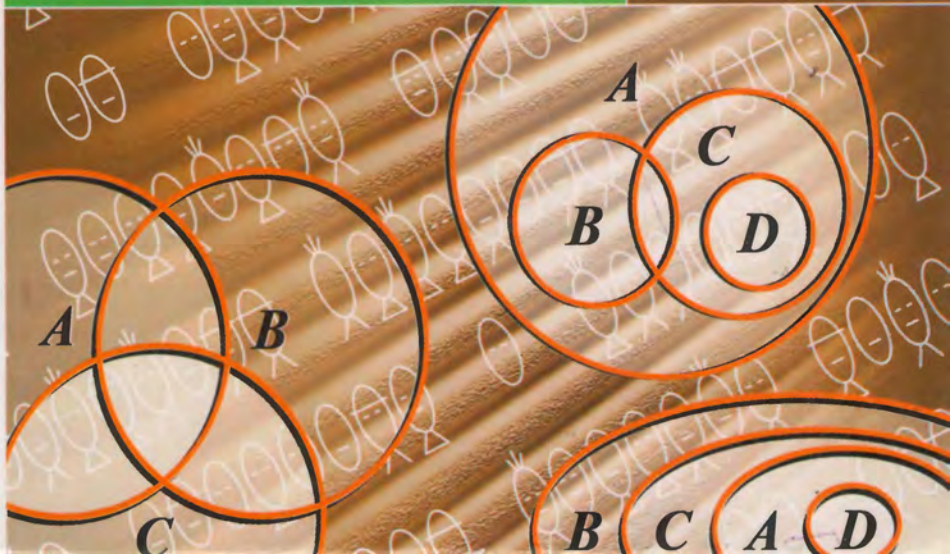
СРЕДНЕЕ ПРОФЕССИОНАЛЬНОЕ ОБРАЗОВАНИЕ

М. С. Спирина
П. А. Спирин

ДИСКРЕТНАЯ МАТЕМАТИКА

ИНФОРМАТИКА
И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

ACADEMIA



УЧЕБНИК

СРЕДНЕЕ ПРОФЕССИОНАЛЬНОЕ ОБРАЗОВАНИЕ

М. С. СПИРИНА, П. А. СПИРИН

ДИСКРЕТНАЯ МАТЕМАТИКА

Учебник

Допущено
Министерством образования Российской Федерации
в качестве учебника для студентов образовательных учреждений
среднего профессионального образования, обучающихся
по специальностям 2202 «Автоматизированные системы
обработки информации и управления (по отраслям)»,
2203 «Программное обеспечение вычислительной техники
и автоматизированных систем»

Москва
ACADEMA
2004

УДК 51(075.32)
ББК 22.176я723
С722

Рецензенты:

академик МАН ВШ, д-р техн. наук, проф. *П. А. Арутюнов*;
преподаватель Московского электронно-технологического колледжа *Т. Б. Абашева*

Спирина М. С.

С722 Дискретная математика: Учебник для студ. учреждений сред. проф. образования / М. С. Спирина, П. А. Спирин. — М.: Издательский центр «Академия», 2004. — 368 с.

ISBN 5-7695-1496-5

Представляет собой углубленный междисциплинарный курс и содержит теоретический материал по традиционным темам дискретной математики и некоторые вопросы классической логики. В каждой главе есть исторический материал, разобранные задачи с указанием методов их решений, система упражнений для самостоятельной работы.

Для студентов и преподавателей учреждений среднего профессионального образования, связанных с информационными системами, компьютерным моделированием, разработкой программных продуктов и автоматизированных систем.

УДК 51(075.32)
ББК 22.176я723

Учебное издание

Спирина Марина Савельевна,
Спирин Павел Алексеевич

Дискретная математика

Учебник

Редактор *В. А. Новиков*
Технический редактор *О. С. Александрова*
Компьютерная верстка: *Г. Ю. Никитина*
Корректоры *Л. С. Киммель, С. Ю. Свиридова*

Диапозитивы предоставлены издательством

Изд. № А-1075-I. Подписано в печать 05.05.2004. Формат 60×90/16.
Гарнитура «Таймс». Печать офсетная. Бумага тип. № 2. Усл. печ. л. 23,0.
Тираж 5100 экз. Заказ № 13416.

Лицензия ИД № 02025 от 13.06.2000. Издательский центр «Академия».
Санитарно-эпидемиологическое заключение № 77.99.02.953.Д.003903.06.03 от 05.06.2003.
117342, Москва, ул. Бутлерова, 17-Б, к. 222. Тел./факс: (095)330-1092, 334-8337.

Отпечатано на Саратовском полиграфическом комбинате.
410004, г. Саратов, ул. Чернышевского, 59.

© Спирина М. С., Спирин П. А., 2004
© Образовательно-издательский центр «Академия», 2004
© Оформление. Издательский центр «Академия», 2004

ISBN 5-7695-1496-5

ПРЕДИСЛОВИЕ

Математика стала частью нашей культуры. Человек не может считать себя широкообразованным, не имея представления о современной математике, ее роли в повседневной жизни, в науке.

Для понимания излагаемых вопросов достаточно знаний в объеме программы средней школы. Некоторые затруднения может вызвать широкое использование языка теории множеств. Чтобы этого не произошло, изложение курса дискретной математики начинается именно с этой темы. Большое внимание уделяется прикладному характеру содержания этого курса.

Его основу составляют такие разделы математики, как комбинаторный анализ, теория множеств, элементы математической и классической логики.

Дискретная математика для программистов относится к числу общепрофессиональных предметов, формирующих базовый уровень знаний, необходимых для изучения других дисциплин, таких, как «Математическая статистика», «Архитектура ЭВМ, систем и сетей», «Базы данных», «Компьютерное моделирование», «Автоматизированные системы», «Технология разработки программных продуктов», «Основы алгоритмизации и программирования».

Предложенный вариант учебной дисциплины можно назвать «курсом на межпредметной основе». Его создание имело целью сформировать профессионально-прикладную компетенцию будущих программистов. Для этого математические знания были увязаны с общенаучными и учтены современные представления о функциональной роли математического образования. Курс формирует у учащихся систему умений и навыков самостоятельного избирательного восприятия информации и ее переработки. Его задачи — научить систематизации, обобщению, структурированию знаний, а также их адекватному применению как в предметных областях, так и в практической деятельности. Сочетание фундаментальных теоретических знаний с их функциональной направленностью призвано показать учащимся использование универсального математического аппарата применительно к различным предметным областям и разнообразным видам деятельности. Акцент делается на знакомство с разными приемами систематизации знаний и представления информации в сжатом виде.

Главной особенностью данной книги является избыточность информации в изложении теоретического материала и в предложенной системе упражнений. Опыт преподавания дискретной математики студентам Тольяттинского политехнического колледжа показывает, что интерес к этой теме приобретается порой в процессе ее изучения. Содержание и язык, близкие к познавательной литературе, способствуют не только освоению программы, но и установлению собственного отношения к приобретению знаний.

В учебник включено большое количество различных занимательных задач. Они способствуют пробуждению у читателя интереса к процессу познания, к исследовательской деятельности. Избыточность теоретической информации и практических заданий оправдана современной установкой на увеличение доли самостоятельной деятельности студентов.

Для поддержания интереса к предмету в книге использованы многочисленные исторические экскурсы, сочетание математического содержания с элементами классической логики, примеры жизненного опыта учащихся и т. д.

Воплощая на практике идею перехода от математики фундаментальной к математике функциональной, авторы считают необходимым найти компромисс между этими противоположными позициями. В учебнике на достаточном для учреждений системы среднего профессионального образования уровне представлен теоретический материал. С другой стороны, современную математику часто определяют как науку, работающую с готовыми моделями и создающую новые. Поэтому авторы стремились показать широкие возможности применения математики, ее универсальность. В современном мире математика становится методом мышления, и авторы считали необходимым показать, каким образом с помощью языка математики формируются новые понятия. Этому способствует введение элементов классической логики, изучение которой помогает установлению связей между различными и на первый взгляд далекими понятиями. С помощью большого круга нематематических заданий мы формируем интерес к предмету, оказывая влияние на интеллектуальный потенциал читателя.

Теоретический материал состоит из большого числа формулировок вновь вводимых понятий. Для облегчения работы с учебником каждое новое понятие выделяется полужирным шрифтом, а также включается в предметный указатель. Содержательная часть может изучаться на двух уровнях сложности: облегченном и глубоком. Читатель сам вправе сделать выбор глубины погружения в предмет дискретной математики.

Практические задания представлены в большом объеме, но их избыточность оправдана теми целями, которые преследовали авторы.

Подбор задач в упражнениях к главам (кроме авторских) осуществлялся с использованием различных источников, приведенных в списке литературы, а также материалов журнала «Квант» и газеты «Математика» — еженедельного учебно-методического приложения к газете «Первое сентября».

Для лучшего усвоения определений аналогичных и противоположных понятий используется метод укрупненных дидактических единиц П. М. и Б. П. Эрдниевых. Определения представлены в виде дробей: их общая часть записана линейно, тогда как числители и знаменатели дробей соответствуют отличительным особенностям. Например, запись «**Степенью** $\frac{\textbf{входа}}{\textbf{выхода}}$ вершины ориентированного графа называется число ребер, для которых эта вершина является $\frac{\text{концом}}{\text{началом}}$» содержит два следующих определения:

Степенью входа вершины ориентированного графа называется число ребер, для которых эта вершина является концом;

Степенью выхода вершины ориентированного графа называется число ребер, для которых эта вершина является началом.

Следует обратить внимание на эпиграфы и вопросы для размышления. Они способствуют приобретению самостоятельности мышления, дают возможность формировать собственное отношение к изучаемому предмету и окружающей действительности.

Предложенный в учебнике материал основан на тщательном анализе и отборе содержания с глубинным пониманием исходных целей образования. Надеемся, что он будет способствовать становлению нового поколения — поколения многогранных, эрудированных, компетентных, творчески мыслящих специалистов.

Авторы приносят благодарность студентам факультета программирования Тольяттинского политехнического колледжа 1998—2002 гг. обучения за помощь в создании этого учебника.

ПЕРЕЧЕНЬ МАТЕМАТИЧЕСКИХ СИМВОЛОВ И СОКРАЩЕНИЙ

Символы и обозначения

$\subset$ — включение;
$\in$ — принадлежность элемента множеству;
$\notin$ — непринадлежность элемента множеству;
$\cup$ — объединение множеств;
$\cap$ — пересечение множеств;
$\sum$ — суммирование;
$\times$ — декартово произведение;
$\forall$ — квантор общности;
$\exists$ — квантор существования;
$|M|$, $n(M)$ — мощность множества M;
U — универсальное множество;
$\bigcup\limits_i A_i$ — объединение множеств A_i;
$\bigcap\limits_i A_i$ — пересечение множеств A_i;
$\langle\rangle$ — кортеж;
$\deg(A)$ — степень вершины A;
$n!$ — факториал числа n, т. е. произведение натуральных чисел от 1 до n;
$\exists!$ — существует единственный;
$A \backslash B$ — разность множеств A и B;
$A \Delta B$ — симметрическая разность;
$\neg a$, $\bar{a}$ — отрицание a;
$\&$, $\cdot$, $\wedge$ — конъюнкция;
$\vee$ — дизъюнкция;
$\veebar$, $\oplus$ — строгая дизъюнкция, сумма по модулю два;
$\prec$ — частичный порядок;
$\varnothing$ — пустое множество;
$\rightarrow$ — импликация;
$\leftrightarrow$, $\equiv$ — эквиваленция;
$\Rightarrow$ — логическое следование;
$\overset{def}{\Leftrightarrow}$ — равносильность высказывательных форм;
$\overset{def}{=}$ — равно по определению;
$\models$ — непосредственная выводимость;

$\vdash$ — выводимость;
$\vdots$ — деление нацело;
2^M — булеан множества M;
A_n^m — число размещений из n по m;
C_n^m — число сочетаний из n по m;
P_n — число перестановок из n элементов;
B^n — n-я декартова степень множества B;
$\mathbb{N}$ — множество натуральных чисел;
$\mathbb{Z}$ — множество целых чисел;
$\mathbb{Q}$ — множество рациональных чисел;
$\mathbb{R}$ — множество действительных чисел;
$\mathbb{C}$ — множество комплексных чисел.

Сокращения

ДНФ — дизъюнктивная нормальная форма;
ДСС — десятичная система счисления;
КНФ — конъюнктивная нормальная форма;
МЗР — младший значащий разряд;
ММИ — метод математической индукции;
ПСС — позиционная система счисления;
СДНФ — совершенная дизъюнктивная нормальная форма;
СКНФ — совершенная конъюнктивная нормальная форма.

ВВЕДЕНИЕ

...И в мире нет таких вершин,
Что взять нельзя.
Среди нехоженых путей
Один — пусть мой.
Среди невзятых рубежей
Один — за мной.

В. Высоцкий

? Что изучает дискретная математика?

Со школьной скамьи нам известно, что некоторые явления в окружающем мире можно описать с помощью понятий действительного числа и непрерывного трехмерного пространства. Математические модели, связанные со свойствами непрерывности изучаемых объектов, раскрывают многие закономерности материального мира.

Дискретная математика или дискретный анализ — сравнительно новое направление в математике, объединяющее отдельные ее разделы, ранее сформированные как самостоятельные теории. К ним относятся математическая логика и теории множеств, графов, кодирования, автоматов.

Что общего между этими, различными на первый взгляд, теориями?

Дискретной математикой называют совокупность математических дисциплин, изучающих свойства абстрактных дискретных объектов, т.е. свойства математических моделей объектов, процессов, зависимостей, существующих в реальном мире, которыми оперируют в различных областях знаний. Таким образом, *дискретный анализ* — самостоятельный раздел современной математики, изучающий свойства различных структур, имеющих конечный характер. Они могут возникать как в самой математике, так и в ее приложениях. К их числу принято относить объекты, имеющие прерывный (*дискретный*) характер в отличие от объектов, изучаемых классической математикой и носящих *непрерывный* характер.

Математический аппарат дискретного анализа можно определить как взаимосвязанную совокупность языка, моделей и методов математики, ориентированную на решение различных, в том числе инженерных, задач. Использование такого аппарата связано с характером исследуемых моделей — отдельных элементов абстрактных множеств, отдельных чисел в различных системах счисления, отдельных значений 0 и 1 (истина и ложь), булевых функций и т.д.

Вообще говоря, деление математики на дискретную и классическую достаточно условно. Например, аппарат теории множеств и теории графов используется при изучении не только дискретных, но и непрерывных объектов. С другой стороны, сама дискретная математика использует средства, разработанные в классической математике. Однако характер объектов, исследуемых дискретной математикой, настолько своеобразен, что методов классической математики не всегда достаточно для их изучения. Поэтому те специфические методы, которые применяются для очень широкого класса *конечных* (*финитных*) дискретных объектов, и были объединены в общее направление — дискретную математику.

Несмотря на то что отдельные направления дискретной математики зародились в глубокой древности и совершенствовались параллельно с классической математикой, наиболее интенсивно дискретная математика стала развиваться в последнее столетие. В настоящее время знание дискретной математики необходимо специалистам в различных областях деятельности.

XXI в. называют веком информатизации, когда основной объем информации хранится в памяти ЭВМ. Любому человеку, стремящемуся ею воспользоваться, необходимо освоить азы «безбумажной информатики». Применение ЭВМ для комплексной автоматизации информационной деятельности принципиально изменило характер взаимоотношений человека и машины. Если раньше компьютер осваивали только те, кто непосредственно его обслуживал: программисты, электронщики, операторы, то в XXI в. без машинной обработки информации не обойдется ни одна отрасль деятельности.

Стимулом для развития многих направлений дискретной математики явились запросы теоретической кибернетики, непосредственно связанной с развитием ЭВМ.

Теоретическая кибернетика занимается изучением разнообразных практических проблем средствами дискретной математики:

растущий поток информации и проблемы ее передачи, обработки и хранения привели к возникновению и развитию *теории кодирования*;

различные экономические задачи, задачи электротехники стимулировали создание и развитие *теории графов*;

связь релейно-контактных схем с формулами алгебры логики и их использование для описания функционирования автоматов дали начало развитию и применению *математической логики* и *теории автоматов*. Математическая логика в широком смысле изучает основания математики, принципы построения математических теорий.

Дискретная математика изучает объекты, которые порой не имеют ни физической, ни числовой интерпретации. В классической математике характеристики реальных объектов можно пред-

ставить в виде чисел, а закономерности — в виде соотношений. В отличие от реальных характеристиками информационных объектов могут служить понятия «структура», «отношение», «связь». Обычно объекты информатики рассматривают как комбинации некоторых абстрактных символов, над которыми производятся некие манипуляции.

В последнее время раздел математики, называемый «Дискретный анализ», все чаще вводится в программы подготовки не только математиков, инженеров, программистов, но даже юристов. Интерес к этой дисциплине не случаен, так как потребность в знаниях этой области математики объясняется широким кругом ее применения: электроника и информатика, вопросы оптимизации и принятия решений.

Взаимосвязь дискретной математики с другими науками. Кибернетические области информатики используют в качестве аппарата язык как фундаментальной, так и прикладной математики. Однако надо учитывать, что эти науки связаны между собой и их деление условно. *Кибернетика* — наука об общих принципах управления в живых, неживых и искусственных системах. Решая множество разнообразных задач, кибернетика имеет общий стержень, общую методологию, в основе которой лежит понятие системы. Под **системой** понимают некую структуру, объединение некоторого количества обособленных элементов, подчиненных единой взаимосвязи, определенным отношениям. Кибернетика является наукой об управляемых системах любого характера: биологических, социальных, технических, экономических. В связи с этим в различных системах выделяют так называемый *кибернетический подход*, смысл которого заключается в наличии механизма управления этой системой, в существовании *обратной связи*.

Методы, разрабатываемые дискретной математикой, часто используются в различных направлениях информатики. Так, теоретическая информатика (или теоретическая кибернетика) использует математические методы для построения и изучения моделей обработки, передачи и использования информации. Объекты ее изучения — дискретные множества. Теоретическая информатика является как поставщиком задач, так и потребителем методов дискретной математики.

Достижения математической логики используются для анализа процессов переработки информации с помощью ЭВМ. *Теория автоматов* разрабатывает методы, с помощью которых можно на основе моделей логического типа изучать процессы, протекающие в самой машине во время ее работы. Для работы на компьютере информацию представляют в дискретной форме, позволяющей переводить ее в программы, понятные ЭВМ.

Теория информации изучает вид тех форм, в которых информация представляется в компьютере. Формализация любой ин-

формации, реально существующей в живой и неживой природе, происходит через *компьютерное моделирование. Системный анализ* изучает структуру реальных объектов и дает способы их формализованного описания. О*бщая теория систем* как часть системного анализа изучает различные по характеру системы с общих позиций. *Теория массового обслуживания* изучает широкий класс моделей передачи и переработки информации в системах массового обслуживания (СМО).

В настоящее время нашла широкое применение наука *семиотика*, которая исследует знаковые системы самой различной природы. Четко различая понятие знака и знаковой ситуации, семиотика включает такие разделы, как *синтактика* (что связывает знак), *семантика* (что выражает знак), *сигматика* (что обозначает знак) и *прагматика* (что дает знак). Синтактический аспект информации, связанный со структурными и статистическими оценками, в основном рассматривается в информатике и вычислительной технике. Сигматический аспект рассматривает *теория сигналов и кодирования*. Знаковые системы благодаря своей гибкости способны обеспечить разнообразные запросы пользователей. Функциональное единство семантики и прагматики имеет широкие перспективы: появляется возможность установления аналогий между функционированием систем естественного и искусственного происхождения.

Ярким примером взаимодействия естественных, общественных и технических наук является разработка лингвистического обеспечения компьютеризации.

Имитационное моделирование — наука, в которой создаются и используются специальные приемы воспроизведения процессов, протекающих в реальных объектах, в тех моделях этих объектов, которые реализуются в вычислительных машинах.

Теория принятия решений изучает общие схемы, используемые при выборе решения из альтернативных возможностей (в условиях неопределенности). *Теория игр* изучает модели, в которых выбор происходит в условиях конфликта или противоборства. *Математическое программирование* рассматривает проблемы принятия оптимальных решений с помощью математического аппарата.

Искусственный интеллект — одно из молодых и перспективных направлений информатики, появившееся во второй половине XX в. на базе вычислительной техники, математической логики, программирования, психологии, лингвистики и других отраслей знаний. Объектами его изучения являются межпредметные процедуры (метапроцедуры), используемые при решении задач, традиционно называемых интеллектуальными. Проникая в тайны творческой деятельности людей, искусственный интеллект создает программные и программно-аппаратные модели таких метапроцедур.

Информационные системы применяются для анализа и прогнозирования потока информации, исследования способов ее представления, хранения и извлечения. Актуальным является также создание информационно-поисковых систем, систем хранения, обработки передачи информации, в состав которых входят информационные базы данных, терминалы, средства связи. *Операционные системы* связаны с разработкой и производством компьютеров.

Мы стоим на пороге информационной индустриализации общества. Отсюда возникают социальные, правовые, технические проблемы, такие, например, как новые компьютерные технологии обучения, автоматизированные обучающие системы, автоматизированные рабочие места и др.

Для нас представляют интерес все эти направления современной информатики. Теперь мы сможем осознать место дискретной математики в системе знаний, необходимых для тех, кто связал свою жизнь с компьютером.

Для представителей многих специальностей, особенно для программистов, существенное значение в будущей профессиональной деятельности имеет знание *классической логики*, так как она образует математическую основу информатики.

На знаниях законов логики базируются принципы алгоритмизации, которые лежат в основе программирования. Фундаментом всей вычислительной техники и автоматики является преобразование двоичных сигналов, анализ, проектирование и использование логических схем. Основу современной математической логики составляют исчисление высказываний и исчисление предикатов. Любой язык программирования базируется на *исчислении высказываний* и *исчислении предикатов*. В частности, на языке программирования «Пролог» анализируются различные виды дедуктивных умозаключений, выводятся достоверные следствия из них. Широко применяются логические методы для построения *баз данных*. Активно используются знания логики в развитии современных направлений информационных наук. Например, ряд проблем искусственного интеллекта невозможно решить без знаний основ классической логики. Рассмотрим их более подробно.

Представление знаний — методы и приемы формализации информации из различных областей знаний для их хранения, классификации, обобщения и применения при решении конкретных задач.

Моделирование рассуждений — изучение и формализация различных умозаключений и их использование при решении задач средствами ЭВМ.

Методы *диалогового общения* человека и машины. Специфика работы программистов заключается в том, что оппонентом в диалоге выступает компьютер. В него заложены программы, обраба-

тывающие только точно сформулированную информацию. В процессе ее обработки возникают два варианта диалогового общения:

- ЭВМ самостоятельно задает вопросы по полученной информации согласно заложенной в нее программе;
- компьютер задает вопросы, которые заложены заранее в заготовленную программистом модель беседы.

В процессе диалогового общения программист должен знать виды вопросов и ответов для составления программы, владеть правилами построения точных, непротиворечивых, логически выстроенных и адекватных ситуациям формулировок.

Часто приходится обрабатывать информацию, полученную в результате всевозможных *статистических обобщений*, социологических опросов и т.д., с помощью средств теории вероятностей и математической статистики. При этом понадобятся умения формулировать гипотезы о видах распределений, проверять их жизнеспособность, истинность, отражение реальности. Поэтому программистам важно ориентироваться в различных видах *индуктивных умозаключений* и уметь отличать *достоверные* выводы от *вероятностных*, т.е. применять в работе знания *классической логики*.

Объектом исследования дискретной математики являются дискретные **множества** — совокупность, набор некоторых элементов. Поэтому начнем с самого общего глубоко абстрактного раздела этой науки — теории множеств и отношений, которая стала интенсивно развиваться с внедрением вычислительной техники. Простейшие представления о множествах впервые появились в связи с исследованиями в области карточных игр и возникновением комбинаторики и дискретной теории вероятностей.

Глава 1
МНОЖЕСТВА

В этой главе мы систематизируем имеющиеся со школы представления об универсальном языке теории множеств, познакомимся с видами множеств и отношений между ними, узнаем, как сравнивать конечные и бесконечные множества и как подсчитывать число их элементов. Эта глава должна стать тем словарем, с помощью которого можно свободно понимать эту книгу и другие математические тексты.

1.1. Общие понятия теории множеств

> Сегодня мы знаем, что, логически говоря, возможно вывести почти всю современную математику из единого источника — теории множеств.
>
> *Н. Бурбаки*

Язык теории множеств. Множество — одно из основных понятий современной математики, с которым каждый человек знаком со школьной скамьи. «Множество решений уравнения или неравенства», «множество точек на плоскости», «множество действительных чисел» и т.д. — привычные словосочетания, не требующие дополнительных рассуждений и определений.

? Что же такое множество?

Понятия множество, элементы множества — первичные базисные неопределяемые понятия, на которых строится теория множеств.

Совокупность элементов, объединенных некоторым признаком, свойством, составляет понятие **множество**. Например, *множество* книг в библиотеке, *множество* студентов в группе, *множество* натуральных чисел $\mathbb{N}$ и т.д.

Запись $a \in M$ означает: элемент *a принадлежит* множеству *M*, т.е. элемент *a* обладает некоторым признаком. Аналогично $a \notin M$ читаем как: элемент *a не принадлежит* множеству *M*.

Множество считается **заданным**, если или *перечислены* все его элементы, или *указано свойство*, которым обладают те и только те элементы, которые принадлежат данному множеству. Первый вариант будем записывать так: $M = \{m_1, m_2, ..., m_k\}$, например, $M = \{0, 1\}$. Последний вариант будем записывать так: $M = \{b | P(b)\}$. Такая запись читается как: M состоит из тех (всех) элементов b, которые обладают признаком P. Например, $M = \{n | n \in \mathbb{N}, n < 5\}$ означает: M составляют только те натуральные числа, что меньше пяти. Само свойство P будем называть **характеристическим**. В качестве характеристического свойства может выступать указанная для этого свойства *порождающая процедура*, которая описывает способ получения элементов нового множества из уже полученных элементов или из других объектов. Тогда элементами множества считаются все объекты, которые могут быть получены с помощью этой процедуры. Например, множество $M_2^n = \{1, 2, 4, 8, 16, 32, ...\}$ всех чисел, являющихся неотрицательными степенями числа 2 ($M_2^n = \{2^i | i \in \mathbb{Z}, i \geq 0\}$), можно задать с помощью порождающей функции по *индуктивным* (подробно см. в гл. 5) правилам:

- $1 \in M_2^n$;
- если $k \in M_2^n$, то $(2k) \in M_2^n$.

Итак, запись $M = \{x | P(x)\}$ означает: множество M состоит из всех элементов x, обладающих признаком P. Например, запись $M = \{x | x^3 + 3x^2 + 2x = 0\}$ означает, что множество M содержит только корни данного уравнения, т.е. числа $\{0; -1; -2\}$. Запись $Z = \{X | |OX| \leq 4\}$ означает, что для любых X расстояние OX меньше или равно 4, т.е. множество всех точек, для которых расстояние до X не больше 4, есть шар с центром в точке O и радиусом $R = 4$. Запись $A = \{x | x \geq 7, x \in \mathbb{N}\}$ читается так: для любых натуральных x, начиная с 7. Отметим, что в записи $M = \{x | P(x)\}$ переменная x является «немой», т.е. несущественной: от нее ничего не зависит. Можно было бы употребить любую другую букву, например y, и все равно это было бы «множество всех элементов, обладающих признаком P», а как называть элементы — несущественно: главное, чтобы они обладали признаком.

Если множество не содержит элементов, обладающих характеристическим признаком, то оно называется **пустым** и обозначается $\varnothing$. Например, множество целых решений неравенства $5 < x < 6$ является пустым: $K = \{x | x \in \mathbb{Z}, 5 < x < 6\} = \varnothing$. Пустым будет множество действительных решений уравнений $x^2 + 25 = 0$ и $5^{2x-3} = -1$. Множество, не являющееся пустым, называется непустым.

? Всегда ли удается, соблюдая все правила, задать множество? Например, как задать множество всех множеств? Будет ли такое множество содержать себя как отдельный элемент, ведь по указанному характеристическому свойству оно должно содержать все возможные множества, а значит, и себя?

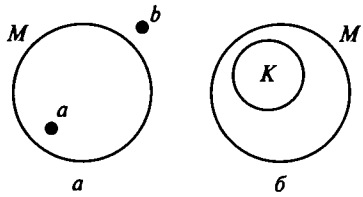


Рис. 1.1. Иллюстрация кругами Эйлера:

а — элемент *a* принадлежит множеству *M*, элемент *b* не принадлежит множеству *M*; *б* — подмножество *K* множества *M*

Изображение множеств. Множества удобно изображать с помощью *кругов Эйлера* (диаграмм Венна). Элементы множества изображаются точками внутри круга, если они принадлежат множеству ($a \in M$ на рис. 1.1, *а*), и точками вне круга, если они множеству не принадлежат ($b \notin M$).

Будем также использовать символы $\forall x$ вместо слов «для любых x», «каждый элемент x» и $\exists x$ вместо слов «существует x». Значение этих символов и правила их употребления будут подробно разъяснены в гл. 4 и 5.

Из множества M можно выделить его часть (также выделением нового характеристического свойства или перечислением элементов) — множество K, все элементы которого обладают таким же признаком, как и элементы множества M. Множество K называют **подмножеством** множества M и обозначают $K \subset M$ (рис. 1.1, *б*). Более строго: множество K называется **подмножеством** множества M ($K \subset M$), если для любого $x \in K$ выполняется $x \in M$ (т.е. $\forall x \in K$ влечет $x \in M$).

Например, добавляя к множеству однозначных целых чисел $A = \{0, 1, ..., 9\}$ признак «число делится на 3», получаем множество $B = \{0, 3, 6, 9\} \subset A$.

Так, множество целых чисел $\mathbb{Z}$ является подмножеством множества рациональных чисел $\mathbb{Q}$. Для числовых множеств справедливо соотношение: $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$, где $\mathbb{N}$ — множество натуральных чисел, $\mathbb{Q}$ — рациональных, $\mathbb{R}$ — действительных, $\mathbb{C}$ — комплексных чисел. Для любого непустого множества M можно сразу указать два его подмножества независимо от состава и структуры M: это оно само и пустое. Очевидно, пустое множество содержится (является подмножеством) в любом множестве.

Также необходимо учитывать различие в употреблении знаков *включения* ($\subset$) и *принадлежности* ($\in$) для множества множеств.

Например, M — множество всех факультетов в нашем колледже, а K — факультет программирования. Хотя K само является множеством (состоит из студентов и сотрудников — преподавателей, администрации и др.), верна запись $K \in M$, так как факультет K является элементом всего множества M. Запись $K \subset M$ неверна, так как множества K и M содержат разные элементы: K — людей, M — факультеты. Однако, если рассмотрим множество O — совокупность людей со всех факультетов (например, при всеобщем голосовании по насущному вопросу), то, безусловно, $K \subset O$.

Универсальным называют множество U, состоящее из всех возможных элементов, обладающих данным признаком. Например,

множество планет Солнечной системы U = {Земля, Марс, Венера, Юпитер, Сатурн, Уран, Плутон, Меркурий, Нептун}. Заметим, что понятие универсального множества четко не определено, т. е. некорректно. U можно включить в другое множество W, и оно тоже будет универсальным. Например, долго считалось, что множество действительных чисел $\mathbb{R}$ универсально (т. е. описывает всю математику), пока не открыли поле комплексных чисел $\mathbb{C}$ и надкомплексные числа и не поняли, что не существует универсального числового множества. Тем не менее там, где область объектов не выходит за рамки некоего множества, иногда бывает удобно оперировать с этим термином. Ведь ржаное поле — вселенная для мыши.

Равными называют два множества A и B, состоящие из одинаковых элементов: $A = B$. Например, равны множества решений уравнений $4x - 8 = 16$, $x/15 = 2/5$ и $5^{x-3} = 125$, так как их решением является одно и то же число 6.

Равны множества букв, из которых составлены слова «навес» и «весна». Равны множества корней уравнения $x^2 = 1$ и множество $M = \{(-1)^k,\ k = 0,\ 1,\ 2,\ \ldots\}$. Поэтому задача «решить уравнение», знакомая с детства, в реальности означает «решить уравнение в каком-то множестве». Так, уравнение $x^2 + 1 = 0$ не имеет действительных корней: $\{x \mid x^2 + 1 = 0,\ x \in \mathbb{R}\} = \varnothing$, но имеет два комплексных корня $x = i$, $x = -i$: $\{x \mid x^2 + 1 = 0,\ x \in \mathbb{C}\} = \{i, -i\}$.

Равенство двух множеств A и B означает также, что $A \subset B$ и $B \subset A$. И наоборот, выполнение свойств $A \subset B$ и $B \subset A$ означает выполнение равенства $A = B$. Эти утверждения равносильны.

Число элементов множества A называется **мощностью** множества и обозначается $|A|$ или $n(A)$. Так, мощность пустого множества равна 0: $n(\varnothing) = 0$, а мощность множества планет Солнечной системы $n(U) = 9$ или $|U| = 9$.

1.2. Основные операции над множествами

Все правила достойного поведения давным-давно известны, остановка за малым — за умением ими пользоваться.

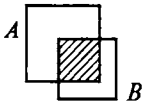
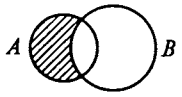
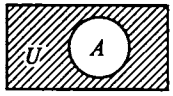
Б. Паскаль

Введение операций над множествами. Из данных множеств A и B можно построить новые множества с помощью операций объединения, пересечения, вычитания и др. (табл. 1.1).

Пример первый. Окружность — множество точек плоскости, равноудаленных от данной (например, O), называемой центром. Математически для ее нахождения надо задать уравнение равно-

Таблица 1.1

Основные операции над множествами

Название операции	Обозначение	Изображение кругами Эйлера	Определение	Символическая запись
Пересечение множеств	$A \cap B$		Те и только те элементы, которые принадлежат *одновременно* A и B	$A \cap B = \{x \mid x \in A \text{ и } x \in B\}$
Объединение множеств	$A \cup B$		Те и только те элементы, которые принадлежат *хотя бы одному* из множеств A и B	$A \cup B = \{x \mid x \in A \text{ или } x \in B\}$
Разность множеств	$A \backslash B$		Те и только те элементы множества A, которые *не* принадлежат B	$A \backslash B = \{x \mid x \in A \text{ и } x \notin B\}$
Дополнение к множеству A	$\overline{A} = A' = U \backslash A$		Те и только те элементы, которые *не* принадлежат множеству A (т. е. дополняют его до универсального U)	$\overline{A} = \{x \mid x \notin A\} = U \backslash A$
Симметрическая разность	$A \Delta B$		Те и только те элементы, которые принадлежат одному из множеств: A *либо* B, но не являются общими элементами	$A \Delta B = (A \backslash B) \cup (B \backslash A) = (A \cup B) \backslash (A \cap B)$

удаленности $L = \{M(x, y, z) \mid (x - a)^2 + (y - b)^2 + (z - c)^2 = r^2\}$ (а это уравнение шара) и уравнение плоскости, проходящей через центр O с координатами $(a; b; c)$: $P = \{M(x, y, z) \mid A(x - a) + B(y - b) + C(z - c) = 0\}$. Окружностью (E) будет множество точек, принадлежащих и шару (L), и плоскости (P), т.е. их пересечение: $E = L \cap P$. Поэтому для нахождения этих точек надо решить систему двух уравнений. Итак, окружность $E = \{M(x, y, z) \mid (x - a)^2 + (y - b)^2 + (z - c)^2 = r^2, A(x - a) + B(y - b) + C(z - c) = 0\}$.

Пример второй. Пусть $A = \{a, b, c, g, e\}$, $B = \{a, c, e, f, r, m\}$, тогда $A \cup B = \{a, b, c, g, e, f, r, m\}$, $A \cap B = \{a, c, e\}$, $A \setminus B = \{b, g\}$, $B \setminus A = \{f, r, m\}$, $A \Delta B = \{b, g, f, r, m\}$. Обратим внимание, что для разности двух множеств не выполняется переместительный закон: $A \setminus B \neq B \setminus A$. Это становится очевидным, если одно множество пустое (например, A), а другое — непустое.

Свойства операций над множествами. Операции над множествами обладают рядом свойств, похожих на свойства операций сложения и умножения чисел. Рассмотрим законы, справедливые для любых множеств A, B, C.

1. $A \cup B = B \cup A$, $A \cap B = B \cap A$ — переместительный закон (коммутативность) для операций объединения и пересечения. Поскольку (а это нетрудно доказать) это свойство справедливо для любого конечного числа множеств, то удобно использовать знаки $\cup$ и $\cap$ для обозначения объединения и пересечения многих множеств. Например, $\bigcup_{i=1}^{n} A_i$ означает объединение n множеств вне зависимости от того, какое из них считать первым, вторым и т.д.

2. $(A \cup B) \cup C = A \cup (B \cup C)$; $(A \cap B) \cap C = A \cap (B \cap C)$ — сочетательный закон (ассоциативность) для операций объединения и пересечения.

3. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ — распределительный закон (дистрибутивность) пересечения относительно объединения множеств.

4. $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ — распределительный закон объединения относительно пересечения множеств.

5. $A \cup A = A$, $A \cap A = A$, $A \subset (A \cup B)$ — законы поглощения.

6. $U = \varnothing'$ и $\varnothing = U'$, т.е. универсальное и пустое множества являются дополнениями друг друга.

7. Если обозначить через A_i все подмножества A_1, A_2, A_3, ..., A_n множества A, то будут справедливы равенства: $A = \bigcup_i A_i$ и $A \setminus \bigcap_i A_i = \bigcup_i (A \setminus A_i)$.

Операция дополнения обладает рядом характерных свойств.

8. Для любого множества $X \subset U$ справедливо $\overline{X'} = (X')' = X$.

9. Для любых двух множеств X и Y справедливо: если $X \subset U$, $Y \subset U$, то $(X \cap Y)' = X' \cup Y'$ или $(X \cup Y)' = X' \cap Y'$.

Докажем последнее свойство.

Пусть $a \in (X \cap Y)'$, что равносильно $a \notin (X \cap Y)$. Это значит, что $a \notin X$ или $a \notin Y$, т.е. $a \in X'$ или $a \in Y'$, поэтому $a \in (X' \cup Y')$.

10. Множество A можно **разбить на классы** непересекающихся подмножеств A_i, если:

- объединение всех подмножеств совпадает с множеством A: $A = \bigcup_i A_i$;
- пересечение любых двух различных подмножеств пусто, т.е. для любых $i \neq j$ выполняется $A_i \cap A_j = \varnothing$.

Рассмотрим примеры.

1. Произвольное множество A разбивается на два дополняющих друг друга подмножества A_1 и $A_2 = A \setminus A_1$, таких, что $A_1 \cup A_2 = A$ и $A_1 \cap A_2 = \varnothing$.

2. Множество двузначных чисел $U = \{10, 11, 12, ..., 98, 99\}$ можно разбить на классы по признаку остатка от деления на 4: класс, порожденный остатком 0: $A_0 = \{12, 16, 20, ..., 96\}$. Аналогично $A_1 = \{13, 17, 20, ..., 97\}$; $A_2 = \{10, 14, 18, ..., 98\}$; $A_3 = \{11, 15, 19, ..., 99\}$.

Разбиение на классы используется для классификации объектов (см. гл. 2 и 3). Так, множество квартир дома может быть разбито на подмножество квартир отдельных подъездов, а множество квартир каждого подъезда — на подмножество квартир одного этажа.

1.3. Соответствия между множествами. Отображения

> Наконец Дронт сказал:
>
> — Победили все! И все получат призы, — добавил он.
>
> ... все наперебой закричали:
>
> — Призы! Где призы? Давай призы!
>
> Бедная Алиса не знала, что ей делать; в растерянности она сунула руку в кармашек и вытащила оттуда коробочку цукатов... Она стала раздавать конфеты всем участникам соревнований, и как раз хватило на всех, кроме самой Алисы...
>
> *Л. Кэрролл*

Основные понятия. Пусть даны два множества $A = \{a_1, a_2, ...\}$ и $B = \{b_1, b_2, ...\}$. Тогда пары (a_i, b_j) задают **соответствие** между множествами A и B, если указано правило R, по которому для элемента a_i множества A выбирается элемент b_j из множества B.

Например, соответствие между элементами множеств $x \in \mathbb{R}$ и $y \in \mathbb{R}$ задает точечное множество $(x_i; y_i)$ координат точек на плос-

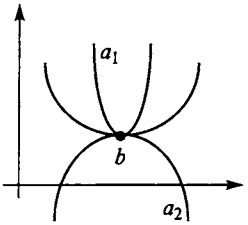


Рис. 1.2. Иллюстрация соответствия R — «вершина параболы»

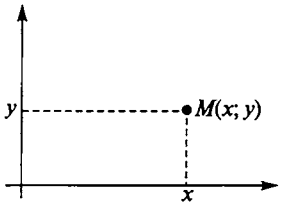


Рис. 1.3. Задание отображения xRy с помощью метода координат

кости; русско-английский словарь устанавливает соответствие значений и написаний слов русского и английского языков.

Пусть задано соответствие R между множествами A и B, т.е. R: $(a; b)$, $a \in A$, $b \in B$. Для некоторого элемента a множества A поставлен в соответствие некоторый элемент b из множества B, который называется **образом** элемента a и записывается $b = R(a)$. Тогда $a = R^{-1}(b)$ — **прообраз** элемента $b \in B$, который обладает свойствами единственности и полноты:

- каждому прообразу соответствует единственный образ;
- образ должен быть полным, так же как полным должен быть и прообраз.

Например, если A — множество парабол, B — множество точек плоскости, а R — соответствие «вершина параболы», то $R(a)$ — точка, являющаяся вершиной параболы a, а $R^{-1}(b)$ состоит из всех парабол a_i с вершиной в точке b (рис. 1.2).

Образ множества A при соответствии R называется **множеством значений** этого соответствия и обозначается $R(A)$, если $R(A)$ состоит из образов всех элементов множества A. Запись: $R(A) = \{b \mid \forall a \in A,\ b = R(a)\}$.

Прообраз множества B при некотором соответствии R называют **областью определения** этого соответствия и обозначают $R^{-1}(B)$, т.е. $R^{-1}(B) = \{a \mid \forall b \in B,\ \exists a \in A\colon R(a) = b\}$; R^{-1} является **обратным** соответствием для R.

Так, для соответствия R, заданного точками координатной плоскости, областью определения является множество точек оси абсцисс, а множеством значений — проекции точек на ось ординат (рис. 1.3). Поэтому для некоторой точки $M(x; y)$ y является образом, а x — прообразом при некотором соответствии R: $Y = R(X)$, $X = R^{-1}(Y)$. Соответствие между множествами X, $Y \subset \mathbb{R}$ удобно представить в виде точки на плоскости с помощью метода декартовых координат.

Пусть задано соответствие R и $Y = R(X)$. Ему соответствует точка M с координатами $(x; y)$ (см. рис. 1.3). Тогда множество точек плоскости, выделяемое отображением R, будет графиком.

Задание отображений. Для описания соответствий между множествами используют понятие **отображения** (**функции**) одного множества на другое.

Для задания отображения необходимо указать:

• множество, которое отображается (**область определения** данного отображения, часто обозначается $D(f)$);

• множество, в (на) которое отображается данная область определения (**множество значений** этого отображения, часто обозначается $E(f)$);

• *закон* или соответствие между этими множествами, по которому для элементов первого множества (прообразов, аргументов) выбраны элементы (образы) из второго множества. Приняты записи $A \xrightarrow{f} B$ или $f\colon A \to B$.

Везде при записи $f\colon A \to B$ будем подразумевать, что отображение f определено *всюду* на A, т. е. A — полный прообраз отображения f, хотя для B такого свойства полноты подразумевать не будем. Запись $f(A)$ означает, что это множество состоит из образов всех элементов множества A: $f(A) = \{f(a) \mid a \in A\}$. Очевидно, что $f(A) \subset B$. Далее будем иметь дело в основном с **однозначными** отображениями, где каждому аргументу поставлено в соответствие не более одного образа.

Напомним, что способ задания отображений в виде формул называется **аналитическим**. Существуют еще табличный и графический способы.

Для задания отображения множеств **табличным** способом принято строить таблицу, в которой первую строку составляют элементы области определения (прообразы вида a), а вторую строку — их образы, т. е. элементы вида $\gamma(x)$ при отображении $\gamma\colon a \to \gamma(a)$, где $a \in A$ (табл. 1.2). Такой способ удобен при достаточно малой мощности прообраза (не более 10). Но все же иногда такой способ задания функции является единственно возможным. Например, публикация в газете многотысячной армии победителей лотереи: аргументом является лотерейный номер, образом — приз.

Графическое представление отображения связано со стрелочными схемами (диаграммами или **графами**), которые подробно рассматриваются в гл. 2. На рис. 1.4 показан пример графического задания отображения множества $A = \{a_1, a_2, a_3\}$ в $B = \{b_1, b_2, b_3, b_4, b_5\}$.

Таблица 1.2

Табличное задание отображения

x	a_1	a_2	...	a_n	...
$\gamma(x)$	$\gamma(a_1)$	$\gamma(a_2)$	...	$\gamma(a_n)$	...

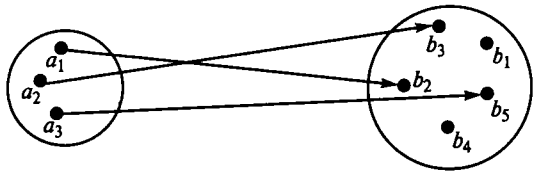


Рис. 1.4. Графическое задание инъективного отображения множества A в B

Отображения $f\colon A \to B$ и $g\colon A \to B$ называются **равными**, если $\forall x \in A\ f(x) = g(x)$.

Виды отображений. Различают два основных вида однозначных отображений (функций). По мощности они делятся на *сюръективные* и *инъективные* (рис. 1.5).

Отображение множества A **на** множество B, при котором каждому элементу множества B соответствует единственный элемент множества A, называется **взаимно-однозначным** соответствием между двумя множествами, или **биекцией**.

На рис. 1.6 изображен график биективного отображения $y = f(x)$, $f\colon \mathbb{R} \to \mathbb{R}$.

Пусть множество A отображается *взаимно-однозначно* на множество B, т.е. $f\colon A \to B$. Тогда отображение f^{-1}, при котором каждому элементу множества B ставится в соответствие его прообраз из множества A, называется **обратным отображением** для f и записывается $B \xrightarrow{f^{-1}} A$ или $f^{-1}\colon B \to A$. Так как одному образу при биекции соответствует в точности один прообраз, обратное отображение будет определено всюду на B и однозначно (отсюда название). Для биекции принята запись: $A \underset{f^{-1}}{\overset{f}{\rightleftarrows}} B$.



Рис. 1.5. Классификация отображений по мощности

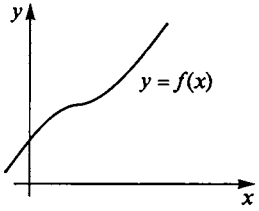


Рис. 1.6. График непрерывного биективного отображения $y = f(x)$

Говорят, что между двумя множествами A и B установлено взаимно-однозначное соответствие f, если элементы этих множеств можно представить в виде пар (a_i, b_k), для которых выполняются два условия:

- $\forall a_i \in A\, \exists b_k \in B$ так, что $f(a_i) = b_k$; $\forall b_j \in B\, \exists a_l \in A$ так, что $f(a_l) = b_j$: все элементы множеств попали хотя бы в одну из пар;
- каждый элемент a_i и b_i попал только в одну из пар.

Если между элементами множеств установлено взаимно-однозначное соответствие, то эти множества имеют одинаковое количество элементов. Говорят, что они **равносильны**, **равномощны**, или **эквивалентны**. На рис. 1.7 показаны примеры прямого и обратного биективного отображения.

Рассмотрим примеры отображений.

1. Отображение γ: $a \to a/2$, где $a \in \mathbb{Z}$, является биекцией множества $\mathbb{Z}$ целых чисел на некоторое множество B. Табличное задание такой биекции можно представить в виде табл. 1.3.

Из табл. 1.3 видно, что каждому элементу множества $\mathbb{Z}$ ставится в соответствие единственный элемент множества B. И наоборот, каждому элементу множества B можно поставить в соответствие единственный элемент из $\mathbb{Z}$. Обратное отображение можно представить аналитически: γ^{-1}: $a \to 2a$ и таблично, поменяв местами строки в таблице.

2. Каждому действительному числу поставим в соответствие его квадрат. Отображение $x \to x^2$ не является взаимно-однозначным соответствием, так как для любого образа $y = x^2$ можно найти два прообраза в области определения: $x = +\sqrt{y}$ и $x = -\sqrt{y}$.

3. Англо-русский словарь устанавливает соответствие между множествами слов английского и русского языков. Такое соответствие не является однозначным, так как каждому английскому понятию соответствуют различные варианты перевода на русский язык, и наоборот.

4. Различные виды кодирования (азбука Морзе, представление чисел в различных системах счисления, шифрованные сообще-

Таблица 1.3

Табличное задание биекции γ: $a \to a/2$

x	...	–2	–1	0	1	2	...
$\gamma(x)$	...	–1	–0,5	0	0,5	1	...

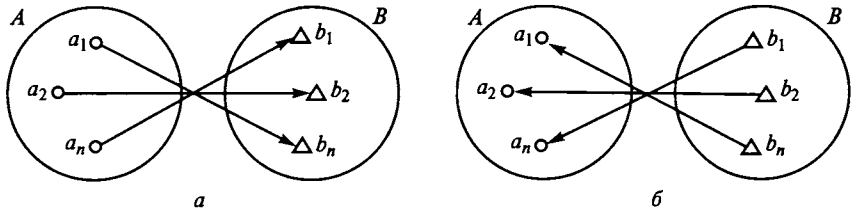


Рис. 1.7. Иллюстрация биекции:
а — прямое отображение *A* в *B*; *б* — обратное отображение *B* в *A*

ния) являются чаще всего примерами взаимно-однозначного соответствия между множествами.

Композиция функций. Пусть заданы отображения f_1: $A \to B$ и f_2: $B \to C$. Отображение f: $A \to C$, при котором каждому элементу $x \in A$ соответствует определенный элемент $z \in C$, такой, что $z = f_2(y)$, где $y = f_1(x)$, называется **произведением**, **композицией**, или **суперпозицией** отображений f_1 и f_2. Также часто употребляют синоним «сложная функция». Обозначается композиция двух отображений f_1 и f_2 как $f = f_2 \circ f_1$, причем именно в таком порядке, поскольку по определению должно быть $z = f(x) = (f_2 \circ f_1)(x) = f_2(f_1(x))$, т.е. сначала производится первое отображение (стоит справа и действует на аргумент x), затем второе (слева).

Пусть $h = g \circ f$. Рассмотрим сначала значение функции $h(x)$ при $x = x_0$:

$f(x_0) = y_0$, f определена в точке x_0 и отображает ее в точку y_0;

$z_0 = g(y_0) = g(f(x_0))$, g определена в точке y_0 и отображает ее в точку z_0.

Таким образом, функция h «транзитом» через точку y_0 отображает точку x_0 в точку z_0: $h(x_0) = g(y_0) = g(f(x_0))$.

С помощью терминов и символов соответствий между множествами такую зависимость можно изобразить в виде схемы (рис. 1.8).

Теперь рассмотрим всю совокупность точек из области определения первой функции f. Здесь все приведенные формулы должны выполняться в каждой точке. На рис. 1.9 с помощью кругов Эйлера показано произведение двух отображений f_1 и f_2, каждое из которых является инъекцией. Пунктирными линиями изображена связь образов и прообразов. Область определения функции f_1 $D(f_1) = A$, область значений функции f_1 $E(f_1) = H \subset B$, область определения функции f_2 $D(f_2) = B$, область значений функции f_2 $E(f_2) = G \subset C$, область значений функции $f_2 \circ f_1$ $E(f_2 \circ f_1) = f_2 \circ f_1(A) = F \subset G$.

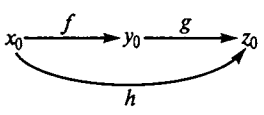


Рис. 1.8. Схема суперпозиции функции $h = g \circ f$ в точке x_0

Рассмотрим суперпозицию числовых функций. Пусть даны произвольные число-

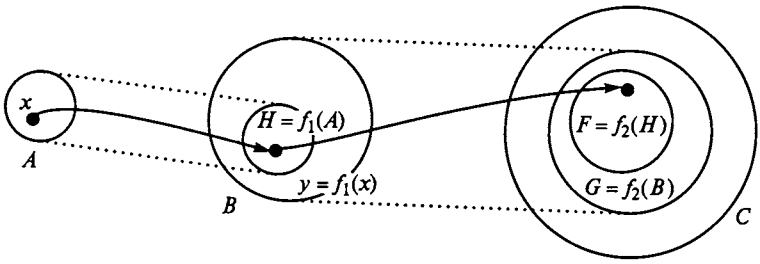


Рис. 1.9. Схема, иллюстрирующая композицию двух инъекций

вые функции $f(x)$, $g(x)$: $\mathbb{R} \to \mathbb{R}$. Суперпозицией $g \circ f$ функций $f(x)$ и $g(x)$ будет новая функция $h(x)$, для которой:

- область определения $h(x)$ состоит из чисел x_0, для которых $f(x_0)$ принадлежит области определения $g(x)$: $D(h(x)) = \{x_0 \mid f(x_0) \in \in D(g(x))\}$;
- значение функции $h(x)$ $\forall x_0 \in D(h(x))$ связано со значениями $f(x)$ и $g(x)$ равенством $h(x_0) = g(f(x_0))$.

Таким образом, только множество $D(g(x)) \cap E(f(x))$ будет далее отображаться посредством функции g, т.е. будет полным прообразом $g^{-1}(E(h))$.

Рассмотрим примеры композиций.

1. Пусть $f(x) = \cos x$, $g(x) = \ln x$, тогда $h(x) = g \circ f$. Разберем подробнее. Каждому x соответствует значение $x \to f(x) = \cos x$. Далее, каждому числу y соответствует $y \to g(y) = \ln y$. Но $y = \cos x$, поэтому, подставляя значение y через x, получаем $g(y) = \ln y = \ln(\cos x) = g(f(x)) = h(x)$. Итак, $h(x)$: $x \to \ln(\cos x)$. Найдем область определения и значения этой функции: $D(f) = \mathbb{R}$, $E(f) = [-1, 1]$, $D(g) = (0, +\infty)$, $E(g) = \mathbb{R}$. Полным прообразом g будет $G = E(f) \cap D(g) = [-1, 1] \cap (0, +\infty) = (0, 1]$. Полным образом тогда будет (так как ln — монотонная функция) $H = g(G) = (-\infty, 0]$. Областью определения будет полный прообраз множества G — множество $I = f^{-1}(G) = \{x \mid 0 < \cos x \leq 1\} = (-\pi/2 + 2\pi n; \pi/2 + 2\pi n]$, $n \in \mathbb{Z}$.

Для сравнения напишем в явном виде отображение $t(x) = f \circ g$. Аналогично предыдущим рассуждениям получим $t(x)$: $x \to \cos(\ln x)$. Для того чтобы доказать нетождественность двух отображений, достаточно указать такое x, что $t(x) \neq h(x)$. Например, $t(1) = \cos(\ln 1) = \cos 0 = 1$. Однако $h(1) = \ln(\cos 1) \approx -0{,}62$.

Поэтому суперпозиция двух функций зависит от порядка записи, т.е. не подчиняется переместительному закону.

Следует сразу указать распространенную ошибку: запись вида $s(x) = f(x) \cdot g(x)$, где знак «$\cdot$» — умножение в $\mathbb{R}$ (например, $(\cos x) \times (\ln x)$), не является композицией или суперпозицией этих *функций*. Дело в том, что произведение функций определено на множестве функций, а выражение $f(x) \cdot g(x)$, знакомое со школы, обо-

значает произведение *значений* каждой из функций в каждой конкретной точке ($x = x_0$). Для сравнения найдем $s(x) = (\cos x) \cdot (\ln x)$ при $x_0 = 1$: $s(1) = (\cos 1) \cdot (\ln 1) = 0$, т.е. это значение не совпадает с уже найденными значениями $t(1)$ и $h(1)$.

Сложные функции представляют собой композицию нескольких простых функций. Но всякая сложная функция, заданная через более простые, разлагается в произведение отображений, т.е. может быть выражена аналитически. Для любых отображений f, g, h справедливо групповое свойство бинарной операции композиции: $f \circ (g \circ h) = (f \circ g) \circ h$, называемое сочетательным законом, т.е. для любого x из области определения выполняется $f(g(h(x))) = (f \circ g)(h(x))$. Это равенство справедливо для тех элементов из $D(h)$, для которых $f(g(h(x))) = (f \circ g)h(x)$ имеет смысл, т.е. определены $h(x)$, $g(h(x))$, $f(g(h(x)))$ и $h(x) \in D(f \circ g)$.

2. Пусть A — это множество людей. Обозначим через f_1 отображение множества A в A, при котором каждому человеку ставится в соответствие его мать, а через f_2 — его отец. Тогда суперпозиции отображений f_1 и f_2 каждому человеку ставят в соответствие:

$g_1 = f_2 \circ f_1$ — дед по материнской линии (отец матери);
$g_2 = f_1 \circ f_2$ — бабушка по отцовской линии (мать отца);
$g_3 = f_1 \circ f_1$ — бабушка по материнской линии (мать матери);
$g_4 = f_2 \circ f_2$ — дед по отцовской линии (отец отца).

? Какими отображениями — инъективными, сюръективными или биективными — являются отображения f_2, f_1 и g_1—g_4?

Отображение e: $A \to A$ называется **тождественным** (**единичным**), если каждому аргументу оно ставит в соответствие себя. Очевидно, такое отображение можно задать на любом непустом множестве. Если $e(x) = x$, то $E(e) = D(e) = A$. Очевидно, что отображение, обратное единичному, также единичное. Рассмотрим произвольную функцию f: $A \to C$. Поскольку и на A, и на C можно ввести тождественное отображение, то, беря композицию отображений e и f, получим два отображения $f_1 = e \circ f$ и $f_2 = f \circ e$. Докажем, что, хотя в первом случае e определено на C, а во втором — на A, отображения f_2 и f_1 равны. Как говорилось, равенство функций определяется действием их на *всех* элементах. Проверим: $\forall x \in A\ f_1(x) = e(f(x)) = f(x)$; и $f_2 = f(e(x)) = f(x)$. Поэтому функции равны. Полученное равенство позволяет сделать вывод, что для такой композиции любой функции с тождественной справедлив переместительный закон: $f \circ e = e \circ f$.

Единичное отображение, естественно, получается как композиция произвольной функции f: $A \to B$ и ей обратной: $e = f^{-1} \circ f$. Действительно, $\forall x \in A\ f^{-1}(f(x)) = x = e(x)$. Но это справедливо, только если обратное отображение является однозначным. Например, при $f(x) = \mathbf{e}^x$, $x \in \mathbb{R}$, $x = \ln(\mathbf{e}^x)$. В противном случае, если $\exists x$,

$y \in A$ такие, что $x \neq y$, но $f(x) = f(y)$, окажется не только $f^{-1}(f(x)) = x = e(x)$, но и $f^{-1}(f(x)) = y \neq e(x)$. Например, $\sqrt{x^2} = |x| \neq x$. Поэтому можно ограничить область определения так, чтобы прямое и обратное отображения были взаимно-однозначными. Тогда будет справедливо $e = f^{-1} \circ f = f \circ f^{-1}$.

Рассмотрим примеры тождественных отображений.

1. $\sqrt{x^2} = x = \left(\sqrt{x}\right)^2$, $x \in [0, +\infty)$.
2. $x = \ln(\mathbf{e}^x) = \mathbf{e}^{\ln x}$, $x \in [0, +\infty)$.
3. $\arcsin(\sin x) = \sin(\arcsin x) = x$, $x \in [-\pi/2, \pi/2]$.

Суперпозиция обладает важным свойством: часто, если функции $f(x)$ и $g(x)$ принадлежат определенному классу, то и их суперпозиция $(f \circ g)(x)$ принадлежит этому классу.

Например, пусть $f(x) = 9x + 1$, $g(x) = 5x - 2$ — линейные отображения, определенные на $\mathbb{R}$. Рассмотрим $z = (g \circ f)$. Тогда $\forall x_0$ имеем: $z(x_0) = g(y_0) = g(f(x_0)) = 5(9x_0 + 1) - 2 = 45x_0 + 5 - 2 = 45x_0 + 3$. Результат также является линейной функцией.

В общем виде для линейных функций $f(x) = ax + b$ и $g(x) = cx + d$ суперпозиция $h(x) = (g \circ f)(x) = c(ax + b) + d = (ac)x + bc + d$ является линейной функцией.

Понятие суперпозиции отображений и соответствий играет важную роль, так как дает возможность установления определенной зависимости одного объекта Z от двух других X и Y.

1.4. Классификация множеств. Мощность множества

— Скажите, пожалуйста, куда мне отсюда идти?

— Это во многом зависит от того, куда ты хочешь прийти, — ответил Кот.

— Да мне почти все равно, — начала Алиса.

— Тогда все равно, куда идти, — сказал Кот.

— Лишь бы попасть куда-нибудь, — пояснила Алиса.

— Не беспокойся, куда-нибудь ты обязательно попадешь, — конечно, если не остановишься на полпути.

Л. Кэрролл

Основной характеристикой множеств является **количество элементов**, содержащихся в этом множестве.

Как известно, число элементов множества M называется его **мощностью** и обозначается $|M|$. Множества A и B называются **экви-**

валентными, или **равномощными**, $A \sim B$, если между их элементами можно установить взаимно-однозначное соответствие (биекцию). Тогда $|A| = |B|$.

Пусть даны два множества A и B.

? Как сравнить эти множества? Каковы критерии оценки множеств?

Если они конечны, то сравнивают их мощности, т.е. количество элементов этих множеств. Обозначим через A, B, C, D соответственно множества букв слов «крот», «корт», «кран» и «рот». Тогда $|A| = |B| = |C| = 4$, $|D| = 3$. Отсюда делаем вывод, что A, B и C имеют равные мощности, а мощность D меньше, чем, например, мощность A: $|D| < |A|$, так как $3 < 4$. Множества A, B и C равномощны. Понятие «равномощные множества» не означает, что они обязательно равны. Например, среди перечисленных равномощных множеств A, B и C множества A и B равны, так как они состоят из одних и тех же элементов, а множества A и C различны. Говорят, что такие равномощные, но не равные множества эквивалентны между собой.

Эталоном для сравнения множеств служит натуральный ряд чисел. Поэтому все числовые последовательности, содержащие различные элементы, эквивалентны натуральному ряду чисел, что видно по индексам их членов. Если $f\colon A \to B$ — инъекция множества A в B, будет справедливо неравенство $|A| \geq |B|$, а если $f\colon A \to B$ — сюръекция множества A на B, то справедливо неравенство $|A| \leq |B|$.

Множества можно классифицировать в зависимости от количества элементов (их мощности) и характера соответствия натуральному ряду чисел (рис. 1.10).

Множество, содержащее конечное число элементов, называется **конечным**. Например, конечным является множество однозначных натуральных чисел {1, 2, 3, 4, 5, 6, 7, 8, 9}. Мощность конечного множества из n элементов равна n. Пустое множество $\varnothing$ по определению не содержит элементов. Оно также является *конечным* и имеет мощность, равную нулю, т.е. $|\varnothing| = 0$. Множество, не являющееся конечным, называется **бесконечным**.

? Но как сравнить бесконечные множества? Ведь для них нельзя указать точное число элементов!

Бесконечное множество, эквивалентное множеству натуральных чисел $\mathbb{N}$, называется **счетным**. Говорят, что все элементы счетного множества можно пронумеровать. В противном случае бесконечное множество будет **несчетным**. Г. Кантор (1878) доказал, что несчетно множество точек, расположенных на отрезке между 0 и 1. По определению Б. Больцано (1837) и Р. Дедекинда (1887) множество называется **бесконечным**, если оно равномощно одному из своих собственных подмножеств. Можно доказать эквивалентность

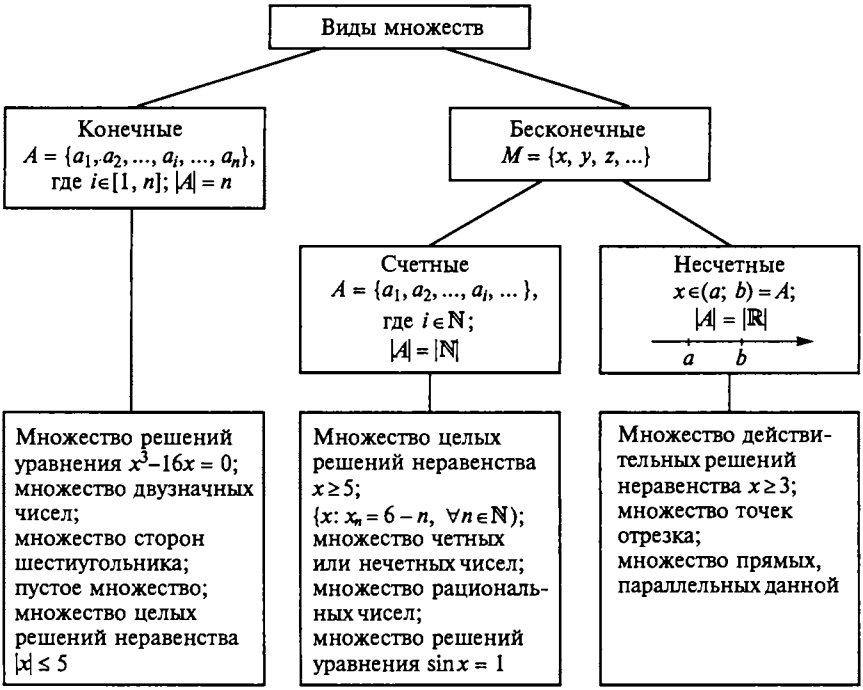


Рис. 1.10. Классификация множеств в зависимости от их мощности и характера соответствия натуральному ряду чисел

этих определений данному нами, но в рамках некоторых допущений (подробнее о подобных доказательствах см. в гл. 5).

Множество A можно отобразить во множество B различными способами. Найдем, сколько существует различных отображений A в B.

Пусть A и B конечны, причем $|A| = n$, а $|B| = k$. Каждый из n аргументов может независимо отобразиться в любой из k элементов B. Поскольку характер множества B не влияет на число функций, возьмем для удобства записи символов множество $B = \{1, 2, ..., k\}$. Зададим функцию g таким образом, что на первом месте стоит $g(a_1)$, ..., на n-м $g(a_n)$. Таким образом, функция может меняться по n «направлениям» (последние два — в плоскости листа) и принимать любые из k значений. Поэтому всего таких записей может быть $\underbrace{k \cdot k \cdot ... \cdot k}_{n} = k^n$. Это и есть искомое число функций.

$$
\begin{array}{llll}
(1, 1, ..., 1, 1, 1) & (1, 1, ..., 1, 2, 1) & ... & (1, 1, ..., 1, k, 1) \\
(1, 1, ..., 1, 1, 2) & (1, 1, ..., 1, 2, 2) & ... & (1, 1, ..., 1, 2, 2) \\
\quad ... & \quad ... & ... & \quad ... \\
(1, 1, ..., 1, 1, k) & (1, 1, ..., 1, 2, k) & ... & (1, 1, ..., 1, 2, k)
\end{array}
$$

Эту же важную формулу можно доказать, используя для каждого элемента множества A отдельный символ (подробнее об алфавитах см. в гл. 6).

Булеаном множества M назовем множество всех его подмножеств, которое обозначается 2^M, т.е. $2^M = \{A \mid A \subset M\}$. Можно доказать, что для конечного множества M мощность булеана $|2^M| = 2^{|M|}$.

В частности, множество всех подмножеств любого конечного множества, состоящего из n элементов, является конечным множеством, состоящим из 2^n элементов. Так как множество целых чисел $\mathbb{Z}$ счетно, то надо показать, что оно эквивалентно натуральному ряду чисел, т.е. надо пронумеровать все его элементы.

? Но как это сделать, если числовая прямая бесконечна и влево от нуля, и вправо (..., –3, –2, –1, 0, 1, 2, 3, ...)?

Для установления соответствия между целыми и натуральными числами будем чередовать положительные и отрицательные числа, начиная от нуля. Тогда индекс элементов полученной последовательности указывает на соответствующее натуральное число: $a_1 = 0$, $a_2 = 1$, $a_3 = -1$, $a_4 = 2$, $a_5 = -2$, $a_6 = 3$, $a_7 = -3$, ..., $a_{2n} = n$, $a_{2n+1} = -n$...

Для конечных множеств справедливо утверждение, которое называется **основной теоремой о конечных множествах**.

Теорема. *Любое конечное множество не эквивалентно никакому его собственному подмножеству, кроме самого себя.*

Следствие. Всякое непустое конечное множество эквивалентно одному и только одному отрезку натурального ряда чисел (1, ..., n).

Тогда мощность конечного множества совпадает с количеством его элементов. Например, мощность множества сторон пятиугольника равна 5.

Из любого бесконечного множества можно выделить счетное подмножество. Так, *счетными* являются множество $\mathbb{Z}$ целых чисел и $\mathbb{Q}$ рациональных чисел. Но множество $\mathbb{R}$ действительных чисел *несчетно*. Счетным будет и множество квадратов натуральных чисел {1, 4, 9, 16, ...}, так как каждому квадрату можно поставить в соответствие единственное натуральное число n, соответствующее его порядковому номеру.

Поскольку для бесконечных множеств нельзя указать число, которое является его мощностью, то будем их сравнивать по эквивалентным им основным множествам $\mathbb{N}$ и $\mathbb{R}$.

Всякое бесконечное множество, равносильное множеству действительных чисел, называется множеством **мощности континуума** (от лат. *continuum* — непрерывный). Поэтому для сравнения мощности множеств с множеством действительных чисел $\mathbb{R}$ будем употреблять условное обозначение $|\mathbb{R}|$ как символ мощности континуума.

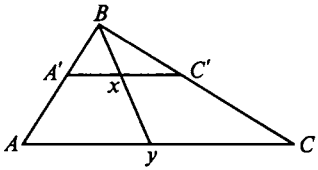


Рис. 1.11. Биекция между средней линией и основанием треугольника

Так, множество точек прямой несчетно и имеет мощность континуума.

? Может ли часть быть эквивалентной целому?

Сравним мощности множества $\mathbb{Z}$ — целых и множества $\mathbb{Q}$ — рациональных чисел. Так как оба они счетные, то имеют одинаковую мощность. Но $\mathbb{Z} \subset \mathbb{Q}$, и получается, что часть ($\mathbb{Z}$) эквивалентна целому ($\mathbb{Q}$)!

Этот удивительный парадоксальный вывод привел в замешательство математиков на рубеже XIX и XX вв. И это не единственный парадокс теории множеств.

? Может ли отрезок быть эквивалентным своей половине?

Еще древнегреческий мыслитель Зенон обратил внимание на удивительный факт, который следует из сравнения длин средней линии треугольника и той стороны, которой эта средняя линия параллельна (рис. 1.11).

Он провел лучи из вершины B на сторону AC и *каждой* точке y отрезка AC поставил в соответствие точку x на средней линии. Видно, что лучи Зенона задают отображение отрезка AC в среднюю линию, причем нетрудно доказать его биективность. Следовательно, одной точке y соответствует единственная точка x, и наоборот. Значит, отрезки *эквивалентны* и мощности множества точек, на них расположенных, равны. Но из геометрии известно, что средняя линия $A'C' = AC/2$. Получается, что отрезок «равен» своей половине! Аналитически такое отображение выглядит еще проще: $y = 2x$ или $x = y/2$. Заметим, что нельзя путать обозначение $|\ |$, символизирующее длину (модуль) в алгебре, с таким же знаком, употребляемым для обозначения мощности множества.

Еще более удивительный вывод получается при сравнении количества всех точек отрезка и квадрата. Георг Кантор (1845—1918) доказал, что множества точек отрезка и квадрата равномощны. Можно занумеровать даже точки плоскости, имеющие целочисленные координаты. Воспользуемся для этого методом решетки: паре натуральных чисел будем ставить в соответствие номер так, как показано в табл. 1.4. При такой нумерации уголком ни одна клетка не останется без номера. Таким образом доказана счетность этого множества.

Таблица 1.4

Метод решетки для нумерации точек плоскости

	1	2	3	4	...
1	$f(1,1) = 1$	$f(1,2) = 2$	$f(1,3) = 4$	$f(1,4) = 7$	...
2	$f(2,1) = 3$	$f(2,2) = 5$	...		
3	$f(3,1) = 6$	...			
...	...				

Задолго до создания самой теории множеств, еще в XVIII в., выдающийся математик, член Петербургской академии наук, уроженец Швейцарии Леонард Эйлер (1707—1783) предложил изображать множества в виде кругов. Затем Кантор открыл, что не все бесконечные множества имеют одинаковую мощность (т.е. эквивалентны между собой). Существуют разные степени эквивалентности, причем мощность любого счетного множества меньше мощности несчетного. Так, мощность множества действительных чисел $\mathbb{R}$ больше мощности множества рациональных чисел $\mathbb{Q}$.

? Как же практически сравнивать множества?

Прямым образом установить равномощность можно, если только множества действительно равномощны. Но реально не всегда возможно, как учили в школе, совместить и проверить их равенство.

Для этого могут помочь некоторые теоремы о мощности множеств:

1. Если $A \subset B$, то $|A| \leq |B|$.
2. Если $A \sim C \subset B$, $B \sim D \subset A$, то $|A| = |B|$.
3. Если $K \subset M$ и K несчетно, то M тоже несчетно.

? Какой вывод сделали ученые, сравнив мощности множеств $\mathbb{Z}$ и $\mathbb{Q}$?

Сравните мощности множеств точек единичного отрезка и множества точек единичного квадрата ($a = 1$). Где больше точек — на отрезке (одномерный континуум) или на плоскости — в квадрате или круге радиуса $r = 1$ (двумерный континуум)?

Так возникает еще один парадокс теории множеств. Нас подводит наша интуиция, так как в теории множеств не срабатывают аналогии. Оказывается, между отрезком и квадратом можно установить взаимно-однозначное соответствие, и эти множества будут эквивалентны.

Кроме математических парадоксов, теория множеств содержит и общелогические парадоксы (рис. 1.12).

Проблемы бесконечности, дискретности и непрерывности интересовали древнегреческих философов начиная с VI в. до н. э. Основные положения теории множеств были заложены в конце XIX в. Георгом Кантором (1845—1918). Важные результаты в области теории множеств были получены Ричардом Дедекиндом (1831—1916). Но абстрактная теория множеств встретила резкое неприятие в математических кругах конца XIX в., так как противоречила привычным математическим представлениям. На рубеже XIX и XX вв. математики столкнулись с так называемыми парадоксами теории множеств, которые противоречили здравому смыслу. Следствием бурного обсуждения вопроса о противоречиях, связанных с теорией множеств, был глубокий кризис основ математики, который затронул весь научный мир. На попытки

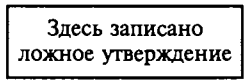


Рис. 1.12. Общелогический парадокс

некоторых математиков отказаться от использования теории множеств в математических рассуждениях один из самых известных ученых этого периода Д. Гильберт заявил: «Никто не может изгнать нас из рая, который создал нам Кантор...»

Теория множеств явилась основой для развития науки в XX в. Ее методы используются в различных математических дисциплинах: теории функций действительного переменного, функциональном анализе и т.д. В настоящее время с нее начинается серьезное изучение математики, поскольку многие математические дисциплины используют ее аппарат.

1.5. Кортежи. Декартовы произведения

Иначе расставленные слова обретают другой смысл, иначе расставленные мысли производят другое впечатление.

Б. Паскаль

В повседневной жизни и математике нам часто приходится иметь дело с упорядоченными множествами — кортежами. Слово кортеж переводится с французского *cortege* как торжественная процессия (например, свадебный кортеж). Треугольник ABC на плоскости задается кортежем из 6 чисел $\langle x_1, y_1, x_2, y_2, x_3, y_3\rangle$, где $A(x_1; y_1)$, $B(x_2; y_2)$, $C(x_3; y_3)$ — координаты вершин. Слова в предложении, буквы в слове, предложения в тексте — все это примеры кортежей. Двоичный код является кортежем, состоящим из цифр 0 и 1.

Пусть A — конечное множество, состоящее из n элементов, f: $A \to \{1, 2, ..., n\}$ — функция, задающая порядок на A, т.е. правило, по которому каждому элементу множества A ставится в соответствие натуральное число от 1 до n, причем одному числу из $\{1, 2, ..., n\}$ соответствует один элемент из A.

Пару $\langle A, f\rangle$ назовем **упорядоченным множеством**, или **перестановкой**, из n элементов. **Кортежем длины** n из элементов множества A (или *n-кой*) называется упорядоченная последовательность $\langle a_1, a_2, ..., a_n\rangle$ элементов этого множества, причем на первом месте стоит прообраз единицы: $a_1 = f^{-1}(1)$, $a_2 = f^{-1}(2)$, ..., $a_n = f^{-1}(n)$, $a_k \in A$, $\forall k \le n$.

Кортежи $\langle a_1, a_2, ..., a_k\rangle$ и $\langle b_1, b_2, ..., b_n\rangle$ называются **равными**, если они имеют одинаковую длину и их элементы с одинаковыми номерами совпадают, т.е. $\langle a_1, a_2, ..., a_k\rangle = \langle b_1, b_2, ..., b_n\rangle$, если $(k = n)$ и для $\forall i$ $a_i = b_i$.

Например, равны кортежи $\langle 2^1, 2^2, 2^3, 2^4, 2^5\rangle = \langle 2, 4, 8, 16, 32\rangle$, так как оба кортежа длины 5 и равны все пары соответствующих элементов данных множеств, $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, $2^4 = 16$, $2^5 = 32$.

Из двух данных кортежей $\langle a_1, a_2, ..., a_i, ..., a_k\rangle$, где $a_i \in A$, длины k и $\langle b_1, b_2, ..., b_j, ..., b_m\rangle$, где $b_j \in B$, длины m можно составить новый кортеж длиной $k + m$, элементы которого $\langle a_1, a_2, ..., a_k, b_1, b_2, ..., b_m\rangle$ принадлежат множеству $A \cup B$. Эта операция называется **соединением кортежей**. Кортеж можно образовать двумя способами, поэтому важно, какой кортеж назван первым. Так, соединив кортежи четных и нечетных однозначных чисел $\langle 0, 2, 4, 6, 8\rangle$ и $\langle 1, 3, 5, 7, 9\rangle$, получим кортеж всех однозначных чисел $\langle 0, 2, 4, 6, 8, 1, 3, 5, 7, 9\rangle$.

Пусть A — конечное множество, элементами которого являются некоторые символы, например цифры, буквы, знаки препинания. Такие множества принято называть **алфавитом над заданным множеством** символов. **Алфавит** есть кортеж попарно различимых символов, называемых **буквами** алфавита. Элементы множества A^n принято называть **словами длины *n*** в алфавите A. Слово над алфавитом есть просто некоторая конечная последовательность символов. Так, шестизначный телефонный номер является словом длины 6 над алфавитом цифр $\{0, 1, 2, ..., 9\}$.

Если число элементов кортежа длины n можно представить в виде суммы $n = n_1 + n_2 + ... + n_k$, то кортеж длины n можно разбить на k кортежей, имеющих соответственно длины $n_1, n_2, ..., n_k$.

Рассмотрим множество B, состоящее из двух элементов: 0 и 1. Кортежи длины m из этих элементов обозначим B^m. Тогда $n(B^m) = 2^m$. Такие кортежи называют упорядоченными наборами или **векторами**. Они имеют широкое применение в дискретной математике.

Каждый такой n-мерный вектор единственным образом определяет *вершину куба*, построенного на *единичных векторах* (рис. 1.13).

В зависимости от величины n кубы могут быть одномерными (*а*), двумерными (*б*), трехмерными (*в*) и т.д.

Вектор из нулей и единиц можно рассматривать как *двоичное представление натурального числа.*

Вектор, состоящий из единиц и нулей, описывает *состояние памяти вычислительных машин*, причем память может содержать числа, тексты, команды и т.д.

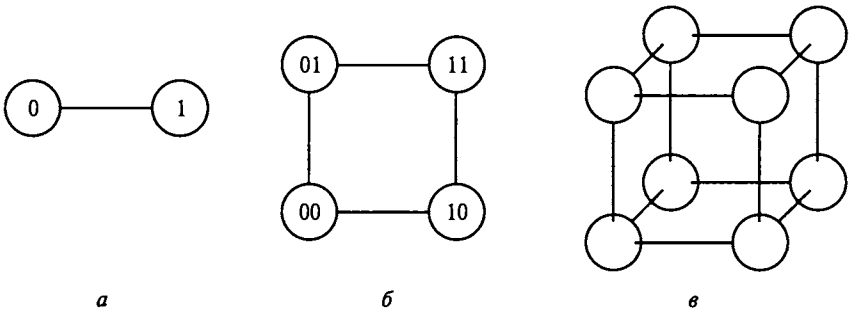


Рис. 1.13. Иллюстрация n-мерного куба:
а — одномерного; *б* — двумерного; *в* — трехмерного

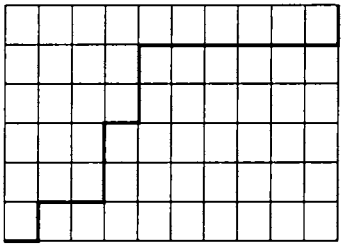


Рис. 1.14. Движение по прямоугольной решетке

Кортежи из нулей и единиц могут быть сообщениями, передаваемыми по некоторому каналу связи с помощью импульсов, каждый из которых принимает одно из двух значений. Сообщения могут характеризовать результаты экспериментов: успех (1) или неудачу (0).

Можно описать путь по некоторой прямоугольной решетке, определив, например, шаг направо через единицу, а шаг наверх — через нуль. Тогда любой путь по такой прямоугольной решетке можно задать кортежем (рис. 1.14).

Кортеж из нулей и единиц используется также для кодировки геометрического изображения. Такая двухцветная черно-белая прямоугольная сетка, состоящая из черных столбцов на белом фоне, может быть представлена в виде вектора. В вычислительной технике широко используется точечное рисование как растровый рисунок.

