S$. Поскольку f — сюръекция, то существует $y \in S$ такой, что $f(y) = z$. Поскольку g — сюръекция, то существует $x \in S$ такой, что $g(x) = y$. Таким образом,

$$(f \circ g)(x) = f(g(x)) = f(y) = z$$

и $f \circ g$ сюръективна;

б) пусть a и b принадлежат образу ретракционного отображения f и $a \neq b$. Следовательно, $a = f(c)$ и $b = f(d)$ для $c, d \in A$.

$$f(a) = f(f(c)) = f(c) = a$$

и

$$f(b) = f(f(d)) = f(b) = b,$$

так что f не только задает взаимно однозначное соответствие, но и является тождественным отображением;

в) поскольку $|x|$ положительно, то его абсолютное значение опять равно $|x|$, и f — ретракционное отображение. Ретракт — множество неотрицательных действительных чисел.
Поскольку $\lfloor x \rfloor$ — целое число, то $g(\lfloor x \rfloor) = \lfloor x \rfloor$, и g — ретракция. Ретракт — множество целых чисел.
Поскольку $\lceil x \rceil$ — целое число, то $h(\lceil x \rceil) = \lceil x \rceil$, и h — ретракция. Ретракт — множество целых чисел;

г) нет;

д) нет.

11. Пусть A и B — $(n \times n)$-матрицы с ненулевым детерминантом. Поскольку

$$\det(AB) = \det(A)\det(B),$$

то $\det(AB)$ не равен нулю. Множество $(n \times n)$-матриц с ненулевым детерминантом замкнуто относительно умножения. Поэтому они образуют полугруппу на множестве $(n \times n)$-матриц. Поскольку $\det(I) = 1$, где I — единица относительно умножения, то множество $(n \times n)$-матриц с ненулевым детерминантом образует моноид. Если $\det(A) = 0$ и $\det(B) = 0$, то $\det(AB) = 0$. Следовательно, множество $(n \times n)$-матриц с нулевым детерминантом замкнуто относительно умножения и также образует полугруппу.

13. Предположим, что $I * a = a = a * I$ для всех $a \in S$ и $e * a = a = a * e$ для всех $a \in S$. Следовательно,

$$\begin{aligned} e = I * e = &\quad \text{поскольку } I \text{ — единица} \\ = I &\quad \text{поскольку } e \text{ — единица.} \end{aligned}$$

Раздел 9.3.

1. (б); (г).

3. Подрешетка, полученная удалением e, изоморфна первой решетке в упражнении 2.

5. (а); (б); (г); (д).

7. **а)** $\{a\}; \{b\}; \{c\}$;

б) элементарными конъюнкциями являются $p \wedge q \wedge r$, $p \wedge q \wedge \sim r$, $p \wedge \sim q \wedge r$, $p \wedge \sim q \wedge \sim r$, $\sim p \wedge q \wedge r$, $\sim p \wedge q \wedge \sim r$, $\sim p \wedge \sim q \wedge r$, $\sim p \wedge \sim q \wedge \sim r$. Пересечение любых двух из них порождает противоречие и, следовательно, равно F, нулевому элементу;

в) (б),(в); **г)** (б),(в),(г).

Раздел 9.4.

1.

$\oplus$	[0]	[1]	[2]	[3]
[0]	[0]	[1]	[2]	[3]
[1]	[1]	[2]	[3]	[0]
[2]	[2]	[3]	[0]	[1]
[3]	[3]	[0]	[1]	[2]

$\odot$	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

3. [2].

5. [3].

7. Да.

9. Пусть $a = 3m$ и $b = 3n$. Тогда

$$a + b = 3k + 3n = 3(k + n),$$

поэтому сумма a и b кратна 3. Несомненно, $0 = 3 \cdot 0$ принадлежит множеству. Обратным элементом для $3k$ будет $(-3)k$. Следовательно, это подгруппа.

11. Согласно задаче 5 раздела 9.2, это полугруппа. Поскольку

$$[0] + [a] = [0 + a] = [a],$$

то [0] является единицей. Поскольку $[a] + [6 - a] = [6] = [0]$, то каждый элемент имеет обратный элемент. Следовательно, Z_6 — группа относительно сложения. Это множество не является группой относительно умножения, так как элементы [0], [2], [3] и [4] не имеют обратных элементов.

13. Элемент g^p — единица, принадлежит множеству и лучше обозначить ее через g^0. Согласно этим обозначениям $g^q \cdot g^r = g^s$, где $0 \leq s < p$ и $s \equiv q + r \pmod{p}$. Обратным элементом для g^q является g^{p-q}.

Раздел 9.5.

1. Имеем $f(1) = f(1 \cdot 1) = f(1) \cdot f(1)$, поэтому $f(1)$ — идемпотент. Вполне очевидно, это единица группы H, поскольку в группе H только идемпотент является единицей.

3. Если $f(h) = f(h')$, то $ah = ah'$. Следовательно, $a^{-1}ah = a^{-1}ah'$ и $h = h'$. Таким образом, функция f — инъекция. Пусть $h \in H$. Тогда $f(a^{-1}h) = aa^{-1}h = h$. Следовательно, f — сюръекция.

5. Определим $f : Z_5^R \to Z_4$ соотношениями $f(0) = 1$, $f(1) = 3$, $f(2) = 4$ и $f(3) = 2$. Соотношение легко угадать, поскольку мы отображаем 1, что порождает Z_5^R на 3, а это, в свою очередь, порождает Z_4. В изоморфности также легко убедиться, проанализировав следующие таблицы

$\oplus$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\odot$	1	3	4	2
	1	3	4	2
	3	4	2	1
	4	2	1	3
	2	1	3	4

7. Пусть $= \{1, \delta_1\}$. $\rho_1\{1, \delta_1\} = \{\rho_1, \phi_1\}$ и $\{1, \delta_1\}\rho_1 = \{\rho_1, \phi_2\}$, поэтому группа H не является нормальной.

9. Было уже показано, что $\phi(1)$ — единица группы H и, следовательно, группы $\phi(K)$. Пусть $h, h' \in \phi(K)$. Тогда $h = \phi(g)$ и $h' = \phi(g')$ для некоторых $g, g' \in H$.

$$h \cdot h' = \phi(g)\phi(g') = \phi(g \cdot g'),$$

поэтому $h \cdot h' \in \phi(K)$. Далее, $\phi(g^{-1}) \cdot \phi(g) = \phi(g \cdot g^{-1}) = \phi(1)$, поэтому если $h \in \phi(K)$, то $h^{-1} \in \phi(K)$. Следовательно, $\phi(K)$ — подгруппа группы H.

11. **а)** $(1,5,2,3)(4)$; **б)** I;
в) $(1,3,5,4)(2)$; **г)** $(1,3)(2)(4,6,7)(5)(8)$.

13. Предположим, что $gH = Hg$ для всех $g \in G$. Пусть $h \in H$. Поэтому $hg \in Hg$. Поскольку $Hg = gH$, то существует элемент h' такой, что $hg = gh'$ и $h = gh'g^{-1} \in gHg^{-1}$. Следовательно, $H \subseteq gHg^{-1}$. Предположим, что $ghg^{-1} \in gHg^{-1}$ Поскольку $Hg = gH$, то существует элемент h' такой, что $gh = gh'$ и $h' = ghg^{-1}$. Следовательно, $ghg^{-1} \in H$ и $gHg^{-1} \subseteq H$. Таким образом, $gHg^{-1} = H$.

Раздел 10.1.

1. **а)** $x = -20$ $y = 6$; **б)** $x = 15$ $y = -78$;
в) $x = 180$ $y = -87$; **г)** $x = -55$ $y = 22$;
д) $x = 12$ $y = -4$.

3. **а)** $x = -28$ $y = 36$; **б)** $x = -64$ $y = 36$;
в) решение не существует; **г)** $x = -2$ $y = 6$;
д) решение не существует.

5. Процедура **Целочисленное решение** (m, n, c):
//Решение уравнения $x \cdot m + y \cdot n = c$//
Положить $t = m$, $u = n$, $x = 0$, $x' = 1$, $y = 0$, $y' = 1$;
Условный цикл: до тех пор, пока $u \neq 0$,
$q = \lfloor t \div u \rfloor$, $t = u$, $u = t - qu$, $x = x'$, $x' = x - qx'$, $y = y'$, $y' = y - qy'$;
Конец условного цикла;
Положить $d = u$, $e = c \div d$;
Если $e \neq \lfloor e \rfloor$, то решения не существует;
Иначе $x = ex$; $y = ey$
Конец процедуры.

Раздел 10.2.

1. $n = 1, 2, 4, 5, 8, 10, 20, 40$.

3. Если a нечетное, то $a = 2m + 1$ для некоторого целого числа m. Следовательно,

$$a^2 = (2m+1)^2 = 4m^2 + 4m + 1$$

и

$$a^2 - 1 = 4m^2 + 4m = 4m(m+1).$$

Но или m, или $m + 1$ является четным. Очевидно, $a^2 - 1$ делится на 8.

5. Если $a \equiv b \pmod{n}$, где $n = \text{НОК}(n_1, n_2, n_3, \ldots, n_k)$, то $a - b$ делится на $\text{НОК}(n_1, n_2, n_3, \ldots, n_k)$, так что $a-b$ делится на n_i для $1 \leq i \leq k$. Таким образом, $a \equiv b \pmod{n_i}$. Обратно, предположим, что $a \equiv b \pmod{n_i}$, для некоторого $1 \leq i \leq k$. Тогда $a - b$ делится на n_i для некоторого $1 \leq i \leq k$. По определению НОК, $a - b$ делится на $\text{НОК}(n_1, n_2, n_3, \ldots, n_k)$, и $a \equiv b \pmod{n}$.

7. Докажем индукцией по m. Несомненно, $a \equiv b \pmod{n}$. Поэтому утверждение истинно для $m = 1$. Предположим, что оно истинно для $n = k$, поэтому $a^k \equiv b^k \pmod{n}$. По теореме 3.54, если $a \equiv b \pmod{n}$ и $a^k \equiv b^k \pmod{n}$, то $a^{k+1} \equiv b^{k+1}$, поэтому утверждение истинно для $n = k + 1$.

9. Если $a \equiv b \pmod{\frac{n}{\text{НОД}(c,n)}}$, то $a - b = k\frac{n}{\text{НОД}(c,n)}$ для некоторого k. Поэтому $ac - bc = k\frac{n}{\text{НОД}(c,n)}c$. Поскольку $\text{НОД}(c,n) | c$, то $\frac{c}{\text{НОД}(c,n)}$ — целое число, скажем, q. Таким образом, $ac - bc = kqn$ и $ac - bc \bmod n$. Обратно, если $ac \equiv bc \bmod n$, то $ac - bc = kn$ для некоторого k. Поскольку $(a-b)c = kn$, $(a-b)\frac{c}{\text{НОД}(c,n)} = \frac{n}{\text{НОД}(c,n)}$. Поскольку $\frac{c}{\text{НОД}(c,n)}$ и $\frac{n}{\text{НОД}(c,n)}$ взаимно простые числа, $\frac{n}{\text{НОД}(c,n)}$ делит $(a-b)$, поэтому $a \equiv b \pmod{\frac{n}{\text{НОД}(c,n)}}$.

Раздел 10.3.

1. **а)** $x \equiv 81 \pmod{300}$;
б) $x \equiv 815 \pmod{1260}$;
в) $x \equiv 16a - 15b \pmod{240}$;
г) $x \equiv 1272 \pmod{2310}$.

3. 1459 разноцветных шариков; 97 рядов по 15 шариков, 182 рядов по 8 шариков и 63 рядов по 23 шариков.

5. **а)** $x \equiv 21 \pmod{72}$;
б) $x \equiv 104 \pmod{180}$;
в) нет решения.

Раздел 10.4.

1. $m = 4$ $n = 2$.

3. Поскольку 1009 — простое число, то по теореме Уилсона $1008! \equiv 1 \pmod{(1009)}$. Поскольку $1008 \equiv -1 \pmod{(1009)}$, то $1007! \equiv -1 \pmod{1009}$.

5. Докажем по индукции, что если $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_t^{\alpha_t}$, то $\phi(n) = \prod_{i=1}^{t} p_i^{\alpha_i - 1}(p_i - 1)$. При $n = 1$ теорема верна в силу теоремы 10.20. Предположим, теорема верна при $t = k$, так что, если $m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$, то $\phi(m) = \prod_{i=1}^{k} p_i^{\alpha_i - 1}(p_i - 1)$. Допустим, что $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_{k+1}^{\alpha_{k+1}}$. Тогда $\phi(n) = \phi(m)\phi(p_{k+1}^{\alpha_{k+1}})$ по теореме 10.19. Но $\phi(p_{k+1}^{\alpha_{k+1}}) = p_{k+1}^{\alpha_{k+1}-1}(p_{k+1} - 1)$ по теореме 10.20, поэтому $\phi(n) = \prod_{i=1}^{k} p_i^{\alpha_i - 1}(p_i - 1)(p_{k+1}^{\alpha_{k+1}-1}(p_{k+1} - 1)) = \prod_{i=1}^{k+1} p_i^{\alpha_i - 1}(p_i - 1)$.

7. 1080.

9. По теореме Уилсона, число p простое тогда и только тогда, когда $(p-1)! \equiv -1 \pmod{p}$. Последнее имеет место тогда и только тогда, когда $(p-1)! + 1 \equiv 0 \pmod{p}$. А это имеет место тогда и только тогда, когда $(p-1)! + 1$ делится на p.

Раздел 10.5.

1. Если a не равно 0 по модулю p, то $a \equiv b \pmod{p}$, где $1 \leq b \leq p-1$. По малой теореме Ферма, $b^{p-1} \equiv 1 \pmod{p}$. Но $a^{p-1} \equiv b^{p-1} \pmod{p}$. Следовательно, $a^{p-1} \equiv 1 \pmod{p}$.

3. Поскольку $J \equiv 0 \pmod{p-1}$, $J = k(p-1)$ для некоторого целого k. Пусть $1 \leq a \leq p-1$. Поскольку

$$a^{p-1} \equiv 1 \pmod{p},$$
$$a^J = a^{k(p-1)} = (a^{(p-1)})^k \equiv 1^k \equiv 1 \pmod{p},$$

то

$$1^J + 2^J + \cdots + (p-1)^J \equiv 1 + 1 + \cdots + 1 \pmod{p} \equiv$$
$$\equiv p - 1 \pmod{p} \equiv$$
$$\equiv -1 \pmod{p}.$$

7. Поскольку $a^p \equiv a \pmod{p}$, $b^p \equiv b \pmod{p}$ и $(a+b)^p \equiv a+b \pmod{p}$, то $(a+b)^p \equiv a^p + b^p \pmod{p}$.

9. **а)**

a	1	2	4	5	7	8
$\operatorname{ord}_n a$	1	6	3	6	3	2

б)

a	1	3	7	9	11	13	17	19
$\operatorname{ord}_n a$	1	4	4	2	2	4	4	2

в)

a	1	2	4	5	7	8	10	11	13	14	16	17	19	20	22	23	25	26
$\operatorname{ord}_n a$	1	18	9	18	9	6	3	18	9	18	9	6	3	18	9	18	9	2

11. **а)** $2401 \equiv 1 \pmod 5$;
б) $2401 \equiv 1 \pmod 2$;
в) $2401 \equiv 1 \pmod{10}$;
г) поскольку $7^4 \equiv 1 \pmod{10}$, $(7^4)^k \equiv 1^k \pmod{10}$, то $7^{4k} \equiv 1 \pmod{10}$. Последняя десятичная цифра числа 7^{4000} есть 1.

13. **а)** 7; **б)** 376; **в)** 68; **г)** 961; **д)** 535044134.

15. **а)** Если $a^{\frac{n-1}{2}} \equiv -1 \pmod n$, то $a^{n-1} \equiv 1 \pmod n$. Поскольку $a^{\frac{n-1}{2}} \equiv -1 \pmod n$, оно не сравнимо с 1 по модулю n. Таким образом, все требования критерия Лукаса удовлетворены, поэтому n — простое число;
б) $2^{18} \equiv -1 \pmod{37}$, $2^{12} \equiv 26 \pmod{37}$, поэтому число 37 простое;
$3^{198} \equiv -1 \pmod{199}$, $3^{66} \equiv 106 \pmod{199}$, $3^{18} \equiv 125 \pmod{199}$, поэтому число 199 — простое.

17. **а)** $3^{256} \equiv 1 \pmod{257}$; $3^{128} \equiv -1 \pmod{257}$. Поэтому число 257 простое.
б) $3^{65536} \equiv 1 \pmod{65537}$;
$3^{32768} \equiv -1 \pmod{65537}$;
$5^{65536} \equiv 1 \pmod{65537}$;
$5^{32768} \equiv -1 \pmod{65537}$.

Раздел 11.1.

1. (a), (c), (e).

3. **а)** $a_n = c_1 \cdot 3^n$;
б) $a_n = c_1 \cdot (-3)^n$;
в) $a_n = c_1 \cdot (-3)^n + c_2 \cdot 2^n$;
г) $a_n = c_1 \cdot \left(\frac{1+\sqrt{13}}{2}\right)^n + c_2 \cdot \left(\frac{1-\sqrt{13}}{2}\right)^n$.

5. **а)** $a_n = (-4)^n$;
б) $a_n = 6^{n-1}$;
в) $a_n = 3^n + 2^n$;
г) $a_n = 4 \cdot 3^n - 2 \cdot 4^n$.

7. **а)** $a_n = 3 \cdot 3^n + 4n \cdot 3^n$;
б) $a_n = 2 \cdot 2^n + n \cdot 2^n$;
в) $a_n = (-2)^n + 3n \cdot (-2)^n$;
г) $a_n = -3 + 4n$;
д) $a_n = 2 \cdot (-1)^n - 8n \cdot (-1)^n$.

9. **а)** $a_n = c_1 \cdot 4^n + c_2$;
б) $a_n = c_1 + c_2 n + c_3 \cos\left(\frac{n\pi}{2}\right) + c_4 \sin\left(\frac{n\pi}{2}\right)$;
в) $a_n = c_1 + c_2 \cdot (-1)^n + c_3 \cos\left(\frac{n\pi}{2}\right) + c_4 \sin\left(\frac{n\pi}{2}\right)$;
г) $a_n = c_1 + c_2 \cdot n + c_3 n^2 + c_4 \cos\left(\frac{n\pi}{4}\right) + c_5 \sin\left(\frac{n\pi}{4}\right) + c_6 \cos\left(\frac{3n\pi}{4}\right) + c_7 \sin\left(\frac{3n\pi}{4}\right)$.

11. Воспользуемся индукцией.
При $n = 3$ имеем $L_3 = a + b = a\,\text{Фиб}(2) + b\,\text{Фиб}(1)$
Допустим, $L_k = a\,\text{Фиб}(k-1) + b\,\text{Фиб}(k-2)$ для любого $3 \le k < m$.

$$\begin{aligned} L_m &= L_{m-1} + L_{m-2} = \\ &= a\,\text{Фиб}(m-2) + b\,\text{Фиб}(m-3) + a\,\text{Фиб}(m-3) + b\,\text{Фиб}(m-4) = \\ &= a(\text{Фиб}(m-2) + \text{Фиб}(m-3)) + b(\text{Фиб}(m-3) + \text{Фиб}(m-4)) = \\ &= a\,\text{Фиб}(m-1) + b\,\text{Фиб}(m-2). \end{aligned}$$

Раздел 11.2.

1. **а)** $a_n = c_1 \cdot 2^n - 5$;
б) $a_n = c_1 \cdot (-3)^n + \frac{1}{4}n + \frac{3}{16}$;
в) $c_1 \cdot 3^n + c_2 \cdot (-4)^n - \frac{2}{3} \cdot 2^n$;
г) $a_n = c_1 \cdot \left(\frac{-3+\sqrt{5}}{2}\right)^n + c_2 \cdot \left(\frac{-3-\sqrt{5}}{2}\right)^n + \frac{9}{19} \cdot (3)^n$;
д) $a_n = c_1 \cdot 2^n + c_2 n \cdot 2^n + n^2 + 8n + 32$.

3. **а)** $a_n = c_1 \cdot 2^n + n \cdot 2^n$;
б) $a_n = c_1 \cdot (-3)^n + c_2 n \cdot (-3)^n + \frac{1}{2}n^2 \cdot (-3)^n$;
в) $a_n = c_1 + c_2 n \cdot 2^n - 5n$;
г) $a_n = c_1 \cdot 3^n + c_2 \cdot (-2)^n + \frac{9}{15}n \cdot 3^n$;
д) $a_n = 2^n \left(c_1 \cos\left(\frac{n\pi}{6}\right) + c_2 \sin\left(\frac{n\pi}{6}\right)\right) + \frac{3}{5-2\sqrt{3}}$.

5. **а)** $a_n = \dfrac{n(n+1)(2n+1)}{6}$; **б)** $a_n = \dfrac{n(3n-1)}{2}$;
в) $a_n = 2^{n+1} - 2$; **г)** $a_n = \frac{1}{3}n(n+1)(n+2)$.

Раздел 11.3.

1. 12.

3. 644.

5. $3x^2 + 8x - 6$.

7. $\dfrac{10}{(3x+4)(3x+7)}$.

9. По определению, $\Delta(x!) = (x+1)! - x! = x!(x+1-1) = x \cdot x!$.

11.
$$\Delta\frac{f(x)}{g(x)} = \frac{f(x+1)}{g(x+1)} - \frac{f(x)}{g(x)} = \frac{f(x+1)g(x) - f(x)g(x+1)}{g(x+1)g(x)} =$$
$$= \frac{f(x+1)g(x) - f(x)g(x) + f(x)g(x) - f(x)g(x+1)}{g(x+1)g(x)} =$$
$$= \frac{(f(x+1) - f(x))g(x) - f(x)(g(x+1) - g(x))}{g(x+1)g(x)} =$$
$$= \frac{\Delta f(x) \cdot g(x) - f(x) \cdot \Delta g(x)}{g(x+1)g(x)}.$$

Раздел 11.4.

1. $18x - 10$.

3. $60x^{(2)} + 72x - 18$.

5. $(x^{(3)} - 2x^{(2)} - 4)(4x^{(3)} + 6x^{(2)} - 12x + 1) +$
$+((x+1)^4 + 2(x+1)^{(3)} - 6(x+1)^{(2)} + x - 2)(3x^{(2)} - 4x)$.

7. $\dfrac{(x^{(4)} - 3x^{(3)} + 6x^{(2)})(4x^{(3)} - 6x + 3) - (x^{(4)} - 3x^{(2)} + 3x)(4x^{(3)} - 9x^{(2)} + 12x)}{(x^{(4)} - 3x^{(3)} + 6x^{(2)})((x+1)^{(4)} - 3(x+1)^{(3)} + 6(x+1)^{(2)})}$.

9. $3!$.

11. x.

13. $x^{(4)} - 7x^{(2)} - 5x + 5$.

15. $x^{(5)} + 9x^{(4)} + 20x^{(3)} + 10x^{(2)} + x - 1$.

17. $x^4 - 3x^3 - 2x^2 + 2x + 1$.

19. $x^5 - 11x^4 + 42x^3 - 65x^2 + 34x - 1$.

21. -3.

Раздел 11.5.

1. 58137.

3. -8988.

5. 62132.

7. $\displaystyle\sum_{i=1}^{n} 2i = \frac{i^{(2)}}{2}\Big|_1^{n+1} = \frac{(n+1)^{(2)} - 1^{(2)}}{2} = \frac{n(n+1)}{2}$.

9. $\sum_{i=1}^{n} i(i+1) = \sum_{i=1}^{n} i^2 + i = \sum_{i=1}^{n} i^{(2)} + 2i = i^{(3)} + i^{(2)}\Big|_1^{n+1} =$

$$= \frac{(n+1)(n)(n-1)}{3} + (n+1)(n) = \frac{(n+1)(n)(n+2)}{3}.$$

11. 177144.

13. 22520.

15. 26106.

17. 6203013120.

19. 28385280.

21. Пусть $\Delta F(x) = f(x)$, тогда

$$\Delta \sum_{i=x}^{c} f(i) = \Delta(F(c+1) - F(x)) = \Delta(F(c+1) - \Delta F(x) = 0 - f(x) = -f(x).$$

Раздел 12.1.

1. **а)** $k^n = 3^7 = 2187$;
б) $k!S_k^{(n)} = 3!S_3^{(7)} = 6 \cdot 301 = 1806$;
в) $C(n+k-1, n) = C(7+3-1, 7) = C(9, 7) = 36$;
г) $C(n-1, k-1) = C(6, 2) = 15$;
д) $S_1^{(7)} + S_2^{(7)} + S_3^{(7)} = 1 + 63 + 301 = 365$;
е) $S_k^{(n)} = 301$.

3.

$n\backslash k$	0	1	2	3	4	5	6	7	8	9	10	11	12
11	0	1	1023	28501	145750	246730	179487	63987	11880	1155	55	1	
12	0	1	2047	86526	611501	1379400	1323652	627396	159027	22275	1705	66	1

5. **а)** $k^n = 4^{12} = 16777216$;
б) $k!S_k^{(n)} = 4!S_4^{(12)} = 24 \cdot 611501 = 14676024$;
в) $C(n+k-1, n) = C(15, 12) = 455$;
г) $C(n-1, k-1) = C(11, 3) = 165$;
д) (е) $S_1^{(12)} + S_2^{(12)} + S_3^{(12)} + S_4^{(12)} = 1 + 2047 + 86526 + 611501 = 70075$;
е) $S_k^{(n)} = S_4^{(12)} = 611501$.

7. **а)** $S(n, 0) = \frac{1}{0!}(-1)^0 \binom{0}{0}(0-0)^n = 0$;
б) $S(n, k-1) + kS(n, k) =$

$$= \frac{1}{(k-1)!}\sum_{i=0}^{k-1}(-1)^i \binom{k-1}{i}(k-i-1)^n + k \cdot \frac{1}{k!}\sum_{i=0}^{k}(-1)^i \binom{k}{i}(k-i)^n =$$

$$= \frac{1}{(k-1)!}\sum_{i=0}^{k-1}(-1)^i \binom{k-1}{i}(k-i-1)^n + \frac{1}{(k-1)!}\sum_{i=0}^{k}(-1)^i \binom{k}{i}(k-i)^n =$$

$$= \frac{1}{(k-1)!}\left(-\sum_{i=1}^{k}(-1)^i \binom{k-1}{i-1}(k-i)^n + \sum_{i=0}^{k}(-1)^i \binom{k}{i}(k-i)^n\right) =$$

$$= \frac{1}{(k-1)!}\left(-\sum_{i=1}^{k}(-1)^i \binom{k-1}{i-1}(k-i)^n + \sum_{i=1}^{k}(-1)^i \binom{k}{i}(k-i)^n + k^n\right) =$$

$$= \frac{1}{(k-1)!}\left(\sum_{i=1}^{k}(-1)^i \left[\binom{k}{i} - \binom{k-1}{i-1}\right](k-i)^n + k^n\right) =$$

$$= \frac{1}{(k-1)!}\left(\sum_{i=1}^{k}(-1)^i\binom{k-1}{i}(k-i)^n + k^n\right) =$$

$$= \frac{1}{(k-1)!}\left(\sum_{i=1}^{k}(-1)^i\frac{(k-i)!}{i!(k-i-1)!}(k-i)^n + k^n\right) =$$

$$= \frac{1}{(k)!}\left(\sum_{i=1}^{k}(-1)^i\frac{(k)!}{i!(k-i-1)!}(k-i)^n + k^{n+1}\right) =$$

$$= \frac{1}{(k)!}\left(\sum_{i=1}^{k}(-1)^i\frac{(k)!}{i!(k-i)!}(k-i)^{n+1} + k^{n+1}\right) =$$

$$= \frac{1}{(k)!}\left(\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^{n+1} =\right)$$

$$= S(n+1,k).$$

Раздел 12.2.

1. $C_5 = \frac{1}{6}\binom{10}{5} = \frac{252}{6} = 42;$ $C_6 = \frac{1}{7}\binom{12}{6} = \frac{924}{7} = 132;$

$C_7 = \frac{1}{8}\binom{14}{7} = \frac{3432}{8} = 429.$

3. $C_{15} = \frac{1}{16}\binom{30}{15} = 9694845.$

Раздел 12.3.

1. $0 = (1-1)^n = \binom{n}{1} - \binom{n}{2} + \binom{n}{3} - \cdots + (-1)^n\binom{n}{n}.$

3. $D_{15} = 15!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} + \cdots + \frac{1}{14!} - \frac{1}{15!}) \approx$
$\approx 4.1 \times 10^{11}.$

5. **а)** $7!(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \frac{1}{720} - \frac{1}{5040}) = 1854;$
б) $7! - 1855 = 3185;$
в) $7!(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \frac{1}{720}) = 1855;$
г) 0.

7. **а)** $7!(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \frac{1}{720} - \frac{1}{5040}) = 1854;$
б) $7!(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \frac{1}{720}) = 1855;$
в) $7! - 1855 = 3185;$
г) $7! - 1855 - 1854 = 1331.$

9. $\binom{n}{k}D_{n-k}.$

11. Число, которое делится на 7, 2 или 5, есть $2730 + 1092 + 780 - 546 - 390 - 156 + 78 = 3588.$ Число, взаимно простое с 700, равно $5640 - 3588 = 1872.$

Раздел 12.4.

1. $R(x,C) = 1 + 6x + 6x^2.$

3. $R(x,C) = (1 + 4x + 2x^2)^2.$

5. $R(x,C) = (1 + 3x + x^2)(1 + 2x) + x(1 + 3x + 2x^2).$

7. $R(x,C) = (1 + 3x + x^2)^2(1 + x).$

9. $R(x,C) = 1 + 8x + 15x^2 + 6x^3$.

11. $R(x,C) = (1+2x)(1+x)(1+4x+2x^2)$.

13. $R(x,C) = 1 + 4x + 68x^2 + 148x^3 + 144x^4 + 48x^5$,
количество размещений равно $8! - 14\cdot 7! + 68\cdot 6! - 148\cdot 5! + 144\cdot 4! - 48\cdot 3! = 4128$.

15. $R(x,C) = 1 + 6x + 11x^2 + 6x^3$,
количество размещений равно $5! - 6\cdot 4! + 11\cdot 3! - 6\cdot 2! = 30$.

17. $R(x,C) = 1 + 12x + 54x^2 + 112x^3 + 108x^4 + 48x^5 + 8x^6$, количество
размещений равно $6! - 112\cdot 5! + 54\cdot 4! - 112\cdot 3! + 108\cdot 2! - 48\cdot 1! + 8\cdot 0! = 80$.

19. $R(x,C) = 1 + 8x + 20x^2 + 16x^3 + 4x^4$,
количество размещений равно $4! - 8\cdot 3! + 20\cdot 2! - 16\cdot 1! + 4\cdot 0! = 4$.

21. $R(x,C) = 1 + 13x + 59x^2 + 115x^3 + 100x^4 + 32x^5$,
количество размещений равно $6! - 13\cdot 5! + 59\cdot 4! - 115\cdot 3! + 100\cdot 2! - 32 = 54$.

23. $R(x,C) = 1 + 8x + 20x^2 + 16x^3 + 4x^4$,
количество размещений равно $4! - 8\cdot 3! + 20\cdot 2! - 16\cdot 1! + 4 = 4$.

25. $R(x,C) = 1 + 10x + 35x^2 + 50x^3 + 26x^4 + 4x^5$,
количество размещений равно $5! - 10\cdot 4! + 35\cdot 3! - 50\cdot 2! + 26\cdot 1! - 4 = 12$.

27. $R(x,C) = 1 + 10x + 35x^2 + 50x^3 + 25x^4 + 2x^5$,
количество размещений равно $5! - 10\cdot 4! + 35\cdot 3! - 50\cdot 2! + 25\cdot 1! - 2 = 13$.

Раздел 13.2.

1. **а)** $\dfrac{-1}{x+4} + \dfrac{1}{x+3}$; **б)** $\dfrac{-1}{x-2} + \dfrac{3}{x-3}$; **в)** $\dfrac{-10}{x-1} - \dfrac{6}{(x-1)^2} + \dfrac{11}{x-2}$;
г) $\dfrac{\frac{1}{2}}{x-1} - \dfrac{2}{x-2} + \dfrac{\frac{3}{2}}{(x-3)}$; **д)** $\dfrac{14}{x-1} + \dfrac{4}{(x-1)^2} - \dfrac{14}{x-2} + \dfrac{13}{(x-2)^2}$.

3. **а)** $\frac{1}{4}(1 - \frac{1}{4}x + \frac{1}{16}x^2 - \frac{1}{64}x^3 + \cdots + (-\frac{1}{4})^n x^n + \cdots)$;

б) $\frac{1}{4}(1 + x + \frac{3}{4}x^2 + \frac{1}{2}x^3 + \cdots + (\dfrac{n+1}{2^n})x^n + \cdots)$;

в) $-1 - 3x - 6x^2 - 10x^3 - \cdots - \dfrac{n(n+1)}{2}x^n - \cdots$;

г) $\frac{1}{2}x + \frac{3}{4}x^2 + \frac{7}{8}x^3 + \cdots + \dfrac{2^n-1}{2^n}x^n + \cdots$;

д) $-\frac{1}{2} - \frac{5}{4}x - \frac{29}{8}x^2 - \cdots - (4n - 6 + \dfrac{13}{2^{n+1}})x^n - \cdots$.

5. **а)** $a_n = 3^{n+1} - 2^{n+1}$; **б)** $a_n = \frac{11}{4}3^n - \frac{1}{2}n - \frac{3}{4}$; **в)** $a_n = 6\cdot 3^n + 2(-1)^n$;
г) $a_n = 2^n - n2^n$; **д)** $a_n = 2^n + 3$.

7. **а)** $a_n = \dfrac{n^2+n+2}{2}$; **б)** $a_n = -2\cdot 2^n + 4\cdot 3^n$;
в) $a_n = (n+3)2^n - n - 2$; **г)** $a_n = (3n-3)3^n + 4\cdot 2^n$.

Раздел 13.3.

1. **а)** $(1 + x + x^2 + \cdots)^2(1 + x + x^2)(1 + x + x^2 + x^3 + x^4)$;
б) $(1 + x + x^2 + \cdots)^2(1 + x + x^2 + x^3)(1 + x^2 + x^4 + x^6 + \cdots)$;
в) $(x + x^3 + x^5 + x^7 + \cdots)(1 + x^2 + x^4 + x^6 + \cdots)(1 + x + x^2 + \cdots)(x^4 + x^5 + x^6 + x^7 + \cdots)$;
г) $(x + x^2 + x^3 + \cdots)(x^2 + x^3 + x^4 + \cdots)(x^3 + x^4 + x^5 + \cdots)(x^4 + x^5 + x^6 + x^7 + \cdots)$;
д) $(1 + x + x^2 + \cdots)(1 + x^2 + x^4 + \cdots)(1 + x^3 + x^6 + \cdots)(1 + x^4 + x^8 + \cdots)$.

3. **а)** $(1 + x + x^2 + x^3 + x^4)(1 + x + x^2 + x^3)(1 + x + x^2 + \cdots + x^6)(1 + x + x^2)$;
б) $(x + x^2)(x + x^2 + x^3 + x^4 + x^5)(x + x^2 + x^3 + x^4)(x + x^2 + x^3)$;
в) $(1 + x + x^2 + \cdots + x^6)(x^4 + x^5 + \cdots + x^{11})(1 + x + x^2 + \cdots + x^7)(1 + x^2 + x^4 + \cdots x^{10})$;
г) $(x^2 + x^4 + +x^6)((x + x^2 + x^3 \cdots + +x^{12})(x + x^3 + x^5 + x^7)(x^2 + x^4 + \cdots x^{10})$.

5. $(1+x+x^2+\cdots)(1+x^5+x^{10}+\cdots)(1+x^{10}+x^{20}+\cdots)((1+x^{25}+x^{50}+\cdots)$.

7. $(1+x^3+x^6+\cdots)(1+x^5+x^{10}+\cdots)$.

9. $(1+x+x^2+x^3+x^4+x^5+x^6)^k$.

11. $\binom{10}{6}$.

13. 1.

15. $\binom{7}{4}$.

17. 51.

19. $\binom{14}{11}\cdot 3^{12}$.

21. **а)** 36;
б) 30.

23. $\binom{15}{12}-3\binom{10}{7}+3\binom{5}{2}$.

25. $F(x)-xF(x)-x^2F(x)=$
$=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5+\cdots-$
$-a_0x-a_1x^2-a_2x^3-a_3x^4-a_4x^5+\cdots-$
$-a_0x^2-a_1x^3-a_2x^4-a_3x^5+\cdots=$
$=a_0+(a_1-a_0)x=$
$=x,$

поскольку $a_0=0$ и $a_1=1$.

Следовательно, $(1-x-x^2)F(x)=x$, так что

$$F(x)=\frac{x}{1-x-x^2}=\frac{x}{1-(x+x^2)}.$$

Воспользовавшись разложением, получаем

$$F(x)=x(1+(x+x^2)+(x+x^2)^2+(x+x^2)^3+(x+x^2)^4+\cdots=$$
$$=x+x^2(1+x)+x^3(1+x)^2+x^4(1+x)^3+x^5(1+x)^4+\cdots.$$

Коэффициент при x^{n+1} в выражении $x^{k+1}(1+x)^k$, если существует, то равен $\binom{k}{n-k}$. Суммируя по $k=n,n-1,n-2,\ldots$, имеем

$$a_n=\binom{n}{0}+\binom{n-1}{1}+\binom{n-2}{2}+\binom{n-3}{3}+\cdots.$$

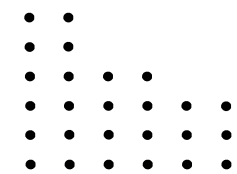
Раздел 13.4 .

1. 15.

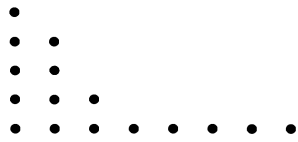
3. $\dfrac{1}{1-x}\cdot\dfrac{1}{1-x^2}\cdot\dfrac{1}{1-x^3}\cdot\dfrac{1}{1-x^4}\cdot\dfrac{1}{1-x5}\cdot\dfrac{1}{1-x^6}$.

5. $\dfrac{1}{1-x^2}\cdot\dfrac{1}{1-x^4}\cdot\dfrac{1}{1-x^6}\cdot\dfrac{1}{1-x^8}\cdot\dfrac{1}{1-x^{10}}\cdot\cdots$

7. В строках, начиная с верхней, представим каждую часть разбиения на четные величины. Теперь, глядя на столбцы, видим, что представленные части разбиения присутствуют парами и убывают по размеру. Поскольку рост в строке влечет за собой увеличение четного числа, в столбцах это порождает пару или пары одного размера.



9. Пусть столбцы изображают разбиение $2n$ объектов на n частей.



Поскольку ни одна из этих частей не пуста, то нижняя строка содержит n элементов. Если отбросить эту строку, то легко увидеть, что другие строки образуют разбиение n элементов.

11. $\dfrac{1}{1-x}\cdot\dfrac{1}{1-x^2}\cdot\dfrac{1}{1-x^3}\cdot\dfrac{1}{1-x^4}\cdot\dfrac{1}{1-x^5}\cdot\dfrac{1}{1-x^6}\cdot\dfrac{1}{1-x^7}$.

Раздел 13.5.

1. **а)** $\left(\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots+\frac{x^{10}}{10!}\right)\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots+\frac{x^{10}}{10!}\right)$;

б) $\left(\frac{x^5}{5!}+\frac{x^6}{6!}+\frac{x^7}{7!}+\cdots\frac{x^{10}}{10!}\right)\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}\right)\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}\right)$;

в) $\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\cdots\frac{x^{10}}{10!}\right)\left(x+\frac{x^3}{3!}+\frac{x^5}{5!}+\frac{x^7}{7!}+\frac{x^9}{9!}\right)\left(1+\frac{x^2}{2!}+\frac{x^4}{4!}+\cdots\frac{x^{10}}{10!}\right)$.

3. e^x.

5. Производящая функция $e^{5x}-2e^{4x}+e^{3x}$, количество способов $5^n-2\cdot4^n+3^n$.

7. Производящая функция $e^{3x}-3e^{2x}+3e^x+1$, количество способов $3^{30}-3\cdot2^{30}+3$.

9. Производящая функция $\frac{1}{4}\left(e^{4x}-2e^{3x}+e^{2x}-1+2e^{-x}-3e^{-2x}\right)$, количество способов $\frac{1}{4}(4^{25}-2\cdot3^{25}+2^{25}+2(-1)^{25}-(-2)^{25})$.

11. $e^x\cdot e^{5x}\cdot e^{10x}\cdot e^{25x}=e^{41x}$.

13. $(e^x-1)(e^{2x}-1)(e^{3x}-1)\cdots(e^{kx}-1)$.