Практический прикладной характер кортежей проявляется в использовании штриховых кодов (*barcodes*), которые широко применяются в различных информационных системах для сообщения определенной информации о характеристике объекта. Например, *штрих-кодом* снабжены товары на базе или в магазине. Кассовый компьютер быстро считывает зашифрованную в них информацию. Каждый символ специальным образом однозначно кодируется с помощью полосок белого и черного цвета. Кортеж таких полосок однозначно переводится в вектор из 0 и 1.

Декартово произведение. Пусть заданы множества $A_1, A_2, ..., A_n$. **Декартовым (прямым) произведением** этих множеств называется множество $A_1 \times A_2 \times ... \times A_n$, состоящее из всех кортежей $\langle a_1, a_2, ..., a_n \rangle$ длины k, в которых $a_k \in A_k$, где $1 \le k \le n$. Поскольку для задания кортежа важен порядок, то порядок множителей важен и в декартовом произведении.

Например, декартовым произведением множеств $A = \{0,1\}$ и $B = \{X, Y, Z\}$ будет являться множество пар $A \times B = \langle (0; X), (0; Y), (0; Z), (1; X), (1; Y), (1; Z) \rangle$. Скобки для указания пар опускают там, где это не может привести к затруднениям: $A \times B = \langle 0X, 0Y, 0Z, 1X, 1Y, 1Z \rangle$. Если множества $A = \{a_1, a_2, ..., a_k\}$ и $B = \{b_1, b_2, ..., b_m\}$ конечны, то их декартово произведение может быть представлено в общем виде таблицей из m столбцов и k строк.

(a_1, b_1)	(a_1, b_2)	...	(a_1, b_m)
(a_2, b_1)	(a_2, b_2)	...	(a_2, b_m)
...	...	...	...
(a_k, b_1)	(a_k, b_2)	...	(a_k, b_m)

Таблица 1.5

Табличное задание декартова произведения $X \times Y$

	Y		
X	Y_1	Y_2	Y_3
X_1	(x_1, y_1)	(x_1, y_2)	(x_1, y_3)
X_2	(x_2, y_1)	(x_2, y_2)	(x_2, y_3)
X_3	(x_3, y_1)	(x_3, y_2)	(x_3, y_3)
X_4	(x_4, y_1)	(x_4, y_2)	(x_4, y_3)

Например, декартово произведение $X \times Y$, где $X = \{x_1, x_2, x_3, x_4\}$, а $Y = \{y_1, y_2, y_3\}$, можно представить в виде табл. 1.5.

Число элементов в декартовом произведении конечных множеств A и B равно произведению числа элементов множества A на число элементов множества B. Варианты записи: $|A \times B| = |A| \cdot |B|$ или $n(A \times B) = n(A) \cdot n(B)$.

Если $A_1 = A_2 = \ldots = A_n = A$, то пишут $A^n = \underbrace{A \times A \times \ldots \times A}_{n}$ и называют ***n*-й декартовой степенью** множества A.

Например, плоскость является декартовым квадратом двух прямых и обозначается соответственно $\mathbb{R}^2$. В физике пространственно-временной континуум есть декартово произведение $\mathbb{R}^3 \times T$, где $\mathbb{R}^3$ — трехмерное пространство, а T — числовая ось времени.

Декартово произведение не обладает переместительным законом, т.е., вообще говоря, пары (a, b) и (b, a) различны: $X \times Y \neq Y \times X$. Так, различны точки плоскости с координатами (5; 3) и (3; 5). Но для произвольного и пустого множеств справедливо $X \times \varnothing = \varnothing \times X = \varnothing$.

• Примерами декартовых произведений являются таблицы сложения и умножения, все возможные наборы пар координат на плоскости, троек координат некоторой точки в пространстве. Железнодорожный билет тоже является кортежем, а совокупность всех билетов — декартовым произведением множеств паспортов, посадочных станций, станций прибытия, времени и других множеств.

Если число элементов множества X обозначить $|X|$, то справедливо соотношение $|X|^n = |X^n|$.

Свое название декартово произведение получило в честь выдающегося французского математика и философа Рене Декарта (1596—1650), являющегося автором знаменитого метода координат. Вспомните выражение «прямоугольная *декартова* система координат», причем координаты точек в этой системе также являются кортежами. На плоскости двумерные кортежи — это пара вида $(x; y)$, а в пространстве — трехмерные кортежи в виде тройки

чисел $(x; y; z)$, где элементами кортежа являются соответствующие координаты точки.

В программировании декартово произведение встречается в некоторых способах представления данных (массивы, одно-, двух-, трех- и многомерные таблицы и др.).

Изоморфизм. Пусть дано множество M с введенной на нем функцией двух переменных f: $M \times M \to M$. Такая функция называется **бинарной операцией** на M. Обозначим ее знаком $\oplus$: $f(x_1, x_2) \overset{def}{=} x_1 \oplus x_2 \in M$. Аналогично пусть задано множество N со своей введенной бинарной операцией, которую обозначим $\otimes$. Множества $\langle M, \oplus\rangle$ и $\langle N, \otimes\rangle$ называются **изоморфными**, если существует биекция s: $M \to N$, такая, что для любых элементов $x_1, x_2 \in M$ выполнено условие $s(x_1 \oplus x_2) = s(x_1) \otimes s(x_2)$. Обозначение: $\langle M, \oplus\rangle \sim \langle N, \otimes\rangle$. В таком случае взаимно-однозначное соответствие s называется **представлением** множества M во множестве N. Пользуясь изоморфизмом, можно выявлять структуру одного множества (M) по структуре другого, с которым легче работать, которое уже было изучено. Например, рассмотрим множество целых степеней числа 3: $M = \{3^k \mid k \in \mathbb{Z}\} \subset \mathbb{R}$. Умножение (обычное в $\mathbb{R}$, обозначим его $\times$) двух элементов из M не выводит из M. Каждому элементу $x \in M$ поставим в соответствие $\log_3 x$. Таким образом построено взаимно-однозначное отображение f: $x \to \log_3 x = \log_3 3^k = k \in \mathbb{Z}$. Посмотрим, как меняется бинарная операция при действии f: $\forall x_1, x_2 \in M\ f(x_1 \times x_2) = f(3^{k_1} \times 3^{k_2}) = f(3^{k_1+k_2}) = \log_3 3^{k_1+k_2} = k_1 + k_2 = f(x_1) + f(x_2)$, где $+$ означает обычное сложение в $\mathbb{Z}$. Очевидно, что множество $\langle M, \times\rangle$ изоморфно множеству целых чисел $\mathbb{Z}$ с операцией сложения: $\langle M, \times\rangle \sim \langle \mathbb{Z}, +\rangle$. С $\langle \mathbb{Z}, +\rangle$ работать удобнее, так как целые числа, например, занимают меньше места в ЭВМ.

1.6. Отношения. Бинарные отношения и их свойства

> — Почему ты не пьешь больше чаю? — спросил Заяц заботливо.
>
> — Что значит «больше»? — обиделась Алиса. — Я вообще ничего тут не пила!
>
> — Тем более! — сказал Шляпа. — Выпить больше, чем ничего, — легко и просто. Вот если бы ты выпила меньше, чем ничего, это был бы фокус!
>
> *Л. Кэрролл*

Основные понятия. Соответствие между равными множествами $A = B$ называется **отношением** на данном множестве (A). Отношения

в некоторых числовых множествах могут выражаться терминами: «быть равным», «быть больше», «быть не меньше», «быть делителем» и т.д.

Отношения во множестве линий на плоскости могут выражаться терминами: «быть параллельными», «пересекаться», «касаться» и т.д.

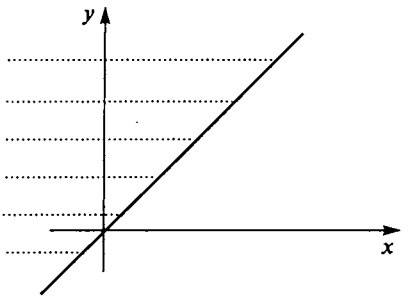


Рис. 1.15. Иллюстрация бинарного отношения R ($x \le y$, где $x, y \in \mathbb{R}$)

В подразд. 1.3 было рассмотрено отношение композиции на множестве всевозможных функций. Отношения являются частным случаем отображения, когда область определения и множество значений совпадают, поэтому все сказанное в подразд. 1.3 справедливо и для отношений.

Назовем ***n*-местным отношением** R на непустом множестве M подмножество $R \subset M^n$. При $n = 2$ отношение R называется **бинарным**. То есть **бинарным** отношением между элементами множеств A и B называют любое подмножество R множества $A \times B$ и записывают $R \subset A \times B$. Для отношения R **обратным** является отношение $R^{-1} \subset B \times A$.

Бинарные отношения принято записывать в виде aRb, где $a, b \in M$. Запись читается как «a и b находятся в отношении R». Например, $a \| b$ (параллельные прямые), $a \le b$ (действительные числа), $a = \log_c b$ и т.д.

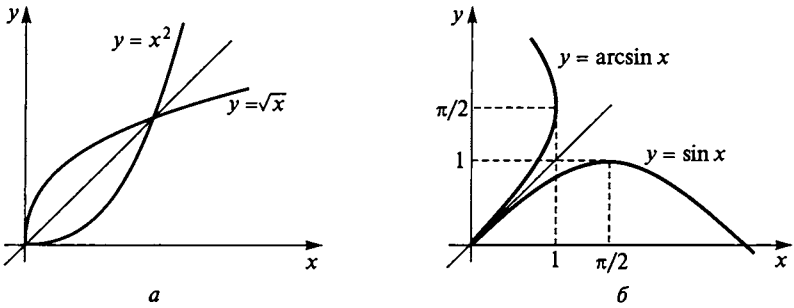
Рассмотрим примеры бинарных отношений.

Бинарное отношение R: $x \le y$, часто встречающееся в практических задачах, показано на рис. 1.15. Заштриховано множество точек, для координат которых это отношение выполняется (истинно).

В школе подробно изучают отношения $x = y$, $x \ne y$, $x > y$, $x \le y$, $x < y$, $y = a^x$, $y = \sin x$, $y = \cos x$, $y = \operatorname{tg} x$, $y = \operatorname{ctg} x$, $y = \log_2 x$, $y = x^n$ и др.

Графики прямых и обратных бинарных отношений, определенных на множестве действительных чисел, симметричны относительно биссектрисы I и III квадрантов. Это свойство обратных бинарных отношений используют при построении графиков обратных функций $y = \log_2 x$ и $y = 2^x$; $y = x^2$ и $y = \sqrt{x}$, где $x \ge 0$ (рис. 1.16, *а*); $y = \sin x$ и $y = \arcsin x$, где $0 \le x < \pi/2$ (рис. 1.16, *б*).

Построение однозначной обратной функции возможно лишь для монотонных функций, поэтому при построении графиков функций, обратных квадратичной и тригонометрической, были введены ограничения. Для функции $y = x^2$ обратную строили не для всей области определения, а лишь для неотрицательных значений x, т.е. на интервале, где функция возрастает.



а *б*

Рис. 1.16. Графики прямых и обратных бинарных отношений:
$а$ — $y = \sqrt{x}$ и $y = x^2$; $б$ — $y = \sin x$ и $y = \arcsin x$

Свойства бинарных отношений. Приведем характерные свойства бинарных отношений, причем заметим, что каждое конкретное отношение может обладать или не обладать некоторыми из указанных свойств.

1. *Рефлективность*: aRa. Например, «быть не больше» на $\mathbb{R}$.

2. *Антирефлективность*. Имеет место, когда отношение не обладает свойством 1 для любых a, например «быть больше», «быть младше» и др.

3. *Симметричность* любых двух элементов. Отношение R на множестве M называется **симметричным**, если для любых $a, b \in M$ одновременно справедливо aRb и bRa (т.е. $R = R^{-1}$). Симметрична параллельность прямых, так как если $a \parallel b$, то $b \parallel a$. Симметрично отношение «быть равным» на любом множестве или «быть взаимно-простым» на $\mathbb{N}$.

4. *Антисимметричность*. Если для несовпадающих элементов $a \neq b$ верно отношение aRb, то ложно bRa. Антисимметричными являются отношения «быть больше», «не меньше» на $\mathbb{R}$, «быть делителем» на $\mathbb{N}$ и др.

5. *Транзитивность*. Если aRb и bRc, то aRc для любых $a, b, c \in M$.
Транзитивны отношения «быть больше», «быть параллельным», «быть равным» и др.

6. *Антитранзитивность*. Имеет место, когда отношение не обладает свойством 5. Например, «быть перпендикулярным» на множестве прямых плоскости ($a \perp b$, $b \perp c$, но неверно $a \perp c$).

7. *Асимметричность*. Ни для одной пары a и b не выполняется одновременно aRb и bRa.

8. *Связность*. Для любых a и b, если $a \neq b$, то aRb или bRa.

Некоторые свойства конкретных бинарных отношений приведены в табл. 1.6.

Рассмотрим основные виды бинарных отношений.

Отношение эквивалентности. Бинарное отношение R называется отношением **эквивалентности**, если оно одновременно обладает

Таблица 1.6

Свойства бинарных отношений

Множества	Отношение	Рефлективность	Симметричность	Асимметричность	Антисимметричность	Транзитивность	Антитранзитивность
Любые	$A \subset B$	+	–	–	+	+	–
Любые непустые	$A \cap B \neq \varnothing$	+	+	–	–	–	–
Любые	$B = A'$	–	+	–	–	–	–
Любые	$a = b$	+	+	–	–	+	–
Любое	$a \neq b$	–	+	–	–	–	+
$\mathbb{N}$	$a \vdots b$, $a = bq$	+	–	–	+	+	–
$\mathbb{R}$	$a > b$	–	–	+	–	+	+
$\mathbb{R}$	$a \geq b$	+	–	–	+	+	+
$\mathbb{R}$	$a < b$	–	–	+	–	+	+
$\mathbb{R}$	$a \leq b$	+	–	–	+	+	+
Прямые плоскости	$a \parallel b$	+	+	–	–	+	–
Прямые плоскости	$a \perp b$	–	+	–	–	–	–
Векторы $\forall a$, $\forall b$	Коллинеарность $a = \lambda b$	+	+	–	–	+	–
Окружности	Касание	+	+	–	–	–	–
Окружности	Концентричность	+	+	–	–	+	–
$\mathbb{N}$	Взаимная простота	–	+	–	–	–	–
$\mathbb{N}$	$a = b \pmod m$ (сравнение по модулю m)	+	+	–	–	+	–

тремя свойствами: рефлективностью, симметричностью и транзитивностью, т.е. если для любых x, y, z выполняется:

- xRx (рефлективность);
- если xRy, то yRx (симметричность);
- если xRy, а yRz, то xRz (транзитивность).

Обозначение эквивалентных отношений: $a\ Q\ b$ или $a \sim b$, что означает «a эквивалентно b в отношении Q», например, «быть равным на множестве чисел», быть подобным на множестве геометрических фигур.

Непересекающиеся подмножества, на которые разбивается множество M отношением эквивалентности, называются **классами эквивалентности**. Множество классов эквивалентности множества A относительно Q называется **фактор-множеством** и обозначается $A \backslash Q$.

Например, множество всех рациональных чисел $\mathbb{Q}$ можно разбить на классы эквивалентности, для которых a/b — рациональная дробь, где $a \in \mathbb{Z}$; $b \in \mathbb{N}$. Любая дробь c/b будет отнесена к тому же классу тогда и только тогда, когда $ad = bc$, т.е. a/b и c/d эквивалентны, если $ad = bc$ (например, $-2/4 \sim -3/6$). Проверим выполнимость свойств для такого отношения.

Рефлективность. Для любой дроби a/b выполняется равенство $ab = ba$, значит $a/b\ Q\ a/b$.

Симметричность. Если $a/b\ Q\ c/d$, то $ad = bc$, но $bc = ad$, значит $c/d\ Q\ a/b$.

Симметричность равенства произведений влечет за собой симметричность отношений между дробями.

Транзитивность. Известно, что $a/b\ Q\ c/d$, $c/d\ Q\ m/n$. Докажем, что $a/b\ Q\ m/n$, т.е. $an = bm$. Действительно, так как $a/b\ Q\ c/d$, то $ad = bc$, аналогично $c/d\ Q\ m/n$, то $cn = md$. Умножим первое равенство на n, а второе на b, тогда имеем $adn = bcn$ и $bcn = mdb$. По свойству транзитивности $adn = mdb$ или $an = mb$. Известно, что такие дроби классифицируются по элементу, порождающему класс эквивалентности, которым в этом примере является несократимая дробь (например, для $2/4 \sim 3/6 \sim 4/8$ таковой будет $1/2$).

Отношение толерантности. Отношение A на множестве M называется отношением **толерантности**, если оно рефлективно и симметрично. Очевидно, что отношение эквивалентности есть частный случай толерантности, когда к двум перечисленным свойствам добавляется транзитивность. Например, отношение «быть другом» рефлективно, симметрично, но не транзитивно. Таким образом, толерантность является более слабой мерой сходства, чем эквивалентность, но тем не менее помогает выявлять различия в схожих вещах.

Пусть A и B имеют некоторые сходные признаки. Тогда A рефлективно A (признаки не только схожи, но и совпадают). Очевид-

но, что выполняется и симметричность, т.е. порядок рассмотрения сходных объектов не важен.

Однако накопление несущественных различий у некогда сходных объектов может впоследствии привести к их полному различию. Сложно разбить на классы множество, состоящее из сходных элементов, так как размыты границы признаков, по которым они объединяются в подмножества. Как известно, каждый элемент множества несет определенную информацию обо всех его элементах. В случае отношения эквивалентности такая информация об одном элементе достаточно полно характеризует свойства всего множества, а отношение сходства малоинформативно. Тогда предельным случаем сходства является неразличимость (но не одинаковость).

При изучении отношения сходства сначала определяется мера сходства — критерии, а затем исследуется взаимное расположение сходных объектов. Отношение толерантности дает интуитивное представление о сходстве объектов. Понятно, что для толерантности свойство транзитивности излишне: сходство между парами a_1 и a_2, a_3 и a_4, ..., a_{n-1} и a_n не означает, что сходны между собой a_1 и a_n, так как размыты критерии сходства, и для каждой пары они могут быть разными. Например, вспомните эффект детской игры в «испорченный телефон».

Пусть $p \in \mathbb{N}$. Обозначим через S_p совокупность всех непустых подмножеств множества $\{1, 2, ..., p\}$. Два таких подмножества будут толерантными, если у них есть хотя бы один общий элемент.

Отношение порядка. Отношение R называется **отношением порядка** на множестве M, если оно обладает свойствами антисимметричности и транзитивности. Для произвольного отношения порядка принято обозначение $\prec$, означающее **предшествование**.

Множество M, которое обладает отношением порядка, называется **упорядоченным**. Такое определение не противоречит определению конечного упорядоченного множества, а является его обобщением на бесконечные множества. И наоборот, старое определение является частным случаем этого, так как сравнение на множестве происходит за счет естественного упорядочения натуральных чисел, заложенного в определении $\mathbb{N}$. Упорядочено множество цифр в моем телефонном номере и множество букв в вашем имени. Например, слова «ракета» и «карета» составлены из одних и тех же букв, взятых в разной последовательности, и имеют различное значение.

$\frac{\text{Рефлективное}}{\text{антирефлективное}}$ отношение порядка называют отношением $\frac{\textbf{нестрогого}}{\textbf{строгого}}$ порядка и обозначают знаком $\frac{\leq}{<}$.

На множестве M задано отношение $\frac{\textbf{полного}}{\textbf{частичного}}$ порядка, если сравнимы $\frac{\text{все}}{\text{не все}}$ элементы этого множества. Множество M, на котором установлено отношение $\frac{\text{частичного}}{\text{полного}}$ порядка, называется $\frac{\text{частично}}{\text{вполне}}$ упорядоченным.

Отношение нестрогого порядка должно удовлетворять трем условиям:

- рефлективности, т.е. xRx;
- антисимметричности, т.е. если xRy и yRx, то $x = y$;
- транзитивности, т.е. если xRy, а yRz, то xRz.

Отношение нестрогого порядка является объединением отношения строгого порядка и отношения тождественности. Каждому отношению порядка R на множестве M можно поставить в соответствие обратное отношение порядка R^{-1}. Например, отношения «больше» и «меньше» на множестве действительных чисел. Для связного отношения порядка R на множестве M существует ему противоположное $\overline{R}$, причем, если R — отношение нестрогого порядка, то $\overline{R}$ — отношение строгого порядка, и наоборот. Во множестве $\mathbb{R}$ отношение нестрогого порядка $\leq$ противоположно отношению строгого порядка $>$.

Отношение порядка дает возможность *сравнивать* между собой различные элементы множества M. Пусть M — упорядоченное множество с отношением строгого порядка $<$. Об упорядоченной паре $x < y$ говорят, что элемент x **предшествует** элементу y. Пусть M — вполне упорядоченное множество. Тогда, если для элемента x не нашлось предшествующего, то он называется **минимальным**: т.е. не существует элементов y, «меньших», чем x. Символически это записывается так:

$$\overline{\exists y} \in M \; y < x \text{ и } y \neq x.$$

На множестве $\mathbb{N}$ натуральных чисел выполняются лишь свойства антисимметричности и транзитивности. Поэтому на нем установлено отношение полного порядка: для любой пары натуральных чисел единица является предшествующим числом, т.е. минимальным. Можно доказать, что конечное вполне упорядоченное множество содержит единственный минимальный элемент. Например, на множествах чисел $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ отношения $\leq$ и $\geq$ есть отношения нестрогого полного порядка, а отношения $<$ и $>$ есть отношения строгого полного порядка. Отношение $\subset$ есть отношение нестрогого частичного порядка на множестве 2^M (булеан).

Всякий частичный порядок на конечном множестве может быть доведен до полного. То есть существует такое отношение полного порядка, для которого заданное отношение частичного порядка является подмножеством.

Если известно, что A и B — упорядоченные множества с отношениями порядка, аналогичными $<$ и $\leq$, и φ: $A \to B$, то можно дать определение для монотонных функций: функция φ называется

монотонной / **строго монотонной**, если из $x_1 < x_2$ следует, что $\dfrac{\varphi(x_1) \leq \varphi(x_2)}{\varphi(x_1) < \varphi(x_2)}$.

На множестве пар действительных чисел (множество точек плоскости) частичный порядок можно задать с помощью покоординатного предшествования, а полный — если поставить порядок одной из координат как приоритетный.

Функциональные отношения. Функциональными отношениями, или **функцией**, называется бинарное отношение $\varphi \subset A \times B$, для которого каждому элементу x из области определения ставится в соответствие единственный элемент y из области значений. Если отношения φ и ему обратное φ^{-1} являются функциональными, то функция φ определяет **взаимно-однозначное** соответствие, или **биекцию**.

1.7. Элементы комбинаторики

За чертогами формул, забыв о весне,
В мире чисел бродя, как лунатик,
Вдруг гармонию выводов дарит струне,
К звучной скрипке прильнув, математик.

М. В. Бромлей

Теория множеств, несмотря на свою глубокую обобщенность, имеет большое практическое значение. Прикладной аспект теории множеств связан с необходимостью подсчитывать количества различных объектов, учитывая или не учитывая их порядок. Например, метеорологу для составления прогноза погоды на 12 ноября этого года приходится анализировать не только характер последних данных о природных явлениях, но и варианты характеристик погоды на 12 ноября в предыдущие годы. Контролеру ОТК при оценке качества продукции приходится рассматривать подмножество данного множества изделий для определения по нему качества всей партии. Фармацевтам приходится рассматривать различные подмножества химических реактивов для изготовления лекарств, а затем по подмножеству экспериментов делать вывод об их эффективности.

Раздел математики, занимающийся подсчетами количества различных комбинаций между объектами, называется **комбинаторикой**. Все комбинаторные задачи сводятся к подсчету мощности конечных множеств и их отображений.

При решении комбинаторных задач в первую очередь будем применять два важных и полезных правила: правило суммы и правило произведения.

Правило суммы. Обозначим число элементов конечного множества A через $n(A)$. Как известно из подразд. 1.2, чтобы найти число элементов в объединении непересекающихся множеств A и B, надо их сложить. То есть, если множества A и B конечны и $A \cap B = \varnothing$, то $n(A \cup B) = n(A) + n(B)$.

На практике будем применять другой вариант этого правила: пусть элемент α можно выбрать k способами, а элемент β — m способами, причем, если любой способ выбора α отличается от любого способа выбора β (независим), то выбор «α или β» можно сделать $k + m$ способами. Например, если в группе 16 юношей и 14 девушек, то преподаватель может вызвать к доске одного учащегося $16 + 14 = 30$ способами.

Если два множества пересекаются и имеют общие элементы, то количество элементов в их объединении можно найти по формуле $n(A \cup B) = n(A) + n(B) - n(A \cap B)$.

Круги Эйлера помогают увидеть, что пересечение $A \cap B$ содержит одни и те же элементы дважды, поэтому необходимо убрать лишние $n(A \cap B)$ (рис. 1.17, *а*).

Задача 1. Сколько человек в группе занимается спортом, если 9 человек занимаются лыжами и плаванием, а 12 человек — плаванием и волейболом, причем в секцию по плаванию ходят 4 человека из группы (рис. 1.17, *б*)?

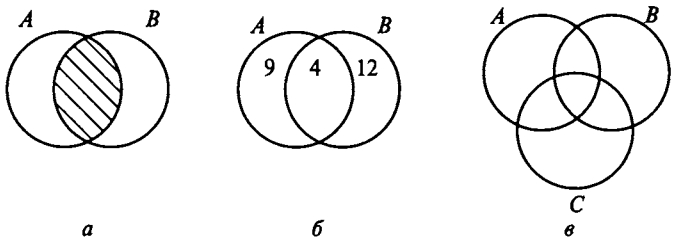
Решение. Имеем: $n(A) = 9$, $n(B) = 12$, $n(A \cap B) = 4$, $n(A \cup B) = n(A) + n(B) - n(A \cap B) = 9 + 12 - 4 = 17$.

Правило суммы для трех множеств имеет вид (рис. 1.17, *в*)

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C).$$

? Как можно объяснить эту формулу с помощью кругов Эйлера?

Напомним, что ключевое слово для применения правила суммы — слово **или**.



а *б* *в*

Рис. 1.17. Схемы, иллюстрирующие правило суммы для объединения множеств:

а — для двух множеств; *б* — к задаче 1; *в* — для трех множеств

Правило произведения. Правило декартова произведения множеств (см. подразд. 1.5) можно распространить и на комбинаторику: если элемент α можно выбрать k способами, а элемент β — m способами, то пару (α, β) можно выбрать $k \cdot m$ способами.

Так как в декартово произведение входят и элементы первого, и элементы второго множества, то при решении задач важно помнить ключевое слово **и**.

Задача 2. В колледже есть три варианта занятий по интересам: творческие объединения (ТО), спортивные секции (СС) и научное студенческое общество (НСО). Каждое направление содержит по четыре вида коллективов: ТО — театральный, музыкальный, танцевальный и КВН; СС — легкая атлетика, лыжи, спортивные игры и плавание. В состав НСО входят естественно-математическое, гуманитарное, техническое и информационное направления. Сколькими способами студенты могут разнообразить свой досуг в колледже после занятий, выбрав коллектив по интересам?

Решение. Так как надо учесть и три основных направления, и то, что в каждом из них по четыре коллектива, то для подсчета общего числа вариантов их нужно перемножить: $n(A) = 3$, $n(B) = 4$, $n(A \times B) = n(A) \cdot n(B) = 3 \cdot 4 = 12$.

Частным случаем правила произведения является число **размещений с повторениями** $\widehat{A_m^k} = m^k$ для подсчета кортежей длины k, составленных из элементов множества X, мощность которого m.

Если $n(X) = m$, то $\underbrace{n(X \times X \times \ldots \times X)}_{k} = (n(X))^k = m^k$.

Перестановки. Рассматривая различные подмножества множества X, состоящего из n элементов, выделим три принципиально разных вида задач для подсчета числа элементов подмножеств.

- Сколькими способами можно переставлять элементы множества, чтобы получить различные кортежи длины n?
- Сколькими способами из всего множества можно выбрать различные кортежи (упорядоченные подмножества) длиной m $(m < n)$?
- Сколькими способами из всего множества можно выбрать различные подмножества длиной m $(m < n)$?

Для решения первой задачи необходимо найти число *перестановок* длины n.

Упорядоченные множества (кортежи), состоящие из n различных элементов, называются **перестановками** (без повторений). Их число обозначается P_n. Фактически эту задачу можно заменить задачей расфасовки n шаров разного цвета в ящик с n свободными местами.

Для определения числа полученных кортежей длины n используем правило произведения. Первый из n элементов можно поставить на n свободных позиций, для второго элемента (при фикси-

рованном первом) останется $n-1$ способов, для третьего (при фиксированных втором и первом элементах) найдется $n-2$ способа и т. д. Для последнего элемента останется единственный вариант (единственное оставшееся место).

Необходимо учесть варианты и для первого, и для второго, и для n-го элемента, поэтому количество вариантов надо перемножить. Тогда $P_n = n(n-1)(n-2) \ldots 1$.

Такое произведение кратко записывается $n!$ (читается «n-факториал»). Формула для числа перестановок имеет вид $P_n = n!$.

Например, из цифр 3, 5, 7, 9 можно составить 4! кортежей, так как $n = 4$, то $P_4 = 4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, т. е. существует 24 различных четырехзначных числа, составленных из этих цифр: 5379, 7359, 9357, ...

Очевидно, $P_n = n! = nP_{n-1}$. Такая формула называется *рекуррентной* и дает возможность подсчитывать число перестановок во множестве $n+1$ элемента через перестановки во множестве n элементов.

Заметим, что $P_1 = 1!$, а $P_0 = 0! = 1$. Если множество содержит один элемент, то и вариант составления кортежей тоже единственный, но если во множестве нет элементов, то это тоже единственный вариант: «кортеж длины 0».

Размещения (без повторений). Упорядоченное подмножество m элементов (кортеж), составленное из всего множества, содержащего n элементов, называется **размещением** (без повторения). Число таких размещений обозначается A_n^m (от фр. *arrangement* — размещение).

Число всех кортежей длины m подсчитаем с помощью перестановок из всех n элементов ($n!$), без тех кортежей, которые не вошли в кортеж длины m. Таких «лишних» кортежей длиной $n-m$ насчитывается $(n-m)!$. Поскольку порядок в «лишнем» подмножестве не зависит от «нужного», действует правило произведения. Поэтому формула размещений (без повторений) примет вид

$$A_n^m = \frac{n!}{(n-m)!}.$$

Задача 3. Сколькими способами из различных нечетных цифр можно составить различные трехзначные числа?

Решение. Нечетных цифр пять: 1, 3, 5, 7, 9. Тогда количество различных трехзначных чисел найдем по формуле $A_5^3 = \frac{5!}{2!} = 60$.

Сочетания без повторений. Сочетания во множестве из n элементов отличаются от размещений тем, что в подмножестве, состоящем из m элементов, в размещениях порядок определен, а в сочетаниях — не важен.

Поэтому **сочетаниями** из n элементов по m называется неупорядоченное подмножество (**выборка**), состоящее из m элементов,

взятых из множества, состоящего из n элементов. Число сочетаний обозначается C_n^m (от фр. *combinaison* — сочетание).

Для подсчета числа сочетаний возьмем формулу размещений A_n^m и поделим ее на число перестановок в подмножестве (так как порядок в сочетаниях не важен).

Тогда формула сочетаний без повторений имеет вид

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n!}{m!(n-m)!}.$$

Отсюда видно, что

$$C_n^m = \frac{n!}{m!(n-m)!} = C_n^{n-m}.$$

Отметим важные частные случаи: при $m = 0$ и $m = n$ имеем $C_n^0 = C_n^n = 1$; при $m = 1$ $C_n^1 = C_n^{n-1} = n$; при $m = 2$ $C_n^2 = C_n^{n-2} = \frac{n(n-1)}{2}$.

Сочетания имеют большое практическое применение в различных разделах математики, например в формуле бинома Ньютона:

$$\forall a,\ b \in \mathbb{R},\ \forall n \in \mathbb{N} \cup 0\ (a+b)^n = \sum_{k=0}^{n} C_n^k a^{n-k} b^k.$$

Задача 4. Сколькими способами могут взойти 3 зерна пшеницы, если посажено 7 зерен?

Решение. По условию задачи порядок в подмножестве из 3 зерен не важен, поэтому по формуле сочетаний имеем: $C_7^3 = \frac{7!}{3!4!} = \frac{5 \cdot 6 \cdot 7}{23} = 35$.

Перестановки с повторениями. Если в кортеже имеются повторяющиеся элементы, то формула $P_n = n!$ уже неприменима, поскольку при переставлении одинаковых элементов кортеж не изменится. Кортеж, имеющий повторяющиеся элементы, называется **перестановкой с повторениями**.

Пусть дан кортеж длины n, состоящий из элементов множества $A = \{a_1, a_2, \ldots, a_m\}$ $(m \le n)$. Отсортируем в кортеже элементы: каждому числу k $(1 \le k \le m)$ ставится в соответствие число n_k, показывающее сколько раз элемент a_k встретился в этом кортеже. Тогда $n = \sum_{k=1}^{m} n_k$. Очевидно, что перестановка одинаковых элементов внутри некоторого k-го сорта не изменит кортежа, т.е. среди $n!$ всех перестановок будет $n_k!$ тождественных. Уменьшим число перестановок из n $(n!)$, убрав повторяющиеся кортежи обратным

действием (разделив на $n_1!$, $n_2!$, ..., $n_m!$). Тогда число перестановок с повторениями, имеющих такой состав, найдем по формуле

$$\widehat{P}_{n_1,n_2,\ldots,n_m} = \frac{(n_1 + n_2 + \ldots + n_m)!}{n_1!\,n_2!\ldots n_m!} \text{ или } \widehat{P}_{n_1,n_2,\ldots,n_m} = \frac{n!}{\prod\limits_{k=1}^{m} n_k!}.$$

Частные случаи

1. Если $n_1 = n_2 = \ldots = n_k = \ldots = n_m = 1$, то $m = n$. Тогда формула должна дать число перестановок без повторений. Действительно, $\widehat{P}_{1,1,\ldots,1} = \frac{n!}{1!\ldots1!} = n! = P_n$.

2. Если $n_1 = n_2 = \ldots = n_{m-1} = 1$, $n_m = n - (m - 1)$, то $\widehat{P}_{1,\ldots,1,n-m+1} = \frac{n!}{(n-m+1)!} = A_n^{m-1}$.

3. Пусть $m = 2$, $n_1 = p$, а $n_2 = n - p$. Тогда число перестановок с повторениями, где все множество из n элементов разделено на две группы из $n_1 = p$ и из $n_2 = n - p$ элементов, — это фактически число сочетаний, так как порядок не важен ни в одном из них. Отсюда сразу получается формула $C_n^p = \widehat{P}_{p,n-p} = \frac{n!}{p!(n-p)!} = C_n^{n-p}$.

Задача 5. Сколькими способами можно расставить белые фигуры на первой линии шахматной доски?

Решение. На первой линии могут находиться король, ферзь, 2 ладьи, 2 коня и 2 слона. Без учета общепринятых шахматных правил составим кортежи длины 8, имеющие указанный состав (1, 1, 2, 2, 2). Тогда число перестановок с размещениями найдем по формуле $\widehat{P}_{(1,1,2,2,2)} = \frac{8!}{1!1!2!2!2!} = 5040$.

Задача 6. Найти число точек пересечения диагоналей выпуклого n-угольника, если никакие три из них не пересекаются в одной точке.

Решение. Минимальный многоугольник с диагоналями имеет четыре вершины ($n \geq 4$). Если взять любые 4 вершины многоугольника, то через них можно провести $C_4^2 = 6$ диагоналей, имеющих пять точек пересечения, но только одна из них лежит внутри многоугольника (рис. 1.18). Наоборот, каждой точке пересечения диагоналей поставим в соответствие четверку вершин многоугольника. Такое соответствие является взаимно-однозначным, а порядок вершин не важен. Поэтому число внутренних точек пересечения диагоналей многоугольника (и число четверок-вершин) можно найти с помощью сочетаний C_n^4.

Применение комбинаторики. Комбинаторный анализ имеет практическое применение в программировании при вычислениях дискретных конечных математических структур.

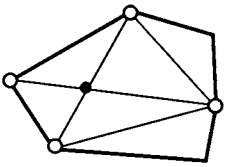


Рис. 1.18. Точки пересечения диагоналей многоугольника

Задача 7. Служба занятости населения располагает базой данных из n записей, каждая из которых содержит одно предложение — наличие вакансий, и один запрос — что востребовано на рынке труда. Необходимо найти вариант трудоустройства безработных в соответствии с имеющимися вакансиями. На языке математики это означает, что нужно найти такие пары записей, в которых предложение первой записи совпадает с запросом второй, и наоборот. Какое количество вариантов пересмотрит компьютер при поиске таких пар?

Решение. Так как надо учесть все n записей предложений, а также $n-1$ запросов, то при сравнении каждого предложения с каждым запросом потребуется $\frac{n(n-1)}{2}$ сравнений. Можно узнать время t, затраченное на обработку информации. Если проверка одного варианта осуществляется за одну миллисекунду, то при заданном n соответствующие результаты можно представить в табл. 1.7.

Для конкретной службы занятости эти результаты означают, что машинная обработка запросов ста клиентов потребует всего 5 с непрерывной работы компьютера. Если на бирже труда 20 000 безработных, то необходимо 55,6 ч для непосредственного поиска вариантов трудоустройства с помощью ЭВМ. Эти знания помогут руководителям службы определить потребность в кадрах и ЭВМ в соответствии с реальной обстановкой в конкретном городе. Точнее говоря, мы получили ответ на вопрос, какое минимальное число сравнений сделает компьютер. Но многое зависит от квалификации программиста, писавшего программу: не заставил ли он машину делать дублирующие сравнения.

Таблица 1.7

Время поиска вакансии в зависимости от объема базы данных

n	100	1000	10 000	15 000	20 000
$\frac{n(n-1)}{2}$	4950	499 500	4 999 500	112 492 500	19 999 000
t, с	4,95	499,5	4999,5	112 492,5	199 990
t, ч	0,0014	0,14	1,4	31,2	55,6

Решение задач, связанных с перебором вариантов, требует предварительного анализа комбинаторных вычислений, поиска соответствующего алгоритма и учета необходимых временных затрат.

Задачи подсчета возможных комбинаций объектов, удовлетворяющих определенным условиям, часто встречаются в практической деятельности и получили название *комбинаторных*.

Многообразие таких задач не всегда удается описать с помощью математических формул. Однако для стандартных распространенных ситуаций способы подсчета вариантов определены.

Задача 8. Разложить n различных деталей в m ящиков. Сколько вариантов таких размещений можно перебрать?

Решение. Поскольку все n деталей могут поместиться и в один ящик, то на языке функций задано сюръективное соответствие между множеством X деталей ($|X| = n$) и множеством Y ящиков ($|Y| = m$), т.е. $\varphi: X \to Y$.

Такое часто встречающееся число подсчетов вариантов называют **размещением с повторением** и обозначают $\widehat{A}_m^n$. Так как каждую из n деталей можно разместить в m ящиков, то необходимо n раз умножать число m, т.е. $\widehat{A}_m^n = m^n$. Напомним, что такой же результат был получен и в подразд. 1.4, когда находили полное число функций $\varphi: X \to Y$.

Задача 9. Сколько различных двоичных чисел длиной 6 можно записать с помощью цифр 0 и 1?

Решение. Размещаем две цифры (0, 1) на шесть мест, т.е. на каждом из шести мест ($m = 6$) может быть одна из двух двоичных цифр. Всего таких вариантов будет $\widehat{A}_2^6 = 2^6 = 64$ двоичных чисел: на каждом из шести мест по два варианта цифр.

Задача 10. Сколько проводится матчей в Чемпионате РФ по футболу в премьер-лиге (16 команд) за сезон?

Решение. Поскольку один матч проводится между двумя командами, каждый матч выделяет подмножество (упорядоченное, так как одна команда играет дома, вторая — на выезде) из двух элементов. По формуле размещений $A_{16}^2 = 16 \cdot 15 = 240$.

Все комбинаторные операции удобно представить в виде схемы (рис. 1.19).

Во второй половине XX в. теория множеств пополнилась новым, перспективным и актуальным для современной практической деятельности направлением, связанным с понятием *нечетких множеств*.

Благодаря создателю этого направления профессору университета г. Беркли (США) Лофти Заде с середины 1960-х гг. появилась реальная возможность приближения математических рассуждений к тем принятым схемам, которые используют люди в процессе непосредственного общения.

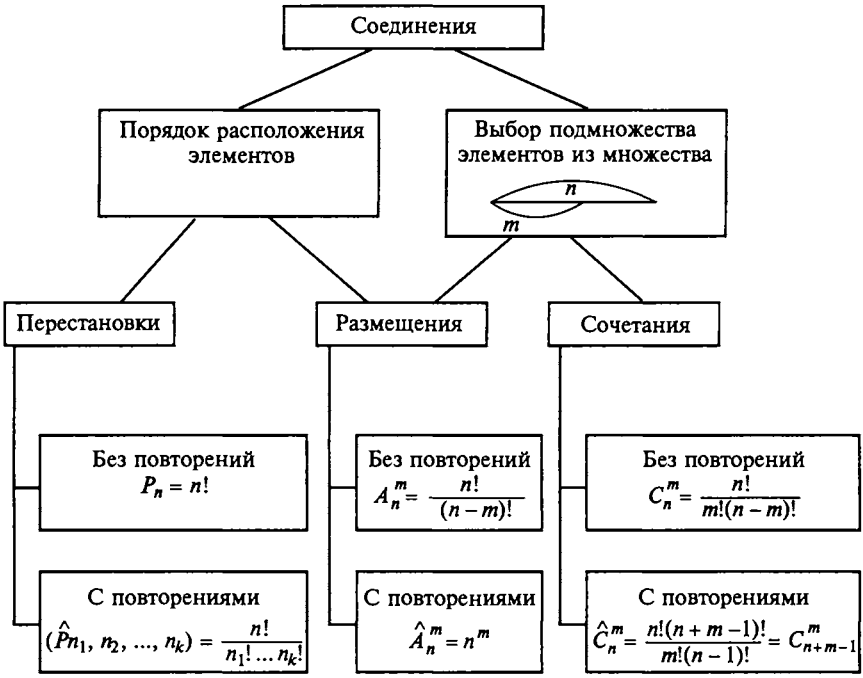


Рис. 1.19. Схема комбинаторных операций

С помощью нечетких множеств удалось преодолеть разрыв между математическими моделями, методами, рассуждениями и принципами решения задач в гуманитарных областях с низким уровнем формализации, оперируя такими новыми математическими понятиями, как *лингвистическая переменная*, *нечеткий вывод* и др.

В отличие от общепринятых нечеткие множества обладают размытыми границами. Их элементы, лингвистические переменные, в повседневной речи служат качественными характеристиками объектов (высокий, умный, новый, большой, сильный и др.).

О принадлежности к некоторому множеству M в теории нечетких множеств можно судить по *функции принадлежности* $\mu = f(m)$, значения которой распределены на отрезке $\mu \in [0; 1]$. Установить принадлежность элемента m множеству удается с помощью схемы

$$f(m) = \begin{cases} 0, & \text{если } m \notin M, \\ \mu, & \text{если } m \in M \text{ с вероятностью } \mu, \\ 1, & \text{если } m \in M. \end{cases}$$

Математика нечетких множеств в настоящее время играет значительную роль в теории принятия решений, искусственном интеллекте и т.д. Анализ и решение плохо формализованных задач требуют экспертных оценок, следовательно, знаний из теории принятия решений.

Нечеткая логика (*fuzzy logic*) — это надмножество классической булевой логики, расширяющее ее возможности и позволяющее вводить неопределенность в логические выводы. Основное отличие нечеткой логики от обычной состоит в том, что она использует не только понятия «истина» и «ложь», но и промежуточные значения истинности.

Не всегда реальные жизненные ситуации дают нам задачи с точным аналитическим описанием их содержания. Задачи, не поддающиеся строгой формализации, привели к появлению новой математической дисциплины — нечеткой логики. Она дает возможность устранить противоречие между неопределенностями реального мира и строгими требованиями четкости в математике.

В настоящее время средствами нечеткой логики решаются задачи из различных областей знаний: распознавание образов, исследование операций, моделирование сложных систем, принятие решений в условиях неопределенности и т.д.

Например, рассмотрим понятие «рост человека». Если рост ниже 150 см, то такого человека считают низкорослым, а выше 180 см — высоким. В зависимости от ситуаций и требований человека с ростом в промежутке от 150 до 180 см могут считать и высоким и низким с определенной долей вероятности. Поэтому четкой границы при делении людей на высоких и низких установить нельзя. Например, человек с ростом 175 см будет низеньким в команде волейболистов, но высоким среди акробатов.

Для выхода из такой неоднозначной ситуации поставим в соответствие каждому числу, характеризующему рост человека (в см), некоторую величину, обозначающую степень принадлежности некоторому множеству. Тогда, характеризуя принадлежность элемента данному множеству, необходимо указать степень принадлежности ему.

Рассмотрим возможности формализации такой задачи.

Определим универсальное множество U, содержащее все возможные элементы. В нашем примере это рост человека, например, в интервале (0; 300).

Установим множество значений μ, выступающее в качестве меры степени принадлежности некоторого элемента из U нечеткому подмножеству M. Промежуточные значения соответствуют определенной степени вероятности (принадлежности), что этот элемент принадлежит подмножеству.

Например, человек с ростом 175 см будет входить в подмножество высоких людей со степенью принадлежности μ_1, а в подмножество низкорослых людей — μ_2. Однако человек с ростом 200 см будет входить в подмножество высоких людей со степенью принадлежности $\mu_1 = 1$, а в подмножество низкорослых людей — $\mu_2 = 0$.

Сформулируем правило, определяющее состав нечеткого множества, — правило, по которому каждому элементу из множества U ставится в соответствие элемент из множества B. Таким образом, определено отображение μ: $U \rightarrow B$.

Комбинаторный анализ, или комбинаторика, — одна из наиболее успешно применяемых в настоящее время областей математики.

Прикладной характер комбинаторики проявляется в ее применении в операционных исследованиях, в статистике и др. В операционных исследованиях она применяется для составления расписаний, определения последовательности промышленных операций, распределения материала, маршрутов транспортных средств и т.д.

1.8. Подстановки

Формулы — не душа математики. Но все-таки язык! Родной язык!

Ю. В. Пухначев, Ю. П. Попов

Рассмотрим множество $E_n = \{1, 2, \ldots, n\}$, $\forall n \in \mathbb{N}$, каждый элемент в котором представлен только один раз. Тогда взаимнооднозначное отображение σ_n: $E_n \to E_n$ множества E_n на себя называется **подстановкой степени** n.

Множество подстановок n-й степени обозначается S_n.

Отношение $E_n \to E_n$ бинарное, поэтому подстановки принято записывать в виде двухрядной матрицы, в первой строке которой записаны прообразы, а во второй — их образы:

$$\sigma = \begin{pmatrix} x_1 & x_2 & & x_i & & x_j & & x_n \\ \sigma(x_1) & \sigma(x_2) & \ldots & \sigma(x_i) & \ldots & \sigma(x_j) & \ldots & \sigma(x_n) \end{pmatrix} \quad i \neq j \leq n,\ x_i \neq x_j.$$

Если прообразы (аргументы) расположены в порядке возрастания (при $i < j$, $x_i < x_j$), запись подстановки такого вида называют **канонической**. Если аргументы не записаны в порядке возрастания, то, переставляя столбцы (при этом сама подстановка не меняется, а изменяется лишь порядок произношения соответствий), можно верхнюю строку привести к упорядоченному виду:

$$\sigma = \begin{pmatrix} 1 & 2 & & x_i & & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(x_i) & \ldots & \sigma(n) \end{pmatrix} \text{ или } (\sigma_1, \sigma_2, \ldots, \sigma_n).$$

По возможности будем пользоваться именно канонической записью. Напомним, что такая запись встречалась, когда мы записывали перестановку из n элементов. Отметим, что прообразом перестановки служит произвольное конечное множество, а прообразом подстановки — обязательно E_n.

Найдем число возможных различных подстановок степени n. Поскольку каждой канонической записи подстановки эквивалентна соответствующая перестановка, то число подстановок n-й степени равно числу перестановок из n элементов, т.е. множество S_n состоит из $P_n = |S_n| = n!$ элементов.

Подстановку вида $\sigma = (1, 2, ..., n) = e$ называют **тождественной**, так как $\forall x \in E_n$, $e(x) = x$. Нетрудно видеть, что это частный случай уже рассмотренного в подразд. 1.3 единичного отображения.

Пусть заданы две подстановки σ_1 и σ_2, причем $E_n \xrightarrow{\sigma_1} \xrightarrow{\sigma_1} E_n' \xrightarrow{\sigma_2} E_n''$. Штрихи означают то же множество E_n, только служат для различия между областями определения и значений. Тогда произведение отображений $\sigma_2 \circ \sigma_1$: $E_n \xrightarrow{\sigma_2 \circ \sigma_1} E_n''$ также является подстановкой и называется **произведением подстановок** σ_1 и σ_2 и записывается $\sigma = \sigma_2 \circ \sigma_1$. Очевидно, что произведение определено для подстановок одинаковой степени.

Как и в других функциях, в подстановке $\sigma = \sigma_2 \circ \sigma_1$ сначала выполняется первая подстановка, а затем вторая действует на результат первой: $\sigma(x) = (\sigma_2 \circ \sigma_1)(x) = \sigma_2(\sigma_1(x))$. Поэтому (за исключением совпадений) для произведения подстановок не выполняется переместительный закон.

Задача 11. Даны две подстановки: $\sigma_1 = \begin{pmatrix} 1 & 5 & 4 & 3 & 2 \\ 3 & 5 & 2 & 4 & 1 \end{pmatrix}$; $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}$. Привести подстановки к канонической записи и найти их произведения.

Решение. Вторая подстановка записана в каноническом виде, первая — нет. Поэтому в верхней строке запишем числа от 1 до 5, а в нижней $\sigma_1(1) = 3$, $\sigma_1(2) = 1$, ..., $\sigma_1(5) = 5$. Итак, $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 4 & 2 & 5 \end{pmatrix}$. Найдем $\sigma_2 \circ \sigma_1$. Сначала выполняется первая подстановка $\sigma_1(1) = 3$, а затем вторая $\sigma_2(3) = 4$, т.е. $\sigma_2 \circ \sigma_1(1) = \sigma_2(\sigma_1(1)) = \sigma_2(3) = 4$. Поэтому можно сразу записывать в матрицу: $\sigma_2 \circ \sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & \dots & \dots & \dots & \dots \end{pmatrix}$. Аналогично найдем остальные образы: $\sigma_2 \circ \sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 5 & 1 \end{pmatrix}$. Теперь найдем $\sigma_1 \circ \sigma_2$: $\sigma_2(1) = 2$, а $\sigma_1(2) = 1$, т.е. $\sigma_1 \circ \sigma_2(1) = \sigma_1(\sigma_2(1)) = \sigma_1(2) = 1$.

В итоге получим $\sigma_1 \circ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 2 & 4 & 3 \end{pmatrix} \neq \sigma_2 \circ \sigma_1$.

Натуральной степенью подстановки σ называется подстановка $\sigma^n = \underbrace{\sigma \cdot \ldots \cdot \sigma}_{n}$, т.е. произведение n функций, каждая из которых есть σ.

Очевидно, что степень подстановки не зависит от порядка множителей.

Задача 12. Дана подстановка $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$. Найти ее степени.

Решение. Имеем: $\sigma^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$; $\sigma^3 = \sigma^2 \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \circ$

$\circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = e$. Тогда очевидно, что $\sigma^4 = \sigma^3 \circ \sigma =$

$= e \circ \sigma = \sigma$ и т.д.

Свойства умножения подстановок.

1. Умножение выполняется только для подстановок одинаковой степени.

2. Для умножения подстановок не выполняется переместительный закон.

3. Для умножения подстановок выполняется сочетательный закон: $\sigma_1(\sigma_2 \circ \sigma_3) = (\sigma_1 \circ \sigma_2)\sigma_3$. Поэтому в дальнейшем не оговаривается количество подстановок при умножении.

4. Подстановка не изменяется, если ее умножить на тождественную:

$$\sigma \circ e = e \circ \sigma = \sigma.$$

5. Поскольку любая подстановка σ — биекция, для нее существует обратная функция, тоже являющаяся подстановкой: можно указать обратную $\sigma^{-1} \in S_n$, т.е. такую, что $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = e$.

Чтобы из подстановки $\sigma = \begin{pmatrix} 1 & 2 & \dots & x_i & & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(x_i) & \dots & \sigma(n) \end{pmatrix}$ получить обратную, нужно поменять местами образы и прообразы, т.е. верхнюю и нижнюю строчки, и, если требуется, привести к каноническому виду.

Например, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 2 & 5 & 3 & 1 & 7 & 4 \end{pmatrix}$. Тогда $\sigma^{-1} = \begin{pmatrix} 6 & 2 & 5 & 3 & 1 & 7 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix} =$

$= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 2 & 4 & 7 & 3 & 1 & 6 \end{pmatrix}$.

6. Обратная подстановка единственна.

7. Подстановка, обратная единичной, также является тождественной.

8. Для того чтобы перемножить степени подстановки σ, достаточно эту подстановку возвести в степень, равную сумме показателей множителей, т.е. $\sigma^m \circ \sigma^n = \sigma^{m+n}$, $\forall m, n \in \mathbb{Z}$.

Следствия

1. Для отрицательных значений показателя k справедливо $\sigma^{-|k|} = (\sigma^{-1})^{|k|}$, $\forall k \in \mathbb{Z}$.

2. Единичная подстановка является нулевой степенью любой подстановки: $e = \sigma \circ \sigma^{-1} = \sigma^0$.

3. n-я степень тождественной подстановки равна тождественной: $e^n = e$, $\forall n \in \mathbb{Z}$.

4. При возведении подстановки в степень справедливо $\sigma^{nm} = (\sigma^n)^m = (\sigma^m)^n$, $\forall n, m \in \mathbb{Z}$.

5. Назовем **порядком** подстановки наименьшее натуральное число λ, такое, что $\sigma^\lambda = e$, $\lambda(\sigma) = \min\{p | \sigma^p - e\}$. Порядок определен для любой подстановки (докажите). В задаче 12 для подстановки

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \quad \lambda = 3.$$

Тогда справедливо соотношение $\sigma^k = \sigma^{k+n\lambda}$, $\forall k, \lambda, n \in \mathbb{Z}$.

Действительно, $\sigma^k = \sigma^k \circ e = \sigma^k \circ \sigma^{n\lambda} = \sigma^{k+n\lambda}$.

6. Подстановка, обратная произведению подстановок σ_1 и σ_2, равна произведению обратных подстановок, взятых в обратном порядке, т.е. $(\sigma_2\sigma_1)^{-1} = (\sigma_1)^{-1} \circ (\sigma_2)^{-1}$.

Инверсией подстановки называется переставление (рокировка) двух соседних значений в нижнем ряду канонической записи. Следует помнить, что при этом подстановка меняется.

Пусть n — число инверсий, приводящих подстановку к единичной. Тогда функция ε: $\sigma \to (-1)^n$ называется четностью (знаком) подстановки σ. Иногда употребляется запись $\operatorname{sgn}(\sigma)$. Если $\varepsilon(\sigma) = 1$, то подстановка называется **четной**, если $\varepsilon(\sigma) = -1$, то подстановка называется **нечетной**. Очевидно, единичная подстановка является четной. К единичной подстановку можно приводить разными способами с разным числом инверсий, но при этом четность числа инверсий и четность подстановки не изменятся.

Задача 13. Найти число инверсий и четность подстановки $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 3 & 1 & 2 & 5 \end{pmatrix}$.

Решение. Имеем

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 3 & 1 & 2 & 5 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 1 & 3 & 2 & 5 \end{pmatrix} \xrightarrow{2}$$

$$\xrightarrow{2} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 1 & 4 & 3 & 2 & 5 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 4 & 3 & 2 & 5 \end{pmatrix} \xrightarrow{4} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 4 & 2 & 3 & 5 \end{pmatrix} \xrightarrow{5}$$

$$\xrightarrow{5} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 2 & 4 & 3 & 5 \end{pmatrix} \xrightarrow{6} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 6 & 4 & 3 & 5 \end{pmatrix} \xrightarrow{7} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 6 & 3 & 4 & 5 \end{pmatrix} \xrightarrow{8}$$

$$\xrightarrow{8} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 6 & 4 & 5 \end{pmatrix} \xrightarrow{9} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 6 & 5 \end{pmatrix} \xrightarrow{10} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}.$$

Тогда $n = 10$, поэтому $\varepsilon(\sigma) = (-1)^{10} = 1$, т.е. подстановка четная.

Следствие. Одна инверсия меняет знак подстановки и знак функции $\varepsilon(\sigma)$.

Любая перемена σ_i и σ_j местами называется **транспозицией**. Очевидно, что инверсия — частный случай транспозиции. Поэтому $\frac{\text{нечетное}}{\text{четное}}$ число транспозиций $\frac{\text{меняет}}{\text{не меняет}}$ знак подстановки. Приведем σ из задачи 13 к единичной подстановке с помощью транспозиций:

$$\sigma = \begin{pmatrix}1&2&3&4&5&6\\6&4&3&1&2&5\end{pmatrix} \xrightarrow{1} \begin{pmatrix}1&2&3&4&5&6\\1&4&3&6&2&5\end{pmatrix} \xrightarrow{2} \begin{pmatrix}1&2&3&4&5&6\\1&2&3&6&4&5\end{pmatrix} \xrightarrow{3}$$

$$\xrightarrow{3} \begin{pmatrix}1&2&3&4&5&6\\1&2&3&4&6&5\end{pmatrix} \xrightarrow{4} \begin{pmatrix}1&2&3&4&5&6\\1&2&3&4&5&6\end{pmatrix}.$$

В этом случае четное число транспозиций ($n = 4$) также приводит к четной подстановке. Отсюда видно, что с помощью транспозиций находить четность проще.

Следствие. Число четных и нечетных подстановок n-й степени одинаково и равно $n!/2$.

Подстановки тесно связаны с одним из самых общих высокоабстрактных разделов математики — с теорией групп. Этот раздел математики дал возможность осуществить перенос фундаментальных базовых выводов теории множеств на отдельные множества, обладающие иной природой, но по своим свойствам схожих между собой. В частности, дискретная симметрия в природе описывается различными подмножествами множества всех подстановок $S_2 \times S_3 \times \ldots \times S_n \times \ldots$. Рассмотрим симметричное отношение R. Симметричность означает, что, если в выражение xRy с произвольными x, y вместо x вставить y, а вместо y — x, то отношение также будет выполняться: $xRy = yRx$. Это означает, что выражение xRy инвариантно (не меняет своего значения) относительно подстановки $\begin{pmatrix}x&y\\y&x\end{pmatrix}$. Аналогично будем называть выражение $F(x_1, x_2, \ldots, x_n)$ **симметричным** по своим переменным, если *любая* подстановка $\sigma \in S_n$ оставляет значение F неизменным: $F(x_1, x_2, \ldots, x_n) = F(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$. Например, числовое выражение $F = x^3 + y^3 + z^3 - 5(x + y + z)$ симметрично по x, y, z. Также выражение может быть симметрично по некоторым из своих переменных.

В настоящее время подстановки и композиции находят широкое применение в теории алгоритмов.

Так, российский математик А.А. Марков (1856—1922) использовал систему допустимых подстановок для формализации понятия «алгоритм», что служит универсальным методом переработки информации.

Алгоритмы А.А.Маркова являются не только средством некоторых теоретических построений, но и основой специализированного языка программирования, который применяется в качестве языка символьных преобразований при разработке систем искусственного интеллекта.

Наибольшее распространение для формализации понятия «алгоритм» получили так называемые **рекурсивные** функции, широко использующие элементы теории множеств, особенно композицию функций, декартова произведения, операции с подмножествами.

Рассмотрим несколько задач, решаемых с помощью теории множеств.

Задача 14. Пятьдесят лучших студентов колледжа наградили за успехи поездкой в Англию и Германию. Из них 5 не владели ни одним разговорным иностранным языком, 34 знали английский язык и 27 — немецкий. Сколько студентов владели двумя разговорными иностранными языками?

Решение. Введем обозначения множеств:

E — множество студентов, не владеющих ни одним иностранным языком;

A — множество всех студентов, $|A| = 50$;

B — множество студентов, владеющих английским языком, $|B| = 34$;

C — множество студентов, владеющих немецким языком, $|C| = 27$;

D — множество студентов, владеющих английским и немецким языками, $|D| = x$.

Представим множества графически с помощью кругов Эйлера (рис. 1.20).

Способ 1. Составим уравнение: $34x - x + 27 - x + x + 5 = 50$, откуда $x = 16$.

Способ 2. Найдем $|D|$ из уравнения $|B| + |C| - |D| = |A| - 5$ или $34 + 27 - x = 50 - 5$, отсюда $x = 16$.

Следовательно, 16 студентов свободно общались на двух иностранных языках.

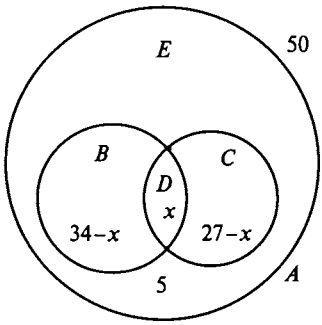


Рис. 1.20. Схематическое изображение условия задачи 14

Задача 15. Каждый студент группы программистов занимается в свободное время либо в НСО, либо спортом. Сколько студентов в группе, если 23 увлекаются спортом, 12 занимаются в НСО, а 7 совмещают занятия в НСО и увлечение спортом?

Решение. Изобразим множества кругами Эйлера, обозначив множества «спортсменов» — A и «исследователей» — B (рис. 1.21). Тогда $n(A \cap B) = 7$, $n(A) = 23$, $n(B) = 12$; $n(A \cup B) = n(A) + n(B) - n(A \cap B) = 23 + 12 - 7 = 28$.

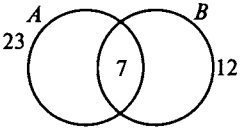


Рис. 1.21. Схематическое изображение условия задачи 15

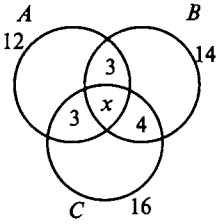


Рис. 1.22. Схематическое изображение условия задачи 16

Итак, в группе 28 студентов.

Задача 16. Из 35 студентов, побывавших на каникулах в Москве, все, кроме двоих, делились впечатлениями. О посещении Большого театра с восторгом вспоминали 12 человек, Кремля — 14, а 16 — о концерте, по три студента запомнили посещение театра и Кремля, а также театра и концерта, а четверо — концерта и пребывания в Кремле. Сколько студентов сохранили воспоминания одновременно о театре, концерте и Кремле?

Решение. Введем обозначения:

A — множество студентов, вспоминающих о театре, $n(A) = 12$;

B — о Кремле, $n(B) = 14$;

C — о концерте, $n(C) = 16$;

D — множество всех студентов, побывавших в поездке.

Изобразим множества графически с помощью кругов Эйлера (рис. 1.22).

$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$.

Обозначим $n(A \cap B \cap C) = x$, тогда $35 - 2 = 12 + 14 + 16 - 3 - 3 - 4 + x$, отсюда $x = 3$.

Всего 3 студента рассказывали о трех культурных мероприятиях поездки.

Упражнения

1.1. Укажите множество действительных чисел, соответствующее записи:

а) $A = \{x \mid 3x - 2 > 0\}$;

б) $B = \{x \mid x^2 + x + 1 > 0\}$;

в) $X = \{x \mid -3 \le x < 9,\ x \in \mathbb{Z}\}$;

г) $M = \{x \mid 5 \le x \le 6,\ x \in \mathbb{N}\}$;

д) $C = \{x \mid x^2 - 5x + 6 = 0\}$;

е) $Y = \{x \mid x^2 - 3x - 4 \le 0\}$.

1.2. Опишите множество M точек плоскости, заданных характеристическим свойством:

а) $X = \{M \,|\, |AM| < 4\}$;
в) $B = \{M \,|\, |MK| = |MQ|\}$;
б) $A = \{M \,|\, |MO| \geq 5\}$;
г) $Y = \{M \,|\, |AM| = |BM| = |CM|\}$;
д) $C = \{M \,|\, |MK| + |ML| \leq 6\}$;
е) $Z = \{M \,|\, |MQ| = |MP| = 3\}$.

1.3. Дано множество M_i:

а) $M_1 = \{n^2 + 1 \,|\, n \in \mathbb{N}\}$;
в) $M_3 = \left\{\frac{1}{n} \,\middle|\, n \in \mathbb{N}\right\}$;
б) $M_2 = \{n^3 - 2 \,|\, n \in \mathbb{N}\}$;
г) $M_4 = \left\{\frac{1}{n^2} \,\middle|\, n \in \mathbb{N}\right\}$;
д) $M_5 = \left\{\frac{1}{n-1} \,\middle|\, n \in \mathbb{N}\right\}$;
е) $M_6 = \left\{\frac{1}{2+n^2} \,\middle|\, n \in \mathbb{N}\right\}$.

1. Приведите по три примера элементов множества M_i.
2. Укажите, каким из множеств принадлежат числа 3, 4, 5, 13, 25, $\frac{1}{9}$, $\frac{1}{6}$, $\frac{1}{4}$.
3. Укажите, каким из множеств не принадлежат указанные числа. Запишите эти утверждения символически.

1.4. Приведите по три примера конечных и бесконечных множеств.

1.5. Задайте характеристическим свойством множество:
а) всех параллелограммов;
б) всех прямоугольников;
в) всех квадратов;
г) всех равнобедренных треугольников;
д) всех ромбов;
е) всех прямоугольных треугольников.

1.6. Составьте различные новые слова из букв слова:

а) апельсин;
в) стационар;
б) норматив;
г) ромашка;
д) множества;
е) ратификация.

1. Представьте буквы новых слов в виде множеств и найдите мощность множеств, состоящих из кортежей длины n, где $n \geq 5$.
2. Какое слово, полученное из букв данного, имеет наибольшую длину?

3. Найдите новое слово, состоящее из всех тех букв, которые входят в данное слово.

1.7. Какие из следующих соотношений справедливы:

а) $A \cup \varnothing = A$; в) $A \cap \varnothing = \varnothing$; д) $A \cup \overline{A} = A$;
б) $A \cup \varnothing = \varnothing$; г) $A \cap \varnothing = A$; е) $A \setminus A = \varnothing$?

1.8. Дано множество $A = \{a, b, c, \{a, b\}, \{a\}, \{a, b, c, d\}, \{a, b, c\}\}$.

1. Какие из элементов этого множества являются множествами?
2. Какие из следующих записей верны:

а) $a \in A$; в) $a \subset A$; д) $\{a, b, c, d\} \subset A$;
б) $\{a\} \in A$; г) $\{a\} \subset A$; е) $\{a, b, c, d\} \in A$?

1.9. На множестве U всех букв русского алфавита заданы множества A, B, C: $A = \{$ё, к, л, м, н$\}$; $B = \{$к, о, з, ё, л$\}$; $C = \{$б, ы, ч, о, к$\}$.

Найдите следующие множества и изобразите их кругами Эйлера:

а) $A \cap B$; в) $(A \cap B) \cup C$; д) $D = U \setminus (A \cup B \cup C)$;
б) $A \cup B$; г) $(A \cup C) \cap B$; е) $D = U \setminus (A \cap B \cap C)$.

1.10. Докажите, используя определения и круги Эйлера:

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

1.11. Даны отрезки $A = [-4; 5]$, $B = (2; 6]$, $C = (5; 10]$. Найдите следующие множества и изобразите их кругами Эйлера:

а) $(A \cup B) \cup C$; в) $A \cap B$; д) $(C \cup B) \setminus (A \cap B)$;
б) $(A \cap B) \cup C$; г) $(A \cup B) \setminus (A \cap B)$; е) $(A \cup C) \setminus (A \cap B)$.

1.12. Результаты статистических исследований реакции на очередную кадровую перестановку занесены в табл. 1.8.

Введем обозначения:

M — множество опрошенных лиц мужского пола, C — сомневающиеся, $П$ — множество преподавателей, O — множество тех, кто одобряет.

Таблица 1.8

Результаты статистических исследований

Социологические группы	Одобряют безоговорочно	Одобряют с некоторыми сомнениями	Сомневаются	Негативная реакция
Мужчины-преподаватели	3	4	2	10
Женщины-преподаватели	8	9	7	11
Юноши-студенты	5	4	4	7
Девушки-студентки	6	6	8	9

Изобразите следующие множества кругами Эйлера и найдите число элементов в них:

а) M; в) O; д) Π'; ж) $M \setminus C$; и) $(M \cap \Pi)'$;
б) Π; г) C; е) O'; з) $(M \cap \Pi') \setminus O$; к) $M' \cup C'$;
л) $M' \setminus (\Pi \cap O \cap C')$;
м) $(M \cap C) \setminus (\Pi \cap C)$.

1.13. Решите задачу Льюиса Кэрролла, автора книг «Алиса в стране чудес» и «Алиса в Зазеркалье»: «В ожесточенном бою из 100 пиратов потеряли по одному глазу — 70, по одному уху — 75, по одной руке — 80, по одной ноге — 85 пиратов. Каково минимальное число пиратов, потерявших одновременно глаз, ухо, ногу и руку?»

1.14. В результате социологического опроса студентов факультета программирования о занятиях в свободное от уроков время выяснилось, что из 100 человек:

18 — любят только читать книги;
24 — читают книги, но не ходят в театр;
7 — читают книги и посещают театр;
28 — читают книги;
47 — ходят на дискотеки;
9 — посещают театр и дискотеки;
13 — лежат на диване перед телевизором, занимаются только просмотром всех возможных каналов телевидения.

1. Сколько студентов любят ходить в театр?
2. Сколько студентов читают книги, посещают театр, но не дискотеки?
3. Сколько студентов посещают либо дискотеки, либо театр?
4. Сколько студентов, посещая дискотеки и театр, не любят читать книги?
5. Сколько студентов предпочитают только дискотеки?
6. Сколько студентов посещают либо дискотеки, либо театр, либо читают книги?

1.15. Определите вид множества по его мощности:

а) $\{x \mid x \in \mathbb{R}$ и $\sqrt{x-5},\ x < 5\}$;
б) $\{x \mid x \in \mathbb{R}$ и $x^2 < 0\}$;
в) множество четырехугольников с пятью вершинами;
г) $\{x \mid x \in \mathbb{R}$ и $\frac{5}{\sqrt{4-x}},\ x > 4\}$;
д) множество городов нашей страны, начинающихся с буквы Ъ.

1.16. Определите мощность множества:

а) $\{\{a\}, a\}$; б) $\{0\}$;
в) состоящего из букв слова «математика»;
г) состоящего из букв слова «перпендикулярные»;
д) состоящего из цифр числа 635252;
е) состоящего из цифр числа 1010111.

1.17. Найдите среди элементов множества B элементы, соответствующие элементам множества A, и укажите правило, по которому установлено это соответствие:

A = {спаниель, ромашка, минотавр, тарификация, карета, жеманство, соратница, коршун, весна};

B = {множества, ракета, стационар, шнурок, мошкара, норматив, ратификация, апельсин, навес}.

1.18. Выполните действия и определите мощность полученного множества:

а) $A = \{5, 7, 9\} \cup \{12, 15\}$, $B = \{5, 7, 9\} \cap \{12, 15\}$;

б) $A = \{5, 7, 9\} \cap \{5, 57, 59\}$, $B = \{5, 7, 9\} \cup \{5, 57, 59\}$;

в) $A = \{x | x$ — звонкий согласный звук$\}$, $B = \{x | x$ — глухой согласный звук$\}$, $A \cup B = ?$; $A \cap B = ?$;

г) $\{1, 2, 3\} \setminus \{2, 3\}$; д) $\{1, 2, 3\} \setminus \{4, 5\}$; е) $\{x^2 + y^2 \le 1\} \setminus \{x^2 + y^2 = 1\}$.

1.19. Даны множества $A = \{1, 2, 3\}$, $B = \{x, y, z\}$, $C = \{\ ; \otimes\}$.

1. Запишите декартовы произведения множеств:

а) $A \times B$; б) $B \times A$; в) $B \times C$; г) $C \times B$; д) $A \times C$; е) $C \times A$.

2. Верно ли, что для декартова произведения справедлив переместительный закон? Докажите вывод.

1.20. Постройте множество A^2, если:

а) $A = \{0, 1\}$; в) $A = \{0, 2, 4, 6, 8\}$; д) $A =$ {день, ночь};

б) $A = \{x, y, z\}$; г) $A = \{1, 3, 5, 7\}$; е) $A = \{a, b, c, d\}$.

1.21. Даны множества $X = \{-2, -1, 4, 5\}$ и $Y = \{-2, 0, 6\}$. Известно, что $x \in X$, $y \in Y$.

Составьте отношения и постройте графики указанных отношений:

а) $x + y > 0$; б) $x - y < 0$; в) $x + y \le 0$; г) $x - y \ge 0$; д) $xy > 0$; е) $xy < 0$.

1.22. Постройте графики отношений, заданных множеством пар:

а) $A = \{(x, y) | -3 \le x \le 4, y = 5\}$; б) $D = \{(x, y) | 2 \le x \le 8, 1 \le x \le 6, x - y < 0\}$.

1.23. Объясните, будут ли выполнимы свойства отношений на заданном множестве людей и почему:

а) «быть знакомым»;

б) «быть отцом»;

в) «поздравлять с праздником»;

г) «встречаться»;

д) «быть другом»;

е) «быть ровесником».

1.24. Существует ли взаимно-однозначное соответствие между:

а) множеством букв и множеством звуков русского языка;

б) множеством чисел, записанных в арабской и двоичной системах счисления;

в) множеством чисел и множеством студентов вашей группы;

г) словом на русском языке и его значением;

д) автором и литературным произведением;

е) названием и музыкальным произведением.

1.25. Секретарь факультета программирования составил отчет, в котором сказано, что из 100 абитуриентов английский язык в

школе изучали 50 человек, немецкий — 23, а французский — 30. С английским и французским языками знакомы 8 абитуриентов, с французским и немецким — 10, а с английским и немецким — 20. Все три языка изучали 5 абитуриентов.

1. Почему секретарь получил выговор за этот отчет? Докажите, что в отчете имеется ошибка.

2. Как выяснилось, что при обработке данных произошла потеря информации. Исправьте допущенную ошибку.

1.26. Из 100 студентов факультета программирования 42 посещают спортивные секции, 30 — занятия НСО, а 28 — кружки художественной самодеятельности. На занятия НСО и спортом успевают ходить 5 студентов, спортом и художественной самодеятельностью занимаются 10, НСО и художественной самодеятельностью — 8, а сразу все три увлечения имеют три студента.

Сколько студентов:

а) не посещают ни одно из этих объединений по интересам;

б) занимаются только спортом;

в) занимаются либо в НСО, либо в художественной самодеятельности;

г) занимаются либо спортом, либо художественной самодеятельностью, но не в НСО;

д) занимаются или спортом, или художественной самодеятельностью, но не в НСО;

е) занимаются или в НСО, или художественной самодеятельностью, но не спортом?

1.28. Найдите σ_1^{-1}, σ_2^{-2}, $\sigma_1 \circ \sigma_2$, $\sigma_2 \circ \sigma_1$, σ_1^3, σ_2^4, σ_1^{55}, σ_2^{-99} и порядок каждой из подстановок:

а) $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}$;

б) $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 2 & 5 & 3 \end{pmatrix}$;

в) $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}$;

г) $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 4 & 3 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 5 & 4 \end{pmatrix}$;

д) $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{pmatrix}$;

е) $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 2 & 4 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 2 & 3 & 1 \end{pmatrix}$.

1.29. Решите задачи.

1. В группе из 26 человек выбирают актив: старосту, физорга, профорга и культорга. Сколькими способами могут избрать актив группы?

2. Сколько различных спортивных прогнозов могут дать болельщики перед началом первенства по футболу, если в высшей лиге участвуют 15 команд и разыгрываются три медали: золотая, серебряная, бронзовая?

3. Сколькими способами в бригаде из шести операторов можно распределить 3 путевки в профилакторий, на турбазу и в дом отдыха?

4. Для проведения итогов олимпиады по компьютерному моделированию избрали жюри в составе председателя, заместителя председателя и трех членов жюри. Сколькими способами можно выбрать жюри из 15 преподавателей кафедры информатики?

5. Сколькими способами можно устроить на работу 8 выпускников факультета программирования на различные должности в 5 вычислительных центрах?

6. Сколькими способами можно устроить на летнюю практику 10 студентов на 3 предприятия города?

7. Придумайте и решите аналогичную задачу.

1.30. Решите задачи.

1. Из всех студентов вашей группы на беседу с деканом приглашены пятеро. Сколькими способами это можно сделать?

2. Сколько существует вариантов, чтобы из букв слова «студент» составить всевозможные кортежи длиной 5?

3. Сколько существует различных шестизначных телефонных номеров?

4. Сколькими способами можно доставить в колледж 12 новых компьютеров на 2 машинах, если на каждой машине можно разместить не более 5 ЭВМ?

5. Группу из 16 студентов должны разбить на подгруппы для работы в разных компьютерных классах. Сколько существует всех возможных вариантов формирования подгрупп, если в трех компьютерных классах соответственно 5, 4 и 7 работающих ЭВМ?

6. Из 15 красных и 7 белых гладиолусов формируют букеты. Сколькими способами можно составить букеты из 4 красных и 3 белых гладиолусов?

7. Придумайте и решите аналогичную задачу.

1.31. Решите задачи.

1. Сколько существует способов поставить на книжную полку в беспорядке собрание сочинений, состоящее из 7 томов?

2. Из цифр 3, 4, 5, 6 составлены четырехзначные числа. Сколько вариантов таких чисел можно найти, если среди найденных четверок нет чисел, заканчивающихся на 36?

3. Сколько всевозможных кортежей длиной 7 можно составить из слова «кислота»?

4. Сколькими способами можно построить кортежи из букв слова «грамматика»?

5. Сколькими способами можно поставить на полку четырехтомник Пушкина, двухтомник Ахматовой и трехтомник Лермонтова так, чтобы книги каждого автора стояли рядом?

6. На полке стоят 10 книг, 5 из них — Собрание сочинений Л. Н. Толстого. Сколько существует вариантов расстановки книг на полке при условии, что все 5 томов Л. Н. Толстого должны стоять рядом?

7. Придумайте и решите аналогичную задачу.

1.32. Решите задачи.

1. После окончания колледжа 12 выпускников-программистов решили ежегодно в день встречи с выпускниками посещать кафе и обмениваться впечатлениями. Так как в кафе столики рассчитаны на четырех человек, то друзья решили, что при каждой новой встрече за столиками будет четверка, не повторяющая прошлогодние. За сколько лет каждый из выпускников побеседует с каждым из друзей, сидя за каким-нибудь одним столом?

2. Сколько различных трехзначных номеров автомобилей можно составить в одной серии?

3. Сколько различных трехзначных чисел состоит только из четных цифр?

4. Три девушки и семь юношей распределились на практику в два вычислительных центра. Сколько способов разместить их поровну так, чтобы в каждый вычислительный центр попала хотя бы одна девушка?

5. В состав хоккейной команды входит три нападающих, два защитника и один вратарь. Сколько различных команд из них может составить тренер, если у него занимаются семь нападающих, пять защитников и три вратаря?

6. Когда Гулливер попал в Лилипутию, он обнаружил, что там все вещи ровно в 12 раз короче, чем на родине. Сколько лилипутских спичечных коробков поместится в спичечном коробке Гулливера?

1.33. Сколько нужно провести матчей в групповом этапе финала Кубка мира по футболу, если участвует 32 команды, разбитые на 8 групп?

1.34. Даны стандартные множества: $[a, b]$ — отрезок, S^1 — окружность, D^2 — диск (круг), $\mathbb{R}_+$ — положительные числа. Пользуясь методом декартовых координат, определите, какие фигуры представляют следующие множества:

а) $\mathbb{R}_+^2$; б) $\mathbb{R} \times \mathbb{R}_+$; в) $\mathbb{R} \times S^1$; г) $[a, b] \times [c, d]$; д) $S^1 \times S^1$; е) $\mathbb{R} \times [a, b]$; ж) $[a, b]^2$; з) $\mathbb{R} \times \{2, 3, 4\}$; и) $D^2 \times [c, d]$; к) $D^2 \times \{x, y\}$; л) $(a, b) \times$ $\times (c, d) \times (e, f)$.

Глава 2
ГРАФЫ

В этой главе описывается еще один универсальный и наглядный язык, графический, который применяется во многих областях науки и техники. Знакомство с элементами теории графов ограничится определениями и иллюстрациями к основным положениям этого раздела математики. Теория графов дает исключительно удобный аппарат для моделирования структурных свойств различных систем и отношений между объектами разной природы, в том числе программных моделей.

2.1. Основные понятия и определения графа и его элементов

> Так о великих вещах помогают составить понятье
> Малые вещи, пути намечая для их постижения.
>
> *Лукреций*

Впервые понятие «граф» ввел в 1936 г. венгерский математик Денни Кёниг. Но первая работа по теории графов принадлежала перу великого Леонарда Эйлера и была написана еще в 1736 г. С помощью графов изображаются схемы различных дорог, линии воздушных сообщений, газопроводов, теплотрасс, электросетей, а также микросхемы, дискретные многошаговые процессы, системы различных бинарных отношений, химические структурные формулы и другие диаграммы и схемы. Применяются графы для решения задач химии, экономики, электротехники и автоматики. Также они широко используются в информатике и строительстве. Без графов сложно анализировать классификации в различных науках.