15. **а)** $(e^x-1)=\left(1+\dfrac{x}{1!}+\dfrac{x^2}{2!}+\cdots\right)-1=$

$$=\frac{x}{1!}+\frac{x^2}{2!}+\cdots;$$

б) $$\frac{e^x+e^{-x}}{2}=\frac{1}{2}(1+\frac{x}{1!}+\frac{x^2}{2!}+\cdots)+\frac{1}{2}(1-\frac{x}{1!}+\frac{x^2}{2!}-\frac{x^3}{3!}\cdots)=$$
$$=\frac{1}{2}(2+2\frac{x^2}{2!}+2\frac{x^4}{4!}\cdots)=$$
$$=1+\frac{x^2}{2!}+\frac{x^4}{4!}+\cdots;$$

в) $$\frac{e^x-e^{-x}}{2}=\frac{1}{2}(1+\frac{x}{1!}+\frac{x^2}{2!}+\cdots)-\frac{1}{2}(1-\frac{x}{1!}+\frac{x^2}{2!}-\frac{x^3}{3!}\cdots)=$$
$$=\frac{1}{2}(2x+2\frac{x^3}{3!}+2\frac{x^5}{5!}\cdots)=$$
$$=x+\frac{x^3}{3!}+\frac{x^5}{5!}+\cdots.$$

Раздел 14.1.

1. а) $f(v_1) = f(v_2) = a$;
$f(v_3) = f(v_4) = b$;
$f(v_5) = c$;

б) гомоморфизм не существует;

в) $f(v_3) = c$;
$f(v_2) = b$;
$f(v_1) = f(v_4) = a$;

г) $f(v_1) = f(v_3) = f(v_5) = a$;
$f(v_2) = f(v_4) = f(v_6) = b$;

д) $f(v_1) = f(v_5) = c$;
$f(v_3) = f(v_6) = b$;
$f(v_2) = f(v_4) = a$;

е) $f(v_1) = f(v_3) = f(v_5) = f(v_7) = c$;
$f(v_8) = f(v_4) = a$;
$f(v_2) = f(v_6) = b$.

3. а) графы не изоморфны, разное количество вершин;

б) $f(v_1) = a$;
$f(v_2) = b$;
$f(v_3) = c$;
$f(v_4) = d$;

в) $f(v_6) = a$;
$f(v_1) = d$;
$f(v_3) = e$;
$f(v_4) = b$;
$f(v_2) = c$;
$f(v_5) = f$;

г) $f(v_1) = e$;
$f(v_4) = f$;
$f(v_2) = a$;
$f(v_3) = b$;
$f(v_6) = d$;
$f(v_5) = c$;

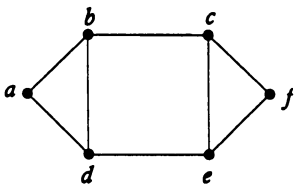
д) графы не изоморфны, вершина d имеет степень 4, ни одна из вершин первого графа не имеет степень 4;

е) графы не изоморфны, первый граф несвязный, второй граф — связный.

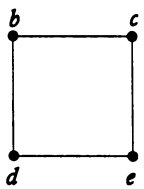
5. (a), (c).

7. (a), (c), (e).

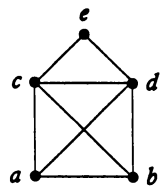
9. а) объединение;



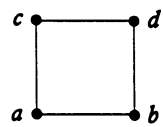
пересечение;



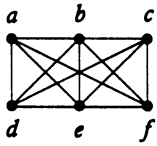
б) объединение;



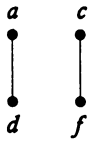
пересечение;



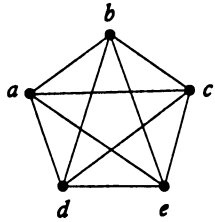
в) объединение;



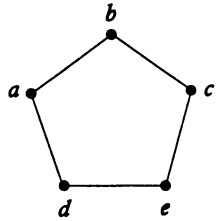
пересечение;



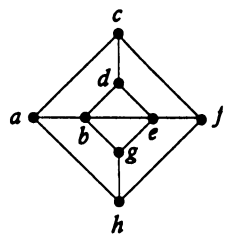
г) объединение;



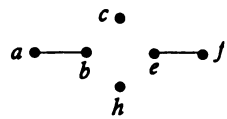
пересечение;



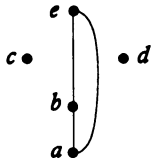
д) объединение;



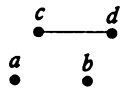
пересечение.



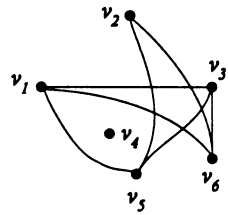
11. а)



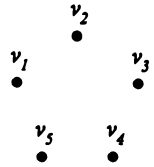
б)



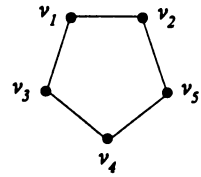
в)



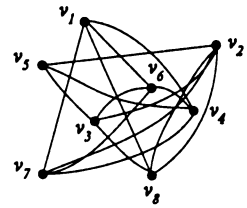
г)



д)

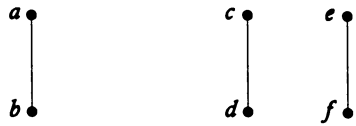


е)



13. (a); (b).

15. G имеет вид G' имеет вид



$f : G \to G'$ задана соотношениями $f(a) = c; \quad f(b) = d.$

17. Пусть $f : G \to G'$. Пусть $a, b \in f(V)$. Тогда $a = f(x), b = f(y)$ для некоторых $x, y \in V$. По-

скольку граф G полный, то существует $e = \{x, y\}$. Поэтому $f(e) = \{a, b\}$. Следовательно, между двумя любыми вершинами графа $f(V)$ имеется ребро, и граф $f(G)$ — полный.

19. Пусть $v \in V$. Имеется наибольший связный граф, содержащий v. Поэтому v принадлежит компоненте. Пусть $e \in E$. Имеется наибольший связный граф, содержащий e. Поэтому e принадлежит компоненте. Следовательно, граф $G(E, V)$ — объединение компонент, и, по теореме 12.17, они являются попарно непересекающимися.

21. Эквивалентно показать, что если $G'(E', V')$ — расширение графа $G(E, V)$, то один является связным тогда и только тогда, когда связным является второй. Предположим, что граф G связный и ребро ab заменено путем acb. Пусть v_1 и v_2 — вершины в графе G. Поскольку G связный, то существует путь из v_1 в v_2. Если ребро ab входит в путь, заменим его на acb, и v_1, v_2 остаются связанными в графе G'. Обратно, если G' связный и $v_1, v_2 \in G'$, то в графе G' существует путь из v_1 в v_2. По определению расширения, только ребра, инцидентные вершине c, являются инцидентными вершинам a и b. Следовательно, любой путь, включающий c, должен включать путь acb или bca. В графе G путь acb можно заменить на ab, а путь bca можно заменить на ba, так что путь из v_1 в v_2 по-прежнему существует.

23. Если $f : G(E, V) \to G'(E', V')$ — изоморфизм, то он задает взаимно однозначное отображение из E на E' и взаимно однозначное отображениее из V на V'. Таким образом, G и G' содержат одинаковое количество ребер и вершин.

25. Пусть $f : G(E, V) \to G'(E', V')$ — изоморфизм. Вершины a и b принадлежат одной и той же компоненте графа G тогда и только тогда, когда существует путь $av_1v_2 \cdots v_{k-1}b$ из a в b, тогда и только тогда, когда существует путь $f(a)f(v_1)f(v_2) \cdots f(v_{k-1})f(b)$ из $f(a)$ в $f(b)$, тогда и только тогда, когда $f(a)$ и $f(b)$ принадлежат одной и той же компоненте.

27. Пусть $f : G(E, V) \to G'(E', V')$ — изоморфизм. Поскольку $\deg(v) = \deg(f(v))$ для всех $v \in V$, то все вершины в V имеют четную степень тогда и только тогда, когда все вершины в $f(V)$ имеют четную степень.

29. Пусть $f : G(E, V) \to G'(E', V')$ — изоморфизм. Если граф G двудольный, то $V = A \cup B$; при этом не существует ребро между вершиной из A и вершиной из B. Если существует ребро e' между $f(a) \in f(A)$ и $f(b) \in f(B)$, то $e' = f(e)$, где e — ребро между a и b. Но это противоречит предположению, что граф G двудольный. Следовательно, граф G' двудольный.

31. По определению, каждая вершина графа G принадлежит его дополнению. Таким образом, пересечение графа G и его дополнения содержит все вершины графа G. Ни одно из ребер графа G не принадлежит его дополнению. Следовательно, пересечение графа G и его дополнения не содержит ни одного ребра графа G. Значит, пересечение графа и его дополнения представляет собой граф, состоящий только из вершин графа G и не содержащий ребер.

33. Пусть e — разрезающее ребро графа G и $\hat{G}$ — граф G с удаленным ребром e. Пусть $\hat{G}'$ — граф G' с удаленным ребром $f(e)$. Пусть f — изоморфизм из $\hat{G}$ в $\hat{G}'$. По определению разрезающего множества, существуют вершины a, b, которые не связаны в $\hat{G}$. Поэтому они принадлежат различным компонентам графа $\hat{G}$. Следовательно, как показано в предыдущем упражнении, $f(a)$ и $f(b)$ принадлежат различным компонентам графа $\hat{G}'$, и граф $\hat{G}'$ несвязный. Вполне очевидно, что $f(e)$ — разрезающее ребро графа G'.

35. В матрице смежности графа имеется единица, если в матрице смежности его дополнения имеется нуль.

37. Пусть $a_1, a_2, a_3, \ldots, a_n$ — вершины графа G, и P — перестановочная матрица размерности $n \times n$. Пусть $v_1 = (1, 0, 0, \ldots, 0)$ представляет a_1, $v_2 = (0, 1, 0, \ldots, 0)$ представляет a_2 и $v_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ представляет a_i, где 1 находится на i-ом месте. Аналогично, пусть $w_1 = (1, 0, 0, \ldots, 0)$, представляет a_1, $w_2 = (0, 1, 0, \ldots, 0)$, представляет a_2 и $w_i =$

$(0,\ldots,0,1,0,\ldots,0)$ представляет a_i, где 1 находится на i-ом месте. Определим $f(a_i) = b_j$, если $Pv_i = w_j$. Предположим, что $f(a_i) = b_j$ и $f(a_k) = b_l$. Тогда $P_{ji} = 1$ и $P_{lk} = 1$ Рассмотрим произведение PBP^t, где B — матрица смежности графа G. Пусть $T = BP^t$. Тогда, поскольку $P^t_{kl} = P_{lk} = 1$, то $B_{ik}P^t_{kl} = T_{il}$. Пусть $B' = PT$. Поскольку $P_{ji} = 1$, то $B'_{jl} = P_{ji}T_{il}$, поэтому число в j-ой строке и в l-ом столбце матрицы B' равно B_{ik}, и ребро между b_j и b_l существует тогда и только тогда, когда существует ребро между a_i и a_k. Следовательно, G и G' изоморфны. Обратно, если они изоморфны, то матрицы B и B', определенные выше, являются матрицами смежности для соответствующих графов.

39. Разрезающие множества включают $\{\{a,d\},\{a,b\}\}$, $\{\{d,e\},\{d,a\}\}$, и $\{\{f,g\},\{f,i\}\}$. Разрезающими ребрами являются $\{b,c\}$, $\{e,f\}$, $\{i,j\}$.

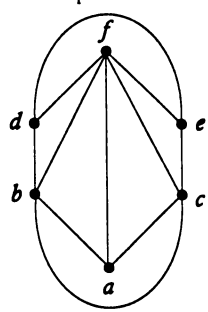
41. Пусть $e = \{v,v'\}$ входит в цикл $vv'v_1\cdots v_{k-1}v$ и пусть a,b — вершины графа G. Если имеется путь из a в b, который не включает e, то путь из a в b существует, если ребро e удалено. Если e не входит в путь, то заменим ребро vv' путем $vv_{k-1}\cdots v_1v'$, и путь из a в b по-прежнему существует. Обратно, допустим, что $e = \{v,v'\}$ не является разрезающим ребром. В этом случае, если ребро e удалено, то по-прежнему имеется путь $vv_{k-1}\cdots v_1v'$ из v в v', и $vv_{k-1}\cdots v_1v'v$ является циклом.

43. Если $G_i(V_i,E_i)$ состоит из единственного ребра e, то не существует простой цикл графа G, содержащий e, поэтому e — разрезающее.

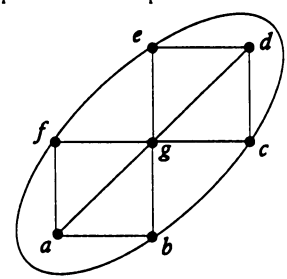
45. Если воспользоваться определением объединения, которое дано в определении 14.14, то вполне очевидно, что любой граф G_j является подграфом объединения $\bigcup_{i=1}^{n} A_i$. Также, согласно этому определению, если любой граф G_i принадлежит графу G'', то каждое ребро и каждая вершина любого из графов G_i принадлежат G''. Следовательно, каждое ребро и каждая вершина объединения графов G_i принадлежат G''. Таким образом, определение объединения, которое дано в определении 14.14, удовлетворяет определению, данному в задаче. Два определения эквивалентны, если существует только один граф, удовлетворяющий определению, данному в задаче. Но если имеются два графа G' и G'', удовлетворяющие определению, то $G' \preceq G''$ и $G'' \preceq G'$, поэтому $G'' = G'$.

Раздел 14.2.

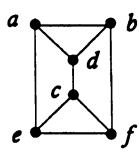
1. **а)** да, граф нарисован так, что линии не пересекаются;
б) нет, граф содержит $K_{3,3}$ как подграф;
в) нет, граф содержит K_5 как подграф;
г) да, граф можно изобразить как



д) да, граф можно изобразить как

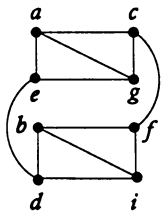


3. **а)** да, граф можно изобразить как

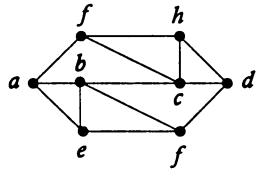


б) нет, граф содержит $K_{3,3}$ как подграф.

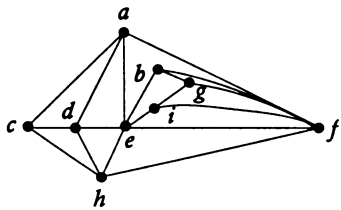
в) да, граф можно изобразить как



г) да, граф можно изобразить как



д) да, граф можно изобразить как



5. Если имеется 12 вершин степени 3, то степень вершин равна 36, поэтому имеется 18 ребер. Поскольку $v - e + f = 2$, то $12 - 18 + f = 2$ и $f = 8$.

7. **а)** 4;

б) 6;

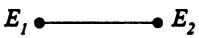
в) $2k$.

9. Поскольку граф с 4 вершинами не может содержать граф, гомеоморфный подграфу, гомеоморфному K_5 или $K_{3,3}$, то он должен быть планарным. С другой строны, легко показать, что граф K_4 планарный. Поскольку любой граф с 4 вершинами — подграф графа K_4, он является планарным.

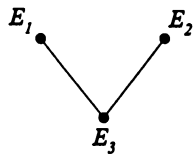
11. Будем использовать индукцию по количеству ребер дерева. Очевидно, что любое дерево с одним ребром — планарное. Предположим, что любое дерево, имеющее k ребер, планарное. Пусть T — дерево, имеющее $k+1$ ребер. Любое дерево имеет вершину степени 1. Выберем такую вершину и удалим ее и ребро, инцидентное к ней. Оставшееся дерево по индуктивному предположению является планарным. Удаленное ребро и вершина могут быть снова добавлены так, что ребро не пересечет ни одно из других ребер. Следовательно, дерево T — планарный граф.

Раздел 14.3.

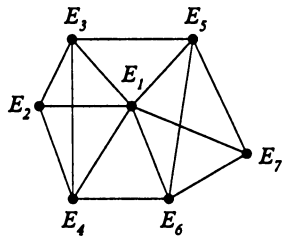
1. а)



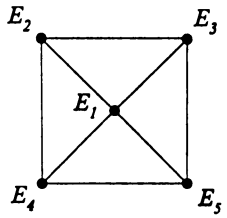
б)



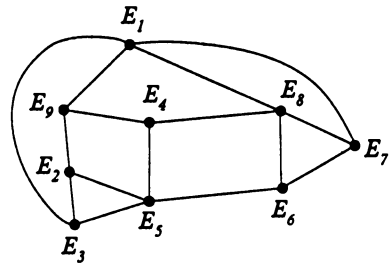
в)



г)



д)



3. **а)** 2; **б)** 3; **в)** 4; **г)** 3; **д)** 4.

5. $K_{m,n}$ содержит множество m вершин и множество n вершин. Ребра имеются только между двумя множествами. Поэтому можно раскрасить одно множество одним цветом, а второе множество — другим.

7. **а)** $\lambda(\lambda-1)(\lambda-2)^2$; **б)** $\lambda(\lambda-1)^2(\lambda-2)$;

в) $\lambda(\lambda-1)(\lambda-2)^2(\lambda-3)$; **г)** $\lambda(\lambda-1)(\lambda-2)^2$;

д) $\lambda(\lambda-1)(\lambda-2)^2$.

9. Если планарный граф изоморфен своему двойственному графу, то количество граней равно количеству вершин. Поскольку $v-e+f=2$, то $n-e+n=2$ и $e=2n-2$.

Раздел 14.4.

1. **а)** $acgfedba$; **б)** $abfecda$; **в)** $abihfdcega$;

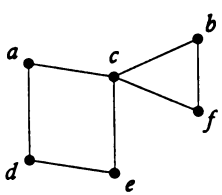
г) не существует; **д)** не существует.

3. **а)** $abcde$; **б)** $abcdef$; **в)** $fdcegabih$;

г) $egacdfhb$; **д)** $adbcefgh$.

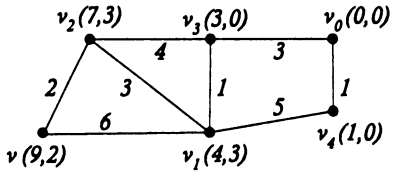
5. Если граф G имеет гамильтонов цикл, то любое ребро является частью цикла. Но разрезающее ребро не может быть ребром, входящим в цикл. Следовательно, G не содержит разрезающее ребро. Пусть $\{a,b\}$ — разрезающее ребро и компоненты C_1 и C_2 — компоненты, имеющие гамильтоновы циклы, например, $av_1v_2v_3\cdots v_{k-1}a$ и $bv'_1v'_2v'_3\cdots v'_{m-1}b$. Тогда $v_1v_2v_3\cdots v_{k-1}abv'_1v'_2v'_3\cdots v'_{m-1}$ — гамильтонов путь.

7.

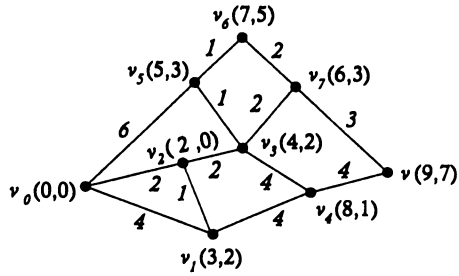


Раздел 14.5.

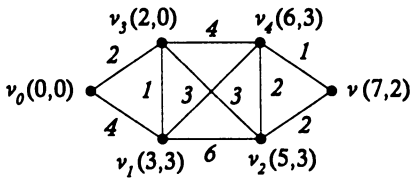
1. а)



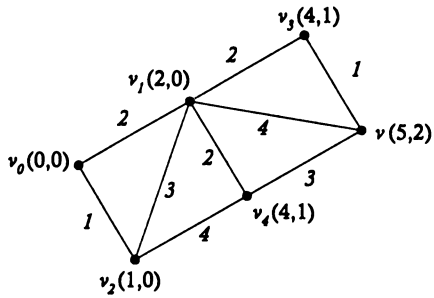
б)



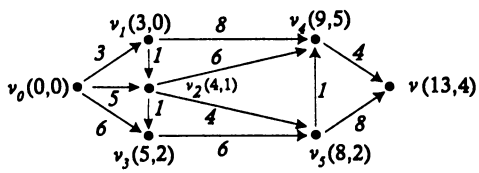
в)



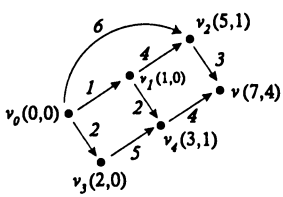
г)



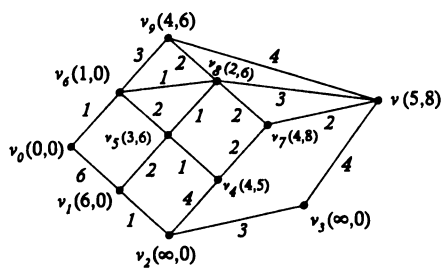
д)



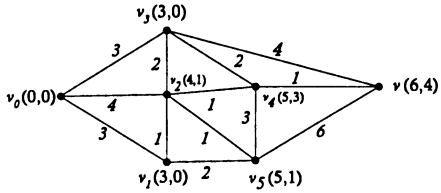
3. а)



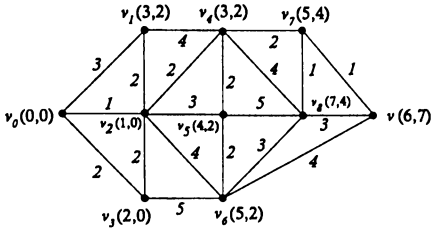
б)



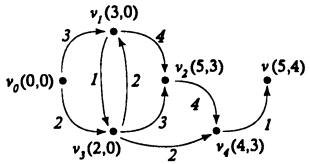
в)



г)



д)

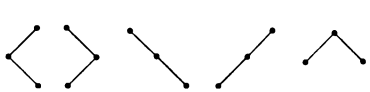


Раздел 15.1 .

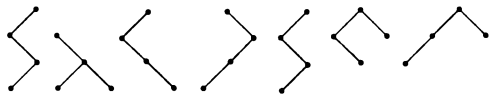
1. а)



б)



в)



3. а) 63 вершины, 32 листа и 31 внутренняя вершина;
б) 141 вершина, 81 лист и 60 внутренних вершин;
в) 511 вершин, 256 листьев и 255 внутренних вершин;
г) 85 вершин, 64 листа и 21 внутренняя вершина;
д) 10 вершин, 1 лист и 9 внутренних вершин.

5. а) (i) 5, (ii) 2 (iii) 4, (iv) 2, (v) d, (vi) e, a;
б) (i) 5, (ii) 4 (iii) 2, (iv) 2, (v) d, (vi) e, d;
в) (i) 4, (ii) 3 (iii) 1, (iv) 1, (v) d, (vi) e, d;
г) (i) 6, (ii) 5 (iii) 3, (iv) 3, (v) d, (vi) e, d;
д) (i) 4, (ii) 1 (iii) 3, (iv) 1, (v) d, (vi) a, e, d.

7. а) сбалансированное; **б)** сбалансированное; **в)** сбалансированное.

9. а) 7; **б)** 127.

11. а) $l = 2^h$. Поэтому $\log_2 l = \log_2 2^h = h$;
б) поскольку количество листьев меньше или равно 2^h, и равно 2^h только в том случае, когда дерево полное, то

$$2^h \geq l$$
$$\therefore \log_2 2^h \geq \log_2 l \qquad \text{поскольку } l \text{ — возрастающая функция}$$
$$\text{и} \quad h \geq \log_2 l$$

в)

$$v = 2^{h+1} - 1;$$
$$\therefore v + 1 = 2^{h+1};$$
$$\therefore \log_2(v+1) = h + 1;$$
$$\text{и} \quad h = \log_2(v+1) - 1;$$

г)

$$v \le 2^{h+1} - 1, \qquad \text{поскольку полное дерево имеет максимальное количество вершин;}$$
$$\therefore v + 1 \le 2^{h+1}, \qquad \text{поскольку } l \text{ — возрастающая функция;}$$
$$\therefore \log_2(v+1) \le h + 1, \qquad \text{поскольку } l \text{ — возрастающая функция}$$
$$\text{и} \quad h \ge \log_2(v+1) - 1.$$

13. Предположим, что граф G — дерево. Пусть ребро $\{a, b\}$ добавлено к дереву. Поскольку дерево было связным, уже существовал путь $av_1v_2v_3 \cdots v_{k-1}b$ из a в b. Таким образом, $av_1v_2v_3 \cdots v_{k-1}ba$ является циклом.

Предположим, что добавление ребра $\{a, b\}$ создает два цикла , $av_1v_2v_3 \cdots v_{k-1}b$ и $av'_1v'_2v'_3 \cdots v'_{m-1}b$, тогда $av_1v_2v_3 \cdots v_{k-1}bbv'_{m-1} \cdots v'_3v'_2v'_1a$ уже является циклом, что противоречит предположению, что G — дерево.

Предположим, что граф G не является деревом. Тогда он уже содержит цикл. Предположим, что ребро $\{a, b\}$ добавлено в G. Путь из a в b уже существовал, поскольку граф G связный. Теперь имеются два пути из a в b, поэтому $\{a, b\}$ — часть цикла, но не часть исходного цикла, поэтому в графе G с добавленным ребром теперь имеется, по крайней мере, два цикла.

15. Утверждение ошибочное. На дереве лист не является разрезающей вершиной.

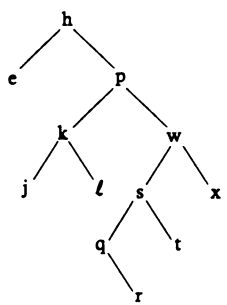
17. G содержит эйлеров цикл.

19. Нужно показать, что $\hat{a}(vv') = \hat{a}(v)\hat{a}(v')$.

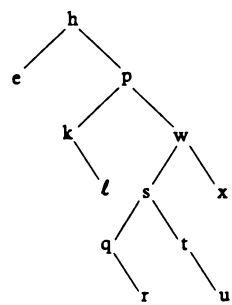
$$\hat{a}(vv') = a(vv') =$$
$$= (aa)(vv') = \qquad \text{поскольку элементы полурешетки являются идемпотентами}$$
$$= avav' = \qquad \text{поскольку решетка ассоциативна и коммутативна}$$
$$= \hat{a}(v)\hat{a}(v').$$

Раздел 15.2.

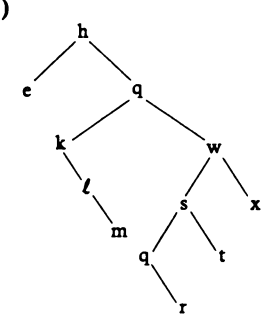
1. а)



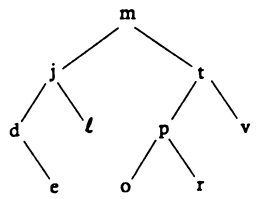
б)



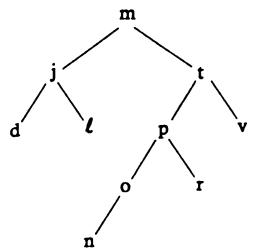
в)



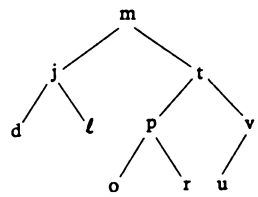
3. а)



б)

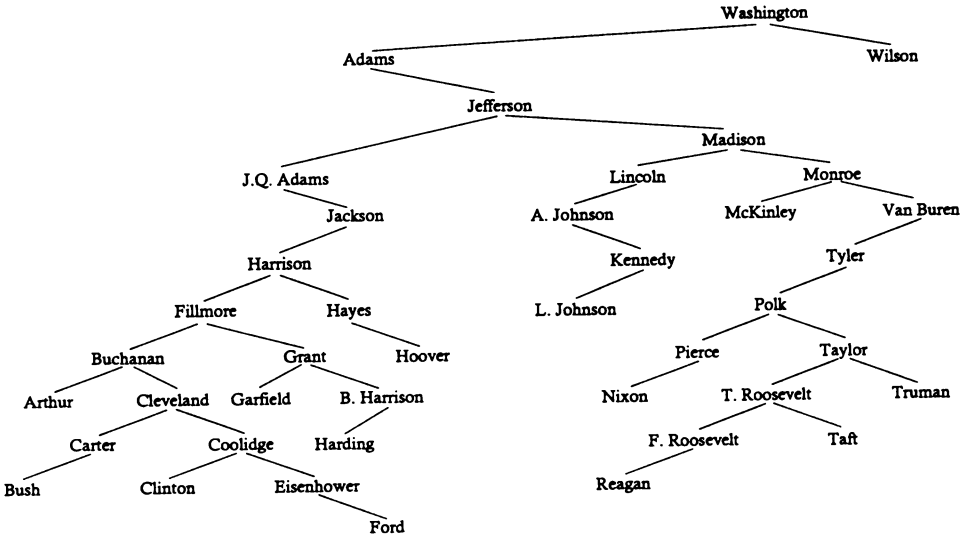


в)

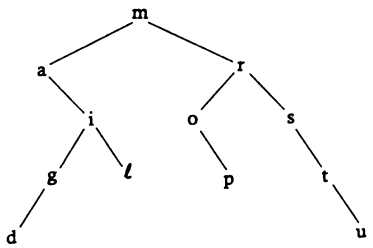


5. Цепь.

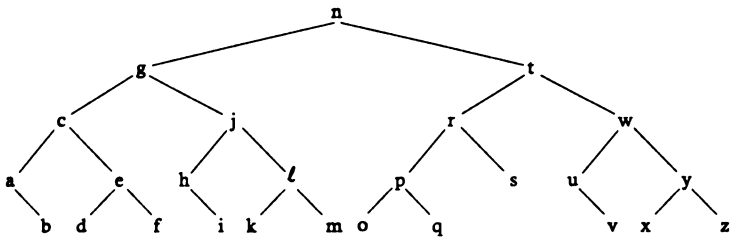
7.



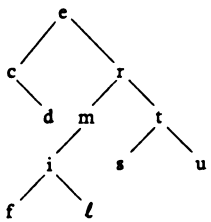
9.



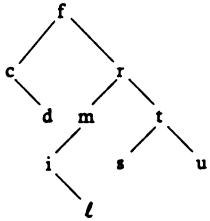
11.



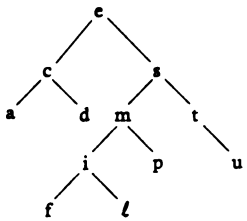
13. а)



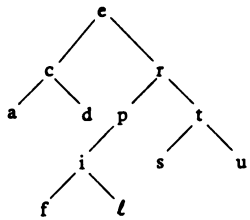
б)



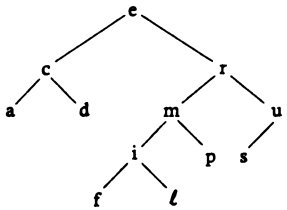
в)



г)



д)



15. Процедура **Поиск** (a, r)

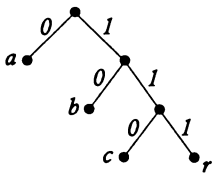
Если $a = r$, то
Положить *Найдено* = *да*;
Если $a < r$, то
Если r имеет левого сына u, то вызвать Поиск(a, u);
Положить *Найдено* = *нет*;
Если $a > r$, то
Если r имеет правого сына w, то вызвать Поиск(a, w)
Положить *Найдено* = *нет*;
Конец процедуры.

Раздел 15.3.

1. а) 10111000100101111110100100010;
б) 001011010110101;
в) decided;
г) issues.

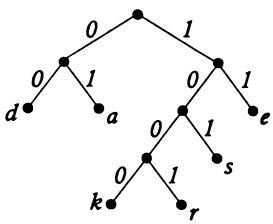
3. Вес кода 1 равен 127. Вес кода 2 равен 112. Вес кода 3 равен 136. Выбираем код 2.

5. а)



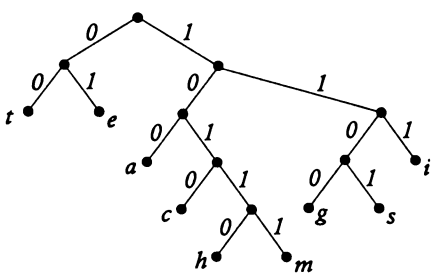
б) $a = 0$, $b = 10$, $c = 110$, $r = 111$;
в) вес равен 41;
г) 110111010;
д) 100111;
е) car;
ж) barb.

7. а)



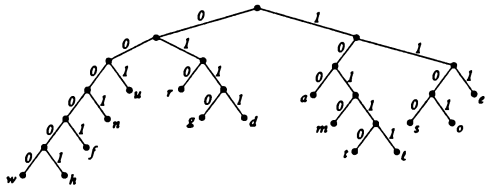
б) $a = 01$, $d = 00$, $e = 11$, $k = 1000$, $r = 1001$, $s = 101$;
в) вес равен 92;
г) 11100101101111001;
д) 000110011000111001;
е) raked;
ж) dearer.

9. а)

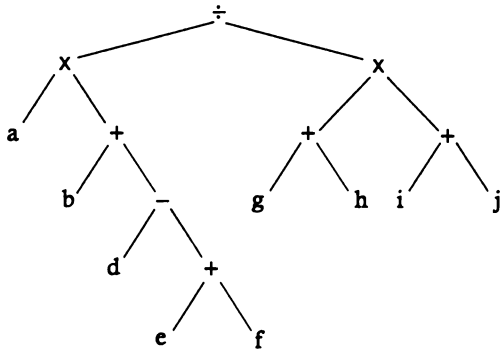


б) $a = 100$; $c = 1010$; $e = 01$; $i = 111$; $g = 1100$; $h = 10110$; $m = 10111$; $s = 1101$; $t = 00$;
в) вес равен 118.

11. а)

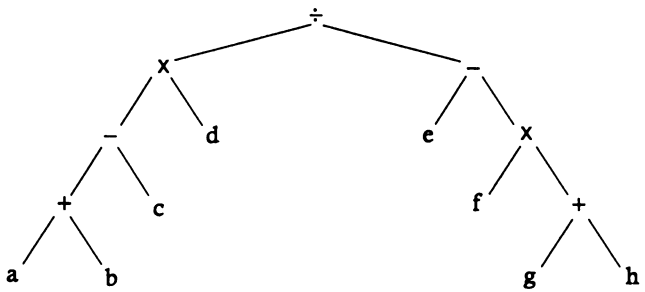

0
1
0
1
0
1

11. а)



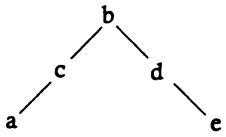
б) $\div \times a + b - d + ef \times + gh + ij$; в) $abcdef + - + \times gh + ij + \times \div$.

13. а)

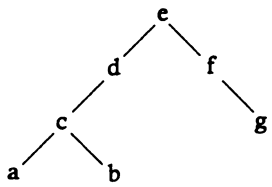


б) $\div \times - + abcd - e \times f + gh$; в) $ab + c - d \times efgh + x - \div$.

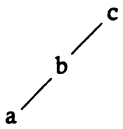
15.



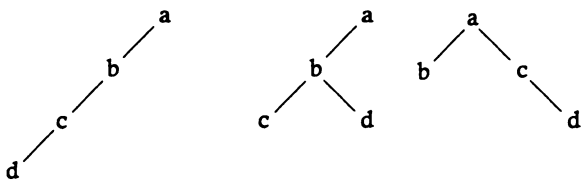
17.



19.



21.



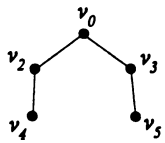
23. (д).

25. Для каждой арифметической операции количество предшествующих ей чисел на два или больше превышает количество предшествующих ей арифметических операций, и последний символ должен быть арифметической операцией.

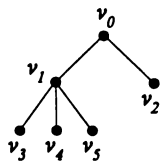
27. После того как вершина b была обработана на обоих деревьях, на шаге 4 вызывается алгоритм ИБД для обработки на обоих деревьях левого сына вершины b. На обоих деревьях это вершина c, и на обоих деревьях у вершины c нет сыновей. ИБД возвращается в обоих деревьях к вершине b, и на шаге 5 вызывается ИБД для обработки правого сына вершины b. Когда на шаге 2 будет $Iso = F$, алгоритм установит, что деревья не изоморфны.

Раздел 15.5.

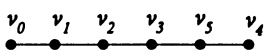
1. а)



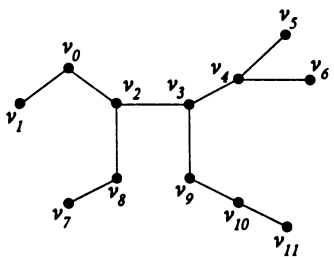
б)



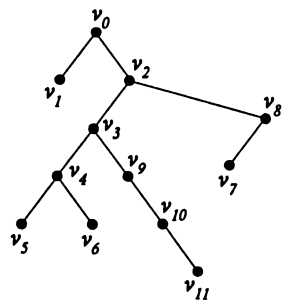
в)



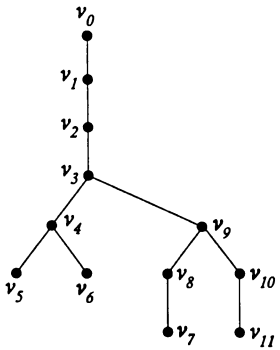
3. а)



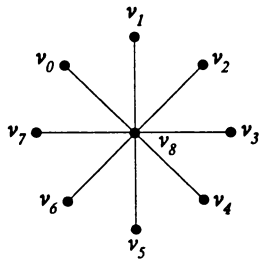
б)



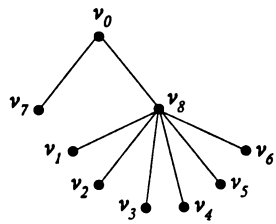
в)


v_0
v_1
v_2
v_3
v_4
v_9
v_5
v_6
v_8
v_{10}
v_7
v_{11}

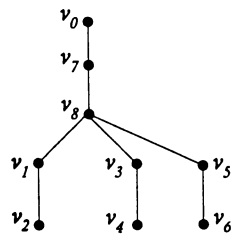
5. а)


v_1
v_0
v_2
v_7
v_8
v_3
v_6
v_4
v_5

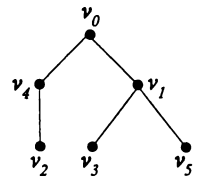
б)


v_0
v_7
v_8
v_1
v_6
v_2
v_5
v_3
v_4

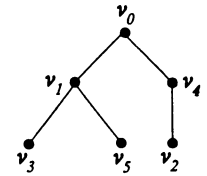
в)


v_0
v_7
v_8
v_1
v_3
v_5
v_2
v_4
v_6

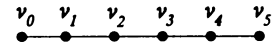
7. а)


v_0
v_4
v_1
v_2
v_3
v_5

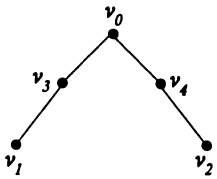
б)


v_0
v_1
v_4
v_3
v_5
v_2

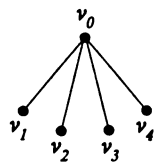
в)


v_0
v_1
v_2
v_3
v_4
v_5

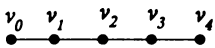
9. а)



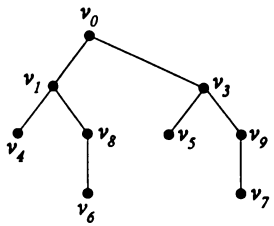
б)



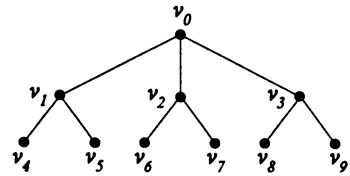
в)



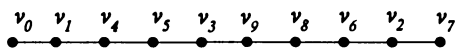
11. а)



б)



в)



13. v_8 точка — сочленения; $v_8v_6v_5, v_8v_4v_3, v_8v_1v_2, v_0v_7v_8$ — компоненты двусвязности

15. Если $v_0v_1v_2\cdots v_{k-1}$ — цикл и если начать в вершине v_0, то дерево поиска в ширину будет состоять из двух простых путей, $v_0v_1v_2\cdots$ и $v_0v_{k-1}v_{k-2}v_{k-3}\cdots$, поэтому оно будет иметь вид $\cdots v_2v_1v_0v_{k-1}v_{k-2}v_{k-3}\cdots$. Дерево поиска в глубину будет иметь вид $v_0v_1v_2\cdots v_{k-1}$. Имеется k деревьев. В каждом одно из k ребер удалено.