Графом $G = (V, X)$ называется пара двух конечных множеств: множество точек и множество линий, соединяющих некоторые пары точек. В терминах декартова произведения (подразд. 1.5) множество

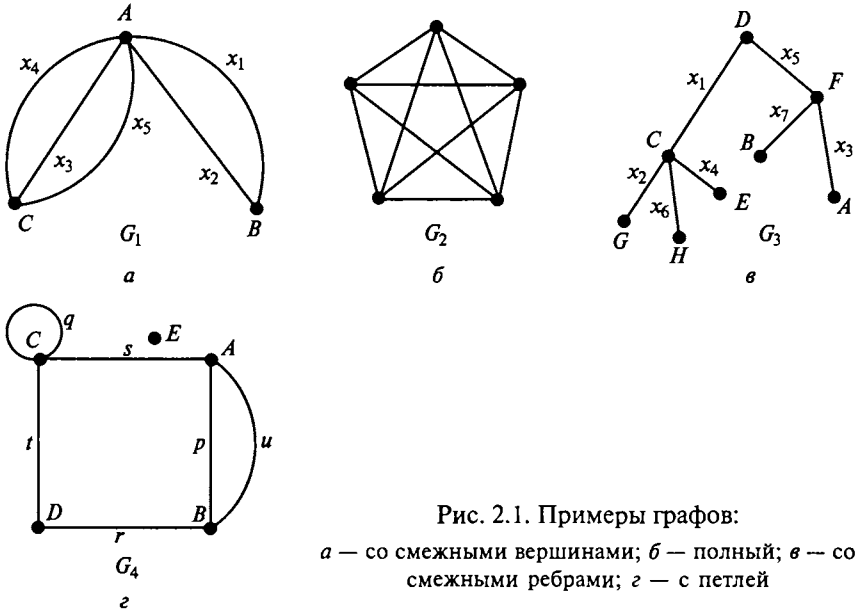


Рис. 2.1. Примеры графов:

а — со смежными вершинами; *б* — полный; *в* — со смежными ребрами; *г* — с петлей

линий X, соединяющее пару точек, — это некоторое подмножество множества $V \times V$: $X \subset (V \times V)$.

Точки называются **вершинами**, или **узлами**, графа, линии — **ребрами** графа. Примеры графов приведены на рис. 2.1.

Пусть дан граф $G = (V, X)$, где $V = \{V, W, ...\}$ — конечное непустое множество его вершин, а $X(V, W)$ — его ребра. Если ребро графа G соединяет две его вершины V и W (т.е. $\langle V, W \rangle \in X$), то говорят, что это ребро им **инцидентно**. Две вершины графа называются **смежными**, если существует инцидентное им ребро: на рис. 2.1, *а* смежными являются вершины A и B, A и C. Если граф G имеет ребро $X(V, V)$, у которого начало и конец совпадают, то это ребро называется **петлей.** На рис. 2.1, *г* петля — $q(C, C)$. Два ребра называются **смежными**, если они имеют общую вершину. На рис. 2.1, *в* смежными являются, например, ребра x_1 и x_2 с общей вершиной C.

Граф $G(V, X)$ может иметь ребра с одинаковыми парами вида $X(V, W)$. Такие ребра называются **кратными**, или **параллельными**. На рис. 2.1, *а* кратными являются, например, ребра $x_1(A, B)$, $x_2(A, B)$. Вершинам A и B инцидентны ребра x_1, x_2, x_3. Количество одинаковых пар вида $x(V, W)$ называется **кратностью** ребра (V, W). На рис. 2.1, *а* ребро AC имеет кратность, равную 3, а ребро AB — кратность, равную 2. Число ребер, инцидентных вершине A, называется **степенью** этой вершины и обозначается $\deg(A)$ (от англ. *degree* — степень). Если вершине инцидентна петля, она дает вклад в степень, равный двум, так как оба конца приходят в эту вершину.

На рис. 2.1, *в* вершина A имеет степень, равную 1, вершина C — 4, вершина D — 2. Записывается это в виде: $\deg(A)=1$, $\deg(C)=4$, $\deg(D)=2$. Граф G_4 (рис. 2.1, *г*) содержит четыре вершины $V=\{A, B, C, D\}$ и шесть ребер $X=\{p, q, r, s, t, u\}$.

Вершина графа, имеющая степень, равную нулю, называется **изолированной**. Граф, состоящий из изолированных вершин, называется **нуль-графом**. Для нуль-графа $X=\varnothing$. Вершина графа, имеющая степень, равную 1, называется **висячей**. На рис. 2.1, *г* вершина E — изолированная: $\deg(E)=0$, а вершины A, B, E, G, H на рис. 2.1, *в* — висячие.

Теорема 2.1. *В графе $G(V,X)$ сумма степеней всех его вершин — число четное, равное удвоенному числу ребер графа*:

$$\sum_{i=1}^{n}\deg(V_i)=2m,$$

где $n=|V|$ — число вершин; $m=|X|$ — число ребер графа.

Вершина называется $\dfrac{\textbf{четной}}{\textbf{нечетной}}$, если ее степень — $\dfrac{\text{четное}}{\text{нечетное}}$ число.

На рис. 2.1, *в* $\deg(D)=2$, $\deg(F)=3$, значит, у графа G_3 вершина D является четной, а F — нечетной. В теории графов доказана следующая теорема.

Теорема 2.2. *Число нечетных вершин любого графа — четно.*

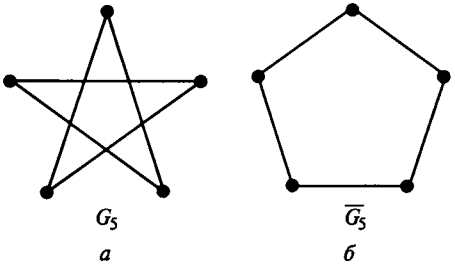
Следствие. Невозможно начертить граф с нечетным числом нечетных вершин.

Граф G называется **полным**, если любые две его различные вершины соединены одним и только одним ребром. Полным является граф G_2 на рис. 2.1, *б*. Таким образом, полный граф определяется только своими вершинами. Пусть число вершин полного графа n. Тогда степень любой вершины, очевидно, равна $\deg(V)=n-1$, а число ребер равно числу сочетаний из n по 2, т. е. $m=C_n^2$. Число ребер также можно найти по теореме 2.1:

$$m=\frac{1}{2}\sum_{i=1}^{n}\deg(V_i)=\frac{1}{2}n\deg(V)=\frac{n(n-1)}{2}=C_n^2.$$

Дополнением графа $G(V, X)$ называется граф $\overline{G}(V, X')$ с теми же вершинами V, что и граф G, и имеющий те и только те ребра X', которые необходимо добавить к графу G, чтобы он стал полным. Очевидно, что граф с кратными ребрами не имеет дополнения. Например, дополнением графа G_5 до графа G_2 на рис. 2.1, *б* является граф $\overline{G}_5$ (рис. 2.2).

Как отмечалось в подразд. 1.3, дополнением универсального множества является пустое, и наоборот. Поскольку граф и его дополнение отличаются только ребрами (множества X и X') и дополнение графов сводится к дополнению множества X, то частным



а *б*

Рис. 2.2. Дополнение $\overline{G}_5$ графа G_5 до графа G_2, изображенного на рис. 2.1, *б*

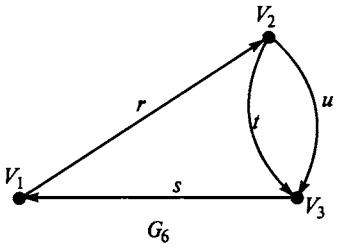


Рис. 2.3. Ориентированный граф

случаем этого свойства будет следующее правило: дополнением полного графа будет нуль-граф, и наоборот.

Если все пары (V_i, V_j) во множестве X являются упорядоченными, т.е. кортежами длины 2, то граф называется **ориентированным**, **орграфом**, или **направленным**. Поскольку сразу может быть не известно о каком графе идет речь, в этой главе мы будем употреблять круглые скобки для обозначения ребра вместо угловых, как это должно было быть для кортежей. В них будет помещаться соответствующая пара вершин. В таком случае ребра принято изображать стрелками (рис. 2.3). **Началом** ребра называется вершина, указанная в кортеже первой, **концом** — вторая вершина этой пары (графически она указана стрелкой). Ребра ориентированного графа имеют определенные фиксированные начало и конец и называются **дугами**. Очевидно, дуги (V_1, V_3) и (V_3, V_1), если они обе существуют, различны: $(V_1, V_3) \neq (V_3, V_1)$.

Степенью $\frac{\textbf{входа}}{\textbf{выхода}}$ вершины ориентированного графа называется число ребер, для которых эта вершина является $\frac{\text{концом}}{\text{началом}}$.

Степень входа вершины V будем обозначать $\deg_+(V)$, а степень выхода — $\deg_-(V)$. На рис. 2.3 $\deg_+(V_1) = 1$, $\deg_+(V_2) = 1$, $\deg_+(V_3) = 2$, $\deg_-(V_1) = 1$, $\deg_-(V_2) = 2$, $\deg_-(V_3) = 1$.

Дуги орграфа называются **кратными**, если они имеют одинаковые начальные и конечные вершины, т.е. одинаковые направления. Например, кратны дуги $u(V_2, V_3)$ и $t(V_2, V_3)$ на рис. 2.3.

Последовательность попарно инцидентных вершин $V_{i_1}, V_{i_2}, \ldots, V_{i_k}$ неориентированного графа, т.е. последовательность ребер неориентированного графа, в которой вторая вершина предыдущего ребра совпадает с первой вершиной следующего, называется **маршрутом**. Число ребер маршрута называется **длиной** маршрута. Например, на рис. 2.1, *в* *HCDFD* — маршрут длиной 4. Обозначение: $|HCDFD| = 4$. Очевидно, что если $V_{i_1}, V_{i_2}, \ldots, V_{i_k}$ — маршрут длины $k-1$, то и $V_{i_k}, V_{i_{k-1}}, \ldots, V_{i_2}, V_{i_1}$ также будет являться маршру-

том длины $k - 1$. Маршрут принято задавать как последовательность ребер, поскольку это более удобно при наличии кратных ребер. Если начальная вершина маршрута совпадает с конечной, то такой маршрут называется **замкнутым** или **циклом**. В графе G_4 (рис. 2.1, *г*) (t, s, p, r), (u, s, t, r) — циклы длиной 4, (r, t, q, s, u) — цикл длиной 5, (t, s, u, r, t, s, p, r) — 8-цикл, (p,u) — 2-цикл, петля (q) — 1-цикл.

Расстоянием между двумя вершинами называется минимальная длина из всех возможных маршрутов между этими вершинами при условии, что существует хотя бы один такой маршрут. Обозначается как $d(V_1V_2)$ (от лат. *distantio* — расстояние) $d(V_1V_2) = \min|V_1...V_2|$.

Поскольку рассматриваются конечные графы, минимум можно найти всегда. Очевидно, что $d(V_1V_2) = d(V_2V_1)$. Формально можно ввести расстояние $d(V'V') = 0$ между любой вершиной и ей же самой, что соответствует нулевому маршруту, у которого начало и конец в одной вершине.

В маршруте одно и то же ребро может встретиться несколько раз. Если ребро встретилось только один раз, то маршрут называется **цепью**. Например, в графе G_4 (рис. 2.1, *г*) (t, s, p) — 3-цепь. Если $(x_1, x_2, ..., x_{k-1}, x_k)$ — k-цикл, то любая циклическая перестановка, например $(x_2, ..., x_{k-1}, x_k, x_1)$, также будет k-циклом, поскольку сведется лишь к выбору начальной вершины. Частным случаем этого утверждения будет следующее: если k-цикл $(x_1, x_2, ..., x_{k-1}, x_k)$ является цепью, то для любой циклической подстановки $\sigma \in S_k$ последовательность $(x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(k)})$ также будет k-циклом и цепью.

В орграфе маршрут является ориентированным и называется **путем**. На путь сразу налагаются важные требования, являющиеся частью определения:

- направление каждой дуги должно совпадать с направлением пути;
- ни одно ребро пути не должно встречаться дважды.

Другими словами, **путь** — упорядоченная последовательность ребер ориентированного графа, в которой конец предыдущего ребра совпадает с началом следующего и все ребра единственны. На рис. 2.3 (u, s, r, t) — 4-путь, (r, u) — 2-путь, (s, r, t, s) путем не является. Тогда **цикл** в орграфе — путь, у которого совпадают начало и конец. На рис. 2.3 (s, r, t) и (u, s, r) — 3-циклы. Для циклов орграфа также справедлива теорема о циклических подстановках.

Цепь, путь и цикл в графе называются **простыми**, если они проходят через любую из вершин не более одного раза. Неориентированный граф называется **связным**, если между любыми двумя его вершинами есть **маршрут**. Для связного графа ориентация дуг не обязательна. Так, граф G_2 (рис. 2.1, *б*) является связным, а граф G_4 (рис. 2.1, *г*) — **несвязным**. Также можно ввести понятие

связности для вершин графа: две вершины называются **связными**, если существует маршрут между ними. Понятно, что связность между вершинами является бинарным отношением. Это отношение будет отношением *эквивалентности*. Действительно, отношение связности обладает известными свойствами (см. подразд. 1.6), т.е. оно:

- рефлективно — каждая вершина (включая изолированные) связна сама с собой;
- симметрично — любой маршрут (V', ..., V'') можно представить в обратном порядке: (V'', ..., V');
- транзитивно — если вершина V соединена с вершиной V' маршрутом $M_1(X_1, ..., X_p)$, а вершина V' соединена с вершиной V'' маршрутом $M_2(X_{p+1}, ..., X_n)$, то вершина V соединена с вершиной V'' маршрутом $M_3(X_1, ..., X_p, X_{p+1}, ..., X_n)$, в котором сначала идут все ребра маршрута M_1, а затем все ребра маршрута M_2.

Граф G можно разбить на непересекающиеся подмножества V_i по признаку связности. Вершины одного множества являются связными между собой, а вершины различных множеств — несвязны. Тогда все подграфы V_i (классы эквивалентности) графа G называют **связными компонентами**, или **компонентами связности**. Связный граф имеет одну компоненту связности.

Доказано, что в конечном связном графе всегда можно построить ориентированный цикл, проходящий через каждое ребро по одному разу в двух направлениях. Такой цикл называют *способом обхода* всего графа и используют при решении многих прикладных задач. В частности, разработаны специальные алгоритмы обхода ребер графа, которые можно использовать при решении задач вида «поиска выхода из лабиринта». В лабиринте дорожки служат ребрами графа, а их разветвления, начало движения (вход) и конец (выход) — вершинами графа. Если вход и выход принадлежат одной компоненте связности, то такой лабиринт принципиально проходим, если разным — то непроходим. Во втором случае не поможет даже мифологическая «нить Ариадны».

Теорема 2.3. *Для того чтобы связный граф G являлся простым циклом, необходимо и достаточно, чтобы каждая его вершина имела степень, равную 2.*

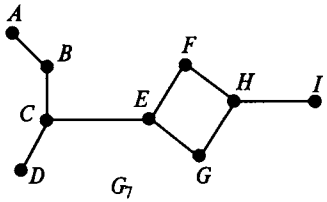


Рис. 2.4. Граф G_7 с мостами BC и CE

Ребро (V, W) связного графа G называется **мостом**, если после его удаления G станет несвязным и распадется на два связных графа G' и G''. На рис. 2.4 мост (CE) разделил связный граф G_7 на два различных связных графа: G' с вершинами (A,B,C,D) и G'' с вершинами (E, F, G, H, I). Также мостом является ребро BC.

Теорема 2.4. *Ребро графа является мостом тогда и только тогда, когда не принадлежит ни одному циклу.*

? Какие графы можно считать различными, а какие не различаются?

Графы G' и G'' называются **изоморфными**, если существует взаимно-однозначное соответствие между их ребрами и вершинами, причем соответствующие ребра соединяют соответствующие вершины. Поскольку в данном издании исследуются конечные множества, такая биекция должна быть подстановкой. Между названием вершины и ее номером различия нет. Смежность двух вершин есть бинарное отношение, поэтому изоморфизм графов можно рассматривать как изоморфизм множеств их вершин (см. подразд. 1.4 и 1.6), на котором введено отношение смежности. Итак, графы $G_1 = (V_1, X_1)$ и $G_2(V_2, X_2)$ называются **изоморфными**, если $|V_1| = |V_2| = n$ и существует подстановка $\sigma \in S_n$, такая, что $V_2 = \sigma(V_1)$, а $X_2 = \{(\sigma(V_i); \sigma(V_j)) \mid (V_i, V_j) \in X_1\}$. Иными словами, можно так переобозначить вершины первого графа, что в новых обозначениях вершины и ребра будут совпадать со вторым графом, причем кратным ребрам первого G'_8 должны соответствовать кратные ребра второго G''_8 такой же кратности (рис. 2.5).

Аналогично устанавливается изоморфизм между ориентированными графами. При этом следует помнить, что ребро является упорядоченным множеством, и надо быть особенно внимательным, соблюдая порядок.

Важно, что изоморфизм графов является отношением эквивалентности. Это удобнее всего заметить исходя из свойств подстановок. На практике такие различные по внешним признакам изоморфные графы не различают, рассматривая их с точностью до изоморфизма.

Граф G называется **планарным** (**плоским**), если существует изоморфный ему граф G', в изображении которого на плоскости ребра пересекаются только в вершинах. Иными словами, у планарного графа никакие два ребра не имеют общих точек, кроме общих вершин. На рис. 2.1 графы G_1 и G_3 являются планарными, а G_2 — нет. **Областью** назовем подмножество плоскости, пересекающееся с планарным графом только по некоторому простому

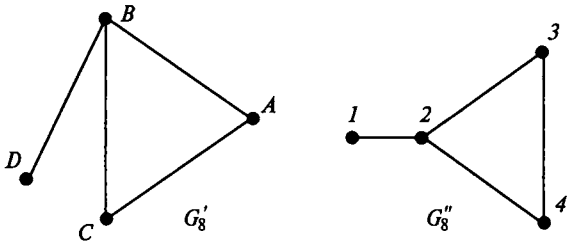


Рис. 2.5. Изоморфные графы

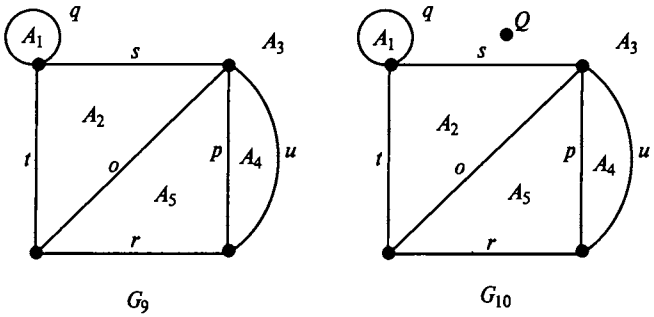


Рис. 2.6. Графы:

а — G_9 (с областями A_1—A_5); *б* — G_{10} (с областями A_1, A_2, A_4, A_5)

циклу графа, являющемуся границей области. Поскольку рассматриваются конечные графы, то плоский граф всегда является ограниченным подмножеством плоскости. Пусть M — планарный граф с внутренними областями. Тогда дополнение M, т.е. $M' = \mathbb{R}^2 \setminus M$ также может являться областью. Например, граф G_9 (рис. 2.6, *а*) выделяет в плоскости следующие области: A_1 с границей q; A_2 с границей (o, s, t); A_3 с границей (q, s, u, r, t); A_4 с границей (p, u); A_5 с границей (o, p, r). Множество A_3 на рис. 2.6, *б* областью не является, так как пересечение $A_3 \cap G_{10}$ содержит точку Q, не принадлежащую никакому циклу.

Теорема 2.5 (Эйлера). *Связный плоский граф с n вершинами и m ребрами разбивает плоскость на r областей (включая внешнюю), причем*

$$n - m + r = 2.$$

Задача 17 «О трех колодцах». Проложить дорожки от трех домов к каждому из трех колодцев так, чтобы никакие две дорожки не пересекались (рис. 2.7).

Граф называется **двудольным**, если его вершины разбиты на два непересекающихся подмножества: $V = V_1 \cup V_2$, а ребра связывают вершины только из разных классов (не обязательно все пары). Если каждая вершина множества V_1 связана ребром с каждой вершиной множества V_2, то двудольный граф называется **полным двудольным** и обозначается $K_{m,n}$, где $m = |V_1|$, $n = |V_2|$. Примером двудольного полного графа $K_{3,3}$ является граф к задаче 17, которую также называют «Три дома — три колодца», показывая этим два непересекающихся множества вершин графа.

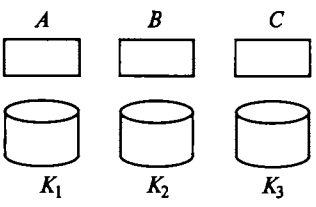


Рис. 2.7. Иллюстрация к задаче «О трех колодцах»

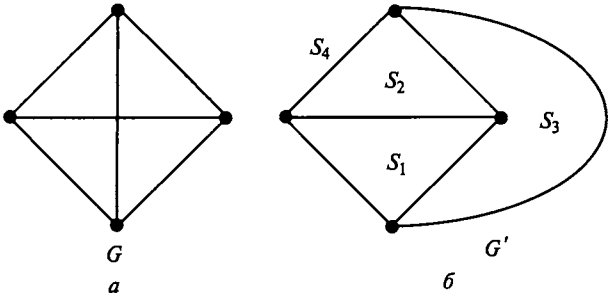


Рис. 2.8. Планарные графы:
а — первоначальный; *б* — изображенный иначе

На (рис. 2.8, *а*) граф G является планарным, так как его можно изобразить иначе (G' на рис. 2.8, *б*). При таком изображении плоскость разбивается на области S_1, S_2, S_3, S_4, которые могут быть раскрашены в разные цвета. Видно, что $n = 4$, $m = 6$, $r = 4$, и справедлива формула Эйлера $n - m + r = 4 - 6 + 4 = 2$.

Эйлеровым $\frac{\textbf{путем}}{\textbf{циклом}}$ графа называется $\frac{\text{путь}}{\text{цикл}}$, который содержит все ребра графа только один раз. Граф, обладающий эйлеровым циклом, называется **эйлеровым**. Плоские эйлеровы графы можно изобразить «одним росчерком пера», причем процесс изображения начинается и заканчивается в одной вершине. Такой граф называют также **уникурсальным**.

Теорема 2.6. *Граф G является эйлеровым тогда и только тогда, когда G — связный граф, имеющий все четные вершины.*

На рис. 2.9 изображен пример эйлерова графа. Граф G_9 на рис. 2.6, *а* не будет эйлеровым, так как не все его вершины являются четными.

Гамильтоновым $\frac{\textbf{путем}}{\textbf{циклом}}$ графа G называется $\frac{\text{путь}}{\text{цикл}}$, проходящий через каждую его вершину только один раз. Граф, содержащий гамильтонов цикл, называется **гамильтоновым**.

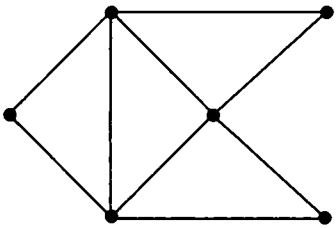


Рис. 2.9. Изображение эйлерова графа

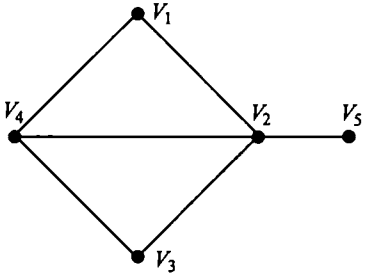
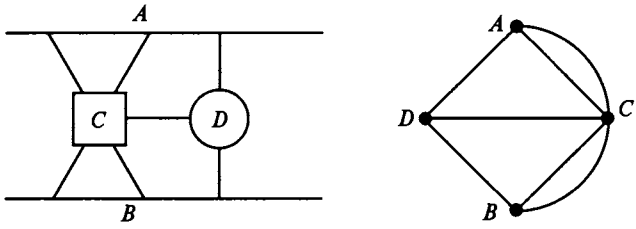


Рис. 2.10. Изображение гамильтоновых путей



а *б*

Рис. 2.11. Иллюстрации к задаче «О кёнигсбергских мостах»:
а — условие; *б* — граф

Например, в графе G путь (V_3, V_4, V_1, V_2, V_5) является гамильтоновым, а путь (V_2, V_3, V_4, V_1, V_2, V_5) не является гамильтоновым (рис. 2.10).

Задача 18 «О кёнигсбергских мостах» (Эйлера). Необходимо обойти все 7 мостов так, чтобы на каждом побывать только один раз и вернуться к началу пути (рис. 2.11, *а*).

Граф к задаче о мостах изображен на рис. 2.11, *б*.

Задача Эйлера «О кёнигсбергских мостах». Историческим поводом для создания математической науки, получившей впоследствии название «Теория графов», явилось решение в 1736 г. Л. Эйлером задачи о кёнигсбергских мостах. В XVIII в. город Кёнигсберг располагался на двух берегах реки Преголи, имеющей два острова, соединенных с берегами и между собой семью мостами. Жители города на практике решали задачу: можно ли пройти по всем семи мостам так, чтобы на каждом из них побывать по одному разу и вернуться к началу пути.

Задача Кэли «О четырех красках». Принята такая формулировка этой задачи: на любой политико-административной карте раскрасить страны четырьмя красками так, чтобы никакие две страны, имеющие общую границу, не были раскрашены одинаковым цветом, причем общей границей не может служить одна точка.

Сформулируем задачу на языке теории графов: в произвольном неориентированном плоском графе G четырьмя красками раскрасить вершины так, чтобы никакие две смежные вершины не были раскрашены одним цветом. Оказывается, условие задачи выполнимо для пяти красок. Даже доказано, что такая раскраска возможна для всех плоских графов с числом вершин, не превосходящим 38. Но проблема четырех красок остается нерешенной с 1878 г., когда английский математик Кэли привел ее формулировку на заседании Английского королевского научного общества.

Задача «О трех домах и трех колодцах». Условие этой задачи нам известно: «Проложить дорожки от трех домов к каждому из трех колодцев так, чтобы никакие две дорожки не пересекались». Эта задача была решена Куратовским в 1930 г.

2.2. Операции над графами

Кто из богов придумал этот знак,
Какое исцеленье от унынья
Дает мне сочетанье этих линий!
Расходится томивший душу мрак.

И. В. Гёте

Объединением графов $G_1 = (V_1, X_1)$ и $G_2 = (V_2, X_2)$ называется граф $G = G_1 \cup G_2$, множество вершин которого $V = V_1 \cup V_2$, а множество ребер $X = X_1 \cup X_2$.

Пересечением графов G_1 и G_2 называется граф $G = G_1 \cap G_2$, для которого $X = X_1 \cap X_2$ — множество ребер, а $V = V_1 \cap V_2$ — множество вершин.

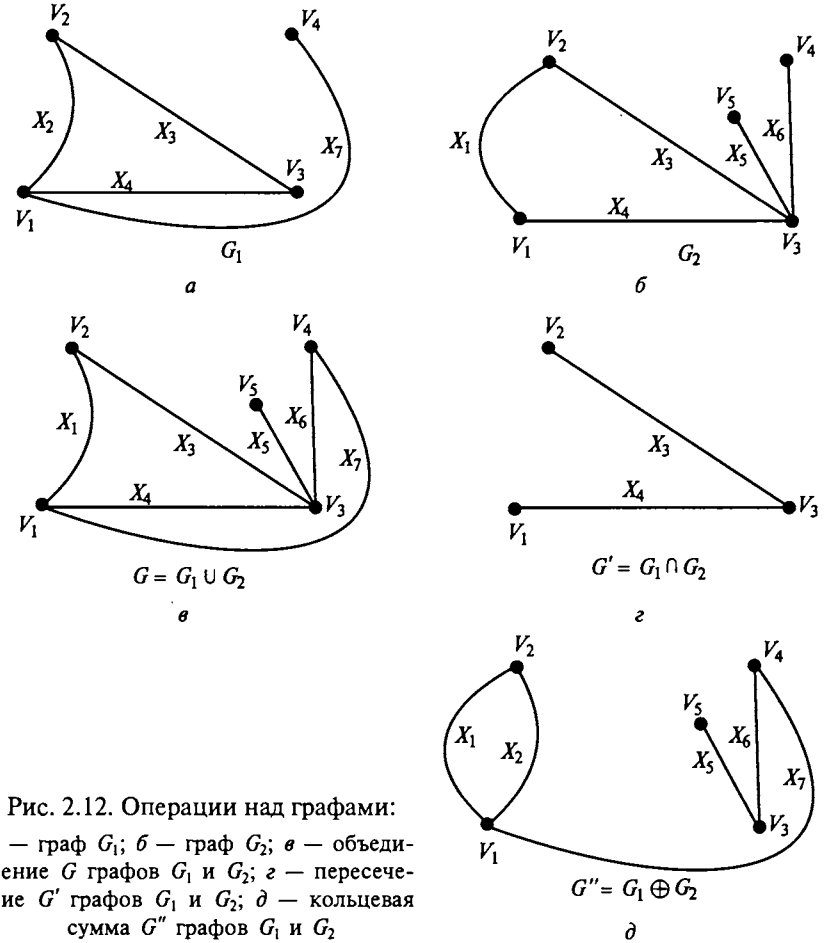


Рис. 2.12. Операции над графами:
а — граф G_1; *б* — граф G_2; *в* — объединение G графов G_1 и G_2; *г* — пересечение G' графов G_1 и G_2; *д* — кольцевая сумма G'' графов G_1 и G_2

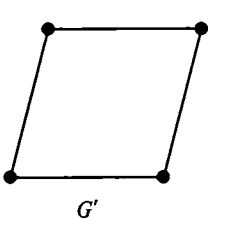


Рис. 2.13. Несвязный граф G с двумя компонентами связности G' и G''

Подграфом графа $G=(V, X)$ называется граф $G_1=(V_1, X_1)$, все вершины и ребра которого являются подмножествами множества вершин и ребер графа G. Обозначения такие же, как и для множеств: $G_1 \subset G$, если $V_1 \subset V$, $X_1 \subset X$. Отметим, что не любая пара (V_1, X_1), где $V_1 \subset V$, $X_1 \subset X$, будет являться подграфом. Для этого еще необходимо, чтобы она являлась графом, т.е. $X_1 \subset \{(V_i, V_j) \mid V_i, V_j \in V_1\}$. В противном случае у подграфа могло появиться ребро, не инцидентное ни одной его вершине.

Кольцевой суммой двух графов называется граф $G = G_1 \oplus G_2$, порожденный множеством вершин $V = V_1 \cup V_2$ и множеством ребер $(X_1 \cup X_2) \setminus (X_1 \cap X_2)$, т.е. множеством ребер, содержащихся либо в G_1, либо в G_2, но не в $G_1 \cap G_2$.

Пусть графы G_1 и G_2 имеют вид, представленный на рис. 2.12, *а*, *б*. Тогда их объединением, пересечением и кольцевой суммой будут соответственно графы G, G' и G'' (рис. 2.12, *в*, *г*, *д*).

Напомним, что **компонентой связности** неориентированного графа $G(V, X)$ называется его подграф G' (V', X) с множеством вершин $V' \subset V$ и множеством ребер X, инцидентных только вершинам из множества V', причем ни одна вершина из $V_i \in V'$ не смежна с вершинами из множества $V' \cap V$. Тогда несвязный граф состоит из нескольких отдельных подграфов — компонент связности, т.е. связных графов. Например, несвязный граф G (рис. 2.13) состоит из двух компонент связности, т.е. из двух подграфов G' и G''.

2.3. Деревья. Лес. Бинарные деревья

С вершины дорога вперед — только вниз.

П. Таранов

Деревом называют конечный связный граф с выделенной вершиной (**корнем**), не имеющий циклов (рис. 2.14).

Для каждой пары вершин дерева — **узлов** — существует единственный маршрут, поэтому вершины удобно классифицировать по степени удаленности от корневой вершины. Расстояние до корневой вершины V_0 называется **ярусом** s вершины, $s = d(V_0 V)$.

Поскольку маршрут между двумя вершинами единственный, то, применяя это свойство к смежным вершинам, можно заключить, что любая ветвь является мостом. Действительно, при удалении ребра этот единственный маршрут прерывается. Тогда граф

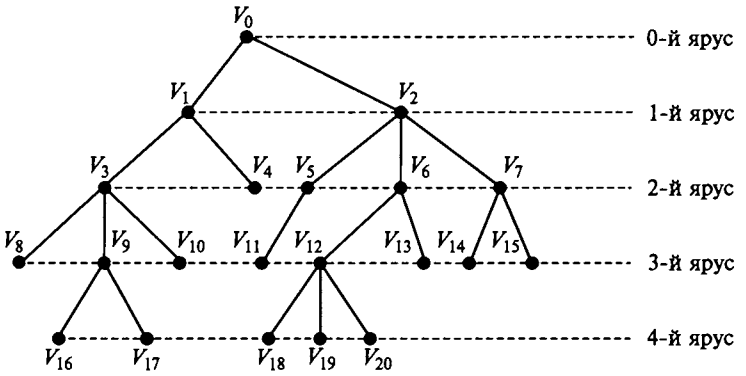


Рис. 2.14. Иллюстрация графа-дерева

распадается на два подграфа. В одном из них остается корневая вершина, и этот граф G_1 тоже будет являться деревом. В другом графе $\overline{G}_1$ выделим вершину, инцидентную удаленному мосту. Тогда второй подграф также будет являться деревом. Если в исходном графе вершина F принадлежала s-му ярусу, а дерево «обрубили» по ребру, соединявшему вершины t-го и $(t-1)$-го ярусов, причем $s \geq t$, то тогда

$$s'(F) = \begin{cases} s, & F \in G_1, \\ s - t, & F \in \overline{G}_1. \end{cases}$$

В частности, если $s = t$ и $F \in \overline{G}_1$, то вершина F будет корневой для $\overline{G}_1$ и $s'(F) = s - t = 0$. Если $s < t$, то вершина заведомо принадлежит подграфу G_1.

Наиболее характерные свойства деревьев, которые одновременно служат эквивалентными определениями дерева, сформулируем в следующей теореме.

Теорема 2.7. *Граф $G(V, X)$ ($|V| = n > 1$) является деревом тогда и только тогда, когда выполняется хотя бы одно из условий:*

граф $G(V, X)$ связен и не содержит циклов;

граф $G(V, X)$ не содержит циклов и имеет $n - 1$ ребро;

граф $G(V, X)$ связен и имеет $n - 1$ ребро;

граф $G(V, X)$ не содержит циклов, но добавление ребра между несмежными вершинами приводит к появлению одного и только одного элементарного цикла;

граф $G(V, X)$ связный, но утрачивает это свойство после удаления любого ребра;

в графе $G(V, X)$ всякая пара вершин соединена цепью, и только одной.

Итак, дерево с n вершинами имеет $n - 1$ ребро, поэтому оно будет минимальным связным графом. Висячие вершины, за ис-

ключением корневой, называются *листьями*. На рис. 2.14 листьями являются, например, вершины V_4, V_{13} и V_{20}. При $n = 2$ дерево состоит из корня и листа и имеет вид отрезка.

Пусть G_1, G_2, ..., G_k — непересекающиеся деревья, т.е. $\forall i, j \in \in (1, ..., k)$ $G_i \cap G_j = \varnothing$. Тогда **упорядоченное** объединение деревьев $G = \bigcup_{i=1}^{k} G_i$ представляет собой несвязный граф, называемый **лесом**. Компонентами связности леса являются деревья. **Остовом** связного графа G называется любой его подграф, содержащий все вершины графа G и являющийся деревом (говорят: «покрывающим его деревом»).

Кодеревом T' остова T графа G называется дополнение T до G, т.е. такой его подграф, который содержит все его вершины и только те ребра, которые не входят в T. Тогда $G = T \cup T' = T \oplus T'$. Иначе говоря, кодеревом остова $T(V, X_1)$ графа $G(V, X)$ будет остов T' $(V, X \setminus X_1)$. Очевидна двойственность: $(T')' = T$.

Дерево может быть представлено расслоенным на ярусы (уровни), при этом ветвям, попавшим в один ярус, соответствует одинаковая длина пути исходного графа. Число путей в каждом дереве соответствует числу висячих вершин (листьев). Например, в графе на рис. 2.14 двадцать листьев и двадцать путей от V_0.

При описании деревьев принято использовать термины: *отец*, *сын*, *предок*, *потомок*.

Каждая вершина дерева называется **узлом**, причем каждый узел является корнем дерева, имеющего n поддеревьев ($n \in [0, n)$). Тогда узел без поддеревьев называется листом и является висячей вершиной. Узел k-го яруса называется **отцом** узла $(k+1)$-го яруса, если они смежны. Узел $(k+1)$-го яруса называется **сыном** узла k-го яруса. Два узла, имеющие одного отца, называются **братьями** (рис. 2.15).

Упорядоченным деревом называется дерево, в котором поддеревья каждого узла образуют упорядоченное подмножество. Для упорядоченных деревьев принята терминология: *старший* и *младший* сын для обозначения соответственно первого и последнего сыновей некоторого узла.

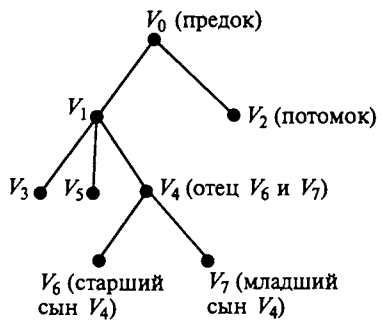


Рис. 2.15. Иллюстрация родственных связей — упорядоченное дерево

В информатике принято использовать подмножество множества деревьев, когда каждый узел либо является листом, либо образует два поддерева: левое и правое. Такой вид деревьев называется **бинарными деревьями** и используется при делении множества на два взаимо-

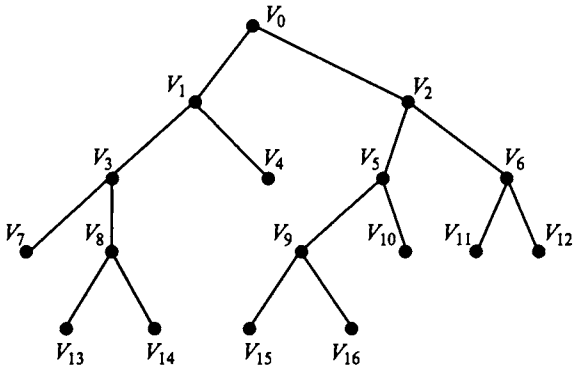


Рис. 2.16. Бинарное дерево

исключающих подмножества по какому-то признаку (так называемое дихотомическое деление). Для отца A — сыновья B и C, причем B — левый, а C — правый потомки. **Строго** бинарным деревом называется такой граф (рис. 2.16), у которого каждый узел, не являющийся листом, содержит два и только два поддерева — левое и правое.

Бинарное дерево уровня n называется **полным,** если каждый его узел уровня n является листом, а каждый узел уровня меньше, чем n, имеет непустое левое и правое поддеревья. Примером полного бинарного дерева служит таблица розыгрыша соревнования по олимпийской системе («плей-офф»). На рис. 2.17 приведена таблица розыгрыша Кубка мира по футболу 2002 г., начиная со стадии четвертьфиналов (указан корень V_0 — Бразилия).

Бинарные деревья применяются в информатике для поиска одного из двух возможных вариантов ответа. Например, при поиске данных, когда необходимо сравнить каждый элемент списка с образцом, и если значения совпадают, то процесс поиска завершен,



Рис. 2.17. Бинарное дерево для представления розыгрыша Кубка мира по футболу 2002 г.

а если не совпадают, то поиск данных продолжается. Впервые понятие двоичного дерева ввел в III в. римский философ Порфирий.

Цикломатическое число графа. Пусть задан неориентированный граф G. **Цикломатическим числом** графа называется число $\nu(G) = m(G) + c(G) - n(G)$, где $m(G)$ — число его ребер; $c(G)$ — число связных компонент графа; $n(G)$ — число вершин. Цикломатическое число дерева равно нулю. Цикломатическое число леса равно сумме цикломатических чисел составных связных компонент — деревьев и, следовательно, тоже равно нулю. Для остальных графов цикломатические числа — положительные.

Например, для полного графа K_5 (имеющего пять вершин и $C_5^2 = 10$ ребер) цикломатическое число равно $\nu = 10 + 1 - 5 = 6$.

2.4. Способы задания графа. Изоморфные графы

> В математике значительно сильнее, чем в других дисциплинах, обнаруживается черта организованности, стремление находить скрытый порядок во всем, что нас окружает...
>
> *А. Сухотин*

Существуют различные способы задания графа: геометрический (рисунки, схемы, диаграммы), простое перечисление вершин и ребер, табличный. Человеку удобно работать с графом-рисунком, так как он может легко установить связь между вершинами в наглядном виде с помощью ребер, изображаемых непрерывными линиями. Такое геометрическое представление плоского графа называется его **реализацией**. Для машинной обработки удобнее задать граф в алгебраической форме — перечислением (списком) вершин или ребер.

Например, орграф на рис. 2.3 можно задать с помощью пар (V_1, V_2), (V_2, V_3), (V_2, V_3) и (V_3, V_1), что соответствует дугам (r, u, t, s). При переходе от алгебраического способа к геометрическому одному и тому же графу могут соответствовать различные изображения — **изоморфные графы**, при этом от правильного изображения зависит, например, свойство плоской реализуемости. Для этого нужно правильно задать сам граф.

Основным способом задания графа является перечисление всех его вершин и ребер. Но такое представление, во-первых, несимметрично (с ним трудно работать, особенно ЭВМ), во-вторых, для указания каждого ребра нужно еще раз выписывать соответствующие вершины, что плохо с точки зрения сжатия и хранения информации. Иногда граф задается **таблицей**, состоящей из n строк (вершины) и m столбцов (ребра). Главным во всех способах зада-

ния графа (диаграммой, матрицей, таблицей) является указание соответствия между множествами n вершин V_i и m ребер X_i.

Одним из самых распространенных способов задания графа является **матричный способ**. Пусть дан граф $G(V, X)$, где $V = \{V_1, V_2, ..., V_n\}$ — вершины, а $X = \{X_1, X_2, ..., X_m\}$ — ребра графа.

Назовем **матрицей инцидентности** таблицу B, состоящую из n строк (вершины) и m столбцов (ребра), в которой:

- для неориентированного графа:

$b_{ij} = 1$, если вершина V_i инцидентна ребру X_j;

$b_{ij} = 0$, если вершина V_i не инцидентна ребру X_j;

- для ориентированного графа:

$b_{ij} = 1$, если вершина V_i является началом дуги X_j;

$b_{ij} = 0$, если вершина V_i не инцидентна дуге X_j;

$b_{ij} = -1$, если вершина V_i является концом дуги X_j.

Очевидно, что в каждом столбце матрицы инцидентности должно быть только два ненулевых числа, так как ребро инцидентно двум вершинам. Число ненулевых элементов каждой строки — степень соответствующей вершины. Но в математике удобнее работать с квадратными матрицами, так как для них хорошо разработан соответствующий алгебраический аппарат.

Назовем **матрицей смежности** графа $G(V, X)$ без кратных ребер квадратную матрицу A порядка n, в которой:

$a_{ij} = 1$, если $(V_i, V_j) \in X$;

$a_{ij} = 0$, если $(V_i, V_j) \notin X$.

Поскольку для неориентированного графа ребра (V_i, V_j) и (V_j, V_i) одновременно принадлежат или не принадлежат графу, так как символизируют одно и то же ребро, то $a_{ij} = a_{ji}$. Матрица смежности неориентированного графа является симметрической и не меняется при транспонировании.

Хотя формально каждая вершина всегда смежна сама с собой, в матрице смежности мы будем ставить $a_{kk} = 0$, если у нее нет петли, и $a_{kk} = 1$, если есть одна петля. Итак, если граф имеет матрицу смежности и не имеет петель, на главной диагонали у него всегда стоят нули.

Например, орграф на рис. 2.3 можно задать такой таблицей инцидентности (табл. 2.1).

Таблица 2.1

Таблица инцидентности орграфа

V_i	X_j			
	s	t	r	u
V_1	−1	0	1	0
V_2	0	1	−1	1
V_3	1	−1	0	−1

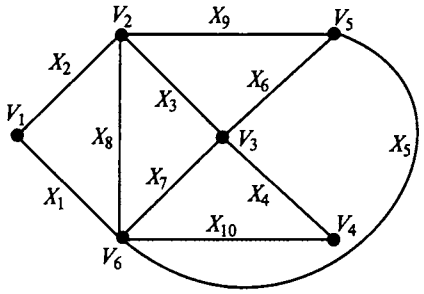


Рис. 2.18. Графическая интерпретация графа G, заданного табл. 2.2

Его же можно задать матрицей $B = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & -1 & 0 & -1 \end{pmatrix}$.

Поскольку ребра изначально не упорядочены, то, например, записывая сначала инцидентность ребра t (1-й столбец), а потом ребра s (2-й столбец), получим матрицу с переставленными столбцами 1 и 2. Тогда при решении обратной задачи — восстановлении графа по его матрице инцидентности — можно получить граф лишь *с точностью до изоморфизма*. Поэтому в графах важно лишь отношение между вершинами (т. е. смежность), а их название и порядок не столь важны.

Граф, изображенный на рис. 2.18, задается таблицей инцидентности (табл. 2.2).

Таблица 2.2

Таблица инцидентности графа

	X_j									
V_i	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
V_1	1	1	0	0	0	0	0	0	0	0
V_2	0	1	1	0	0	0	0	1	1	0
V_3	0	0	1	1	0	1	1	0	0	0
V_4	0	0	0	1	0	0	0	0	0	1
V_5	0	0	0	0	1	1	0	0	1	0
V_6	1	0	0	0	1	0	1	1	0	1

Матрица инцидентности для него имеет вид

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Этому же рисунку соответствуют таблица и матрица смежности (табл. 2.3).

Таблица 2.3

Таблица смежности графа G

V_i	V_j					
	V_1	V_2	V_3	V_4	V_5	V_6
V_1	0	1	0	1	0	1
V_2	1	0	1	1	0	1
V_3	0	1	0	1	1	1
V_4	1	1	1	0	0	1
V_5	0	0	1	0	0	1
V_6	1	1	1	1	1	0

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Граф с кратными ребрами (особенно орграф) сложно задать с помощью матрицы смежности. Сделаем это формально. Если граф неориентированный, то справедливо $a_{ij} = a_{ji}$ и равно кратности ребра (V_j, V_i). В частности, если $i = j$, то a_{ij} — число петель при V_i. Недостаток подобного подхода заключается в том, что остается неучтенным взаимное расположение кратных ребер. Так, ребра могут переплетаться между собой, что, к сожалению, не отразится на матрице смежности.

Заметим, что для ориентированного графа данное определение графа без кратных ребер является частным случаем графа с кратными ребрами при кратности любого ребра, равной 1 или 0. Очевидно, что для двух вершин V_j и V_i $(i \neq j)$ существуют две принципиальные возможности: если все ребра выходят из одной и входят в другую вершину или если для каждой вершины существуют как входящие, так и исходящие ребра.

Пусть полная кратность ребра равна n, при этом из вершины V_i в вершину V_j исходят $m \leq n$ ребер, а из V_j в V_i исходят $n - m$ ребер. Тогда в клетке a_{ij} напишем m, а в клетке a_{ji} напишем $n - m$. Если есть кратные петли, то все они связаны с одной вершиной V_i, поэтому в клетке a_{ii} напишем кратность петли при V_i.

Такому заданию графа присущи те же недостатки, что и неориентированному, и еще неучет взаимного расположения направлений. Однако главным недостатком служит то, что при таком

определении матрицы смежности (как графа с кратными ребрами, так и без них) не всегда возможно определить по матрице смежности ориентированный граф или нет.

В матрицах инцидентности такой проблемы нет, так как наличие элемента вида –1 является критерием ориентированности графа. Для матрицы смежности несимметричность может являться достаточным условием ориентированности, но не критерием. Например, графу с матрицей смежности $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ может соответствовать отрезок V_1V_2 (и две вершины) — неориентированный граф или кольцо с двумя ребрами $V=\{(V_1, V_2); (V_2, V_1)\}$ — орграф. Это существенный недостаток, и возник он как раз при попытке определения матрицы смежности для графа с кратными ребрами. Поэтому для задания ориентированного графа с помощью матрицы смежности (если она получается симметричной) надо или указывать это отдельно, например $A_{\text{ор}}$, или у любого элемента матрицы написать «–».

Задача 19. Пусть граф G задан матрицей смежности A. Построить диаграмму этого графа, если

$$A=\begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

Решение. Поскольку матрица A несимметрична (например, $a_{35} \neq a_{53}$) и указания на ориентированность нет, A не может являться матрицей смежности реального графа.

Задача 20. Пусть граф G задан матрицей смежности A. Построить диаграмму этого графа, если

$$A=\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Решение. Диаграмму графа, имеющего шесть вершин, представим на рис. 2.19.

Любой ориентированный граф является бинарным отношением A под V, где V — множество вершин графа, а пары из X — ребра.

Для конечного числа V вершин отношение X можно представить тремя способами:

- графически, т.е. диаграммой (рис. 2.19);
- с помощью таблиц, в которых представлены 1 и 0;
- с помощью матриц (в случае матриц смежности).

Такая форма записи отношений удобна при решении многих логических и производственных задач. Она также используется при машинной обработке для систематизации информации.

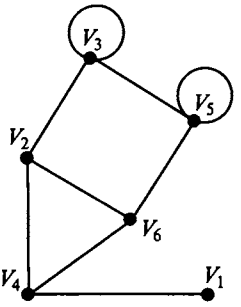


Рис. 2.19. Граф к задаче 20

2.5. Сети. Сетевые модели представления информации

> В этом проглядывается талант исследователя охватить значительные районы явлений с помощью немногочисленных допущений, представить разносторонние совокупности предметов и процессов в сжатой, компактной форме.
>
> *А. Сухотин*

Граф называется **взвешенным** или **сетью**, если каждому его ребру поставлено в соответствие некоторое число (**вес**). Взвешенными графами могут быть схемы в электронике, электрические схемы, карты автомобильных и железных дорог и др. Например, на картах автодорог вершины являются населенными пунктами, ребра — дорогами, а весом — числа, равные расстоянию между населенными пунктами.

В строительстве сетевые графы применяются для наглядного изображения некоторого комплекса работ или производственных процессов. Ребрам графа могут соответствовать числа, означающие длину, уклон, запланированное время и другие характеристики.

Например, последовательность работ для монтажа каркаса здания изображена в виде графа (рис. 2.20).

Числами обозначены технологические операции:

1 — рытье котлована;
2 — монтаж фундамента;
3 — завоз металлоконструкций;
4 — монтаж подъемного крана;
5 — монтаж каркаса здания.

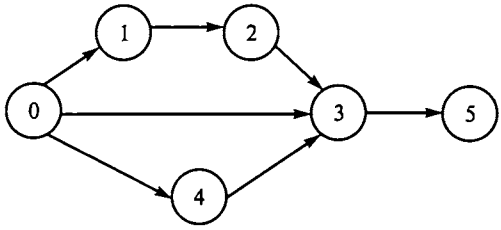


Рис. 2.20. Диаграмма последовательности работ при строительстве здания

На рис. 2.21 изображен сетевой граф некоторого комплекса работ в виде взвешенного графа с указанием времени, затраченного на выполнение этой работы (в минутах).

В основе процесса планирования лежит некоторый сценарий, представляющий собой сеть, состоящую из вершин — пошагового описания действий и дуг — отношений между ними. Такой граф дает возможность, сравнивая альтернативы, планировать действия для достижения поставленной цели.

Сети широко используются в качестве моделей для представления знаний в интеллектуальных системах. Сетевая модель представления информации основана на том, что любые знания можно рассматривать как множества объектов (понятий) и связи между ними (отношения).

Понятия-объекты и другие элементы предметной области могут быть графически изображены в виде вершин, а отношения между ними — в виде дуг, связывающих эти вершины. Такое графическое представление информации (знаний) в интеллектуальных системах носит название **семантических сетей**. Они являются универсальным средством для представления знаний в интеллектуальных системах. Понятия, входящие в сети, можно описать с помощью фрейма. **Фреймом** называется минимально возможное описание сущности некоторого явления, объекта, события или процесса. Состоит фрейм из набора стандартных единиц — *слотов*, содержащих определенный минимум информации о его содержании и назначении. Семантическая сеть в виде некоторой со-

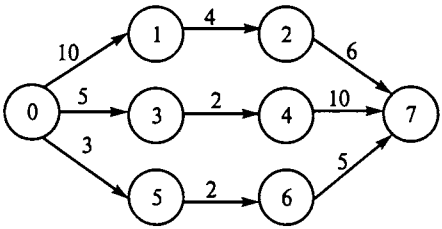


Рис. 2.21. Взвешенный сетевой граф проведения комплекса работ

вокупности фреймов нуждается в указании отношения между ее вершинами, что тоже возможно осуществить в виде слота. Семантические сети широко применяются в информатике, например, для операций **поиска по образцу**, где в виде сетей представляется база данных. Результат такого поиска можно изобразить графом. Используются сети и для графической иллюстрации системы отношений базы данных.

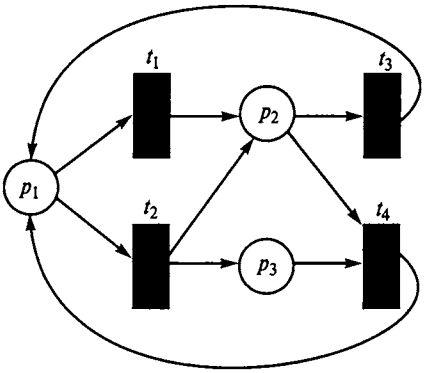


Рис. 2.22. Сеть Петри

Широко применяются сети для графического изображения различных логических схем в теории автоматов, например схемы с памятью, у которых каждый узел $F(i)$ — функция алгебры логики (см., например, рис. 4.17, 4.18).

Для формального описания совокупности процессов, протекающих одновременно, используют **сети Петри**. Они представляют собой ориентированные графы, состоящие из вершин двух видов: некоторых позиций и переходов, причем позиции изображают кружочками, а переходы — «планками» (рис. 2.22). Сети Петри предназначены для описания действия дискретных процессов во времени. Такие сети дают возможность моделировать ситуации протекания параллельных процессов, прослеживать возможные варианты их взаимодействия, выявлять нежелательные ситуации. Также в виде сетей изображаются схемы устройств, например радиоприемника или телевизора.

2.6. Применение графов и сетей

Храни порядок, и порядок сохранит тебя.

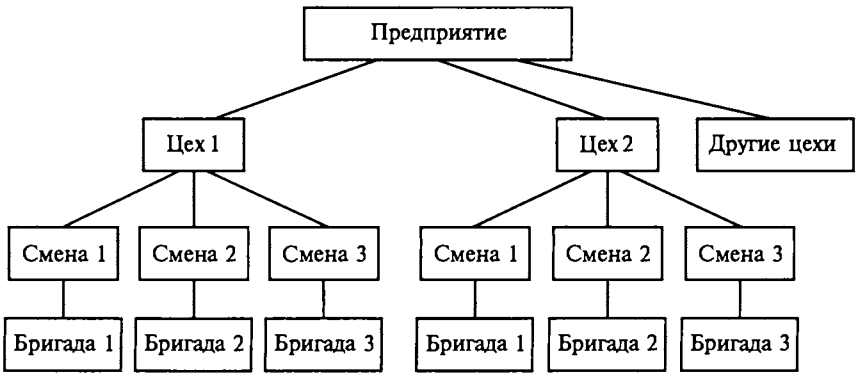
Латинская формула

Сети получили широкое практическое применение потому, что они являются естественным и удобным способом изображения и дальнейшего анализа различных сложных систем. Одним из первых применений сетевого изображения было создание американскими учеными баллистических ракет «Поларис» для оснащения атомных подводных лодок военно-морского флота США. В реализации этого грандиозного проекта участвовали 600 фирм из 48 штатов. Сам сетевой граф содержал 10 000 событий. В основе

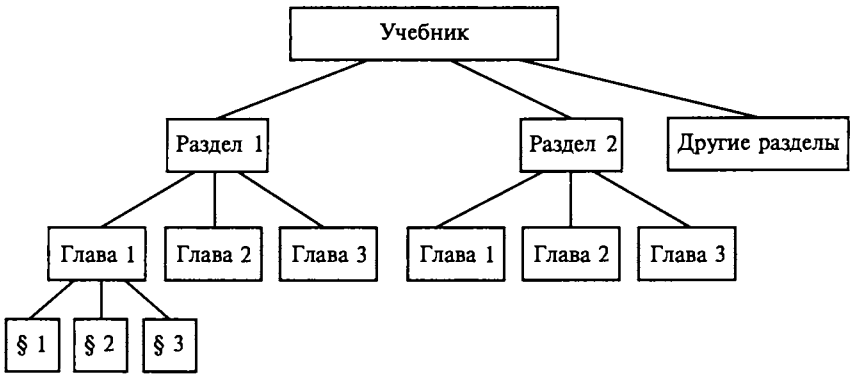
системы планирования и управления лежат в США система «Перт», а в нашей стране — «СПУ», широко и успешно применяющие графы, так как они позволяют обрабатывать на ЭВМ проекты с большим количеством событий.

С помощью графов-деревьев решают задачи планирования (дерево целей, дерево переборов вариантов). Графы используют также для иллюстрации классификаций в различных областях знаний при построении **иерархических структур** сложных систем.

Рассмотрим пример. Для классификации некоторого объекта (варианты его названия — $\frac{\text{предприятие}}{\text{учебник}}$) выберем в качестве корня дерева сам объект (рис. 2.23). Компоненты, составляющие этот объект, разместим на первом уровне (ярусе) графа: это будут $\frac{\text{цехи завода}}{\text{разделы книги}}$. Второй уровень (ярус) графа содер-



а



б

Рис. 2.23. Блок-схема иерархической структуры:
а — предприятие; *б* — учебник

жит компоненты, непосредственно связанные с компонентами первого уровня (его подмножества: $\frac{\text{рабочие смены}}{\text{главы учебника}}$). Аналогично на третьем уровне разместим подмножества компонентов второго уровня — $\frac{\text{бригады}}{\text{параграфы}}$. От нижнего уровня к высшему ведет лишь одна дуга, поэтому граф является деревом.

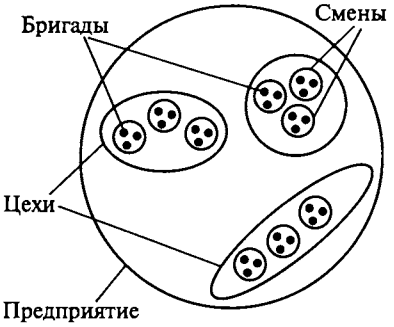


Рис. 2.24. Иллюстрация иерархической структуры предприятия с помощью кругов Эйлера

Так как иерархические структуры представляют собой подчиненные отношения (подмножества), то графически их можно изобразить либо в виде графа-дерева, либо с помощью кругов Эйлера (рис. 2.24).

Структуру такого типа имеют, например, предприятия (его составные части: цехи, бригады, участки и т.д.), учебные заведения (колледж состоит из факультетов, которые, в свою очередь, состоят из курсов, курсы — из групп). Такой же зависимости подчинены армейские соединения (дивизия, полк, батальон, рота, взвод, отделение, отдельные солдаты). Аналогично устроена любая административно-территориальная структура: республика, области, районы, населенные пункты. По путям этих деревьев движутся информационные потоки: сверху вниз — распоряжения, руководящие указания, снизу вверх — отчеты о работе, оперативная информация. Так как путь от листа к корню единственный, то его можно использовать для опознания компонентов системы. Например, почтовый адрес представляет собой «путь в дереве», аналогичный административно-территориальному. В разделе «Кому» указывается страна, республика, область, район, населенный пункт, улица, дом, квартира. Аналогично классифицируют объекты в любой науке. Получаемая классификация служит примером **иерархической структуры.** Например, в биологии: класс, отряд, семейство, род, вид. Соответствующий граф содержит элементы разных уровней, корень — класс, а листья — отдельные виды животных.

Иногда связи между объектами образуют не дерево, но все же их можно представить в виде графа. Это бывает в тех случаях, когда, например, происходит подчинение не одному, а нескольким независимым службам (соподчинение между собой).

В информатике иерархические структуры применяют при описании базы данных, вычислительных сетей, сетей связи, организационных систем. С помощью графа можно графически изображать родословные (генеалогическое дерево или древо).

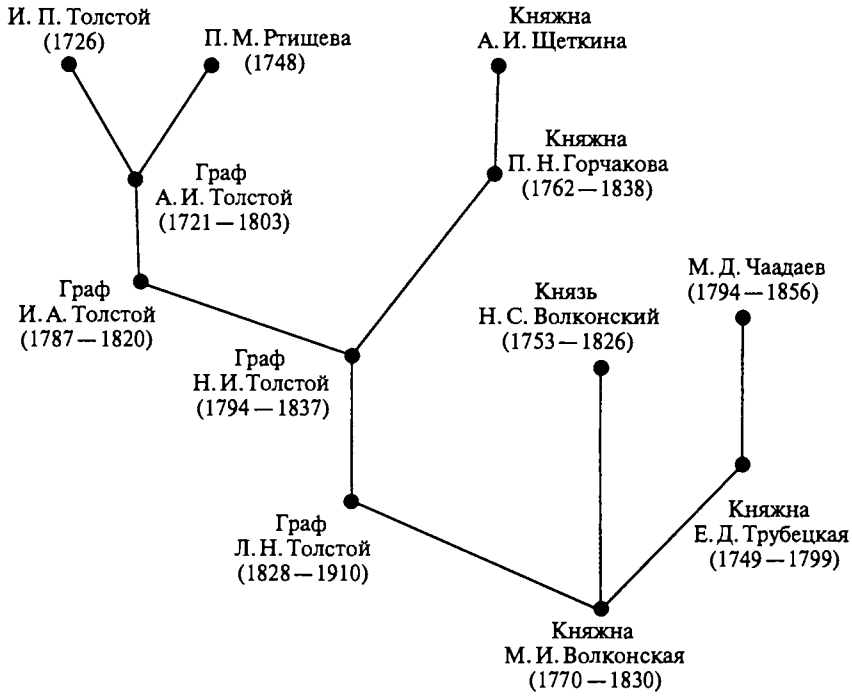


Рис. 2.25. Генеалогическое древо Л. Н. Толстого

Генеалогическое древо графа Льва Николаевича Толстого имеет вид, изображенный на рис. 2.25.

Бинарный поиск. Бинарные деревья применяются в информатике для одной из самых распространенных в прикладных науках операций — поиску. К нему прибегают, когда необходимо найти в некотором упорядоченном массиве (множестве) определенную информацию. Например, в телефонном справочнике — номер какого-нибудь абонента, в словаре — определенное слово, в файле — сведения о зарплате сотрудников некоторого предприятия, сведения о зарплате отдельного работника и т. д.

Последовательный, или линейный, поиск является наиболее общим и простым методом выявления интересующей информации: каждый элемент множества проверяется на соответствие заданным условиям. Если множество неупорядочено, то такой процесс будет носить случайный характер. Если множество было предварительно упорядочено (распределение по алфавиту фамилий в телефонном справочнике, слов в словаре, распределение служащих предприятия по табельным номерам в информации о зарплате), то удобнее использовать другой, более эффективный метод — бинарный поиск.



Рис. 2.26. Блок-схема иерархического представления видов вычислительных машин

Бинарный поиск основан на методе половинного деления. Поиск начинается с середины множества. Если первый же элемент удовлетворяет условию, то процесс поиска завершен, если не удовлетворяет, то процесс продолжается в любой из половин. Если искомый элемент так и не найден, то анализируется вторая половина до тех пор, пока не обнаружится соответствие найденного элемента заданным условиям. Бинарные деревья необходимы, когда делается выбор одной альтернативы из двух.

Пусть надо найти дубликаты всех чисел в некотором списке. Можно применить сравнение каждого числа с предшествующим на предмет «был — не был». Но тогда необходимо большое число сравнений. Использование бинарных деревьев (БД) укорачивает процедуру сравнения. Считывается новое число и помещается в узел — будущий корень нового БД. Каждое последующее число из списка сравнивается с числом в корне БД. Если эти значения совпадают, то дубликат найден, если число меньше корня, то процесс продолжается в правом поддереве, если больше — в левом поддереве.

Различные виды графов и деревьев находят широкое применение в учебном процессе. Поэтому первоначальные сведения о них необходимы для успешного обучения в различных областях, так как с их помощью часто передается учебная информация. Например, основные классы ЭВМ можно изобразить в виде графа (рис. 2.26).

На процедуре бинарного поиска с использованием графа-дерева основана работа ЭВМ с базой данных. Информация о базе может быть представлена несколькими способами, в том числе матричным. Так называемые реляционные базы хранят различную информацию в форме таблиц, причем порядок строк и столбцов задается при вводе данных. В базах возможна длительная работа с информацией, ее реорганизация и обновление. С помощью автоматического поиска данных происходит их

отбор на основании запроса по определенным характерным признакам.

Запросы в виде сложносоставных высказываний образуются из простейших с помощью *логических связок* и ряда *логических операций*. Все методы работы с информацией — поиск, обработка и накопление — основаны на широком применении законов *математической логики*, знание которой необходимо для понимания принципов работы вычислительных машин, различных электронных устройств.

? Что изучает новая для нас наука логика? Почему она математическая? Какие еще могут быть логики? Чем разные логики отличаются друг от друга?

Логикой называют науку о законах и формах *правильного* мышления.

А разве можно научиться мыслить правильно? Что значит правильное и неправильное мышление? Возможно ли постичь секреты идеально точных рассуждений?

Познакомившись с первыми разделами дискретной математики, вы, конечно, обратили внимание на множество новых понятий и их определений.

Так ли уж важно определять новые понятия и требовать точности формулировок?

На эти и другие неожиданные вопросы попытаемся ответить в следующей главе, которая приоткроет дверь в хранилище человеческой мудрости — первоначальные представления об основах логики как классической, так и математической. Причем связь между логикой и математикой не случайна: на протяжении двух с половиной тысячелетий они не только параллельно развивались, но и обогащали друг друга. Недаром над входом в философскую школу Платона — Академию — было начертано: «Пусть сюда не входит тот, кто не знает геометрии». Построение математики, в первую очередь геометрии, как точной науки, имеющей достоверные выводы, дает возможность учиться на ее примерах логически мыслить. И это делали еще древнегреческие философы. Знание законов логики, философских законов, например диалектики, помогает при изучении самой математики, так как способствует установлению предельно обобщенного, системного взгляда на отношения в окружающем мире как в природе, так и в обществе.

Упражнения

2.1. Пусть орграф задан матрицей смежности. Постройте изображение этого графа, укажите степени вершин графа. По матрице смежности постройте матрицу инцидентности этого графа:

а)

V	V_1	V_2	V_3	V_4	V_5	V_6
V_1		1			1	1
V_2	1		1		1	
V_3		1	2			
V_4				2		
V_5	1	1				1
V_6	1				1	

б)

V	V_1	V_2	V_3	V_4	V_5	V_6
V_1			1	1		
V_2				1		1
V_3	1				1	1
V_4	1	1			1	
V_5			1	1	2	
V_6		1	1			

в)

V	V_1	V_2	V_3	V_4	V_5	V_6
V_1			1	1		
V_2		2	1			1
V_3	1	1		1		
V_4	1		1		1	1
V_5				1		
V_6	1	1		1		

г)

V	V_1	V_2	V_3	V_4	V_5	V_6
V_1	2			1		
V_2			1			1
V_3		1		1	1	
V_4	1		1			1
V_5			1			1
V_6		1		1	1	

д)

V	V_1	V_2	V_3	V_4	V_5	V_6
V_1					1	1
V_2		2				1
V_3				1		
V_4			1		1	1
V_5	1			1		
V_6	1	1		1		

е)

V	V_1	V_2	V_3	V_4	V_5	V_6
V_1			1	1		
V_2					1	1
V_3	1			1		1
V_4	1		1		1	
V_5		1		1		
V_6		1	1			2

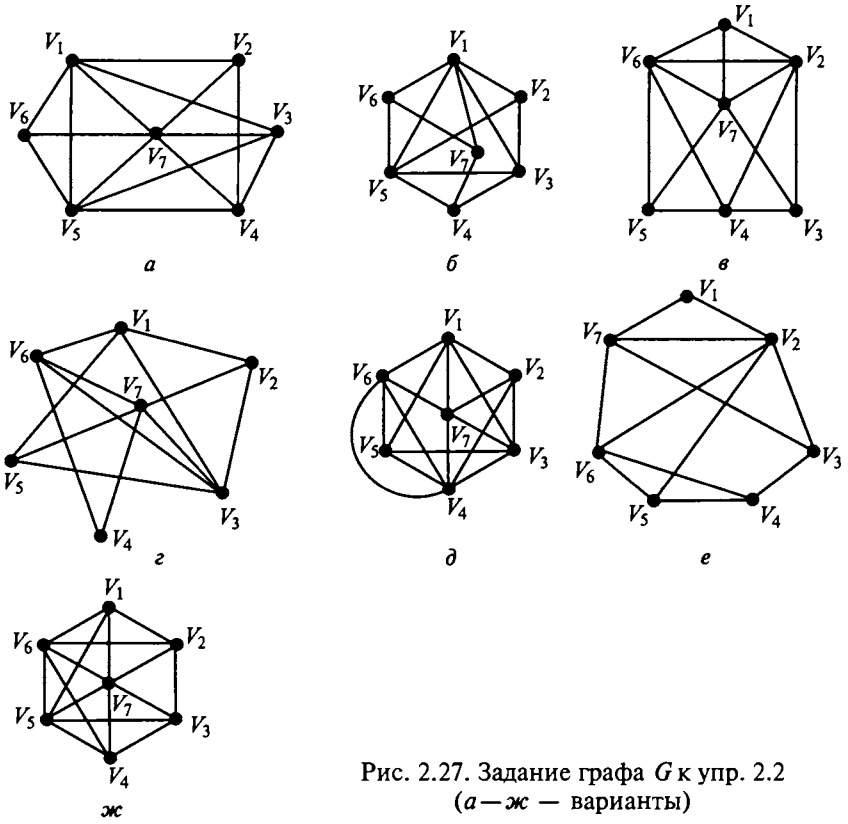


Рис. 2.27. Задание графа G к упр. 2.2 ($а$—$ж$ — варианты)

2.2. Граф G задан диаграммой (рис. 2.27).

1. Составьте для него матрицу смежности.
2. Постройте матрицу инцидентности.
3. Укажите степени вершин графа.
4. Найдите длину пути из вершины V_2 в вершину V_5, составьте маршруты длины 5, цепь и простую цепь, соединяющие вершину V_2 и вершину V_5.
5. Постройте простой цикл, содержащий вершину V_4.
6. Найдите цикломатическое число графа G.
7. Определите вид заданного графа.

2.3. Найдите объединение и пересечение графов G_1 и G_2, дополнение для графа G_1 (рис. 2.28).

2.4. Постройте матрицу смежности и матрицу инцидентности для отношений, заданных графом G. Найдите число степеней входа и выхода этого графа, дайте ему характеристику (рис. 2.29).

2.5. Орграф задан матрицей смежности. Постройте его рисунок (схему, диаграмму), определите степени вершин графа и найдите маршрут длины 5. Есть ли среди них изоморфные?

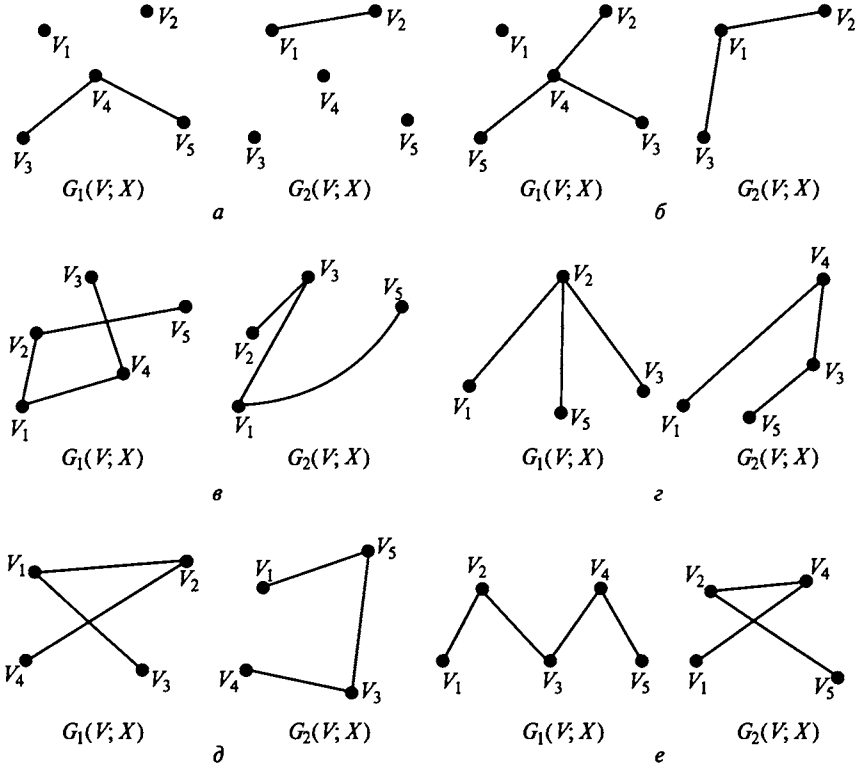


Рис. 2.28. Задание графов G_1 и G_2 к упр. 2.3 ($а$ — $е$ — варианты)

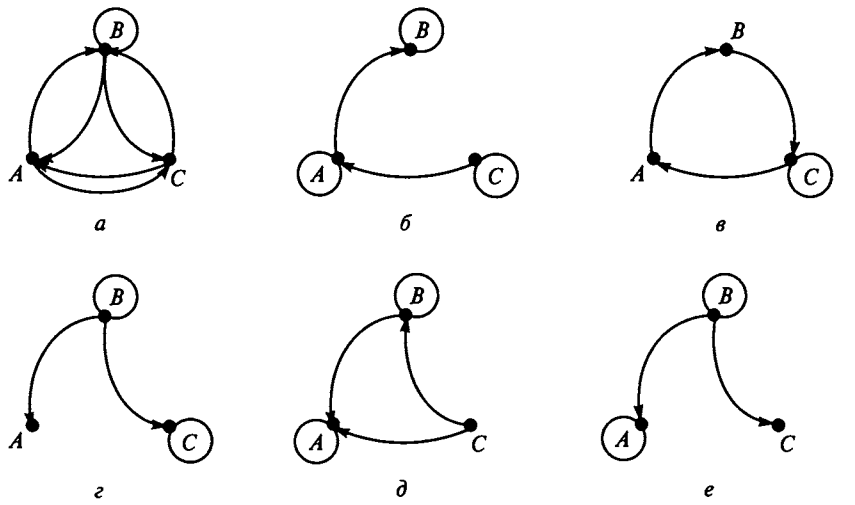


Рис. 2.29. Задание графа G к упр. 2.4 ($а$ — $е$ — варианты)

а) $G = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$;

б) $G = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$;

в) $G = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$;

г) $G = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$;

д) $G = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$;

е) $G = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$.

2.6. Составьте все возможные планы маршрута путешествия по историческим местам, если автотуристам надо проехать из пункта M в пункт N, осмотрев все памятники архитектуры не более одного раза. Как называется такой маршрут (рис. 2.30)?

2.7. Ориентированный граф $G(V, X)$ с множеством вершин $V = \{1, 2, 3, 4, 5, 6, 7\}$ задан списком дуг X.

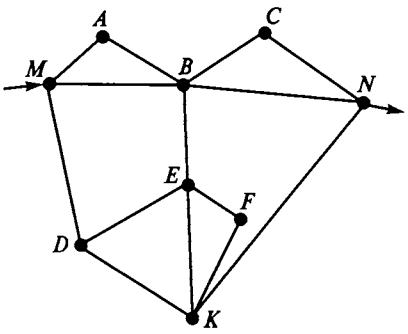


Рис. 2.30. Граф путешествия по историческим местам

1. Постройте реализацию графа G.
2. Постройте матрицу инцидентности графа G.
3. Постройте матрицу смежности G.
4. Задайте соответствующий неориентированный граф матрицей смежности.
5. Укажите степени вершин полученных графов, найдите цикломатическое число графа G.

а) $X = \{(1, 2), (2, 3), (4, 3), (4, 5), (6, 5), (7, 6), (7, 1), (7, 7), (7, 2), (6, 4), (4, 4), (2, 7), (6, 4), (5, 3)\}$;

б) $X = \{(1, 4), (2, 1), (4, 3), (4, 5), (2, 6), (2, 6), (7, 1), (7, 6), (3, 2), (5, 4), (3, 4), (2, 2), (6, 2), (5, 5)\}$;

в) $X = \{(1, 5), (2, 3), (2, 3), (4, 5), (4, 6), (5, 6), (5, 1), (6, 6), (3, 2), (5, 4), (6, 4), (7, 2), (6, 7), (7, 5)\}$;

г) $X = \{(1, 1), (2, 2), (2, 3), (3, 5), (4, 6), (4, 6), (5, 1), (5, 6), (5, 2), (6, 4), (7, 4), (7, 2), (7, 2), (7, 5)\}$;

д) $X = \{(1, 1), (1, 3), (1, 3), (2, 5), (2, 6), (3, 6), (3, 1), (3, 6), (3, 7), (4, 4), (4, 6), (5, 2), (6, 3), (6, 5)\}$;

е) $X = \{(1, 3), (2, 3), (2, 3), (3, 5), (3, 6), (2, 7), (4, 1), (4, 6), (4, 2), (6, 4), (6, 4), (7, 2), (6, 6), (7, 6)\}$.

2.8. Составьте сценарий и по нему постройте сетевой граф, иллюстрирующий порядок выполнения операций, для того чтобы:

а) выпустить газету;

б) провести шахматный турнир на первенство колледжа;

в) подготовить и провести в колледже КВН;

г) посадить и вырастить картофель;

д) организовать работу торговой точки;

е) изготовить табурет.

2.9. Решите задачи «о переправах», изобразите решение графом.

1. Три генерала — Строгий, Лихой и Грозный — со своими адъютантами переправлялись через реку с помощью двухместной лодки. Адъютант может либо перевозить своего генерала, либо переправляться с другим адъютантом. Однако ни один из генералов не разрешил своему адъютанту ни оставаться с другим генералом вдвоем на берегу, ни переправляться с ним через реку. Как они переправились через реку?

2. Трое мужчин и три женщины должны переправиться через реку. У них была одна лодка, которая вмещала только двух человек. Грести умели все мужчины и только одна женщина. Кроме того, женщины требовали, чтобы ни на одном берегу не оставалось больше женщин, чем мужчин. Как им переправиться через реку?

3. Муж, жена и двое детей должны переправиться на противоположный берег реки при помощи лодки. Муж и жена весят по 100 кг, а дети — по 50. Как им быть, если лодка вмещает до 100 кг и каждый из них умеет грести?

4. Человеку необходимо было переправить через реку с помощью лодки волка, козу и капусту. В лодке мог поместиться только человек, а с ним или волк, или коза, или капуста. Но если оставить волка с козой без человека, то волк съест козу, если оставить козу с капустой, то она съест капусту, а в присутствии человека никто никого не ел. Человек все-таки перевез через реку и волка, и козу, и капусту. Как он это сделал?

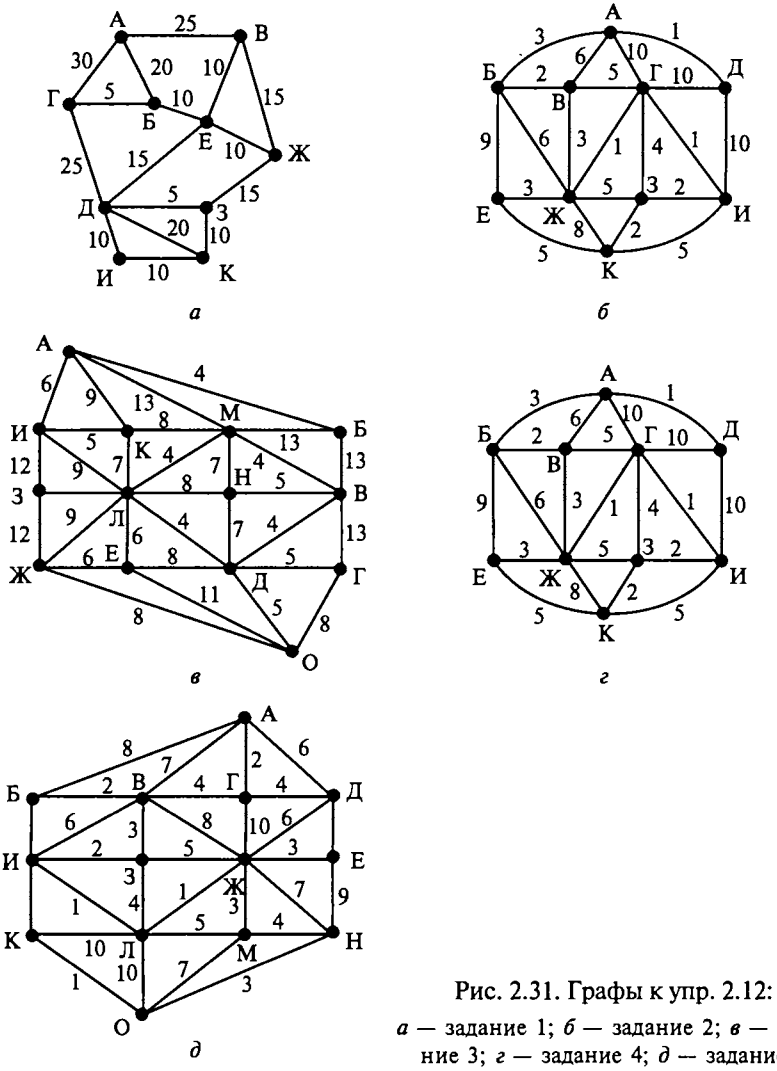


Рис. 2.31. Графы к упр. 2.12:
а — задание 1; *б* — задание 2; *в* — задание 3; *г* — задание 4; *д* — задание 5

2.10. Задачи на поиск «фальшивой монеты» решите с помощью графов.

1. Из 9 монет одна фальшивая (более легкая). Как двумя взвешиваниями на чашечных весах определить фальшивую монету?

2. Из 80 одинаковых по виду монет одна более легкая (фальшивая). Как четырьмя взвешиваниями на чашечных весах определить фальшивую?