17. Пусть $\{a,b\}$ — ребро графа. Если это не разрезающее ребро, то после удаления $\{a,b\}$ граф остается связным. Следовательно, у графа, не содержащего $\{a,b\}$, имеется остовное дерево. Если $\{a,b\}$ — разрезающее ребро, то существуют такие вершины v_0 и v_1, что любой путь из v_0 в v_1 содержит ребро $\{a,b\}$. Поскольку v_0 и v_1 принадлежат любому остовному дереву и дерево связно, то на дереве должен существовать путь из v_0 в v_1, который должен содержать ребро $\{a,b\}$. Следовательно, ребро $\{a,b\}$ принадлежит всем остовным деревьям графа G.

19. Остовное дерево, полученное поиском в глубину для K_n, — это простой путь вида $v_1v_2\cdots v_k$ или в любом другом порядке. Поскольку из любой вершины можно попасть в любую другую вершину, то при поиске в глубину каждая вершина будет в какой-то момент выбрана. Если при поиске в ширину вершина v_i выбрана как начальная, то другие вершины являются сыновьями этой вершины, поскольку все являются смежными к ней.

21. Пусть E' — разрезающее множество графа G. Если E' удалено из графа, то граф перестает быть связным, и существуют такие вершины a и b, что путь из a в b отсутствует. Следовательно, в графе G любой путь из a в b должен содержать ребро из E'. Для любого остовного дерева графа G вершины a и b принадлежат дереву. Поскольку остовное дерево связно, то существует путь из a в b. Но этот путь должен содержать ребра из E'.

23. $n \times 2^{n-1}$.

25. Предположим, что граф G имеет $n+1$ вершин. Следовательно, T и T' — оба имеют по n ребер. Предположим, что они имеют k общих ребер, так что каждое дерево имеет $n-k$ ребер, не принадлежащих другому дереву. Пусть G' — граф, который содержит n ребер дерева T' и $n-k-1$ ребер дерева T_e, не принадлежащих T'. Граф G' содержит $n+n-k-1=2n-k-1$ ребер. При формировании остовного дерева $n-k-1$ ребер могут быть удалены из цепи. Ни одно из ребер дерева T_e не может быть удалено, поскольку в цепи всегда имеется ребро, не принадлежащее T_e. Поэтому в остовном дереве остаются все ребра дерева T_e и одно ребро из T'.

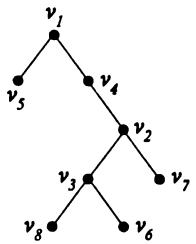
27. **а)** 1, 1, 0, 2, 2, 6, 6;

б) 2, 2, 3, 4, 5, 4, 5;

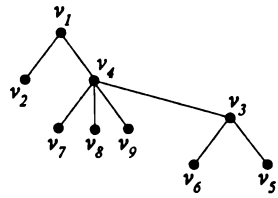
в) 2, 2, 3, 4, 5, 4, 5;

г) 0, 1, 1, 1, 0, 3, 11, 4, 14, 4, 3, 0, 5, 15, 5.

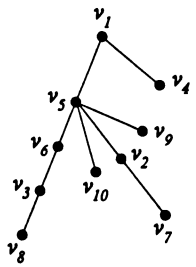
29. **а)**



б)



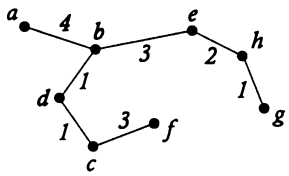
в)

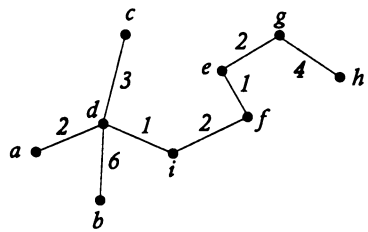


31. **а)** 4; **б)** 8; **в)** 24.

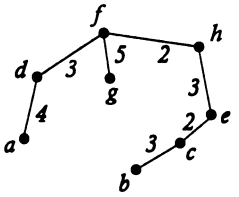
Раздел 15.6.

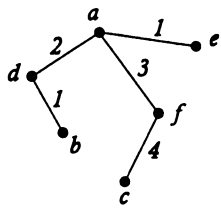
1.





3.





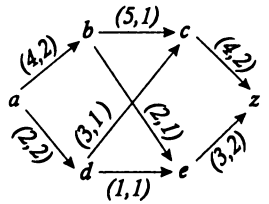
Раздел 16.1.

1. а)

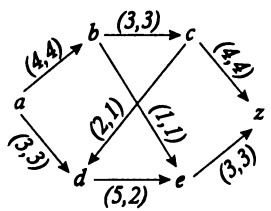
вершина	вход. поток	ребро	выход. поток	ребро
b	2	(a, b)	3	(b, c)
	1	(e, b)		
c	3	(b, c)	2	(c, z)
			1	(c, d)
d	3	(a, d)	4	(d, e)
	1	(c, d)		

б) 5; **в)** 10; **г)** 10; **д)** 11.

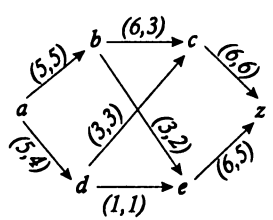
3.



5.



7.

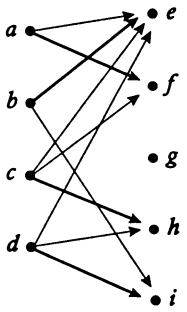


9. {a}.

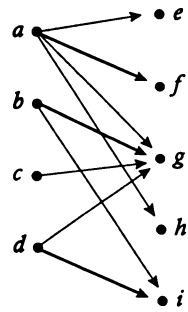
11. {a,d}.

Раздел 16.2.

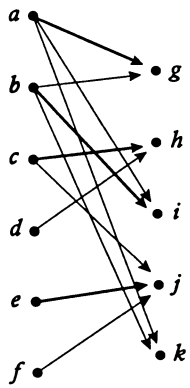
1. а)



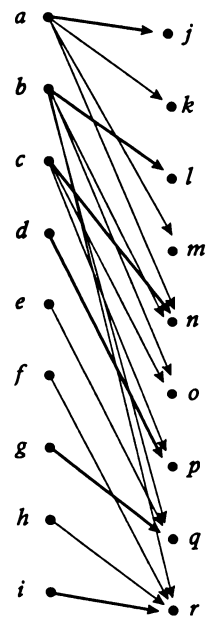
б)



в)



г)



3. Да.

5. Да.

7. Алгоритм **(A,B,M)**:
Отметить все $a \in A$, для которых имеется ребро, инцидентное к a, в паросочетании M, знаком $*$. Отметить все $b \in B$ такие, что не существует ребро, инцидентное к b, в паросочетании M, знаком $\#$.

1) Для каждого вновь отмеченного $a_i \in A$ отметить все b_j, для которых существует ребро из a_i в b_j, знаком i, если они уже не отмечены целым числом.

2) Если некоторое b_j, отмеченное знаком $\#$, теперь отмечено целым числом, то паросочетание расширено. Если никакое b_j не является вновь отмеченным, то паросочетание максимизировано.

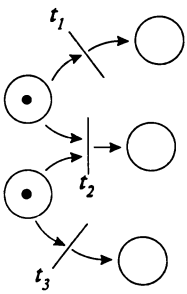
3) Для каждого вновь отмеченного $b_j \in B$, если существует a_i такое, что существует ребро из a_i в b_j в паре, отметьте его знаком j.

4) Вернуться к шагу (1).

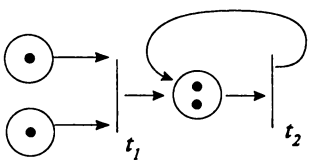
Раздел 16.3.

1. **(i)** (а), (б), (в);
(ii) (б), (в), (е);
(iii) (а), (в), (г), (е), (ж);
(iv) (а), (б), (в), (г), (д), (е), (ж);
(v) (а), (в), (г), (е), (ж)
(vi) (б), (в), (г), (д), (е), (ж);
(vii) (а), (б), (в), (г), (д), (е), (ж).

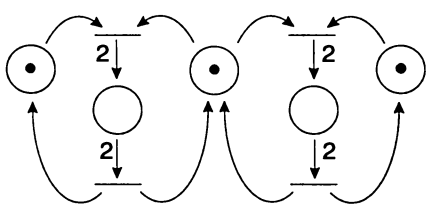
3.



5.

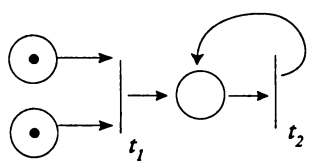


7.



9. Да.

11.



Раздел 17.1.

1. **а)** $\{aba, aca, ada\}$;
б) $\{c, abc, aabc, abbc, aabbc, ac, bc, aac, bbc, aaabbbc\}$;
в) $\{ac, ad, bc, bd\}$;
г) $\{a, ab, abb, abbb, \lambda, cd, cdcd, cdcdcd, cdcdcdcd, abbbb\}$;
д) $\{ad, abcd, abcbcd, abcbcbcd, abcbcbcd, abcbcbcbcbcd, abcbcbcbcbcbcd, abcbcbcbcbcbcbcbcd\}$.

3. **а)** $\mathbf{a(b \vee c \vee d)}$;
б) $\mathbf{(a \vee b)(b \vee c)}$;
в) $\mathbf{abab^*}$;
г) $\mathbf{(ab)^*}$;
д) $\mathbf{a(\lambda \vee b \vee a \vee ab)b}$.

5. **а)** $\mathbf{(a \vee c)^*b(a \vee c)^*b(a \vee c)^*}$;
б) $\mathbf{a^*ba^*ba^*ca^*\ ca^* \vee a^*ba^*ca^*\ ba^*ca^* \vee a^*ca^*ba^*ba^*ca^* \vee \vee a^*ca^*ca^*ba^*ba^* \vee a^*ca^*ba^*ca^*\ ba^* \vee a^*ba^*ca^*\ ca^*ba^*}$;
в) $\mathbf{(a \vee b \vee c)^*b(\ a \vee b \vee c)^*b(a \vee b \vee c)^*}$;
г) $\mathbf{a(a \vee b \vee c)^*b(a \vee b \vee c)^*c(a \vee b \vee c)^* \vee a(a \vee b \vee c)^*c(a \vee b \vee c)^*b(a \vee b \vee c)^*}$;
д) $\mathbf{aa^*bb^*cc^*}$.

7. (б), (в), (д).

9. Однозначно декодируемые коды (а), (б), (в), (г), (д).
Суффиксные коды (а), (в), (д).

11. **а)** код не является однозначно декодируемым, т.к. в 111011 за 11 следует 1011 и за 1110 следует 11;
б) однозначно декодируемый префиксный код;
в) код не является однозначно декодируемым, т.к. в 10101010 за 10101 следует 010 и за 1010 следует 1010;
г) код однозначно декодируемый. Предположим, что существует два различных способа комбинирования кода для образования одной и той же строки. Конечное слово кода одной строки должно быть заключительной подстрокой конечного слова кода другой строки. Но тогда слово кода, предшествующее самому короткому слову, должно заканчиваться на 0, что невозможно;
д) однозначно декодируемый префиксный код.

13. (а), (б).

15. (а), (б), (г).

17. Нет, 110011, 1001 является кодом, который является суффиксным и префиксным, но инфиксным не является.

19. Никакой.

21. (б).

23. Нет.

25. Пусть w и w' — элементы ключевого кода. Существует элемент алфавита, который принадлежит w и не принадлежит w'. Существует также элемент алфавита, который принадлежит w' и не принадлежит w. Следовательно, ни одна из строк не может быть начальной или конечной строкой другого элемента кода.

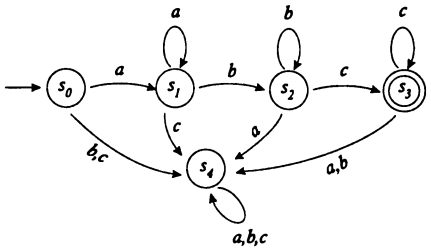
Раздел 17.2.

1. (б), (в), (г).

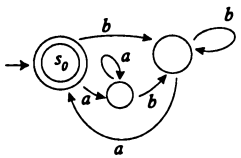
3. $\mathbf{a^*b(aa^*b)^*}$.

5. $\mathbf{(a \vee ba)^*}$

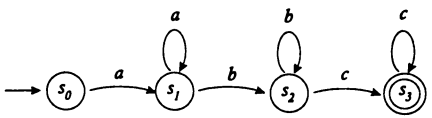
7.


a
b
c
a
b
c
s_0
s_1
s_2
s_3
b,c
c
a
a,b
s_4
a,b,c

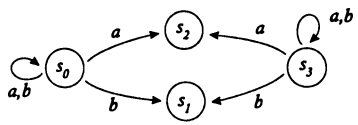
9.


b
b
s_0
a
a
b
a

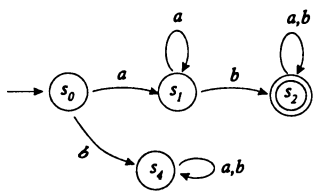
11.


a
b
c
a
b
c
s_0
s_1
s_2
s_3

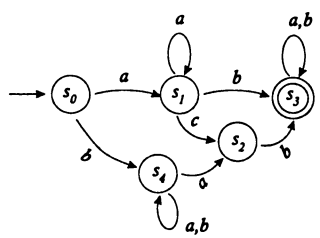
13.


a
a
a,b
s_2
s_0
s_3
a,b
b
b
s_1

15.

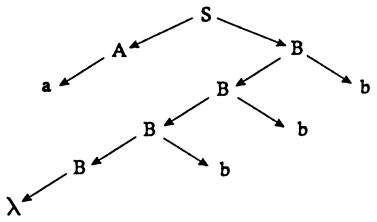

a
a,b
a
b
s_0
s_1
s_2
b
s_4
a,b

17.

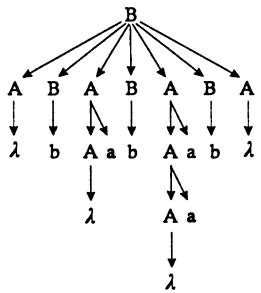

a
a,b
a
b
s_0
s_1
s_3
c
s_2
b
b
s_4
a
a,b

Раздел 17.3.

1.



3.



5. **ab(ab)***.

7. **b*a b*a (b *ab*ab*)***.

Раздел 18.2.

1. (c).

3. **а)** $G^{\perp} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$;

б) 1111110, 0101101, 1001100, 1010011;

в) 111, 110, 010, 011;

г) да, нет, нет, нет.

5. **а)** $G^{\perp} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$;

б)

000000	100011	010101	001110	110110	011011	101101	111000
100000	000011	110101	101110	010110	111011	001101	011000
010000	110011	000101	011110	100110	001011	111101	101000
001000	101011	011101	000110	111110	010011	100101	110000
000100	100111	010001	001010	110010	011111	101001	111100
000010	100001	010111	001100	110100	011001	101111	111010
000001	100010	010100	001111	110111	011010	101100	111001
100100	000111	110001	101010	010010	111111	001001	011100;

в) вместо 011001 должно стоять 011011; 111000 — правильно; вместо 110111 должно стоять 110110; вместо 101100 должно стоять 101101.

7. а) $G^{\perp} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$;

б)

000000	0 0 0	100101 010011 001111 110110 011100 101010 111001
100000	1 0 1	000101 110011 101111 010110 111100 001010 011001
010000	0 1 1	110101 000011 001111 100110 001100 111010 101001
001000	1 1 1	101101 011011 000111 111110 010100 100010 110001
000100	1 0 0	100001 010111 001011 110010 011000 101110 111101
000010	0 1 0	100111 010001 001101 110100 011110 101010 111001
000001	0 0 1	100100 010010 001110 110111 011101 101011 111000
110000	1 1 0	010101 100011 111111 000110 101100 011010 001001

в) вместо 111101 должно стоять 111001; 111001 — правильно; вместо 110010 должно стоять 110110; вместо 101001 должно стоять 111001.

9. 101101, 011011, 110110;

$$G^{\perp} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

11. Ассоциативность. Пусть $a = (a_1, a_2, \cdots, a_n)$, $b = (b_1, b_2, \cdots, b_n)$ и $c = (c_1, c_2, \cdots, c_n)$ — элементы из B_n

$$\begin{aligned}(a+b)+c &= ((a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n)) + (c_1, c_2, \cdots, c_n) = \\ &= (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n) + (c_1, c_2, \cdots, c_n) = \\ &= ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2, \cdots, (a_n + b_n) + c_n) = \\ &= (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2), \cdots, a_n(+b_n + c_n)) = \\ &= (a_1, a_2, \cdots, a_n) + (b_1 + c_1, b_2 + c_2, \cdots, b_n + c_n) = \\ &= (a_1, a_2, \cdots, a_n) + ((b_1, b_2, \cdots, b_n) + (c_1, c_2, \cdots, c_n)) = \\ &= a + (b + c).\end{aligned}$$

Единицей является элемент $(0, 0, \cdots, 0)$. Каждый элемент является своим обратным элементом. Замкнутость следует из определения сложения. Следовательно, B_n — группа.

13. Поскольку $C^{\perp}$ — подмножество множества B_n, оно ассоциативно. Каждый элемент является своим обратным элементом. Поэтому необходимо показать только замкнутость. Пусть $u, v \in C^{\perp}$. Тогда $u \circ c = v \circ c = \mathbf{0}$ для всех $c \in C$. Поэтому $(u+v) \circ c = u \circ c + v \circ c = \mathbf{0} + \mathbf{0} = \mathbf{0}$ для всех $c \in C$ и $u + v \in C^{\perp}$. Следовательно, $C^{\perp}$ — группа.

Раздел 18.3.

1. $G_H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$.

3. $\delta(110010101, 010101111) = 5$.

5. 001100110, 101100111, 100000111.

7. **а)** $G_H^{\perp} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$;

б) вместо 1110110 должно быть 1110010, вместо 1011001 должно быть 1010001, вместо 1101100 должно быть 1100100, и вместо 1111110 должно быть 1111111.

9. Порождающая функция в упражнении 7 меняет 1001001 на 1101001. Порождающая функция в упражнении 8 меняет 1001001 на 1011001. Обе порождающие функции "исправляют" код, генерируя ошибки.

11. **а)** если $c = c'$, то нет такой позиции, в которой соответствующие биты отличались бы. Следовательно, отсутствуют позиции, в которых один бит равен 1, а другой равен 0. Таким образом, $\delta(c, c') = 0$. Обратно, если $\delta(c, c') = 0$, отсутствует позиция, в которой один бит равен 1, а соответствующий бит равен 0. Вполне очевидно, что все соответствующие биты должны совпадать, и $c = c'$.

б) пусть $\delta(c, c') = n$. В таком случае имеется n позиций, где соответствующие биты строк c и c' отличаются. Но тогда But имеется n позиций, где соответствующие биты строк c' и c отличаются, и $\delta(c', c)$=n.

13.

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix};$$

$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix};$$

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix};$$

$$H_{16} = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1
\end{bmatrix}$$

Раздел 19.1.

1. Пусть

$$I = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}; \quad \rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}; \quad \rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix};$$

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

В таком случае имеем следующую таблицу умножения.

	I	ρ_1	ρ_2	σ_1	σ_2	σ_3
I	I	ρ_1	ρ_2	σ_1	σ_2	σ_3
ρ_1	ρ_1	ρ_2	I	σ_3	σ_1	σ_2
ρ_2	ρ_2	I	ρ_1	σ_2	σ_3	σ_1
σ_1	σ_1	σ_2	σ_3	I	ρ_1	ρ_2
σ_2	σ_2	σ_3	σ_1	ρ_2	I	ρ_1
σ_3	σ_3	σ_1	σ_2	ρ_1	ρ_2	I

3.

$$I = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{pmatrix};$$

$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 1 \end{pmatrix};$$

$$\rho_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 5 & 6 & 7 & 8 & 1 & 2 \end{pmatrix};$$

$$\rho_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 6 & 7 & 8 & 1 & 2 & 3 \end{pmatrix};$$

$$\rho_4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 8 & 1 & 2 & 3 & 5 \end{pmatrix};$$

$$\rho_5 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 7 & 8 & 1 & 2 & 3 & 4 & 5 \end{pmatrix};$$

$$\rho_6 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 8 & 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix};$$

$$\rho_7 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix};$$

$$\delta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 1 & 8 & 7 & 6 & 5 & 4 & 3 \end{pmatrix};$$

$$\delta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 3 & 2 & 1 & 8 & 7 & 6 & 5 \end{pmatrix};$$

$$\delta_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 5 & 4 & 3 & 2 & 1 & 8 & 7 \end{pmatrix};$$

$$\delta_4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix};$$

$$\eta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 8 & 7 & 6 & 5 & 4 & 3 & 2 \end{pmatrix};$$

$$\eta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 2 & 1 & 8 & 7 & 6 & 5 & 4 \end{pmatrix};$$

$$\eta_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 4 & 3 & 2 & 1 & 8 & 7 & 6 \end{pmatrix};$$

$$\eta_4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 & 8 \end{pmatrix}.$$

5. 8.

7. 7.

Раздел 19.2.

1. 8.

3. 4.

5. 18.

7. 39.

9. $r^4 + r^3b + 3r^2b^2 + rb^3 + b^4$.

11. $r^5 + r^4b + 2r^3b^2 + 2r^2b^3 + rb^4 + b^5$.

13. $r^8 + r^7b + 4r^6b^2 + 5r^5b^3 + 8r^4b^4 + 5r^3b^5 + 4r^2b^6 + rb^7 + b^8$.

15. $r^{10} + 4r^9b + 15r^8b^2 + 42r^7b^3 + 72r^6b^4 + 84r^5b^5 + 72r^4b^6 + 42r^3b^7 + 15r^2b^8 + 4rb^9 + b^{10}$.

Раздел 20.1.

1. Если $ax = b$, то $a^{-1}ax = a^{-1}b$ и $x = a^{-1}b$. Если $am = b$ и $an = b$, то $am = an$, поэтому $a^{-1}am = a^{-1}an$ и $m = n$.

3. Предположим, что ab — делитель нуля. Тогда $(ab)c = 0$ для некоторого c. Следовательно, $a(bc) = 0$. Если $bc = 0$, то b — делитель нуля. Если $bc \neq 0$, то a — делитель нуля.

5. 0.

7. Предположим, что из $ab = ac$ следует, что $b = c$ для всех ненулевых a из R. Если $ad = 0 = a0$ и $a \neq 0$, то $d = 0$, и R — область целостности. Обратно, если R — область целостности и $ab = ac$ для $a \neq 0$, то $a(b - c) = 0$. Следовательно, $b - c = 0$ и $b = c$.

9. Пусть $i + j$ и $i' + j'$ принадлежат $I + J$. Тогда $i + j + i' + j' = i + i' + j + j'$. Поскольку $i + i' \in I$ и $j + j' \in J$, то $i + j + i' + j' \in I + J$. Также $a(i + j) = ai + aj$. Поскольку $ai \in I$ и $aj \in J$, то $a(i + j) \in I + J$.

11. $\langle 3 \rangle$.

13. $f(0) = 0$. Поэтому $0 \in f(R)$. Пусть $a, b \in f(R)$. Тогда $a = f(c)$ и $b = f(d)$ для некоторых $c, d \in R$. Поэтому $c - d \in R$ и $f(c - d) = f(c) - f(d)$. Следовательно, $f(c) - f(d) \in f(R)$. Аналогично, если $a, b \in f(R)$, $a = f(c)$ и $b = f(d)$ для некоторых $c, d \in R$, то $cd \in R$ и $f(cd) = f(c)f(d)$. Таким образом, $ab \in f(R)$. Следовательно, $f(R)$ — кольцо.

15. Нет.

17. Неверное утверждение. Элементы [2] и [3] — нулевые делители кольца Z_{12}. Однако, [2] + [3] = [5] не является нулевым делителем кольца Z_{12}.

19. Пусть R — область целостности. Если $a^2 = a$, то $a \cdot a = a \cdot 1$ и, согласно результату предыдущей задачи, $a = 1$.

21. Для всех $a, b \in R$ имеем $(-a)(b) + ab = (-a + a)b = 0 \cdot b = 0$. Поэтому $(-a)(b) = -ab$. Аналогично, $a(-b) = -ab$. Для всех $a, b \in R$

$$(-a)(-b) + (-a)(b) = (-a)(-b + b) = (-a) \cdot 0 = 0.$$

Поэтому $(-a)(-b) + (-a)(b) = ab + (-a)(b)$ и $(-a)(-b) = ab$.

Раздел 20.2.

1. Покажем сначала, что пересечение подколец с единицей является кольцом с единицей. Поскольку мультипликативная единица принадлежит всем подкольцам, то она принадлежит их пересечению. Пусть a, b принадлежит пересечению подколец, тогда a, b принадлежит каждому из подколец. Поэтому $a - b$ принадлежит всем подкольцам и, следовательно, пересечению подколец. Аналогично, ab принадлежит пересечению всех подколец. Следовательно, пересечение подколец с единицей есть кольцо с единицей. Далее покажем, что пересечение областей целостности есть область целостности. Предположим, что a и b принадлежат пересечению всех подколец, $ab = 0$ и $a \neq 0$. Но a и b принадлежат каждой из областей целостности. Поэтому $b = 0$.

3. Или $a > 0$, или $-a > 0$. Если $a > 0$, то, согласно определению упорядочения, $a^2 > 0$. Если $-a > 0$, то $(-a)^2 > 0$. Но $(-a)^2 = a^2$, поэтому $a^2 > 0$. Поскольку $1^2 = 1$, то $1 > 0$.

Раздел 20.3.

1. Пусть $f = (a_i)^*$, $g = (b_i)^*$ и $h = (c_i)^*$.

$$\begin{aligned} f + (g + h) &= (a_i)^* + (b_i + c_i)^* = \\ &= (a_i + (b_i + c_i))^* = \\ &= ((a_i + b_i) + c_i)^* = \\ &= (a_i + b_i)^* + (c_i)^* = \\ &= (f + g) + h, \end{aligned}$$

поэтому сложение ассоциативно. Аддитивной единицей является $(d_i)^*$, где $d_i = 0$ для всех i. Обратным элементом относительно сложения для $(a_i)^*$ будет $(-a_i)^*$.

$$\begin{aligned} f + g &= (a_i)^* + (b_i)^* = \\ &= (a_i + b_i)^* = \\ &= (b_i + a_i)^* = \\ &= (b_i)^* + (a_i)^* = \\ &= g + f, \end{aligned}$$

и сложение коммутативно. Следовательно, множество многочленов — коммутативная группа относительно сложения.

$$\begin{aligned} f(gh) &= (a_i)^*((b_i)^*(c_i)^*) = \\ &= (a_i)^*(\sum_{i+j=k} b_i c_j)^* = \\ &= (\sum_{m+k=n} a_m \sum_{i+j=k} b_i c_j)^* = \\ &= (\sum_{m+i+j=n} a_m b_i c_j)^* = \\ &= (\sum_{s+j=n} (\sum_{m+i=s} a_m b_i) c_j)^* = \\ &= (\sum_{m+i=s} a_m b_i)^*(c_j)^* = \\ &= ((a_i)^*(b_i)^*)(c_i)^* = \\ &= (fg)h, \end{aligned}$$

и умножение ассоциативно. Окончательно,

$$\begin{aligned} f.(g+h) &= (a_i)^*(b_i + c_i)^* = \\ &= (\sum_{i+j=k} a_i(b_j + c_j))^* = \\ &= (\sum_{i+j=k} (a_i b_j + a_i c_j))^* = \\ &= (\sum_{i+j=k} a_i b_j)^* + (\sum_{i+j=k} a_i c_j{}^* = \\ &= fg + fh. \end{aligned}$$

3. Неверное утверждение. Пусть $f = (1,1,1,1,0,0,\cdots)$ и $g = (1,1,-1,-1,0,0,\cdots)$. Тогда $f+g = (1,1,0,0,0,0,\cdots)$. $\deg(f) = \deg(g) = 3$, тогда как $\deg(f+g) = 1$.

5. В Z_5, $f(x) = x^4 - 1$ — нулевая функция, тогда как f имеет степень 4.

Раздел 20.4.

1. Пусть $r = a+bi$ и $s = c+di$ — целые комплексные числа. Тогда $r-s = (a-c)+(b-d)i$ и $rs = (a+bi)(c+di) = (ac-bd)+(ad+bc)i$ — также целые комплексные числа. Поэтому множество целых комплексных чисел является подкольцом.

3. $\sin(n\pi) = 0$ для любого целого числа n. Следовательно, уравнение имеет бесконеное количество нулей, и $\sin(x)$ нельзя представить в виде многочлена.

5. Пусть $A = Z$ и $f, g \in A[x]$ определены соотношениями $f(x) = x^2 + 1$ и $g(x) = x^2 - 3$ соответственно, тогда f и g — простые многочлены, но $f+g$ простым многочленом не является, поскольку $f + g(x) = 2x^2 - 2$.

7. Пусть $c = a + bi$ — комплексное целое число. Тогда

$$\begin{aligned} f(x) &= (x - (a+bi))(x - (a-bi)) = \\ &= (x-a) + bi)(x-a) - bi) = \\ &= (x-a)^2 + b^2, \end{aligned}$$

поэтому $f \in Q[x]$ и $f(c) = 0$.

9. Подкольцо, состоящее из всех чисел вида $a + b\sqrt{5}$, где a и b — рациональные числа.

11. Подкольцо, состоящее из всех чисел вида $a + b\sqrt{5}i$, где a и b — рациональные числа.

13. Подкольцо, состоящее из всех чисел вида $a + b\sqrt{3}i$, где a и b — рациональные числа.

Раздел 21.2.

1. $1 = \chi(gg^{-1}) = \chi(g)\chi(g^{-1}) = (a + bi)\chi(g^{-1})$. Следовательно, $\chi(g^{-1}) = \frac{1}{a+bi} = \frac{a-bi}{a^2+b^2}$. Но поскольку $a + bi$ лежит на единичной окружности, то $a^2 + b^2 = 1$, поэтому $\chi(g^{-1}) = a - bi$.

3. Если $g \neq g'$, то $g_1^{j(1)} g_2^{j(2)} g_3^{j(3)} \cdots g_m^{j(m)}$ и $g' = g_1^{j'(1)} g_2^{j'(2)} g_3^{j'(3)} \cdots g_m^{j'(m)}$, где для некоторого i имеем $g_i^{j(i)} \neq g_i^{j'(i)}$. Тогда $\chi_g(g_i) = a_i^{j(i)}$ и $\chi_{g'}(g_i) = a_i^{j'(i)}$. Поскольку $a_i^{j(i)} \neq a_i^{j'(i)}$, то $\chi_g(g_i) \neq \chi_{g'}(g_i)$ и $\chi_g \neq \chi_{g'}$.

Раздел 21.3 .

1. Пусть $i \in I$ и $s \in S$. $\chi(i \cdot s) = \chi(i) \cdot \chi(s) = 0 \cdot \chi(s) = 0$ и $i \cdot s \in I$. Следовательно, I — идеал. Пусть s, $s' \in S - I$. Тогда $\chi(ss') = \chi(s)\chi(s') = 1 \cdot 1 = 1$. Таким образом, $ss' \in S - I$ и S принадлежат идеалу. Следовательно, I — простой идеал. Обратно, если I — простой идеал, то искомый характер может быть получен путем отображения. Если χ — отображение из

$$\chi(s) = \begin{cases} 1, & \text{если } s \in S - I; \\ 0, & \text{если } s \in I. \end{cases}$$

Если χ — характер на $S - I$, то определим χ' на S соотношениями

$$\chi'(s) = \begin{cases} \chi(s), & \text{если } s \in S - I; \\ 0, & \text{если } s \in I. \end{cases}$$

3. Пусть $C = [C_{ij}] = \sigma\sigma'$. Тогда

$$\begin{aligned} C_{ij} = \sum_{k=1}^{n} \sigma_{ik}\sigma' &= \sum_{k=1}^{n} \sigma_{ik}\overline{(\sigma_{jk})} = \\ &= \sum_{k=1}^{n} \chi_i(g_k)\overline{(\chi_j(g_k))} = \\ &= \sum_{k=1}^{n} \chi_i(g_k)(\chi_j(g_k))^{-1} = \\ &= n\delta_{ij} \end{aligned}$$

согласно теореме 21.19.

Раздел 22.1.

1. Пусть $N(W) = x_1 2^{n-1} + x_2 2^{n-2} + \cdots + x_{n-1} 2^1 + x_n$ и $N(V) = y_1 2^{n-1} + y_2 2^{n-2} + \cdots + y_{n-1} 2^1 + y_n$. Тогда

$$\begin{aligned} |N(W) - N(V)| &\leq |x_1 - y_1|2^{n-1} + |x_2 - y_2|2^{n-2} + \cdots \\ &\qquad \cdots + |x_{n-1} - y_{n-1}|2^1 + |(x_n - y_n)| = \\ &= \sum_{i=1}^{n} |x_i - y_i| 2^{n-i} \leq \sum_{i=1}^{n} 2^{n-i} = \\ &= 2^{n-1}. \end{aligned}$$

3. $w = 7, 8, 11, 12, 13, 14, 15, 16$.

5. k слов для хранения k простых чисел;
k слов для хранения $N_p(X)$ для k простых чисел p;
k слов для хранения $N_p(Y(i))$ для k простых чисел p;
k слов для хранения $[[2^{n-1}]]_p$;
незначительное количество слов для циклов и т.п.;
всего приблизительно $4k$ слов.

Раздел 22.2.

1. Если $d|a$ и $d|b$, то $d|a-b$. Поэтому $d|\,\text{НОД}(a-b,a)$ и $\text{НОД}(a,b)|\,\text{НОД}(a-b,a)$. Обратно, если $d|a$ и $d|a-b$, то $d|b$, и, следовательно, $d|\,\text{НОД}(a,b)$. Таким образом, $\text{НОД}(a-b,a)|\,\text{НОД}(a,b)$. Значит, $\text{НОД}(a-b,a) = \text{НОД}(a,b)$.

3. $\text{НОД}(6,10,15) = 1$, однако $\text{НОД}(6,10) \neq 1$. Фактически, наибольший общий делитель любой пары этих целых чисел не равен единице.

5. **а)**

i	k_i	m_i	M_i	b_i	N_i	$M_ib_iN_i$
1	12	145	26551252	532	0	0
2	14	169	22780660	5	43	4897841900
3	15	181	21270340	711	91	1376212268340
4	18	217	17741620	81	163	234242608860

$D = 12$; $C = 2231403840$;

$$f(12) = \left[\left[\left\lfloor \frac{2231403840}{12\times 12+1} \right\rfloor\right]\right] = 0;$$
$$f(14) = \left[\left[\left\lfloor \frac{2231403840}{12\times 14+1} \right\rfloor\right]\right] = 1;$$
$$f(15) = \left[\left[\left\lfloor \frac{2231403840}{12\times 15+1} \right\rfloor\right]\right] = 2;$$
$$f(18) = \left[\left[\left\lfloor \frac{2231403840}{12\times 18+1} \right\rfloor\right]\right] = 3.$$

7. **а)**

i	k_i	m_i	M_i	b_i	N_i	$M_ib_iN_i$
1	5	301	6496934404	1066	0	0
2	15	901	2170452004	232	226	113801139473728
3	20	1201	1628290804	379	601	370890451044316
4	25	1501	1302849604	5832	1126	8555594470734528

$D = 60$; $C = 1607985850884$;

$$f(5) = \left[\left[\left\lfloor \frac{1607985850884}{60\times 5+1} \right\rfloor\right]\right] = 0;$$
$$f(15) = \left[\left[\left\lfloor \frac{1607985850884}{60\times 15+1} \right\rfloor\right]\right] = 1;$$
$$f(20) = \left[\left[\left\lfloor \frac{1607985850884}{60\times 20+1} \right\rfloor\right]\right] = 2;$$
$$f(25) = \left[\left[\left\lfloor \frac{1607985850884}{60\times 25+1} \right\rfloor\right]\right] = 3.$$

Раздел 22.3.

1.

$$[[1683^{217}]]_{2993} = 77;$$
$$[[79^{217}]]_{2993} = 79;$$
$$[[2560^{217}]]_{2993} = 78;$$
$$[[872^{217}]]_{2993} = 69;$$
$$[[2571^{217}]]_{2993} = 89.$$

3. Для шифрования необходимо иметь значения e и n. Если известно $\phi(n)$, то, используя формулы из теоремы 22.15, можно найти d, что обеспечит возможность дешифровки.

Предметно-именной указатель

А

Абсолютное значение — Absolute value, 820
Автомат — Automaton, 731
  детерминированный — deterministic, 735
  диаграмма состояний — state diagram, 732
  допускаемый язык — accepted language, 731
  допускание — accepting, 733
  недетерминированный — nondeterministic, 735
  состояние зацикливания — sink state, 734
  функция переходов — transition function, 731
Адлеман, Л. М. (Adleman, L. M.), 844
Аксиоматическая система — Axiom system, 67
  Гильберта-Бернайса-Геделя — Gödel-Hilbert-Bernays, 67
  Рассела-Уайтхеда — Russell-Whitehead, 67
  Цермело-Френкеля-фон Неймана — Zermelo-Fraenkel-von Neumann, 67
  равенство — equality, 123
  целые числа — integers, 124
Алгебраический элемент — Algebraic element, 815
Алгебраическое число — Algebraic number, 815
  целое — Algebraic integer, 815
Алгоритм — Algorithm, 184, 298
  Горнера — Horner's algorithm, 185
  Дейкстры — Dijkstra's algorithm, 612, 615
  Евклида — Euclidean algorithm, 302
  Крускала — Krushkal's algorithm, 682
  Прима — Prim's algorithm, 684, 686
  Флойда-Уоршолла — Floyd-Warshall algorithm, 619, 621
  Форда-Фалкерсона — Ford-Fulkerson algorithm, 699
  Хаффмана — Huffman's algorithm, 640, 642
  быстрой сортировки — quick sort, 208
  деления — Division algorithm, 140, 301
  перевода последовательности в дерево — Sequence to tree algorithm, 672
  поиска дерева в глубину — Depth-first tree search, 661
  поиска дерева в ширину — Breadth-first tree search, 659
  преобразования дерева в последовательность — Tree to sequence algorithm, 670
  пузырьковой сортировки — bubble sort, 207
  сортировки — sorting algorithm, 205
  сортировки вставками — insertion sort, 211
  сортировки выбором — selection sort, 206
  сортировки слиянием — merge sort, 209
Алфавит — Alphabet, 399, 725

Б

Бернулли, Иоганн (Bernoulli, Jean), 438
Бесконечный спуск — Infinite descent, 139
Бинарная операция — Binary operation, 163, 397

Бинарное дерево поиска — Binary search tree, 631
обход — traversing, 649
Биномиальная теорема — Binomial theorem, 337
Биномиальное распределение — Binomial distribution, 375
Битовая строка — Bit String, 318
Блоковый код — Block code, 753
Боер, Р. С. (Boyer, R. S.), 836
Булева алгебра — Boolean algebra, 84, 406
дополнение законов тождества — complement of identities laws, 87
законы ассоциативности — associative properties, 85
законы де Моргана — De Morgan's laws, 87
законы дистрибутивности — distributive properties, 85
законы дополнения — complement properties, 85
законы идемпотентности — idempotent laws, 86
законы инволюции — involution laws, 87
законы коммутативности — commutative properties, 85
законы поглощения — absorption laws, 86
законы тождества — identity properties, 85
свойства констант — null laws, 86
Бурали-Форти (Burali-Forti), 67

В

Вероятность — Probability, 348, 369, 833
биномиальное распределение — binomial distribution, 375
математическое ожидание — mathematical expectation, 376
множество элементарных событий — sample space, 347
независимые события — independent events, 374
непересекающиеся множества — disjoint set, 348
ожидаемое значение — expected value, 376
среднее значение — mean, 379
условная — conditional, 370
Взаимоисключающие множества — Mutually exclusive sets, 110
Включение-исключение — Inclusion-exclusion, 502
Внутреннее произведение — Inner product, 169
Вполне упорядоченная область целостности — Well ordered integral domain, 798
Выводимость — Deducibility, 34
Выражение — Expression, 726
регулярное — regular, 726
Высказывание — Proposition, 15
Вычет — Residue, 153

Г

Гамильтонов — Hamiltonian
путь — path, 601
цикл — cycle, 601
Гауссово кольцо — Unique factorization domain, 797
Гиперкуб — Hypercube, 290
Гипотеза — Hypothesis, 35
Гомоморфизм — Homomorphism
колец — ring homomorphism, 790
полугрупп — semigroup homomorphism, 415
Грамматика — Grammar, 741
контекстная — context-sensitive, 749
начальный символ — start symbol, 742
непосредственных составляющих — phrase structure, 742
нетерминальные символы — nonterminal symbols, 741, 742
порождаемый язык — language generated by, 742
продукции — productions, 741
регулярная — regular, 751
соответствующее дерево — corresponding tree, 747
терминальные символы — terminal symbols, 741, 742
формальная — formal, 742
Граф — Graph, 95, 244, 556
Петерсена — Petersen graph, 584
вершина — vertex, 95, 244
взвешенный — weighted, 611
гомеоморфный — homeomorphic, 557