3. Из 28 монет одна более легкая. Как при помощи 4 взвешиваний определить ее?

4. Из 27 монет одна более легкая. Как при помощи 3 взвешиваний определить ее?

5. Из 81 монеты одна более легкая. Показать, что 4 взвешиваний достаточно, чтобы ее определить.

6. Из 82 монет одна более легкая. Какое наименьшее число взвешиваний необходимо для определения этой монеты?

7. Из m одинаковых по виду монет одна фальшивая (более легкая). Указать наименьшее число взвешиваний, необходимых для определения фальшивой монеты.

2.12. Решите задачи, используя графы (рис. 2.31).

1. Винни-Пух вышел на прогулку, взяв с собой карту (рис. 2.31, *а*). Числа на рисунке обозначают время движения (в минутах) от пункта до пункта. Помогите Винни-Пуху найти кратчайший путь от своего дома в пункте А до дома Пятачка в пункте К. Перечислите пункты, через которые должен пройти Винни-Пух, и подсчитайте время, которое он затратит на весь путь.

2. Атос поскакал в гости к Портосу, взяв с собой карту (рис. 2.31, *б*). Числа на рисунке обозначают время движения (в часах) от пункта до пункта. Помогите Атосу найти кратчайший путь от своего поместья в пункте Е до поместья Портоса в пункте Д. Перечислите пункты, через которые должен проехать Атос, и подсчитайте время, которое он затратит на весь путь.

3. Рыцарь, находясь в пункте А, узнал, что Прекрасной Даме, в пункте О, ровно через сутки может грозить опасность. Взяв с собой карту (рис. 2.31, *в*), он немедленно выехал на помощь. Числа на рисунке обозначают время движения (в часах) от пункта до пункта. Успеет ли рыцарь спасти Прекрасную Даму? Обоснуйте ответ, указав кратчайший маршрут.

4. Рыцарь, находясь в пункте А, узнал, что Прекрасной Даме, в пункте К, через 14 часов может грозить опасность. Взяв с собой карту (рис. 2.31, *г*), он немедленно выехал на помощь. Числа на рисунке обозначают время движения (в часах) от пункта до пункта. Успеет ли рыцарь спасти Прекрасную Даму? Обоснуйте ответ, указав кратчайший маршрут.

5. Перед вами — карта (рис. 2.31, *д*). Числа на карте обозначают время движения (в часах) от пункта до пункта. Можно ли успеть доехать из пункта А в пункт О за 22 часа? В случае положительного ответа укажите маршрут, в случае отрицательного ответа обоснуйте его.

6. Придумайте и решите аналогичную задачу.

Глава 3
ПОНЯТИЯ

В этой главе мы узнаем, как теоретико-множественный подход поможет познакомиться с такой важной наукой, как логика. Знание логических приемов формирования понятий и логических операций над ними обогащает тех, кто изучает обе эти науки. Дело в том, что дискретная математика оперирует большим количеством новых понятий, таких, как «отношения», «классификации», «подмножество» и др., и знание основ логики поможет разобраться в их особенностях.

Место отдельных понятий в системе знаний устанавливается благодаря различным операциям над понятиями: ограничению, обобщению, определению, а также делению понятия, которые рассматриваются в этой главе.

3.1. Понятие как форма мышления

...Настоящий математик работает не
с числами, а с понятиями.

Я. Стюарт

Логика как наука возникла в трудах выдающегося древнегреческого мыслителя Аристотеля (384—322 до н. э.). Традиционная аристотелева логика называется **классической** или формальной и соответствует *первому периоду* развития этой науки. Она создана как орудие или метод исследования формальной правильности рассуждений в зависимости от логических операций, производимых в процессе мышления. Объектами изучения такой логики являются различные формы мышления — понятия, суждения, умозаключения, их виды и операции над ними, законы логики, доказательства и опровержения, гипотезы и т. д. Наряду с Аристотелем, тщательно разработавшим теорию дедуктивных умозаключений и доказательств, в классическую логику входит индуктивная логика, родоначальником которой является крупнейший английский философ и естествоиспытатель Фрэнсис Бэкон (1561—1626). Значительный вклад в развитие классической логики внес также великий французский философ и математик Рене Декарт (1596—1650).

Второй период развития логики как науки связан с работами знаменитого немецкого философа и математика Готфрида Вильгельма Лейбница (1646—1716), целью которого было создание «азбуки понятий» и применение логики для теоретического обоснования математики. С другой стороны, у Лейбница возникла идея придать рассуждениям вид вычислений. Для этого он хотел поставить символы в соответствие понятиям и получить верные выводы в рассуждениях с помощью некоторых алгебраических операций. Стремясь облегчить процесс мышления, он старался смотреть на классическую логику через призму математики.

Мечты Лейбница частично удалось воплотить в жизнь ирландскому математику и логику Джорджу Булю (1815—1864), который создал алгебру логики. В этой науке свои операции и законы похожи на привычные математические. В этом разделе математики, получившем название *булева алгебра*, буквами обозначают высказывания как элементы множества высказываний. С помощью логических операций из простых элементарных высказываний строят составные, определяют, истинны они или ложны, т.е. определяют их **семантическую** характеристику. Алгебра Буля — часть математической логики, в состав которой входят логика высказываний и логика предикатов. Математическая логика в отличие от традиционной изучает логические связи и отношения, лежащие в основе достоверных логических выводов. В математической логике для получения достоверных выводов выявляется их структура, и на этой основе строятся логические исчисления. Второй период развития логики связан с использованием различных математических символов, поэтому его называют **символическим**.

Наиболее характерной особенностью современного *третьего периода* развития логики является применение новых методов для исследования как традиционных, так и вновь выявленных логических проблем, связанных с появлением ЭВМ и информатизацией общества.

Кроме перечисленных, существует еще и *диалектическая* логика, отличающаяся от формальной своим специфическим подходом к мышлению. У ее истоков стояли немецкие философы Иммануил Кант (1724—1804) и Георг Гегель (1770—1831). Особенностью диалектической логики является учет противоречивости мышления, основанный на законах диалектики.

Все перечисленные разновидности логики тесно взаимосвязаны между собой и важны как для освоения других наук, так и для формирования современного эрудированного специалиста. Поэтому и необходимо наряду с изучением элементов математической логики познакомиться с основами формальной и диалектической логики, уделив основное внимание взаимосвязям с математикой и практическому применению.

Связь между математикой и логикой. И логика, и математика уделяют большое внимание работе с понятиями. В логике **понятие** рассматривается как форма мышления, отражающая предметы в их существенных признаках. Понятия — это наши знания о некоторых предметах и явлениях окружающего мира или о множестве сходных объектов. Глубинный смысл этих знаний заключается в установлении *отношений* между исследуемым понятием и другими

объектами или идеями. Причем для выявления этих отношений мы используем *суждения*, т.е. мысли, связывающие понятия между собой. Именно с помощью суждений раскрывается сущность исследуемых понятий. В свою очередь, суждения и их упрощенная модель — высказывания — являются объектом изучения математической логики — одного из важных разделов дискретной математики. Из основных функций, свойственных понятиям, наиболее значимыми являются познавательная и коммуникативная. В процессе развития науки знания, приобретенные *опытным* путем (эмпирические знания), углубляются, систематизируются, обобщаются и уточняются. Они становятся, с одной стороны, научными данными, с другой — служат средством общения между людьми в их совместной деятельности, осуществляя преемственность поколений.

Различные *отношения* (или связи) между понятиями были изучены в гл. 1 и 2 с точки зрения теоретико-множественного подхода. Отношения рассматриваются также в логике предикатов — составной части математической логики. Так, с логико-математической точки зрения предикаты есть пропозиционные функции, где аргументами являются различные высказывания, т.е. логические объекты. Следовательно, изучение отношений между понятиями должно сочетать в себе как теоретико-множественный, так и логические подходы.

В связи с тем, что понятия естественного языка многозначны, часто приходится уточнять их определения в информации разного вида.

Так, «квадрат» может быть и геометрической фигурой (прямоугольником с равными сторонами), и второй степенью числа или алгебраического выражения. Например, плоскость $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ — декартово произведение двух прямых, а сам квадрат (фигура) — это квадрат (декартов) двух отрезков: $[0, a] \times [0, a]$.

Аналогично, под «моделью» можно понимать и алгебраическую систему, в которой определены только отношения, а множество операций пусто, и некоторую интерпретацию формального языка и аналог предмета или явления, сохраняющего его существенные черты. Часто используются физические, экономические, математические и другие модели.

Различные науки изучают одни и те же предметы и явления в зависимости от собственных интересов. Они как окна, через которые люди смотрят на объект изучения. Так, математика будет изучать количество кусочков мела, их форму, установит их размеры, площадь поверхности, объем. Физика будет изучать, с какой скоростью мел упадет на пол, как при этом он деформируется, какие силы действуют на мел при его скольжении по поверхности доски и т.д. Химия заинтересуется строением мела как вещества, его способностью изменяться при воздействии высокой температуры и т.д. Иначе говоря, любой предмет обладает различными

признаками, но для каждой конкретной науки одни признаки будут существенными, а другие — несущественными. Признаки предмета — это те особенности, которые присущи данному предмету и отличают его от других.

Предметы составляют объективную реальность, существующую независимо от сознания человека. Но человек каждому предмету поставил в соответствие слово, термин, понятие, отличающее этот предмет от других. Причем реальные предметы и предметы мысли принадлежат к разным областям действительности: реальные предметы к объективному, а предметы мысли — к субъективному миру человека.

Итак, **понятие** — это форма мышления, отражающая предметы в их существенных общих признаках. Каждый из признаков необходим для описания некоторого понятия, а все вместе они достаточны для того, чтобы с их помощью отличить данное понятие от других из общего множества однородных объектов.

Под **признаком** обычно понимают свойства и отношения реальных вещей, выраженные в этом понятии. Отношение между понятиями называют предикатами (подробно см. в гл. 4).

Логические приемы формирования понятий. Для формирования понятий используются следующие мысленные логические приемы.

Анализ — мысленное расчленение объектов (явлений, процессов) на их составные части, выделение в них признаков.

Синтез — соединение частей или признаков в единое целое.

Сравнение — установление сходства или различия объектов по существенным или несущественным признакам.

Абстрагирование — выделение одних признаков и отвлечение от других (чаще всего — выделение существенных и отказ от несущественных признаков).

Обобщение — объединение однородных объектов в класс.

Например, теория множеств с высокой степенью абстракции оперирует различными знаками — символами, которые представляют некоторые объекты, но зачастую никак с ними не связаны по смыслу. *Сравнивая* между собой естественный и математический языки, мы анализируем их характерные черты и *обобщаем* в единый класс: **теорию знаков и знаковых систем**, науку **семиотику**. В семиотике *синтезируются* три вида отношений: синтактические, семантические и прагматические, с некоторыми из которых мы будем в дальнейшем встречаться. Поэтому, чтобы понимать смысл последующих глав, необходимо раскрыть содержание этих понятий.

Под **синтактическими** понимают отношения между знаками, которые зафиксированы в правилах и способах образования языковых выражений. **Семантические** отношения характеризуют смысл языковых выражений: они могут быть либо истинными,

либо ложными. Семантические правила определяют способы придания значений выражениям языка.

Логические характеристики понятий. Любое понятие имеет логическую структуру, включающую в себя объем и содержание — основные логические характеристики понятия.

Содержание понятия характеризует совокупность существенных признаков предмета, отраженных в понятии. Например, содержание понятия «молекула» составляет ее основной отличительный признак — быть мельчайшей частицей вещества, отражающей его физические и химические свойства.

Содержание понятия «квадрат» составляют его отличительные признаки: «быть прямоугольником и иметь равные стороны», либо «быть ромбом и иметь прямые углы».

Объем понятия есть множество всех предметов, которые мыслятся в данном понятии. Объем понятия «столицы государств» определяется перечислением всех столиц: Москва, Париж, Лондон, Пекин и т.д. Объем понятия «стороны горизонта» составляют восток, запад, север и юг.

Для краткой лаконичной записи взаимоотношений, взаимосвязей между различными понятиями будем использовать язык теории множеств и теории графов.

Пусть объем одного понятия входит в объем другого. Из двух понятий A и B назовем **родовым** понятием то из них (A), для которого объем *больше,* а **видовым** то (B), объем которого *меньше*: $B \subset A$ (рис. 3.1, *а*).

Закон обратного отношения между объемом и содержанием понятий. Взаимосвязь между логическими характеристиками понятия, его объемом и содержанием, выражает закон обратного отношения. *Чем шире объем понятия, тем уже его содержание, и, наоборот, чем шире содержание, тем уже объем.* Это значит, что чем меньше информации о предмете, заключенном в понятии, тем шире класс предметов, соответствующих этому понятию. Например, «игры» ($И$) — понятие более широкое, чем «компьютерные игры» ($К$). Поэтому объем понятия $И$ шире, чем объем понятия $К$: $К \subset И$ (рис. 3.1, *б*).

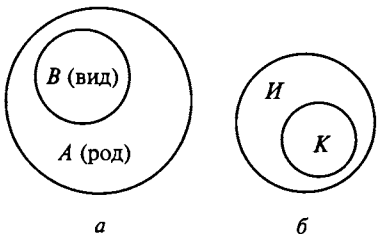


Рис. 3.1. Закон обратного отношения между объемом и содержанием понятия:

а — соотношение между родовыми и видовыми понятиями; *б* — $К \subset И$

Другой пример. Свидетель ДТП видел, что автомобиль, сбивший пешехода и скрывшийся с места преступления, был «девяткой» цвета «мокрый асфальт». А другой свидетель запомнил две первые цифры номера: эта дополнительная информация (содержание

шире) значительно сузила круг возможных автомобилей, причастных к ДТП (объем уже). На рис. 3.2 показано: A — множество автомобилей, B — множество «девяток» цвета «мокрый асфальт» (это пересечение множества всех «девяток» и множества машин цвета «мокрый асфальт»), C — множество автомобилей с подобными первыми цифрами. Очевидно, что $|C| \leq 10 \cdot 89 \cdot 33^3$ (с учетом числа регионов) при трехцифровом номере. Поскольку $B \cap C \subset B$ и $B \cap C \subset C$, то объем возможных правонарушителей сужается при добавлении новых признаков, т.е. расширении содержания. Более того, добавляя новые очевидные признаки, например регион, можно добиться разумного числа возможных вариантов, которые легко проверить на предмет алиби и технического состояния автомобиля.

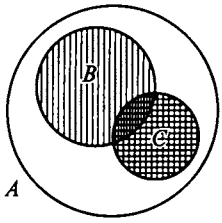


Рис. 3.2. Иллюстрация отношения между понятиями

3.2. Логические операции над понятиями: обобщение и ограничение понятий

> В основе любого понятия лежит абстракция, то есть отвлечение конкретных, интересующих исследователя признаков, когда все остальные признаки опускаются... Математика объединяет предметы в классы, совершенно не считаясь с их природными признаками.
>
> *А. Сухотин*

В наших рассуждениях вообще, при изучении точных наук в частности, велика роль *ограничений* и *обобщений* понятий.

Под **ограничениями** понимается переход от рассмотрения изучаемого *множества* к рассмотрению его *подмножества*. Так, первое предложение этого подраздела иллюстрирует логическую операцию *ограничения* понятий: от множества рассуждений, в которых велика роль ограничений понятия, мы перешли к его подмножеству, которое обладает тем же характерным свойством — рассуждением, связанным с изучением точных наук.

Все множество имеет больший объем, чем его подмножество, и, следовательно, его характеризует мало признаков (рассуждения вообще). Для его подмножества ограничивается объем, а значит, согласно закону обратного отношения между объемом и содержанием увеличивается количество признаков (рассуждений, связанных с изучением точных наук).



Рис. 3.3. Схема взаимно-обратных переходов ограничения и обобщения понятий

Обобщением является обратная логическая операция, которая осуществляет переход от рассмотрения данного множества к рассмотрению более широкого множества, содержащего данное. Та же мысль (суждение), высказанная в первом предложении этого подраздела, могла быть представлена в виде операции обобщения: в рассуждениях, применяемых при изучении точных наук, да и вообще в рассуждениях, велика роль ограничений и обобщений понятий. В данном предложении от множества *рассуждений, применяемых в точных науках*, мы осуществили переход к более широкому множеству *рассуждений вообще*, т.е. перенесли признаки, которыми обладало некоторое подмножество, на все множество.

Эти взаимно-обратные переходы будем иллюстрировать кругами Эйлера. При этом все множество является родовым понятием, а его подмножество — видовым (рис. 3.3).

Ограничение →

A (наука)	*Aa* (математика)	*Aab* (геометрия)	*Aabc* (планиметрия)

← Общее — Обобщение — Единичное

Рис. 3.4. Схема взаимно-обратных переходов ограничения и обобщения

Тогда операция обобщения, заключающаяся в переходе от видового к родовому понятию, сопровождается отказом от отдельных существенных признаков понятия. Так, если общее понятие — наука (все множество различных наук), то одно из ее подмножеств — математика. В свою очередь математика разделяется на отдельные предметные области, одной из которых является геометрия — наука, изучающая пространственные фигуры и отношения между ними. Та часть геометрии, которая изучает фигуры и их отношения на плоскости, называется планиметрией. Совершим с перечисленными понятиями логические операции обобщения и ограничения и изобразим их с помощью кругов Эйлера. Как известно, ограничить — значит добавить признаки *a*, *b*, *c*, *d*, обобщить — отбросить признаки. Общим будет понятие наука, единичным — планиметрия. Обозначим прописными буквами видовые признаки понятия. Поставим в соответствие отдельным понятиям символы, соответствующие некоторым признакам: наука (*A*), математика (*a*), геометрия (*b*), планиметрия (*c*) (рис. 3.4).

Значение логических операций обобщения и ограничения понятий состоит в своеобразном способе закрепления полученных знаний, способе, важном для правильного мышления любого человека.

3.3. Отношения между понятиями

— Какая разница между пуганой вороной и письменным столом?..

Л. Кэрролл

Понятия, как и предметы или идеи, им соответствующие, находятся в определенных отношениях, взаимосвязях друг с другом.

Если содержания понятий имеют общие признаки, они называются **сравнимыми**. Например, телевидение и радио — понятия сравнимые, так как имеют общий признак — они являются средствами массовой информации. Компьютер и снегопад — далекие, несравнимые понятия.

Виды отношений между понятиями по объему удобно представить в виде графа, а взаимосвязи между его частями проиллюстрировать с помощью кругов Эйлера (рис. 3.5).

Одной из наиболее характерных особенностей современной математики является ее высокая степень *абстрагирования*. Понятие объединяет множество объектов, обладающих определенными свойствами.

Некоторая абстрактная теория выводит следствия из этих свойств, которые впоследствии можно будет применить к любому объекту этого множества. Абстрагирование достигается за счет выполнения логической операции *обобщения*. Благодаря обобщению мы переносим свойства одного объекта на свойство другого, с ним родственного.

Так, доказав, что в данном прямоугольном треугольнике, имеющем стороны длиной 3, 4 и 5, квадрат гипотенузы равен сумме квадратов катетов, мы обобщаем этот вывод сначала на все прямоугольные треугольники со сторонами, пропорциональными этим числам («египетский треугольник» вида $(3k)^2 + (4k)^2 = (5k)^2$), а затем на произвольный прямоугольный треугольник, т.е. $a^2 + b^2 = c^2$.

Рассматривая некоторое множество (1) (рис. 3.6), можно совершить операцию *обобщения* (перейти к предку (2) — «надмножество»), провести операцию *ограничения* и рассматривать частный случай (потомок (5) — «подмножество») или рассмотреть аналогичные понятия, расположенные на одном уровне (братья (3) и (4)).

Для понятия «параллелограмм» аналогичными, т.е. обладающими сходными свойствами, будет понятие «параллелепипед» (и в меньшей степени — «прямая призма»).

Обобщением понятия «параллелограмм» является многоугольник, а ограничением — виды параллелограмма, например ромб или прямоугольник (рис. 3.7, *а*).

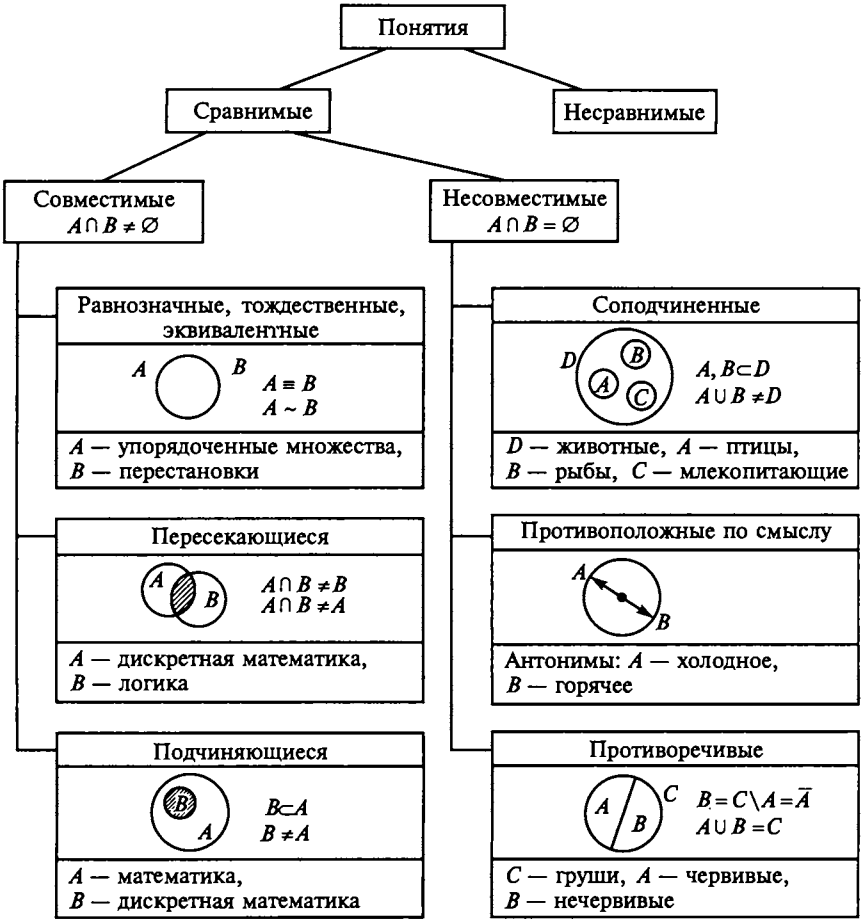


Рис. 3.5. Классификация отношений между понятиями

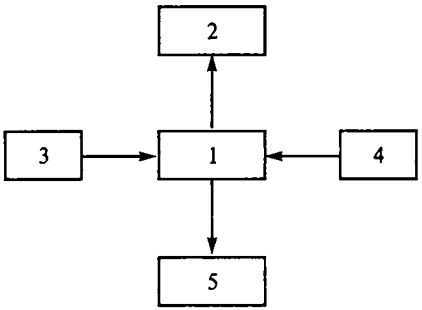
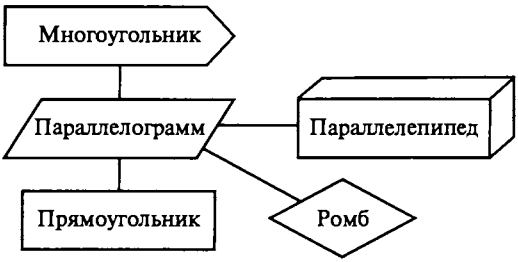
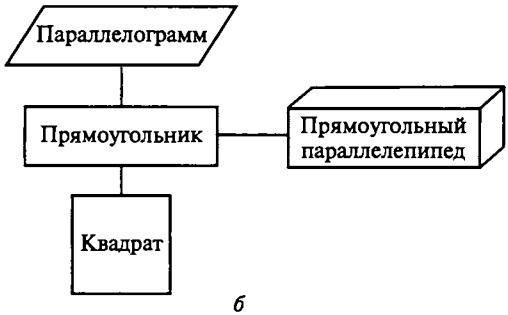


Рис. 3.6. Схема логических переходов к подмножеству (ограничение понятия), надмножеству (обобщение понятия) и аналогичным (соподчиненным) понятиям



а



б

Рис. 3.7. Примеры изображения логических операций для понятий: *а* — параллелограмм; *б* — прямоугольник

Действительно, $\frac{\text{параллелограмм}}{\text{параллелепипед}}$ обладают сходными свойствами:

- противоположные $\frac{\text{стороны}}{\text{грани}}$ равны и параллельны;
- $\frac{\text{противоположные углы}}{\text{углы с соответственно параллельными сторонами}}$ равны.

Аналогично, для понятия «прямоугольник» обобщением является параллелограмм, ограничением — квадрат. Сходными свойствами обладает прямоугольный параллелепипед (рис. 3.7, *б*).

К перечисленным свойствам, рассмотренным для параллелограмма, мы добавим еще одно.

Справедливо утверждение: $\frac{a^2 + b^2 = d^2}{a^2 + b^2 + c^2 = d^2}$,

где a, b, c — длины сторон $\frac{\text{прямоугольника}}{\text{прямоугольного параллелепипеда}}$, а d — длина его диагонали, т.е. сумма квадратов всех измерений $\frac{\text{прямоугольника}}{\text{прямоугольного параллелепипеда}}$ равна квадрату его диагонали.

Логические операции *обобщения*, *ограничения*, а также *аналогия* являются необходимым «мыслительным инструментарием» любого ученого. Владение этими операциями может пригодиться каждому человеку для переноса информации с одного объекта на другой. Особенно важны эти приемы для учащихся, так как помогают приобретать знания с меньшими затратами, используя не только память, но и логику. Мы применяем аналогию для сжатия представленной информации в тех случаях, когда записываем аналогичные определения в виде дроби.

Прислушаемся к Лапласу, который утверждал: «В самой математике главные средства достигнуть истины — индукция и аналогия».

3.4. Операции над понятиями. Определение понятий

> Определите значение слов, и вы избавите человечество от половины его заблуждений.
>
> *Р. Декарт*

Язык, на котором мы говорим, многозначен. У большинства слов в нем есть несколько возможных значений. В повседневной речи мы понимаем из контекста, какой смысл придавался тому или иному понятию. Но для строгих математических выводов надо *определить* понятие, дать точное *определение* значения терминов, используемых в науке, через набор *базисных*, неопределяемых, но всем понятных (очевидных) или ранее определенных слов.

? Почему необходимо научиться правильно определять понятия? Насколько важно давать правильные определения понятиям?

В первую очередь, определения нужно дать вновь вводимым понятиям, содержание которых ранее было неизвестно. С помощью определений фиксируется смысл научного термина. Так, например, в гл. 1 введены операции над множествами.

Во-вторых, определения необходимы для тех понятий, которые употребляются в новом значении. Например, в гл. 2 были определены понятия графа, вершины, ребра, так как ранее их употребляли в ином, бытовом, а не математическом смысле. Термины «сравнение», «масса», «вес», «соль», «угол», «рациональное» (число), «действие» и др. в научном понимании отличаются от бытового употребления.

В-третьих, точное определение понятий имеет большое значение в науке вообще, а в математике — в особенности, так как точность формулировок приучает к конкретности и однозначности мыслей. Точные формулировки облегчают и упрощают процесс доказательств не только в математике.

Определения терминов нужны еще и потому, что в некоторых случаях они в обычной речи не встречаются, например тригонометрия, логарифм, валентность и т.д. Иногда в разных науках один и тот же термин имеет разное значение, например «медиана» в геометрии и математической статистике.

Заметим, что, многократно употребляя понятие «информация», мы ни разу его не определили. В информатике под *информацией* понимают содержание некоторого сообщения или сигнала как совокупность сведений об определенной физической системе, например сообщение о качестве и количестве изготовленной продукции, экологическом состоянии окружающей среды и т.д.

Информация, как и другие фундаментальные понятия современной науки — вещество, энергия, является основой, сутью нашей действительности. В силу своей общности и доступности термин «информация» в науке является неопределяемым. Однако в математике, с вероятностной точки зрения, под *информацией* принято понимать совокупность сведений, уменьшающих неопределенность в выборе различных возможностей.

Определения нужны для того, чтобы наша речь и мысли были восприняты однозначно, чтобы не было двусмысленностей в их толковании.

Возможны различные виды определения понятий. Для того чтобы понимать смысл терминов, важно знать способы их определения. Хотя в логике нет однозначной характеристики различных видов определений, мы представим наиболее существенные из них в виде графа. Заметим, что понятие «определение» произошло от слова «делить», т.е. устанавливать *предел*, некоторую *границу*. Для этого нужно указать *родовое понятие* — все *необходимые* признаки этого понятия, а затем ограничить *видовыми — достаточными* для вычленения этого понятия из других, принадлежащих этому роду.

Итак, **определение** понятия — это логическая операция, с помощью которой раскрывается его содержание.

Представим виды определений в виде графа, не претендуя на исчерпывающую полноту (рис. 3.8).

В отдельных случаях для раскрытия многозначности понятий используют *контекстуальные* определения. Обычно для этого берут отрывок текста такой длины, в котором встретилось исследуемое понятие и из которого ясно о чем идет речь.

Еще нам понадобятся также *индуктивные* и *аксиоматические* определения. Индуктивные определения обязательно содержат **базисные** пункты, определяющие объекты, которые именуют определяемым термином, и **индуктивные** пункты, которые задают правила образования новых объектов, также именуемых этим термином. Примерами индуктивных определений могут служить определения *терма* и *формулы* в подразд. 5.1.



Рис. 3.8. Классификация видов определения понятий

В отличие от явных прямых определений неявные, косвенные, т. е. описательные, определения называют **дескриптивными** (от лат. *descriptio* — описание). Например, к дескриптивным определениям относится определение: «арифметической прогрессией называется последовательность чисел, в которой разность двух соседних чисел имеет одно и то же значение».

Самый распространенный способ определения понятий — классический — через ближайший род и видовые отличия. Род (родовой признак в дальнейшем будем обозначать B) указывает на тот круг предметов и понятий, из числа которых надо выделить опре-

деляемое понятие. Видовыми отличиями (далее — C) называются признаки, существенные для данного понятия.

Известно, что для описания множества достаточно указать некое характеристическое свойство, которым обладают все элементы этого множества, и только они. Поэтому мы не будем далее делать буквенного различия между множеством и его характеристическим признаком. Поскольку вид A — это множество всех элементов, имеющих видовой признак C (т.е. принадлежащих соответствующему множеству C) из всего рода B (т.е. имеющих признак, характерный для множества B), то A состоит из тех и только тех элементов, что принадлежат как множеству B, так и множеству C, т.е. A — пересечение (произведение) множеств B и C, $A = B \cap C$.

Очевидно, что $A \subset B$. Приведем примеры.

Квадратом (A) называется прямоугольник (B) с равными сторонами (C).

Окружностью (A) называется множество точек плоскости (B), равноудаленных от заданной точки M этой плоскости (C).

Информатикой (A) называется наука (B), изучающая технологию сбора, хранения и переработки информации (C).

Правила определений. Возможные ошибки. Примеры определения математических понятий

Правило 1. *Определение должно быть соразмерным*: $V_A = V_{B \cap C}$, т.е. объем определяемого понятия A равен объему определяющих понятий $B \cap C$.

Рассмотрим некоторые примеры.

«Алгоритмом называется точное формальное предписание, однозначно определяющее содержание и последовательность операций, переводящих заданную совокупность исходных данных в искомый результат». Это генетическое определение.

«Арифметикой (A) называется раздел математики, изучающий числа и простейшие действия над ними». Это классическое определение через род (B — «раздел математики») и видовое отличие (C — «изучающий числа и простейшие действия над ними»).

В разных разделах дискретной математики используется понятие «базис». «**Базисом** называется множество элементов, порождающих все математические объекты заданного вида с помощью определенных операций». Это пример генетического определения понятия.

«Дугой графа называется упорядоченная пара связанных вершин графа». Это номинальное определение (введение термина).

«Дифференцирование — математическая операция нахождения производной или дифференциала данной функции». Это пример генетического определения.

«Бит (A) — это единица количества информации (B), численно равная количеству информации в испытании с двумя равновероятными взаимно-исключающими исходами (C)». Это пример классического определения, где «бит» — определяемое понятие A, родовое, определяющее понятие «единица количества информации» — B, которая обладает характерной особенностью, т.е. видовым отличием C, — «численно равна количеству информации в испытании с двумя равновероятными взаимно исключающими исходами».

Рассмотрим возможные ошибки.

Широкое определение $V_A < V_{B \cap C}$, где B — ближайшее к A родовое понятие.

Определения «Окружностью называется фигура, образованная движущимся концом циркуля» или «Ромбом называется четырехугольник, у которого все стороны равны» являются примерами ошибочных конструкций. Так, четырехугольник — широкое, отдаленное понятие, не ближайший род.

Известен исторический пример такой ошибки. Однажды древнегреческий философ Платон дал определение понятию «человек»: «Человек — это двуногое животное без перьев». Другой древнегреческий философ Диоген принес на лекцию ощипанного петуха со словами: «Вот человек Платона». Признав свою ошибку, Платон «уточнил»: «Человек — это двуногое животное без перьев и с широкими ногтями».

Узкое определение: $V_A > V_{B \cap C}$.

Например, «Вершина — самая высокая часть холма» (но вершина есть, например, у многоугольника, графа, горы, параболы); «Совесть — это осознание человеком ответственности перед самим собой за свои действия и поступки» (есть еще ответственность перед обществом, близкими).

Определение в одном смысле широкое, в другом — узкое.

Например, «Ящик — тара для хранения овощей». Но овощи хранят не только в ящиках, а и в холодильнике, в мешках. Поэтому это определение узкое. В ящиках хранят не только овощи, но и стройматериалы, почту и т.д. Значит, данное определение одновременно и широкое.

Правило 2. *Определение не должно заключать в себе круга.*

Это происходит, когда определяемое и определяющее понятия отражают одно и то же.

Определение должно быть *логически корректным.* Нельзя «изобретать изобретение», «играть в игру», «итожить итоги». Например, ошибочными являются определения: «Параллельными называются прямые, расположенные параллельно», «Смешное — это то, что вызывает смех», «Символика — совокупность применяемых в математике символов и правил их употребления». Тавтологией, «кругом» может служить фраза, заменяющая определяемое поня-

тие, но не вносящая новых сведений о нем. Классический пример такой ошибки — в пьесе Мольера «Мнимый больной», где один из героев объяснял усыпляющее действие опиума: «Опиум имеет усыпляющую силу, так как он усыпляет, а усыпляет он потому, что имеет усыпляющую силу».

Правило 3. *Определение положительного понятия не должно содержать отрицания.*

Так, «Окружностью называется фигура, не имеющая углов», «Сравнение — не доказательство» — ошибочные определения.

Отрицание не может указать существенные признаки определяемого понятия, а лишь исключает несущественные.

Однако в отдельных случаях определение через отрицание допустимо, поскольку определяются отрицательные понятия через независимо определенные положительные. Например в определении инертных газов (аргона, неона и др.) указывается их химическая неактивность. Отрицательные понятия также можно определить, используя приставки отрицания *дис-*, *а-*, *без-*, *де-*, *ир-*, *не-* и др.

Рассмотрим примеры.

«Иррациональными числами называются числа, не соизмеримые ни с единицей, ни с ее частями.»

«Бездарным является человек, не обладающий способностями.»

«Аморальными называются поступки, не соответствующие моральным нормам общества.»

«Нелогичными называются поступки, не соответствующие нормам логики мышления и поведения представителей данного сообщества.»

Правило 4. *Определение должно быть четким и ясным, без двусмысленностей.*

«Природа — это наука, способствующая пониманию вопросов, относящихся к духовной истине» — пример неправильного определения.

Правило 5. *Определение неизвестного понятия должно производиться через известные.*

Соответственно, ошибкой будет определение неизвестного понятия через неизвестные. Например, «Дивергент — разновидность аллофона».

Правило 6. *В науке недопустима подмена определения метафорами и сравнениями.*

Она возможна лишь в художественной литературе и речи как образный прием.

Например, «Лень — мать всех пороков», «Такт — это разум сердца» (К. Гуцков), «Неблагодарность — род слабости» (И. В. Гёте), «Повторение — мать учения» и др. Эти поговорки не являются определениями или аксиомами, но могут являться теоремами, устанавливающими неочевидную эквивалентность понятий. Подобные утверждения называются предикатами (см. гл. 4 и 5).

Приемы, заменяющие определения. Такие приемы допустимы в художественной и научно-популярной литературе, в повседневной речи.

Сравнение используется для образной характеристики предмета через метафоры. Например, «Архитектура — окаменевшая музыка», «Лев — царь зверей», «Книги — величественные маяки в океане времени» (Ф. Ницше), «Слово — полководец человеческой мысли» (В. Маяковский).

Описания используются в литературе, следственной практике, науке для наиболее точного указания признаков предмета. В научной литературе обязательно описание эксперимента. Возможно описание процесса (общественного), физического явления либо химической реакции.

Характеристика используется для указания отличительных признаков предмета: как единичного (характеристика студента Петрова), так и общего явления: «признаками волевых действий человека является преодоление им препятствий».

В определении, в отличие от приемов, заменяющих определения, всегда указаны такие общие и существенные признаки предмета и явления, каждый из которых в отдельности *необходим*, а все вместе они *достаточны* для выделения этого предмета или явления из множества аналогичных. Определение понятий выполняет две важные функции:

- в нем закреплены наиболее важные результаты процесса познания;
- оно является основой для понимания предмета или явления.

3.5. Деление понятий. Классификация

> Хорошая классификация важна; она сводит разнообразие объектов наблюдения к относительно небольшому числу ясно характеризуемых и хорошо упорядоченных типов.
>
> *Дж. Пойа*

Делением понятия называется логическая операция, раскрывающая его объем. К делению понятия прибегают тогда, когда необходимо:

- раскрыть весь объем понятия, т. е. выявить все компоненты, составляющие это понятие (так, рассматривая множества, мы делили их на виды в зависимости от мощности);
- указать все формы проявления и развития понятия, а не только сущность самого явления или предмета, составляющего понятие;

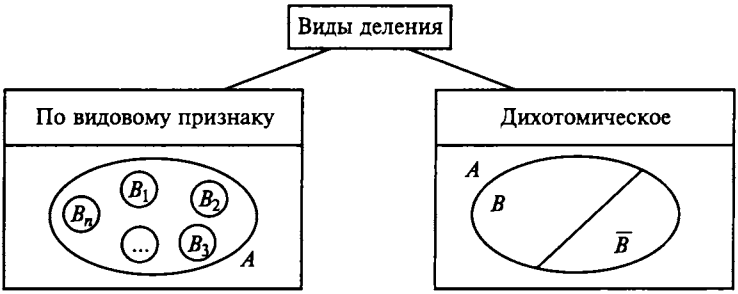


Рис. 3.9. Классификация видов деления понятий

• уточнить сферу применения понятия (так, отношения между понятиями различаются по сфере их применения в зависимости от взаимосвязей между отдельными составляющими).

К делению прибегают, когда понятие многозначно и требуется уточнить его отдельные значения. Ранее мы уточняли новое понятие «определение понятий», разделив его на виды.

При делении необходимо выявить делимое понятие, основание деления, члены деления.

Правила и ошибки при делении понятий представлены в табл. 3.1.

При делении по **видовому признаку** (рис. 3.9) объем делимого (родового) понятия A раскрывается через перечисление *всех* его видов $B_1, B_2, ..., B_n$ по выбранному признаку — основанию деления. В основе деления лежит *изменение* этого признака. Члены деления — подмножества, соподчиненные группы (виды), A — делимое (родовое) понятие — все множество.

При **дихотомическом делении** (от греч. διχα — две части, τομη — сечение) объем делимого понятия A делится на два противоречивых B и $\overline{B}$ (см. рис. 3.9). В основе лежит наличие или отсутствие признака, характеристического для B или $\overline{B}$. Таким образом, дихотомия также является делением по видовому признаку при $n = 2$.

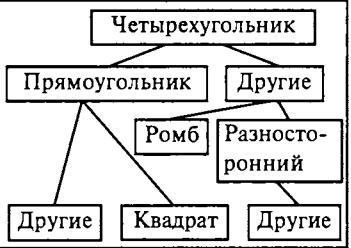
Графически схема деления понятия может быть представлена или кругами Эйлера, или в виде графа-дерева. Основанием (корнем) такого дерева является само понятие A. Ветви дерева — члены деления.

Принцип дихотомического деления широко применяется в моделировании и конструировании ЭВМ, так как электронно-вычислительные машины «понимают» язык алгебры логики (булевой алгебры), построенный на основе дихотомического деления классов.

С помощью дихотомического деления можно описать множество процессов реальной жизни: принадлежит — не принадлежит, существует — не существует, моральный — аморальный, нравственный — безнравственный, рациональный — иррациональный.

Таблица 3.1

Правила и ошибки при делении понятий

Правила	Ошибки	Примеры ошибок	Примеры правильного деления
Деление должно быть соразмерным: $V_{рода} = \sum V_{видов}$ $V_B = \sum_i V_{BC_i}$	Неполное деление	Личные документы: паспорт, автобиография, страховой полис, аттестат об образовании, водительские права	Личные документы: паспорт, автобиография, страховой полис, аттестат об образовании, водительские права и т. д.
	Деление с лишними членами (в основе — разные признаки деления)	Углы бывают прямые, острые, тупые, вертикальные, накрест лежащие	Углы бывают прямые, острые, тупые (признак — величина угла)
Деление должно производиться лишь по одному признаку	При делении одного понятия одновременно в основу ложатся несколько признаков	Треугольники: равнобедренные; равносторонние; прямоугольные	По величине сторон треугольники: равнобедренные; равносторонние; разносторонние. По величине углов треугольники: прямоугольные; тупоугольные; остроугольные
Члены деления должны исключать друг друга	Члены деления — пересекающиеся, а не соподчиненные понятия	Четырехугольники: параллелограмм; прямоугольник; ромб; трапеция	
Деление непрерывно, без скачков	От родового к ближайшему видовому понятию	Члены предложения: подлежащее; сказуемое; второстепенные	

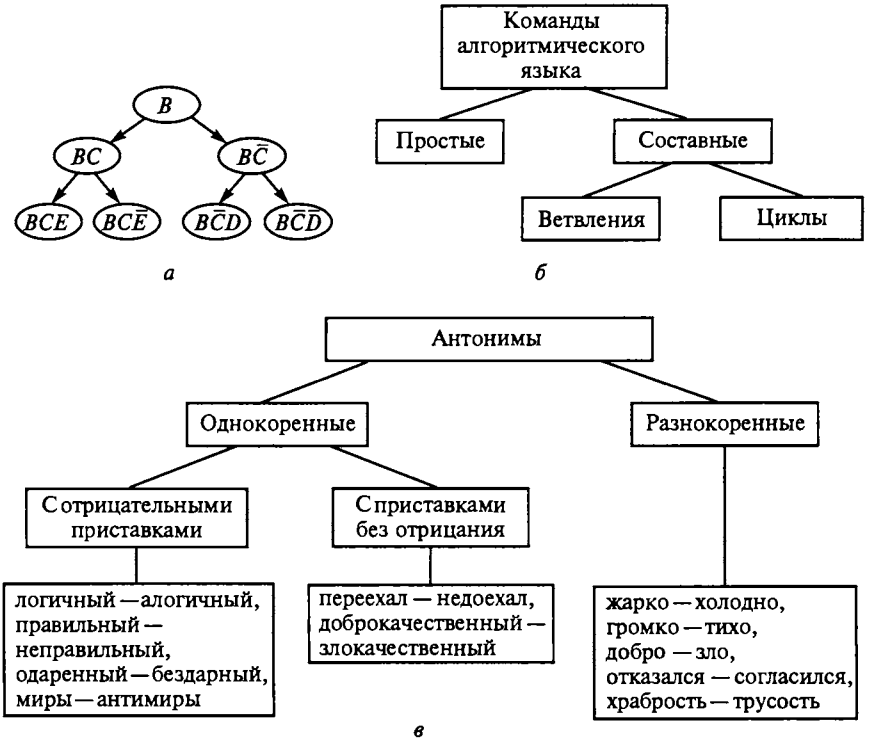


Рис. 3.10. Схема дихотомической классификации:

а — бинарное дерево; *б* — команды алгоритмического языка; *в* — понятие «антонимы»

Классификация. Частным случаем деления понятий является классификация. Она как разновидность операции деления представляет собой последовательное деление элементов множества на подмножества, в которых все элементы обладают своей уникальной особенностью, т. е. относительно устойчивыми характеристиками. Схема дихотомической классификации имеет вид бинарного дерева (рис 3.10, *а*).

Например, в информатике осуществляется классификация команд алгоритмического языка (рис. 3.10, *б*).

Умение человека обобщать, ограничивать и классифицировать понятия характеризует системное мышление.

Итак, **классификацией** называется распределение объектов по определенному существенному признаку, в результате чего каждый из них попадает в точно указанный класс, подмножество или группу.

Классификация бывает *естественная*, *искусственная*, *рабочая*, *научная*. В науке принято классифицировать объекты, начиная от

общего понятия — всего множества, и рассматривая его подмножества согласно иерархии. Тогда от общего понятия к частному можно выстроить цепочку видовых понятий: классы → подклассы → → семейства → отряды → роды → виды или {виды} ⊂ {роды} ⊂ ⊂ {отряды} ⊂ {семейства} ⊂ {подклассы} ⊂ {классы} ⊂ {общие понятия (типы)}.

Со школьной скамьи известны примеры научных классификаций: Периодическая система элементов, предложенная Д. И. Менделеевым, или классификация животных и растений К. Линнея. Научная классификация имеет огромное значение, как теоретическое, так и практическое. Она не только сохраняет знания, накопленную информацию об окружающем мире, но и дополняет ее новыми, ранее неизвестными элементами. Например, согласно старой классификации элементарных частиц в настоящее время известно свыше 200 частиц. Это не могло удовлетворить физиков и привело к переопределению понятия «элементарная частица», поиску новых теорий, минимизирующих их число, открытию кварков и глюонов.

С помощью научной классификации упорядочивается огромный, накопленный, быть может, многими поколениями исследователей, научный материал, который систематизируется и распределяется по устойчивым множествам (классам, родам, видам).

Другим важным достоинством является возможность выявить существенные отличия между отдельными подмножествами (в рамках известного сходства общего свойства выявить характерные различия). Это не только помогает различать отдельные подмножества и видеть их специфические особенности, но и открывает новые возможности для познания объективных закономерностей.

Примерами искусственной классификации являются каталоги книг, список вашей группы по алфавиту в журнале, список телефонных номеров.

Рассмотрим примеры классификаций из различных областей знаний.

Понятие «антонимы» встречается не только на уроках русского языка. Оно широко используется при построении дедуктивных умозаключений для формирования отрицания и будет часто применяться при изучении дискретной математики и математической статистики (рис. 3.10, *в*).

Изучая дискретную математику, мы одновременно будем учиться различным способам *систематизации* учебного материала. Они помогут увидеть весь вопрос в целом и во взаимосвязи этого вопроса с его частями и научиться сжимать информацию об изучаемом вопросе, кратко ее излагать.

Существуют различные приемы сжатия учебного материала: блок-схемы, графы, деревья, а также таблицы или матрицы. Каждый элемент матрицы — клетка — содержит информацию, соот-

ветствующую пересечению строки и столбца. Поэтому важно тщательно продумать направляющие элементы строк и столбцов. Примером такой матрицы может служить табл. 3.2. Другие примеры, а также более подробное обсуждение происхождения научных знаний приведены в книге Ю. В. Пахомова «Логика естествознания».

? Какое практическое значение имеют отношения между понятиями, операции над ними?

Оказывается, от отношений между понятиями зависит, например, правописание сложных слов.

Если сложное слово объединяет соподчиняющиеся понятия, то они как равноправные по степени обобщения пишутся через дефис: северо-западный циклон, научно-популярная литература.

Если сложное слово объединяет понятия, находящиеся в отношении подчинения относительно друг друга как родовое и видовое понятия (смысловая зависимость), то такое сложное слово пишется слитно: Среднерусская возвышенность, западноевропейское искусство, железнодорожный вокзал.

Важно знать различные определения одного понятия, выявлять тождественные понятия.

При употреблении противоположных или противоречивых понятий необходимо учитывать различие между ними. Это особенно актуально в математической логике при построении доказательств, например, методом от противного, при решении большого круга задач, содержащих отрицание. Так, для понятия «не менее» противоречивым будет «менее», а противоположным — «не более». Но понятия «не более» и «менее» не тождественны между собой: они имеют разный объем.

Операции над понятиями имеют особое значение для тех, кто связал свою профессиональную деятельность с компьютером в различных сферах его применения. При составлении программ, особенно в тех случаях, когда прибегают к методу перебора, зачастую необходимо увидеть реальный объем некоторого понятия. В таких случаях нужно произвести операцию деления понятия.

Исторический пример первой машины, осуществлявшей перебор всех вариантов словосочетаний, связан с именем рыцаря-монаха Раймунда Луллия, который в XIII в. сконструировал машину для механического доказательства всех основных положений христианской религии. Такое изобретение понадобилось Луллию для приобщения «диких» мавров (Северная Африка) к истинной вере. В пятой главе знаменитой книги «Путешествие Гулливера» сатирик Дж. Свифт едко высмеял ученого, прообразом которого являлся Луллий. Однако мечты Луллия о механическом (автоматическом) доказательстве, о машине, способной заменить рассуждения человека, оказались не такими уж бесплодными и фантастическими. Фактически его механическая машина была прообразом современных ЭВМ. Понятия, включенные в его машину, были глубоко

Таблица 3.2

Происхождение научных знаний

Компоненты научного предмета	Объекты		Операции над объектами	Научные законы	
Научные дисциплины	основные	производные		основные	производные
Геометрия	Точка, прямая	Другие геометрические фигуры	Выделение части из целого, разложение целого на части, синтез целого из частей и др.	Аксиомы планиметрии	Теоремы планиметрии
Арифметика натуральных чисел	Единица	Натуральные числа	Сложение, вычитание, умножение, деление	Таблицы сложения в пределах до 10, законы арифметических действий	Алгоритмы действий (+, –, ×, :) над многозначными числами
Классическая механика	Пары взаимодействующих или взаимно перемещающихся тел	Более сложные механизмы	Разложение механических систем на более простые и синтез простых систем в более сложные	2-й закон Ньютона, 3-й закон Ньютона, закон всемирного тяготения	1-й закон Ньютона, закон свободного падения, закон движения маятника
Неорганическая химия	Атомы химических элементов	Молекулы и атомные группы	Разложение молекул на атомы, синтез молекул и атомных групп из атомов	Нулевая валентность молекул, сохранение общего количества атомов при их перекомбинировании в новые молекулы в ходе реакции	Закон сохранения массы в ходе химических реакций, постоянство весовых отношений реагирующих веществ и продуктов реагирования

религиозны. Они располагались на вращающихся дисках и в процессе движения образовывали различные новые словосочетания. Понятия, использованные Луллием в своей машине, дошли до нашего поколения сквозь века и эпохи.

? Какие десять «вечных» понятий могли бы вы оставить будущему поколению или, например, инопланетянам?

Упражнения

3.1. Определите, правильно ли произведена операция ограничения понятий (изобразите их кругами Эйлера):

а) персональный компьютер — монитор;
б) системный блок — дисковод;
в) Web-страничка — Internet;
г) язык программирования — Си+;
д) Си+ — Borland Pascal;
е) мышка — коврик.

3.2. Найдите общие понятия для следующих пар понятий (изобразите их кругами Эйлера):

а) монитор, системный блок — это...;
б) Pascal, Си+ — это...;
в) принтер, сканер — это...;
г) Windows, Unix, OS/2 — это...;
д) постановка задачи, составление алгоритма — это...;
е) клавиатура, мышь — это....

3.3. Исключите одно из понятий ряда так, чтобы оставшиеся понятия объединялись в некоторый ряд. Укажите этот ряд и изобразите кругами Эйлера:

а) дисковод, процессор, жесткий диск, принтер, звуковая карта;
б) сложение, вычитание, деление, переменная, умножение;
в) Pascal, Си+, Basic, Unix;
г) абак, арифмометр, ЭВМ, магнитофон, ПК;
д) прямоугольник, окружность, треугольник, куб, трапеция;
е) шахматы, нарды, шашки, баскетбол.

3.4. Правильно ли произведено обобщение понятий? Изобразите кругами Эйлера их отношения:

а) программный язык — английский язык;
б) микропроцессор — главная часть компьютера;
в) алгоритм — строгая определенность последовательных действий над задачей;
г) код — шифрованная информация;
д) кибернетика — наука об искусственном интеллекте;
е) Windows — операционная система.

3.5. Какая произведена операция — обобщения или ограничения:

а) частичное решение — общее решение;

б) число, делящееся на 6, — четное число;

в) деление — деление на 10;

г) системный блок — персональный компьютер;

д) алгоритм решения — постановка задачи;

е) несчетные множества — бесконечные множества?

3.6. Расположите понятия в порядке уменьшения объема (изобразите их кругами Эйлера):

а) программный язык, процедуры, библиотеки, Pascal;

б) системный блок, микропроцессор, материнская плата, микросхемы;

в) Тольятти, европейский город, город Самарской области, город России;

г) байт, килобайт, бит, мегабайт, гигабайт;

д) грамм, центнер, тонна, миллиграмм, килограмм;

е) контейнер, упаковка, пачка.

3.7. 1. Подберите понятия, находящиеся с данными в отношении:

а) тождественности; б) пересечения; в) подчинения.

Изобразите кругами Эйлера и впишите в табл. 3.3.

Таблица 3.3

Таблица отношений

Понятие	Отношения		
	тождественные	пересекающиеся	подчинения
Микропроцессор			
Алгоритм			
Натуральное число			
Компьютер			
Информатика			

2. Подберите понятия, находящиеся с данными в отношении:

а) соподчинения; б) противоположности; в) противоречия.

Изобразите в виде кругов Эйлера и впишите в табл. 3.4.

3.8. Определите, правильно ли изображены кругами Эйлера отношения между понятиями на рис. 3.11.

Если схема ошибочна, нарисуйте правильную:

а) А — время, Б — минута, В — секунда, Г — час, Д — сутки (рис. 3.11, *а*);

Таблица 3.4

Таблица отношений

Понятие	Отношения		
	соподчинения	противоположности	противоречия
Команда			
ЭВМ			
Сложение			
Положительное число			
Натуральное число			

б) А — системный блок, Б — монитор, В — процессор, Г — принтер (рис. 3.11, *б*);

в) А — бит, Б — байт, В — килобайт, Г — мегабайт (рис. 3.11, *в*);

г) А — компьютер, Б — Pentium 333, В — Pentium 333/32 Mb, Г — Pentium 333/32 Mb/4,3 Gb (рис. 3.11, *г*);

д) А — абстрактное, Б — конкретное, В — положительное, Г — понятия (рис. 3.11, *д*).

3.9. Подберите понятия, отношения между которыми соответствуют приведенным на рис. 3.12 кругам Эйлера.

3.10. Определите, правильно ли проведено ограничение понятий. Если допущена ошибка — исправьте (изобразите их кругами Эйлера):

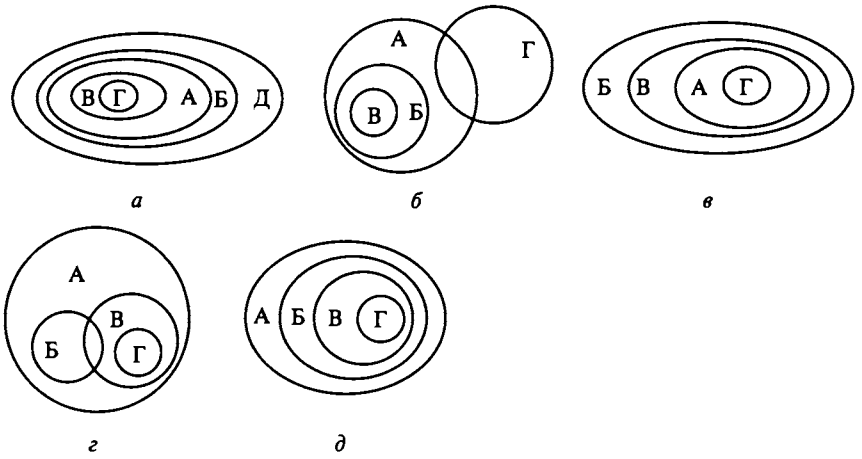


Рис. 3.11. Схемы к упр. 3.8 (*а*—*д* — варианты)

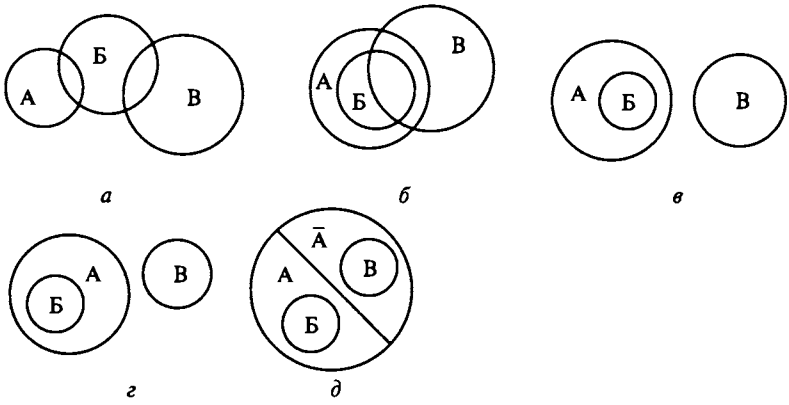


Рис. 3.12. Схемы к упр. 3.9 (*а*—*д* — варианты)

а) время, час, минута, секунда;
б) программный язык, Pascal, Basic;
в) бит, байт, килобайт;
г) системный блок, материнская плата, процессор;
д) игра, спортивная игра, футбол;
е) граф, плоский граф, дерево.

3.11. Определите, правильно ли проведено обобщение понятий. Если допущена ошибка — исправьте (изобразите их кругами Эйлера):

а) группа, факультет, колледж, учебное заведение;
б) системный блок, материнская плата, процессор;
в) DOS, командная строка, команда, символ;
г) правильная дробь, десятичная дробь, дробь;
д) алгоритм, определение входных и выходных данных, математическая модель;
е) множество, конечное множество, пустое множество.

3.12. Определите, в каких из приведенных примеров производится деление понятия, а в каких — деление предмета:

а) запись алгоритма бывает словесной и графической;
б) граф состоит из ребер и вершин;
в) графы делятся на ориентированные и неориентированные;
г) дискеты бывают 3- или 5-дюймовые;
д) CD-ROM бывают пишущие и непишущие;
е) система бывает: линейная, иерархическая, сотовая, кольцеобразная.

Глава 4
МАТЕМАТИЧЕСКАЯ ЛОГИКА

В этой главе мы познакомимся с элементами математической логики. В ее основе лежит логика высказываний, в которой высказывание рассматривается как особое буквенное исчисление — **алгебра логики**. Математическая логика изучает схемы (формы) истинных высказываний, имеющих наибольшую степень общности, схемы математических доказательств и правила их вывода. Изучение исчисления высказываний как алгебраической системы составляет предмет алгебры логики, или *булевой* алгебры. Мы освоим язык алгебры логики, ее законы, научимся строить и упрощать булевы функции, выполнять операции над сложными высказываниями, а также узнаем, как язык алгебры логики применяется в процессе рассуждений.

4.1. Суждение как форма мышления. Простые высказывания

> Я знаю только то, что я ничего не знаю, другие не знают и этого.
>
> *Сократ*

Суждением называется форма мышления, в которой что-либо утверждается или отрицается о *существовании* предмета, *связях* между предметом и его свойствами или об *отношениях* между предметами.

Словам естественного языка, например русского, в логике соответствуют понятия. Слова объединяются в предложения. По интонации предложения делятся на вопросительные, восклицательные и повествовательные. Но информацию несут только повествовательные предложения. Таким предложениям в логике соответствуют суждения. Они выражают наши знания о связях между понятиями. Суждение характеризуют две стороны: его форма и его истинность. Отвлечемся от формы и рассмотрим вопрос о том, *истинно* оно или *ложно*.

Всякое суждение, утверждающее что-либо о чем-либо, называют **высказыванием**, если можно сказать, истинно оно или ложно в данных условиях места и времени. Напомним, что характеристика суждения по признаку истинности-ложности называется **семантической**. Рассмотрим примеры повествовательных предложений.

1. Умение грамотно использовать логические операции повышает эффективность программирования.

2. История логики насчитывает около двух с половиной тысячелетий.

3. Знание математической логики необходимо любому специалисту.

4. Математическая логика — увлекательная наука.

5. $x > 5$.

6. Была метель.

7. Он — программист.

Предложения 1, 2, 3 являются высказываниями, а 4 и 5 — нет. Однозначно для всех людей определить отношение к науке невозможно, так же как невозможно определить истинность неравенства $x > 5$, не зная значений переменной x, входящей в него. Не является высказыванием и предложение 6, так как нет достаточной информации, чтобы установить, истинно оно или ложно (где и когда?). Однако предложение $x > 5$ может стать высказыванием, если будут известны конкретные значения переменной x. Так, если задать множество значений x: $x \in \{0, 2, 5, 7, 12\}$, то высказывания $0 > 5$, $2 > 5$, $5 > 5$ будут ложными, а $7 > 5$, $12 > 5$ — истинными. Истинными будут высказывания 1 и 2, ложным — третье, так как есть множество профессий, для которых знания математической логики не обязательны. Предложение «Он — программист» станет высказыванием при подстановке вместо местоимения «он» имени отдельного человека. Используя кванторные слова «всякий», «некоторый», «есть» и др., тоже можно получить высказывания. Например, «Всякий человек есть программист» — ложное высказывание. Заметим, что предложения вида 5 и 7 подробно изучает логика предикатов, с которой мы познакомимся в гл. 5.

Формализацией высказываний называют операцию замены высказывания естественного языка формулой математического языка, включающего высказывательные переменные и символы тех логических операций, которые соответствуют структуре самого высказывания.

Предложения, употребляемые в повседневной речи, приобретают вид формулы, если каждую смысловую составляющую проверять на предмет истина-ложь. Сам же результат в зависимости от того, реализовалось событие или нет, также имеет дихотомическую основу (двоичный исход). Поэтому фактически речь идет о функциях над множеством {истина, ложь}. Но из гл. 1 известно,

что все равномощные множества эквивалентны. Более того, если существует бинарная операция на них, сохраняющая свой вид при переходе от одного множества к другому, то эти множества изоморфны. Поскольку мы стремимся придать высказываниям вид исчислений и создать алгебру высказываний, то уместно развить аппарат на более понятном (например, числовом) множестве. Затем, построив биекцию на семантическое множество {истина, ложь}, получим алгебру логики.

Таким образом, мы должны выбрать числовое множество, где числовые символы будут соответствовать истине и лжи. Очевидно, что наиболее подходящим множеством будет {0, 1}. Во-первых, эти символы занимают мало места, во-вторых, удобна техническая реализация: есть сигнал — (1), нет сигнала — (0).

4.2. Булевы функции

> Мы употребляем знаки не только для того, чтобы передавать наши мысли другим людям, но и для того, чтобы облегчить сам процесс нашего мышления.
>
> *Г. В. Лейбниц*

Логические функции. Алгебра логики, выстроенная в XIX в., долго существовала как абстрактная, хотя и очень красивая наука. Но в середине XX в. оказалось, что она имеет конкретное и очень важное применение в современной жизни. Булева алгебра в настоящее время служит основой для описания логики работы аппаратных и программных средств ЭВМ. Дело в том, что алгебра логики использует логические переменные, которые принимают лишь два значения 0 и 1. Аналогично ЭВМ, используя лишь сигналы 0 и 1, воспринимает их как двоичные числа или логические переменные.

Итак, рассмотрим множество {0, 1}, которое будем обозначать буквой B. Отображение f: $B^n \to B$ назовем **булевой функцией** n переменных.

Областью определения булевой функции являются кортежи длиной n, состоящие из символов 0 и 1. При этом каждому варианту кортежа должен быть поставлен в соответствие единственный элемент из B — значение булевой функции, например $f(0, 1, 1, 0) = 0$. Поскольку в B всего два элемента, то снятие разделяющих запятых не нарушит однозначного восприятия кортежа, например $f(0, 1, 1, 0) = f(0110) = 0$. Аргумент булевой функции фактически представляет собой целое двоичное число (см. подразд. 6.2). По этому принципу соответствующие аргументы можно упорядо-

чить и при задании булевой функции располагать их в виде двоичных кортежей.

Очевидно, что всего у булевой функции n переменных может быть 2^n аргументов, соответствующих двоичным числам от $\underbrace{00\ldots0}_{n}$ до $\underbrace{11\ldots1}_{n}$ включительно. Итак, $|B^n| = |B|^n = 2^n$.

Определим, сколько может быть различных булевых функций n переменных. Из раздела 1.5 известно, что если $|A| = N$, а $|C| = k$, то существует k^N различных функций, действующих из A в C. Поскольку $A = B^n$, а $C = B$, то $N = 2^n$, а $k = 2$. В итоге заключаем, что всего существует $2^{(2^n)}$ различных булевых функций n переменных.

Для рассмотрения свойств булевых функций воспользуемся их геометрической интерпретацией. Поставим в соответствие различным наборам значений аргументов функции $f(x_1, x_2, \ldots, x_n)$ определенные точки n-мерного пространства. Тогда множеству 2^n двоичных наборов соответствует множество вершин n-мерного единичного куба. Припишем каждой вершине конкретное значение 0 или 1, которое принимает булева функция при соответствующем наборе аргументов. Если значение функции равно 0, то точка не рисуется («выкалывается»), если значение функции равно 1, то точка рисуется (выделяется полужирно). Такой n-мерный куб будет однозначно задавать любую булеву функцию $f(x_1, x_2, \ldots, x_n)$ с n переменными. Например, для двух переменных геометрическая интерпретация булевой функции будет на плоскости иметь вид квадрата или «двумерного куба» (рис. 4.1, *а*), а для трех переменных — куба в пространстве (рис. 4.1, *б*).

Равенство функций. Как для отображений вообще, для булевых функций справедливо определение равенства функций. С учетом области определения оно будет выглядеть так: две булевы функции f и g, зависящие от n переменных, называются **равными (эквивалентными)**, если для любого $x \in B^n$ справедливо соотношение

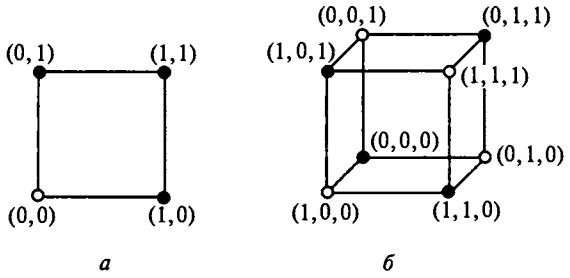


Рис. 4.1. Геометрическая интерпретация булевых функций:
$а - f(x_1, x_2)$; $б - f(x_1, x_2, x_3)$

$f(x) = g(x)$. Проверить эквивалентность двух функций можно с помощью таблиц истинности. Преобразование булевой функции F_1 с применением заданных эквивалентностей (формул) дает новую функцию F_2, эквивалентную исходной.

Формулы. Как известно, композицией функций $f_1, f_2, ..., f_n$ называется функция f, полученная с помощью подстановок этих функций друг в друга и переименования переменных. Тогда **формулой** называется выражение, *соответствующее* этой композиции, описывающее ее. Всякая формула, выражающая некоторую функцию f как суперпозицию других функций (подфункций), задает способ ее вычисления в том случае, если известны способы вычислений этих подфункций. Каждому набору значений аргументов формула ставит в соответствие значение функции, т. е. может служить способом задания и вычисления логической функции. Говорят, что формула *представляет* или *реализует* соответствующую логическую функцию. Практически **композицией** двух булевых функций f (от n переменных) и g (от m переменных) будет функция $h(x)$, такая, что $\forall x \in B^n\ h(x) = (g(f_1, ..., f_m)(x) = g(f_1(x), ..., f_m(x))$.

В процессе формализации учтем следующие правила.

• Простому высказыванию соответствует элементарная формула.

• Для построения формулы, соответствующей составному высказыванию, надо:

выделить все элементарные высказывания и логические связки, соответствующие структуре высказывания;

заменить их соответствующими символами;

расставить скобки в соответствии с логической структурой высказывания.

Для упрощения операций в логике высказываний принято не заключать в скобки формулы, стоящие под знаком отрицания и не являющиеся частями других формул.

Пусть дано множество (конечное или бесконечное) некоторых аргументов $(x_1, ..., x_n, ...)$ и исходных функций $\{f_1, f_2, ..., f_m\}$, зависящих от них. Формулы могут состоять из самих аргументов $x_1, ..., x_n, ...$ или включать в себя другие функции (или **подформулы**). Тогда, например, $f_1(x_1, x_2, x_3)$ — булева функция или формула, непосредственно зависящая от аргументов x_i $(i = 1, 2, 3)$, а $f_1(f_2(x_1, x_3); f(x_2); f(x_1))$ — сложная функция, образованная суперпозицией функций f (одного переменного), f_1 (трех переменных) и f_2 (двух переменных).

Всякая формула, выражающая функцию f как суперпозицию других (исходных) функций, одновременно задает и способ ее вычисления. Сначала вычисляются значения подформул, а затем сама формула, согласно установленному порядку действий.

Булевы функции одной переменной. При $n = 1$, согласно общей формуле, существуют четыре различные булевы функции. Поскольку при упорядоченных (по принципу целого двоичного числа) аргу-

Таблица 4.1

Виды булевых функций одной переменной

x	$f_1 = 0$ — «тождественный нуль»	$f_2(x) = x$ — «тождественная функция»	$f_3(x) = \bar{x}$ — «отрицание, инверсия»	$f_4 = 1$ — «тождественная единица»
0	0	0	1	1
1	0	1	0	1

ментах $x_1, x_2, \ldots, x_{2^n}$ строки значений $f(x_1), f(x_2), \ldots, f(x_{2^n})$ для разных функций также можно упорядочить по этому же принципу, то подобным способом упорядочиваются и сами функции (табл. 4.1). Далее можно вообще не выписывать значения в виде таблицы, а задавать их одной строкой. Например, запись $F = (11000110)$ расшифровывается следующим образом. Поскольку значений этой функции восемь, то функция F зависит от $\log_2 8 = 3$ переменных (например, x, y, z): $F(x = 0, y = 0, z = 0) = F(000) = 1$, $F(001) = 1$, $F(010) = 0$, $F(011) = 0$, $F(100) = 0$, $F(101) = 1$, $F(110) = 1$, $F(111) = 0$.

Если *любому* значению аргумента ставится в соответствие значение функции $\frac{0}{1}$, то подобное отображение называется **тождественным $\frac{\text{нулем}}{\text{единицей}}$**. Такое название справедливо для булевых функций произвольного числа переменных.

Если любому значению аргумента ставится в соответствие он сам, то такая логическая функция называется **тождественной**.

Если любому значению аргумента ставится в соответствие противоположный элемент, то такая функция называется **отрицанием** или **инверсией**. Поскольку значение инверсии зависит только от выражения, к которому она применена, то подобная операция является отношением, притом унарным. Не составит труда проверить, что отрицание отрицания есть тождественное отображение, т.е. $\forall x \in B\ \bar{\bar{x}} = \overline{(\bar{x})} = x$. Действительно, если $x = 0$, то $\bar{x} = 1$, а $\bar{\bar{x}} = 0$, аналогично при $x = 1$ $\bar{x} = 0$, а $\bar{\bar{x}} = 1$. Более удобно (особенно для громоздких формул) подобные действия, где операции применяются к промежуточным значениям формул, представлять в виде

Таблица 4.2

Таблица истинности булевых функций одной переменной

x	$y = \bar{x}$	$\bar{\bar{x}} = \bar{y}$	$f_2(x) = x$
0	1	0	0
1	0	1	1

таблицы, которую называют **таблицей истинности**. В данном случае она примет вид (табл. 4.2).

Булевы функции двух переменных. При $n = 2$ существуют 16 булевых функций. Их виды приведены в табл. 4.3.

Приведенные в табл. 4.3 функции можно разделить на три категории.

1. **Симметрические функции.** Функция $f_2 = (0001)$, которая равна 1, только если оба аргумента равны 1, называется **конъюнкцией** (от лат. *conjungere* — объединять) и обозначается как $f(x, y) = x \wedge y$ или $x \,\&\, y$. Однако чаще употребляется $f(x, y) = x \cdot y = xy$. В дальнейшем будем пользоваться именно последними обозначениями. Иногда удобно представление $xy = \min(x, y)$.

Функция $f_8 = (0111)$, значение которой равно 0, только если оба аргумента равны 0, называется **дизъюнкцией** (от лат. *disjunctio* — разделять). Обозначается как $f(x, y) = x \vee y$. Иногда удобно представление $xy = \max(x, y)$.

Функция $f_{10} = (1001)$, равная 1 только при совпадающих аргументах, называется **эквиваленцией** (от лат. *aequivalens* — равнозначный). Обозначается как $f(x, y) = (x \equiv y)$. Также может использоваться знак $\sim$.

Функция $f_7 = (0110)$, равная 0 только при совпадающих аргументах, называется **суммой по модулю два**. Обозначается как $f(x, y) = x \oplus y$. Также применяются знаки + и $\veebar$. Другое название — строгая дизъюнкция: значение функции равно 1, если либо первый, либо второй аргументы равны 1, но никак не оба одновременно. Следуя традиции, будем употреблять термин *сумма по модулю два* и знак $\oplus$, когда речь идет только о булевых функциях, и термин *строгая дизъюнкция* и знак $\veebar$ для высказываний, когда надо определить именно истинность, хотя фактически это одно и то же. Очевидно, что эта функция является отрицанием эквиваленции: $\forall (x, y) \in B^2 \; x \oplus y = \overline{x \equiv y}$.

Функция $f_9 = (1000)$, равная 1, только если оба аргумента равны 0, называется **стрелкой Пирса**. Обозначается как $f(x, y) = x \downarrow y$. Стрелка Пирса является отрицанием дизъюнкции.

Таблица 4.3

Таблица истинности булевых функций двух переменных

x_1	x_2	f_1 0	f_2 $\wedge$	f_3	f_4 x_1	f_5	f_6 x_2	f_7 $\oplus$	f_8 $\vee$	f_9 $\downarrow$	f_{10} $\equiv$	f_{11} $\overline{x_2}$	f_{12}	f_{13} $\overline{x_1}$	f_{14}	f_{15} $\mid$	f_{16} 1
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

Функция $f_{15} = (1110)$, равная 0, только если оба аргумента равны 1, называется **штрихом Шеффера**. Обозначается как $f(x, y) = x \mid y$. Штрих Шеффера является отрицанием конъюнкции.

Все перечисленные функции двух булевых переменных замечательны не только тем, что являются логическими операциями, сходными с операциями над множествами, но и тем, что они симметричны по своим аргументам. Запись соответствующей операции между символами переменных позволяет рассматривать ее как бинарное отношение, причем симметричное. Поэтому в табл. 4.3 указан только разделяющий знак, а порядок следования аргументов несуществен.

Общее правило. Поскольку подстановка $0 \to 1$, $1 \to 0$ в строке значений (т.е. инверсия значений функции) отображает двоичное число в дополнительное до 11...1, то любые две функции, расположенные симметрично относительно центральной вертикальной оси таблицы значений, будут взаимно инвертированными. Поэтому $f_k = \overline{f_{2^{2^n}+1-k}}$. Эта формула является простым переводом чисел из двоичной в десятичную систему записи. Например, штрих Шеффера $f_{15}(x, y) = x \mid y = \overline{xy} = \overline{f_2(x, y)}$.

2. **Импликации.** Функция $f_{14} = (1101)$, значение которой равно 0, только если первый аргумент равен 1, а второй 0, называется **импликацией**, или **следуемостью** (от лат. *implico* — тесно связывать). Обозначается как $f(x, y) = x \to y$. Очевидно, что подобная операция несимметрична. Тогда функция $f_{12} = (1011)$ будет также импликацией, но уже $f(x, y) = y \to x$.

Функции $f_3 = (0010)$ и $f_5 = (0100)$ являются отрицаниями импликаций f_{14} и f_{12} соответственно.

3. **Функции, явно зависящие от одной переменной.** Функция $f_4 = (0011)$, все значения которой совпадают со значениями первой переменной, обозначается $f(x, y) = x$. Аналогично $f_6(x, y) = (0101) = y$.

Функции $f_{13} = (1100)$ и $f_{11} = (1010)$ являются отрицаниями f_4 и f_6 соответственно.

В последнем примере уже из формулы было очевидно, что от переменной y функция не зависит, так как она явно не входит в формулу. Однако возможен случай, когда переменная явным образом входит, но значение функции от нее все равно не зависит.

Докажем два тождества с помощью таблицы истинности (см. табл. 4.3).

1. $x \cdot 1 = x$.

2. $y \vee \overline{y} = 1$.

Рассмотрим таблицу значений (таблицу истинности) функции $g(x, y) = x(y \vee \overline{y})$. Из тождеств видно, что $g(x, y) = x(y \vee \overline{y}) = x \cdot 1 = x$. Обе переменные в формулу входят явно, а фактически получаем ту же функцию f_4, зависящую только от x. Итак, булева функция

$f(x_1, x_2, ..., x_{k-1}, x_k, x_{k+1}, ..., x_n)$ зависит от переменной x_k **существенно**, если $\exists z = (z_1, z_2, ..., z_{k-1}, z_k, z_{k+1}, ..., z_n) \in B^n$ такой, что $f(z_1, z_2, ..., z_{k-1}, z_k, z_{k+1}, ..., z_n) \neq f(z_1, z_2, ..., z_{k-1}, \overline{z}_k, z_{k+1}, ..., z_n)$. В таком случае сама переменная x_k называется **существенной**. Если x_k не является существенной, т.е. $\forall z = (z_1, z_2, ..., z_{k-1}, z_k, z_{k+1}, ..., z_n) \in B^n$ выполняется равенство $f(z_1, z_2, ..., z_{k-1}, z_k, z_{k+1}, ..., z_n) = f(z_1, z_2, ..., z_{k-1}, \overline{z}_k, z_{k+1}, ..., z_n)$, то переменная x_k называется **фиктивной**. Например, тождественные нуль и единица любого числа переменных ни от одной из них не зависят существенно. Для лучшего осознания несущественности можно привести аналогичный пример из геометрии. Функция $z(x, y) = 5x + 4$ не зависит от переменной y явно, и ее график состоит из прямых, параллельных оси y. Вдоль каждой такой прямой значение x не меняется, y — меняется, а значение функции z не меняется. В двоичной логике та же ситуация, но рассматриваются только значения аргументов 0 и 1. Например, в функции (11000011) фиктивной является третья переменная, в функции (11001100) — первая и третья, в функции (00110010) все переменные существенны.

Очевидно, что к любой функции n переменных можно искусственно добавить k фиктивных переменных, например, приведенным выше способом, домножив на $x_{n+1} \vee \overline{x_{n+1}}$. Соответственно из таблицы истинности можно вычеркнуть столбец фиктивной переменной x_k и повторяющиеся строки, например, вида $(z_1, z_2, ..., z_{k-1}, 1, z_{k+1}, ..., z_n)$ или $(z_1, z_2, ..., z_{k-1}, 0, z_{k+1}, ..., z_n)$ по выбору. Полученные таким способом функции будем считать равными.

Способы задания булевых функций. Мы убедились, что логическую функцию можно задать **аналитически**, т.е. в виде формулы, или с помощью таблицы истинности, в левой части которой выписаны все возможные значения аргумента, а в правой — соответствующие им значения функции. Таблица истинности для булевой функции n переменных будет иметь вид (табл. 4.4).

Мы выписали сводные таблицы для функций при $n = 1$ и $n = 2$ и показали, что этим функциям соответствуют унарные и бинарные отношения. Можно было бы выписать такую масштабную

Таблица 4.4

Таблица истинности для булевой функции n переменных

x_1	x_2	...	x_n	$f(x_1, x_2, ..., x_n)$
0	0	0	0	$f(00,...,0)$
0	0	0	1	$f(00,...,01)$
...	...	...	...	...
1	1	1	0	$f(11,...,10)$
1	1	1	1	$f(11,...,1)$

таблицу и для $n = 3$, но тогда она содержала бы $2^8 = 256$ столбцов, что технически невозможно реализовать. С другой стороны, функциям трех и более переменных соответствовали бы трехместные отношения и отношения более высоких порядков. Но природа устроена согласно принципу минимальной целесообразности: там, где это возможно, в частности в дискретных множествах, одно- и двухместные отношения являются достаточными для описания всех других. Далее будет показано, что все булевы функции выражаются в виде формул через уже знакомые функции одного и двух переменных.

Соглашения о написании формул. Определены правила компактной записи булевых функций и порядок их упрощения.

- Внешние скобки не используются.
- Действия в скобках выполняются в первую очередь.
- Учитываются приоритеты логических операций в порядке возрастания: отрицание, следствие, конъюнкция, дизъюнкция.

По аналогии с непрерывными функциями можно ввести условное понятие элементарных булевых функций: отрицание, конъюнкция, дизъюнкция, константы. Тогда некоторые известные эквивалентности можно объединить в свойства элементарных булевых функций.

Теперь можно описать законы логики на уже доступном языке булевых функций. В предисловии к этой главе указывались причины, побудившие нас говорить сначала об абстрактных символах 0 и 1, а не об интуитивно знакомых, хотя и относительных понятиях «истины» и «лжи». Далее символ 0 будет означать «ложь» (или отсутствие сигнала), а 1 — «истину». Тем самым мы в явном виде построили взаимно-однозначное соответствие между множеством $B = \{0, 1\}$ и множеством {истина, ложь}.

4.3. Сложные высказывания

> Рано или поздно всякая правильная математика находит применение в том или ином деле.
>
> *А. Н. Крылов*

Сложные предложения в русском языке образуются из простых путем связок (*и*, *а*, *если... то* и т. д.). Сложные высказывания в логике образуются из простых путем логических операций, соответствующих определенным связкам естественного языка. Истинность или ложность сложного суждения являются функциями простых суждений, входящих в его состав. Зная истинность простых суждений, можно установить истинность сложных суждений.

4.3.1. Операции над сложными высказываниями

В русском языке сложные предложения получают из двух других с помощью союзов *и*, *а*, *если... то*, *либо*, *или*, *тогда и только тогда*, *когда* и др. Назовем такие и аналогичные им союзы **логическими связками**. Новые предложения появляются также с употреблением частицы *не* или слов *неверно*, *что*. Эти слова также будем называть логическими связками.