гомоморфизм — homomorphism, 556
грань — face, 581
двойственный — dual, 587
двудольный — bipartite, 250
двусвязный — biconnected, 562
диаметр — diameter, 290
дополнение — complement, 559
изоморфизм — isomorphism, 556
компонента — component, 250
компонента двусвязности — bicomponent, 563
матрица инцидентности — incidence matrix, 278
матрица смежности — adjacency matrix, 279
мультиграф — multigraph, 245
объединение — union, 559
остовное дерево — spanning tree, 560
остовной граф — spanning graph, 559
пересечение — intersection, 559
планарный граф — planar graph, 580
полный — complete, 250
полный двудольный — complete bipartite, 251
попарно непересекающиеся — pairwise disjoint, 559
производный — derived, 557
простой путь — simple path, 248
простой цикл — simple cycle, 250
путь — path, 248
разрезающее множество — cut set, 561
разрезающее ребро — cut edge, 561
раскраска — coloring, 589
расстояние — distance, 290
расширение — extension, 557
ребро — edge, 95, 244
связный — connected, 249
сетка — grid, 295
смежные вершины — adjacent vertices, 244
смежные ребра — adjacent edges, 244
степень вершины — degree of vertex, 247
точка сочленения — articulation point, 561
хроматический многочлен — chromatic polynomial, 589
хроматическое число — chromatic number, 589
цикл — cycle, 250
эйлеров цикл — Euler path, 270
Группа — Group, 410
абелева — abelian, 410
группа симметрий — symmetric group, 419
единичный элемент — identity, 410
коммутативная — commutative, 410
конечная циклическая — finite cyclic group, 412
левый смежный класс — left coset, 413
нормальная подгруппа — normal subgroup, 417
обратный элемент — inverse, 410
подгруппа — subgroup, 412
порядок — order, 410
порядок элемента — order of an element, 411
прямая сумма — direct sum, 821
симметрий квадрата — symmetries of the square, 776
смежный класс — coset, 413
фактор-группа — factor group, 418
характер — character, 821
циклическая — cyclic group, 412
ядро — kernel, 417

Д

Деление
остаток — remainder, 140
частное — quotient, 140
Делитель единицы — Unit, 795
Дерево — Tree, 259, 624
m-арное — m-ary, 263
бинарный поиск — binary search, 631
братья — sibling, 263
вес — weight, 639
внутренняя вершина — internal vertex, 261
высота — height, 262
корень — root, 262
корневое — rooted, 262
корневое ориентированное — rooted directed, 262
лес — forest, 259
листья — leaves, 261
ориентированное — directed, 260, 624
остовное — spanning tree, 264, 658
подсчета — counting tree, 316

поиск в глубину — depth-first search, 661
поиск в ширину — breadth-first search, 659
потомок — descendent, 263
предок — ancestor, 263
родитель — parent, 262
сын — child, 262
уровень — level, 262
Детерминированный автомат — Deterministic automata, 735
Дешифровка — Decryption, 844
Диаграмма — Diagram
Венна — Venn diagram, 77
Ферре — Ferrer's diagram, 545
Эйлера — Euler diagram, 119
Дизъюнктивная нормальная форма — Disjunctive normal form, 47
Диффи, У. (Diffie, W.), 847
Доказательство — Proof
контрпример — counterexample, 129
перебором — case proof, 128
приведение к абсурду — reductio as absurdum, 127

Е

Евклид (Euclid)
теорема о простых числах — theorem on primes, 145
Евклидово кольцо — Euclidean domain, 809
норма — norm, 809
Единичная окружность — Unit circle, 820

З

Запрещенные позиции — Forbidden positions, 509
Зашифрованный текст — Ciphertext, 844

И

Идеал — Ideal, 793, 825
максимальный — maximal, 810
простой — prime, 795
Идемпотент — Idempotent, 401
Идентифицирующая функция — Fingerprint, 831
Изоморфизм — Isomorphism
колец — ring isomorphism, 790
полугрупп — semigroup isomorphism, 415
Импликация — Conditional, 24, 33
Инверсия — Inverse, 33
Инверсная польская запись — Reverse Polish notation, 651, 653
Индукция — Induction, 136, 137, 799
Инфиксная запись — Infix notation, 215
Инфиксный код — Infix code, 731
Испытания Бернулли — Bernnoulli trials, 377

К

Кан, Дэвид (Kahn, David), 849
Кантор, Георг (Cantor, Georg), 67
Карп, Р. М. (Karp, R. M.), 830
Карта Карно — Karnaugh map, 50
Квантор — Quantifier, 114, 115
существования — existential, 115
универсальный — universal, 114
Китайская теорема об остатках — Chinese Remainder Theorem, 429
Класс смежности — Coset, 414
левый — left, 811
Ключ — Key
криптографический — cryptographic, 843
общий шифрующий — enciphering, 847
секретный дешифрирующий — deciphering, 847
Кнут, Д. Э. (Knuth, D. E.), 836, 837
Код — Code, 638, 727
ASCII — ASCII code, 756
RSA — RSA code, 845
Адамара — Hadamard code, 773
Голе — Golay code, 771
Грея — Gray code, 294, 755
Морзе — Morse code, 754
Рида-Мюллера — Reed-Muller code, 772
Хаффмана — Huffman code, 754
Хаффмана — Huffman's code, 640
Хемминга — Hamming code, 768
блоковый — block code, 753
вес — weight, 639, 758
групповой — group code, 757
двойственный — dual code, 761
инфиксный — infix code, 731
исправляющий ошибки — error correcting code, 755

ключевой — key code, 728
кома-код — comma code, 754
линейный — linear code, 757
мгновенный — instantaneous code, 754
обнаруживающий ошибки — error detecting code, 754
однозначно декодируемый — uniquely decipherable, 638, 727
префиксный — prefix, 638
префиксный — prefix code, 728, 754
суффиксный — suffix code, 728
Кольцо — Ring, 788
главный идеал — principal ideal, 794
группа делителей единицы — group of units, 795
делитель единицы — unit, 795
идеал — ideal, 793
коммутативное — commutative, 789
неприводимый элемент — irreducible element, 795
подкольцо — subring, 791
полиномов — ring of polynomials, 802
с единицей — ring with unity, 789
упорядоченное — ordered, 798
Комбинаторный принцип
сложения — Addition counting principle, 319, 324
умножения — Multiplication counting principal, 317
Коммутационная схема — Circuit Diagram, 56
инвертор — negation gate, 58
логический элемент *и* — and gate, 58
логический элемент *или* — or gate, 58
логический элемент *не* — negation gate, 58
логический элемент *не-и* — nand gate, 60
логический элемент *не-или* — nor gate, 60
полный сумматор — full-adder, 62
полусумматор — half-adder, 61
схема переключения — switching circuit, 57
Комплексная плоскость — Complex plane, 820
Комплексное число — Complex number, 819
абсолютное значение — absolute value, 820
действительная часть — real part, 820
длина — length, 820
мнимая часть — imaginary part, 820
сопряженное — conjugate, 820
Конверсия — converse, 33
Конечные разности — Finite differences, 469
правило произведения — product rule, 471
правило частного — quotient rule, 471
суммирование — summation, 483
суммирование по частям — summation by parts, 485
факториальный многочлен — factor polynomials, 473
Конкатенация — Concatenation, 399, 725
Константа — Constant, 802
Контрапозиция — Contrapositive, 29, 33
Конъюнктивная нормальная форма — Conjunctive normal form, 48
Корректирующая формула — Update formula, 832
Криптоанализ — Cryptanalysis, 844
Криптография — Cryptography, 844
Криптосистемы с ключами общего пользования — Public-key cryptosystems, 847
Критерий простоты Лукаса — Lucas' Primality Test, 443

Л

Ладейный полином — Rook polynomial, 509
Лексикографический порядок — Lexicograhical ordering, 343
Лемма Бернсайда о подсчете — Burnside counting lemma, 780
Линейное рекуррентное отношение — Linear recurrence relation, 448, 451
неоднородное — nonhomogeneous, 460
одноролное — homogeneous, 449
с постоянными коэффициентами — constant coefficients, 449
Линейные уравнения — Linear equations, 422
целочисленные решения — integral solutions, 422
Логическая истинность — Logical truth, 30
Логическая ложность — Logical falsity, 30

Логическая эквивалентность — Logical equivalence, 27, 28
закон двойного отрицания — double negation, 29
закон идемпотентности — idempotent law, 29
закон контрапозиции — contrapositive, 29
законы де Моргана — de Morgan's Laws, 29
свойства ассоциативности — associative properties, 29
свойства дистрибутивности — distributive properties, 29
свойства коммутативности — commutative properties, 29
Логическое умозаключение — Logical argument, 35

M

Максимальное паросочетание — Maximal matching, 707
Максимальный идеал — Maximal ideal, 810
Малая теорема Ферма — Fermat's Little Theorem, 440
Матиас (Matyas, S. M.), 848
Матрица — Matrix, 167, 186
Адамара — Hadamard matrix, 773
Хемминга — Hamming matrix, 767
булева — boolean, 172
детерминант — determinant, 238
единичная — identity, 240
контроля четности — parity check matrix, 762
матричное представление — matrix representation, 171
обратная — inverse matrix, 240
перестановок — permutatiom, 174
порождающая — generator matrix, 758
произведение — product, 169, 170, 186
симметричная — symmetric, 171
степеней — degree matrix, 676
сумма — matrix sum, 169, 186
транспонированная — transpose, 171
умножение на скаляр — multiplication by a scalar, 168
Матричная формула Кирхгофа для дерева — Kirchoff's matrix tree formula, 676
Мгновенное безумие — Instant insanity, 267
Мейер (Meyer, C. H.), 848
Метод — Method
Ферма выделения множителей — Fermat's Factorization, 300
двоичного дополнения — two's compliment, 231
Минимальное остовное дерево — Minimal spanning tree, 682
алгоритм Крускала — Krushkal's algorithm, 682
алгоритм Прима — Prim's algorithm, 684, 686
Множество — Set, 67
всех подмножеств — power set, 74
декартово произведение — Cartesian product, 75
дополнение — complement, 72
индексированное — indexed, 76
индексное — indexing, 76
пустое — empty set, 69
универсальное — universal, 69
элемент — element of, 68
Моноид — Monoid, 398
Мономорфизм — Monomorphism
колец — ring monomorphism, 790
полугрупп — semigroup monomorphism, 415
Моррис, Дж. Х. (Morris, J. H.), 836
Мощность — Cardinality, 177
счетная — countable, 177, 180
Мур, Дж. С. (Moore, J. S.), 836

H

Наибольший общий делитель — Greatest common divisor, 142
Наименьшее общее кратное — Least common multiple, 143, 304
Недетерминированный автомат — Nondeterministic automata, 735
Неоднородное рекуррентное отношение — Nonhomogeneous recurrence relation, 460
Непересекающиеся множества — Disjoint sets, 110
Неполное частное — Partial quotient, 309
Неприводимый элемент — Irreducible element, 795

Неравенство Чебышева — Chebyshev's inequality, 381

О

Область главных идеалов — Principal ideal domain, 795
Область значений — Range, 156
Область определения — Domain, 156
Область потенциальных значений — Codomain, 156
Область целостности — Integral domain, 789
  вложенная — embedded, 793
  вполне упорядоченная — well-ordered, 798
  минимальная область — minimal, 801
  неприводимый элемент — irreducible element, 808
  приводимый элемент — reducible element, 808
Образ — Image, 156
Обратная польская запись — Reverse Polish notation, 215
Обратное отношение — Inverse relation, 158
Обратное ребро — Back edge, 661
Обход дерева — Transversal, 650
  в обратном порядке — postorder transversal, 653
  в прямом порядке — preorder transversal, 651
  в центрированном порядке — inorder transversal, 650
Однозначно декодируемый код — Uniquely decipherable code, 727
Однородное рекуррентное отношение — Homogeneous recurrence relation, 449
Операции над множествами — Set operations
  обобщенное объединение — generalized union, 76
  обобщенное пересечение — generalized intersection, 76
  объединение — union, 71
  пересечение — intersection, 71
  разность — difference, 72
  симметрическая разность — symmetric difference, 72
Ориентированное дерево — Directed tree, 624
  m-арное — m-ary, 626
  высота — height, 626
  гомоморфизм — homomorphism, 627
  изоморфизм — isomorphism, 627
  корневое — rooted directed tree, 624
  полное m-арное — full m-ary, 626
  сбалансированное — balanced, 626
  уровень — level, 626
Ориентированный граф — Directed graph, 96, 252
  источник — source, 253
  конечная вершина — terminal vertex, 96, 252
  начальная вершина — initial vertex, 96, 252
  ориентированное ребро — directed edge, 96, 252
  ориентированный путь — directed graph, 255
  ориентированный цикл — directed cycle, 273
  подграф — subgraph, 255
  размеченный — labeled, 254
  связный — connected, 256
  сильно связный — strongly connected, 256
  соотнесенный граф — underlying graph, 255
  степень входа вершины — indegree of vertex, 253
  степень выхода вершины — outdegree of vertex, 253
  сток — sink, 253
Основная теорема
  арифметики — Fundamental Theorem of Arithmetic, 146
Основная теорема алгебры — Fundamental Theorem of Algebra, 820
Остовное дерево — Spanning tree, 560, 658
  минимальное — minimal, 682
Открытый текст — Plaintext, 844
Отношение — Relation, 90
  антирефлексивное — antireflexive, 93
  антисимметричное — antisymmetric, 93
  эквивалентность — equivalence, 107
  композиция — composition, 92
  конгруэнтности — congruence, 811

множество значений — range, 91
область определения — domain, 91
обратное — inverse, 91, 158
рефлексивное — reflexive, 93
рефлексивное замыкание — reflexive closure, 94
симметрическое замыкание — symmetric closure, 94
симметричное — symmetric, 93
транзитивное — transitive, 93
транзитивное замыкание — transitive closure, 94
частичного порядка — partial ordering, 103, 392
Отношение конгруэнтности — Congruence relation, 400
Отношение множеств — Set relation
взаимоисключение — mutual exclusiveness, 110
непересечение — disjointness, 110
подмножество — subset, 68
равенство — equality, 69
собственное подмножество — proper subset, 69
функция — function, 156
Отображение — Mapping, 156
инъективное — injective, 158
сюръективное — surjective, 158

П

Парадокс Рассела — Russell's paradox, 182
Пардум, Джек (Purdum, Jack), 836
Паросочетание — Matching, 707
максимальное — maximal, 707
полное — complete, 707
Паросочетающие ребра — Matching edges, 707
Первообразный корень n-ой степени — Primitive n-th root, 821
Переменная — Indeterminate, 805
Перестановка — Permutation, 161, 331, 332, 418
обобщенная — generalized, 354
орбита — orbit, 420
с запрещенными позициями — with forbidden positions, 509
с повторением — with repetition, 359
цикл — cycle, 420
Перечень раскрасок — Inventory of colorings, 784
Планарный граф — Planar graph, 580
По модулю — Modulo, 149
Подграф — Subgraph, 248
Подходящая дробь — Convergent fraction, 309
Поле — Field, 791
подполе — subfield, 808, 819
расширение — extension, 808
рациональных функций — field of rational functions, 815
частных — field of fractions, 793
Полином — Polynomial, 801, 802
деление — division, 803
делители — factors, 803
коэффициент — coefficient, 802
ладейный — rook polynomial, 509
неприводимый — irreducible, 808
нормированный — monic, 802
приводимый — reducible, 808
примитивный — primitive, 802
произведение — product, 802
простой — prime, 809
равенство — equality, 803
старший коэффициент — leading coefficient, 802
степень — degree, 802
сумма — sum, 802
функция — function, 805
Полиномиальная форма — Polynomial form, 801, 805
Полиномиальное уравнение
решение — Solution of a polynomial equation, 806
Полугруппа — Semigroup, 398
гомоморфизм — homomorphism, 400
замкнутость — closure, 398
идеал — ideal, 825
минимальное порождающее множество — minimal generating set, 400
моноид — monoid, 398
свободная полугруппа — free semigroup, 399
Полурешетка — Semilattice, 401
верхняя — upper semilattice, 393
нижняя — lower semilattice, 394
Польская запись — Polish notation, 215, 652
Порядок целого числа — Order of an integer, 440

Последовательность — Sequence, 163
Фибоначчи — Fibonacci sequence, 449, 452
Постфиксная запись — Postfix notation, 215, 653
Посылка — Premise, 35
Поток — Flow, 692
величина — value, 695
максимальный — maximal, 696
алгоритм нахождения — maximum flow algorithm, 699
сохранение — conservation, 692
Правило Горнера — Horner's rule, 444
Пратт, В. Р. (Pratt, V. R.), 836
Предикат — Predicate, 113
n-местный — n-place, 113
двуместный — two-place, 113
одноместный — one-place, 113
Предметная область — Domain of discourse, 114
Префиксная запись — Prefix notation, 215, 651, 652
Префиксный код — Prefix code, 638, 754
Принцип индукции — Induction principle, 129
второй — second principle, 136, 799
для целых чисел — for integers, 137
первый — first principle, 799
принцип полного упорядочения — well-ordering principle, 799
Принцип клеток — Pigeonhole principle, 363
сильная форма — strong form, 365
Принцип полного упорядочения — Well-ordering principle, 136, 799
Проблема обедающих философов — Dining philosophers' problem, 721
Производящая функция — Generating function, 523–525
подсчет — counting, 535
разбиения — partitions, 542
рекуррентное отношение — recurrence relation, 524
рекуррентное отношение — recurrence relation, 525
экспоненциальная — exponential, 549
Прообраз — Preimage, 156
Прогрессия — Progression
арифметическая — arithmetic, 164
геометрическая — geometric, 164, 193
Простое число — Prime
взаимно простое — relatively prime, 143
псевдопростое — pseudo-prime, 440
целое — integer, 144
Противоречие — Contradiction, 30
Прямая сумма — Direct sum, 821

P

Рабин, М. О. (Rabin, M. O.), 830
Разбиение — Partition, 110, 542
Разделяй и властвуй — Divide and conquer, 212
Размещение — Occupancy, 489
Разупорядочение — Derangement, 504
Райвест, Р. Л. (Rivest, R. L.), 844
Раскраска — Coloring, 775
перечень — inventory, 784
Раскраска графа — Coloring of a graph, 589
Рассел, Бертран (Russell, Bertrand), 67
Расстояние Хемминга — Hamming distance, 768
Рациональные числа — Rational numbers, 793
Регулярная грамматика — Regular grammar, 751
Регулярное выражение — Regular expression, 726
Рекуррентное отношение
линейное — linear, 448
производящая функция — generating function, 524, 525
Рекурсивная функция — Recursive function
решение — solution, 195
Рекурсия — Recursive, 188
определение — definition, 188
решение — solving, 195
Ретракт — Retract, 402, 728
Ретракционное отображение — Retraction map, 728
Решетка — Lattice, 403
дистрибутивная — distributive, 405
ограниченная — bounded, 404
пересечение — meet, 403
подрешетка — sublattice, 404
соединение — join, 403
Решето Эратосфена — Sieve of Eratosthenes, 298
Россер, Дж. Б. (Rosser, J. B.), 834