Тогда утвердительные предложения, $\frac{\text{не содержащие}}{\text{содержащие}}$ логические связки, называют $\frac{\textbf{элементарными}}{\textbf{составными}}$ высказываниями. Например, предложение «Все программисты имеют высшее или среднее специальное образование» состоит из двух элементарных предложений: «Все программисты имеют высшее специальное образование» и «Все программисты имеют среднее специальное образование». Соединены предложения связкой *или*.

Предложение «Если программист не имеет специального образования, то он не будет конкурентоспособен на рынке труда» состоит из простых предложений, соединенных связками *если... то* и отрицаний *не*. Однако не всегда грамматическая форма предложений, в которых выражено суждение, совпадает с его логической формой. Так, сложное суждение может быть представлено простым предложением. Например: «Дискретная математика включает разделы: «Элементы теории множеств», «Элементы теории графов», «Элементы классической и математической логики», «Элементы теории автоматов». Это сложное суждение хотя и представлено простым предложением, но содержит пять простых суждений.

Поскольку языком алгебры логики служат булевы функции, то естественно смысловым связкам поставить в соответствие операции над булевыми переменными.

Если простое высказывание является *истинным*, то ему соответствует значение логической переменной 1. Если простое высказывание является *ложным*, то ему соответствует значение логической переменной 0. Если выполняется все (составное) высказывание, то оно называется истинным и ему соответствует 1, если не выполняется, то высказывание считается ложным и ему соответствует 0. Составное (сложное) высказывание можно рассматривать как булеву функцию, где аргументами являются семантические характеристики простых высказываний, входящих в сложное, а значением будет семантическая характеристика всего высказывания. Остается только установить соответствие между синтактическими связками русского языка и булевыми операциями.

Для того чтобы определить, истинно или ложно некоторое сложное высказывание, используют таблицы истинности, в которых латинскими буквами обозначаются сами высказывания, а

цифрами 1 и 0 их семантическая характеристика — соответственно «истина» или «ложь». Например, высказывание «Город X расположен на берегу Волги» (A) может быть истинным (при X = Самара A = 1) или ложным (при X = Омск A = 0).

Основными операциями являются: *отрицание* (не), *дизъюнкция* (*строгая* (либо) и *нестрогая* (или)), *конъюнкция* (и), *импликация* (если... то), *эквиваленция* (тогда и только тогда). Естественно, они определяются на основе введенных ранее соответствующих булевых функций с учетом чисто логической специфики.

Введем определения логических операций.

Отрицанием, или **инверсией**, высказывания A называется высказывание $\bar{A}$, которое истинно, когда высказывание A ложно, и ложно, когда A истинно.

Дизъюнкцией (нестрогой или соединительной) высказываний A и B называется высказывание $A \vee B$, которое истинно тогда и только тогда, когда истинно *хотя бы одно из этих высказываний.*

Конъюнкцией высказываний A и B называется высказывание AB, которое истинно тогда и только тогда, когда истинны *оба высказывания.*

Строгой дизъюнкцией высказываний A и B называется высказывание $A \veebar B$, которое истинно тогда и только тогда, когда истинно *только одно из этих высказываний.*

Импликацией высказываний A и B называется высказывание $A \to B$, которое ложно тогда и только тогда, когда *из истины следует ложь.*

Эквиваленцией высказываний A и B называется высказывание $A \leftrightarrow B$, которое истинно тогда и только тогда, когда либо истинны, либо ложны одновременно *оба высказывания.*

Например, для простых высказываний «Максимов — хороший программист» (A) и «Он побеждает на олимпиадах» (B) можно составить сложные высказывания.

Конъюнкция AB: «Максимов — хороший программист, *и* он побеждает на олимпиадах» или «Максимов — хороший программист, он побеждает на олимпиадах».

Нестрогая дизъюнкция $A \vee B$: «*Или* Максимов хороший программист, *или* побеждает на олимпиадах».

Строгая дизъюнкция $A \veebar B$: «*Либо* Максимов хороший программист, *либо* побеждает на олимпиадах».

Импликация $A \to B$: «*Если* Максимов хороший программист, *то* он побеждает на олимпиадах»; «Для того чтобы Максимов побеждал на олимпиадах, *достаточно*, чтобы он был хорошим программистом».

Эквиваленция $A \leftrightarrow B$: «Максимов хороший программист *тогда и только тогда, когда* он побеждает на олимпиадах». «Максимов хороший программист *только когда* он побеждает на олимпиадах».

Заметим, что операция отрицания не всегда выражена в явном виде с помощью слова *неверно* или частицы *не* перед глаголом. Например, для высказывания «$X > 5$» (A) отрицание может быть представлено вариантами: «неверно, что $X > 5$», «X не более 5» или «$X \leq 5$», а также «$X < 5$ или $X = 5$» ($\overline{A}$).

Для высказывания «Прямая a, лежащая на плоскости, не пересекает окружность» (A) отрицанием служат высказывания: «Неверно, что прямая a, лежащая на плоскости, не пересекает окружность», «Прямая a, лежащая на плоскости, пересекает окружность», «Прямая a, лежащая на плоскости, имеет с окружностью общие точки» ($\overline{A}$).

Однако добавление частицы *не* и слова *неверно* может не превратить высказывание A в отрицательное $\overline{A}$: «Максимов — хороший программист» и «Неверно, что Максимов — хороший программист», так как речь может идти о разных Максимовых. Поэтому для получения высказываний, противоположных данным, необходимо учитывать общие правила построения противоречивых суждений A и $\overline{A}$. Если высказывание имеет вид элементарного предложения без квантора общности, то *не* перед глаголом действительно заменит высказывание A на его отрицание $\overline{A}$.

Составим **словарь перевода** с русского языка на язык алгебры логики и включим в него таблицы истинности логических операций (табл. 4.5).

Из сравнения определений логических операций, операций над множествами и соответствующих кругов Эйлера можно заключить, что дизъюнкции (логической сумме) соответствует объединение множеств, конъюнкции (логическому произведению) — пересечение (произведение) множеств, строгой дизъюнкции (сумме по модулю 2) — прямая сумма, отрицанию — разность, импликации — подмножество. Обратим внимание, что фактически в определении множеств были использованы логические связки *и*, *или*, *не* и др. В таком случае возникают проблемы с определениями, ведь получается замкнутый круг. Именно поэтому мы считаем необходимым введение логических операций только через соответствующие булевы функции, значения которых вводятся постулативно.

Определения логических операций целиком базируются на соответствующих булевых функциях и теории множеств. Все они, кроме импликации, являются понятными и бесспорными. Что касается импликации, то на первый взгляд не очевидно, почему заведомо ложное высказывание (например, «2,53 является натуральным числом») может приводить к любому утверждению (например, «Первый полет человека в космос был в 1961 г.»). Также смущает, что истинна импликация двух не связанных логически между собой истинных высказываний, например «17 — простое число» и «XXI в. начался с 2001 г.». Поэтому именно эта логическая операция будет рассмотрена подробно в следующем подразделе.

Таблица 4.5

Словарь перевода на язык алгебры логики

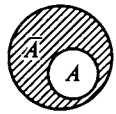
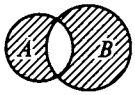
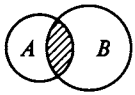
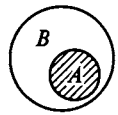
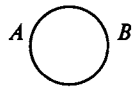
Высказывание на русском языке	Соответствие в алгебре логики	Название операции	Таблица истинности	Круги Эйлера
Не A; **неверно**, что A; A не имеет места	$\overline{A}$	Отрицание, инверсия	A $\overline{A}$ 0 1 1 0	
A **или** B; или A, или B, или оба вместе	$A \vee B$	Нестрогая дизъюнкция (соединительная)	A B $A \vee B$ 0 0 0 0 1 1 1 0 1 1 1 1	
Либо A, **либо** B; A или B, но не оба вместе	$A \veebar B$	Строгая дизъюнкция (разделительная)	A B $A \veebar B$ 0 0 0 0 1 1 1 0 1 1 1 0	
A **и** B; как A, так и B; не только A, но и B; A вместе с B; A, несмотря на B; A, в то время как B	$A \wedge B$ AB $A \,\&\, B$	Конъюнкция	A B AB 0 0 0 0 1 0 1 0 0 1 1 1	
Если A, **то** B; A, только если B; A достаточно для B; B необходимо для A; A влечет B; B тогда, когда A; из A следует B	$A \to B$	Импликация, следование	A B $A \to B$ 0 0 1 0 1 1 1 0 0 1 1 1	
A эквивалентно B; A **тогда и только тогда, когда** B; A, если и только если B; A необходимо и достаточно для B	$A \equiv B$ $A \leftrightarrow B$ $A \sim B$	Тождественность, равносильность, эквиваленция	A B $A \leftrightarrow B$ 0 0 1 0 1 0 1 0 0 1 1 1	

Таблица 4.6

Таблица истинности

X	Y	$X \to Y$	$(X \to Y)X$	$((X \to Y)X) \to Y$
0	0	1	0	1
0	1	1	0	1
1	0	0	0	1
1	1	1	1	1

Все точные знания, такие, как математические теоремы, физические, химические или биологические законы, а также события повседневной жизни, ситуации, связанные с управлением различными системами, можно сформулировать в виде высказываний и перевести на язык алгебры логики. С помощью математической логики устанавливается истинность или ложность сложных высказываний на основе истинности или ложности простых.

Если доказательство основано на применении таблиц истинности, то говорят, что использован *семантический* способ доказательств, если в процессе доказательства использовались равносильности алгебры логики, то способ называют *синтактическим*.

Задача 21. Поможем синоптикам определить прогноз погоды. Известно, что если атмосферное давление понижается, то возможен дождь. В настоящее время атмосферное давление понижается. Возможен ли дождь?

Решение. Формализуем условия задачи, введя обозначения. Пусть X — атмосферное давление понижается; Y — возможен дождь.

Высказывание «Если давление понижается, то возможен дождь» имеет вид: $X \to Y$. Тогда «Если давление понижается, то возможен дождь. Давление понижается» можно записать в виде конъюнкции $(X \to Y) \cdot X$. Сформулируем теорему: $((X \to Y)X) \to Y = 1$. Докажем истинность вывода, используя оба способа.

Синтактический способ. $((X \to Y)X) \to Y = ((\overline{X} \vee Y) \cdot X) \to Y = (Y \cdot X) \to Y = \overline{X \cdot Y} \vee Y = \overline{X} \vee \overline{Y} \vee Y = \overline{X} \vee 1 = 1$.

Семантический способ представлен в табл. 4.6.

Из табл. 4.6 видно, что теорема истинна при любом наборе значений X и Y. Таким образом, доказана справедливость утверждения «Возможен дождь». Получение тождественной единицы подтверждает справедливость доказываемой теоремы.

4.3.2. Необходимое и достаточное условия импликации

Все логические операции имеют широкое применение в обычной речи. Но особенно важное значение в познании имеют операции импликации и эквиваленции. Рассмотрим пример. Высказы-

вание A: «Выполнил домашнее задание» и B: «Получил пятерку». Попробуем составить импликацию $A \to B$. Тогда имеем $A \to B$: «Если выполнил домашнее задание, то получил пятерку». Таблица истинности для этого высказывания примет вид (табл. 4.7).

Аналогичные выводы можно сделать с любыми простыми высказываниями. Например, A: «Через проводник пропустили ток»; B: «Длина проводника увеличилась».

Действительно, вывод «Если через проводник пропустили ток, то его длина увеличится», будет ложным, если ток пропущен, а длина проводника не увеличилась. А вот если и ток не пропускали, и длина не увеличилась, то вывод станет истинным. Ведь отсутствие у предмета некоторого свойства — тоже свойство. Но если ток пропускать не стали, а длина увеличилась, то мы говорим, что вывод все равно истинный, так как длина могла увеличиться по другой причине, например под действием тепла или неупругой деформации.

Импликация не обладает переместительным свойством, т. е. если известно, что $A \to B$, то это не равносильно импликации $B \to A$, что непосредственно следует из сравнения таблиц истинности этих двух булевых функций. Например, высказывание «Если через проводник пропущен ток, то его длина увеличится» истинно. Но про высказывание «Если длина проводника увеличилась, то через него пропустили ток» нельзя сказать, истинно оно или ложно, так как увеличение длины проводника не является причиной того, что через этот проводник стали пропускать ток.

А для высказывания «Если выполнил домашнее задание, то получил пятерку» обратное — «Если получил пятерку, то выполнил домашнее задание» — тоже, вообще говоря, неверно.

Таблица 4.7

Таблица истинности

A	B	$A \to B$	Примеры
1	1	1	«Если задание выполнил (1), то пятерку получил (1)» — вывод $A \to B$ истинный
1	0	0	«Задание выполнил (1), а пятерку не получил (0)» — вывод $A \to B$ ложный
0	1	1	«Задание не выполнил (0), но пятерку получил (1)» — вывод $A \to B$ истинный (1), так как пятерку поставили, например, за другой вид работы на уроке
0	0	1	«Если задание не выполнил (0), то пятерку не получил (0)» — вывод $A \to B$ истинный, ведь он полностью противоречит выводу, т. е. не ложь, значит, истина

Таблица 4.8

Таблица истинности

Поезд a	Сигнал b	
	Путь закрыт ($b = 1$)	Путь открыт ($b = 0$)
Прибывает ($a = 1$)	1	0
Не прибывает ($a = 0$)	1	1

Рассмотрим пример. Обозначим через a высказывание «Некоторый поезд прибывает на данную станцию», а через b — высказывание «Подается сигнал "Путь закрыт"». Тогда импликация $a \rightarrow b$ означает «Если некоторый поезд прибывает на станцию, то подается сигнал "Путь закрыт"» или «Как только поезд прибывает на станцию, подается сигнал "Путь закрыт"», а также «Для того чтобы был подан сигнал "Путь закрыт", достаточно, чтобы поезд прибыл на станцию». Определим истинность производных высказываний с помощью табл. 4.8.

Из табл. 4.8 видно, что импликация истинна в трех случаях из четырех:

- поезд прибывает, и подается сигнал «Путь закрыт»;
- поезд не прибывает, и подается сигнал «Путь открыт»;
- поезд не прибывает, и подается сигнал «Путь закрыт» (так как путь мог быть закрыт независимо от прибытия этого поезда).

Импликация ложна только в одном случае: поезд прибывает, а подается сигнал «Путь открыт». Этот и другие аналогичные примеры подробно рассмотрены в книге Х. Фрейденталя «Язык логики».

В разговорной речи и при изучении различных дисциплин необходимо хорошо понимать, что является причиной (или посылкой), а что — следствием (или заключением).

Причина, посылка, т.е. высказывание, являющееся первым аргументом импликации, называется **достаточным условием**. В разговорной речи это то высказывание, которое находится между словами *если... то*, перед союзами *поэтому*, *следовательно*, *значит* или после союзов *так как*, *потому что* и т.д. «Для того чтобы ученик получил пятерку», *достаточно* «правильно выполнить домашнее задание». «Ученик получил пятерку», *так как* «правильно выполнил домашнее задание». Эти примеры — словесная реализация импликации $A \rightarrow B$, где A — высказывание «Ученик правильно выполнил домашнее задание», B — «Ученик получил пятерку».

Следствие, заключение, т.е. высказывание, являющееся вторым аргументом импликации, называется **необходимым условием**. Оно стоит после союза *то* в предложениях с оборотами *если... то*, *поскольку... то*, *так как... то*, а также по другую от достаточного

условия сторону по отношению к соединительным союзам. Надо научиться отличать необходимое условие от достаточного, так как от этого зависит правильность доказательства теорем и выводов во всех областях естествознания.

Условие теоремы — то, что дано в ее формулировке, — нельзя путать с заключением — тем, что требует доказательства. А для того чтобы выяснить, где достаточное условие, а где то, что необходимо доказать, надо переформулировать теорему, используя слова «если..., то».

Например, в теореме Пифагора «В прямоугольном треугольнике квадрат гипотенузы равен сумме квадратов катетов», надо выделить условие и заключение. Новая формулировка этой теоремы, согласно наложенным требованиям, имеет вид: «Если дан прямоугольный треугольник, то квадрат гипотенузы равен сумме квадратов его катетов». Сравним с привычной записью условия теоремы в геометрии.

Дано: ΔABC — прямоугольный, $\angle C = 90°$, c — гипотенуза, a, b — катеты.

Доказать: $c^2 = a^2 + b^2$.

Подчеркнем еще раз, что, вообще говоря, импликация не обладает переместительным свойством: $(A \to B) \neq (B \to A)$. Но учащиеся во время рассуждений об этом часто забывают. Например, равносильны ли два математических высказывания «Вертикальные углы равны» и обратное: «Углы равны, значит, они вертикальные»?

Первое утверждение истинно, так как доказывается соответствующая теорема. Второе неверно, так как в некоторых случаях равные углы действительно могут быть вертикальными, а в некоторых они могут иметь другую причину равенства. Например, они могут быть углами при основании равнобедренного треугольника или внутренними накрест лежащими углами при параллельных прямых и секущей.

Поэтому и говорят, что высказывание «Углы вертикальные» является *достаточным* для утверждения «Углы равны». А равенство углов *необходимо*, чтобы утверждать, что речь идет о вертикальных углах. Невыполнение *необходимого* условия (его ложность) влечет за собой невыполнение *достаточного* условия, например «Если углы не равны, то они (в принципе) не могут быть накрест лежащими».

Назовем $B \to A$ **обратным высказыванием** для высказывания $A \to B$, а высказывание $\overline{A} \to \overline{B}$ — **противоположным** к импликации $A \to B$. Рассмотрим правила построения обратных высказываний (рис. 4.2).

Диагональные стрелки на рис. 4.2 показывают одновременную истинность (т.е. эквиваленцию) соответствующих высказываний. Докажем, например, равенство $A \to B = \overline{B} \to \overline{A}$, которое назовем

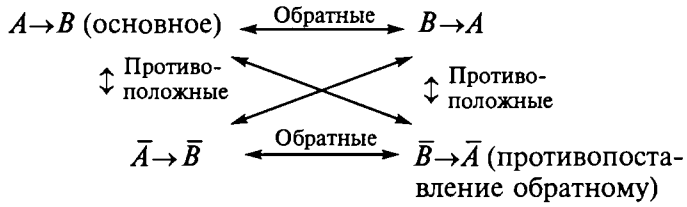


Рис. 4.2. Правила построения обратных и противоположных высказываний

правилом **контрапозиции** (от лат. *contrapositio* — противопоставление).

Доказательство проведем с помощью таблицы истинности и проверим на примере (табл. 4.9), где A: «Вертикальные углы», B: «Равные углы», $A \to B$: «Если углы вертикальные, то они равны», $\bar{B} \to \bar{A}$: «Если углы не равны, то они не могут быть вертикальными».

Совпадение третьего и шестого столбцов таблицы истинности подтверждает вывод о том, что высказывание, противоположное обратному ($\bar{B} \to \bar{A}$), равносильно основному высказыванию $A \to B$. Формула $A \to B = \bar{B} \to \bar{A}$ справедлива для любых высказываний A и B. Поэтому, беря в качестве A высказывание B и наоборот, получаем формулу $B \to A = \bar{A} \to \bar{B}$, соответствующую другой диагональной стрелке на рис. 4.2.

Ошибки при применении импликации. $(A \to B) = (\bar{A} \to \bar{B})$ — это высказывание — импликация, противоположная исходному высказыванию, но такая операция, вообще говоря, неверна. Ошибка в том, что положительный смысл изменили на отрицательный, но местами не поменяли, т. е. не заменили на высказывание, обратное противоположному. Например, путь C — «Если $ABCD$ — квадрат (A), то его стороны равны (B)»: $C = A \to B$. Противоположное формулируется как «Если четырехугольник — не квадрат ($\bar{A}$), то его стороны не равны ($\bar{B}$). Это неверно. Например, ромб не квадрат, но его стороны равны. Для исправления ошибки надо сформулировать высказывание, обратное противоположному: $\bar{B} \to \bar{A}$:

Таблица 4.9

Таблица истинности

A	B	$A \to B$	$\bar{A}$	$\bar{B}$	$\bar{B} \to \bar{A}$
1	1	1	0	0	1
1	0	0	0	1	0
0	1	1	1	0	1
0	0	1	1	1	1

«Если стороны не равны ($\overline{B}$), то $ABCD$ — не квадрат ($\overline{A}$)» или «Для того чтобы $ABCD$ был квадратом, *необходимо* чтобы его стороны были равны».

Равенство сторон — необходимое свойство квадрата.

Эту же ошибку можно переформулировать иначе: $(\overline{A} \to \overline{B}) = (A \to B)$. Проще говоря, из невыполнения достаточного условия ничего определенного не следует, так как из ложного высказывания может следовать все, что угодно.

Тот факт, что противоположные и обратные высказывания не всегда дают истинный результат, лежит в основе многих математических ошибок.

Рассмотрим пример. Даны простые высказывания:

A: «$ABCD$ — квадрат»;

B: «Диагонали $ABCD$ равны».

Сформулировать прямую, обратную, противоположную и противоположную обратной теоремы (или импликации).

Выводы:

- для того чтобы диагонали четырехугольника были равны, *достаточно*, чтобы четырехугольник был квадратом;
- равенство диагоналей четырехугольника — *необходимое* условие для существования прямоугольника;
- для того чтобы четырехугольник $ABCD$ был квадратом, *достаточно*, чтобы его диагонали были равны, перпендикулярны и делились точкой пересечения пополам (при этом каждое из трех последних условий является необходимым для квадрата).

Итак, прямая и обратная импликации не равны между собой и в рассмотренном примере имели противоположные семантические характеристики. Но прямая и обратная теоремы *не всегда* имеют противоположный смысл.

Существуют такие высказывания, для которых одновременно справедливы и прямая, и обратная импликации: т. е. $A \to B$ и $B \to A$. Такие логические операции называются **эквиваленцией**. Покажем, что это определение не противоречит определению эквиваленции, данному в подразд. 4.3. На языке логических операций это можно записать так: $A \leftrightarrow B = (A \to B) \cdot (B \to A)$.

Прямой подстановкой $\begin{pmatrix} A & B \\ B & A \end{pmatrix}$ в формулу можно убедиться в том, что эквиваленция симметрична по своим аргументам.

Справедливость этой формулы видна из таблицы истинности (табл. 4.10).

Рассмотрим еще один пример. Даны высказывания:

A: «Сумма цифр целого числа в десятичной записи делится на 3»;

B: «Число делится на 3».

Сформулировав высказывания $A \to B$, $B \to A$, $\overline{A} \to \overline{B}$, $\overline{B} \to \overline{A}$ и определив их истинность, самостоятельно убедитесь, что обрат-

Таблица 4.10

Таблица истинности для эквиваленции

A	B	$A \to B$	$B \to A$	$(A \to B) \cdot (B \to A)$	$A \leftrightarrow B$
0	0	1	1	1	1
0	1	1	0	0	0
1	0	0	1	0	0
1	1	1	1	1	1

ные импликации тождественны. Эти же высказывания можно было сформулировать и с оборотом *необходимо и достаточно*: «Для того чтобы число делилось на 3, необходимо и достаточно, чтобы сумма цифр делилась на 3».

В некоторых случаях удобнее использовать слова *тогда и только тогда*. Например, с этими словами формулируются определения всех логических операций.

В математике, когда аргументы каждой импликации сами являются сложными высказываниями, эквивалентность двух утверждений называется **критерием**. По формуле $A \leftrightarrow B = (A \to B) \cdot (B \to A)$ его доказывают сначала в одну сторону $(A \to B)$, затем в обратную $(B \to A)$.

Эквиваленцию можно использовать в определении всех логических операций. Приведем два примера.

Импликацией называется такая логическая операция, которая ложна *тогда и только тогда*, когда из истинного высказывания следует ложь.

Нестрогой дизъюнкцией двух высказываний называется такая логическая операция, которая ложна *тогда и только тогда*, когда оба высказывания ложны одновременно.

Следует помнить, что подобные определения будут верными, если операция эквиваленции установлена независимо. Если же ее определить по формуле $A \to B = \overline{B} \to \overline{A}$, то эти определения будут содержать в себе порочный круг: эквиваленция определяется через конъюнкцию и импликацию, а импликация и конъюнкция определяются словами *тогда и только тогда*, т.е. через эквиваленцию.

Поясним различие между знаком =, применяемым в формулах, и знаком ≡, символизирующим эквиваленцию. Рассмотрим два выражения: $y \vee \overline{y} = 1$ и $y \equiv 1$. Второе выражение — это булева функция, которая может принимать разные значения в зависимости от значений переменной y. Первое выражение является тождеством, оно истинно *всегда*, т.е. отличается от второго случая. В общей алгебре, где нет понятия булевой функции эквиваленции, наоборот, знак ≡ обозначает тождественность, т.е. истинность

при любых значениях. Знак = входит в уравнение, которое, например, нужно решить. При этом *только* корни уравнения, будучи подставлены в него, обратят его в тождество.

4.3.3. Формулы алгебры логики

В математической логике, так же как и в алгебре, операции подчиняются определенным законам. С помощью этих законов (тождеств, равносильностей) можно упрощать составные высказывания.

Рассмотрим их для операций дизъюнкции и конъюнкции, учитывая свойство **двойственности.** Двойственность операций (табл. 4.11) заключается в том, что если в формуле, содержащей лишь операции дизъюнкции, конъюнкции и отрицания, заменить операции $\wedge$ и $\vee$ на $\vee$ и $\wedge$ соответственно, а символы 1 на 0 и 0 на 1, то получатся новые равносильности.

Законы Де Моргана называют переносом отрицания через логические связки.

Приведем формулы с другими операциями, которые также будем считать основными:

Таблица 4.11

Законы алгебры логики

Дизъюнкция	Законы	Конъюнкция
$a \vee b = b \vee a$	Переместительный закон	$ab = ba$
$a \vee (b \vee c) = (a \vee b) \vee c$	Сочетательный закон	$a(bc) = (ab)c$
$a(b \vee c) = ab \vee ac$	Распределительный закон	$a \vee (bc) = (a \vee b)(a \vee c)$
$a \vee a = a$	Правила идемпотентности	$a \cdot a = a$
$\overline{a \vee b} = \overline{a} \cdot \overline{b}$	Законы Де Моргана	$\overline{ab} = \overline{a} \vee \overline{b}$
$a \vee 0 = a$	Правила операций с константами	$a \cdot 0 = 0$
$a \vee 1 = 1$		$a \cdot 1 = a$
$a \vee ab = a$	Законы поглощения	$a(a \vee b) = a$
$a \vee (\overline{a}b) = a \vee b$		$a(\overline{a} \vee b) = ab$
$a \vee \overline{a} = 1$	Законы инверсии (отрицания)	$a \cdot \overline{a} = 0$
$ab \vee a\overline{b} = a$	Законы склеивания	$(a \vee b) \cdot (a \vee \overline{b}) = a$

• $\overline{\overline{a}} = a$ — снятие двойного отрицания;
• $a \to b = \overline{a} \vee b$ — снятие импликации;
• $a \sim b = ab \vee \overline{a} \cdot \overline{b}$ — снятие эквиваленции;
• $a \veebar b = \overline{a}b \vee a\overline{b}$ — снятие строгой дизъюнкции.

Все эти формулы получаются простой проверкой по таблице истинности с учетом истинности каждой операции, правильного раскрытия скобок и выполнения операций по приоритету. Читателю предлагается проверить самому эти тождества.

Знание законов математической логики помогает не только упрощать высказывания, но и правильно, логически рассуждать. Так, некоторые формулы помогают понять законы правильного мышления (см. далее в подразд. 4.4).

С помощью введенных логических операций над высказываниями можно строить различные сложные высказывания. Использование скобок дает возможность установить порядок выполнения операций.

Рассмотрим несколько примеров на установление истинности сложного высказывания.

1. $(\overline{X} \vee Y) \to \overline{\overline{Y}} \to X$.

Истинность сложного высказывания можно установить с помощью таблицы истинности (табл. 4.12).

Введем порядок действий, определяющий последовательность столбцов в таблице истинности для высказывания $(\overline{X} \vee Y) \to \to \overline{\overline{Y} \to X}$:

1) $\overline{X}$; 2) $\overline{X} \vee Y$; 3) $\overline{Y}$; 4) $\overline{Y} \to X$; 5) $\overline{\overline{Y} \to X}$; 6) $(\overline{X} \vee Y) \to \overline{\overline{Y} \to X}$.

Любые простые и сложные высказывания, полученные из элементарных высказываний с помощью конечного числа введенных логических операций, называются **формулами алгебры логики**.

Для упрощения формул, содержащих скобки и различные логические операции, будем учитывать ряд правил. Так, при опускании скобок:

• самой первой выполняется конъюнкция между элементарными высказываниями и их отрицаниями;
• дизъюнкция выполняется раньше импликации и эквиваленции;
• знак отрицания над формулой дает возможность опустить скобки, в которых эта формула заключена.

Таблица 4.12

Таблица истинности для двух переменных

X	Y	$\overline{X}$	$\overline{X} \vee Y$	$\overline{Y}$	$\overline{Y} \to X$	$\overline{\overline{Y} \to X}$	$(\overline{X} \vee Y) \to \overline{Y} \to X$
0	0	1	1	1	0	1	1
0	1	1	1	0	1	0	0
1	0	0	0	1	1	0	1
1	1	0	1	0	1	0	0

Таблица 4.13

Таблица истинности для трех переменных

X	Y	Z	$\overline{Y}$	$X \vee \overline{Y}$	$(X \vee \overline{Y})Z$
0	0	0	1	1	0
0	0	1	1	1	1
0	1	0	0	0	0
0	1	1	0	0	0
1	0	0	1	1	0
1	0	1	1	1	1
1	1	0	0	1	0
1	1	1	0	1	1

2. $(X \vee \overline{Y}) \cdot Z$. Так, в формуле $(X \vee \overline{Y}) \cdot Z$ определим порядок действий: 1) $\overline{Y}$; 2) $X \vee \overline{Y}$; 3) $(X \vee \overline{Y}) \cdot Z$ и истинность (табл. 4.13) для набора значений трех переменных X, Y и Z.

С помощью таблицы истинности (или таблицы состояний) можно описать все возможные комбинации трех входных переменных, что имеет широкое практическое значение. Такая таблица позволяет анализировать контактную технологическую структуру с помощью записи в систематизированном виде состояний ее выхода для всех возможных состояний входов. Математическая логика является тем прочным теоретическим фундаментом, на котором основаны многие технические проекты.

3. $((X \vee Y)X) \vee \overline{XY}$. Для формулы $((X \vee Y)X) \vee \overline{XY}$ установим порядок действий: сначала уберем скобки, заключающие отрицание $X \vee Y$ как излишние, и скобки для операции конъюнкции, так как она в любом случае выполняется раньше дизъюнкции. Тогда формула примет вид $\overline{X \vee Y} \cdot X \vee \overline{XY}$. Действия будем выполнять в следующем порядке: 1) $X \vee Y$; 2) $\overline{X \vee Y}$; 3) $\overline{X \vee Y} \cdot X$; 4) $\overline{X}$; 5) $\overline{XY}$; 6) $\overline{X \vee Y} \cdot X \vee \overline{XY}$. Таблица истинности примет следующий вид (табл. 4.14):

Таблица 4.14

Таблица истинности

X	Y	$X \vee Y$	$\overline{X \vee Y}$	$\overline{X \vee Y} \cdot X$	$\overline{X}$	$\overline{XY}$	$\overline{X \vee Y} \cdot X \vee \overline{XY}$
0	0	0	1	0	1	0	0
0	1	1	0	0	1	1	1
1	0	1	0	0	0	0	0
1	1	1	0	0	0	0	0

Таким образом, логические значения формул полностью определяются логическими значениями элементарных высказываний, из которых эти формулы состоят. Так, формула из примера 3 будет истинна, только если X — ложно, а Y — истинно. Во всех остальных случаях она — ложна.

Две формулы алгебры логики A и B называются **равносильными**, если они принимают одинаковые логические значения при любом наборе значений элементарных высказываний, входящих в них. Обозначают равносильности (тождества) с помощью знака =.

Формула A называется **тождественно-истинной**, или **тавтологией**, если она принимает значение «истинно» при всех значениях переменных, входящих в нее. Иными словами, тавтологией является функция, где все переменные фиктивны и хотя бы при одном наборе значений аргументов ее значение равно 1. Так, тавтологиями будут: 1) $a \vee \overline{a}$; 2) $a \to (b \to a)$; 3) $a \vee (a \to (b \to a))$.

Первое утверждение $a \vee \overline{a} = 1$ было доказано в подразд. 4.2. Докажем второе $a \to (b \to a) = 1$ (табл. 4.15).

Таблица 4.15

Таблица истинности

a	b	$b \to a$	$a \to (b \to a)$
0	0	1	1
0	1	0	1
1	0	1	1
1	1	1	1

Третье $a \vee (a \to (b \to a)) = 1$ следует из второго и формулы $a \vee 1 = 1$.

Формула называется **тождественно-ложной**, если она принимает значение нуль при всех значениях переменной, входящих в нее. Так, тождественно-ложны формулы $(a \leftrightarrow \overline{a})$ и $(a \leftrightarrow \overline{a}) \cdot (a \vee b \to a)$ (табл. 4.16).

Таблица 4.16

Таблица истинности

a	b	$\overline{a}$	$a \leftrightarrow \overline{a}$	$b \to a$	$a \vee b \to a$	$(a \leftrightarrow \overline{a}) \cdot (a \vee b \to a)$
0	0	1	0	1	1	0
0	1	1	0	0	0	0
1	0	0	0	1	1	0
1	1	0	0	1	1	0

Таблица 4.17

Распределительный закон для операции дизъюнкции

a	b	c	bc	$a \vee (bc)$	$(a \vee b)$	$(a \vee c)$	$(a \vee b)(a \vee c)$	$a \vee (bc) \leftrightarrow (a \vee b)(a \vee c)$
0	0	0	0	0	0	0	0	1
0	0	1	0	0	0	1	0	1
0	1	0	0	0	1	0	0	1
0	1	1	1	1	1	1	1	1
1	0	0	0	1	1	1	1	1
1	0	1	0	1	1	1	1	1
1	1	0	0	1	1	1	1	1
1	1	1	1	1	1	1	1	1

Попутно убеждаемся, что $b \to a = a \vee b \to a$, т.е. формула $b \to \to a \leftrightarrow a \vee b \to a$ является тавтологией.

Докажем некоторые равносильности, проверив их с помощью таблиц истинности.

Распределительный закон (дистрибутивность) дизъюнкции относительно конъюнкции непривычен, так как аналогичного закона в обычной алгебре нет. В качестве доказательства сравним результаты таблиц истинности (табл. 4.17) левой и правой части уравнения $a \vee (bc) \leftrightarrow (a \vee b)(a \vee c)$.

Результаты пятого и восьмого столбцов совпали, следовательно, девятый столбец состоит из всех единиц, что соответствует тождественной истине.

4.4. Законы правильного мышления

> Суров закон, но он закон. Ложное направление ума заключается исключительно в привычке рассуждать из плохо определенных принципов.
>
> *Э. Б. де Кондильяк*

Известно, что классическая логика изучает характер связи мыслей в процессе рассуждений. На протяжении всего своего развития человечество выработало законы, при которых эта связь между мыслями оптимальна.

Логические законы универсальны, едины для всех людей, объективны, т.е. не зависят от наших знаний или желаний. Три первых логических закона были сформулированы Аристотелем в IV в. до н. э. в «Аналитике» и «Метафизике». Четвертый закон был сформулирован Лейбницем в «Монадологии».

Законы логики применяются как принципы правильных рассуждений в ходе доказательства верных суждений и опровержения мысли. При правильном применении законов мышления люди однозначно понимают друг друга, однозначно толкуют информацию и делают выводы, адекватные этой информации.

Законы мышления не устанавливаются волей людей, а лишь сформулированы ими на основе изучения свойств и отношений между мыслями, т.е. носят нормативный характер. Мысли, сформулированные без учета законов мышления, порой многозначны и неопределенны. Нарушение законов логики в общении ведет к взаимному непониманию.

Итак, **законы мышления** — это необходимая, существенная и устойчивая связь между мыслями. Они отражают в сознании людей связь между предметами объективного мира, имеют общечеловеческий характер и сформулированы на основании свойств мышления — определенности, непротиворечивости, обоснованности, четкости выбора. Законы правильного мышления, как наиболее общие, лежат в основе операций с понятиями, суждениями и умозаключениями, причем не только в общении между людьми, но и в основе работы ЭВМ.

1. **Закон тождества** задается формулой $a = a$ (*а* есть *а*).

В процессе рассуждений всякое понятие и суждение тождественно самому себе.

Закон тождества выражает коренное свойство правильного мышления — его качественную определенность. Он устанавливает правила функционирования отдельно взятой мысли, которая носит конкретный, качественно определенный однозначный характер, соответствующий предмету мысли. Закон универсален для всех форм мышления.

Правила соблюдения закона тождества.

Мысль или понятие не подменяется другой мыслью или понятием. Особенно часто проявляется в спорах, где нередки подмена тезиса и другие ошибки в доказательствах. («Кто про Фому, а кто про Ерему».)

Недопустима замена слов омонимами (*игра слов*). Примером такой игры слов может служить эпиграмма Николая Минского: «Переводимы все — прозаик и поэт. Лишь переводчики — им перевода нет», а также знаменитая фраза «Ученик *прослушал* рассказ учителя».

Нельзя путать формальное тождество с содержательным. На этой ошибке основаны многие софизмы: «Сидящий встал. Кто встал, тот стоит. Значит, сидящий стоит»; «2 и 3 есть четное и нечетное. 2 и 3 есть 5. Значит, 5 есть четное и нечетное».

Нельзя использовать неопределенные термины или некорректно их формулировать. Например, нечеткий вопрос спровоцировал неадекватный ответ:

— У вас *затопили*?

— Да, со всех сторон, даже сверху...

Тонко подметил такие логические ошибки Льюис Кэрролл в книге «Алиса в Стране чудес»: «Все понятно! — с торжеством сказал Шляпа. — Провести время? Ишь чего захотела! Время не проведешь! Да и не любит он этого! Ты бы лучше постаралась с ним подружиться — вот тогда бы твое дело было... в шляпе!»

Широко применяется закон тождества в точных науках, например в математике, информатике и т.д. Рассмотрим несколько примеров его применения.

• Замена различными буквами значения некоторой переменной, например *подстановка*. Переменная заменяется на свое конкретное значение, выражение на тождественное себе и т.д. Примером может служить тождественность при использовании символов отрицания $\overline{a}$ и $\neg a$.

• Тождественные преобразования выражений: $(a + b)^2 = a^2 + ab + ba + b^2$; $\cos 2\alpha = \cos^2\alpha - \sin^2\alpha$; $x^3 = x \cdot x \cdot x$ и т.д.

• Тождественны свойства отношений эквивалентности (см. гл. 1):

рефлективность aRa; например, $a \parallel a$;

симметричность $aRb = bRa$; например, $(a \parallel b) \Rightarrow (b \parallel a)$; $(a \perp b) \Rightarrow$ $\Rightarrow (b \perp a)$; $(a = b) \Rightarrow (b = a)$ и т.д.;

транзитивность: если aRb, а bRc, тогда aRc; например, $a \parallel b$ и $b \parallel c$, значит, $a \parallel c$.

• В теории алгоритмов и теории кодирования (см. гл. 6) равны числа в различных алфавитах: $2_{10} = 10_2$; $3_{10} = 11_2$; $4_{10} = 100_2$ и т.д.

С помощью закона тождества даются формулировки теорем, определения понятий, различные формулы и т.д.

Красота русского языка выражается в его многозначности, в возможности употребления *синонимов*, отождествления смысла различных слов, выражающих одно и то же понятие.

В процессе движения происходит изменение образа. Например, процесс старения человека (Иванова), изменение погодных условий в данном городе (М), но при этом сам Иванов или город М остаются в нашем восприятии и в восприятии государства тем же человеком и тем же городом. У Иванова сохраняется номер паспорта, а город М является определенным административным центром и получает столько-то полномочий и бюджетных денег.

Отождествление понятий происходит, когда человек, слушая или читая, переносит эти понятия на себя, т.е. рассматривает применительно к своему опыту.

Мы используем закон тождества при решении любой задачи по заданному алгоритму, отождествляя условие конкретной задачи с известными приемами ее решения. Аналогично, отождествляя данное слово с правилами грамматики, мы применяем правила и не допускаем ошибок в правописании.

Если закон тождества выражает логическую однозначность и определенность мышления, то закон противоречия выражает его последовательность и противоречивость.

2. **Закон противоречия** задается формулой $a \cdot \bar{a} = 0$.

У Аристотеля формулировка закона имела вид: «Невозможно что-либо *одновременно* утверждать и отрицать».

Современная формулировка выглядит так: «Два противоположных или противоречащих суждения, взятых *одновременно*, не могут быть истинными в одном и том же отношении, в одном и том же месте». Например, известное выражение «Я — не я, и лошадь не моя».

Смысл этого закона заключается в требовании не допускать противоречивых мыслей. Универсальность закона противоречия проявляется в работе с понятиями и суждениями (A и $\bar{A}$ противоречивые), а также с умозаключениями, которые должны быть непротиворечивы. На использовании этого закона основаны, например, непосредственные умозаключения через превращение суждений. Порой закон противоречия называют законом непротиворечия, отождествляя название закона с его действием (и при этом нарушают законы логики).

Правила соблюдения закона противоречия.

Нельзя допускать взаимоисключающие суждения в структуре одного предложения.

Необходимо давать истинные определения и оценки суждений.

Необходимо выявлять как явные, так и скрытые противоречия, но при этом различать реальные и мнимые противоречия.

Особое значение имеют **логические противоречия** — взаимоисключающие суждения, зафиксированные в одном контексте.

Например, не могут быть одновременно истинными или ложными общеутвердительное и общеотрицательное суждение об одном и том же: «Все люди применяют законы логики», «Ни один человек не применяет законов логики» или общеутвердительные и общеотрицательное суждения (см. подразд. 5.5): «Все студенты сдали экзамены за I семестр», «Некоторые студенты сдали экзамены за I семестр».

В науке закон противоречия имеет особое значение, так как одним из основных требований к ее построению как формальной системы является непротиворечивость ее исходных данных (подробно см. в подразд. 5.1 и 5.2). При построении формальных теорий непротиворечивой должна быть система аксиом, а также непротиворечивыми должны быть выводы в этой теории. Однако противоречия научной системы могут иметь и позитивный характер, являясь двигателем прогресса. Так, противоречие между незнанием о чем-то и стремлением узнать больше лежит в основе движения и развития всей теории познания.

Например, в математике каждое новое математическое действие возникало на основе противоречия между знаниями той исторической эпохи и неудовлетворенностью при решении конкретных задач. В результате расширялись знания человека о числах.

Потребность решить уравнение типа $15 + x = 9$ привела к открытию отрицательных чисел ($x = -6$). Стремление решить уравнение вида $5x = 9$ привело к открытию обыкновенных дробей ($x = 9/5$). Попытки решить некоторые квадратные уравнения вида $x^2 = 2$ привели к открытию иррациональных чисел ($x = \pm\sqrt{2}$). Потребность в решении уравнений вида $x^2 = -1$ привела к открытию комплексных чисел $x = \pm i$.

В природе и в науке вообще, независимо от места и времени, противоположности могут существовать вместе (как, например, рациональные и иррациональные числа). Тогда противоположности требуют разделения и отдельного изучения. Разрешение противоречий открывает дорогу для развития мысли, для творчества.

Особое значение имеет закон противоречия в ходе проведения доказательств. С этим же связан и четвертый закон правильного мышления. Он лежит в основе требований, накладываемых на доказательства. Непротиворечивыми должны быть как используемые в доказательствах посылки (тезисы и аргументы), так и выводы этой теории.

В математике закон противоречия используется и при доказательстве методом от противного.

Пусть дано истинное суждение A и надо доказать справедливость B. Метод от противного заключается в том, что вместо B берут его противоречие $\overline{B}$ и затем с помощью математических аргументов доказывается, что $A \cdot \overline{B} = 0$ (ложь). Если удается это доказать, то в результате делается вывод, что B — истинно. Проверим это: $A \cdot \overline{B} = 0$ равносильно $\overline{A \cdot \overline{B}} = 1$, значит, $\overline{A} \vee \overline{\overline{B}} = \overline{A} \vee B = 1$, но $\overline{A}$ — ложно, т.е. для того чтобы значение дизъюнкции было равно 1, *необходимо*, чтобы $B = 1$.

Методом от противного доказываются теоремы, которые нельзя или трудно доказать непосредственно.

Доведение до абсурда — прием, который используется не только в математике. В споре (полемике) метод приведения к абсурду заключается в том, что с помощью аргументов оппонента доказывается, что он противоречит сам себе. Применяют этот прием, например, адвокаты в процессе защиты на суде. В диалектике закону противоречия соответствует закон единства и борьбы противоположностей.

Нельзя мириться с логическими противоречиями в обыденной жизни людей. Однако иногда юристы употребляют фразу «Отсутствие следа на месте преступления — тоже след». Имеется в виду след опытного преступника, умеющего не оставлять следов.

Закон противоречия дополняет закон тождества. Это как бы крайние проявления различия вещей, их нетождественности. Например, день и ночь, зима и лето, живой и мертвый, горячий и холодный и т.д.

С помощью приема синтеза двух противоречивых понятий авторы — художники, писатели — могут придумать броское, привлекающее внимание и в то же время обнажающее проблемы название произведения: «Яблоки на снегу», «Живой труп», «Горячий снег».

Закон противоречия устанавливает, что утверждение и отрицание одной и той же мысли не могут быть одновременно истинными. На вопрос, могут ли они быть одновременно ложными, дает ответ закон исключенного третьего.

3. **Закон исключенного третьего** задается формулой $A \vee \overline{A} = 1$.

Из двух противоречивых суждений одно истинно, другое ложно. Третьего не дано. Он сформулирован в рамках двузначной логики Аристотеля, признающей только две альтернативы: истинно — ложно.

Закон исключенного третьего работает в процессе построения умозаключений и доказательств. На его основе строятся непосредственные умозаключения в виде превращений, а также противоречащие суждения (см. подразд. 5.5). С помощью закона исключенного третьего проводятся косвенные доказательства, так как для того чтобы установить истинность доказываемого тезиса, надо выявить ложность противоречащего суждения.

Правила соблюдения закона исключенного третьего.

Существуют две альтернативы A и $\overline{A}$, между которыми надо сделать выбор.

Третьей альтернативы, т. е. некоторого суждения $B = \overline{A} = \overline{\overline{A}}$, не существует.

Между альтернативами установлено отношение противоречия — одна является отрицанием другой.

Этот закон очень важен в дискуссиях при выборе альтернатив, во время голосования.

Он же лежит в основе дихотомического деления, в котором также существуют две взаимоисключающие друг друга альтернативы A и $\overline{A}$. На законе исключенного третьего основано кодирование информации для ЭВМ, а также вся математическая логика.

Но жизнь многообразнее аристотелевой логики. И в житейских ситуациях, науке, политике часто бывает недостаточно двух альтернатив. Например, на выборах каждый человек голосовал по закону исключенного третьего — либо «за» (A), либо «против» (B). Но возникает потребность в варианте «воздержался» (рис. 4.3).

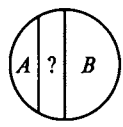


Рис. 4.3. Схема для трехзначной логики

В природе и обществе существуют так называемые случайные явления: результаты голосования на выборах, стихийные бедствия, результаты экспериментов, поиска, в которых недостаточно двумерного аристотелева языка оценок явлений типа «холодно—жарко». Нельзя однозначно оценивать людей по принципу киногероев: хороший — плохой. Часто

приходится искать третью характеристику. Например, противоположные умозаключения (см. гл. 5): индукция (от частного к общему) и дедукция (от общего к частному) не исключают и третьего вида — аналогии (от частного к частному). Та же лампочка, о которой нам хочется думать, что она имеет два режима: горит — не горит, имеет переходное время. Просто это время у лампочки — 0,1 с, а у электроплитки это время 5 мин, что уже сравнимо со временем ее эксплуатации, например кипячением воды.

Аналогичные проблемы возникли и в философии, и в математике (парадоксы теории множеств, подрад. 1.4), и в квантовой механике (корпускулярно-волновой дуализм), и в генетике (наряду с мужским и женским началом есть гермафродит). Как охарактеризовать, например, феномен белых, или полярных, ночей? Поэтому в последнее время, переосмыслив двузначную аристотелеву логику, в некоторых случаях используют закон «исключенного четвертого». Даже в тестах стали даваться не двузначные, а трехзначные варианты ответов.

Из этого закона можно сделать несколько выводов.

Не использовать закон исключенного третьего при оценке прогноза, связанного с природными явлениями. Надо анализировать предметную область применения этого закона.

При решении некоторой проблемы в условии неопределенности надо выявить все альтернативные варианты с помощью здравого смысла, логики, закона исключенного третьего (*четвертого* и т.д.) и комбинаторики.

Оценить вероятность осуществления каждой альтернативы и ее последствий, опираясь на здравый смысл, интуицию, логику и математику.

При выявлении альтернатив в житейских ситуациях надо стараться свести их к двум.Так, все булевы функции трех аргументов сводятся к функциям двух переменных, чего уже нельзя сказать о других классах функций.

Приведем некоторые требования, которые накладывает на наши мысли закон исключенного третьего.

При голосовании либо «за», либо «против» избегать варианта «воздержался».

При подсчете голосов, если возникает паритет мнений, необходимы повторные выборы и повторное голосование.

В ходе судебного расследования подозреваемый либо виновен, либо не виновен, но если вина на данный момент не доказана и есть сомнения, то дело отдается на доследование.

В медицинской практике пациент либо болен, либо не болен (здоров), но если диагноз не установлен, а симптомы болезни выявлены, то врач отправляет пациента на дообследование.

Во время учебы в учебном заведении учащийся либо знает конкретную тему, либо не знает. Если преподавателю в ходе ответа не

удалось определить факт наличия знаний, или не удалось четко определиться с оценкой, то он задает дополнительные вопросы.

При отстаивании позиций в споре необходимо определить интеллектуальные границы, в которых осуществляется осознанный поиск истинного вывода.

4. **Закона достаточного основания** в логике Аристотеля не было, его сформулировал Лейбниц.

Всякая истинная мысль должна быть достаточно обоснована. Ни одно суждение не может считаться истинным без достаточного на то основания.

Правила применения закона достаточного основания.

В доказательствах можно использовать только истинные или обоснованные суждения.

Обоснования суждений могут быть логического и фактического характера.

Аргументами в обосновании могут служить:

- истинные суждения, аксиомы, нормы морали;
- фактический материал;
- законы науки, законы общества;
- теоремы.

Для обоснования применяются законы тождества, непротиворечивости и исключенного третьего, а также правила построения дедуктивных умозаключений (см. подразд. 5.5). Выводы в таком случае будут *достоверными.*

Если же при обосновании кроме перечисленных законов логики использовались построения по аналогии и индуктивные, то выводы будут *вероятностными* (подробно см. в подразд. 5.6 и 5.7).

Закон достаточного основания — логический, но он имеет прямое отношение к общефилософскому закону о причинно-следственных связях. В природе, обществе и любой науке, как в отражении законов природы и общества, все взаимосвязано. Любое событие, явление, суждение имеет свои причины и порождает следствия, — это объективные законы действительности. В XVIII в. М. В. Ломоносов в работе «Элементы математической химии» использовал аксиому: «Ничто не происходит без достаточного основания».

Поскольку мышление есть отражение действительности в нашем сознании, то и наши рассуждения должны подчиняться этому правилу. Говорят, что человек мыслит логично тогда, когда цепочка рассуждений от причины к следствию неразрывна при переходе от одного суждения к другому. Выражения «выстроить логический ряд» и «логическую цепочку» означают, что внутри одного рассуждения мысли должны быть связаны друг с другом причинно-следственными связями.

Но из-за двусмысленности толкования слова «следовательно» философское и житейское применение этого термина не всегда

совпадает. Классический пример: в природе «*Пошел дождь, и (поэтому) крыша стала мокрой*», тогда как в жизни, посмотрев в окно, мы сделали вывод: «*Крыша стала мокрой, значит, пошел дождь*».

Чтобы избежать ошибок в заключениях, нужно придерживаться простых правил.

Вывод, основанный на ложных посылках, — ложен, точнее, не определен и не может служить дальнейшим аргументом.

При нарушении связи между единичным и общим вопреки закону достаточного основания делается поспешное обобщение (не все посылки учтены) и происходит путаница между причинной связью и элементарной последовательностью во времени.

Так, не является достоверным вывод «Студент получил на экзамене двойку, значит, не выучил заданные вопросы». На самом деле студент не выучил вопросы и поэтому на экзамене получил двойку. А первое утверждение неправильно, так как он мог, например, выучить билеты, но нарушить дисциплину, передавая кому-то шпаргалки.

Цель познания — достижение истины. В ходе информационного обмена между участниками процесса возникает потребность в доказательстве этой истины. Доказательство непосредственно связано с аргументацией, обоснованием суждений.

Доказательство есть совокупность логических приемов, применяемых для обоснования истинности некоторого утверждения (суждения) с помощью уже установленных истин или аксиом в рамках некоторой формальной системы. Доказательные рассуждения характеризуют научный стиль мышления.

В структуру доказательства входят следующие понятия.

Тезис — суждение, истинность которого надо доказать.

Аргументы основания — истинные суждения, которые используются при доказательстве тезиса.

Демонстрация или форма доказательства — способ логической связи между тезисом и аргументами.

В математике тезисом является формулировка теоремы. Доказать можно лишь внутренне непротиворечивый тезис с помощью набора аксиом и правил логики.

Сформулируем требования к видам аргументов.

Определения. Не следует давать определения очевидным понятиям. Существуют неопределяемые понятия, например, в геометрии за неопределяемые понятия можно принять точку, пространство, в математике — множество, элемент, соответствие, в жизни — любовь и т.д.

Не следует считать неопределяемыми те понятия, которым можно дать определения через неопределяемые и уже введенные понятия. Например, в математике: «Параллельными называются прямые, лежащие в одной плоскости и *не имеющие общих точек*». Сравните с ошибочным определением, в котором используется

не определенный ранее термин «пересекающиеся»: «Параллельными называются прямые, лежащие в одной плоскости и не *пересекающиеся* между собой».

Аксиомы. Аксиома — это суждение, которое в рамках некоторой науки или теории принимается истинным без доказательств.

Число аксиом должно быть необходимо и достаточно для выстраивания рассуждений.

Одна аксиома не должна зависеть от всех других.

Система аксиом должна быть внутренне непротиворечивой.

Система аксиом должна быть полной. Это означает, что можно доказать любую истинную формулу с помощью имеющегося набора.

Примером полной теории может служить исчисление высказываний (см. подразд. 5.2). Неполными теориями являются арифметика и теория множеств.

Факты. В различных науках, в юриспруденции и т.д. фактический материал подтверждает или опровергает тезис. Для опровержения тезиса достаточно иметь *один* противоречащий ему факт. Например, формула простых чисел $n = x^2 + x + 41$ неверна при $x = 41$. О факте можно говорить лишь в прошедшем времени: он свершился, и есть *достаточное* количество надежных свидетелей, могущих это подтвердить. Вопрос заключается в том, какого свидетеля считать надежным. Пусть, например, над средневековым Базелем пролетает самолет. Принципиальная возможность существования самолетов, разумеется не в средние века, доказана действительностью. Тогда весь город, скажем, десять тысяч человек, бóльшая часть из которых находилась в здравом уме, скажут, что они были свидетелями божественного вмешательства: пролета нечистой силы или, наоборот, ангела-хранителя. Так же и в физике, где нет оснований считать людей глупыми, любой эксперимент не является таковым, пока не будет получена удовлетворительная трактовка его результатов в рамках существующих теорий. Если ее нет, то появляются новые теории, способные этот эксперимент объяснить. Так, ньютонова механика могла объяснить все явления, пока не стали изучать электромагнитные явления.

Теоремы. В науке новые теоремы опираются на уже доказанные.

Кассификация видов рассуждений в процессе доказательства представлена на рис. 4.4.

Критика или опровержение — логическая операция установления ложности или необоснованности тезиса. Ложность тезиса доказывается с помощью аргументов опровержения. Опровержение должно показать неправильно построенное доказательство или ложность либо недоказуемость выдвинутого тезиса.

Критика может быть деструктивной и конструктивной.

Все способы опровержения демонстрируют **деструктивную** критику, так как они, выявляя недостатки, ничего не предлагают взамен. В искусстве критика — явление уникальное и необходимое.

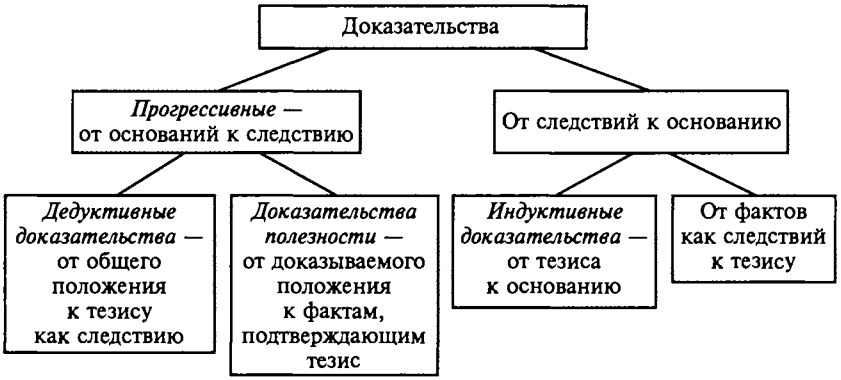


Рис. 4.4. Виды рассуждений в процессе доказательства

В отношении новаторов критика часто имеет деструктивный характер. История знает тысячи примеров откровенной травли гения, ломающего стереотипы.

Конструктивная критика предлагает взамен опровергаемого тезиса альтернативный. При этом альтернативный тезис сопровождается необходимой и достаточной аргументацией.

Отметим частный вид критики в современных политтехнологиях. Здесь конструктивную критику позволяют себе единицы, и, как правило, до избирателя она не доходит. А вот деструктивная, называемая черным пиаром (от англ. *PR — public relations* — общественные связи), встречается очень часто. Как правило, она носит провокационный и дискредитирующий характер.

4.5. Логика вопросов и ответов

Есть у меня шестерка слуг,
Проворных, удалых.
И все, что вижу я вокруг, —
Все знаю я от них.
Они по знаку моему
Являются в нужде.
Зовут их: Как и Почему,
Кто, Что, Когда и Где.

Р. Киплинг

Велика роль вопроса в процессе познания истины, ведь путь от незнания к знанию идет через постановку вопросов и поиска ответов на них.

Этот процесс состоит из трех этапов: постановки вопроса, поиска новой информации, конструирования ответа.

Грамматической формой вопроса является вопросительное предложение.

Вопрос — это выраженная в вопросительном предложении мысль, направленная на уточнение или дополнение знаний.

Характерные особенности вопросов состоят в том, что, во-первых, нет семантической характеристики (нет значений истины и ложности), во-вторых, опора делается на начальную *исходную* информацию (предпосылка вопроса).

В зависимости от содержания предпосылки, ее истинности и непротиворечивости вопросы можно классифицировать на корректные, некорректные и риторические.

Некорректные, *провокационные* вопросы, несмотря на противоречивую основу, также применяют, например, в судопроизводстве, а также во время различных дискуссий, диспутов с целью вызвать участников на обсуждение.

С появлением компьютеров традиционная тема классической логики, логика вопросов и ответов, приобрела новое значение. Проблема «обучения» машины правилам общения с человеком стоит не только перед создателями ЭВМ. Ведь обучить машину ведению диалога оказалось легче, чем научить этому всех пользователей ЭВМ. А грамотное ведение диалога становится необходимым для того, чтобы машина и человек говорили на одном, причем правильном, логически выверенном языке. Современный синтетический подход к диалогу с ЭВМ заключается в максимальном приближении этого диалога к естественно-речевому общению, что увеличивает эффективность восприятия информации в системе человек — машина. Рассмотрим возможные варианты такого диалога.

ЭВМ самостоятельно генерирует диалог согласно алгоритмам, заложенным разработчиками на основе строгих законов логики, т.е. самостоятельно моделирует ответы на реакцию партнера (человека).

Вполне определенная модель беседы программируется заранее, а ЭВМ ведет диалог согласно этой программе.

В компьютерном диалоге ведущая роль принадлежит вопросам. Общий вид структуры вопроса включает известную информацию, т.е. знания как логические предпосылки, неизвестный материал, логический переход от знания к незнанию и наоборот.

Рассмотрим виды и *познавательные* функции вопросов (табл. 4.18).

Правила постановки простых и сложных вопросов.

Корректность постановки вопроса. В обычной речи недопустимы провокационные, риторические, некорректные вопросы.

Альтернативность ответов на уточняющие вопросы должна быть предусмотрена в самом вопросе.

Перечисление всех альтернатив сложных дизъюнктивных вопросов.

Краткость, ясность и простота формулировок.

Таблица 4.18

Виды и познавательные функции вопросов

Виды вопросов	Грамматический признак	Определение	Характеристика	Символическая схема записи	Область поиска ответа	Структура вопросов				
						Простые		Сложные		Смешанные
						Условный	Безусловный	Конъюнктивный	Дизъюнктивный	
Уточняющие	«*ли*-вопрос»: верно ли; надо ли; правда ли?	Направлены на выявление истинности суждения, в нем содержащегося	Прямые, определенные, закрытые	$?(P)$, где $?$ — оператор вопроса; P — суждение, про которое надо уточнить его истинность	$P_1 \vee P_2$	$?(p \to q)$	$?(p)$	$?(p \wedge q)$ (верно ли p и верно ли q)	$?(p \vee q)$ (верно ли p или q)	$?(p \wedge q) \vee (m \wedge n)$; $?(p \vee q) \wedge (m \vee n)$ (верно ли p и q или m и n)
Восполняющие	«*что*-вопрос»; «*когда*-вопрос»: кто; что; где; когда; как; почему?	Направлены на выявление новых свойств исследуемого явления	Неопределенные, непрямые	$?Q(P)$, где $?$ — оператор вопроса; Q — переменная для вопросительного слова; P — суждение, нуждающееся в дополнении	Выбор суждений из множества возможных	—	—	$?(q_1 \wedge \wedge q_2)(p)$ $?q(p \wedge q)$	$?(q_1 \vee \vee q_2)(p)$ $?q(p \vee q)$ $?(Q_1 \wedge Q_2)$ $(Q_1 \wedge Q_2)$	$?Q_1(p \wedge q) \vee \vee Q_2(m \wedge n)$ $?Q_1(p \vee q) \wedge \wedge Q_2(m \vee n)$ (где могут быть обнаружены p и q или когда появятся m и n)

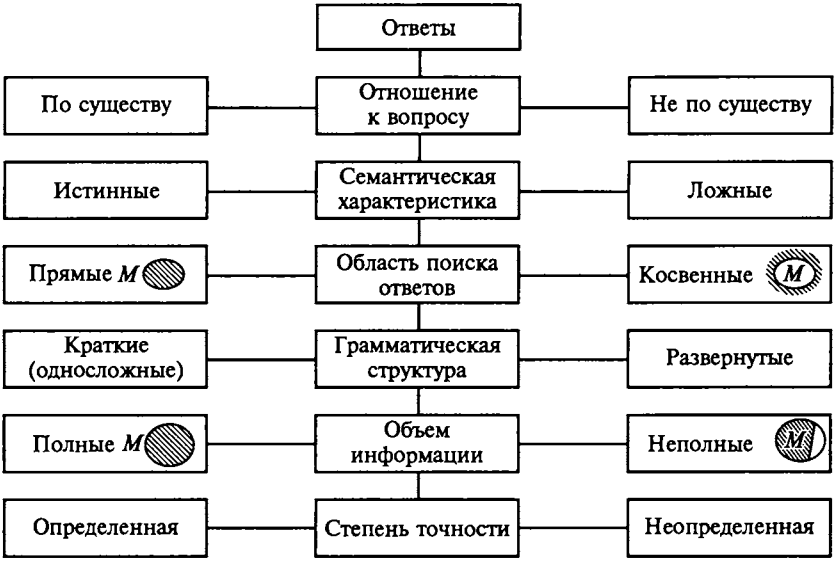


Рис. 4.5. Классификация ответов

Ответом называют новое суждение, уточняющее или дополняющее в соответствии с вопросом прежние знания. Представления о видах ответов можно получить по следующей классификации, где в каждом отдельном случае указаны основания для дихотомического деления (рис. 4.5).

Велика роль вопросов и ответов в процессе познания. Родоначальник индуктивной логики Фрэнсис Бэкон говорил, что мы должны уметь ставить вопросы природе.

Не менее важна точность формулировок вопросов и ответов при работе с компьютером. Диалоговый режим общения с ЭВМ обязывает человека задавать компьютеру корректные и правильно сформулированные вопросы, предполагающие однозначные ответы на них. В компьютерном диалоге большое значение имеет соответствие ответа содержанию вопроса, так называемая адекватность ответов. А для того чтобы человек говорил с ЭВМ на одном языке, он должен знать этот язык.

В настоящее время логическая теория вопросов и ответов интенсивно развивается именно в связи с применением ЭВМ. Так, современная компьютерная программа «Персональный доктор» основана на постановке диагноза «врачом-компьютером», который следует после диалога ЭВМ с больным. Разработана специальная программа, предусматривающая возможные виды заболеваний с учетом реальных симптомов пациентов. Диалог ЭВМ с больным человеком заканчивается тем, что компьютер выдает конкретный диагноз и рекомендации по лечению. Конечно, он не

может лечить или оперировать больного, но он помогает врачу определиться с выбором направления лечения. Причем скорость реакции «врача-компьютера» и энциклопедические знания, заложенные в него человеком, делают его достойным помощником врача-человека. Но для работы с ним необходимо уметь правильно задавать вопросы и находить адекватную форму ответа.

4.6. Минимизация булевых функций

Если исключить невозможное, то, что останется, сколь бы невероятным оно ни было, должно быть истиной.

А.Конан-Дойль

Прежде не раз упоминалась возможность представления любой булевой функции в виде суперпозиции булевых функций одного и двух аргументов. Кроме того, хотелось, чтобы подобное разложение включало в себя простейшие с точки зрения интуитивного понимания операции: отрицание, конъюнкцию и дизъюнкцию. Так, в подразд. 4.3 строгая дизъюнкция, импликация и эквиваленция были выражены через эти три элементарные функции. Кроме того, упоминалась двойственность конъюнкции и дизъюнкции. Подобное представление полезно, например, в электротехнике, где микросхема реализует одну из этих простейших операций. Далее мы докажем, что любую булеву функцию можно выразить через отрицание (НЕ), конъюнкцию (И) и дизъюнкцию (ИЛИ), а пока будем считать, что функция уже эквивалентна композиции этих трех функций.

4.6.1. Разложение функций по переменным. Нормальные формы

Рассмотрим булевы функции, представленные в виде суперпозиции элементарных функций И, ИЛИ, НЕ. Используя законы алгебры логики, можно заменить громоздкие булевы функции им равносильными, но более простыми. Такой процесс называется **минимизацией** булевых функций. Его проводят для упрощения сложных логических выражений в программах, а также для того, чтобы построенные на их основе функциональные схемы не содержали лишних элементов.

Минимизировать булевы функции надо, приводя их к так называемой *нормальной форме.*

Существуют две разновидности нормальных форм — дизъюнктивные (ДНФ) и конъюнктивные (КНФ). Введем определения.

Элементарной $\frac{\text{конъюнкцией}}{\text{дизъюнкцией}}$ называется выражение, состоящее из конечного числа переменных и их отрицаний, взятых в этом выражении не более одного раза и разделенных операциями $\frac{\text{конъюнкции}}{\text{дизъюнкции}}$:

$\bigwedge_{i=1}^{s} y_i$, где $y_i = x_i$ или $\overline{x_i}$, например $x_1x_2\overline{x_3}$ — элементарная конъюнкция;

$\bigvee_{i=1}^{s} y_i$, где $y_i = x_i$ или $\overline{x_i}$, например $x_1 \vee x_2 \vee \overline{x_3}$ — элементарная дизъюнкция.

$\frac{\text{Дизъюнктивной}}{\text{Конъюнктивной}}$ **нормальной формой** называется $\frac{\text{дизъюнкция}}{\text{конъюнкция}}$ конечного числа элементарных $\frac{\text{конъюнкций}}{\text{дизъюнкций}}$. Сокращенно они обозначаются ДНФ и КНФ соответственно.

Нормальная форма называется **совершенной**, если в каждой ее элементарной дизъюнкции (конъюнкции) представлены все переменные, входящие в данную функцию (либо сами, либо с отрицанием).

Любая булева функция и любая формула алгебры логики могут быть представлены множеством различных *дизъюнктивных форм*, равносильных между собой. Например: $F_1(x_1, x_2, x_3) = x_1 \vee x_1x_3 \vee \vee x_1x_2x_3 = x_1 \vee x_1x_2x_3 = x_1 \vee \overline{x_2}\overline{x_3} = x_1x_2\overline{x_3} \vee x_1\overline{x_2}x_3$ и т. д.

Из всех различных ДНФ функции $F_1(x_1, x_2, x_3)$ особо выделяется последнее логическое выражение, которое является **совершенной** ДНФ, сокращенно СДНФ. На СДНФ накладываются следующие требования:

- формула не является тождественно-ложной;
- формула приведена к одному из видов ДНФ;
- из формулы удалены элементарные конъюнкции, включающие одновременно переменную и ее отрицание, согласно закону инверсии;
- из формулы удалены одинаковые элементарные конъюнкции, кроме одной, согласно правилу идемпотентности;
- каждая элементарная конъюнкция в ДНФ включает все логические переменные, входящие в эту формулу.

Если в логической функции не выполняется последнее требование, то в неполную элементарную конъюнкцию необходимо ввести дополнительный множитель, включающий дизъюнкцию отсутствующей переменной и ее отрицание. Это всегда можно сделать, так как согласно закону инверсии $\overline{a} \vee a = 1$, а $1 \cdot a = a$.

Так, в рассмотренном примере $x_1 \vee x_1\overline{x_2} \cdot \overline{x_3}$ является ДНФ. Для того чтобы привести ее к СДНФ, необходимо в первой элементарной конъюнкции иметь переменные x_2 и x_3 или их отрицания. Для этого дважды умножим x_1 на 1, с тем чтобы затем эти единицы заменить дизъюнкциями $x_2 \vee \overline{x_2}$ и $x_3 \vee \overline{x_3}$ соответственно: $F_1(x_1, x_2, x_3) = x_1 \vee x_1\overline{x_2}\overline{x_3} = x_1 \cdot 1 \cdot 1 \vee x_1\overline{x_2}\overline{x_3} = x_1(x_2 \vee \overline{x_2})(x_3 \vee \overline{x_3}) \vee \vee x_1x_2x_3 = x_1x_2x_3 \vee x_1x_2\overline{x_3} \vee x_1\overline{x_2}x_3 \vee x_1\overline{x_2}\overline{x_3} \vee x_1x_2x_3 = x_1x_2x_3 \vee x_1x_2\overline{x_3} \vee \vee x_1\overline{x_2}x_3 \vee x_1\overline{x_2}\overline{x_3}$.

Полученная ДНФ является совершенной, так как соответствует всем перечисленным требованиям.

Аналогично любую формулу, имеющую вид ДНФ, можно привести к совершенной конъюнктивной нормальной форме (КНФ), для которой выполняются требования:

- формула не является тождественно-ложной;
- формула приведена к одному из видов КНФ;
- из формулы удалены одинаковые элементарные дизъюнкции, кроме одной;
- каждая элементарная дизъюнкция в КНФ включает все логические переменные, входящие в эту формулу.

Если логическая функция имеет вид КНФ, то привести ее к виду совершенной КНФ (сокращенно СКНФ) можно, дополнив каждую элементарную дизъюнкцию логическим нулем, который в следующем шаге заменяется на конъюнкцию недостающей переменной и ее отрицание.

Для примера рассмотрим $F_2(x_1, x_2, x_3) = (\overline{x_1} \vee x_3) \cdot (x_2 \vee \overline{x_3}) \cdot (x_1 \vee \vee \overline{x_2} \vee x_3)$. Она имеет вид КНФ, но в первой скобке нет переменной x_2, а во второй — x_1. Воспользовавшись приведенным правилом, приведем ее к виду СКНФ. Тогда имеем: $F_2(x_1, x_2, x_3) = (\overline{x_1} \vee \vee 0 \vee x_3) \cdot (0 \vee x_2 \vee \overline{x_3}) \cdot (x_1 \vee \overline{x_2} \vee x_3) = (\overline{x_1} \vee x_2\overline{x_2} \vee x_3) \cdot (\overline{x_1}x_1 \vee \overline{x_2} \vee \overline{x_3}) \cdot \cdot (\overline{x_1} \vee \overline{x_2} \vee x_3) = (\overline{x_1} \vee x_2 \vee x_3) \cdot (\overline{x_1} \vee \overline{x_2} \vee x_3) \cdot (x_1 \vee x_2 \vee \overline{x_3}) \cdot (\overline{x_1} \vee x_2 \vee \vee \overline{x_3}) \cdot (x_1 \vee \overline{x_2} \vee x_3)$.

Полученная конъюнктивная нормальная форма является совершенной.

Для сравнения приведем примеры нормальных форм (табл. 4.19).

Для того чтобы привести булеву функцию к совершенной нормальной форме, надо выполнить операции в следующем порядке.

Таблица 4.19

Примеры нормальных форм

Формы	Конъюнктивные	Дизъюнктивные
Нормальные	КНФ $(x_1 \vee \overline{x_2}) \cdot (\overline{x_1} \vee x_3)$	ДНФ $x_1x_2 \vee \overline{x_1}x_3$
Совершенные нормальные	СКНФ $(x_1 \vee \overline{x_2} \vee x_3) \cdot (\overline{x_1} \vee x_2 \vee \overline{x_3})$	СДНФ $x_1\overline{x_2}x_3 \vee \overline{x_1}x_2\overline{x_3} \vee x_1x_2\overline{x_3}$

Используя равносильности алгебры логики, заменить все имеющиеся операции на конъюнкцию, дизъюнкцию и отрицание.

Применяя законы Де Моргана, снять отрицание с логических операций конъюнкции и дизъюнкции.

Используя распределительный и другие законы, привести функции к нормальной форме.

Используя законы идемпотентности, склеивания и др., минимизировать полученные булевы выражения.

Применяя правила операций с константами, привести минимизированные нормальные формы к совершенному виду.

Рассмотрим пример. Привести булеву функцию $F = (x_1, x_2, x_3) = \overline{x_1 \to x_2 x_3} \cdot (x_2 \to (x_1 \to x_3))$ к совершенной нормальной форме.

Сначала приведем функцию к виду СКНФ. Имеем: $F(x_1, x_2, x_3) = \overline{x_1 \to x_2 x_3} \cdot (x_2 \to (x_1 \to x_3)) = \overline{\overline{x_1} \vee x_2 x_3} \cdot (\overline{x_2} \vee (\overline{x_1} \vee x_3)) = (\overline{\overline{x_1}} \cdot \overline{x_2 x_3}) \cdot (\overline{x_2} \vee \overline{x_1} \vee x_3) = x_1(\overline{x_2} \vee \overline{x_3}) \cdot (\overline{x_2} \vee \overline{x_1} \vee x_3) = (x_1 \vee 0) \cdot (0 \vee \overline{x_2} \vee \overline{x_3}) \cdot (x_2 \vee \overline{x_1} \vee x_3) = (x_1 \vee x_2 \overline{x_2}) \cdot (x_1 \overline{x_1} \vee \overline{x_2} \vee \overline{x_3}) \cdot (\overline{x_2} \vee \overline{x_1} \vee x_3) = (x_1 \vee \overline{x_2}) \cdot (x_1 \vee \overline{x_2}) \cdot (x_1 \vee \overline{x_2} \vee \overline{x_3}) \cdot (x_1 \vee x_2 \vee \overline{x_3}) \cdot (x_1 \vee \overline{x_2} \vee \overline{x_3}) \cdot (\overline{x_2} \vee \overline{x_1} \vee x_3) = (x_1 \vee x_2 \vee x_3 x_3) \cdot (x_1 \vee \overline{x_2} \vee x_3 \overline{x_3}) \cdot (x_1 \vee \overline{x_2} \vee \overline{x_3}) \cdot (x_1 \vee x_2 \vee \overline{x_3}) \cdot (\overline{x_2} \vee \overline{x_1} \vee x_3) = (x_1 \vee \vee \overline{x_2} \vee x_3) \cdot (x_1 \vee x_2 \vee \overline{x_3}) \cdot (x_1 \vee \overline{x_2} \vee x_3) \cdot (x_1 \vee x_2 \vee x_3) \cdot (\overline{x_1} \vee \overline{x_2} \vee \overline{x_3}) \cdot (\overline{x_2} \vee \overline{x_1} \vee x_3)$.

Приведем упрощенную функцию к виду СДНФ:

$F(x_1, x_2, x_3) = x_1(\overline{x_2} \vee \overline{x_3}) \cdot (\overline{x_1} \vee \overline{x_2} \vee x_3) = (x_1 \overline{x_2} \vee x_1 \overline{x_3}) \cdot (\overline{x_1} \vee \overline{x_2} \vee x_3) = x_1 \overline{x_2} x_1 \vee x_1 \overline{x_2} x_2 \vee x_1 \overline{x_2} x_3 \vee x_1 x_3 \overline{x_1} \vee x_1 x_3 \overline{x_2} \vee x_1 x_3 x_3 = x_1 \overline{x_2} \vee x_1 \overline{x_2} x_3 \vee \vee x_1 \overline{x_2} x_3 = x_1 \overline{x_2}(x_3 \vee \overline{x_3}) \vee x_1 \overline{x_2} x_3 \vee x_1 \overline{x_2} x_3 = x_1 \overline{x_2} x_3 \vee x_1 \overline{x_2} x_3$.

Полученное булево выражение действительно имеет вид СДНФ, так как каждая элементарная конъюнкция содержит все переменные этой булевой функции, которые встретились по одному разу. Зная СДНФ, можно составить таблицу истинности соответствующей функции.

Пусть $F(x_1, x_2, x_3) = x_1 \overline{x_2} x_3 \vee x_1 x_2 \overline{x_3}$. Для получения таблицы истинности подставим в функцию значения логических переменных в каждой строке и, выполнив действия с логическими константами, найдем значение функции. Это гораздо проще, чем расписывать каждый столбец в исходной функции (табл. 4.20).

Совершенные нормальные формы нужны, например, в тех содержательных задачах, где необходимо рассмотреть все возможные варианты информации о каждой из переменных, входящих в данную булеву функцию.

Любая булева функция, не являющаяся тождественным нулем или единицей, имеет только одну СДНФ с точностью до расположения переменных.

Мы рассмотрели примеры, когда булева функция была уже выражена через логические операции. Теперь посмотрим, как составить совершенные нормальные формы тогда, когда известна лишь таблица значений.

Таблица 4.20

Таблица истинности функции F

x_1	x_2	x_3	$F(x_1, x_2, x_3)$
0	0	0	$F(0, 0, 0) = 0 \cdot \overline{0} \cdot 0 \vee 0 \cdot 0 \cdot \overline{0} = 0 \vee 0 = 0$
0	0	1	$F(0, 0, 1) = 0 \cdot \overline{0} \cdot 1 \vee 0 \cdot 0 \cdot \overline{1} = 0 \vee 0 = 0$
0	1	0	$F(0, 1, 0) = 0 \cdot \overline{1} \cdot 0 \vee 0 \cdot 1 \cdot \overline{0} = 0 \vee 0 = 0$
0	1	1	$F(0, 1, 1) = 0 \cdot \overline{1} \cdot 1 \vee 0 \cdot 1 \cdot \overline{1} = 0 \vee 0 = 0$
1	0	0	$F(1, 0, 0) = 1 \cdot \overline{0} \cdot 0 \vee 1 \cdot 0 \cdot \overline{0} = 0 \vee 0 = 0$
1	0	1	$F(1, 0, 1) = 1 \cdot \overline{0} \cdot 1 \vee 1 \cdot 0 \cdot \overline{1} = 1 \vee 0 = 1$
1	1	0	$F(1, 1, 0) = 1 \cdot \overline{1} \cdot 0 \vee 1 \cdot 1 \cdot \overline{0} = 0 \vee 1 = 1$
1	1	1	$F(1, 1, 1) = 1 \cdot \overline{1} \cdot 1 \vee 1 \cdot 1 \cdot \overline{1} = 0 \vee 0 = 0$

Задача 21. Пусть при $n = 3$ булева функция задана таблицей истинности (табл. 4.21). Составить СДНФ и СКНФ для данной функции.

Решение. Каждому набору переменных (строке) поставим в соответствие элементарную конъюнкцию, причем каждому аргументу $x_k = 1$ будет соответствовать x_k, а каждому $x_k = 0$ будет соответствовать $\overline{x_k}$ в соответствующей элементарной конъюнкции.

Составим булеву функцию, объединив дизъюнкцией те элементарные конъюнкции, которые дают значение соответствующего булева выражения, равное единице: $F(x_1, x_2, x_3) = \overline{x_1}x_2\overline{x_3} \vee x_1\overline{x_2}x_3 \vee \vee x_1x_2\overline{x_3}$. В результате получили СДНФ. Это правило справедливо для любой булевой функции.

Чтобы построить СКНФ для некоторой функции F, надо построить $\overline{F}$ в СДНФ, т. е. взять дизъюнкцию элементарных конъюн-

Таблица 4.21

Таблица истинности булевой функции трех переменных

x_1	x_2	x_3	$F(x_1, x_2, x_3)$	Элементарные конъюнкции
0	0	0	0	$\overline{x_1} \cdot \overline{x_2} \cdot \overline{x_3}$
0	0	1	0	$\overline{x_1} \cdot \overline{x_2} \cdot x_3$
0	1	0	1	$\overline{x_1} \cdot x_2 \cdot \overline{x_3}$
0	1	1	1	$\overline{x_1} \cdot x_2 \cdot x_3$
1	0	0	0	$x_1 \cdot \overline{x_2} \cdot \overline{x_3}$
1	0	1	0	$x_1 \cdot \overline{x_2} \cdot x_3$
1	1	0	1	$x_1 \cdot x_2 \cdot \overline{x_3}$
1	1	1	0	$x_1 \cdot x_2 \cdot x_3$

кций, дающих значение, равное нулю, и взять отрицание от $\overline{F}$, так как $\overline{\overline{F}} = F$.

В нашем случае $\overline{F}(x_1, x_2, x_3) = \overline{x}_1\overline{x}_2\overline{x}_3 \vee \overline{x}_1\overline{x}_2x_3 \vee x_1\overline{x}_2\overline{x}_3 \vee x_1\overline{x}_2x_3 \vee x_1x_2x_3$. С помощью закона Де Моргана отрицание дизъюнкции превратим в конъюнкцию отрицаний и минимизируем выражение, сохраняя его вид, т.е. конъюнктивную форму. Тогда

$$F(x_1, x_2, x_3) = \overline{\overline{F(x_1, x_2, x_3)}} =$$

$$= \overline{\overline{x}_1\overline{x}_2\overline{x}_3 \vee \overline{x}_1\overline{x}_2x_3 \vee x_1\overline{x}_2\overline{x}_3 \vee x_1\overline{x}_2x_3 \vee x_1x_2x_3} =$$

$$= \overline{\overline{x}_1\overline{x}_2\overline{x}_3} \cdot \overline{\overline{x}_1\overline{x}_2x_3} \cdot \overline{x_1\overline{x}_2\overline{x}_3} \cdot \overline{x_1\overline{x}_2x_3} \cdot \overline{x_1x_2x_3} = (x_1 \vee x_2 \vee x_3) \cdot$$

$$\cdot (x_1 \vee x_2 \vee \overline{x}_3) \cdot (\overline{x}_1 \vee x_2 \vee x_3) \cdot (\overline{x}_1 \vee x_2 \vee \overline{x}_3) \cdot (\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3).$$

Полученное выражение является СКНФ данной функции $F(x_1, x_2, x_3)$.

Практическим построением СДНФ показано, что любая булева функция, не являющаяся константой, по таблице истинности разлагается в совершенную нормальную форму, т.е. в композицию отрицания, конъюнкции и дизъюнкции. А поскольку то же было сделано и для констант, то утверждение доказано для любой булевой функции произвольного числа переменных.

Существование СДНФ позволяет провести процедуру, называемую **разложением** булевой функции по переменной x_k. Разложение позволяет представить произвольную функцию $f(x_1, ..., x_n)$ в виде $f(x_1, ..., x_n) = x_k \cdot p(x_1, ..., x_{k-1}, x_{k+1}, ..., x_n) \vee \overline{x}_k \cdot q(x_1, ..., x_{k-1}, x_{k+1}, ..., x_n)$.

Это легко сделать, отделяя в СДНФ все слагаемые с x_k от $\overline{x}_k$ и вынося x_k или $\overline{x}_k$ за скобку. Тогда выражения в скобках будут некими функциями p и q, не зависящими от x_k. В задаче 21 имеем $F(x_1, x_2, x_3) = \overline{x}_1x_2x_3 \vee \overline{x}_1x_2x_3 \vee x_1x_2\overline{x}_3 = \overline{x}_1(x_2x_3 \vee x_2x_3) \vee x_1(x_2\overline{x}_3) = \overline{x}_1(x_2) \vee x_1(x_2\overline{x}_3)$. Следовательно, мы произвели разложение $F(x_1, x_2, x_3)$ по x_1. Аналогично можно производить разложение по группе переменных. Проще всего это можно сделать последовательным разложением по одной переменной, входящей в эту группу, а затем привести оставшиеся выражения (в частности, p и q) к СДНФ от меньшего числа переменных.

4.6.2. Логические схемы

Одной из целей булевой алгебры является описание поведения и структуры логических схем. **Логическая схема** имеет вид «черного ящика», в котором вход — набор булевых переменных, а выход — булева функция $F(x_1, ..., x_n)$ (рис. 4.6).

Метод «черного ящика» используется в тех случаях, когда предметом изучения является поведение сложных систем, а не их устройство.

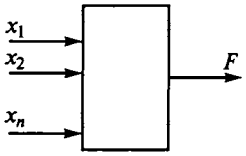


Рис. 4.6. Логическая схема

Исследователя интересуют лишь входные и выходные сигналы, а не процессы, происходящие внутри самого устройства. Впервые понятие «черный ящик» ввел английский ученый У. Р. Эшби для изучения отношений между экспериментом и окружающей средой, когда предметом исследования служат потоки информации.

Для того чтобы описать поведение «черного ящика», достаточно выразить выход F в виде функции от переменных $x_1, x_2, ..., x_n$ или построить истинные выражения, соответствующие логической связи между входными переменными, или минимизировать аналитическую формулу этих связей.

Примерами логических схем служат обыкновенные микросхемы, которые в большом количестве присутствуют в электробытовых приборах и компьютерах. В электротехнике принята маркировка микросхем по той функции, которую они реализуют. Она имеет вид И—НЕ, 2И—НЕ и т.д. Если элемент имеет входное напряжение в пределах от 0 до 0,4 В, то оно рассматривается как логический 0, если напряжение в пределах от 0,7 до 1,5 В, то оно рассматривается как 1. Примерно такие же характеристики имеет выходное напряжение. Работу подобной микросхемы удобнее всего анализировать с помощью осциллографа. Логическая схема НЕ, реализующая отрицание, называется **инвертором**. Она переворачивает сигнал на экране осциллографа.

Комбинационная схема — это логическая схема, в которой значения входных переменных в данный момент времени полностью определяются значениями выходных переменных. С развитием вычислительной техники математическая логика оказалась тем инструментом, который дает возможность анализировать электрические цепи при проектировании ЭВМ. Логическая схема устройства основывается на объединении электронных элементов, реализующих конкретные логические операции. Процесс построения функциональных схем для разработки устройства ПК можно представить в виде следующего алгоритма.

1. Анализ функций.
2. Составление таблиц истинности по результатам п. 1.
3. Синтез логической функции по таблице истинности.
4. Минимизация полученной логической функции.
5. Построение логической схемы устройства по результатам п. 4.

В свою очередь, алгоритм синтеза логической функции имеет следующий вид.

1. В заданной таблице истинности находятся наборы переменных (строки), в которых $F(x_1, ..., x_n) = 1$.
2. Для каждого набора записывается конъюнкция всех входных переменных, значение которых равно 0.

3. Все полученные конъюнкции объединяются дизъюнкцией в логическую функцию и минимизируются.

Понятно, что такая логическая функция представляет собой ДНФ, минимизированную так, чтобы было меньше логических операций.

Задача 22. По заданной таблице истинности (табл. 4.22) найти логическую функцию.

Таблица 4.22

Заданная таблица истинности

x_1	x_2	x_3	$F(x_1x_2x_3)$
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	0

Решение. Найдем основные конъюнкции, исходя из истинных значений данной функции (табл. 4.23).

Таблица 4.23

Таблица основных конъюнкций

x_1	x_2	x_3	$F(x_1x_2x_3)$	Основные конъюнкции
0	0	0	1	$\overline{x_1} \cdot \overline{x_2} \cdot \overline{x_3}$
0	0	1	1	$\overline{x_1} \cdot \overline{x_2} \cdot x_3$
0	1	0	1	$\overline{x_1} \cdot x_2 \cdot \overline{x_3}$
0	1	1	0	
1	0	0	0	
1	0	1	1	$x_1 \cdot \overline{x_2} \cdot x_3$
1	1	0	0	
1	1	1	0	

Тогда $F(x_1x_2x_3) = \overline{x_1}\,\overline{x_2}\,\overline{x_3} \vee \overline{x_1}\,\overline{x_2}x_3 \vee \overline{x_1}x_2\overline{x_3} \vee x_1\overline{x_2}x_3$.

Минимизируем полученную формулу:

$F(x_1x_2x_3) = (\overline{x_1}\,\overline{x_2}\,\overline{x_3} \vee \overline{x_1}x_2\overline{x_3}) \vee (\overline{x_1}\,\overline{x_2}x_3 \vee x_1\overline{x_2}x_3) = \overline{x_1}\,\overline{x_3}(\overline{x_2} \vee x_2) \vee \vee \overline{x_2}x_3(\overline{x_1} \vee x_1) = \overline{x_1}\,\overline{x_3} \cdot 1 \vee \overline{x_2}x_3 \cdot 1 = \overline{x_1}\,\overline{x_3} \vee \overline{x_2}x_3$.

Задача 23. По заданной таблице истинности (табл. 4.24) составить логическую схему.

Таблица 4.24

Заданная таблица истинности

x_1	x_2	x_3	$F(x_1x_2x_3)$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Решение. Найдем основные конъюнкции и составим булеву функцию $F(x_1x_2x_3) = x_1\overline{x_2}x_3 \vee x_1x_2\overline{x_3} \vee x_1x_2x_3$.

Минимизируем результат:

$F(x_1x_2x_3) = (x_1\overline{x_2}x_3 \vee x_1x_2x_3) \vee x_1x_2\overline{x_3} = x_1x_3(\overline{x_2} \vee x_2) \vee x_1x_2\overline{x_3} = x_1x_3 \vee \vee x_1x_2\overline{x_3} = x_1(x_3 \vee x_2\overline{x_3}) = x_1(x_3 \vee x_2) = x_1x_3 \vee x_1x_2$.

Построим два варианта логических схем (рис. 4.7) по булеву выражению: *а* — $F(x_1x_2x_3) = x_1(x_3 \vee x_2)$; *б* — $F(x_1x_2x_3) = x_1x_3 \vee x_1x_2$.

Далее будем следовать электротехническому правилу: если в узле на рисунке сходятся три линии, то соединение есть, а если четыре — то нет, т.е. перекрещивающиеся линии не связаны, поэтому выколотую точку рисовать не будем.

Логическая схема на рис. 4.7, *а* состоит из двух элементов, а на рис. 4.7, *б* — из трех. Сравнивая логические схемы, построенные для одной и той же булевой функции, представленной в различных формах, видим, что более рациональна запись через КНФ.

Рассмотрим пример. По заданной логической схеме (рис. 4.8, *а*) составить булеву функцию и минимизировать ее до ДНФ.

Поставим в соответствие каждому входу булеву переменную x_1, x_2, x_3. Можно в произвольном порядке, но тогда итоговая функция будет определена по-другому, а именно с точностью до соответствующей подстановки аргументов.

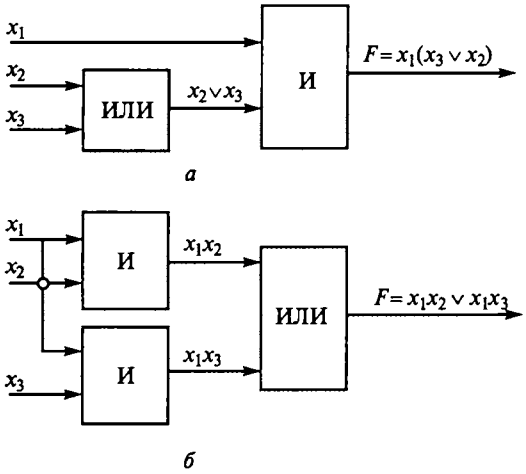


Рис. 4.7. Логическая схема булевой функции F:
а — $F(x_1x_2x_3) = x_1(x_3 \vee x_2)$; *б* — $F(x_1x_2x_3) = x_1x_3 \vee x_1x_2$

На выходах значения компонентов, составляющих булеву функцию, имеют вид: 1) $x_1 \vee x_3$; 2) $\overline{x_2}$; 3) $\overline{x_2} \vee x_3$; 4) $\overline{x_3}$; 5) $x_1 \vee \overline{x_3}$; 6) $(x_1 \vee \vee x_3)(x_1 \vee \overline{x_3})(\overline{x_2} \vee x_3)$.

Минимизируем результат, доведя его до ДНФ: $F(x_1x_2x_3) = (x_1 \vee \vee x_3)(x_1 \vee \overline{x_3})(\overline{x_2} \vee x_3) = x_1(\overline{x_2} \vee x_3) = x_1\overline{x_2} \vee x_1x_3$. Из полученного

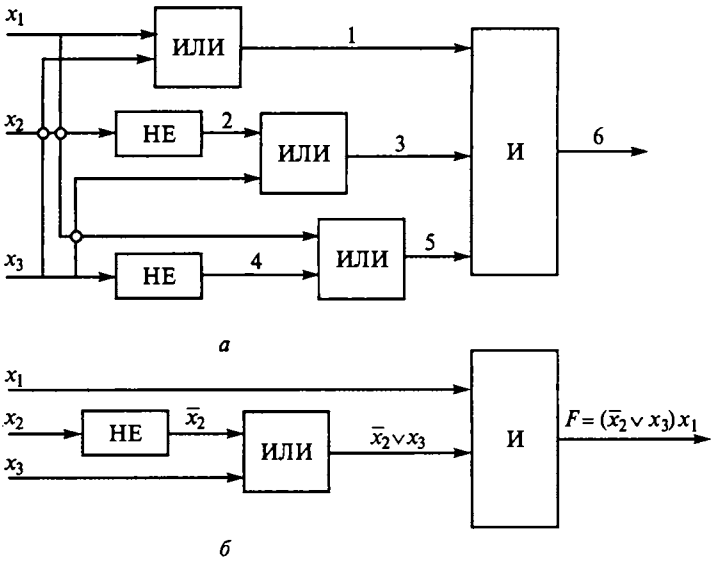


Рис. 4.8. Логические схемы булевой функции F:
а — заданная схема; *б* — минимизированная схема

результата видим, что эта же булева функция может быть задана другой схемой, содержащей всего три элемента (рис. 4.8, *б*).

Потребность в минимизации булевых функций возникает как при синтезе логических схем (в целях экономии), так и при решении логических задач для поиска минимального числа переменных и связок между ними. При решении логических задач ответ удобнее представлять или в СДНФ, или в виде КНФ. Если результат должен быть представлен в виде условного предложения (импликации), то переход к ней возможен из дизъюнкции.

4.6.3. Карты Карно

Минимизировать нормальные формы можно различными способами: методом каскадов, с помощью карт Карно и другими. **Минимальная** или **сокращенная** нормальная форма получается из совершенной $\frac{\text{конъюнктивной}}{\text{дизъюнктивной}}$ нормальной формы удалением некоторых элементарных $\frac{\text{дизъюнкций}}{\text{конъюнкций}}$. **Тупиковой** нормальной формой называется $\frac{\text{КНФ}}{\text{ДНФ}}$, из которой нельзя удалить ни одной элементарной $\frac{\text{конъюнкции}}{\text{дизъюнкции}}$ так, чтобы сохранить неизменной заданную булеву функцию. Для представления булевой функции в таком виде необходимо сначала представить ее в совершенном виде и только затем минимизировать до минимальной из всех тупиковых форм.

Карты Карно являются одним из наиболее удобных способов минимизации. Они впервые появились в одной из статей Мориса Карно в 1953 г. Это специальные таблицы, дающие возможность упростить процесс поиска минимальной формы булева выражения с помощью графического представления для $n \le 6$. Они имеют вид прямоугольника, разделенного на 2^n клеток, в каждой из которых — двоичный n-мерный набор значений функции F из таблицы истинности.

Для $n = 2$ карта Карно имеет вид таблицы, состоящей из $2^2 = 4$ клеток.

x_2 \ x_1		
	$\overline{x_1}\,\overline{x_2}$	$x_1\overline{x_2}$
	$\overline{x_1}x_2$	x_1x_2

или

00	10
01	11

Логическая функция F на карте Карно представлена совокупностью клеток, заполненных единицами (1) или пустотами (0), если известны ее значения при всем наборе аргументов, т.е. известна

таблица истинности или СДНФ. При $n = 3$ карты Карно имеют вид таблицы с $2^3 = 8 = 2 \cdot 4$ клетками (табл. 4.25).

Таблица 4.25

Карта Карно для булевых функций трех переменных

	Изменение x_2x_3			
Изменение x_1	$\overline{x}_1\overline{x}_2\overline{x}_3$ $x_1\overline{x}_2\overline{x}_3$	$\overline{x}_1\overline{x}_2x_3$ $x_1\overline{x}_2x_3$	$\overline{x}_1x_2x_3$ $x_1x_2x_3$	$\overline{x}_1x_2\overline{x}_3$ $x_1x_2\overline{x}_3$

Для $n = 4$ карты Карно имеют $2^4 = 16 = 4 \cdot 4$ клеток (табл. 4.26).

Таблица 4.26

Карта Карно для булевых функций четырех переменных

	Изменение x_1x_2			
Изменение x_3x_4	$x_1\overline{x}_2\overline{x}_3\overline{x}_4$	$\overline{x}_1x_2\overline{x}_3\overline{x}_4$	$x_1x_2\overline{x}_3\overline{x}_4$	$x_1\overline{x}_2x_3x_4$
	$x_1\overline{x}_2\overline{x}_3x_4$	$\overline{x}_1x_2\overline{x}_3x_4$	$x_1x_2\overline{x}_3x_4$	$x_1\overline{x}_2\overline{x}_3x_4$
	$\overline{x}_1\overline{x}_2x_3x_4$	$\overline{x}_1x_2x_3x_4$	$x_1x_2x_3x_4$	$x_1\overline{x}_2x_3x_4$
	$\overline{x}_1\overline{x}_2x_3\overline{x}_4$	$\overline{x}_1x_2x_3\overline{x}_4$	$x_1x_2x_3\overline{x}_4$	$x_1\overline{x}_2x_3\overline{x}_4$

Для построения минимальной ДНФ производится «склеивание» единиц. Склеиваются только соседние клетки, которые отличаются значением одной переменной. Процесс сводится к объединению в группы единичных клеток карт Карно. При этом общие переменные сохраняются, а различные опускаются.

Рассмотрим более подробно процедуру минимизации с помощью карт Карно. Алгоритм «склеивания» с помощью карт Карно имеет следующий вид.

1. Привести булеву функцию к ДНФ.
2. Нанести единицы на карту Карно.
3. Объединить соседние единицы контурами, охватывающими 2^m клеток, где $m = 0, 1, 2, 3$. При этом может оказаться, что единица попадает одновременно в два контура. Если контур охватывает более одной пары единиц одновременно, то предпочтительнее его не дробить на пары, а рассматривать как единый целый контур, например квадрат.
4. Провести упрощения, т.е. исключить члены, дополняющие друг друга до 1 внутри контура, следя за тем, чтобы переменные внутри контура были связаны операцией конъюнкции.
5. Объединить оставшиеся члены (по одному в каждом контуре) функцией ИЛИ, т.е. дизъюнкцией.

6. Записать полученное упрощенное булево выражение в ДНФ.

При заполнении карт Карно необходимо обратить внимание на порядок заполнения строк и столбцов значениями переменных. Последовательность значений переменных должна сохраняться неизменной. При таком заполнении каждые две соседние клетки отличаются лишь значением одной переменной. Нарушение порядка заполнения строк или столбцов не запрещается, но может не дать ожидаемого результата.

Рассмотрим примеры минимизации булевых функций с помощью карт Карно.

Пусть $f(ABC) = \overline{A}B\overline{C} \vee \overline{A}BC \vee AB\overline{C} \vee ABC$.

Нанесем единицы на карту (рис. 4.9) и обведем их сначала попарно двумя контурами. Такое действие соответствует заключению в скобки слагаемых $(\overline{A}B\overline{C} \vee \overline{A}BC)$ и $(AB\overline{C} \vee ABC)$. Вынося за скобки одинаковые конъюнкции согласно распределительному закону, в скобках получаем дизъюнкцию противоположных значений одной из переменных. В данном примере этому шагу соответствуют конъюнкции $\overline{A}B(\overline{C} \vee C)$ и $AB(\overline{C} \vee C)$. Поэтому объединение двух соседних единиц всегда приводит к закону инверсии, согласно которому дизъюнкция противоположных значений переменной равна 1.

Поэтому при записи ответа после применения карты Карно переменные, заключенные в общий контур, связываются конъюнкцией (как и общий множитель при вынесении за скобки), а такие отдельные конъюнкции, т.е. различные контуры, объединяются между собой дизъюнкцией.

Если записать полученный результат, то, очевидно, к нему вновь можно применить то же правило: $f(A, B, C) = \overline{A}B \vee AB = B$.

Однако в данном примере удобнее рассмотреть целиком весь квадрат из четырех единиц и сравнить переменные, записанные на горизонтальных и вертикальных клетках. Очевидно, общие множители сохранятся после упрощения (ведь их можно было вынести за скобки), а инвертируемые уйдут согласно закону инверсии. Поэтому целесообразнее опустить инвертируемые пары $\overline{A} \vee A$ и $\overline{C} \vee C$, а в ответе сохранить общую для всех клеток переменную B.

Синтактический способ минимизации дает тот же результат.

Проверка: $f(ABC) = (\overline{A}B\overline{C} \vee \overline{A}BC) \vee AB\overline{C} \vee ABC = \overline{A}B \vee AB = B(\overline{A} \vee A) = B$.

	$\overline{A}\overline{B}$	$\overline{A}B$	AB	$A\overline{B}$
C		1	1	
$\overline{C}$		1	1	

Рис. 4.9. Карта Карно для $f(ABC) = \overline{A}B\overline{C} \vee \overline{A}BC \vee AB\overline{C} \vee ABC$

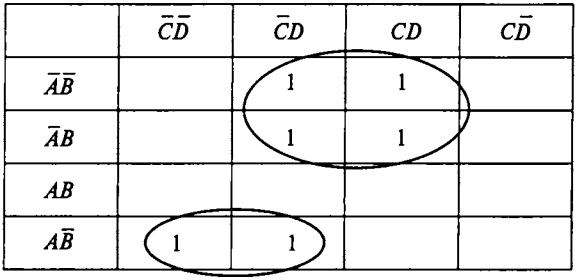


Рис. 4.10. Карта Карно для $f(A\overline{B}CD) = A\overline{B}\,\overline{C}\,\overline{D} \vee \overline{A}B\overline{C}D \vee \overline{A}\,\overline{B}\,\overline{C}D \vee \overline{A}\,\overline{B}CD \vee \vee \overline{A}BCD \vee A\overline{B}CD$

Рассмотрим примеры минимизации булевой функции, содержащей четыре переменные.

$f(ABCD) = A\overline{B}C\overline{D} \vee \overline{A}B\overline{C}D \vee \overline{A}\,\overline{B}\,\overline{C}D \vee \overline{A}\,\overline{B}CD \vee \overline{A}BCD \vee A\overline{B}\,\overline{C}D.$

Занесем единицы в соответствующие клетки карты Карно (рис. 4.10).

Рассмотрим переменные, заключенные контуром в квадрат. Среди них есть повторяющиеся в соседних клетках (это $\overline{A}$ и D) и инвертируемые (это пары B, $\overline{B}$ и C, $\overline{C}$). Повторяющиеся переменные как общий множитель мы сохраним, а инвертируемые — опустим. Из контура, содержащего две единицы, вынесем переменные A, $\overline{B}$, расположенные на одной строке, а также одинаковую для первых двух столбцов переменную $\overline{C}$, при этом опустим инвертируемую пару D, $\overline{D}$. После этих преобразований булева функция примет вид: $f(ABCD) = \overline{A}D \vee A\overline{B}\,\overline{C}$.

Проверка: $f(A\overline{B}CD) = A\overline{B}\,\overline{C}(\overline{D} \vee D) \vee \overline{A}CD(\overline{B} \vee B) \vee \overline{A}\,\overline{C}D(\overline{B} \vee \vee B) = A\overline{B}\,\overline{C} \vee AD(\overline{C} \vee C) = A\overline{B}\,\overline{C} \vee \overline{A}D$, т.е. результаты минимизации совпали.

Пусть $f(x_1x_2x_3x_4) = x_1x_2\overline{x}_3\overline{x}_4 \vee \overline{x}_1x_2\overline{x}_3\overline{x}_4 \vee \overline{x}_1x_2x_3\overline{x}_4 \vee x_1x_2x_3\overline{x}_4$.

Как обычно, занесем единицы в табл. 4.27.

Но чередование переменных в строке и столбце ничем не ограничено. Такой порядок был введен для удобства последующего упрощения. Поэтому можно сделать так, чтобы все единицы в

Таблица 4.27

Карта Карно

	$\overline{x}_3\overline{x}_4$	$\overline{x}_3x_4$	x_3x_4	$x_3\overline{x}_4$
$\overline{x}_1\overline{x}_2$				
$\overline{x}_1x_2$	1			1
x_1x_2	1			1
$x_1\overline{x}_2$				

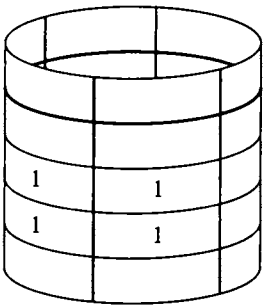


Рис. 4.11. Иллюстрация склейки карты Карно в цилиндр для $f(\underline{x_1}x_2\underline{x_3}\underline{x_4}) = \underline{x_1}x_2\underline{x_3}\underline{x_4} \vee \vee x_1x_2x_3x_4 \vee \underline{x_1}\underline{x_2}x_3x_4 \vee \vee x_1x_2x_3x_4$

данном случае оказались рядом. Для этого достаточно отождествить, т.е. «склеить», отмеченные на рис. 4.11 жирной линией стороны. Фактически это соответствует свертыванию карты в вертикальный цилиндр, в котором левый край совмещается с правым (также отмечено жирной линией). При их совмещении единицы первого и последнего столбцов окажутся соседними и для них можно будет применить алгоритм склеивания.

Таким образом, и без мысленного сворачивания карты можно *циклически переставлять* аргументы в строках и столбцах. Из всей четверки единиц по вертикали сохранится общая переменная второй и третьей строк x_2, а по горизонтали — общая переменная первого и последнего столбцов $\overline{x_4}$: $f(x_1x_2x_3x_4) = x_2\overline{x_4}$.

Такие нестандартные приемы минимизации булевых выражений упрощают саму процедуру и экономят время.

Пусть $f(x_1x_2x_3x_4) = \overline{x_1}\,\overline{x_2}x_3\overline{x_4} \vee \overline{x_1}\,\overline{x_2}\,\overline{x_3}x_4 \vee x_1\overline{x_2}\,\overline{x_3}\,\overline{x_4} \vee x_1\overline{x_2}\,\overline{x_3}x_4 = \overline{x_2}\,\overline{x_3}$.

Карта Карно для этой функции представлена на рис. 4.12.

Аналогично предыдущему случаю здесь удобно переставить переменные в столбце, что соответствует свертыванию карты в горизонтальный цилиндр.

В этом примере контуры объединены «через край» при свертывании карты Карно в горизонтальный цилиндр так, чтобы совместились верхний и нижний края карты и образовался квадрат. В первой и последней строках сохраняется общая переменная $\overline{x_2}$, а в первом и втором столбцах — общая переменная $\overline{x_3}$. Их конъюнкция и является ответом.

Пусть $f(\underline{x_1}x_2x_3\underline{x_4}) = \overline{x_1}\,\overline{x_2}x_3\overline{x_4} \vee \overline{x_1}\,\overline{x_2}x_3\overline{x_4} \vee x_1\overline{x_2}\,\overline{x_3}\,\overline{x_4} \vee x_1\overline{x_2}x_3\overline{x_4} = \overline{x_1}\,\overline{x_2}\,\overline{x_4} \vee \vee x_1\overline{x_2}\,\overline{x_4} = \overline{x_2}\,\overline{x_4}$.

Нанесем единицы на карту (табл. 4.28).

	$\overline{x_3}\,\overline{x_4}$	$\overline{x_3}x_4$	x_3x_4	$x_3\overline{x_4}$
$\overline{x_1}\,\overline{x_2}$	1	1		
$\overline{x_1}x_2$				
x_1x_2				
$x_1\overline{x_2}$	1	1		

Рис. 4.12. Горизонтальный цилиндр на карте Карно

Таблица 4.28

Карта Карно

	$\overline{x_3}\,\overline{x_4}$	$\overline{x_3}x_4$	x_3x_4	$x_3\overline{x_4}$
$\overline{x_1}\,\overline{x_2}$	1			1
$\overline{x_1}x_2$				
$x_1\overline{x_2}$				
x_1x_2	1			1

В свете описанных сверток ясно, что подобную карту нужно свернуть в шар и одновременно склеить четыре единицы, расположенные в углах. Для полученного квадрата вновь применим тот же алгоритм, сохранив переменные, одинаковые для первой и последней строк ($\overline{x_2}$), а также первого и последнего столбцов ($\overline{x_4}$), объединив их в конъюнкцию. Однако, если есть неуверенность в правильности совершаемых действий, лучше циклически переставить переменные (в любую сторону) и упрощать по обычным правилам. Один из таких вариантов изображен на рис. 4.13.

Горизонтальные аргументы сдвинули на 1 вправо, вертикальные — на 2 вниз. В данном случае был допущен большой произвол в действиях. Там, где на карте больше единиц, нужно так переставлять аргументы, чтобы как можно меньше единиц «рассекалось» границами карты и не нужно было дальше сворачивать.

Рассмотрим еще несколько примеров на упрощение булевых функций с помощью карт Карно.

Пусть $F(x_1x_2) = \overline{x_1}x_2 \vee \overline{x_2}x_1 \vee x_1x_2$.

Составим карту Карно для двух переменных (табл. 4.29).

Нанесем на карту Карно единицы. Объединим единицы в контуры по два элемента. Так как получилось два контура, один из которых дает x_1, а второй x_2, то упрощенный вариант булевой функции содержит дизъюнкцию двух членов: x_1 и x_2. Попадание единицы в два и более контуров соответствует закону идемпотен-

	$x_3\overline{x_4}$	$\overline{x_3}\,\overline{x_4}$	$\overline{x_3}x_4$	x_3x_4
x_1x_2				
$x_1\overline{x_2}$	1	1		
$\overline{x_1}\,\overline{x_2}$	1	1		
$\overline{x_1}x_2$				

Рис. 4.13. Шар на карте Карно

Таблица 4.29

Карта Карно

	$\overline{x_2}$	x_2
$\overline{x_1}$		1
x_1	1	1

тности $x = x \vee x$, поэтому каждое слагаемое (элементарная конъюнкция) может быть представлено столько раз, сколько нужно для упрощения. При этом они группируются с другими слагаемыми, с которыми попали в один контур, независимо. Фактически сделано следующее: $\overline{x_1}x_2 \vee \overline{x_2}x_1 \vee x_1x_2 = \overline{x_1}x_2 \vee \overline{x_2}x_1 \vee x_1x_2 \vee x_1x_2 = x_2(\overline{x_1} \vee x_1) \vee x_1(\overline{x_2} \vee x_2) = x_1 \vee x_2$. Получили $F(x_1x_2) = x_1 \vee x_2$.

Проверим результат аналитически: $\overline{x_1}x_2 \vee \overline{x_2}x_1 \vee x_1x_2 = \overline{x_1}x_2 \vee (\overline{x_2}x_1 \vee x_1x_2) = \overline{x_1}x_2 \vee x_1(\overline{x_2} \vee x_2) = \overline{x_1}x_2 \vee x_1 = x_2 \vee x_1$.

Ответы совпали, что подтверждает правильность преобразований.

Пусть $F(x_1x_2x_3) = \overline{x_1}\,\overline{x_2}\,\overline{x_3} \vee \overline{x_1}\,\overline{x_2}x_3 \vee x_1\overline{x_2}\,\overline{x_3} \vee x_1x_2\overline{x_3}$.

Составим карту Карно для функции трех переменных (табл. 4.30).

Объединим контурами полученные единицы.

Таблица 4.30

Карта Карно

	$\overline{x_3}$	x_3
$\overline{x_1}\,\overline{x_2}$	1	1
$\overline{x_1}x_2$		
x_1x_2	1	
$x_1\overline{x_2}$	1	

Опустим дополняющие друг друга переменные $x_3 \vee \overline{x_3}$ и $x_2 \vee \overline{x_2}$.

Запишем упрощенный вариант логической функции $F(x_1x_2x_3) = x_1\overline{x_3} \vee \overline{x_1}\,\overline{x_2}$.

Проверим полученный результат аналитически: $F(x_1x_2x_3) = (\overline{x_1}x_2\overline{x_3} \vee \overline{x_1}x_2x_3) \vee (x_1\overline{x_2}x_3 \vee x_1x_2\overline{x_3}) = \overline{x_1}x_2 \vee x_1\overline{x_3}$.

Ответ: $F(x_1x_2x_3) = x_1\overline{x_3} \vee \overline{x_1}\,\overline{x_2}$.

Пусть $F(x_1x_2x_3x_4) = x_1\overline{x_2}x_3\overline{x_4} \vee \overline{x_1}x_2x_3x_4 \vee \overline{x_1}\,\overline{x_2}\,\overline{x_3}x_4 \vee \overline{x_1}\,\overline{x_2}x_3x_4 \vee \overline{x_1}x_2x_3\overline{x_4} \vee x_1\overline{x_2}\,\overline{x_3}x_4$.

Таблица 4.31

Карта Карно

	$\overline{x_3}\,\overline{x_4}$	$\overline{x_3}x_4$	x_3x_4	$x_3\overline{x_4}$
$\overline{x_1}\,\overline{x_2}$		1	1	
$\overline{x_1}x_2$		1	1	
x_1x_2				
$x_1\overline{x_2}$				

Составим карту Карно для функции четырех переменных (табл. 4.31).

Объединим контурами полученные единицы.

Опустим дополняющие друг друга переменные $(x_4 \vee \overline{x_4})$ и $(x_2 \vee \overline{x_2})$ с $(\overline{x_3} \vee x_3)$.

Получим упрощенную булеву функцию $F(x_1x_2x_3x_4) = \overline{x_1}x_4$.

При желании можно проверить аналитически.

$F(x_1x_2x_3x_4) = (\overline{x_1}\,\overline{x_2}\,\overline{x_3}x_4 \vee \overline{x_1}\,\overline{x_2}x_3x_4) \vee (\overline{x_1}x_2x_3x_4 \vee \overline{x_1}x_2\overline{x_3}x_4) = \overline{x_1}\,\overline{x_2}x_4 \vee \vee \overline{x_1}x_2x_4 = \overline{x_1}x_4$.

Из последних примеров особенно наглядно видно, что карты Карно дают более рациональный путь минимизации булевых функций.

4.7. Сумма по модулю два

Решением всякой проблемы служит новая проблема.

Народная мудрость

Напомним, что **суммой по модулю два** (М2 или строгой дизъюнкцией) двух переменных x_1 и x_2 называется булева функция $x_1 \oplus x_2$, которая равна 1 тогда и только тогда, когда равна 1 *только* одна переменная. Приведем таблицу ее значений еще раз (табл. 4.32).

Приведем свойства, которыми обладает сумма по модулю два.

Она подчиняется законам:

- переместительному $x_1 \oplus x_2 = x_2 \oplus x_1$;
- сочетательному $(x_1 \oplus x_2) \oplus x_3 = x_1 \oplus (x_2 \oplus x_3)$, поэтому скобки можно не писать;
- распределительному конъюнкции относительно М2 $x_1(x_2 \oplus x_3) = = x_1x_2 \oplus x_1x_3$.

Таблица 4.32

Значения функции М2

x_1	x_2	$f(x_1, x_2) = x_1 \oplus x_2$
0	0	0
0	1	1
1	0	1
1	1	0

Операции с константами имеют вид:

- $x \oplus x = 0$;
- $x \oplus 0 = x$;
- $x \oplus \overline{x} = 1$;
- $x \oplus 1 = \overline{x}$.

Возможно разложение в СДНФ:

- $x_1 \oplus x_2 = \overline{x_1} x_2 \vee x_1 \overline{x_2}$.

Для суммы по модулю два справедлива формула отрицания

$\overline{x_1 \oplus x_2} = \overline{x_1} \oplus x_2 = x_1 \oplus \overline{x_2}$.

Проверим справедливость последней формулы с помощью таблицы истинности (табл. 4.33).

Совпадение пятого, шестого и восьмого столбцов подтверждает справедливость формулы. Докажем аналитически эту парадоксальную на первый взгляд равносильность. Для этого выведем формулы, связывающие сумму по модулю два с другими операциями. Из свойства $x \oplus 1 = \overline{x}$ видно, что операции отрицания и М2 связаны формулой $\overline{x} = x \oplus 1$. Поэтому $\overline{x_1 \oplus x_2} = (x_1 \oplus x_2) \oplus 1 = x_1 \oplus (x_2 \oplus 1) = x_1 \oplus \overline{x_2}$. В силу симметрии исходной функции по своим аргументам $\overline{x_1 \oplus x_2} = \overline{x_1} \oplus x_2$.

Таблица 4.33

Снятие отрицания с операции М2

x_1	x_2	$\overline{x_1}$	$\overline{x_2}$	$\overline{x_1} \oplus x_2$	$x_1 \oplus \overline{x_2}$	$x_1 \oplus x_2$	$\overline{x_1 \oplus x_2}$
0	0	1	1	1	1	0	1
0	1	1	0	0	0	1	0
1	0	0	1	0	0	1	0
1	1	0	0	1	1	0	1

Связь между дизъюнкцией и суммой по модулю два (строгой дизъюнкцией) $x_1 \vee x_2 = x_1 \oplus x_2 \oplus x_1x_2$ показана на рис. 4.14.

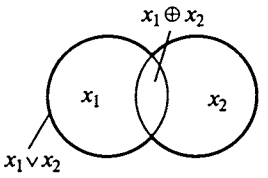


Рис. 4.14. Связь между дизъюнкцией $x_1 \vee x_2$ и суммой по модулю два $x_1 \oplus x_2$

Конъюнкцию x_1x_2 можно выразить через сумму по модулю два и дизъюнкцию по формуле $x_1x_2 = x_1 \oplus x_2 \oplus (x_1 \vee x_2)$, что следует из таблицы истинности (табл. 4.34).

Формула в последнем столбце истинна при любых значениях x_1 и x_2 , т.е. является тавтологией.

Теперь представим отрицание суммы по модулю два в виде, аналогичном $x_1 \oplus x_2 = \overline{x_1}x_2 \vee x_1\overline{x_2}$: $\overline{x_1 \oplus x_2} = \overline{\overline{x_1} \oplus x_2} = \overline{\overline{x_1}x_2 \vee x_1\overline{x_2}} = \overline{\overline{x_1}} \cdot \overline{x_2} = x_1x_2 \vee \overline{x_1}\,\overline{x_2}$.

Получено разложение в СДНФ и самой функции и ее отрицания.

Как видно из таблиц истинности, для двух переменных М2 реализует неравнозначность, т.е. исключающее *или*. Можно попытаться сконструировать функцию для трех переменных $x_1\overline{x_2}x_3 \vee \vee x_1x_2\overline{x_3} \vee \overline{x_1}x_2x_3$ или $x_1\overline{x_2}\overline{x_3} \vee \overline{x_1}x_2\overline{x_3} \vee \overline{x_1}\,\overline{x_2}x_3$, аналогичную $x_1 \oplus x_2 \oplus \oplus x_3$ и $A(x_1, x_2, x_3)$, принимающую значение 0 только при совпадении всех аргументов.

Проверим гипотезу о том, что и для трех переменных операции дадут одинаковый результат.

Для этого сравним таблицы истинности трех этих операций (табл. 4.35).

Как видно из таблицы, несовпадение последних трех столбцов говорит о том, что гипотеза не подтвердилась: для трех переменных результаты операций различны. Это является прямым следствием того, что мы рассматриваем двузначную логику: невозможно сделать так, чтобы все три булевы переменные принимали разные значения.

Логический элемент, соответствующий операции М2, показан на рис. 4.15, *а*. Треугольником обозначен инвертор НЕ.

Таблица 4.34

Связь между конъюнкцией и суммой по модулю два

x_1	x_2	$x_1 \vee x_2$	$x_1 \oplus x_2$	$x_1 \oplus x_2 \oplus (x_1 \vee x_2)$	x_1x_2	$x_1x_2 \equiv x_1 \oplus x_2 \oplus (x_1 \vee x_2)$
0	0	0	0	0	0	1
0	1	1	1	0	0	1
1	0	1	1	0	0	1
1	1	1	0	1	1	1

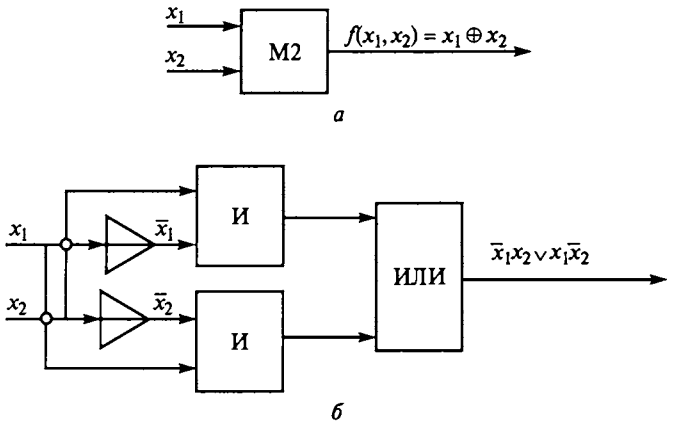


Рис. 4.15. Элемент и схема для операции М2:
а — логический элемент; *б* — логическая схема

Реализовать такой элемент в электронике достаточно сложно. Поэтому на практике он реализуется из набора простых и хорошо работающих элементов: И, НЕ, ИЛИ.

Воспользовавшись свойством разложения в СДНФ, имеем схему, изображенную на рис. 4.15, *б*.

Сравнивая операции двоичного сложения и суммы по модулю два, можно увидеть аналогию. Операция двоичного сложения в пределах последнего двоичного разряда имеет ту же последовательность символов, что и сумма по модулю два. Действительно,

Таблица 4.35

Таблица истинности $x_1 \oplus x_2 \oplus x_3$ и исключающего *или* для трех переменных

x_1	x_2	x_3	$x_1 \oplus x_2$	$x_1 \oplus x_2 \oplus x_3$	A	Исключающее *или* $x_1\bar{x}_2\bar{x}_3 \vee \bar{x}_1x_2\bar{x}_3 \vee \bar{x}_1\bar{x}_2x_3$
0	0	0	0	0	0	0
0	0	1	0	1	1	1
0	1	0	1	1	1	1
0	1	1	1	0	1	0
1	0	0	1	1	1	1
1	0	1	1	0	1	0
1	1	0	0	0	1	0
1	1	1	0	1	0	0

$0 + 0 = 0$,
$0 + 1 = 1$,
$1 + 0 = 1$,
$1 + 1 = (1)\ 0$.

Эти наблюдения дают основания для выдвижения еще одной гипотезы: возможно применение суммы по модулю два в двоичном сумматоре для системы контроля и исправления ошибок.

Действительно, если из-за неисправности в схеме один из аргументов функции М2 исказится, то одновременно и значение функции изменится на противоположное, что сразу можно будет обнаружить на выходе.

Поэтому операция сложения по модулю два имеет особое значение для организации работы компьютера: в схемах контроля и исправления ошибок используется это ее специфическое свойство.

Теперь поясним, откуда берется название «сумма по модулю два». Обозначим остаток от деления целого числа M на число N как $M \bmod N$. Рассмотрим множество {0, 1}, причем не абстрактных булевых символов, а обычных чисел (в математике его обозначают $\mathbb{Z}_2$). Произведем на этом множестве операцию $(a + b) \bmod 2$. Имеем: $(0 + 0) \bmod 2 = 0$, $(1 + 0) \bmod 2 = (0 + 1) \bmod 2 = 1 \bmod 2 = 1$, $(1 + 1) \bmod 2 = 2 \bmod 2 = 0$. Видим, что операция $(a + b) \bmod 2$ на множестве $\mathbb{Z}_2$ совпадает с нашей строгой дизъюнкцией на множестве B. Поэтому эти множества изоморфны: $\langle \mathbb{Z}_2, \bmod 2 \rangle \sim \langle B, \oplus \rangle$.

Геометрическая интерпретация суммы по модулю два с различным числом переменных может быть представлена на плоскости для функции $F(x_1, x_2) = x_1 \oplus x_2$ (рис. 4.16, *а*) и в трехмерном пространстве для функции $F(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$ (рис. 4.16, *б*). Закрашенные точки символизируют 1, выколотые или отсутствующие точки обозначают 0. Отрицание функции $x_1 \oplus x_2 \oplus x_3$ изображено на рис. 4.1, *б*.

При любом обходе такого куба по его ребрам нельзя попасть из одного единичного значения в другое непосредственно, так как между этими значениями расположен хотя бы один нуль. За счет

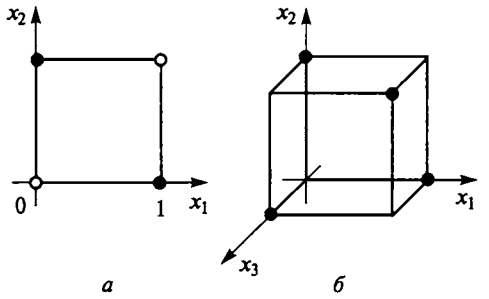


Рис. 4.16. Геометрическая интерпретация суммы по модулю два:
а — на плоскости; *б* — в пространстве

этого и происходит смена значения суммы по модулю два на противоположное при изменении значения одного из аргументов.

По этой же причине сумма по модулю два не поддается минимизации через операции отрицания, дизъюнкции и конъюнкции, так как ее единичные значения не образуют ни общих ребер, ни общих граней. На практике если булева функция имеет много единиц и сложно минимизируется, то не исключено, что она имеет вид М2 или близкий к нему.

Приведенные соотношения позволяют выражать одни булевы функции через другие.

? Какое минимальное число булевых функций позволяет выразить все остальное множество булевых функций в виде композиции заданных?

4.8. Полином Жегалкина. Функционально замкнутые классы

Верность теорий проверяется не опытом, а временем.

Н. Векшин

Как было отмечено ранее, в СДНФ булевой функции f только одна из элементарных конъюнкций равна 1. Это дает основание представить любую булеву функцию с помощью операции сложения по модулю два.

4.8.1. Канонический полином Жегалкина

Заменив в СДНФ $\overline{x_i}$ на $1 \oplus x_i = \overline{x_i}$ и используя распределительный закон для конъюнкции относительно сложения по модулю два, имеем $(1 \oplus x_i)(1 \oplus x_j) = 1 \oplus x_i \oplus x_j \oplus x_i x_j$. Тогда, учитывая, что $f \oplus f = 0$, а $f \oplus 0 = f$, булеву функцию f можно представить в виде $f(x_1 x_2 \ldots x_n) = f_0 \oplus f_1 x_1 \oplus f_2 x_2 \oplus \ldots \oplus f_n x_n \oplus f_{12} x_1 x_2 \oplus \ldots \oplus f_{12\ldots n} x_1 x_2 \ldots x_n$, причем $f_{i_1} \ldots i_n \in B = \{0, 1\}$.

Такое представление булевой функции называется **каноническим полиномом Жегалкина.**

Например, представим в виде полинома Жегалкина булеву функцию, заданную таблично (табл. 4.36).

Найдем ее СДНФ и преобразуем:

$f(x_1 x_2 x_3) = \overline{x}_1 x_2 \overline{x}_3 \vee x_1 \overline{x}_2 \overline{x}_3 \vee x_1 \overline{x}_2 x_3 \vee x_1 x_2 x_3 = (1 \oplus x_1) x_2 (1 \oplus x_3) \oplus$
$\oplus x_1 (1 \oplus x_2)(1 \oplus x_3) \oplus x_1 (1 \oplus x_2) x_3 \oplus x_1 x_2 x_3 = (x_2 \oplus x_1 x_2)(1 \oplus x_3) \oplus$
$\oplus (x_1 \oplus x_1 x_2)(1 \oplus x_3) \oplus (x_1 \oplus x_1 x_2) x_3 \oplus x_1 x_2 x_3 = x_2 \oplus x_1 x_2 \oplus x_2 x_3 \oplus x_1 x_2 x_3 \oplus$
$\oplus x_1 \oplus x_1 x_2 \oplus x_1 x_3 \oplus x_1 x_2 x_3 \oplus x_1 x_3 \oplus x_1 x_2 x_3 = x_1 \oplus x_2 \oplus x_2 x_3 \oplus x_1 x_2 x_3 =$
$= 0 \oplus x_1 \oplus x_2 \oplus 0 \cdot x_1 x_3 \oplus 0 \cdot x_1 x_2 \oplus x_2 x_3 \oplus x_1 x_2 x_3.$

Таблица 4.36

Булева функция, заданная таблично

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1

Как видим, $f_0 = f_{13} = f_{12} = 0$, остальные полиномиальные коэффициенты равны 1. Нули можно не выписывать, а отсутствие слагаемого (конъюнкции) вида $x_1 x_2 ... x_k$ будет означать, что $f_{12...k} = 0$. Таким образом, предпоследнее выражение тоже является полиномом Жегалкина.

Итак, с помощью конъюкции и М2 любую логическую функцию можно *единственным* образом представить в виде многочлена. Иногда можно представить функцию в виде полинома Жегалкина более простым способом. Например, функцию $x\overline{y}$ будем искать в виде многочлена с неопределенными коэффициентами: $x\overline{y} = a \oplus bx \oplus cy \oplus dxy$.

При $x = 0$, $y = 0$ будет $0 = a$; при $x = 0$, $y = 1$ имеем $0 = 0 \oplus c$, откуда $c = 0$; при $x = 1$, $y = 0$ имеем $1 = b \oplus 0$, т.е. $b = 1$. Наконец, при $x = y = 1$ получим $0 = 1 \oplus d$ и $d = 1$. В итоге $x\overline{y} = x \oplus xy$.

4.8.2. Функциональная замкнутость

Обозначим через P_2 множество всех функций алгебры логики.

Класс функций $R \subset P_2$ называется **функционально замкнутым**, если любая суперпозиция функций этого класса R принадлежит этому же классу.

Это означает, что вместе с каждой функцией $f(x_1 x_2 ... x_m) \in R$ классу R принадлежит и функция $f(y_1 y_2 ... y_m)$, где y_i — необязательно различные функции или аргументы и $f(f_1 f_2 ... f_m) \in R$, если $f(x_1 x_2 ... x_m) \in R$ и для всех $i = 1, 2, ..., n$ либо $f_j(x_{i_1} x_{i_2} ... x_{i_m}) \in R$, либо $f_j(x_{i_1} x_{i_2} ... x_{i_n}) \equiv x_{i_k}$.

Рассмотрим наиболее важные функционально замкнутые классы функций алгебры логики.

Класс функций, сохраняющих константу 0 (K_0). Так называют функции, для которых выполняется $f(00\ldots0) = 0$. Итак, $K_0 = \{f \mid f(00\ldots0) = 0\}$. В строчной записи функции через ее значения они начинаются на 0, например $f = (01111111)$.

Покажем, что $F = x_1\overline{x}_2 \vee x_1x_2\overline{x}_3 \in K_0$. Действительно, $F(000) = 0 \cdot \overline{0} \vee 0 \cdot 0 \cdot \overline{0} = 0$.

Для n аргументов имеется 2^{2^n-1} функций от n аргументов, сохраняющих константу, поскольку на одном из двоичных наборов, а именно на $x = (0\ldots0)$, значение f фиксировано, и для функции с n аргументами возможно подобрать лишь $2^n - 1$ произвольных наборов аргументов.

Такой класс функционально замкнут по определению, поскольку если $f_j, f_{i_1}, \ldots, f_{i_m} \in K_0$ и $x = (0\ldots0)$, то $f_j(f_{i_1}(x), \ldots, f_{i_m}(x)) = f_j(0\ldots0) = 0$.

Класс функций, сохраняющих константу 1 (K_1). Так называют функции, для которых выполняется $f(11\ldots1) = 1$. Итак, $K_1 = \{f \mid f(11\ldots1) = 0\}$. В строчной записи функции, сохраняющие константу 1, оканчиваются на 1, например $f = (11101111)$.

Покажем, что $F(x_1x_2x_3) = x_1 \vee \overline{x}_1x_2 \vee x_1x_2\overline{x}_3$ сохраняет 1. Это можно проверить подстановкой: $F(111) = 1 \vee \overline{1} \cdot 1 \vee 1 \cdot 1 \cdot \overline{1} = 1$.

Для K_1 справедливо все то же, что и для K_0.

Класс самодвойственных функций (K_c). Функция $f^*(x_1x_2\ldots x_n)$, удовлетворяющая условию $f^*(x_1x_2\ldots x_n) = \overline{f(\overline{x}_1\overline{x}_2\ldots\overline{x}_n)}$, называется **двойственной** по отношению к функции $f(x_1x_2\ldots x_n)$. Очевидно, что $(f^*)^* = f$. Таким образом, если $f^* = g$, то $g^* = f$, т.е. множество булевых функций разбивается на пары взаимно-двойственных функций.

Обратим внимание на свойства важнейших функций $(0)^* = 1$, $(x)^* = x$, $(\overline{x})^* = \overline{x}$, $(xy)^* = x \vee y$. Подобная двойственность лежит в основе построения противоположных высказываний.

Функция $f(x_1x_2\ldots x_n) \in P_2$ называется **самодвойственной**, если $f = f^*$. Например, самодвойственна функция $f(x) = x$.

Самодвойственная функция полностью определяется своими значениями на половине наборов аргументов, поэтому число таких функций при n аргументах равно $2^{2^n/2} = 2^{2^{(n-1)}}$.

Множество самодвойственных функций образует функционально замкнутый класс.

Пусть $f_j, f_{i_1}, \ldots, f_{i_m} \in K_c$, а $g = f_j \circ f_{i_k}$. Обозначим противоположный элемент на булевом кубе через $\overline{x} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n) \in B^n$. Тогда

$$g^*(x) = \overline{g(\overline{x})} = \overline{f_j(f_{i_1}(\overline{x}), \ldots, f_{i_m}(\overline{x}))} = f_j^*(\overline{f_{i_1}(\overline{x})}, \ldots, \overline{f_{i_m}(\overline{x})}) =$$

$$= f_j(\overline{f_{i_1}(\overline{x})}, \ldots, \overline{f_{i_m}(\overline{x})}) = f_j(f_{i_1}(x), \ldots, f_{i_m}(x)) = g(x),$$

откуда следует, что любая композиция самодвойственных функций не выходит из этого класса.

Класс линейных функций ($K_л$). Функцию алгебры логики вида $f(x_1x_2\ldots x_n)$ называют **линейной**, если ее канонический полином

Жегалкина имеет вид многочлена первой степени: $f(x_1x_2...x_n) = k_0 \oplus k_1x_1 \oplus k_2x_2 \oplus ... \oplus k_nx_n$, аналогичного обычному алгебраическому многочлену первой степени, но с коэффициентами k_i в виде 0 или 1. Число таких линейных функций от n аргументов равно 2^{n+1}, при $n = 2$ их восемь из шестнадцати. Они образуют функционально замкнутый класс. В гл. 1 было доказано это утверждение для непрерывных линейных отображений одного переменного. Можно доказать, что оно справедливо и для композиции булевых линейных функций $g = f_j \circ f_i$, где $f_j, f_{i_1}, ..., f_{i_m} \in K_л$. Для этого в многочлен $f_j(f_1, f_2, ..., f_m) = k_0 \oplus k_1f_1 \oplus k_2f_2 \oplus ... \oplus k_mf_m$ нужно подставить многочлены $f_i(x) = k_{0i} \oplus k_{1i}x_1 \oplus k_{2i}x_2 \oplus ... \oplus k_{ni}x_n$ и затем, используя переместительный и сочетательный законы для операции суммы по модулю два, привести подобные слагаемые. Тогда при каждом аргументе x_j ($j = 1, ..., n$) будет коэффициент $k_1k_{j1} \oplus k_2k_{j2} \oplus ... \oplus k_mk_{jm}$, который нужно просто вычислить.

Класс монотонных функций ($K_м$). В гл. 1 были определены понятия *порядок* и *упорядоченное множество*. Для бесконечных множеств это означает возможность введения бинарного отношения $\prec$, для не более чем счетных — возможность перенумерования элементов и тем самым также введения $\prec$. При этом эталоном сравнения служат натуральные числа.

? Можно ли сравнивать булевы функции, упорядочить их?

На множестве B введем полный порядок: $0 \prec 0$, $0 \prec 1$, $1 \prec 1$ по аналогии с отношением $\leq$ на множестве целых чисел $\mathbb{Z}$. На множестве B^n введем *частичный* порядок, означающий, что неравенство $\alpha = (\alpha_1\alpha_2...\alpha_n) \prec (\beta_1\beta_2...\beta_n) = \beta$ выполняется тогда и только тогда, когда $\forall i \in \{1, 2, ..., n\}$ $\alpha_i \prec \beta_i$. Например, $001100 \prec 001110$.

Два элемента α и β называются **сравнимыми** между собой, если $\alpha \prec \beta$ или $\beta \prec \alpha$. Частичным такой порядок является потому, что не все элементы B^n сравнимы. Например, кортежи 1100 и 0110 не сравнимы с помощью $\prec$. Поэтому не стоит путать частичный порядок $\prec$ на B^n и полное упорядочивание, использовавшееся для табличного и строчного задания булевой функции (назовем его здесь $\lhd$). Очевидно, что если $\alpha \prec \beta$, то $\alpha \lhd \beta$.

Функция $f(x_1x_2...x_n)$ называется **монотонной**, если для любых двух элементов $\alpha, \beta \in B^n$, сравнимых между собой, из $(\alpha_1\alpha_2...\alpha_n) \prec (\beta_1\beta_2...\beta_n)$ следует, что $f(\alpha_1\alpha_2...\alpha_n) \prec f(\beta_1\beta_2...\beta_n)$.

Монотонными будут x_1x_2 и $x_1 \vee x_2$, а $\bar{x}$ не является монотонной.

Монотонные функции также образуют функционально замкнутый класс. Действительно, пусть $f_i(x_1x_2...x_n)$, $g(x_1x_2...x_m) \in K_м$. Тогда композиция $h = g \circ f$ будет иметь вид $h(x) = g(f_1(x), f_2(x), ..., f_m(x))$, где $x = (x_1x_2...x_n) \in B^n$. Пусть $x \prec y$. Тогда $\forall i \in \{1, ..., m\}$ $f_i(x) \prec f_i(y)$ и $\alpha = (f_1(x), f_2(x), ..., f_m(x)) \prec (f_1(y), f_2(y), ..., f_m(y)) = \beta$, а поскольку g — монотонна, то $g(\alpha) \prec g(\beta)$. Подставляя сюда

явные выражения для α и β, получим $h(x) \prec h(y)$, т. е. монотонную функцию.

Следует отметить свойство монотонных булевых функций одного и двух аргументов. Поскольку на B и B^2 сравнимы лишь $0 \prec 1$ и $00 \prec 11$ соответственно, а значения получаются $0 \prec 1$, то такие монотонные функции будут сохраняющими 0 и 1, т. е. пересечением этих двух подклассов.

Каждый из пяти рассмотренных классов функций обладает очень важным свойством: *любая из функций алгебры логики, полученная с помощью операций суперпозиции и подстановки из функций одного класса, обязательно будет принадлежать тому же классу булевых функций.*

4.8.3. Функционально полные системы функций

Систему S функций f_1, f_2, ..., f_m алгебры логики называют **функционально полной**, если любую функцию алгебры логики можно записать с помощью суперпозиции некоторого набора булевых функций $f_1 f_2 \ldots f_m$.

Очевидно, что если S — функционально полная система, то добавление любого числа булевых функций не изменит статуса системы как функционально полной.

Функционально полная система функций называется **базисом** в пространстве P_2, если удаление хотя бы одной из функций, входящих в нее, превращает эту систему в функционально неполную.

Ранее было доказано, что через функции $x_1 \vee x_2$, $x_1 \wedge x_2$, $\overline{x}$ можно выразить любые функции алгебры логики, приведя их к виду СКНФ или СДНФ. Следовательно, система этих трех функций полна. Так же мы представляли любую функцию в виде композиции суммы по модулю два, конъюнкции и константы 1 (удаление всех 0 приводило к равной функции). Однако это не единственные возможности для представления функций алгебры логики. Оказывается, существует довольно много функционально полных систем.

Например, булевы функции можно выразить только через $x_1 \vee x_2$ и $\overline{x}$, воспользовавшись правилом Де Моргана и представив операцию конъюнкции через дизъюнкцию и отрицание, так как $x_1 \wedge x_2 = \overline{\overline{x_1} \vee \overline{x_2}}$.

Критерий функциональной полноты системы функций сформулирован в *теореме Поста—Яблонского.*

Для того чтобы система S функций $f_1 f_2 \ldots f_m$ алгебры логики была функционально полной, необходимо и достаточно, чтобы она содержала функцию:

не сохраняющую константу 0;

не сохраняющую константу 1;

не являющуюся самодвойственной;
не являющуюся линейной;
не являющуюся монотонной.

Выбор элементарных функций для базиса можно осуществлять с помощью табл. 4.37, где плюсом отмечены свойства, которыми эти функции обладают. В ней приведено распределение различных булевых функций двух переменных по пяти рассмотренным функционально замкнутым классам.

? Какое минимальное число булевых функций должен содержать базис?

Просим не путать обозначения функций в этой таблице, например f_7, с аналогичной нумерацией в подразд. 4.2.

Из табл. 4.37 видно, что каждая из функций $f_2 = x_1 \downarrow x_2$ — стрелка Пирса и $f_6 = x_1 \mid x_2$ — штрих Шеффера является минимальным базисом, так как соответствует условиям теоремы Поста—Яблонского. Это подтверждает и возможность выражать эти функции через отрицание, конъюнкцию и дизъюнкцию по формулам:

$$\overline{x_1 \downarrow x_2} = x_1 \vee x_2 \text{ или } x_1 \downarrow x_2 = \overline{x_1 \vee x_2}, \text{ т.е. } x \downarrow x = \overline{x};$$

$$\overline{x_1 \mid x_2} = x_1 x_2 \text{ или } x_1 \mid x_2 = \overline{x_1 x_2}, \text{ т.е. } x \mid x = \overline{x}.$$

Таблица 4.37

Критерии функциональной полноты булевых функций

Функция	$f(x_1, x_2)$	Значения	Не сохраняет 0	Не сохраняет 1	Несамодвойственная	Нелинейная	Немонотонная
f_1	0	0000		+	+		
f_2	$x_1 \downarrow x_2$	1000	+	+	+	+	+
f_3	$x_1 \overline{x_2}$	0010		+	+	+	
f_4	$\overline{x_2}$	1010	+	+			+
f_5	$x_1 \oplus x_2$	0110		+	+	+	
f_6	$x_1 \mid x_2$	1110	+	+	+	+	+
f_7	$x_1 x_2$	0001			+	+	
f_8	$x_1 \leftrightarrow x_2$	1001	+		+	+	
f_9	$x_2 \rightarrow x_1$	1011	+		+	+	
f_{10}	$x_1 \vee x_2$	0111			+	+	
f_{11}	1	1111	+		+		

Таким образом, для описания булевой алгебры достаточно одной функции. Другое дело, что стрелку Пирса и штрих Шеффера трудно интерпретировать в теории множеств и реализовать в виде элементарных логических схем через элементы И—НЕ и ИЛИ—НЕ, а формулы и логические схемы, составленные только из них, являются громоздкими.

Вообще говоря, в алгебре логики доказано, что минимальный базис содержит не более четырех функций. Хотя существует базис, который содержит ровно четыре функции и образует функционально полную систему: $f_1' = 0, f_2' = 1, f_3' = x_1x_2, f_4' = x_1 \oplus x_2 \oplus x_3$.

Приведем примеры основных функционально полных систем булевых функций:

и, или, не, которые лежат в основе алгебры высказываний или булевой алгебры, $S_1 = (f_7, f_{10}, f_4) = \{\vee, \wedge, \neg\}$;

и, не, $S_2 = (f_4, f_7) = \{\wedge, \neg\}$;

или, не, т.е. $S_3 = (f_4, f_{10}) = \{\vee, \neg\}$;

штрих Шеффера, $S_4 = (f_6) = \{|\}$;

стрелка Пирса, $S_5 = (f_2) = \{\downarrow\}$;

сумма по модулю два, и, 1, $S_6 = (f_5, f_7, f_{11}) = \{\oplus, \wedge, 1\}$;

$x_1\overline{x_2}$, *не*, $S_7 = (f_3, f_4) = \{x_1\overline{x_2}, \overline{x_2}\}$;

импликация, отрицание, $S_8 = (f_9, f_4) = \{\rightarrow, \neg\}$.

Заметим, что полная система функций $S_1 = (f_7, f_{10}, f_4)$ будет избыточной. Действительно, удалив из него одну из функций f_7 или f_{10} и выразив их с помощью законов Де Моргана через f_{10} и f_4 или f_7 и f_4 соответственно, можно получить новую функционально полную систему с базисом S_3 или S_2 соответственно.

Рассматриваемая проблема представления логических функций заключается не только в выборе минимального базиса, но и в выборе формы наиболее экономичного представления этих функций.

Так, совершенные дизъюнктивные и конъюнктивные нормальные формы (СДНФ и СКНФ) менее экономичны, чем некоторые из ДНФ и КНФ, хотя и обладают определенными достоинствами, например дают однозначное представление логической функции. Поэтому одинаково важны обе встречные проблемы: с одной стороны, минимизация булевых функций, т.е. приведение к наиболее экономичному виду, и с другой стороны — придание им совершенного вида, важного в некоторых приложениях, например для получения выводов.

Уточним, что минимальной называется такая форма представления логических функций, которая содержит минимальное количество термов и переменных в термах, а значит, минимальная форма предполагает завершение процесса упрощения.

Можно построить различные формальные системы — *алгебры* — в зависимости от выбранных в качестве базисных логических операций.

Так, булевы функции вида $S_1 = \{\vee, \wedge, \neg\}$ образуют *булеву алгебру*, а операции дизъюнкции, конъюнкции и отрицания образуют *базис* булевой алгебры.

Множество булевых функций в базисе $S_6 = \{\oplus, \wedge, 1\}$ образуют *алгебру Жегалкина*, а операции М2, $\wedge$ и 1 образуют *базис Жегалкина*. Достоинством алгебры Жегалкина является арифметизация логики, благодаря чему двузначная аристотелева логика нашла свое практическое применение. Поскольку в ней справедлив закон исключенного третьего, то алгебра Жегалкина обладает одной существенной особенностью: она непротиворечива, так как непротиворечива ее система аксиом.

Анализ построения различных алгебр с помощью выбора соответствующего языка дает возможность увидеть принципы построения любой науки вообще.

Выбирая, например, язык будущей системы, примем во внимание структуру любой науки, будь то алгебра, геометрия, математическая логика или языки программирования.

В конце попытаемся ответить на некоторые любопытные вопросы.

Что значит говорить с наукой на одном языке? К чему приведет эффект Вавилонской башни?

Что такое формализация языка? Какова роль математики в системе знаний как способе формализации?

Все ли знания поддаются формализации? Достаточно ли законов логики, чтобы формализовать все знания, известные науке на сегодня, и получать новые?

Как должна быть построена теория, чтобы в ней не возникало противоречий?

Какими свойствами должны обладать методы доказательств, чтобы считаться достаточно строгими?

Упражнения

4.1. Проверьте, являются ли булевы функции F_1 и F_2 эквивалентными:

а) $F_1 = X \to (Y \equiv Z)$ и $F_2 = (X \to Y) \equiv (X \to Z)$;
б) $F_1 = X \cdot (Y \equiv Z)$ и $F_2 = (XY) \equiv (XZ)$;
в) $F_1 = X \to (Y \vee Z)$ и $F_2 = (X \to Y) \vee (X \to Z)$;
г) $F_1 = \overline{XZ} \vee XY \vee \overline{XZ}$ и $F_2 = X\overline{YZ} \vee \overline{XZ}$;
д) $F_1 = X \equiv Z$ и $F_2 = (X \vee Y \vee Z) \to (X \vee Y)(Y \vee Z)$;
е) $F_1 = Y \to (X \to Z)$ и $F_2 = X \to (XY \to ((X \to Y) \to Y)Z)$.

4.2. Вычислите значение функции $F(x_1, x_2, x_3)$ при заданных значениях аргументов $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ и при $x_1 = 1$, $x_2 = 1$, $x_3 = 1$; затем приведите функцию к минимальной ДНФ:

а) $F(x_1, x_2, x_3) = x_2\overline{x_3} \vee x_3 \vee (x_1 \cdot \overline{x_2} \to x_3)$;

б) $F(x_1, x_2, x_3) = \overline{x_1 \cdot \overline{x_2} \to x_3 \vee x_1 \overline{x_2} x_3}$;

в) $F(x_1, x_2, x_3) = \overline{x_1 \cdot (\overline{x_2} \vee x_1 x_3)} \to x_1 x_2 \overline{x_3}$;

г) $F(x_1, x_2, x_3) = \overline{x_1 \cdot \overline{x_3} \vee x_3} \to x_1 \cdot \overline{x_2}$;

д) $F(x_1, x_2, x_3) = \overline{x_2 \cdot \overline{x_3} \to x_3} \vee \overline{x_2} \cdot \overline{x_3}$;

е) $F(x_1, x_2, x_3) = x_1 \cdot \overline{x_2} \vee x_3 \to x_1 \cdot \overline{x_3}$.

4.3. По заданной функции постройте таблицу истинности, приведите функцию к минимальной ДНФ:

а) $F(x_1, x_2, x_3) = x_1 \vee \overline{x_2 \cdot x_3} \cdot (x_1 \vee x_2)$;

б) $F(x_1, x_2, x_3) = x_1 \vee x_2 \vee \overline{x_1 \cdot x_3 \vee \overline{x_1 \cdot x_2}}$;

в) $F(x_1, x_2, x_3) = x_1 \cdot x_2 \vee \overline{x_1 x_2 \vee \overline{x_1}} \cdot x_3 \vee x_1$;

г) $F(x_1, x_2, x_3) = (x_1 \vee x_3) \cdot \overline{(x_1 \cdot x_2)} \vee x_1$;

д) $F(x_1, x_2, x_3) = x_1 \cdot \overline{x_2 \cdot x_3} \vee \overline{x_1 \vee x_2}$;

е) $F(x_1, x_2, x_3) = x_1 \cdot x_2 \vee x_1 \vee \overline{x_1 \cdot x_3}$.

4.4. Постройте логическое выражение по заданной таблице истинности (табл. 4.38), приведите его к минимальной ДНФ алгебраически и с помощью карт Карно постройте соответствующий логический элемент.

Таблица 4.38

Варианты таблицы истинности к упр. 4.4

а)

X_1	X_2	X_3	F
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

б)

X_1	X_2	X_3	F
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

в)

X_1	X_2	X_3	F
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	0

г)

X_1	X_2	X_3	F
0	0	0	0
0	0	1	0
0	1	0	1

д)

X_1	X_2	X_3	F
0	0	0	1
0	0	1	0
0	1	0	0

е)

X_1	X_2	X_3	F
0	0	0	0
0	0	1	1
0	1	0	0

Окончание табл. 4.38

X_1	X_2	X_3	F
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	1
1	1	1	1

X_1	X_2	X_3	F
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1

X_1	X_2	X_3	F
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

4.5. Постройте совершенные ДНФ и КНФ и соответствующие минимальные формы для булевых функций, заданных таблично (табл. 4.39).

Таблица 4.39

Варианты таблицы истинности к упр. 4.5

а)

X_1	X_2	X_3	F
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

б)

X_1	X_2	X_3	F
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

в)

X_1	X_2	X_3	F
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	0

г)

X_1	X_2	X_3	F
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

д)

X_1	X_2	X_3	F
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

е)

X_1	X_2	X_3	F
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

4.6. Определите, может ли каждая из данных высказывательных форм стать истинным или ложным высказыванием:

а) число n — четное или число $n + 1$ — четное;

б) число n — четное и число $n + 1$ — четное;

в) $\begin{cases} x > 0, \\ x < 0; \end{cases}$ г) $\begin{cases} x > 1, \\ x < 0; \end{cases}$ д) $\begin{cases} x \geq 0, \\ x \leq 0; \end{cases}$ е) $(x > 0) \vee (x \leq 0)$.

4.7. Переведите предложения на язык алгебры логики и определите, если возможно, их истинность:

а) каждое слагаемое суммы $a + b + c$ делится на 2;

б) все простые однозначные числа больше 3 — четные;

в) хотя бы одно из чисел n, $n + 1$, $n - 1$ — четное;

г) число a принадлежит по крайней мере одному из множеств A и B;

д) существует натуральное число x, которое больше 25, но меньше 52 и которое делится на 3 и на 5;

е) квадратное уравнение имеет не более двух корней.

4.8. Введя обозначения, запишите логическую форму высказываний и определите их вид:

а) «Порок — это не употребление плохого, а злоупотребление хорошим» (древняя мудрость);

б) «Чем честнее человек, тем менее он подозревает других в бесчестности»;

в) «Мастер не учит, а создает ситуации» (древняя мудрость);

г) «Либо все люди должны быть счастливы, либо никто» (Роберт Оуэн);

д) «Хотите подчинить себе других — начинайте с себя» (Л. Вовенарг);

е) «Благие намерения без квалификации дают тот же результат, что и квалификация без благих намерений» (В. Леви, психотерапевт).

4.9. Введя обозначения, запишите логическую форму сложного высказывания, постройте его отрицание, установите семантическую характеристику этих высказываний:

а) «Не может управлять другим тот, кто не в состоянии управлять самим собой» (английская пословица);

б) «Единственный урок, который можно извлечь из истории, состоит в том, что люди не извлекают из истории никаких уроков» (Б. Шоу);

в) «Со счастьем дело обстоит, как и с часами: чем проще механизм, тем реже они портятся» (Н. Шамфор);

г) «Чтобы победить противника, не стремись стать сильнее его, а сделай его слабее себя»;

д) «Чем меньше человек собирается сделать, тем больше он об этом говорит»;

е) «В жизни возможны лишь две трагедии: не осуществить свою страстную мечту и добиться ее осуществления».

4.10. Запишите с помощью логических операций высказывания:
а) $bb = 0$; б) $ab \neq 0$; в) $a^2 + b^2 = 25$; г) $a/b = 0$; д) $|a| = 3$; е) $|a| \geq 3$.

4.11. Сформулируйте отрицание высказывания и определите истинность данного высказывания и его отрицания:

а) $5 \leq 3$; б) $\sqrt{16} = 4$; в) если $x^2 = 9$, то $x = 3$; г) если $\frac{x^2 - 9}{x + 3} \geq 0$, то $x - 3 \geq 0$;

д) если $\frac{5}{8 - x} < 0$, то $x < 8$; е) если $n \in N$, то $(n - 1)n(n + 1)$ делится на 6.

4.12. Установите, какие из следующих пар являются отрицаниями друг друга, а какие не являются:
а) $x > 0$ и $x \leq 0$; б) ΔABC — прямоугольный и ΔABC — тупоугольный;
в) $f(x)$ — четная функция и $f(x)$ — нечетная функция;
г) все простые числа нечетные и все простые числа нечетные;
д) все простые числа нечетные и существуют простые нечетные числа;
е) четырехугольник $ABCD$ — квадрат и четырехугольник $ABCD$ — ромб.

4.13. Запишите эти предложения без знака отрицания:

а) $\overline{a < b}$; б) $\overline{a \leq b}$; в) $\overline{a \geq b}$; г) $\overline{a > b}$; д) $\overline{a \neq b}$; е) $\overline{a \in B}$.

4.14. Пусть A — высказывание «$9 \vdots 3$», B — «$10 \vdots 3$», где знак $\vdots$ — деление нацело. Определите значение истинности и обоснуйте ответ:
а) $A \to B$; б) $\overline{B} \to A$; в) $\overline{A} \to B$; г) $\overline{B} \to \overline{A}$; д) $A \leftrightarrow B$; е) $A \oplus B$.

4.15. Найдите в тексте гл. 1 любые примеры операций над высказываниями, введите обозначения и представьте найденные сложные высказывания в виде формулы алгебры логики.

4.16. Из двух простых высказываний A и B составьте сложные высказывания по формулам:
$\overline{A}$, $A \vee B$, $A \oplus B$, $A \wedge B$, $A \to B$, $A \leftrightarrow B$:
а) A: «Учит разуму» и B: «Быть разумным»;
б) A: «Понять истину» и B: «Додуматься самому»;
в) A: «Рыть яму другому» и B: «Попасть в яму»;
г) A: «Знаешь рецепт» и B: «Лекарство подействовало»;
д) A: «Иметь свою волю» и B: «Иметь свою долю»;
е) A: «Учиться без книг» и B: «Черпать решетом воду».

4.17. Из высказываний упражнения 4.16 составьте сложные высказывания по формулам и сравните их таблицы истинности:
а) $\overline{A} \wedge \overline{B}$ и $\overline{A \vee B}$; б) $\overline{A} \wedge \overline{B}$ и $\overline{A \wedge B}$; в) $(A \wedge B) \vee (\overline{A} \wedge \overline{B})$ и $A \leftrightarrow B$; г) $(A \wedge \overline{B}) \vee (\overline{A} \wedge B)$ и $A \oplus B$; д) $\overline{A} \vee \overline{B} \wedge A$ и $\overline{A} \vee \overline{B}$; е) $\overline{A} \vee B$ и $A \to B$.

4.18. Даны высказывания A: «Я купил компьютер», B: «Я успешно освоил ArchiCAD», C: «Я участвовал в олимпиаде по компьютерной графике». По предлагаемым формулам сформулируйте

высказывания и сравните их таблицы истинности. Сделайте вывод об истинности этих высказываний:

а) $A(B \vee C)$ и $AB \vee AC$; б) $A \vee BC$ и $(A \vee B)(A \vee C)$; в) $\overline{ABC}$ и $A \vee \overline{B} \vee \overline{C}$; г) $\overline{A \vee B \vee C}$ и $\overline{ABC}$; д) $\overline{A \to BC}$ и $\overline{A} \vee BC$; е) $\overline{A \vee BC}$ и $\overline{ABC}$.

4.19. Даны простые высказывания A: «Четырехугольник $ABCD$ — параллелограмм», B: «Диагонали четырехугольника $ABCD$ в точке пересечения делятся пополам». Сформулируйте сложные высказывания по формулам и определите их истинность по таблице. Упростите высказывания и сравните их таблицы истинности:

а) $\overline{A} \to \overline{B}$; б) $\overline{B} \to \overline{A}$; в) $A \leftrightarrow B$; г) $\overline{A} \leftrightarrow \overline{B}$; д) $\overline{A} \to B$; е) $\overline{B} \to A$.

4.20. Составьте таблицы истинности для предлагаемых выражений. Упростите выражения и сделайте вывод об их истинности:

а) $A \to (B \to A)$; б) $\overline{\overline{A} \vee \overline{B}} \leftrightarrow AB$; в) $(A \vee B)\overline{A} \to B$; г) $\overline{\overline{A}}\,\overline{\overline{B}} \leftrightarrow$ $\leftrightarrow A \vee B$; д) $A \to (B \to C) \to ((A \to B) \to (A \to C))$; е) $\overline{A} \to (A \to B)$.

4.21. Для данной импликации сформулируйте обратную, противоположную и противоположную обратной. Какие из этих четырех импликаций истинные, а какие ложные?

а) если число делится на 9, то сумма его цифр делится на 3;

б) если диагонали параллелограмма равны, то это прямоугольник;

в) если натуральное число делится на 3 и на 5, то оно делится на 15;

г) произведение двух положительных чисел положительно;

д) площади равных фигур равны;

е) произведение трех отрицательных чисел отрицательно.

4.22. Введите обозначения и запишите высказывания на языке алгебры логики. С помощью таблицы истинности установите, являются ли они тождественными:

а) «если электростанция прекратит подачу тока, то предприятие остановится, а если оно остановится, то понесет большие убытки» и «если электростанция остановится, то предприятие понесет большие убытки»;

б) «если по проводнику проходит электрический ток, то вокруг проводника образуется магнитное поле, но вокруг проводника не образуется магнитное поле» и «по проводнику не проходит электрический ток»;

в) «если по проводнику проходит электрический ток, то вокруг проводника образуется магнитное поле, но по проводнику не проходит электрический ток» и «вокруг проводника не образуется магнитное поле»;

г) «если стоит туманная погода, то аэропорт закрывают. Стоит туманная погода» и «аэропорт закрыт»;

д) «если поезд прибывает на станцию, то подается сигнал “Путь закрыт”, но поезд на станцию не прибыл» и «подается сигнал “Путь закрыт”»;

е) «как только поезд прибывает на станцию, подается сигнал "Путь закрыт", поезд прибыл на станцию» и «сигнал "Путь закрыт" не подан».

4.23. В формулах опустите излишние скобки и упростите высказывание:

а) $(((A \vee B) \to (\overline{C}D)) \to (\overline{A \vee C}))$;

б) $((\overline{B} \to A \vee C)) \vee ((\overline{A} \vee B) \to C)$;

в) $((\overline{A} \vee (BC)) \to ((\overline{CD}) \vee B))$;

г) $(((A\overline{C}) \to (B \vee D)) \vee ((A \to (D \vee C)) \to \overline{B})$;

д) $((\overline{A} \vee B) \to C) \vee (\overline{D} \to ((BC) \vee A)))$;

е) $((C \to (A \vee B)) \to ((C \vee \overline{D}) \to (\overline{AB})))$.

4.24. Упростите высказывания и сравните таблицы истинности условия и ответа. Сделайте вывод об истинности:

а) $\overline{\overline{AB} \vee \overline{BC} \vee AC}$; б) $\overline{\overline{(A \vee B)C} \vee \overline{AB} \vee \overline{BC}}$;

в) $\overline{\overline{B \vee C} \vee \overline{A \vee C} \vee AB}$; г) $\overline{A \vee \overline{B}(A \vee C) \vee B(\overline{A \vee C})}$;

д) $\overline{\overline{A \vee \overline{B} \vee C} \vee \overline{A \vee C} \vee A\overline{B}}$; е) $\overline{\overline{AC \vee B} \vee \overline{B} \vee B(\overline{A \vee C})}$.

4.25. Найдите двойственные функции и заполните табл. 4.40. Докажите двойственность аналитически и с помощью таблиц истинности.