С

Сандей, Д. М. (Sunday, D. M.), 836
Свойства множеств — Set properties
  двойное дополнение — double complement, 81
  законы де Моргана — de Morgan's Laws, 81
  свойства ассоциативности — associative properties, 81
  свойства дистрибутивности — distributive properties, 81
  свойства дополнения — complement properties, 81
  свойства коммутативности — commutative properties, 81
  свойства тождества — identity properties, 81
  свойство идемпотентности — idempotent properties, 81
Связка — Connective, 16
  дизъюнкция — disjunction, 16
  исключающее или—exclusive or, 18
  конъюнкция — conjunction, 16
  не-и — nand, 46
  не-или — nor, 46
  отрицание — negation, 16, 18
  условная — conditional, 24
  эквиваленция — biconditional, 32
  эквиваленция — biconditional, 25
Сеть — Network, 692
  паросочетаний — matching network, 708
  поток — flow, 692
  сечение — cut, 695
Сеть Петри — Petri net, 716
  безопасная — safe, 722
  достижимое состояние — reachable state, 720
  консервативная — conservative, 722
  метка — token, 717
  непосредственно достижимое состояние — immediately reachable state, 720
  ограниченная — bounded, 722
  переход — transition, 716
  позиция — place, 716
  разметка — marking, 716
  разрешенный переход — enabled *transition, 718*
  срабатывание — firing, 717
  частичный тупик — deadlock, 721
Сечение — Cut, 695
  минимальное — minimal, 696
  пропускная способность — capacity, 695
Символ Кронекера — Kronecker delta, 804
Синдромы — Syndromes, 763
Система вычетов — Residue system
  первичная — primary, 153
  полная — complete, 153
  приведенная — reduced, 153
Система счисления
  двоичная — binary representation, 219
  десятичная — decimal representation, 219
  шестнадцатеричная — hexadecimal representation, 222
Скаляр — Scalar, 167
Скалярное произведение — Dot product, 169
Слово — Word, 725
Сложность — Complexity, 201
  О большое — big-oh, 201
Случайная величина — Random variable, 375
Смежный класс — Coset, 413
Составное число — Composite number, 144
Сочетание — Combination, 335
  обобщенное — generalized, 354
  с повторением — with repetition, 361
Сравнение — Congruence, 149, 424
Стабилизатор — Stabilizer, 778
Стек — Stack, 216
  вталкивание — push, 216
  выталкивание — pop, 216
Стрелка Пирса — Peirce's Arrow, 45
Строка — String, 399, 725
  конкатенация — concatenation, 725

Т

Таблица истинности — Truth table, 17
Тавтология — Tautology, 30
Теорема — Theorem, 35
  Вандермонда — Vandermonde's Theorem, 340
  Клини — Kleene's Theorem, 738
  Куратовского — Kuraotowski's *Theorem, 583*
  Пойа о подсчете — Polya's Counting Theorem, 785

Рамсея — Ramsey's Theorem, 368
Уилсона — Wilson's Theorem, 437
Эйлера — Euler's Theorem, 439
о разложении на простые множители — Prime Factorization Theorem, 146
Теория Рамсея — Ramsey theory, 366
Тотиент-функция Эйлера — Euler's totient function, 434
Точка сочленения — Articulartion point, 561, 668
Треугольник Паскаля — Pascal's triangle, 339
Трихотомия
аксиома — Trichotomy Axiom, 127

У

Умозаключение — Argument, 35
сведение к абсурду — reductio ad absurdum, 247
modus ponens, 39, 40
modus tollens, 40
выбор — cases, 40
гипотеза — hypothesis, 35
доказательство — proof, 42
исключающий выбор — case elimination, 40
контрапозиция — contrapositive, 126
конъюнкция — conjunction, 40
ложная инверсия — inverse fallacy, 42
ложная конверсия — converse fallacy, 41
неправильное — invalid, 41
посылка — premise, 35
правило отделения — law of detachment, 39, 40
правильное — valid, 35
расширение — addition, 40
сведение к абсурду — reductio ad absurdum, 41
силлогизм — syllogism, 40
специализация — specialization, 40
Универс — Universe, 114
Универсальная конкретизация — Universal instantiation, 117
Универсальное обобщение — Universal generalization, 118
Упорядоченная пара — Ordered pair, 75

Ф

Формирование перестановок — Generating permutations, 344
Формирование сочетаний — Generating combinations, 344
Функция — Function, 156
Аккермана — Ackerman's function, 192
Сложность — Complexity, 201
Эйлера — Euler's ϕ function, 434
биекция — bijection, 158
бинарная операция — binary operation, 163, 397
верхнее округление — ceiling, 162
инъективная — injective, 158
нижнее округление — floor, 162
областью значений — range, 156
область определения — domain, 156
область потенциальных значений — codomain, 156
образ — image, 156
перестановка — permutation, 158
прообраз — preimage, 156
сюръективная — surjective, 158
тождественная — identity, 159
факториал — factorial, 162
хеширования — hashing function, 837
минимальная совершенная — minimum perfect, 837

Х

Ханойская башня — Tower of Hanoi, 193
Характер — Character
Дирихле — Dirichlet character, 826
главный — principal, 821
группы — group character, 820
полугруппы — semigroup character, 825
Хеллмэн, М. Э. (Hellman, M. E.), 847, 849
Хроматический многочлен — Chromatic polynomial, 589
Хроматическое число — Chromatic number, 589

Ц

Целое число — Integer
взаимно простые — relatively prime, 143
гауссовы — Gaussian integers, 816

делитель — divisor, 139
кратное — multiple, 139
множество классов вычетов по модулю n — set of integers modulo n, 149
наибольший общий делитель — greatest common divisor, 142
наименьшее — least, 135
наименьшее общее кратное — least common multiple, 143, 304
порядок — order of, 440
простое — prime, 144
простое число — prime, 798
составное — composite, 144
сравнимое — congruent, 149
Цепи Маркова — Markov chains, 386
матрица перехода — transition matrix, 387
состояния — states, 386
Цепная дробь — Continued fraction, 309
конечная — finite continued fraction, 309
простая — simple continued fraction, 309
Цикл — Cycle
гамильтонов — hamiltonian, 601
Цикловый индекс — Cycle index, 783

Ч

Чанг (Chang, C. C.), 839
Частично упорядоченное множество — Partially ordered set, 103, 392
вполне упорядоченное — total ordering, 104, 392
диаграмма Гессе — Hasse diagram, 104
максимальный элемент — maximal element, 392
минимальный элемент — minimal element, 393
наибольшая верхняя грань — greatest lower bound, 392
наименьшая верхняя грань — least upper bound, 392
наименьший элемент — least element, 392
сравнимое — comparable, 104, 392
топологическая сортировка — topological sorting, 185
цепь — chain, 104
Частичный порядок — Partial ordering, 103, 392
Числа Каталана — Catalan numbers, 190, 495
Числа Стирлинга — Stirling numbers
второго рода — second kind, 480, 481, 490, 491, 553
первого рода — first kind, 480, 481, 492, 493
Число Рамсея — Ramsay number, 367

Ш

Шамир, А. (Shamir, A.), 844
Шекерес (Szekeres), 364, 365, 367
Ши (Shieh, J. C.), 839
Шифрование — Encryption, 844
Шонфилд, Л. (Schoenfeld, L.), 834
Штрих Шеффера — Sheffer Stroke, 45

Э

Эйлер, Леонард — Euler, Léonard, 438
Эквивалентность — Equivalence
класс — class, 108
отношение — relation, 107
Экзистенциональная конкретизация — Existential Instantiation, 118
Экзистенциональное обобщение — Existential Generalization, 118
Экспоненциальная производящая функция — Exponential generating function, 549
Элементарная дизъюнкция — Maxterm, 48
Элементарная конъюнкция — Minterm, 47
Элементарные дроби — Partial fractions, 527
Эпиморфизм — Epimorphism
колец — Ring epimorphism, 790
полугрупп — semigroup epimorphism, 415
Эрдош (Erdös), 364, 365, 367

Я

Язык — Language, 725
Яэшке, Г. (Jaeschke, G.), 837

Список обозначений

■ ЛОГИКА

1.	$p \wedge q$	p и q	стр. 16
2.	$p \vee q$	p или q	стр. 16
3.	$\sim p$	не p	стр. 16
4.	$\veebar$	исключающее или	стр. 18
5.	$p \rightarrow q$	если p, то q	стр. 23
6.	$p \leftrightarrow q$	p тогда и только тогда, когда q	стр. 25 и 32
7.	$p \equiv q$	p логически эквивалентно q	стр. 28
8.	$\mathbf{T}$	утверждение-тавтология	стр. 30
9.	$\mathbf{F}$	логически ложное утверждение	стр. 30
10.	$\therefore$	следовательно	стр. 35
11.	$\mid$	штрих Шефера (не-и)	стр. 45
12.	$\downarrow$	стрелка Пирса (не-или)	стр. 45
13.	$\forall$	для всех	стр. 114
14.	$\exists$	существует	стр. 115

■ МНОЖЕСТВА

15.	$\{a_1, a_2, a_3, \ldots, a_n\}$	множество элементов $a_1, a_2, a_3, \ldots, a_n$	стр. 67
16.	$\{x : P\}$	множество элементов, обладающих свойством P	стр. 68
17.	$a \in A$	a есть элемент множества A	стр. 68
18.	$A \subseteq B$	множество A есть подмножество множества B	стр. 68
19.	$\varnothing$	пустое множество, или нуль-множество	стр. 69
20.	$\boldsymbol{U}$	универсальное множество	стр. 69
21.	$A \cap B$	пересечение множеств A и B	стр. 71
22.	$A \cup B$	объединение множеств A и B	стр. 71
23.	$A - B$	разность множеств A и B	стр. 72
24.	$A \triangle B$	симметрическая разность множеств A и B	стр. 72
25.	A'	дополнение множества A	стр. 72
26.	$\mathcal{P}(A)$	множество всех подмножеств множества A	стр. 74
27.	$A \times B$	декартово произведение множеств A и B	стр. 75
28.	(a, b)	упорядоченная пара	стр. 75
29.	aRb	a и b связаны отношением R	стр. 90
30.	R^{-1}	отношение, обратное к R	стр. 91
31.	$S \circ R$	композиция отношений S и R	стр. 92
32.	$(S, \leq)$	множество, с отношением частичного порядка $\leq$	стр. 103
33.	$\bigcap_{i=1}^{n} A_i$	пересечение множеств $A_1, A_2, \ldots, A_n$	стр. 138
34.	$\bigcup_{i=1}^{n} A_i$	объединение множеств $A_1, A_2, \ldots, A_n$	стр. 138

■ ФУНКЦИИ

35.	$f : A \to B$	функция f из A в B	стр. 156
36.	$f(E)$	образ E	стр. 156
37.	$f^{-1}(F)$	прообраз F	стр. 156
38.	$f \circ g$	композиция функций f и g	стр. 157
39.	f^{-1}	обратная функция	стр. 158
40.	$f(x) = \lfloor x \rfloor$	нижнее округление	стр. 162
41.	$f(x) = \lceil x \rceil$	верхнее округление	стр. 162
42.	$f(x) = x!$	факториал	стр. 162
43.	$O(g(n))$	О-большое от $g(n)$	стр. 201
44.	δ_{ij}	символ Кронекера	стр. 240 и 804

■ КОММУТАЦИОННЫЕ СХЕМЫ

45.		логический элемент *и*	стр. 58
46.		логический элемент *или*	стр. 58
47.		логический элемент *не*	стр. 58
48.		логический элемент *не-и*	стр. 60
49.		логический элемент *не-или*	стр. 60
50.	$p \cdot q$	булево произведение (p *и* q)	стр. 56
51.	$p + q$	булева сумма (p *или* q)	стр. 56
52.	p'	булево дополнение (*не* p)	стр. 56

■ ГРАФЫ И ДЕРЕВЬЯ

53.	$G(V, E)$	граф с вершинами V и ребрами E	стр. 95 и 244
54.	$\{a, b\}$	ребро между вершинами a и b	стр. 95 и 244
55.	(a, b)	ориентированное ребро из вершины a в вершину b	стр. 252
56.	$deg(v)$	степень вершины v	стр. 247
57.	K_n	полный граф с числом вершин n	стр. 250
58.	$K_{m,n}$	полный двудольный граф с числом вершин m и n	стр. 251
59.	$v_0 v_1 v_2 \cdots v_n$	путь длины n из v_0 в v_n	стр. 248
60.	$G^c(V, E')$	дополнение к графу $G(V, E)$	стр. 559
61.	$C_G(\lambda)$	число способов раскраски графа $G(V, E)$ с использованием λ цветов	стр. 589
62.	$G_{\hat{e}}$	граф G с удаленным ребром e	стр. 590
63.	$G_{/e}$	граф G с удаленным ребром $e = \{a, b\}$ и отождествленными вершинами a и b	стр. 590
64.	$cl(G)$	замыкание графа $G(V, E)$	стр. 606
65.	$c(e)$	пропускная способность ребра e	стр. 691
66.	$f(e)$	поток через ребро e	стр. 692
67.	$val(f)$	величина потока f	стр. 695
68.	(S, T)	сечение сети на подсети S и T	стр. 695
69.	$C(S, T)$	пропускная способность сечения (S, T)	стр. 695

■ ТЕОРИЯ ЧИСЕЛ

70.	Z	множество целых цисел	стр. 123
71.	N или Z^+	множество целых положительных чисел	стр. 127
72.	$a \mid b$	a делит b	стр. 139

73.	НОД(a, b)	наибольший общий делитель чисел a и b	стр. 142
74.	НОК(a, b)	наименьшее общее кратное	стр. 143
75.	$a \equiv b \pmod{n}$	a сравнимо с b по модулю n	стр. 149
76.	$[a]$	класс эквивалентности, содержащий a	стр. 108
77.	$[a] \oplus [b]$	сумма классов эквивалентности	стр. 152
78.	$[a] \odot [b]$	произведение классов эквивалентности	стр. 152
79.	$[[m]]_n$	наименьшее целое положительное число, сравнимое с m по модулю n	стр. 151
80.	ϕ	функция Эйлера	стр. 434
81.	$\mathrm{ord}_n\, a$	порядок a по модулю n	стр. 440

■ МАТРИЦЫ

82.	$[A_{ij}]$	матрица элементов A_{ij}	стр. 167
83.	$A + B$	сумма матриц	стр. 169
84.	aA	умножение матрицы на скаляр	стр. 168
85.	AB	произведение матриц	стр. 170
86.	A^t	транспонированная матрица A	стр. 171
87.	$A \odot B$	булево произведение матриц	стр. 172
88.	$\overline{A_{ij}}$	матрица, полученная из A удалением i-ой строки и j-го столбца	стр. 237
89.	$\lvert A\rvert = \det(A)$	определитель (детерминант) матрицы A	стр. 238
90.	I_n	единичная $n \times n$ матрица	стр. 240
91.	A^{-1}	матрица, обратная к A	стр. 240
92.	$\mathrm{adj}(A)$	матрица, сопряженная с A	стр. 241

■ РЕКУРСИЯ И ПОДСЧЕТ

93.	$n!$	n-факториал	стр. 162
94.	$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n$	символ суммы	стр. 164
95.	$P(n, r) = \frac{n!}{(n-r)!}$	число перестановок объектов из n по r	стр. 332
96.	$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$	число сочетаний объектов из n по r	стр. 335
97.	$P(n; n_1, n_2, \ldots, n_k) = \frac{n!}{n_1!n_2!\ldots n_k!}$	обобщенная перестановка	стр. 354
98.	$C(n; n_1, n_2, \ldots, n_k) = \frac{n!}{n_1!n_2!\ldots n_k!}$	обобщенное сочетание	стр. 357
99.	$P(A)$	вероятность события A	стр. 348
100.	$P(B \mid A)$	вероятность B при данном A	стр. 371
101.	$E(R)$	ожидаемое значение случайной величины R	стр. 376
102.	μ	среднее значение случайной величины R	стр. 379
103.	σ^2	дисперсия случайной величины R	стр. 379
104.	$R(p, q)$	число Рамсея	стр. 367
105.	Фиб(n)	n-ое число Фиббоначи	стр. 190
106.	Кат$(n) = C_n$	n-ое число Каталана	стр. 190
107.	$G(C)$	стабилизатор раскраски C	стр. 778
108.	$\Delta^k f(x)$	k-я разность функции $f(x)$	стр. 469
109.	E	оператор, определяемый равенством $E(f(x)) = f(x+1)$	стр. 470
110.	$x^{(n)}$	$x(x-1)(x-2)\cdots(x-n+1)$	стр. 473
111.	$a_n x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1 x + a_0$	факториальный многочлен	стр. 473
112.	$x^{(-m)}$	$\frac{1}{(x+m)^m}$	стр. 479

113. $\binom{x}{n}$	$\frac{x^{(n)}}{n!}$	стр. 480
114. $s_k^{(n)}$	число Стирлинга первого рода	стр. 480
115. $S_k^{(n)}$	число Стирлинга второго рода	стр. 480
116. $\sum$	оператор суммирования	стр. 483
117. L_n	n-ое число Лукаса	стр. 459
118. D_n	число разупорядочений n различных упорядоченных символов	стр. 504
119. $r_k(c)$	число возможных расположений k ладей на доске C в неатакующих позициях	стр. 509
120. $R(x, C)$	ладейный полином на доске C	стр. 509
121. $[t_0 : t_1, t_2, \cdots, t_n]$	цепная дробь	стр. 309
122. $T_k^{(n)}$	число способов размещения n различимых объектов в k различимых ящиках при условии, что ни один из ящиков не будет пустым	стр. 553

■ АЛГЕБРА

123. $(A, \vee)$ или $(A, +)$	верхняя полурешетка A	стр. 393
124. $(A, \wedge)$ или $(A, \cdot)$	нижняя полурешетка A	стр. 394
125. $(S, \vee, \wedge)$	решетка S	стр. 403
126. $a \circ H$	левый смежный класс	стр. 413
127. $G_1 \oplus G_2$	прямая сумма групп G_1 и G_2	стр. 821
128. $\langle a \rangle$	главный идеал, порожденный элементом a	стр. 794
129. $A[x]$	множество полиномов с коэффициентами в A	стр. 805
130. χ	характер группы	стр. 821
131. $a_1 a_2 a_3 \cdots a_n$	строка символов алфавита A	стр. 725
132. A^*	множество всех строк над алфавитом A	стр. 725
133. $M = (A, S, s_0, T, F)$	автомат с алфавитом A, множеством состояний S, начальным состоянием s_0, множеством допускающих состояний T и функцией переходов F	стр. 731
134. $M(L)$	язык, допускаемый автоматом M	стр. 731
135. $\Gamma = (N, T, S, P)$	грамматика с множеством нетерминальных символов N, множеством терминальных символов T, начальным символом S и множеством продукций P	стр. 742

■ КОДЫ

136. $wt(c)$	число единиц в строке кода c	стр. 758
137. $C^\perp$	код, двойственный к коду C	стр. 761
138. G_H	порождающая матрица Хемминга	стр. 767
139. $G_H^\perp$	матрица контроля четности Хемминга	стр. 767
140. (c, c')	расстояние Хемминга между строками кодов c и c'	стр. 768
141. $D(C)$	наименьшее расстояние между двумя произвольными элементами C	стр. 769

Научно-популярное издание

Джеймс Андерсон

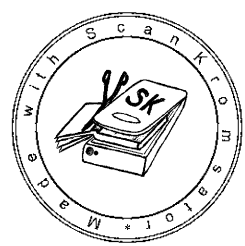
Дискретная математика и комбинаторика

Литературный редактор *Т.Т. Шматко*
Верстка *М.Р. Саит-Аметов*
Художественный редактор *С.А. Чернокозинский*
Корректоры *Л.А. Гордиенко, О.В. Мишутина*

Издательский дом "Вильямс".
101509, Москва, ул. Лесная, д. 43, стр. 1.
Изд. лиц. ЛР № 090230 от 23.06.99
Госкомитета РФ по печати.

Подписано в печать 25.08.2004. Формат 70×100/16
Гарнитура AntiquaPSCyr. Печать офсетная.
Усл. печ. л. 77,4. Уч.-изд. л. 63,17.
Доп. тираж 2500 экз. Заказ № 343.

Отпечатано с диапозитивов в ФГУП "Печатный двор"
Министерства РФ по делам печати,
телерадиовещания и средств массовых коммуникаций
197110, Санкт-Петербург, Чкаловский пр., 15.



Дискретная математика и комбинаторика

Джеймс Андерсон



Книга представляет собой современный учебник по дискретной математике университетского уровня. Она может использоваться как пособие для самостоятельного обучения, а также послужит незаменимым справочником по дискретной математике и смежным разделам математики, в том числе:

- Логика и теория доказательств
- Теория множеств
- Теория алгоритмов, языков и вычислений
- Теория графов (включая теорию сетей)
- Алгебра (включая теорию полей, групп и полугрупп)
- Теория чисел
- Комбинаторика
- Теория вероятностей
- Теория кодирования

Особенности книги

- Необычайно широкий спектр затронутых тем
- Простота и доходчивость изложения
- Практически полная самодостаточность
- Описан ряд практических приложений
- Предметно-именной указатель содержит английские эквиваленты всех встречающихся понятий, что фактически делает его кратким словарем терминов

Об авторе

Джеймс А. Андерсон — известный математик, доктор наук, профессор университета Южной Каролины, автор множества научных работ, включая книгу по теории чисел. В круг его научных интересов также входят алгебра и математические основы информатики. Имеет более чем 30-летний стаж преподавательской работы.

Вильямс
www.williamspublishing.com

ISBN 5-8459-0498-6

Pearson Education
www.prenhall.com