Таблица 4.40

Таблица двойственных функций

F	0	1	X	$\overline{X}$	$X_1 \vee X_2$	$X_1 \wedge X_2$	$X_1 \mid X_2$	$X_1 \downarrow X_2$	$X_1 \to X_2$	$X_1 \equiv X_2$	$X_1 \oplus X_2$
F^*											

4.26. По мишени произведено три выстрела. Для высказывания P_i: «Мишень поражена i выстрелами, где $i = 1, 2, 3$» сформулируйте высказывания:

а) $P_1 \vee P_2 \vee P_3$; в) $(P_1 \vee P_2) \cdot \overline{P_3}$; д) $\overline{P_1 \cdot P_2} \cdot P_3$;

б) $P_1 \cdot P_2 \cdot P_3$; г) $\overline{P_1} \cdot P_3 \vee P_1 \cdot \overline{P_3}$; е) $(\overline{P_1 \vee P_2}) \cdot \overline{P_3}$.

Определите истинность каждого высказывания. Упростите высказывания, если возможно, и сформулируйте результат на русском языке.

4.27. Докажите или опровергните:

а) $A \wedge B = 1 \Leftrightarrow A \vee B = 1$; б) $(AB) \wedge (B \to C) = A \to C$; в) $A \to \to (A \to B) = 1$; г) $A \to B = 1 \Leftrightarrow A \vee B = 1$; д) $A \to B = \overline{B} \to A$; е) $A \to \to (B \to C) \to ((A \to B) \to (A \to C)) = 1$.

4.28. Проверьте, являются ли булевы функции F_1 и F_2 эквивалентными:

а) $F_1 = X \oplus (Y \to Z)$ и $F_2 = (X \oplus Y) \to (X \oplus Z)$;

б) $F_1 = ((X \oplus Y) \to (X \vee Y))((\overline{X} \to Y) \to (X \oplus Y))$ и $F_2 = X \mid Y$;

в) $F_1 = (X \to Y) \to (XY \oplus (X \sim Y))$ и $F_2 = (X \vee Y)(X \vee \overline{Y})$;
г) $F_1 = (X \to Y) \to ((X \& \overline{Y}) \oplus (Z \sim \overline{Y}))$ и $F_2 = X \& \overline{Y} \& (\overline{Y} \to X\overline{Z})$;
д) $F_1 = X \to (Y \to Y)$ и $F_2 = (X \wedge Y) \to Y$;
е) $F_1 = (X \to Y) \wedge (X \to \overline{Y})$ и $F_2 = A \to (B \wedge C)$.

4.29. Определите вид вопроса и дайте на него истинный полный ответ:

а) Правда ли, что если произойдет повреждение таблицы распределения файлов, то произойдет значительная потеря информации?

б) Правда ли, что первый сектор дискеты состоит из таблицы параметров и кода программы загрузчика?

в) Правда ли, что в электронных таблицах EXCEL для редактирования диаграмм можно воспользоваться последовательностью МЕНЮ—ПРАВКА—ОБЪЕКТ или вызвать контекстное меню в области диаграммы?

г) Правда ли, что форм-факторы материнской платы бывают AT или ATX и NLX или LPX?

д) Сколько и каких основных характеристик имеет системная шина и к каким шинам относятся шины PSI и AGP?

е) Где находится и что такое boot-record или что и в каком порядке идет после boot-record?

4.30. Докажите или опровергните:

а) $x \to (y \to z) \equiv (xy) \to z$ (правило импортации);
б) $x \to (y \to z) \equiv y \to (x \to z)$;
в) $x \to (y \to z) \equiv (x \to y) \to (x \to z)$ (правило самодистрибутивности импликации);
г) $(\overline{x} \to y)(y \to x) \equiv x$;
д) $(xy) \to z \equiv (x \to y) \vee (y \to z)$;
е) $(x \vee y) \to z \equiv (x \to y)(y \to z)$.

4.31. Какие законы правильного мышления нарушены в мудрой сказке Льюиса Кэрролла «Алиса в Стране чудес»?

а) «Кто — ты — такая?..

— Видите ли... видите ли, сэр, я... просто не знаю, кто я такая. Нет, я, конечно, примерно знаю, кто такая я была утром, когда встала, но с тех пор я все время то такая, то сякая — словом, какая-то не такая. — И она беспомощно замолчала.

— Выражайся яснее! — строго сказал Червяк. — Как тебя прикажешь понимать?

— Я сама не понимаю, сэр, потому что получается, что я — это не я! Видите, что получается?

— Не вижу! — отрезал Червяк.

— Простите меня, пожалуйста, — сказала Алиса очень вежливо, — но лучше я, наверное, не сумею объяснить. Во-первых, я сама никак ничего не пойму, а во-вторых, когда ты то большой, то маленький, то такой, то сякой, то этакий — то все как-то путается, правда?»

б) ... «Нет, почему, — осторожно начала Алиса, — иногда, особенно на уроках музыки, я думала — хорошо бы получше провести время...

— Все понятно! — с торжеством сказал Шляпа. — Провести время? Ишь чего захотела! Время не проведешь! Да и не любит он этого! Ты бы лучше постаралась с ним подружиться — вот тогда бы твое дело было...в шляпе!»

в) «Деликатес встревожился.

— Каждый день? — повторил он в раздумье. — Да-а, интересно, на каком же уровне твоя школа?

— Простите, я не понимаю, сказала Алиса, — что значит на каком уровне?

— На каком уровне она стоит! — пояснил Деликатес. — Ну от поверхности моря, поняла?

— Там моря нет, — сообщила Алиса. Она стоит в городе. Но я думаю, все-таки выше моря, конечно, над водой!

— Выше? Над водой? — переспросил Деликатес. — Ты серьезно?

Алиса молча кивнула.

— Ну, тогда это не серьезно! — с облегчением сказал Деликатес. — Какое же тогда может быть сравнение с *нашей* школой? Это... это верхоглядство, а не образование, вот что это такое.

Грифон фыркнул.

— Да уж, воображаю, какие вы там получаете поверхностные знания! — сказал он. — У нас мальков — и тех учат гораздо глубже! А уж кто хочет по-настоящему углубиться в науку, тот должен добраться до самого дна! Вот это и называется Законченное Низшее Образование! — Но, конечно, — покачал он головой, — это не каждому дано!..

— Мне вот так и не удалось по-настоящему углупиться! Не хватило меня на это, — сказал Деликатес со вздохом. — Так я и остался при Высшем Образовании...»

4.32. По известным формулам сформулируйте различные виды вопросов, связанных с вашей специальностью, дайте им характеристику. Представьте ответ на каждый вопрос и укажите его характеристику.

Глава 5
ФОРМАЛЬНЫЕ СИСТЕМЫ И УМОЗАКЛЮЧЕНИЯ. ЛОГИКА ПРЕДИКАТОВ

В этой главе мы познакомимся со словарем и правилами построения формальных систем, с одним из содержательных представлений формальных систем — логикой предикатов. Затем рассмотрим методы научного познания и способы построения умозаключений: дедуктивный, индуктивный и по аналогии. Подробно изучим один из важнейших методов математики — метод математической индукции и его применение для построения доказательств. Также в этой главе найдут содержательное применение элементы математической логики, которые рассмотрены в предыдущей главе.

5.1. Формальные системы

> Мудрец не только знает следствия из принципов, но и обладает истинными знаниями самих принципов.
>
> *Аристотель*

Понятие о формальных системах. Для того чтобы знания, которыми обладает один человек, стали доступны всем и однозначно поняты всеми другими людьми, они должны быть изложены с помощью точно определенных понятий. Проще всего это сделать тогда, когда знания формализованы. Под **формализацией** принято понимать математические формулировки теоретических и прикладных проблем в различных областях науки. Особенно важно соблюдать строгие правила формализации при работе с компьютером. Для компьютеризации любой области знаний или деятельности необходимо представить ее в виде системы, которая функционирует по строго определенным правилам и различает объекты только по синтактическим признакам. Это значит, что она однозначно задает некоторую модель. **Формальная система** — это математическая модель, задающая множество дискретных объектов с помощью описания некоторых исходных объектов и правил построения новых объектов из исходных и уже построенных.

Под *объектами* понимают символические и графические представления о ситуациях, состояниях, различных системах связей и информационных структур. Несмотря на разнообразие формальных систем, их объединяют общие свойства.

Дискретность, которая проявляется в том, что дискретный характер носят и сами объекты системы, и правила их соединений друг с другом конечным числом способов.

Формальность, суть которой заключается в соблюдении ряда принципов, таких, как принцип *педантизма* (все, что делается в системе, подчинено строго определенным правилам), *принцип явного описания* (все условия применения правил и действия, которые надо совершать при их применении, формализуются явно) и др.

Такой формальный подход принципиально важен в различных системах, имеющих дело с множеством однотипных объектов, при описании точных наук, а особенно при общении с ЭВМ, когда необходимо четко описать каждый шаг. Функционируя строго по правилам, формальная система задает некоторую модель, используемую в различных ситуациях. При реальном применении модели элементы формальной системы наполняются конкретным содержанием. Этот процесс называют *интерпретацией* формальной системы в некоторой предметной области. Уровень изучения формальных систем есть *метауровень*, т.е. уровень работы с объектами, которые используются для описания иных различных предметных областей.

Представлением системы называется любой способ рассмотрения объектов формальной системы как конкретных объектов, полученных эмпирически из опыта, при условии, что содержательные объекты сохраняют структуру формальных.

Различные науки наряду с другими языками широко используют некоторые искусственные языки, формализация которых осуществляется в рамках дискретной математики, что дает возможность саму математику рассматривать как специально организованный формальный язык. Словарем такого языка служит система символов, обозначающих математические объекты и переменные, а также операции над объектами и отношения между ними. Формулы и любая совокупность символов, отвечающая определенным требованиям, играют роль предложений этого языка. Важнейшая особенность формального математического языка заключается в том, что процесс перехода от одних формул к другим совершается по строго определенным правилам, не допускающим двусмысленного толкования. Характерной особенностью математики как науки является использование строгих доказательств, дедуктивных выводов, в отличие от других наук, таких, как физика, химия, биология, использующих индуктивные методы научного познания.

В отличие от естественного язык формальной системы представляет собой совокупность цепочек в некотором алфавите. К формальным в частности относятся искусственные языки, предназначенные для общения человека с машиной, — языки программирования. Правила формальной грамматики рассматривают правила вывода (продукции) как элементарные операции, которые применяются в определенной последовательности к исходной цепочке (аксиоме) и порождают лишь правильные цепочки. Последовательность правил, используемых в процессе порождения некоторой цепочки, является ее выводом. Определенный таким образом язык представляет собой формальную систему.

? С какой целью изучаются формальные системы?

Каковы возможности объектов, изучаемых в данной теории, т.е. какие множества могут порождаться формальными системами?

Для ответов на эти вопросы необходимо показать конкретные средства, аксиомы и правила выводов, которые могут привести к конкретным моделям. Итак, формальные системы задаются следующим образом.

1. Выделяем алфавит A — совокупность символов, которые следует понимать единственным определенным образом. Каждый элемент (символ) алфавита принято называть **буквой**, а совокупность символов — **словом** или **выражением** над A. Следовательно, слово длины m — кортеж, т.е. произвольный элемент из A^m. Например, пусть A — русский алфавит, дополненный запятой, точкой, арабскими цифрами и пробелом: $A = \{а, б, ..., ю, я, 0, 1, ..., 9, «,», «.», «_»\}$. Тогда $a_1 = 0вода_98{,}1 \in A^{10}$, $a_2 = ыва.ц \in A^5$. Очевидно, что множеством всех выражений над алфавитом A будет бесконечное множество $A \cup A^2 \cup A^3 \cup ... \cup A^m \cup ... = \bigcup_{m=1}^{\infty} A^m$, которое назовем C.

2. Из множества всех слов C выделяем некое подмножество $F \subset C$. Это подмножество весьма произвольно, т.е. вводится «руками», а в качестве характеристического свойства, отличающего F от $C \setminus F$, можно условно указать, что его составляют только те выражения, которые имеют смысл. Например, слова *ладья* и *лютый_мороз* над алфавитом A из п. 1 имеют смысл в русском языке (задание F), а слова *ъlотс* и 2*й*1*к*,*щ*_ — не имеют: *ъlотс* $\notin F$. Элементы множества F называются **формулами** над алфавитом A. Произвол задания F очевиден. Например, для понимающего русский язык человека *ладья* имеет смысл, а для человека, не знакомого с русским языком, нет особой разницы между наборами символов *ладья* и *ъlотс*, т.е. формулы над одним и тем же алфавитом он задаст по-другому.

3. Рассмотрим множество F^n. Пусть $f = \langle f_1, f_2, ..., f_n \rangle \in F^n$ — набор формул. Образуем декартово произведение $F^n \times F$ и выделим из него некое подмножество R: $R \subset F^n \times F$. Из гл. 1 известно, что такая конструкция называется *отношением* между двумя множе-

ствами. Это отношение назовем **правилом вывода**. Если пара (f, g), где $f = \langle f_1, f_2, ..., f_n \rangle \in F^n$, а $g \in F$, принадлежит R (т. е. $(f, g) \in R$ или $f R g$), то формула g называется **непосредственным следствием** формул $f_1, f_2, ..., f_n$ или **непосредственно выводимым выражением** из формул $f_1, f_2, ..., f_n$, полученным по правилу вывода R. Основным требованием считается конечность правил вывода: совокупность отношений $\{R_1, R_2, ..., R_k\}$, которую далее будем обозначать R, должна быть конечной. Если существует $R_l \in R$, такое, что $(\langle f_1, f_2, ..., f_n \rangle, g) \in R_l$, то g называется **непосредственным следствием** формул $f_1, f_2, ..., f_n$ безотносительно к одному определенному правилу вывода. Обозначение как $f_1, f_2, ..., f_n \models g$ или $\frac{f_1, f_1, ..., f_n}{g}$.

4. Во множестве формул F выделим подмножество $P \subset F$, элементы которого назовем **аксиомами** теории. Выбор этого подмножества опять же произволен: это те формулы, которые условно можно считать априорно истинными, т. е. истинными изначально и безоговорочно. Условность заключается прежде всего в том, что мы пока не определили понятие «истинный». В алгебре логики ему соответствовало значение «1» булевой функции или переменной, здесь же речь идет о произвольных алфавитах, не обязательно содержащих букву «1». Обычно стремятся к тому, чтобы число аксиом было по возможности наименьшим. Но может получиться так, что каждая аксиома из меньшего числа будет менее априорно понятной. Например, в подразд. 4.8 было показано, что все булевы функции можно описать с помощью одной функции, например штриха Шеффера. Но нам интуитивно ближе операции отрицание, конъюнкция и дизъюнкция в качестве базовых, так как они опираются на повседневный человеческий опыт и имеют аналогию в теории множеств.

5. Формула h называется **выводимой** из формул $\{f_1, f_2, ..., f_n\} = \Phi$ или **следствием** формул Φ, если существует кортеж формул $\langle g_1, g_2, ..., g_p \rangle \in F^p$, такой, что $g_p = h$ и $\forall i \le p \; g_i \in P \cup \Phi \cup G_i$, где $G_i = \{g \mid \Phi, g_1, g_2, ..., g_{i-1} \models g\}$. Каждая формула в этой цепочке является или аксиомой, или исходной формулой, или непосредственно выводимой из предыдущих. Тогда кортеж $\langle g_1, g_2, ..., g_p \rangle$ называется **выводом**, или **доказательством**, формулы h, а набор $\{f_1, f_2, ..., f_n\}$ — **гипотезами**, или **посылками**, что обозначается как $\Phi \vdash h$ или $f_1, f_2, ..., f_n \vdash h$. Обращаем внимание, что аксиомы *всегда* подразумеваются в качестве посылок, поэтому как гипотезы в явном виде они не пишутся. Очевидно, что если $\Phi \models h$, то $\Phi \vdash h$.

6. Если $\Phi = \varnothing$ и $\Phi \vdash h$, т. е. формула является следствием только аксиом, то h называется **теоремой**. Вместо записи $\varnothing \vdash h$ обычно используют $\vdash h$. Очевидно, что каждый элемент $g_1, g_2, ..., g_p$ вывода теоремы h также является теоремой.

7. Заданные алфавит, множество формул, аксиомы и правила вывода, т. е. совокупность $T = \langle A, F, P, R \rangle$, называются **формальной**

системой (теорией) T. Строго говоря, непосредственная выводимость и выводимость являются свойствами конкретной формальной системы T, поэтому должны употребляться знаки $\vdash_T$ и $\models_T$. Но в тех случаях, когда очевидно, о какой формальной системе идет речь, т.е. все выводы происходят в одной и той же формальной системе, указание на нее мы будем опускать и употреблять знаки $\vdash$ и $\models$.

Допустим, формальная система задана: выделены алфавит, множество формул, аксиомы и правила вывода.

? Как пользоваться построенной формальной системой? Во-первых, она должна быть понятной не только ее автору. Во-вторых, из нее должна следовать некая конструктивная теория, в рамках которой можно делать что-то определенное. Например, мы хотим решать уравнения, формулировать и доказывать теоремы, вычислять выражения, причем не обязательно математического характера, допустим, выводить законы генетики или социологии.

Очевидно, что нельзя универсально задать формальную теорию. Так, определяя и объясняя ее, кроме алфавита самой теории был использован русский язык и язык теории множеств. В этом случае, когда мы вынуждены прибегать к другим знакам, их придется брать из некоторой более общей теории — **метатеории**, предметный язык которой называют **метаязыком**.

Возникает проблема *интерпретации* полученных формул, перевода их на более доступный язык. Но любой перевод есть отображение слов из одного языка в другой, например английского в немецкий. Это языки над одним и тем же латинским алфавитом, но фонетическая транскрипция и лексическое значение слов различны. Итак, отображение x: $F \to L$ называется **интерпретацией** формальной теории $\langle A, F, P, R\rangle$ во множество содержательных высказываний метаязыка L. Метаязык должен содержать, помимо прочего, семантические характеристики высказываний, если таковые имеются. В частности, мы сопоставляли истинным высказываниям символ «1», а ложным — «0», т.е. эти элементы или слова «ложь», «истина» *должны* содержаться в метаязыке. Поскольку некоторые элементы L несут на себе семантическую характеристику, это значит, что определено отображение v: $L \to$ {истина, ложь}, которое отождествим со знакомым $B = \{0, 1\}$. Тогда автоматически определена суперпозиция $v \circ x$: $F \to B$, ставящая в соответствие некоторым формулам их семантическую характеристику. Пусть $f \in F$. Тогда, если высказывание $x(f)$ является истинным в смысле L (т.е. $v(x(f)) = 1$), то считают, что формула f *выполняется* в интерпретации x. Именно в конкретной интерпретации имеет смысл говорить об отрицании формулы.

О формулах теории вообще можно судить прежде всего по ее аксиомам.

? Что делать, если одни аксиомы теории и следствия из них будут выполняться в интерпретации x, а другие — нет?

Такая интерпретация не может являться адекватным представлением теории, ее *моделью*. Интерпретация x: $F \to L$ называется **моделью** формальной теории $T = \langle A, F, P, R \rangle$, если в ней выполняются *все* теоремы T. Модель нужна для более простого и наглядного описания формальной теории, а формальная теория является «законодательным актом» существующей модели.

Также можно допустить, что одна и та же формула будет выполняться в одной интерпретации, но не выполняться в другой. Понятно, что, например, аксиома не должна допускать подобные противоположные интерпретации, т.е. для любой интерпретации ее образ должен быть «тождественно-истинным». В гл. 4, где рассматривались булевы функции, такая формула названа *тавтологией*. Обобщим этот термин на формальные системы. Формула $f \in F$ формальной теории $T = \langle A, F, P, R \rangle$ называется **общезначимой**, или **тавтологией**, если она выполняется в *любой* интерпретации, и **противоречивой**, если в любой интерпретации ее образ ложен.

Если среди теорем формальной теории T нет противоречивых, т.е. $\forall f \in \{h \mid \varnothing \vdash_T h\}\ \exists x$: $v(x(f)) = 1$, то формальная теория T называется **семантически непротиворечивой**.

Требования, предъявляемые к формальным системам. Итак, для семантически непротиворечивой теории существует описывающая ее модель. Но практически понятием семантической непротиворечивости пользоваться трудно. Должен быть однозначный критерий, служащий инструментом выявления непротиворечивости. Им служит понятие *формальной непротиворечивости*. Формальная теория T называется **формально непротиворечивой**, если в ней не являются одновременно выводимыми формула f и ее отрицание. Такое определение непротиворечивости соответствует аристотелевой логике и является обывательски более понятным. Инструментом же практического применения этих понятий служит следующее утверждение: *формальная теория семантически непротиворечива тогда и только тогда, когда она формально непротиворечива*. Итак, формальная и семантическая непротиворечивость теории выполняются или не выполняются одновременно, поэтому далее будем просто употреблять словосочетание «непротиворечивость теории».

Также необходимыми свойствами формальных систем являются их *полнота* и *независимость*. Рассмотрим отдельно каждое из требований, накладываемых на формальную систему. Они называются критериями.

Непротиворечивость формальной системы S означает, что в ней нельзя из некоторого набора аксиом вывести одновременно два противоречивых суждения A и $\bar{A}$, что соответствует одному из

законов логики: $A \cdot \overline{A} = 0$. Покажем это на примере высказываний. То, что это тоже формальная теория, будет показано далее. Непосредственным выводом будет импликация-тавтология из произведения высказываний: $f_1, f_2, ..., f_n \models f$ суть $(f_1 \cdot f_2 \cdot ... \cdot f_n) \to f = 1$. Действительно, пусть получены два противоположных вывода, причем обе импликации истинны. Тогда $(S \to A) \cdot (S \to \overline{A}) = (\overline{S} \vee \vee A) \cdot (\overline{S} \vee \overline{A}) = \overline{S}$, что подразумевает или противоречивость системы S, или невыводимость формулы A.

Иллюстрацией примера того, что из двух противоположных посылок, одна из которых в двоичной логике всегда ложна, следует любая бессмыслица, могут служить рассуждения философа и математика Г.Х.Харди. На просьбу доказать, что из ложной посылки можно вывести все, что угодно, он предложил такие «рассуждения». Рассмотрим одновременно две противоположные посылки: $2 \cdot 2 = 4$ и $2 \cdot 2 = 5$ и докажем, что если они обе истинны, то Мак-Таггарт — папа римский. Итак, если $2 \cdot 2 = 4$ и $2 \cdot 2 = 5$, то $4 = 5$. Вычтем из обеих частей равенства число 3, имеем $1 = 2$, т.е. *двое* — папа римский и Мак-Таггарт суть *один* и тот же человек.

Полнота. В п. 4.8.3 подробно рассматривались полные системы булевых функций. Аналогичные рассуждения применимы к формулам в формальных теориях. Из аксиом должны следовать истинные (в смысле некой модели) высказывания, и, наоборот, для полноты нужно, чтобы *все* истинные высказывания модели выводились из аксиом. Если L — модель формальной теории T (x: $F \to L$), то T называется **полной**, если всякому истинному высказыванию из L соответствует теорема из T, т.е. $\forall l \in L$: $v(l) = 1$ $\exists t = x^{-1}(l) \in \{ h \mid \vdash h\}$.

Полнота такой системы означает, что тот набор элементарных функций или аксиом, который принят за базисный, *достаточен* для того, чтобы с их помощью вывести любые другие функции и суждения этой системы.

Независимость. В определении функционально полных систем функций (см. п. 4.8.3) *базисом* называлась такая функционально полная система, удаление любой функции из которой превращает ее в функционально неполную. Добавление же произвольной функции в базис не влияет на функциональную полноту, но лишает систему свойства базиса. Так, в $\mathbb{R}^3$ есть только три линейно *независимых* вектора, например декартов базис (0, 0, 1), (0, 1, 0) и (0, 0, 1). Любые четыре вектора уже будут зависимы, т.е. четвертый можно разложить по первым трем. То же справедливо для системы аксиом формальной теории.

Независимой системой аксиом непротиворечивой формальной теории $T = \langle A, F, P, R \rangle$ называется такая система, в которой любая аксиома не может быть выведена на основании всех остальных аксиом этой системы, т.е. $\forall a \in P$ $\{h \mid P \setminus \{a\} \vdash h\}$ $P = \varnothing$.

Примером может служить пятый постулат Евклида, который на протяжении двух тысячелетий пытались доказать или опровергнуть математики различных эпох и национальностей. Но лишь в середине XIX в. независимо друг от друга наш гениальный соотечественник, будущий ректор Казанского университета Н. И. Лобачевский и молодой венгерский математик Я. Бойаи, а также великий Карл Гаусс доказали независимость аксиомы параллельных (пятый постулат Евклида) от остальных аксиом и возможность отказа от него. Поэтому и геометрия Евклида, и геометрия Лобачевского являются независимыми системами, аксиоматика которых отличается лишь аксиомой параллельных. Однако это полностью изменяет выводы, получаемые в этих геометриях на основе сформулированной аксиоматики.

Схематически эти геометрии можно представить в виде пересекающихся множеств, общая часть которых включает в себя множество теорем, выводимых из общего для них набора аксиом. Это так называемая абсолютная геометрия.

Различия в геометриях Евклида и Лобачевского, возникшие из-за включения в них разных по смыслу аксиом о параллельных прямых, можно увидеть из сравнительного неполного анализа, приведенного в табл. 5.1.

Эти и другие теоремы геометрии Лобачевского, такие непонятные и противоречивые на первый взгляд, послужили толчком для анализа *непротиворечивости* построения формальной системы. Дело в том, что неприятие противоречивых теорем основано на нашем *личном опыте* и *здравом смысле*, что послужило причиной неприятия неевклидовой геометрии математиками различных уровней. И только когда была установлена возможность реализации неевклидовой геометрии, тогда удалось увидеть, что это стройная, логически выводимая система, которая отвечает требованиям *полноты*, *непротиворечивости*, *независимости*.

Одной из главных задач формальной теории систем является выявление необходимого минимума средств, с помощью которых можно описать любую формальную систему. Для формальных систем доказана теорема: *В любой формальной системе множество доказуемых в ней формул перечислимо.*

Для любого не более чем счетного множества M существует система подстановок, все выведенные слова которой содержатся в M. Наглядно формальную систему можно представить в виде графа с выделенной вершиной — аксиомой, у которого остальные вершины — слова, порождаемые этой системой в заданном алфавите. Ребра такого графа — применения продукций, а путь от корневой вершины к заданной — возможные выводы данного слова.

Полученные выводы вдохновили математиков. Идея построения таких непротиворечивых полных функционально замкнутых систем на рубеже XIX и XX вв. казалась реальной и достижимой в ближайшее время. Математики различных направлений работали

Таблица 5.1

Отличительные особенности геометрии Евклида и геометрии Лобачевского

Геометрия Евклида	Геометрия Лобачевского
Через точку, не лежащую на прямой, проходит одна и только одна прямая, параллельная данной	Через точку, не лежащую на прямой, проходит любое число нетождественных прямых, параллельных данной
Сумма углов любого треугольника равна π	Сумма углов треугольника — величина всегда меньше, чем π, причем дефект угла пропорционален площади треугольника
Существует четырехугольник, имеющий четыре прямых угла	Не существует четырехугольника с четырьмя прямыми $(\pi/2)$ углами
Существуют подобные, не равные между собой треугольники	Любые подобные треугольники обязательно равны между собой
Расстояние между параллельными прямыми — величина постоянная	Расстояние между параллельными прямыми неограниченно уменьшается в одном направлении и увеличивается в другом
Вокруг всякого треугольника можно описать окружность	Существуют треугольники, вокруг которых нельзя описать окружность
Геометрия Евклида реализуется на плоскости, причем справедлива теорема Пифагора $ds^2 = dx^2 + dy^2$	Геометрия Лобачевского реализуется на верхней полуплоскости, причем справедлива «теорема Пифагора» в виде $ds^2 = (dx^2 + dy^2)/y^2$
Евклидова плоскость — пространство постоянной нулевой кривизны	Плоскость Лобачевского — пространство постоянной отрицательной кривизны

над созданием таких систем аксиом, которые бы подчинялись требованиям непротиворечивости, полноты, независимости, выводимости. Так удалось доказать, что система аксиом геометрии Лобачевского непротиворечива, если непротиворечива аксиоматика Евклида. Система аксиом евклидовой геометрии, в свою очередь, непротиворечива, если непротиворечива аксиоматика действительных чисел. Арифметика действительных чисел непротиворечива, если непротиворечива аксиоматика натуральных чисел.

Таким образом, непротиворечивость различных обширных областей, охватывающих по меньшей мере все разделы элементарной математики, оказалась сведенной к проблеме непротиворечивости

арифметики натуральных чисел. Аналогично происходил процесс установления независимости некоторой системы аксиом.

Неполнота. Формальная система S как математический объект полностью определяется заданием языка, системы аксиом и правил вывода.

Напомним, что построенная таким образом система S должна обладать свойствами непротиворечивости, полноты, выводимости, независимости.

? Очень важен также и вопрос о метаязыке: это язык самой системы S или какой-то другой системы, например естественный язык?

В поисках ответа на эти вопросы к 1920 г. была разработана так называемая *программа Гильберта* формализации аксиоматических теорий и использования ее для математических доказательств. Известный немецкий ученый Давид Гильберт (1862—1943) предложил развернутую программу обоснования математики, с помощью которой он собирался доказать непротиворечивость классической математики.

Кроме этого, программа сведения математики к логике, к возможности формализации математической системы, в частности арифметики, была предпринята в фундаментальном трехтомном труде английских математиков, философов и логиков Бертрана Рассела (1872—1970) и Алфреда Уайтхеда (1861—1947), который назывался *Principia Mathematica* (1910—1913).

Однако в 1931 г. в одном из журналов была опубликована статья австрийского ученого Курта Гёделя (1906—1978) «О формально неразрешимых противоречиях *Principia Mathematica* и родственных систем». В ней была доказана невозможность полной формализации не только человеческих знаний и мышления, но и полной формализации арифметики натуральных чисел. Эта статья, содержавшая всего 25 страниц, открыла новую эпоху в истории математики и логики. Рассматривая непротиворечивую систему S, которую удалось свести к системе натуральных чисел, Гедель построил формулу f и доказал, что она не может быть выведена в этой системе аксиом. Таким образом, в этой системе с помощью известных аксиом и правил выводов нельзя доказать ни справедливость формулы f, ни справедливость формулы $\bar{f}$.

Поэтому невозможна *полная* формализация арифметики натуральных чисел — достаточно обширной и известной формальной системы. Вывод о неполноте аксиоматики натуральных чисел одновременно доказывает и невозможность сведения математики к формальной логике.

Смысл второй теоремы, доказанной Гёделем, состоит в том, что непротиворечивость любой достаточно мощной математической системы не может быть выявлена средствами этой системы и описана на языке этой системы. Это значит, что необходимо применение языка другой системы для вывода новых формул, т. е. необходимы средства более мощной формальной системы. Поэтому можно утверждать, что полная формализация не может осуществиться на некотором этапе развития науки.

Эти две теоремы можно обобщить: *в любой формальной системе есть формулы, недоказуемые средствами только этой системы.* Теоремы Гёделя

имеют настолько важное значение в математике и логике, что признаны одним из выдающихся открытий XX в. Но удивительно в открытиях Гёделя то, что несостоятельность, ограниченность логических методов доказаны с помощью этих же формальных логических методов!

? Что это, очередной логический парадокс? Противоречие?

Ответом на этот вопрос служат знаменитые слова О. Хевисайда: «Логика непобедима, потому что одолеть ее можно только с помощью логики!»

В отличие от Гильберта, Рассела и других математиков и логиков К. Гёдель, используя в своих рассуждениях такой сильный метод, как абстракция актуальной бесконечности, смог доказать два важных положения.

1. Всякая естественная непротиворечивая формализация S арифметики натуральных чисел или другой математической теории, содержащей арифметику (например, теории множеств),– неполная и непополнимая, в том смысле, что:

• в S имеются содержательно истинные формулы f, такие, что ни f, ни f не являются выводимыми в S ,т.е. являются **неразрешимыми** в S;

• любые попытки расширить систему S каким бы то ни было конечным множеством дополнительных аксиом ΔS приводят к тому, что в новой системе S' неизбежно появятся свои неразрешимые формулы.

2. Если формализованная арифметика натуральных чисел в действительности непротиворечива, то доказательство этого утверждения, проведенное средствами, формализуемыми в ней самой, невозможно. Хотя утверждение о ее непротиворечивости может быть выражено на ее собственном языке.

Следствием теорем Гёделя явилось установление связи между формализацией доказательства непротиворечивости арифметики и использованием понятия бесконечности.

Дело в том, что когда вопрос о непротиворечивости разных разделов математики был сведен к проблеме доказательства непротиворечивости арифметики натуральных чисел, дальнейшим обобщением могло быть лишь сведение к непротиворечивости теории множеств как к самой абстрактной математической конструкции. А как мы знаем из гл. 1, теория множеств как раз и содержит в себе ряд противоречий (определение операций над множествами через логические функции и наоборот), неразрешимых на данном уровне развития науки.

Познакомившись с правилами построения формальных систем, сравним с ними известные формальные системы для того, чтобы увидеть в действии эти теоретические положения.

Уточним, что такой подход к построению формальных систем называется **аксиоматическим**. Стремление математиков формализовать различные математические теории (иногда их называют математическими конструкциями) вызвано желанием найти общий метод не только для имеющихся теорий, но и для *конструирования*

новых. Использование аксиоматики как модели построения систем и их дальнейшего развития могло бы, с одной стороны, систематизировать все имеющиеся знания, с другой — алгоритмизировать процесс поиска новых знаний. И такие примеры в истории математики есть. Так, А. Н. Колмогоров применил аксиоматический метод к построению теории вероятностей. Благодаря этому из прикладной математической науки — «золушки математики» — теория вероятностей превратилась в XX в. в одну из самых бурно развивающихся фундаментальных наук, имеющих широкое практическое применение.

5.2. Исчисление высказываний

Математика есть способ называть
разные вещи одним именем.

А. Пуанкаре

В этом подразделе будет показано, как можно формализовать систему высказываний, придавая им семантическую характеристику. Подчеркнем сразу разницу между той трактовкой (в частности, в булевых функциях) высказываний и тем, что будет описано в этой главе. В гл. 3—4 рассуждения исходили от постулированных значений булевых функций и интуитивного соответствия, и многие тавтологии являлись следствиями. Здесь же будет предложена жесткая аксиоматика и доказана их формальная непротиворечивость, а затем будет показано, что старые представления являлись одной из возможных интерпретаций.

Исчисление высказываний есть формальная теория T, в которой заданы:

- алфавит A:
  связки $\neg$ (или $\overline{\ }$), $\to$; (дополнительно можно ввести связки $\wedge$ и $\vee$: $A \wedge B \overset{def}{=} \overline{A \to \overline{B}}$, $A \vee B \overset{def}{=} \overline{A} \to B$);
  (,) — служебные символы, позволяющие определить порядок выполнения связок;
  D, B, ..., A_1, B_1, ... — переменные высказывания;
- формулы F:
  переменные есть формулы;
  если A, B — формулы, то $(\overline{A})$ и $(A \to B)$ — формулы;
- аксиомы P:
  P_1: $(A \to (B \to A))$;
  P_2: $((A \to (B \to C)) \to ((A \to B) \to (A \to C)))$;
  P_3: $((\overline{B} \to \overline{A}) \to ((\overline{B} \to A) \to B))$;
- правило *modus ponens* (*mp*): $\dfrac{A,\ A \to B}{B}$ — правило заключения

(от лат. *modus* — способ, правило; *ponens* — отделение, заключение).

Заметим, что символы $\neg$ или $\overline{\ }$, обозначающие отрицание, эквивалентны, но использование черты позволяет компактно и более понятно описать отрицание сложных выражений, а уголок позволяет правильно оценить длину полученного слова. Например, $B \to \overline{\overline{C} \to D}$ и $B \to \neg((\neg C) \to D)$ — суть одно и то же слово, состоящее из 11 букв, где связка $\to$ есть импликация. В исчислении высказываний множество аксиом теории T и множество правил вывода бесконечны, так как зависят от бесконечного числа различных высказываний-переменных, хотя задаются соответственно тремя и одним правилом.

Пусть $A_1, A_2, ..., A_n, B$ — составные высказывания. Говорят, что высказывание B является **логическим следствием** высказываний $A_1, A_2, ..., A_n$, если при всех значениях элементарных высказываний, при которых все $A_1, A_2, ..., A_n$ истинны (в любой интерпретации), будет истинным и высказывание B. Символическая запись логического следования B из $A_1, A_2, ..., A_n$ имеет вид $A_1, A_2, ..., A_n \Rightarrow B$ и читается так: «Если (выполнены) $A_1, A_2, ..., A_n$, то (выполнено) B».

Два утверждения A и B называются **логически эквивалентными**, если $A \Rightarrow B$ и $B \Rightarrow A$.

Сформулируем некоторые *теоремы*, доказываемые в исчислении высказываний:

1. $\vdash A \to A$.
2. $A \vdash B \to A$ — введение импликации.
3. $\Phi, A \vdash B \Leftrightarrow \Phi \vdash A \to B$ — теорема дедукции.
4. $A \vdash B \Leftrightarrow \vdash A \to B$ — следствие из теоремы дедукции (при $\Phi = \varnothing$).
5. $A \to B, B \to C \vdash A \to C$ — правило транзитивности.
6. $A \to (B \to C), B \vdash A \to C$ — правило сечения.
7. $\vdash \overline{\overline{A}} \to A$.

7′. $\vdash A \to \overline{\overline{A}}$.

8. $\vdash \overline{A} \to (A \to B)$.
9. $\vdash (\overline{B} \to \overline{A}) \to (A \to B)$.

9′. $(\overline{B} \to \overline{A}) \vdash (A \to B)$ — противопоставление обратному.

Кроме правила *mp* должно быть такое правило, которое позволило бы обращаться со сложными высказываниями, как с простыми. Ранее было показано, как это удобно, когда сложные булевы формулы назывались одной пропозиционной переменной, что позволяло производить операции над столбцами для нахождения таблицы значений булевой функции. Например, $(a \to b) \vee x = y \vee x$, где $(a \to b) = y$.

Можно показать, что такое правило — правило подстановки — является производным от основных правилом вывода и логически мотивировано.

Правило выводов обычно включает в себя два правила, достаточных для проведения формальных умозаключений.

Правило подстановки (*ms*). Если формула B является частным случаем формулы A, то B непосредственно выводима из A. Другая формулировка правила подстановки имеет вид: если E — выводимая формула, содержащая символ A (т. е. $E(A)$), то выводима формула $E(B)$, полученная из E заменой всех вхождений A на произвольную формулу B: т. е. $\dfrac{E(A)}{E(B)}$.

Правило *modus ponens* (*mp*). Если набор формул A, B, C является частным случаем набора формул A, $A \to B$, то формула C является непосредственно выводимой из формул A и B. Тогда доказанные правила могут стать новыми правилами вывода. Так, теорему 2 о введении импликации можно сделать новым правилом вывода: $A \models B \to A$.

Приведем некоторые общие правила.

Правило 1. Общие свойства символа $\models$ ($\vdash$).

При $n > 1$ $A_1, A_2, \ldots, A_n \models A_1$; $A_1, A_2, \ldots, A_n \models A_2, \ldots, A_1, A_2, \ldots, A_n \models A_n$ — из набора формул непосредственно выводима каждая.

Если $A_1, A_2, \ldots, A_n \models C$, то $A_1, A_2, \ldots, A_n, B_1, \ldots, B_m \models C$, т. е. добавление лишних формул при сохранении непротиворечивости не влияет на выводимость.

Если при $m > 0$ и $n > 0$ $A_1, A_2, \ldots, A_m \models B_1, \ldots, A_1, A_2, \ldots, A_m \models B_n$ и $B_1, B_2, \ldots, B_n \models C$, то $A_1, \ldots, A_m \models C$ — транзитивность непосредственной выводимости.

Те же свойства справедливы для выводимости $\vdash$.

Правило 2 введения и удаления логических знаков.

Пусть A, B, C — произвольные высказывания. Обозначим Φ — конечный список высказываний, быть может пустой. Правила введения и удаления логических знаков представим в виде табл. 5.2.

С помощью этих правил можно доказать другие равносильности.

Таблица 5.2

Правила введения и удаления логических знаков

Операция	Правила введения	Правила удаления
$\to$	$(\Phi, A \vdash B) \Leftrightarrow (\Phi \vdash A \to B)$	$A, A \to B \vdash B$
$\wedge$	$A, B \vdash A \wedge B$	$A \wedge B \vdash A$; $A \wedge B \vdash B$
$\vee$	$A \vdash A \vee B$	Если $\Phi, A \vdash C$ и $\Phi, B \vdash C$, то $\Phi, A \vee B \vdash C$
$\neg$	Если $\Phi, A \vdash B$ и $\Phi, A \vdash \overline{B}$, то $\Phi \vdash \overline{A}$	$\overline{\overline{A}} \vdash A$; $\overline{A} \vdash A \to B$ (слабое удаление)
$\equiv$	$A \equiv B, B \equiv A \vdash A \equiv B$	$A \equiv B \vdash A \to B, A \equiv B \vdash B \to A$

Докажем, например, справедливость правила противопоставления обратного $A \to B \vdash \overline{B} \to \overline{A}$ и проанализируем доказательство:

1) $A \to B,\ \overline{B},\ A \vdash \overline{B}$;

2) $A \to B,\ \overline{B},\ A \vdash B$ — удаление $\to$;

3) $A \to B,\ \overline{B} \vdash \overline{A}$ введение $\neg$;

4) $A \to B \vdash \overline{B} \to \overline{A}$, введение $\to$.

Правило 3 подстановки вместо элементарных высказываний.

Пусть B — некоторое составное высказывание, в которое входят элементарные высказывания $a_1, \ldots, a_n$; B' — высказывание, полученное из B одновременной подстановкой составных высказываний $A_1, \ldots, A_n$ вместо $a_1, \ldots, a_n$. Тогда, если $\vdash B$, то $\vdash B'$.

Благодаря этому правилу разрешается осуществлять подстановку в основные равносильности, логические следования и правила доказательства тавтологий. Замена элементарных высказываний на произвольные составные не нарушит строгости рассуждений.

Правило 4 о замене.

Пусть C_A — составное высказывание, содержащее в качестве своей части составное высказывание A, а C_B — высказывание, полученное из A заменой A на B. Тогда, если $\vdash C_A$, то $\vdash C_B$.

С учетом этих правил, формул и аксиом логика высказываний является **полной теорией**, так как все логические следствия выводимы, а все, что выводится, будет логическим следствием из данных посылок.

Логическое следование есть конечная алгоритмическая операция в отличие от логической выводимости.

С помощью подстановок в любых формулах A и B заменятся пропозиционными переменными либо другими формулами, что не нарушает требований, предъявляемых к теории. При этом логические операции между формулами и переменными сохраняются.

Теоремами теории T являются тавтологии, т.е. общезначимые формулы, и только они: $\vdash A \Leftrightarrow A$ — тавтология. Прямым следствием этой теоремы является формальная непротиворечивость и полнота теории T. Итак, всякая теорема исчисления высказываний является тождественно-истинным высказыванием, и, наоборот, всякая тождественно-истинная формула является теоремой исчисления высказываний. Исчисление высказываний выполняет задачу порождения общелогических законов как тождественно-истинных высказываний.

В построенной формальной теории T выводимы многие важные для практических целей формулы. Они получаются «перефразированием» теорем с добавлением дополнительных связок. Они называются, естественно так же, как и связки булевых переменных: $\vee$ — дизъюнкция, $\wedge$ — конъюнкция.

Приведем некоторые важные формулы.

1. $a \vdash a \vee b$, т.е. к заключению можно добавить другое высказывание с помощью дизъюнкции.

2. $a, b \vdash a \wedge b$.

3. $a \wedge b \vdash a$.

4. $a \to c,\ b \to c \vdash (a \vee b) \to c$ — распределительный закон относительно дизъюнкции посылок.

5. $a \to b, \overline{a} \to b \vdash b$.

6. $\vdash a \vee \overline{a}$ — закон исключенного третьего.

Для доказательства других тавтологий нужно использовать аналитические правила выводов через логические следствия. Например, достаточно часто применяется *теорема дедукции* $\Phi, A \vdash B \Leftrightarrow \Leftrightarrow \Phi \vdash A \to B$, где Φ — список формул, A и B — отдельные формулы. Известный из школьного курса геометрии метод доказательства от противного можно сформулировать как правило исчисления высказываний ($\Gamma, A \vdash B$ и $\Gamma, A \vdash \overline{B}$) $\Rightarrow (\Gamma \to \overline{A})$ — правило введения отрицания. Эти теоремы могут быть доказаны *методом математической индукции* (см. подразд. 5.7).

Построенная теория T не является единственно возможной аксиоматизацией исчисления высказываний. Существуют и другие аксиоматизации исчисления высказываний (табл. 5.3).

Таблица 5.3

Аксиоматизации исчисления высказываний

	Гильберт, Аккерман (1938)	Клини (1952)	Россер (1953)
A	Основные: $\vee$, $\neg$; дополнительные: $(A \to B \overset{def}{=} \overline{A} \vee B)$	$\neg$, $\wedge$, $\vee$, $\to$	Основные: $\neg$, $\wedge$; дополнительные: $(A \to B \overset{def}{=} \overline{A \wedge \overline{B}})$
P	$A \vee A \to A$ $A \to A \vee B$ $A \vee B \to B \vee A$ $(B \to C) \to (A \vee B \to \to A \vee C)$	$A \to (B \to A)$ $(A \to (B \to C)) \to ((A \to \to B) \to (A \to C))$ $A \wedge B \to A$ $A \wedge B \to B$ $A \to (B \to A \vee B)$ $A \to (A \vee B)$ $B \to (A \vee B)$ $(A \to C) \to ((B \to C) \to \to ((A \vee B) \to C))$ $(A \to B) \to ((A \to \overline{B}) \to \to \overline{A})$ $\overline{A} \overset{def}{=} \neg(\overline{A}) \to A$	$A \to A \wedge A$ $A \wedge B \to A$ $(A \to B) \to (\overline{B} \wedge \wedge C) \to \overline{C \wedge A}$
R	*modus ponens*	*modus ponens*	*modus ponens*

Известно, что в исчислении высказываний теоремами являются общезначимые формулы, поэтому автоматизация доказательства осуществляется в виде проверки общезначимости формул с помощью таблиц истинности. Проверка истинности формул производится для любого конечного набора значений переменных. Таким образом можно алгоритмизировать процесс доказательства теорем, но не процесс вывода теорем из аксиом.

5.3. Логика предикатов

> Результат считается красивым, если из малого числа условий удается получить общие заключения, относящиеся к широкому кругу объектов.
>
> *Б. Гнеденко*

С помощью формальных теорий можно описать обширный класс высказываний, называемых предикатами. Сначала дадим определение исчисления предикатов как формальной теории, а затем подробно остановимся на интерпретации. Сразу отметим, что теория будет дана в значительно упрощенном виде: многие моменты мы опустим.

Формальная теория $S = \langle A, F, P, R \rangle$ называется **исчислением предикатов** первого порядка, если заданы алфавит, формулы, аксиомы и правила вывода.

1. Алфавит A:

$x, y, z...$ — предметные переменные, принимающие конкретные значения из некоего множества D. Тогда $x_0, y_0, z_0...$ — значения предметных переменных, т.е. предметные постоянные (константы);

$p, q, r...$ — переменные высказывания, принимающие два значения: 1 (истина) и 0 (ложь). Тогда $p_0, q_0, r_0...$ — фиксированные значения;

P, Q, F — переменные, символизирующие само высказывание; $P_0, Q_0( )$ — постоянные предикаты;

$\rightarrow, \neg$ — символы логических операций; дополнительно используются символы $\wedge, \vee$;

$\forall, \exists$ — кванторы общности и существования;

служебные символы) , (— нужны для установления порядка выполнения связок и области действия кванторов;

можно использовать также знаки ! — единственность, : — «такой, что...» и другие символы метаязыка. Например, $\forall x \in \{3, 4, ..., 25\}\ \exists!\ y \in (0, +\infty)$: $x = y^2$.

2. Формулы: F:

переменные есть формулы;

если A, B — формулы, x — переменная, то $A(x)$, $(\overline{A})$, $(A \to B)$, $\forall x\, A(x, \ldots)$ и $\exists x\, A(x, \ldots)$ — формулы.

3. Аксиомы:

исчисления высказываний: P_1: $(A \to (B \to A))$; P_2: $((A \to (B \to C)) \to ((A \to B) \to (A \to C))$; P_3: $(\overline{B} \to \overline{A}) \to ((\overline{B} \to A) \to B)))$;

кванторные: P_4: $\forall x A(x) \to A(y)$; P_5: $A(x) \to \exists y A(y)$.

4. Правила вывода:

R_1: $\dfrac{A, A \to B}{B}$ — *modus ponens*;

R_2: $\dfrac{A \to B(x)}{A \to \forall x B(x)}$ — введение квантора общности;

R_3: $\dfrac{A(x) \to B}{\exists x A(x) \to B}$ — введение квантора существования.

Построенная формальная теория S описывает весьма общие объекты, поэтому нужно ее интерпретировать в то, с чем можно работать.

Под предикатом будем понимать следующее. Произвольная функция P: $M^n \to B$, заданная на произвольном множестве M, называется n-местным **предикатом** $P(x_1, \ldots, x_n)$. Причем в такой трактовке P уже задает семантическую характеристику.

Заметим, что предикаты дают возможность математически анализировать суждения. В классической логике **предикатом** называется *сказуемое* суждения, т. е. то, что утверждается или отрицается относительно субъекта этого суждения, имени предмета мысли, фиксирующее его определенные свойства. А в математической логике понятие предиката рассматривается как тождественное суждению, содержащему местоимения, т. е. пропозиционная функция, аргументами которой служат *имена*. Например, о высказывательной форме «Он получил специальность программиста» нельзя сказать, истинна она или ложна, пока не произведена замена местоимения «он» на существительное: «М. А. Иванов стал программистом» (истинно), «Дом стал программистом» (ложно).

Далее подробно опишем все составляющие формальной теории исчисления предикатов в такой интерпретации.

Язык логики предикатов. Символами X, Y, Z, X_i, Y_i, ... в логике предикатов принято обозначать *предметные переменные*, т. е. отдельные *предметы — имена*. Они могут быть простыми и сложными. Если такие предметы (имена) не содержат дополнительной информации о себе, то они называются **собственными** (простыми), например «земля», «студент» и др. Если такое имя содержит наряду с самим предметом его отдельные свойства, то оно является **сложным**, например «автор романа «Анна Каренина», «перпендикулярные прямые», «взаимно-однозначное соответствие» и др.

Символами a, b, c, a_i, b_i, ... принято обозначать **константы** или *предметные постоянные*, т. е. конкретные значения имен предметов

из указанной предметной области. Высказывательные формы, входящие в предикаты, называют также **пропозиционными функциями**, или **предикаторами**.

Рассмотрим двухместную высказывательную форму $x \geq y + 2$, где x определено на множестве $M_x = \{3, 5\}$, а y — на множестве $M_y = \{1, 5, 8\}$. Этой форме соответствуют два предиката P_1 и P_2, определенные на множествах $M_x \times M_y$ и $M_y \times M_x$, которые образованы с помощью декартова произведения двух множеств. Сравним предикаты $P_1(x, y)$ и $P_2(x, y) = P_1(y, x)$ по табл. 5.4.

Видно, что двухместная высказывательная форма, вообще говоря, определяет два различных предиката. Для установления взаимно-однозначного соответствия между n-местной ($n \geq 2$) высказывательной формой и соответствующим предикатом принято устанавливать для этих переменных определенный порядок. Так, если для n переменных высказывательной формы $P(x_1, x_2, ..., x_n)$ установлен порядок $(x_{i_1} \prec x_{i_2} \prec, ..., \prec x_{i_n})$, а $M_{i_1}, M_{i_2}, ..., M_{i_n}$ — множества значений этих переменных, то такая форма задает предикат на множестве, составляющем декартово произведение имеющихся множеств: $M = M_{i_1} \times M_{i_2} \times ... \times M_{i_n}$. Принято одноместный предикат называть **предикатом-свойством**, n-местный (для $n > 1$, $n \in \mathbb{N}$) — **предикатом-отношением**, 0-местный предикат — **высказыванием**.

Полный прообраз единицы (1) при P назовем **множеством истинности** $T(P)$ предиката P (от англ. *truth* — истина): $T(P) = P^{-1}(1) = \{x \mid x \in M^n, P(x) = 1\}$.

В рассмотренном примере множество истинности $T(P_1)$ предиката $P_1(x)$: $x \geq y + 2$ равно $T(P_1) = \{(3, 1), (5, 1)\}$, а для предиката $P_2(x)$ множество истинности $T(P_2) = \{(5, 3), (8, 3), (8, 5)\}$.

Как мы знаем из гл. 1, любое непустое множество содержит два подмножества: само себя и пустое. Это свойство автоматически выделяет из области определения два случая.

Таблица 5.4

Сравнение предикатов $P_1(x, y)$ и $P_2(x, y) = P_1(y, x)$

$M_x \times M_y$: (x, y)	$x \geq y + 2$	$P_1(x, y)$	$M_y \times M_x$: (y, x)	$y \geq x + 2$	$P_2(x, y)$
(3, 1)	$3 \geq 3$	1	(1, 3)	$1 \geq 5$	0
(3, 5)	$3 \geq 7$	0	(1, 5)	$1 \geq 7$	0
(3, 8)	$3 \geq 10$	0	(5, 3)	$5 \geq 5$	1
(5, 1)	$5 \geq 3$	1	(5, 5)	$5 \geq 7$	0
(5, 5)	$5 \geq 7$	0	(8, 3)	$8 \geq 3$	1
(5, 8)	$5 \geq 10$	0	(8, 5)	$8 \geq 7$	1

Тождественно-истинным называется предикат, истинный всюду на области определения: $T(P) = D(P)$. **Тождественно-ложным** называется предикат, множество истинности которого пусто: $T(P) = \varnothing$.

Два предиката в одной и той же области определения **различны** в том и только в том случае, если их множества истинности не совпадают. Это определение совпадает с отрицанием обычного определения равенства функций.

Логические операции (связки) над предикатами. Связки, аналогичные связкам булевой алгебры и исчисления высказываний, соответствуют логическим операциям над предикатами. Операции над n-местными предикатами вводятся аналогично одноместным.

Пусть, например, $P(x, ...)$ и $Q(x, ...)$ — предикаты, которые определены на множестве D, причем $T(P)$ и $T(Q)$ — их множества истинности.

Отрицанием предиката $P(x, ...)$ называется предикат $\overline{P(x)}$, также определенный на множестве D и истинный при тех значениях переменной x, при которых $P(x, ...)$ ложен, т.е. $T(\overline{P}) = D \setminus T(P)$ (рис. 5.1).

Например, предикат $P(x)$: «x — простое число» определен на множестве $D = \mathbb{Z}$ целых чисел, а его областью истинности являются только простые числа, т.е. числа, имеющие два делителя: x и 1. Тогда предикат «x — составное (целое) число», также определенный на $\mathbb{Z}$, будет отрицанием предиката $P(x)$, т.е. $\overline{P(x)}$, а его областью истинности будет множество всех целых составных чисел (имеющих три и более делителей): $T(\overline{P(x)}) = D \setminus T(P(x))$.

Аналогично вводятся и остальные операции.

Конъюнкция предикатов $P(x, ...)$ и $Q(x, ...)$ есть новый предикат $P(x) \wedge Q(x)$, определенный на множестве D и истинный при тех значениях переменной x, при которых истинны одновременно *оба* предиката $P(x, ...)$ и $Q(x, ...)$, поэтому $T(P \wedge Q) = T(P) \cap T(Q)$ (рис. 5.2).

Рассмотрим примеры.

1. Для предикатов $P(x)$: «x — четное число» и $Q(x)$: «x кратно 7» конъюнкцией $P(x) \wedge Q(x)$ служит предикат «x — четное *и* кратное 7 число» или «x — число, кратное 14».

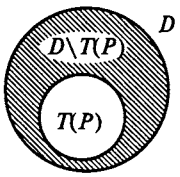


Рис. 5.1. Множество истинности предиката $\overline{P(x)}$

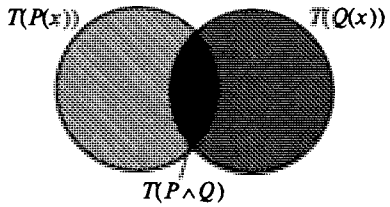


Рис. 5.2. Множество истинности конъюнкции предикатов

2. Решить систему неравенств $\begin{cases}2x \le 16,\\ 3x > 15\end{cases}$ означает: решить первое неравенство, т.е. определить $T(P_1)$, решить второе неравенство — определить $T(P_2)$: $\begin{cases}2x \le 16,\\ 3x > 15\end{cases} \Leftrightarrow \begin{cases}x \le 8,\\ x > 5.\end{cases}$ Определить, при каких x верны *и* первое, *и* второе неравенства. В данном случае система $\begin{cases}x \le 8,\\ x > 5\end{cases}$ означает конъюнкцию высказываний $(x \le 8) \wedge (x > 5) \Leftrightarrow$ $\Leftrightarrow 5 < x \le 8$, а ответ является пересечением $T(P_1)$ и $T(P_2)$ (рис. 5.3), т.е. интервалом $5 < x \le 8$.

Обратите внимание, что в итоговый ответ вошла конъюнкция высказываний, эквивалентных данным в условии, а не самих исходных.

Дизъюнкцией предикатов $P(x, ...)$ и $Q(x, ...)$ называется предикат $P(x) \vee Q(x)$, определенный на множестве D и истинный при тех значениях переменной x, при которых истинен хотя бы один из предикатов $P(x)$ или $Q(x)$. Поэтому $T(P \vee Q) = T(P) \cup T(Q)$ (рис. 5.4).

Например, для предикатов $P(x)$: «x — число, кратное 3» и $Q(x)$: «x — число, оканчивающееся на 3», определенных на $\mathbb{N}$, дизъюнкцией $P(x) \vee Q(x)$ служит предикат: «x — число *или* кратное 3, *или* оканчивающееся цифрой 3».

Так, при решении уравнений (неравенств), левая часть которых есть произведение нескольких множителей, а правая — нуль, они разбиваются на совокупность уравнений (неравенств). Например, $x^2 - 8x - 20 = 0 \Leftrightarrow (x - 10)(x + 2) = 0 \Leftrightarrow x - 10 = 0$ (P_1) *или* $x + 2 = 0$ (P_2). Таким образом, нужно найти $T(P_1) = \{10\}$, $T(P_1) = \{-2\}$ и их объединение: $T(P) = \{-2, 10\}$.

Импликацией предиката $P(x, ...)$ в $Q(x, ...)$ называется предикат $P(x) \to Q(x)$, определенный на множестве D и ложный только при тех значениях переменной x, при которой предикат $P(x, ...)$ истинен, а предикат $Q(x, ...)$ ложен. В полном соответствии с

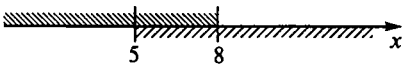


Рис. 5.3. Графическое решение системы неравенств $\begin{cases}2x \le 16,\\ 3x > 15\end{cases}$

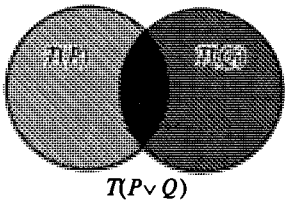


Рис. 5.4. Множество истинности дизъюнкции предикатов

формулой алгебры логики $a \to b = \bar{a} \vee b$ имеем: $P \to Q = \overline{P} \vee Q$ и $T(P \to Q) = (D \setminus T(P)) \cup T(Q)$ (рис. 5.5).

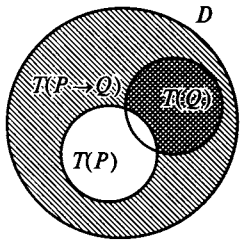


Рис. 5.5. Множество истинности импликации предикатов

Например, импликацией предикатов $P(x)$: «x — нечетное число» и $Q(x)$: «x кратно 5», определенных на $\mathbb{N} \cup \{0\}$, служит предикат $P(x) \to Q(x)$: «Если x — нечетное число, то x кратно 5». Здесь $T(P) = \{y \mid (y \bmod 2) = 1\} = \{1, 3, 5, ...\}$; $T(Q) = \{y \mid (y \bmod 5) = 0\} = \{0, 5, 10, ...\}$. Тогда $D/T(P) = \{y \mid (y \bmod 2) = 0\} = \{0, 2, 4, ...\}$; $T(P \to Q) = (D \setminus T(P)) \cup T(Q) = \{y \mid (y \bmod 2) = 0$ или $(y \bmod 5) = 0\} = \{0, 2, 4, 5, 6, ...\}$. Импликация верна, если число кратно двум или пяти.

Замечание. Поскольку в данном нами алфавите связка $\to$ является основной, а $\vee$ и $\wedge$ — дополнительными, то дадим введение конъюнкции и дизъюнкции через $\to$ и $\neg$: $P \vee Q \overset{def}{=} \overline{P} \to Q$, $P \wedge Q \overset{def}{=} \overline{P \to \overline{Q}}$.

Эквиваленцией предикатов $P(x, ...)$ и $Q(x, ...)$ называется предикат $P(x) \equiv Q(x)$, определенный на множестве D и истинный при тех значениях переменной x, при которых либо оба предиката истинны, либо оба предиката ложны. Поэтому $T(P \Leftrightarrow Q) = (T(P) \cap T(Q)) \cup ((D \setminus T(P)) \cap (D \setminus T(Q)))$. В силу законов Де Моргана $(T(P) \cap T(Q)) \cup (D \setminus (T(P) \cup T(Q))) = D \setminus T(P) \Delta T(Q))$. Если $T(P) = T(Q)$, то $T(P \equiv Q) = D$. Например, эквивалентны предикаты $P(x)$: «x — натуральное число, кратное 3» и $Q(x)$: «x — натуральное число, сумма цифр которого делится на 3».

Следование и эквиваленция. Высказывательная форма Q_2 **следует** из высказывательной формы Q_1, если импликация $Q_1 \to Q_2$ обращается в истинное высказывание при любых наборах значений переменных, входящих в нее. Для операции логического следования принято обозначение $Q_1 \Rightarrow Q_2$.

Пусть даны предикаты $Q_1(x_1, x_2, ..., x_n)$ и $Q_2(x_1, x_2, ..., x_n)$, а их множества истинности соответственно $T(Q_1)$ и $T(Q_2)$. Поскольку $Q_1 \Rightarrow Q_2$, то если $(x_1, x_2, ..., x_n) \in T(Q_1)$, т.е. Q_1 истинна, то должна быть истинна Q_2, т.е. $(x_1, x_2, ..., x_n) \in T(Q_2)$. Поскольку такое свойство должно быть у любого элемента из $T(Q_1)$, то это определение подмножества. Итак, $T(Q_1) \subset T(Q_2)$.

Пусть даны два предиката, определенные на одном множестве. Высказывательные формы Q_1 и Q_2 назовем **равносильными**, если при любом наборе значений переменных, входящих в них, высказывательные формы принимают одинаковые значения истинности: $Q_1 \Leftrightarrow Q_2$. Очевидно, что если $Q_1 \Rightarrow Q_2$, а $Q_2 \Rightarrow Q_1$, то $Q_1 \Leftrightarrow Q_2$. Тогда $T(Q_1) = T(Q_2)$, т.е. множества истинности равносильных предикатов также совпадают. Например, пусть высказывательные формы заданы на множестве действительных чисел $\mathbb{R}$.

$\frac{x^2-5x+6}{x-2}=0$ и $x^2-5x+6=0$ не являются равносильными.

$\frac{3x+8}{x^2+1}=0$ и $3x+8=0$ являются равносильными.

$\ln(x-1)+\ln(x+1)=2$ и $\ln(x^2-1)=2$ не являются равносильными.

$\sqrt{x+3}=x-1$ и $x+3=(x-1)^2$ не являются равносильными.

$\frac{4-8x}{2+x}\geq 0$ и $(4-8x)(2+x)\geq 0$ не являются равносильными.

$\frac{4-8x}{2+x}>0$ и $(4-8x)(2+x)>0$ являются равносильными.

$x^2-x^4\geq 0$ и $1-x^2\geq 0$ не являются равносильными.

В математике нарушение цепочки тождественных преобразований при решении уравнений или неравенств влечет за собой потерю имеющихся или приобретение посторонних корней, т.е. изменение множества истинности исследуемого предиката.

Можно доказать, что *отношение равносильности* высказывательных форм обладает известными свойствами, а именно, оно *рефлективно* и *симметрично*. В том случае, когда одинаковые переменные в каждой из исследуемых форм принимают значения из одного множества, отношение равносильности будет обладать также и свойством *транзитивности*.

Тогда назовем *равносильным преобразованием* высказывательной формы Q_1 ее замену на равносильную форму Q_2. Две равносильные высказывательные формы с одинаковым набором переменных, для которых установлен одинаковый порядок, определяют один и тот же *предикат*.

Эти свойства предиката используются при решении уравнений и неравенств, которые тоже являются некоторыми высказывательными формами. Так, решение любого уравнения или неравенства предусматривает установление множества его истинности, т.е. множества истинности соответствующего ему предиката. В процессе поиска множества истинности производят замену одного предиката другим, равносильным данному, с целью упрощения имеющихся высказывательных форм.

Например, $2x-13+x^2-(6x^2-4x+5-5x^2)=0 \Leftrightarrow 6x=18 \Leftrightarrow x=3$, т.е. множество истинности каждого из этих уравнений состоит из одного числа 3.

Рассмотрим примеры, для которых областью определения является множество действительных чисел: $D=\mathbb{R}$.

Для двух высказывательных форм — уравнений $(x-2)(x-3)=0$ (Q_1) и $x-3=0$ (Q_2) — из $x-3=0$ следует, что $(x-2)(x-3)=0$, т.е. верна запись $Q_2 \Rightarrow Q_1$. Однако из $(x-2)(x-3)=0$ не следует $x-3=0$. Например, $x=2$ является корнем первого уравнения, но не второго.

Из уравнения $(x - 5)(x - 2) = 0$ следует неравенство $x \geq 0$, так как корни уравнения — числа 2 и 5 — удовлетворяют также и неравенству.

Тождественно-истинное высказывание $x^2 + 5 > 0$ может следовать из любой высказывательной формы Q, имеющей непустое множество истинности $T(Q) \subset \mathbb{R}$, т.е. форма $Q \to (x^2 + 5 > 0)$ истинна при любых значениях x.

Отношения *следования* и *равносильности* для высказывательных форм, вообще говоря, зависят от того множества, на котором оно рассматривается. Например, высказывательная форма $x \geq 9$ следует из неравенства $8 \leq x < 12$, если $D = \{2, 0, 4, 5, 7, 9, 10, 11, 13\}$, но не следует, если $D(Q) = \mathbb{N}$. Действительно, при $D = \{2, 0, 4, 5, 7, 9, 10, 11\}$ $T(Q_1) = \{9, 10, 11\}$, а $T(Q_2) = \{9, 10, 11, 13\}$ и выполняется $T(Q_1) \subset T(Q_2)$, т.е. форма $Q_1 \to Q_2$ истинна. Во втором случае $(D(Q) = \mathbb{N})$, $T(Q_1) = \{8, 9, 10, 11\}$, а $T(Q_2) = \{9, 10, 11, 12, 13, 14 \ldots\}$, но отношение $T(Q_1) \subset T(Q_2)$ не выполняется, поскольку $8 \notin T(Q_2)$.

Кванторы. Для количественных характеристик обычно используют понятия «все», «некоторые», «существуют» и др., которые называют **кванторами** (от лат. *quantum* — сколько). Мы часто пользовались символами $\forall$ и $\exists$, заменяющими слова «любой» и «существует». Покажем действие этих кванторов в высказывательных формах. Часть формулы, на которую распространяется действие квантора, называется **областью действия** этого квантора. Вхождение переменной в формулу может быть **связанным,** если переменная расположена либо непосредственно после знака квантора, либо в области действий квантора, после которого стоит переменная. Все прочие вхождения — свободные. Например, в выражении $\forall x\, P(x)$ переменная x связывает свойство предиката и квантор общности. Грубо говоря, от этой переменной, ее конкретного вида и имени, ничего не зависит, т.е. $\forall x\, P(x)$ и $\forall y\, P(y)$ суть одно и то же. Так, можно произвольно называть индекс суммирования в рядах и переменную интегрирования в определенных интегралах. В частности, в определении множества как совокупности *всех* объектов, удовлетворяющих характеристическому свойству, использовалась запись $G = \{x \mid P(x)\}$. Очевидно, что в предикате со связанной переменной ее так же легко можно заменить на любую другую. При этом множество все равно будет совокупностью тех же элементов, удовлетворяющих свойству P. Переменная, не являющаяся связанной, называется **свободной**, если после подстановки вместо нее имени некоторых конкретных объектов предикат превращается в осмысленное предложение.

Между кванторами $\exists$, $\forall$ и логическими операциями существует тесная связь. Пусть предикат $P(x)$ определен на конечном множестве $D = \{a_1, a_2, \ldots, a_n\}$. Тогда высказывание $\forall x \in D(P(x))$ будет

истинным только в том случае, если истинны одновременно все высказывания $P(a_1)$, $P(a_2)$, ..., $P(a_n)$, т.е. если истинна их конъюнкция: $P(a_1) \wedge P(a_2) \wedge \ldots \wedge P(a_n)$. Аналогично высказывание $\exists x P(x)$ означает, что оно истинно, когда истинно хотя бы одно из высказываний $P(a_1)$, $P(a_2)$, ..., $P(a_n)$. Тогда должна быть истинной дизъюнкция высказываний $P(a_1) \vee P(a_2) \vee \ldots \vee P(a_n)$. Поэтому для конечной области определения выполняются равносильности: $(\forall x \in D)P(x) \Leftrightarrow P(a_1) \wedge P(a_2) \wedge \ldots \wedge P(a_n)$ и $(\exists x \in D)P(x) \Leftrightarrow P(a_1) \vee P(a_2) \vee \ldots \vee P(a_n)$.

Таким образом, кванторы общности и существования являются дополнениями и аналогами соответственно логических операций конъюнкции и дизъюнкции.

Поскольку конъюнкцию можно выразить через отрицание и дизъюнкцию, то, вообще говоря, символ $\exists$ можно было не включать в число основных символов, так как квантор существования $\exists x A(x)$ по сути является сокращенной записью формулы $\overline{\forall x \overline{A}(x)}$, выражающей так называемую двойственность.

Например, запишем с помощью формул логики предикатов следующее утверждение: «Для лечения любого известного компьютерного вируса имеются программы. Существуют новые (неизвестные) компьютерные вирусы, для лечения которых программы еще не разработаны».

Введем обозначения элементарных формул:

$A(x)$ — известен компьютерный вирус x;

$B(x)$ — для лечения вируса x существует программа.

Тогда с помощью логических связок и кванторов получим формулы:

$\overline{B}(x)$ — против вируса x нет программы;

$\forall x(A(x))$ — любой вирус известен;

$\exists x(\overline{A}(x))$ — существуют новые (неизвестные) вирусы;

$\forall x(A(x) \to B(x))$ — если вирус давно известен, то имеется программа для его лечения;

$\exists x(\overline{A}(x) \wedge \overline{B}(x))$ — существуют (появились) новые вирусы, для лечения которых программы еще не разработаны.

Тогда формализованное исходное утверждение примет вид

$$(\forall x(A(x) \to B(x))) \wedge (\exists x(\overline{A}(x) \wedge \overline{B}(x))).$$

В эту формулу входит лишь связанная переменная x, а, например, в формуле $A(x_1, x_2) \to \forall x_1 B(x_1)$ первое включение переменной x_1 свободное, а второе — связанное.

Отношения следования и равносильности между высказывательными формами связаны с тождественно-истинными импликацией и эквиваленцией, следовательно, их можно записать с помощью квантора общности:

$Q_1(x) \Rightarrow Q_2(x)$ тождественно $\forall x(Q_1(x) \to Q_2(x))$;

$Q_1(x) \Leftrightarrow Q_2(x)$ тождественно $\forall x(Q_1(x) \equiv Q_2(x))$.

Например, запись $x^2 - 5x = 0 \Leftrightarrow x(x - 5) = 0$ является не формулой, а истинным высказыванием о равносильности двух формул, представленных в виде уравнений. В то же время справедлива запись $\forall x \in \mathbb{R}\ (x^2 - 5x = 0) \equiv (x(x - 5) = 0)$, выражающая истинное высказывание, которое включает операцию эквиваленции в качестве составляющей.

Поэтому логическое следование можно определить через импликацию, а равносильность через эквиваленцию. Так, для эквиваленции справедливо: «Две высказывательные формы Q_1 и Q_2 истинны или ложны $(Q_1 \Leftrightarrow Q_2)$ одновременно с высказыванием $\forall x(Q_1(x) \equiv Q_2(x))$», что и было ранее введено. Итак, существует различие в употреблении знаков «$\Rightarrow$» и «$\rightarrow$», «$\Leftrightarrow$» и «$\equiv$». Как известно, знаки «$\rightarrow$», «$\leftrightarrow$» обозначают логические операции импликации и равносильности и входят составной частью в формулы. Знаки «$\Rightarrow$» и «$\Leftrightarrow$» обозначают определенные отношения между высказывательными формами, не входя в них в качестве составной части.

Квантификация многоместных высказывательных форм. Пусть $Q(x_1, x_2, ..., x_n)$ — n-местная высказывательная форма. Ее замену на высказывательную форму $\forall x_i\, Q(x_1, x_2, x_i, ..., x_n)$ либо на $\exists x_i\, Q(x_1, x_2, x_i, ..., x_n)$ называют **квантификацией** высказывательной формы $Q(x_1, x_2, x_i, ..., x_n)$ по переменной x_i.

В процессе такой квантификации эта i-я переменная связывается одним из кванторов, а n-местная высказывательная форма превращается в $(n - 1)$-местную. Это аналогично тому, что если функцию $f(x_1, x_2, x_i, ..., x_n)$ проинтегрировать от a до b по переменной x_i, то полученный результат будет функцией от $n - 1$ переменной и не будет зависеть от x_i: $I(x_1, ..., x_{i-1}, x_{i+1}, ..., x_n) = \int_a^b f(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n)dx_i$. Так, интеграл от функции одной $(n = 1)$ переменной является константой и вообще ни от чего не зависит.

Пусть дана двухместная высказывательная форма $x - y < 0$, определенная на $\underbrace{\mathbb{R}}_{x} \times \underbrace{(0, +\infty)}_{y}$. Произведем квантификацию по переменной y («навесим» квантор общности). Получим одноместную высказывательную форму $\forall y(x - y < 0)$ со свободной переменной x. Если для фиксированного $x = x_0$ будет выполнено $\forall y\,(x_0 - y < 0)$, то эта высказывательная форма превращается в истинное высказывание, например, при $x = -2$, а при $x = 3$ — в ложное.

Если в одноместной высказывательной форме связать квантором и вторую переменную x, то можно получить высказывание: либо $\exists x \forall y(x - y < 0)$ — истинное высказывание; $\forall x \forall y(x - y < 0)$ — ложное высказывание.

При «навешивании» кванторов на двухместную высказывательную форму $Q(x, y)$ можно получить одну из восьми комбинаций:

1) $\forall x \forall y(Q(x, y))$ — «для любого x и любого y $Q(x, y)$»;

2) $\forall y \forall x(Q(x, y))$ — «для любого y и любого x $Q(x, y)$»;

3) $\exists x \exists y(Q(x, y))$ — «существует x и существует y, такие, что $Q(x, y)$»;

4) $\exists y \exists x(Q(x, y))$ — «существует y и существует x, такие, что $Q(x, y)$»;

5) $\exists x \forall y(Q(x, y))$ — «существует x, такой, что для любого y $Q(x, y)$»;

6) $\forall x \exists y(Q(x, y))$ — «для всякого x существует y, такой, что $Q(x, y)$»;

7) $\exists y \forall x(Q(x, y))$ — «существует y, такой, что для любого x $Q(x, y)$»;

8) $\forall y \exists x(Q(x, y))$ — «для всякого y существует x, такой, что $Q(x, y)$».

Очевидно, что первое и второе высказывания, а также третье и четвертое тождественны между собой, их значения истинности совпадают. Между остальными полученными высказываниями нельзя установить тождественности: если истинно высказывание 5, то истинным будет и высказывание 8, причем обратное неверно. Аналогично, если истинно высказывание 7, то истинно и высказывание 6, но не наоборот. То есть, если кванторы одноименны (1—4), то их порядок не играет роли и полученные высказывания эквивалентны. Если кванторы разноименные (5—8), то их порядок в полученном высказывании принципиально важен.

Например, для двухместного предиката «Город x находится в стране y» высказывание $\forall y \exists x P(x, y)$ имеет вид 0-местного предиката и читается «В каждой стране y есть некоторый город x». Оно будет истинным, в то время как высказывание $\exists x \forall y P(x, y)$ читается «Существует город x, находящийся во всех странах y» будет ложным.

Пусть даны x, y — две различные переменные, $F(x)$, $\Phi(x)$ и $Q(x, y)$ — любые формулы логики предикатов и M — формула, не содержащая свободных вхождений x. Тогда справедливы равносильности, представленные с учетом двойственности кванторов $\exists$ и $\forall$ в табл. 5.5.

Формулы. Определение формулы лежит в основе так называемой *логики предикатов первого порядка*, в которой разрешается квантифицировать (связывать кванторами) только предметные переменные. Логика предикатов первого порядка включает в себя все формулы логики высказываний, все равносильности исчисления высказываний, а также большинство правил вывода умозаключений из классической логики. Поэтому язык логики предикатов дает возможность анализировать рассуждения естественного языка и науки, делать выводы в различных формальных системах.

Таблица 5.5

Следствия и равносильности логики предикатов

Равносильности для $\exists$	Правила	Равносильности для $\forall$
$\exists x \exists y\, Q(x, y) \Leftrightarrow \exists y \exists x\, Q(x, y)$	Правила перестановки кванторов	$\forall x \forall y\, Q(x, y) \Leftrightarrow \forall y \forall x\, Q(x, y)$
$\exists x \forall y\, Q(x, y) \Rightarrow \forall y \exists x\, Q(x, y)$		
$\overline{\exists} x F(x) \Leftrightarrow \forall x \overline{F}(x)$	Перенос отрицания с квантора на предикат	$\overline{\forall} x F(x) \Leftrightarrow \exists x \overline{F}(x)$
$\overline{\exists} x \overline{F}(x) \Leftrightarrow \forall x F(x)$		$\overline{\forall} x \overline{F}(x) \Leftrightarrow \exists x F(x)$
$\exists x (F(x) \vee \Phi(x)) \Leftrightarrow \exists x F(x) \vee \exists x \Phi(x)$	Правила дистрибутивности кванторов	$\forall x (F(x) \wedge \Phi(x)) \Leftrightarrow \forall x F(x) \wedge \forall x \Phi(x)$
$\exists x (F(x) \wedge \Phi(x)) \Rightarrow \exists x F(x) \wedge \exists x \Phi(x)$		$\forall x (F(x) \vee \Phi(x)) \Rightarrow \forall x F(x) \vee \forall x \Phi(x)$
$\exists x (M \wedge F(x)) \Rightarrow M \wedge \exists x F(x)$		$\forall x (M \wedge F(x)) \Rightarrow M \wedge \forall x F(x)$
$\exists x (M \vee F(x)) \Rightarrow M \vee \exists x F(x)$		$\forall x (M \vee F(x)) \Rightarrow M \vee \forall x F(x)$

Так, высказывательная форма $\overline{p \vee Q(x)} \to q$ является формулой. В то же время высказывательная форма $\overline{\exists x\, P(x) \to Q(x)}$ не будет формулой, поскольку в формуле $\exists x P(x)$ переменная x связана квантором существования, тогда как в формуле $Q(x)$ эта же переменная свободна.

? В чем ценность формальных теорий?
Для описания каких объектов используется логика предикатов?

Вообще говоря, ценность любой формальной теории заключается в возможности описывать с ее помощью произвольные объекты и связи между ними.

Теоремы. К числу основных равносильностей логики предикатов относят:

1. $\overline{\forall x P(x)} \Leftrightarrow \exists x \overline{P(x)}$.
2. $\overline{\exists x P(x)} \Leftrightarrow \forall x \overline{P(x)}$.
3. $\forall x P(x) \Leftrightarrow \overline{\exists x \overline{P(x)}}$.
4. $\exists x P(x) \Leftrightarrow \overline{\forall x \overline{P(x)}}$.
5. $\forall x (P(x) \wedge Q(x)) \Leftrightarrow \forall x P(x) \wedge \forall x Q(x)$.
6. $\exists x (P(x) \vee Q(x)) \Leftrightarrow \exists x P(x) \vee \exists x Q(x)$.

Правила вывода исчисления предикатов:

Правило заключения (*modus ponens*) — правило, аналогичное тому, которое введено в исчислении высказываний.

Правило обобщения ($\forall$*-введения*, *ug-правило*) R_2: $\dfrac{F \to G(x)}{F \to \forall x G(x)}$, где $G(x)$ содержит свободные вхождения x, тогда как F не содержит свободных x.

Правило $\exists$*-введения* (*eg*-правило) R_3: $\dfrac{G(x) \to F}{\exists x G(x) \to F}$.

Нарушение этих требований может привести к ложным выводам, полученным из истинных высказываний. Например, даны предикаты $P(x)$: «натуральное x делится на 15», $Q(x)$: «x делится на 5». Высказывание $P(x) \to Q(x)$ истинно для любых $x \in \mathbb{N}$. Применим для него правило обобщения. Имеем $P(x) \to \forall x Q(x)$: «Если x делится на 15, то каждое число x делится на 5». Получили ложное утверждение, так как правило $\forall$-введения применимо к 0-местным, а не к одноместным, как $P(x)$, предикатам.

Можно доказанные теоремы делать новыми правилами вывода. Так, помимо правил $\forall$- и $\exists$-введения можно ввести правила удаления кванторов.

Пусть выведена или дана формула $\exists x F(x)$, например «Существуют студенты, работающие по специальности». Из предметного множества всех студентов выберем такого, о котором действительно известно, что он работает по специальности, и для него

введем константу a. Поэтому $\exists x F(x) \to F(a)$. Это так называемое правило $\exists$-*удаления*, или *es-правило* (правило выбора).

Правило $\forall$-*удаления* снимает квантор общности, осуществляя переход от $\forall x F(x)$ к произвольным формулам $F(a)$, $F(y)$ и др. с учетом того, что эти переменные свободны от x в Fx. Например, из высказывания «Каждый студент колледжа владеет компьютером» будет следовать, что конкретный студент Максимов тоже владеет компьютером, и произвольно выбранный некоторый студент y владеет компьютером, и всякий студент z тоже владеет компьютером. При этом необходимо помнить, что предметные переменные y и z не должны быть связанными. Правило $\forall$-*удаления* называют *правилом универсальной конкретизации*, или *us-правилом.*

Рассмотрим примеры.

«Все металлы (М) — плавятся (П). Цинк (Ц) — металл. Значит, цинк плавится». Формализация в логике предикатов примет вид: $\forall x(\text{М}(x) \to \text{П}(x)) \wedge \forall x(\text{Ц}(x) \to \text{М}(x)) \vdash \forall x(\text{Ц}(x) \to \text{М}(x))$. Снятие квантора общности: $(\text{М}(x) \to \text{П}(x)) \wedge (\text{Ц}(x) \to \text{П}(x))$; тогда на основании транзитивности импликации имеем $(\text{Ц}(x) \to \to \text{М}(x)), (\text{М}(x) \to \text{П}(x)) \models \text{Ц}(x) \to \text{П}(x)$. Поэтому вывод $\forall x(\text{Ц}(x) \to \to \text{П}(x))$ — обобщение по R_2 — верен.

«Все студенты (С) проходят практику (П). Некоторые студенты работают в фирме (Ф), значит, некоторые работающие в фирме — проходят практику». Формализация примет вид: $\forall x(\text{С}(x) \to \to \text{П}(x)) \wedge \exists x(\text{С}(x) \wedge \Phi(x)) \vdash \exists x(\Phi(x) \wedge \text{П}(x))$. Уберем кванторы по правилам *us* и *es*. Имеем $(\text{С}(a) \to \text{П}(a)) \wedge (\text{С}(a) \wedge \Phi(a)) \Leftrightarrow \Leftrightarrow (\text{С}(a) \to \text{П}(a)) \wedge \text{С}(a) \wedge \Phi(a) \Leftrightarrow \text{П}(a) \wedge \Phi(a)$. Вывод: $\exists x(\text{П}(x) \wedge \wedge \Phi(x))$, т.е. существуют студенты, которые проходят практику в фирме.

Свойства отношения классификации. Рассмотрим непустое множество U. Пусть дана одноместная высказывательная форма Ф с переменной, которая принимает значения из U, проявляя свойство некоторых объектов из него и соответствуя некоторому предикату Q. Множество истинности $T(Q)$ таких объектов является подмножеством U как универсального множества. Пусть дано $U_1 = = \{5, 6, 7, 8, 9, 10, 11, 12, 13, 14 \ldots\}$. Высказывательной форме «$5 \le x < 12$» соответствует подмножество $T_1(Q) = \{5, 6, 7, 8, 9, 10, 11\}$ ($T(Q_1) \subset U_1$). Из множества $U_2 = \{1, 3, 5, 7, 9, 11, 13, 15\}$ та же высказывательная форма выделяет множество истинности $T_2(Q) = = \{5, 7, 9, 11\}$, из $U_3 = \{5, 6, 7, 8, 9, 10, 11\}$ — $T_3(Q) = U_3$, из $U_4 = = \{12, 13, 14, 15\}$ рассматриваемая высказывательная форма выделяет пустое подмножество истинности $T_4(Q) = \varnothing$.

Эта высказывательная форма выражает на множестве U единственное свойство, характерное для рассматриваемого предиката на заданном множестве U, т.е. одноместный предикат Q (в данном случае «$5 \le x < 12$») задает свойство данного множества. Тогда

множество элементов, обладающих таким свойством Q, будем называть **объемом** этого свойства.

Если на множестве U задан предикат, выражающий некоторое свойство P, то множество U можно разбить на два подмножества $T(P)$ и $U \setminus T(P)$. Так, в предыдущем примере $T(P) = \{5, 6, 7, 8, 9, 10, 11\}$ множество истинности предиката P: «$5 \le x < 12$» из множества $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$. $U \setminus T(P) = \{1, 2, 3, 4, 12, 13, 14, 15\}$. Такое разбиение на непересекающиеся подмножества мы называем **классификацией** множества U по основанию P.

Пусть на множестве U задано еще одно свойство Q. Тогда все множество U, разбиваясь на четыре подмножества, представляет новую классификацию. С помощью логических операций такую классификацию записывают в виде: $P(x) \cdot Q(x) \vee P(x) \cdot \overline{Q}(x) \vee \overline{P}(x) \cdot Q(x) \vee \overline{P}(x) \cdot \overline{Q}(x)$. Это аналогично разложению булевой функции по двум переменным (см. подразд. 4.7), с той разницей, что каждое слагаемое должно иметь вид $P(x) \cdot Q(x) \cdot 1$, так как мы разлагаем формулу исчисления предикатов, имеющую вид тавтологии.

Пусть в нашем примере предикат Q выражает новое свойство — «быть нечетным числом». Тогда эти два свойства одновременно классифицируют множество U на подмножества:

$T(P) \cap T(Q) = \{5, 7, 9, 11\}$, выполнено $P(x) \cdot Q(x)$;

$T(P) \cap \overline{T(Q)} = \{6, 8, 10\}$, выполнено $P(x) \cdot \overline{Q(x)}$;

$\overline{T(P)} \cap T(Q) = \{1, 3, 13, 15\}$, выполнено $\overline{P(x)} \cdot Q(x)$;

$\overline{T(P)} \cap \overline{T(Q)} = \{2, 4, 12, 14\}$, выполнено $\overline{P(x)} \cdot \overline{Q(x)}$.

Уточним понятие «отношение» с помощью понятия «предикат». Во всех n-местных предикатах ($n \ge 2$) устанавливаются некоторые *отношения* между переменными.

Рассмотрим примеры.

Высказывательная форма «x — друг y» выделяет из всего множества людей пары x и y, которые связаны между собой отношением дружбы.

Высказывательная форма «$x \perp y$» выделяет из множества пар прямых, например на плоскости, те пары, которые связаны отношением перпендикулярности.

Высказывательная форма «$x^2 + y^2 + z^2 = 16$» выделяет из всего множества троек координат те, которые связаны отношением «точка с координатами $(x; y; z)$ лежит на сфере с центром в начале координат и радиусом $R = 4$».

Отрицания в исчислении предикатов. В разговорной речи для построения отрицания обычно перед сказуемым ставят частицу *не*. Например, *1а* «Студент x учится на факультете программирования» имеет отрицание *1б* «Студент x не учится на факультете программирования».

? Но всегда ли построенное таким образом отрицание истинно?

Утверждения *2а* «Все выпускники колледжей продолжили образование в вузе» и *2б* «Все выпускники колледжей не продолжили образование в вузе» не являются отрицанием друг друга, так как они оба ложны.

Пары утверждений *3а* «Некоторые выпускники колледжей продолжили образование в вузе» и *3б* «Некоторые выпускники колледжей не продолжили образование в вузе» тоже не служат отрицанием друг друга, так как они оба истинны.

Вторая и третья пары утверждений отличаются от первых тем, что содержат кванторные слова «все» и «некоторые». А при построении отрицаний для предложений, содержащих кванторы, прием введения отрицания *не* перед сказуемым не срабатывает. Можно воспользоваться другим, универсальным, приемом построения отрицаний предложений, содержащих кванторы, добавив общее отрицание *неверно, что...* Тогда во втором примере «Неверно, что все выпускники колледжей продолжили образование в вузе» совпадает по смыслу с утверждением «Некоторые выпускники колледжей не продолжили образование в вузе». Таким образом, отрицанием предложения *2а* служит *3б*, а отрицанием *3а* служит *2б*.

Символически общее отрицание принято записывать с помощью либо общей черты, либо отрицания самого квантора. Для отрицания предложения $\forall x(\Phi(x))$ возможны записи $\overline{\forall x}(\Phi(x))$, или $\overline{\forall x\ (\Phi(x))}$, или $(\exists x(\overline{\Phi}(x))$: $\overline{\forall x}(\Phi(x)) \Leftrightarrow \exists x(\overline{\Phi}(x))$; для отрицания $\exists x(\Phi(x))$ аналогично: $\overline{\exists x}(\Phi(x))$, или $\overline{\exists x\ (\Phi(x))}$, или $\exists x(\overline{\Phi}(x))$: $\overline{\exists x}(\Phi(x)) \Leftrightarrow \forall x(\overline{\Phi}(x))$.

Эти равносильности являются обоснованием метода построения отрицаний высказываний, содержащих кванторы. Для построения **отрицания** высказываний, содержащих квантор $\frac{\text{общности } (\forall)}{\text{существования } (\exists)}$, достаточно заменить его на другой квантор $\frac{\text{существования } (\exists)}{\text{общности } (\forall)}$ и взять отрицание выражения, на которое этот квантор был «навешан».

Для многоместных кванторов также применяется это правило: осуществляется последовательный перенос отрицания с кванторного слова на предложение, стоящее за квантором, а сам квантор заменяют на двойственный. Например, для формулы $\forall x\,\forall y\,\exists z\,(R(x, y, z))$ построим отрицание: $\overline{\forall x\,\forall y\,\exists z\,(R(x, y, z))} \Leftrightarrow \exists x(\overline{\forall y\,\exists z\,(R(x, y, z))} \Leftrightarrow \exists x\,\exists y(\overline{\exists z\,(R(x, y, z))} \Leftrightarrow \exists x\,\exists y\,\forall z(\overline{R}(x, y, z))$. В подразд. 4.8 была показана булева двойственность конъюнкции и дизъюнкции. Поэтому для сложных высказываний, состоящих из простых, разделенных операциями конъюнкции и дизъюнкции, **отрицание** строится следующим образом: нужно все кванторы $\exists$ заменить на $\forall$, и

наоборот; все связки *и* ($\wedge$) заменить на *или* ($\vee$), и наоборот; и взять отрицание утверждения.

Рассмотрим примеры.

1. Дадим определение непрерывности функции. Функция $f(x)$ называется непрерывной в точке x_0, если для любого положительного числа ε существует положительное число δ, такое, что для любого x, удаленного от x_0 менее чем на δ, значение $f(x)$ отличается от $f(x_0)$ менее чем на ε. Другими словами, из попадания точки x в δ-окрестность x_0 следует попадание значения функции $f(x)$ в ε-окрестность точки $f(x_0)$.

Переведем это предложение на язык кванторов: $f(x) \in C(x_0)$, если $\forall \varepsilon > 0\ \exists \delta > 0$: $\forall x$: $|x - x_0| < \delta(...)\ |f(x) - f(x_0)| < \varepsilon$. Здесь многоточие означает «выполнено» или «следует», т. е. можно было бы поставить знак $\Rightarrow$. Скобки указывают на необязательность этого, а комментарий дан для лучшего понимания читателем этого выражения.

Построим отрицание: $\exists \varepsilon > 0\ \forall \delta > 0$: $\exists x$: $|x - x_0| < \delta$, $|f(x) - f(x_0)| \geq \varepsilon$, тогда $f(x) \notin C(x_0)$. Приведем развернутую формулировку этого отрицания. Функция $f(x)$ не является непрерывной в точке x_0, если существует $\varepsilon > 0$, такое, что для любого $\delta > 0$ существует x, удаленное от x_0 менее чем на δ, но при этом значение $f(x)$ отличается от $f(x_0)$ не менее чем на ε. Например, функция $\theta(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0 \end{cases}$ в точке $x_0 = 0$ не является непрерывной. Проверим с помощью отрицания: $\exists \varepsilon > 0\ \forall \delta > 0$: $\exists x$: $|x - x_0| < \delta$, $|\theta(x) - \theta(x_0)| \geq \varepsilon$. Действительно, достаточно взять $\varepsilon = 1/2$, $x = \delta/2$, тогда $\forall \delta > 0$, $|\delta/2 - 0| < \delta$ выполняется и $|\theta(\delta/2) - \theta(0)| = |1 - 0| = 1 \geq 0{,}5$. Точно так же можно убедиться, что прямое определение непрерывности не выполняется.

2. Заявленный концерт струнного квартета состоится (K), если все исполнители (M_i) будут здоровы (H — множество здоровых людей) и помещение будет соответствовать противопожарным нормам (событие $\overline{F}$). В каком случае концерт не состоится? Запишем все это в алфавите предикатов: $K \Leftrightarrow (\forall i \in \{1, 2, 3, 4\}\ M_i \in H) \wedge \wedge \overline{F}$. Согласно правилу отрицанием будет $\overline{K} \Leftrightarrow (\exists i \in \{1, 2, 3, 4\}$: $M_i \notin H) \vee F$, т. е. концерт струнного квартета не состоится, если хотя бы один участник заболеет или помещение будет пожароопасным. Поскольку вероятность заболевания как минимум одного человека из четырех значительно выше вероятности заболевания одного человека из одного, ансамблевые концерты отменяются гораздо чаще, чем сольные. Замен в таких камерных коллективах, как квартет, практически не бывает, так как это творческий коллектив единомышленников, а не предприятие или спортивная команда.

3. Заявленный концерт симфонического оркестра состоится (K), если будет здоров дирижер (D), все исполнители (M_i) (80 чел.)

будут здоровы (H — множество здоровых людей) и помещение будет соответствовать противопожарным нормам (событие $\overline{F}$), но если заболеет не более 7 чел., то им есть замена (S_i). В каком случае концерт не состоится?

Запишем сначала простые высказывания: $1: D \in H$ — здоров дирижер; $2: \forall i \in \{1, ..., 80\}\ M_i \in H$ — здоровы оркестранты; $3: \overline{F}$ — помещение соответствует противопожарным нормам; $4: |\{i | M_i \notin H\}| \leq 7$ — заболевших не более семи; $5: \forall j \in \{1, ..., 7\}: \exists S_j \in H$ — каждому заболевшему есть дублер на замену. Тогда K эквивалентно $1 \cdot 2 \cdot 3 \cdot 4 \cdot (4 \to 5)$. Вспоминая введение дизъюнкции через отрицание и импликацию $4 \to 5 = \overline{4} \vee 5$ и далее раскрывая скобки, получим $K = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$. Поэтому $\overline{K} = \overline{1} \vee \overline{2} \vee \overline{3} \vee \overline{4} \vee \overline{5}$. Строя отрицания к каждому простому: $\overline{1}: D \notin H$ — заболеет дирижер; $\overline{2}: \exists i \in \{1, ..., 80\}: M_i \notin H$; $\overline{3}: F$; $\overline{4}: |\{i | M_i \notin H\}| > 7$ — заболеют более семи человек; $\overline{5}: \exists j \in \{1, ..., 7\}: S_j \notin H$ — существуют незаменимые позиции в оркестре или дублер также болен. Поэтому концерт отменят, если заболеет дирижер, или помещение будет пожароопасным, или заболеют более семи оркестрантов, или заболевшему нет замены. Поскольку уважающие себя оркестры имеют группу дублеров, а пожароопасность везде одна и та же, то практически отмены обусловлены болезнью дирижера и вероятность отмены симфонического концерта примерно равна вероятности отмены сольного концерта.

Общезначимость. Как известно, в логике предикатов и высказываний формула считается **общезначимой** тогда и только тогда, когда она выводима. Но если в логике высказываний мы можем процесс доказательства провести за конечное число шагов и заранее знаем конечный результат, то в логике предикатов такой уверенности у нас нет. В исчислении предикатов доказано, что всякая выводимая формула тождественно-истинна, т.е. общезначима, и наоборот. Наиболее важная прикладная задача логики предикатов заключается в поиске эффективной процедуры обоснования равносильностей, в ней применяемых. В настоящее время разработан ряд методов, позволяющих стандартизировать обоснование равносильностей. Но в отличие от логики высказываний логика предикатов *неразрешима*, так как невозможно указать алгоритм установления истинного значения формул для многоместных предикатов ($n \neq 1$). Однако логика одноместных предикатов, наиболее полно соответствующая логике исчисления высказываний, *разрешима*.

Практика последних десятилетий XX в. показывает, что логика высказываний является эффективным инструментом для решения как гуманитарных, так и многих технических задач. Например, при конструировании и оснащении ЭВМ инженерам и программистам вместо того чтобы работать с контактными схемами, достаточно произвести необходимые известные логические операции. Однако логики высказываний недостаточно для

конструирования умозаключений, необходимых для работы вычислительных машин. Для обработки текстов естественного языка на ЭВМ необходимо создание особых логических языков — некоторых формальных систем. Они должны удовлетворять известным требованиям непротиворечивости, полноты и др. для того, чтобы избежать различных логических ошибок и парадоксов. Для того чтобы ЭВМ могли выполнять умозаключения, необходимо с помощью аппарата математической логики создать такую теорию, которая удовлетворяла бы всем перечисленным требованиям. Установлено, что такой формальной системой являются исчисления предикатов первого порядка.

Таким образом, проблема построения умозаключений за два тысячелетия своего существования нисколько не устарела, а даже приобрела особое значение. Так, работа будущих интеллектуальных машин нового поколения будет основана на идее логического программирования, в связи с чем востребованность в построении формальных систем будет возрастать. Поэтому необходимо познакомиться с видами умозаключений, известными из классической логики, а также с правилами выводов, которые проще и доступнее сформулированы в исчислении предикатов.

Операции с кванторами особо важны при работе с нечетко выраженными кванторными словами типа «не менее чем n», «по меньшей мере n», «n и только n» (где $n \in \mathbb{N}$), «хотя бы один из n». Такие кванторные слова называются «численными кванторами» и широко применяются при решении текстовых задач в практических приложениях математики, например в теории вероятностей. Эти выражения могут быть заменены на равнозначные, не содержащие числительных и состоящие только из терминов и знака тождественности. Так, для одноместного предиката предложение «По меньшей мере один объект обладает свойством P» имеет тот же смысл, что и предложение «Существует хотя бы один объект, обладающий свойством P»: $\exists x\, P(x)$.

Для двухместного предиката «Не более одного объекта обладает свойством P» тождественно предложение «Если есть объекты, обладающие свойством P, то они совпадают»: $\forall x\, \forall y\, (P(x) \wedge P(y) \rightarrow \rightarrow x = y)$. Тогда предложение «Один и только один объект обладает свойством P» является конъюнкцией этих предложений.

Предложения «По меньшей мере два объекта обладают свойством P» и «Существуют несовпадающие объекты x и y, которые обладают свойством P» тождественны между собой и символически записываются в виде: $\exists x \exists y (P(x) \wedge P(y) \wedge x \neq y)$.

Предложения «Не более чем два объекта обладают свойством P» и «Каковы бы ни были объекты x, y, z, если все они обладают свойством P, то по меньшей мере два из них совпадают» тождественны между собой и могут быть записаны символически: $\forall x, y, z\, (P(x) \wedge P(y) \wedge P(z)) \rightarrow (x = y) \vee (x = z) \vee (y = z)$.

Конъюнкция двух последних предложений будет читаться «Два и только два объекта обладают свойством P». Например, «Уравнение $ax^2 + bx + c = 0$ имеет не более двух действительных корней» и «Уравнение вида $ax^3 + bx^2 + cx + d = 0$ имеет хотя бы один действительный корень».

Символическая запись с помощью кванторов дает возможность не только компактно записать формулировки математических утверждений, но и формулировать противоречивые суждения для определений и теорем. Использование кванторов существенно упростило методы логического анализа высказываний, дало возможность выводить одни высказывания из других.

5.4. Умозаключения как форма мышления. Дедуктивные умозаключения и их виды

Гибкость ума может заменить красоту.

Стендаль

Предметы, явления действительности находятся во взаимодействии. Отображением предметов в наших мыслях служат понятия об этих предметах и суждения, которые формируются с помощью понятий. Поэтому суждения о понятиях, как и их образы в реальном мире, тоже находятся во взаимодействии. Взяв за основу истинные исходные суждения (посылки), мы делаем выводы (умозаключения) о тех понятиях, которые фигурировали в суждениях. Связь между ними наглядно можно представить в виде схемы (рис. 5.6).

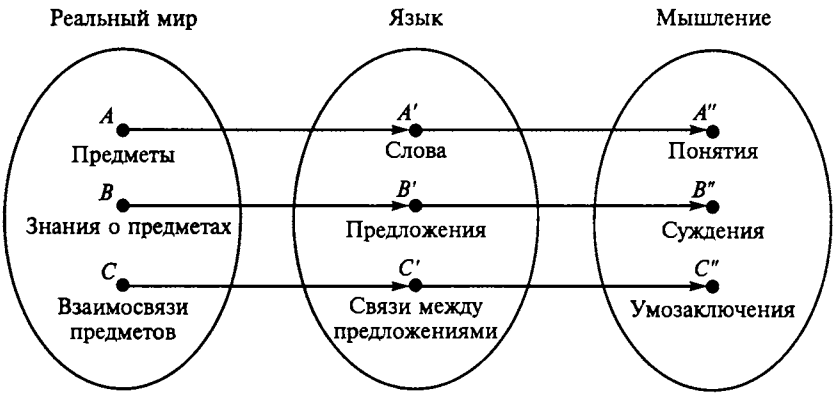


Рис. 5.6. Иллюстрация связи между объектами реального мира, их образами в языке и мышлении

Существует соответствие между объектами действительности, их образами в языке и в мышлении.

Но не всякое сочетание суждений дает умозаключение. Для того чтобы из одного или нескольких исходных суждений (посылок) получились умозаключения, надо знать правила и законы, по которым они образуются.

Приведем три правила образования умозаключений.

Необходимо, чтобы исходные суждения были истинными.

Формирование умозаключений можно проводить только по строго определенным законам, которые необходимо изучить.

Нарушение этих правил ведет к ложным умозаключениям.

Умозаключение — это форма мышления, посредством которой из одного или нескольких суждений на основании правил выводится новое суждение. В состав умозаключения входят посылки, вывод и заключение.

Посылки — это исходные суждения.

Заключение есть новое суждение, полученное из посылки логическим путем.

Вывод — логический переход от посылки к умозаключению.

Окружающий мир имеет многообразные формы связи между своими объектами, поэтому и в мышлении существуют различные виды умозаключений.

По **направлениям логического следования** умозаключения делятся на: дедуктивные — от общих суждений к частным, индуктивные — от частных суждений к общим, по аналогии — от частных суждений к частным.

По **степени достоверности** умозаключения бывают: достоверными (истинными, демонстративными) и вероятностными (правдоподобными, недемонстративными).

Умозаключения являются логическими моделями рассуждений. В зависимости от характера умозаключений выводы при истинных посылках и заключениях могут получаться как *достоверные*, которые обязательно будут иметь место, так и *вероятностные*, которые могут произойти лишь с определенной долей вероятности.

Дедуктивные умозаключения также можно классифицировать в зависимости от *количества* истинных посылок:

непосредственные, $\frac{\text{посылка}}{\text{заключение}}$;

опосредованные, например $\frac{\begin{matrix}\text{большая посылка}\\ \text{меньшая посылка}\end{matrix}}{\text{заключение}}$.

Так, дедуктивное опосредованное умозаключение может иметь вид:

Всякий порок заслуживает наказания

$$\frac{\text{Курение — порок}}{\text{Курение заслуживает наказания}}.$$

Здесь и далее над чертой будем писать посылки, под чертой — заключение.

Для того чтобы заключения были истинными, необходимо знать способы их получения, т.е. логическую связь между посылками и заключением. Незнание законов логики ведет к ложным заключениям. Например, оцените истинность заключений:

$$\frac{\text{Людей много}}{\text{Сократ — человек}} \quad \frac{\text{Сократов много}}{\text{Сократов много}} \quad \text{или} \quad \frac{\text{Платон — человек}}{\text{Ты — не Платон}} \quad \frac{\text{Ты — не Платон}}{\text{Ты — не человек}}.$$

Здесь нет игры слов, просто неправильно построен вывод. Другой пример неправильного построения вывода:

$$\frac{\begin{array}{c}\text{Все математики изучали математическую логику}\\ \text{Все математики имеют высшее образование}\end{array}}{\begin{array}{c}\text{Все, кто имеет высшее образование, изучали}\\ \text{математическую логику}\end{array}}.$$

Правильное заключение имеет вид: «Некоторые, кто имеет высшее образование, изучали математическую логику».

Правильные дедуктивные умозаключения образуются через отношение логического следования между посылкой и заключением. Истинные посылки, если соблюдены все необходимые правила выводов (т.е. импликация истинна), всегда приводят к истинному заключению.

Поэтому дедуктивные умозаключения — самый строгий вид умозаключений, который при соблюдении всех правил всегда дает *достоверный* результат. Дедуктивные рассуждения являются основным видом рассуждений, применяемых в математике. «Высшим долгом физиков является поиск таких элементарных законов, из которых путем чистой дедукции можно получить картину мира», — писал Альберт Эйнштейн о физике, которая для доказательств законов природы использует математический язык.

Рассмотрим отдельные виды дедуктивных умозаключений.

5.4.1. Непосредственные умозаключения по логическому квадрату

Рассмотрим обобщенную классификацию суждений по количеству и качеству, в которой приняты обозначения:

A — общеутвердительные: «Все S есть P», $\forall S \exists P : S \to P$;

E — общеотрицательные: «Ни один S не есть P», $\forall S$ не $\exists P$;

J — частно утвердительные: «Некоторые S есть P»;

O — частно отрицательные: «Некоторые S не есть P».

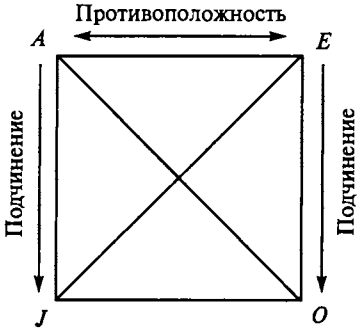


Рис. 5.7. Логический квадрат

«Есть» — связка положительного суждения, «не есть» — связка отрицательного суждения.

Удобнее всего эту классификацию изображать в виде **логического квадрата** (рис. 5.7). Тогда отношения между простыми суждениями можно изобразить на логическом квадрате или в виде графа, в котором две верхние вершины — общие понятия (отношение противоположности), нижние — частные понятия (отношение соподчинения), две правые — отрицательные и две левые — положительные понятия (отношение подчинения), а AO и JE — противоречивые понятия.

Суждение характеризуют две стороны: его форма и его истинность.

Используя свойства отношений между простыми суждениями A, E, J, O, выраженные с помощью логического квадрата, можно делать дедуктивные выводы, определяя истинность или ложность заключений по истинности или ложности посылки. Преобразование простых суждений имеет большое значение в процессе мышления, но не может осуществляться произвольно. Зная правила выводов для непосредственных умозаключений, мы сможем избежать логических ошибок.

Обобщим правила построения умозаключений по логическому квадрату: из истинности общего суждения следует истинность частного, подчиненного ему суждения ($A_и \to J_и$, $E_и \to O_и$), для противоречивых суждений справедлив закон исключенного третьего ($A_и \to \overline{O}_и$, $E_и \to \overline{J}_и$ или $A_л \to O_и$). Индексами «и» и «л» обозначены истинные и ложные суждения соответственно.

Рассмотрим примеры.

A: «Все пьесы — драматургические произведения» (и), но O: «Некоторые пьесы не являются драматургическими произведениями» (л), $A_и \to \overline{O}_и$, значит, $A_и \to O_л$.

A: «Все местоимения — сказуемые» (л), но O: «Некоторые местоимения не являются сказуемыми» (и), $A_л \to O_и$.

E: «Ни одна окружность не является многоугольником» (и), но J: «Некоторые окружности являются многоугольниками» (л), $E_и \to J_л$.

A: «Все союзы — сказуемые» (л), E: «Ни один союз не является сказуемым (и)», $A_л \to \overline{E}_и$.

Истинно или ложно суждение «Ни один воспитанный человек не совершает аморальных поступков»? Это общеотрицательное суждение (E).

Определим истинность соответствующего общеутвердительного суждения. «Все воспитанные люди совершают аморальные поступ-

Таблица 5.6

Непосредственные умозаключения

Операнды	Превращения	Обращения	Противопоставления предикату
Формула	$\dfrac{\text{все}\ \atop \text{некоторые}\ S\ \dfrac{\text{есть}\ \atop \text{не есть}\ P}{}}{\text{некоторые}\ S\ \dfrac{\text{не есть}\ \atop \text{есть}\ \overline{P}}{}}$	$\dfrac{\text{все}\ \atop \text{некоторые}\ S\ \dfrac{\text{есть}\ \atop \text{не есть}\ P}{}}{?\,P\ \dfrac{\text{есть}\ \atop \text{не есть}\ S}{}}$	$\dfrac{\text{все}\ \atop \text{некоторые}\ S\ \dfrac{\text{есть}\ \atop \text{не есть}\ P}{}}{?\,\overline{P}\ \dfrac{\text{не есть}\ \atop \text{есть}\ S}{}}$ $(S \to P) = (\overline{P} \to \overline{S})$
Правила выводов	S сохраняется P — противоречивое Связка — противоположная $A \to E$, $E \to A$ $J \to O$, $O \to J$ — перенос «не» со связки на предикат, и наоборот	S P Связка сохраняется, если $V_S = V_P$ связка $\dfrac{\text{все}}{\text{некоторые}}$, если $V_S > V_P$	Превращаем Обращаем
Противопоказания	Нет	Не выполняется для O	Не выполняется для J
Примеры			
Общеутвердительные A	Все металлы — проводники **электричества** Ни один металл не является неэлектропроводным	**Все квадраты — ромбы** Некоторые ромбы — квадраты Все квадраты — **равносторонние прямоугольники** Все равносторонние прямоугольники — квадраты	Все директора являются **руководителями предприятия** Ни один не руководитель предприятия не является директором

Окончание табл. 5.6

Операнды	Превращения	Обращения	Противопоставления предикату
Общеутвердительные *A*			**Все львы — хищники** Некоторые нехищники — не львы
Общеотрицательные *E*	Ни один хищник не является пернатым Все хищники являются не пернатыми	Ни один кит не является **парнокопытным** Ни одно парнокопытное не является китом	Ни один ядовитый гриб не **является съедобным** Все несъедобные грибы являются ядовитыми
Частноутвердительные *J*	Некоторые вещества — **проводники** Некоторые вещества не являются проводниками (диэлектриками)	**Некоторые поэты — гениальные** Некоторые гениальные люди — поэты	
Частноотрицательные *O*	Некоторые подлежащие не **являются существительными** Некоторые подлежащие являются не существительными		Некоторые высказывания не **являются истинным суждением** Некоторые ложные суждения являются высказыванием

ки (A)». Это суждение ложное, значит, истинно исходное $E \to \overline{A}$ (отношение противоположности). Соответствующее частноотрицательное суждение O: «Некоторые воспитанные люди не совершают аморальных поступков» — истинное, значит, исходное тоже истинное $E \to O$ (отношение подчинения).

Умение определять истинность суждений является важной и необходимой частью мышления человека. Но иногда суждения принимают такой вид, что трудно определить их истинность. Тогда на помощь приходят знания об отношениях между суждениями и умение выполнять логические операции обращения, превращения и противопоставления предикату. Из одной посылки с помощью определенных правил можно получить новое суждение — заключение. При этом за основу берутся виды суждений по количеству и качеству, описанные с помощью логического квадрата.

Классифицируем непосредственные умозаключения в зависимости от *правил выводов* (табл. 5.6).

Умения правильно выполнить непосредственные умозаключения необходимы для получения истинных заключений в различных мысленных построениях, в процессе аргументации, при использовании некоторых приемов косвенных доказательств и опровержений.

5.4.2. Простые категорические силлогизмы

Один из видов опосредованных дедуктивных умозаключений, в котором из двух категорических суждений выводится третье категорическое суждение, термины которого связаны определенным отношением с термином, общим для обеих посылок, называется **простым категорическим силлогизмом** (от греч. συλλογησμος — выведение следствия).

В состав категорического силлогизма входят большая посылка (БП), меньшая посылка (МП), заключение. Общая формула имеет вид:

$$\frac{\text{Все } M \text{ суть } P \text{ (большая посылка)}\quad \text{Все } S \text{ суть } M \text{ (меньшая посылка)}}{\text{Все } S \text{ суть } P \text{ (заключение)}},$$ например

$$\frac{\text{Все люди смертны}\quad \text{Врачи — бесспорно, люди}}{\text{Все врачи (к сожалению) смертны}},$$

где P — предикат вывода, больший термин; S — субъект вывода, меньший термин; M — средний термин, посредник (рис. 5.8).

Классическая логика сконцентрировала свое внимание на тех правилах, по которым делаются достоверные выводы, т.е. на правилах и законах категорических силлогизмов. Она со времен Аристотеля учит строить умозаключения с помощью так называемых

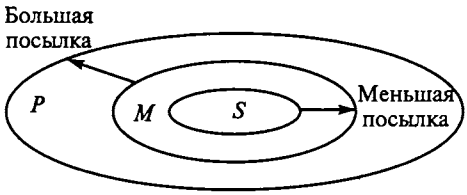


Рис. 5.8. Схема состава категорического силлогизма

фигур категорического силлогизма. Но с развитием математической логики оказалось, что достоверные дедуктивные выводы можно получить не только логическими рассуждениями, но и математическим путем. Приемы математической логики универсальны, более доступны, основаны на строгих законах алгебры высказываний и потому предпочтительнее чисто логических.

5.4.3. Энтимемы

Энтимемы (в переводе с греческого — в уме, в мыслях) — сокращенные категорические силлогизмы, в которых пропущены либо одна из посылок, либо заключение. В повседневной речи энтимемы употребляются достаточно часто, так как очевидные, понятные по контексту посылки либо заключения люди интуитивно опускают.

Например, мы говорим: «Все параллелограммы — четырехугольники, значит, и квадрат — четырехугольник», пропуская при этом МП — «квадрат является параллелограммом». В полной форме такой силлогизм имеет вид

БП: Все параллелограммы (M) являются четырехугольниками (P)
МП: Все квадраты (S) являются параллелограммами (M)
Вывод: Все квадраты являются четырехугольниками.

Из любого силлогизма можно получить три вида энтимем. Пусть дан силлогизм

БП: Все, не знающие правил грамматики (M) допускают ошибки (P)
МП: Я (S) не знаю правил грамматики (M)
Вывод: Я допускаю ошибки.

1. «Раз я не знаю правил грамматики, то допускаю ошибки». Пропущена большая посылка, так как очевидно, что допускают ошибки те, кто не знает правил грамматики.

2. «Так как все, кто не знает правил грамматики, допускают ошибки, то и я допускаю ошибки». Пропущена меньшая посылка, понятная из контекста: я не знаю правил грамматики.

3. «Все, кто не знает правил грамматики, допускают ошибки, а я их не знаю». Пропущен очевидный вывод: значит, я пишу с ошибками.

5.4.4. Умозаключения из сложных суждений

В зависимости от видов посылок в сложных суждениях из них можно получить различные умозаключения, причем в одних случаях они дают достоверные заключения, а в других — вероятностные.

Разделительные силлогизмы содержат хотя бы в одной из посылок разделительное суждение, выраженное через строгую дизъюнкцию, и дают достоверный вывод (рис. 5.9).

Условные силлогизмы содержат условные суждения в посылках или заключении (операция следования): $p \to q$, где p — основание, q — следствие, и дают достоверный вывод (рис. 5.10).

Кроме перечисленных, существуют и другие виды сложных силлогизмов. Для того чтобы проверить, правильно ли сделан вывод в дедуктивных умозаключениях, можно использовать математическую логику.

Схема проверки включает в себя следующие этапы.

1. Проверка справедливости заключения в результате сравнения его с соответствующим правилом.

2. Проверка справедливости заключения с помощью составления таблиц истинности на основании того, что между посылками и выводом дедуктивного умозаключения существуют отношения

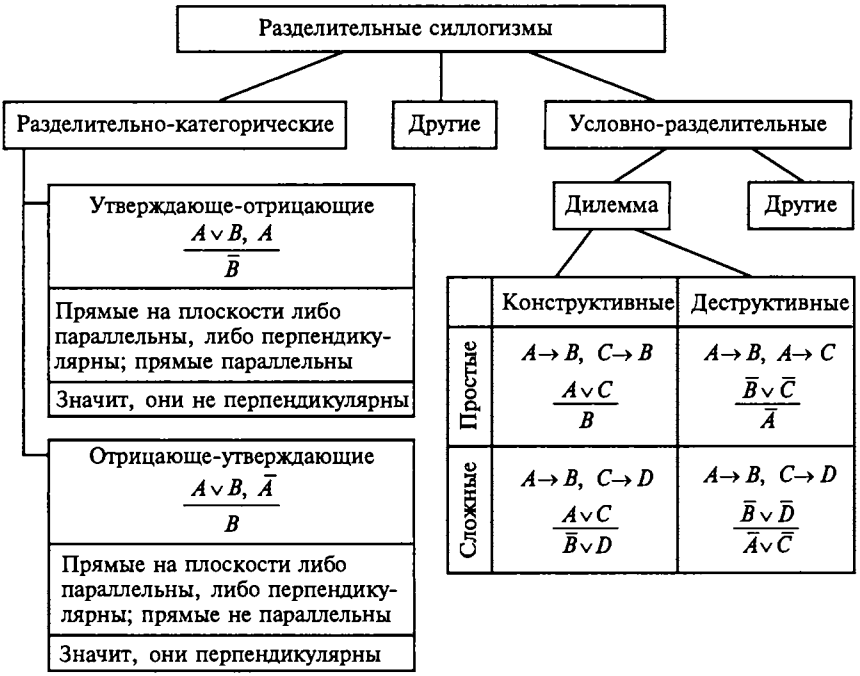
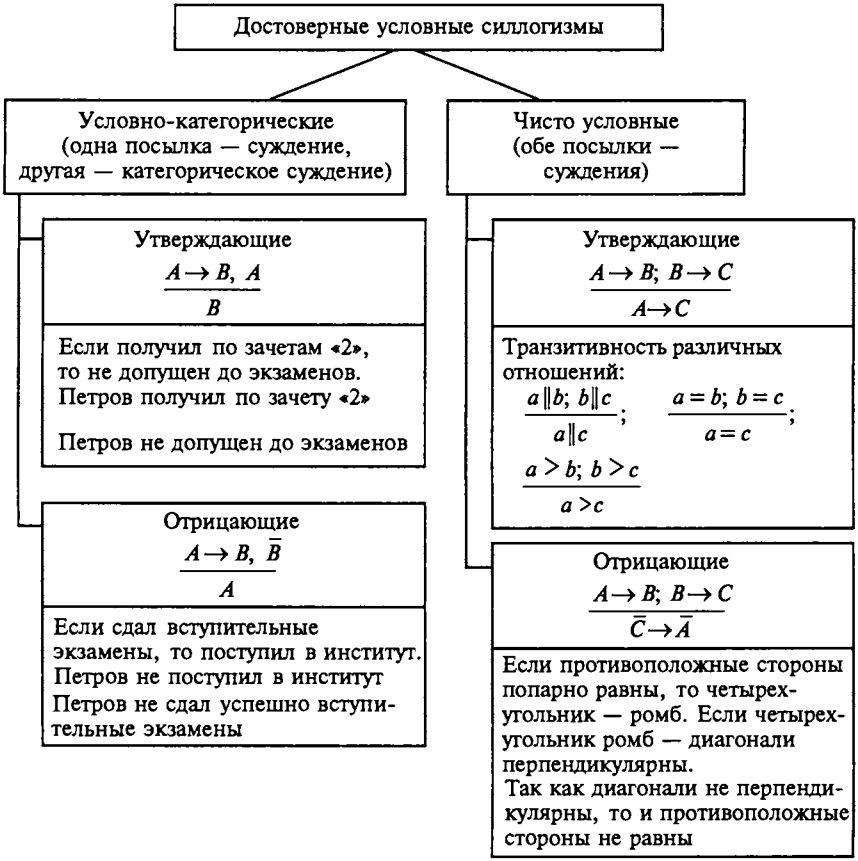


Рис. 5.9. Разделительные силлогизмы



а

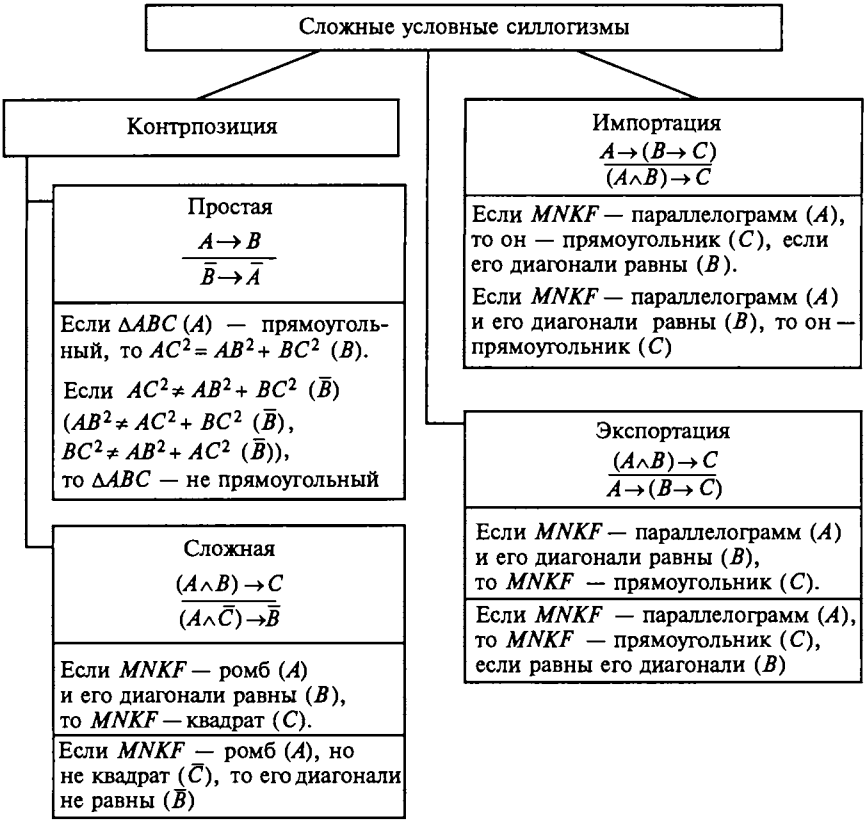
Рис. 5.10. Условные силлогизмы:
а — достоверные; *б* — сложные

логического следования. Заключение не может быть ложным при истинных посылках и правильном ведении вывода.

3. Запись посылок и заключения в виде сложного высказывания, которое с помощью формул алгебры логики упрощается до минимального и затем устанавливается его истинность или ложность.

Итак, умозаключение считается верным, правильным, если из истинных посылок оно не приводит к ложным заключениям.

Задача 24. Необходимо проверить, правильно ли сделан вывод в умозаключении: «Все студенты факультета программирования добросовестны в учебе или талантливы. Если они добросовестны, то систематически готовятся к занятиям. Поэтому, если студенты-программисты не будут готовиться к занятиям, то они должны быть талантливы».



б

Окончание рис. 5.10

Решение. Введем обозначения:
A: студенты-программисты талантливы;
B: студенты-программисты добросовестно относятся к учебе;
C: они систематически готовятся к занятиям.
Тогда данное умозаключение примет вид формулы

$$((A \vee B)(B \to C)) \Rightarrow (\overline{C} \to A) \text{ или } \frac{((A \vee B),\ (B \to C))}{\overline{C} \to A} \text{ или}$$
$$((A \vee B)(B \to C)) \Rightarrow (\overline{C} \to A).$$

Составим таблицу истинности для проверки справедливости этого умозаключения (табл. 5.7).

Строки последнего столбца свидетельствуют о том, что умозаключение истинно при любых значениях пропозиционных переменных A, B, C. Формулу можно было упростить, используя

Таблица истинности

A	B	C	$A \vee B$	$B \to C$	$(A \vee B)(B \to C)$	$\overline{C}$	$\overline{C} \to A$	$(A \vee B)(B \to C) \to (\overline{C} \to A)$
1	1	1	1	1	1	0	1	1
1	1	0	1	0	0	1	1	1
1	0	1	1	1	1	0	1	1
1	0	0	1	1	1	1	1	1
0	1	1	1	1	1	0	1	1
0	1	0	1	0	0	1	0	1
0	0	1	0	1	0	0	1	1
0	0	0	0	1	0	1	0	1

тождество $p \to q = \overline{p} \vee q$. Тогда $((A \vee B)(B \to C)) \to (\overline{C} \to A) = (A \vee B)(\overline{B} \vee C) \to (\overline{\overline{C}} \vee A) = \overline{(A \vee B)(\overline{B} \vee C)} \vee (C \vee A) = \overline{A \vee B} \vee \overline{\overline{B} \vee C} \vee \overline{C} \vee A = \overline{A} \cdot \overline{B} \vee B\overline{C} \vee C \vee A = \overline{A} \cdot \overline{B} \vee B\overline{C} \vee C \vee A = \overline{B} \vee B\overline{C} \vee C = \overline{B} \vee \overline{C} \vee C = 1$. Значит, вывод верен.

5.4.5. Применение аппарата алгебры высказываний для работы с умозаключениями

Проверка правильности умозаключений. Известно, что с помощью равносильностей алгебры логики можно проверять правильность умозаключений. Так, если из истинных посылок получается ложное заключение, то умозаключение и процесс его вывода будут ложными, неправильными. Рассмотрим две похожие задачи.

Задача 25. 1. Про некоторое число известно, что оно делится на 6 или на 15. Но если это число делится на 15, то оно делится на 5. А если число делится на 6, то оно делится на 2. Если же это число делится и на 2, и на 5, то оно делится и на 10. Значит, это число делится на 10.

2. Другой вариант задачи. Известно, что после школы Саша будет поступать или на юридический, или на экономический факультет колледжа. Если Саша закончит юридический факультет, он станет юристом. Если Саша закончит экономический факультет, он станет бухгалтером. А если Саша планирует стать и бухгалтером, и юристом, то сможет работать бухгалтером-юристом.

Правильно ли выполнено умозаключение?

Решение. Введем обозначения для обоих вариантов.

E: $\begin{cases} x \vdots 10, \\ \text{Саша есть бухгалтер-юрист} \end{cases}$

A: $\begin{cases} x \vdots 6, \\ \text{Саша закончит юридический факультет} \end{cases}$

C: $\begin{cases} x \vdots 2, \\ \text{Саша станет юристом} \end{cases}$

B: $\begin{cases} x \vdots 15, \\ \text{Саша закончит экономический факультет} \end{cases}$

D: $\begin{cases} x \vdots 5, \\ \text{Саша станет бухгалтером} \end{cases}$

По условию этих задач составим формулу

$$\frac{(A \vee B)(A \to C)(B \to D)(CD \to E)}{E}.$$

Будем решать задачу методом от противного, т.е. делаем вывод, что истинное заключение $\overline{E}$.

Найдем конъюнкцию условия и нового вывода: $(A \vee B)(A \to C)(B \to D)(CD \to E)\overline{E}$ и упростим полученную формулу:

$$(A \vee B)(A \to C)(B \to D)(CD \to E)\overline{E} = (A \vee B)(\overline{A} \vee C)(\overline{B} \vee D)(\overline{CD} \vee E)\overline{E} = (B\overline{A} \vee AC \vee BC)(\overline{B} \vee D)(\overline{CD} \vee E)\overline{E} = (B\overline{A} \vee AC \vee BC)(\overline{B} \vee D)(\overline{C} \vee \overline{D})\overline{E} = (A\overline{B}C \vee \overline{A}BD \vee ACD \vee BCD)(\overline{C} \vee \overline{D})\overline{E} = (A\overline{B}C\overline{D}\overline{E} \vee \overline{A}\overline{B}CD\overline{E}).$$

Видим, что привести форму к тождественному нулю не удалось, что свидетельствует лишь о том, что методом от противного ее решить нельзя. Нужно брать конъюнкцию прямых высказываний. Видно, что вывод выделяется в виде множителя. Не будем проводить подобные упрощения (читатель это может сделать самостоятельно), чтобы показать, что $(A \vee B)(A \to C)(B \to D)(CD \to E)E = (B\overline{A}D \vee AC\overline{B}\overline{D} \vee ACD \vee BCD)E$. Видно, что это не 1. Это происходит потому, что в условии задачи за обилием посылок, каждая из которых в первой задаче про числа истинна, теряется логическая связка. Если бы в задаче была формулировка «про некоторое число известно, что оно делится на 6 *и* (а в задаче *или*) на 15», то все было бы верно. В случае с числами мы опирались на свойства натуральных чисел с привлечением понятий конъюнкции и дизъюнкции, а в случае карьерного будущего Саши ничего не определено достоверно. Нетрудно подсчитать, что детерминированность в задаче на делимость при истинных $A = B = 1$ (тогда это не отличалось бы от варианта, в котором *или* исправлено на *и*, поскольку $1 \cdot 1 = 1 \vee 1$) приводила к истинному выводу при $A = B = C = D = 1$. Однако в оригинале истинность исходных высказываний не предполагалась.

Энтимемы. С помощью аппарата алгебры высказываний возможно восстановление полного умозаключения из энтимемы.

Например, на судебном заседании адвокат делает вывод: «у нас нет доказательств вины Петрова, значит, он не виноват в преступлении».

Проанализируем вывод, введя обозначения: Д — доказательство вины; П — вина Петрова (преступление). Тогда $\overline{Д} \models \overline{П}$. Ясно, что для восстановления категорического силлогизма нужна еще одна посылка, которая очевидна: если есть преступление, то должны быть доказательства, что оно происходило, т.е. должна быть доказана вина Петрова. В данном и других аналогичных примерах видно, что энтимемы облегчают речь, дают возможность сокращать рассуждения, опуская очевидные высказывания.

Но применение энтимем чревато проблемами непонимания, двусмысленностей, неопределенностей. Ведь не всегда однозначно и легко определяется пропущенное высказывание. Поэтому, используя энтимемы, необходимо уметь восстанавливать полный категорический силлогизм из них. Это можно сделать как с помощью аппарата классической логики, так и с помощью алгебры высказываний.

Задача 26. «Если будет идти снег, трудно будет вести автомобиль. Если трудно будет вести автомобиль, то я опоздаю, если не выеду раньше. Идет снег. Вывод: я должен выехать раньше». Проанализируйте умозаключение и найдите пропущенное высказывание.

Решение. Чтобы восстановить из этой энтимемы категорический силлогизм, введем обозначения: C — идет снег, T — трудно вести автомобиль, O — опоздаю, P — выехать раньше, тогда

$$\frac{(C \to T),(T \to (\overline{P} \to O)),C}{P}.$$

Запишем условия задачи с помощью логических операций $(C \to T)(T \to (\overline{P} \to O))C \models P$. В этой энтимеме не хватает одной посылки.

Используем метод от противного. Докажем, что пропущена одна посылка, но при этом возьмем противоположный данному вывод $\overline{P}$. Тогда формула примет вид $(C \to T)(T \to (\overline{P} \to O))C \models \overline{P}$.

Упростим новую формулу

$(C \to T)(T \to (\overline{P} \to O))C\overline{P} = (\overline{C} \vee T)(\overline{T} \vee (P \vee O))C\overline{P} = C\overline{P}T(\overline{T} \vee (P \vee \vee O)) = C\overline{P}TO$.

Высказывание будет ложным, если будет ложным каждый из множителей, в частности должно быть $\overline{P} = 0$, т.е. $P = 1$ (выехать раньше). Но вывод будет верным, если действительно пойдет снег. Поэтому выехать раньше достаточно для неопоздания.

Для того чтобы из энтимемы сделать правильное полное умозаключение, возьмем отрицание этого вывода: $\overline{TC\overline{P}O} = \overline{T} \vee \overline{C} \vee \vee P \vee \overline{O}$. Любой член этой дизъюнкции может быть пропущенной

посылкой. Соединим его с условием задачи и рассмотрим возможные варианты.

Пусть дополнительная посылка $\overline{T}$, тогда $C \to T$, $T \to (\overline{P} \to O)$, C, $\overline{T}$, но посылки $C \to T$, C, $\overline{T}$ — противоречивы, значит, этот вариант неудачный.

Аналогичный вывод получаем, рассматривая в качестве дополнительной посылки $\overline{C}$.

P — выехал рано, но тогда все остальные посылки излишние, так как в этом случае опоздания не будет.

$\overline{O}$ — в соединении с первыми посылками появление этой посылки оправданно, так как нельзя опаздывать на работу. Заметим, что посылку о том, что «опаздывать на работу нельзя», можно было найти и интуитивно.

Отыскание следствий из заданных посылок. Один из видов энтимем — это нахождение следствий из заданных посылок. В задачах такого рода есть посылки, причем не всегда в полном объеме. Вопросом в этих задачах служит поиск заключения из заданных посылок.

Задача 27. Синоптики объявили прогноз погоды: будет дождь, или ветер, или снег. Если будет дождь с ветром, то будет и снег. Если снега не будет, то будет ветреная погода без дождя. Какая будет погода?

Решение. Введем обозначения: R — дождь, W — ветер, S — снег. Символическая запись условий: $R \vee W \vee S$, $RW \to S$; $\overline{S} \to W\overline{R}$. Поскольку все условия должны быть выполнены, то необходимо найти их конъюнкцию и упростить до конъюнктивно нормальной формы минимизацией формулы:

$$(R \vee W \vee S)(RW \to S)(\overline{S} \to W\overline{R}) = (R \vee W \vee S)(\overline{RW} \vee S)(S \vee W\overline{R}) =$$
$$= (R \vee W \vee S)(\overline{R} \vee \overline{W} \vee S)(S \vee W\overline{R}) = ((R \vee W)(\overline{R} \vee \overline{W}) \vee S)(S \vee$$
$$\vee \overline{W}\overline{R}) = (R\overline{W} \vee W\overline{R} \vee S)(S \vee W\overline{R}) = (S \vee W\overline{R}) = (S \vee W)(S \vee \overline{R}).$$

Видим, что конъюнкция трех условий есть третье ($(S \vee W\overline{R}) = \overline{S} \to W\overline{R}$), т.е. синоптики практически ничего содержательного не сказали. Полученное в конце выражение есть минимальная КНФ. Ее можно заменить совокупностью исходных условий, так как она равносильна конъюнкции условий.

Итак, верными будут умозаключения

$$\frac{(S \vee W)(S \vee \overline{R})}{S \vee W}, \text{ а также } \frac{(S \vee W)(S \vee \overline{R})}{S \vee \overline{R}}.$$

Из предполагаемой истинности исходных условий $R \vee W \vee S$, $RW \to S$; $\overline{S} \to W\overline{R}$ можно сделать вывод $S \vee W$, $S \vee \overline{R}$, что значит «снег или ветер», либо «снег или нет дождя».

Итак, рассуждения на обычном, естественном языке не всегда проводятся с необходимой ясностью и четкостью. Поэтому в обыденной речи так часто бывают двусмысленности в понимании

умозаключений, отсутствие строгости в выводах и рассуждениях. Тогда на помощь приходит логический анализ рассуждений, который удобнее проводить на языке алгебры логики. Установление обоснованности рассуждений — главная цель логического анализа аргументов. Для этого требуется установить логическую структуру умозаключений, что удобнее делать с помощью аппарата математической логики.

5.5. Методы научного познания

Идея вдохновляет, опыт исполняет,
метод царствует.

Н. Векшин

На протяжении более чем двух тысячелетий во всем мире образцом построения науки была математика. В ее основе лежит аксиоматический метод, опирающийся на строгие логические дедуктивные рассуждения. Для выводов в науке используют:

неопределяемые понятия, с помощью которых дают определения всем остальным математическим понятиям;

аксиомы — утверждения, принимаемые без доказательства, играющие роль фундамента, на котором строится здание науки;

теоремы — утверждения, истинность которых необходимо доказать с помощью аксиом и ранее доказанных теорем, образующие строительный материал для построения здания науки.

В зависимости от избранной системы аксиом и неопределяемых понятий выстраивается и сама наука. Так, в настоящее время в геометрии как в учебном предмете в качестве неопределяемых выбраны понятия «точка», «прямая», «плоскость», а также отношения между ними — «принадлежность», «лежать между» и т.д. Знакомые из курса школьной геометрии аксиомы к нам попали в переработанном виде из самого известного учебника геометрии — «Начала», написанного Евклидом (III в. до н. э.) более двух тысяч лет назад.

Доказательства в геометрии носят строго дедуктивный характер и поэтому имеют достоверные выводы.

Аналогичная ситуация и в арифметике, где за основу принята система аксиом Пеано. Здесь в качестве неопределяемых понятий выступают число и множество, а в качестве отношений — «следовать за». На основе этой системы аксиом в арифметике также можно доказывать различные теоремы, т.е. рассуждения строятся от общих к частным и носят дедуктивный характер.

Однако аксиоматический метод используется не во всех науках, и ход рассуждений не всегда имеет дедуктивный характер. Есть много наук, где рассуждения идут от частных к общим и от частных к частным. С помощью аналогий и индуктивных выводов в науке устанавливаются гипотезы, которые затем доказываются дедуктивными методами или опровергаются.

Роль аналогии в научном познании. Велика роль аналогии (в переводе с греческого — сходство) в научном познании. Ведь благодаря аналогии люди получили возможность расширять свои знания. Умозаключение **по аналогии** — это вывод о принадлежности определенного признака исследуемому единичному объекту (предмету, событию, отношению или классу) на основании его сходства в существенных чертах с другими уже известными единичными объектами. Выводы по аналогии имеют огромное значение в науке. Они являются источником многих открытий. Известно, что такие разделы математики, как комбинаторика и теория вероятностей, возникли из анализа азартных игр. Но анализ азартных игр и сейчас имеет принципиально важное значение в науке. Так, процесс тасования карт в ходе игры аналогичен с математической точки зрения технологии перемешивания в химии, а также сходен с некоторыми фундаментальными понятиями термодинамики. В дедуктивных рассуждениях заключение содержит те же предикат и субъект, что и посылки. В индуктивных умозаключениях мы на основании частных знаний о некотором множестве делаем вывод, обобщающий наши знания. Но само множество при этом остается тем же. И лишь умозаключения по аналогии дают возможность людям делать открытия новых свойств неизученных объектов на основании их аналогии с ранее изученными, т.е. происходит переход от изученного множества объектов к новому, исследуемому. Так, обратив внимание на аналогию между принципом действия нервной системы и цифровых вычислительных устройств, Норберт Винер начал свои исследования в области конструирования логических машин.

Процесс умозаключений по аналогии можно осуществить поэтапно:

I этап — операция сравнения объектов с целью установления сходства и (или) различия;

II этап — перенос свойств с оригинала (прототипа) на модель (образ).

Но не всегда аналогия столь легко и убедительно приводит к научным открытиям. Если в основе сравнения лежат не существенные признаки, а иные соображения, то выводы могут оказаться ложными.

Так, пифагорейцы, а затем и Платон провели аналогию между пятью известными правильными выпуклыми многогранниками и строением Вселенной. Их интуитивные представления заключались в том, что четыре стихии — огонь, вода, воздух и земля — состоят из каких-то различных неделимых минимальных элементов — атомов, сохраняющих характерные особенности самой стихии. По их гипотезе, атомы основных элементов должны были иметь формы различных платановых тел: атомы огня — форму тетраэдра (из-за острых углов), земли — форму куба (характерный признак — устойчивость), воздуха — форму октаэдра (самый

воздушный), воды — форму икосаэдра (самый текучий), и объединял все — додекаэдр. Очевидно, такая аналогия оказалась ложной.

Поэтому в естественных, экспериментальных, описательных науках наиболее успешно используются выводы индуктивного характера. Если утверждения, доказанные с применением дедуктивного метода носят достоверный характер, то индуктивные выводы вероятностны по своей природе.

В математике также применяются индуктивные выводы. С их помощью, например, формулируются аксиомы как очевидные утверждения.

Французский математик XVII в. Пьер Ферма, рассматривая и сравнивая числа вида $a_n = 2^{(2^n)} + 1$, пришел к выводу, что для любых натуральных n числа этого вида a_n — простые. Действительно, для первых четырех значений n в результате получаются простые числа: $2^{2^1} + 1 = 5$; $2^{2^2} + 1 = 17$; $2^{2^3} + 1 = 257$; $2^{2^4} + 1 = 6537$. Однако вывод, полученный индуктивно на основании результатов первых четырех значений n, оказался ложным, так как Леонард Эйлер доказал, что число $2^{2^5} + 1$ делится на 641.

В теории чисел часто встречаются задачи на определение количества делителей некоторого числа, заданного формулой.

Например, найдем делители числа k, заданного формулой $k = n^2 + n + 41$. Подставляя вместо n натуральные числа ($n \in \mathbb{N}$), имеем:

$n = 1$, $k_1 = 1^2 + 1 + 41 = 43$ — простое число;
$n = 2$, $k_2 = 2^2 + 2 + 41 = 47$ — простое число;
$n = 3$, $k_3 = 3^2 + 3 + 41 = 53$ — простое число;
$n = 4$, $k_4 = 4^2 + 4 + 41 = 61$ — простое число и т.д.

Возникает гипотеза: для любых натуральных n число $k = n^2 + n + 41$ имеет два делителя 1 и k, т.е. число k является простым.

Справедливость этой гипотезы можно проверить двумя способами: подставляя все натуральные значения n, убедиться в ее справедливости или опровергнуть ее, найдя хотя бы одно значение n, при котором она не выполняется. Очевидно, что подставить в выражение все натуральные n невозможно, так как множество $\mathbb{N}$ бесконечно. А вот найти такое значение $n \in \mathbb{N}$, при котором число k будет составным, несложно. При $n = 41$ имеем $k = 41^2 + 41 + 41 = 41(41 + 2) = 41 \cdot 43$, т.е. число k имеет три множителя — 41, 43, 1 и не является простым. А индуктивный вывод на основании первых сорока значений k оказался ложным.

? Так, может, вывод, полученный путем индуктивных рассуждений, вообще не бывает достоверным?

Дело в том, что в предыдущих рассуждениях вывод делался на основании неполной индукции, поскольку вместо n подставляли отдельные значения $n \in \mathbb{N}$. Если в процессе анализа условия задачи удается проверить все объекты изучаемого множества на соответ-

ствие какому-то свойству, то можно сделать достоверный вывод об обладании этим свойством всеми элементами множества (полная индукция).

Например, проверим, обладает ли множество всех правильных выпуклых многогранников свойством, выражающимся в формуле Эйлера: В + Г – Р = 2, где В — число вершин многогранника; Г — число его граней; Р — число ребер (см. подразд. 2.1). Из геометрии известно, что правильных многогранников в $\mathbb{R}^3$ всего пять: с треугольными гранями — тетраэдр, октаэдр, икосаэдр, с четырехугольными гранями — куб, с пятиугольными — додекаэдр.

А. Проверить гипотезу о том, что формула Эйлера В + Г – Р = 2 справедлива для правильных многогранников, можно с помощью табл. 5.8.

Итак, метод проверки всех элементов множества на соответствие некоторому признаку называется **полной индукцией** (подробнее см. подразд. 5.6) и дает достоверный вывод. Действительно, для всех возможных правильных многогранников справедлива формула Эйлера: В + Г – Р = 2.

Б. Проверим справедливость этой формулы и для других выпуклых многогранников: например, для трехгранной призмы В = 6, Г = 5, Р = 9 формула справедлива. Так же можно проверить все призмы и пирамиды.

? В чем разница между выводами, сделанными в случаях А и Б? Дают ли одинаковые результаты таблиц основания для того, чтобы в случае Б по аналогии со случаем А сделать достоверный вывод?

Между случаями А и Б есть существенное различие. В случае А мы рассмотрели все возможные правильные выпуклые многогранники и сделали вывод на основании *полной индукции*. А в случае Б все возможности неисчерпаемы и вывод не может носить достоверный характер.

Таблица 5.8

Соотношения между числом элементов в правильных многогранниках

Многогранник	Количество			Свойство В + Г – Р = 2
	В — вершин	Р — ребер	Г — граней	
Тетраэдр	4	6	4	4 + 4 – 6 = 2
Октаэдр	6	12	8	6 + 8 – 12 = 2
Куб	8	12	6	8 + 6 – 12 = 2
Додекаэдр	20	30	12	20 + 12 – 30 = 2
Икосаэдр	12	30	20	12 + 20 – 30 = 2

5.6. Индуктивные умозаключения и их виды

Сперва собирать факты и только после этого связывать их мыслью.

Аристотель

Индукция (от лат. *inductio* — наведение) — вид умозаключений, при котором на основании анализа частных суждений о принадлежности признака отдельным элементам множества делается вывод о принадлежности этого признака всему множеству. Обобщим виды индукции (рис. 5.11).

? Как определить истинность умозаключений, полученных с помощью индуктивных рассуждений?

Каков характер выводов в следующих примерах?

• Железо электропроводно. Медь электропроводна. Олово электропроводно. Цинк электропроводен. Железо, медь, олово и цинк — металлы.	• Железо — твердое вещество. Медь — твердое вещество. Олово — твердое вещество. Цинк — твердое вещество. Железо, медь, олово и цинк — металлы.
Все металлы электропроводны.	Все металлы — твердые вещества.

Оказывается, для проверки индуктивных умозаключений необходимо большое число частных случаев, примеров, опытов, подтверждающих данный вывод. Но для опровержения индуктивного умозаключения достаточно одного *единственного* контрпримера, противоречащей инстанции. Так, для подтверждения того, что все жвачные животные имеют рога, надо приводить в качестве примера *все* множество жвачных животных: коз, оленей, коров и т.д. Но для опровержения достаточно в качестве *единственного* примера использовать верблюда.

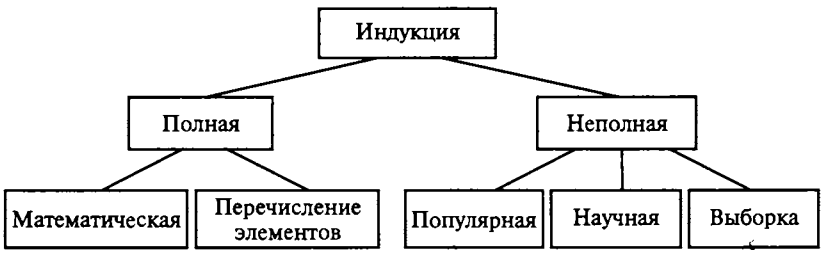


Рис. 5.11. Виды индукции

5.6.1. Виды индукции

Наблюдения в любой области знаний могут привести к определенным индуктивным выводам. Ряд сходных, частных примеров выполнения некоторого свойства дает возможность сформулировать гипотезу о том, что все элементы рассматриваемого множества обладают этим свойством.

? Но справедлива ли гипотеза в каждом конкретном случае? Будут ли наши рассуждения достоверными?

Представим сведения о видах индукции в табл. 5.9.

Таблица 5.9

Виды индукции

Вид	Вывод	Заключение	Условия применения	Примеры
Полная	Достоверный	Перечисление всех элементов	1. Известно конечное число n элементов класса. 2. Проверка всех элементов класса, обладающих данным признаком. 3. Число n элементов счетное	Метод Сократа «Майевтика» — рождение истины из логической цепочки частных вопросов и ответов, единичных примеров обыденной жизни. Индукция через простое перечисление всех элементов конечного счетного множества (перекличка, проверка каталога, перепись жильцов дома и т.д.). S_1 есть P; $\{S_1, S_2, ..., S_n\}$ исчерпывает весь класс P; S_n есть P. Значит, все элементы S есть P
		Математическая	1. Признак имеет место при $n = 1$ $(1 \in P)$. 2. Пусть признак имеет место при $n = k$ $(k \in P)$. 3. Доказать при $n = k + 1$ $(k + 1 \in P)$	Формулы общего члена и суммы арифметической $(a_n = a_1 + (n - 1)d)$ и геометрической $(b_n = b_1 q^{n-1})$ прогрессии. Бином Ньютона $(a + b)^n$. Формулы, определенные на множестве $\mathbb{N}$
Неполная		Научная	По необходимым признакам элементов подмножества делается вывод обо всем множестве	Законы природы в естественных науках физике, химии, биологии. Законы развития общества в общественных науках. Законы логики и философии

Окончание табл. 5.9

Вид	Вывод	Заключение	Условия применения	Примеры
Неполная	Вероятностный	Популярная	Случайные признаки случайных элементов — вывод о принадлежности признака всему классу	Народные суеверия и приметы. Неполная математическая индукция $\{A_1, A_2, A_3, ..., A_n, ...\}$, A_1 имеет признак B, A_2 имеет признак B, A_3 имеет признак B. По-видимому, все A имеют признак B
		Выборка	1. Анализ и отбор элементов планомерного характера. 2. По существенным признакам подмножества вывод о множестве	Математическая статистика (средняя урожайность, проверка на качество в промышленности, медицина). Социологические исследования

Чтобы доказать справедливость операции обобщения в каждом конкретном случае, необходимо иметь информацию о том, что действительно все элементы рассматриваемого множества обладают исследуемым свойством.

? А если множество бесконечно или по некоторым причинам невозможно проверить, все ли элементы множества обладают этим свойством?

В таком случае справедливость гипотезы придется доказывать с помощью неполной индукции, но при этом получать не достоверные, а *вероятностные* выводы. В математике разработан способ, позволяющий сделать достаточно точный *правдоподобный* вывод, не проверяя непосредственно все элементы исследуемого множества. Этот метод называется **методом (полной) математической индукции (ММИ)**.

Как правило, индуктивные выводы осуществляются по следующему алгоритму.

1. *Сравнить* различные элементы некоторого множества.
2. *Подметить* некоторое общее свойство, которым обладают элементы этого множества.
3. *Сформулировать* это свойство для изученных элементов, т.е. сформулировать *гипотезу*.
4. *Обобщить* вывод на более широкий класс элементов, на все множество.

? Как убедиться в том, что наша гипотеза — не иллюзия и действительно имеет место?

Прислушаемся к мнению Д. Пойа, высказанному в работе «Математика и правдоподобные рассуждения».

«В нашей личной жизни мы часто цепляемся за иллюзии, но в науке мы нуждаемся в совершенно ином подходе, в *индуктивном подходе*. Этот подход имеет целью приспособление наших представлений к нашему опыту в такой степени, в какой это возможно. Он требует говорить "*быть может*" и *"возможно*" с тысячей различных оттенков. Он требует многих других вещей, и особенно следующих трех:

Во-первых, мы должны быть готовы пересмотреть любое из наших представлений. Во-вторых, мы должны изменить представления, имеются веские обстоятельства, вынуждающие его изменить. В-третьих, мы не должны изменять представления произвольно, без достаточных оснований.

Первый принцип требует "мужества ума". Вам нужно мужество, чтобы пересмотреть ваши представления. Галилей, бросивший вызов предрассудку своих современников и авторитету Аристотеля, являет собой великий пример мужества ума.

Второй принцип требует "честности ума". Оставаться верным моему предположению, ясно опровергнутому опытом, только потому, что это мое предположение, было бы нечестно.

Третий принцип требует "мудрой сдержанности". Изменить представление без серьезного исследования, например только ради моды, было бы глупо...

Смелость ума, честность ума и мудрая сдержанность — моральные достоинства ученого».

Итак, индуктивный принцип, лежащий в основе не только деятельности ученых, но и обычных разумных людей, заключается в том, что «...предположительное общее утверждение становится более правдоподобным, если оно подтверждается для нового частного случая».

5.6.2. Методы установления причинных связей

Все явления природы и общества находятся в определенных причинных взаимосвязях. Знания о различных явлениях означают, что нам известны причины, породившие это явление, процесс его развития и его последствия — причины следующих явлений.

Всеобщий характер причинно-следственных связей нашел подтверждение в диалектическом единстве случайного и закономерного. В логике еще с XVI в., со времен Фрэнсиса Бэкона и Джона Стюарта Милля, разработаны специальные методы установления причины, породившей то или иное явление. Все методы установления причинных связей носят индуктивный характер. Основной прием, применяемый для установления причинных связей, — сравнение (табл. 5.10).

Таблица 5.10

Методы Милля установления причинно-следственных связей

Наименование	Формула (содержание)	Порядок применения	Примеры
Метод единственного сходства (общего в различном)	$ABCD - a$ $AEFK - a$ $HATD - a$ $\underline{BZXA - a}$ $A \to a$ (по-видимому)	1. Установить все случаи проявления события a. 2. Выявить все причины, обстоятельства появления a A, B, C, $Д$, E, $Г$... 3. Найти в них единственное сходство	1. Определение аллергена у больного аллергией. 2. Построение новых гипотез, версий
Метод единственного различия (различного в сходном)	$ABCD - a$ $\underline{BCD - \bar{a}}$ $A \to a$ (по-видимому)	1. Выявить два случая со сходными обстоятельствами, по результатам. 2. Найти единственное различие в обстоятельствах	1. Прохождение через магнитоконтроль в аэропорту (инородный элемент — причина звукового сигнала). 2. Открытие катализаторов в химии. 3. Эффект удобрения в почве
Объединенный метод сходства и различия	$ABC - a$ $MKB - a$ $MBC - a$ $AC -$ не a $MK -$ не a $\underline{MC -}$ не a По-видимому, $B \to a$	Более высокая вероятность истинного вывода	В криминалистике, в математике
Метод остатков	$ABC - abc$ $A - a$ $\underline{B - b}$ По-видимому, $C - c$	$\dfrac{(A \vee B \vee C)\overline{A} \cdot \overline{B}}{C}$	Открытие планеты Нептун при наблюдении за отклонением Урана. Леверье рассчитал влияние новой планеты на отклонение Урана, а немецкий астроном Галле в 1846 г. открыл Нептун
Метод сопутствующих изменений	$A_1BCD - a_1$ $A_2BCD - a_2$ ..... $\underline{A_n\,BCD - a_n}$ $A - a$	В процессе наблюдений установить, какая из причин изменилась при сохранении других	Любая физическая закономерность

5.6.3. Формальная аксиоматическая теория для арифметики натуральных чисел

В основании арифметики или теории чисел, как и в геометрии, лежит система аксиом, в которой неопределяемыми понятиями являются единица, натуральное число, множество и отношение «следовать за». С их помощью определяются другие понятия теории чисел. Основные аксиомы арифметики сформулировал итальянский математик Д. Пеано (1858—1932). Остальные выводимые свойства, теоремы, доказываются с помощью этих аксиом, а также неопределяемых и определяемых понятий. Следует помнить выводы первой теоремы Гёделя, согласно которой даже в непротиворечивой теории существуют невыводимые формулы.

Формальная аксиоматическая теория *S* для арифметики натуральных чисел.

I. Введем формальные символы в качестве букв алфавита языка теории *S*. Будем использовать алфавит из 12 букв:

- логические символы: $\neg$, $\Rightarrow$, $\forall$, =;
- символы для обозначения предметных переменных языка теории: x, I;
- символы, обозначающие операции в теории *S*: 0, $'$, +, $\cdot$;
- левая и правая скобки: (,).

Других знаков в теории *S* нет, кроме слов метаязыка.

Так, переменные предметного языка системы *S* можно записать с помощью двух букв x и I в виде xI, xII и т.д. Но записи вида x_{IIIIIII} — громоздкие, и для удобства принято в записи переменных использовать арабские цифры 1, 2, ..., 9, ... (т.е. $x_{\text{IIIIIII}} = x_9$) и др. Так как алфавит теории *S* не содержал таких символов, то их взяли из другого языка — метаязыка, поэтому мы называем их метазнаками. Подчеркнем еще раз, что их используют исключительно для удобства, и при желании от них можно отказаться. Рассмотрим семантический смысл знаков.

1. Выбрав в качестве минимального набор логических символов $\neg$ (отрицание), = (равенство), $\Rightarrow$ (логическое следование), $\forall$ (знак общности), мы сможем через них выразить и все остальные логические символы, такие, как $\vee$, $\leftrightarrow$, $\exists$, например $\exists P(x) = \forall x \overline{P}(x)$.

Знак = будем применять при совпадении объектов.

2. Переменные предметного языка (т.е. метаматематические обозначения x_1, x_2, ..., x_n) принимают значения из натурального ряда чисел и числа 0.

3. Функциональная буква $'$ (штрих) обозначает функцию с одним аргументом (унарную операцию) — операцию перехода от числа x к следующему за ним, т.е. $x' \overset{def}{=} x + 1$. Функциональные знаки +, $\cdot$ совпадают по смыслу с привычными бинарными операциями сложения и умножения. Умножение определяется через

сложение. Символ 0 обозначает, что можно дополнить $\mathbb{N}$ одним элементом, называемым нулем, так, что $\forall x \in \mathbb{N}\ x + 0 = x$. Выбор некоего элемента, т.е. приписывание ему какого-то свойства, иногда называют *нулярной* операцией в отличие от унарных и бинарных.

4. Левая и правая скобки предназначены для знаков препинания.

II. Рассмотрим **синтаксис** теории S, т.е. правила образования формул. Существуют два класса слов: термы и формулы. До сих пор мы не оговаривали особо, каким образом можно выделять множество формул, т.е. способы их задания. На этом конкретном примере покажем, что индуктивное задание является основным.

Дадим индуктивное определение **терма**:

1. 0 — терм.
2. $x_1, x_2, \ldots$ — термы.
3. Если t — терм, то (t) — терм.
4. Если t и r — термы, то $t + r$ и $t \cdot r$ — термы.
5. Других термов, кроме определенных согласно 1—4, нет.

В привычном понимании термы — это имена существительные в широком смысле, т.е. имена и именные формы.

Дадим индуктивное определение **формулы**:

1. Если t и r — термы, то $t = r$ — элементарная формула.
2. Если A — формула, то $\overline{A}$ — тоже формула.
3. Если A и B — формулы, то $(A) \to (B)$ — формула.
4. Если A — формула, то $\forall x(A)$ — формула.
5. Других формул, кроме определенных согласно 1—4, нет.

Так, если x_1, x_2, 5 — термы, то $x_1 + x_2 = 5 \cdot x_1$ — формула, включающая в себя термы $x_1 + x_2$ и $5 \cdot x_1$. Если x_1, x_2, 5 — термы по п. 2 определения термов, $x_1 + x_2$ и $5 \cdot x_1$ — термы по п. 4, значит, $x_1 + x_2 = 5 \cdot x_1$ — формула по п.1 определения формул. Формула $(x_1 + x_2 = 5x_1) \to (4 \cdot x_1 = x_2)$ включает в себя две другие, причем скобки можно по договоренности убирать.

Выражение $\forall x(A)$ станет формулой только после того, когда вместо A будет подставлена формула. В теории S недостаточно знаков для выражения всех отношений. Например, отношение «больше» $(x_1 > x_2)$ может быть выражено формулой $\exists x_3(x_1 = x_2 + x_3)$. Тогда недостающие знаки для записи всех возможных отношений можно включать в теорию, но заимствовать их из метаязыка.

Определив синтаксис теории S, включающий алфавит и правила образования термов и формул, необходимо определить и **правила вывода** новых предложений предмета этой теории, т.е. всех *натуральных чисел.*

Так, для любых формул A, B, C теории S справедливы формулы, называемые **логическими аксиомами**:

P_1: $A \Rightarrow (B \Rightarrow A)$;

P_2: $A \Rightarrow (B \Rightarrow C) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$;

P_3: $(A \Rightarrow B) \Rightarrow ((A \Rightarrow \overline{B}) \Rightarrow \overline{A})$;

P_4: $\forall x\, A(x) \Rightarrow A$;

P_5: $\forall x\, (A \Rightarrow B) \Rightarrow (A \Rightarrow \forall x\, B)$.

Выражения $P_1 - P_5$ называют **формальными схемами** аксиом, так как из математической логики нам известна их общезначимость.

Эти схемы выражают те логические законы, которые вместе с правилами вывода и собственно *математическими* аксиомами образуют логику теории S.

В качестве исходных примем **правила вывода**:

1. Из A и $A \Rightarrow B$ выводимо B (*modus ponens*): R_1: $(A(A \Rightarrow B)) \Rightarrow B$.

2. Из A выводимо $\forall xA$ (правило обобщения): R_2: $A \Rightarrow \forall xA$.

К собственным аксиомам теории S относятся следующие формулы:

S_1: $x_1 = x_2 \Rightarrow (x_1 = x_3 \Rightarrow x_2 = x_3)$;

S_2: $x_1 = x_2 \Rightarrow x_1' = x_2'$;

S_3: $x_1' \neq 0$;

S_4: $x_1' = x_2' \Rightarrow x_1 = x_2$;

S_5: $x_1 + 0 = x_1$;

S_6: $x_1 + x_2' = (x_1 + x_2)'$;

S_7: $x_1 \cdot 0 = 0$;

S_8: $x_1 \cdot x_2' = x_1 \cdot x_2 + x_1$;

S_9: $A(0) \Rightarrow \forall x(A(x) \Rightarrow A(x') \Rightarrow \forall xA(x)$ (или $A(0) \wedge \forall x(A(x) \rightarrow \rightarrow A(x')) \Rightarrow \forall xA(x)$,

где $A(x)$ — произвольная формула теории S.

Формулу S_9 называют **аксиомой индукции**.

? Какой семантический смысл вкладывается в эти аксиомы?

Аксиомы S_1 и S_2 выражают свойства отношения равенства с учетом отношения «непосредственно следовать».

Аксиомы S_3 и S_4 соответствуют аксиомам Пеано и выражают отношение «непосредственно не следуют ни за каким натуральным числом».

Аксиомы S_5 и S_6 дают возможность индуктивно (рекурсивно) ввести операцию сложения.

Аксиомы S_7 и S_8 дают возможность индуктивно ввести операцию умножения.

В логику теории S принято вводить два важных понятия: определения формального доказательства и вывода. Тогда с помощью правил вывода можно получать любые новые.

Сформулируем **аксиомы Пеано**.

1. Для каждого натурального числа a существует одно и только одно следующее за ним число. Новое число принято обозначать a'.

2. Единица является натуральным числом, причем она не следует ни за каким натуральным числом: $a' \neq 1$ $(\forall a \in \mathbb{N})$.

3. Ни одно натуральное число не следует за двумя различными числами $(a = b) \Rightarrow (a' = b')$ $\forall a' \in \mathbb{N}$, $\forall b' \in \mathbb{N}$.

4. Если множество A содержит единицу и вместе с каждым числом a содержит следующее за ним число a', то A содержит все натуральные числа: $A = \mathbb{N}$. Эта аксиома дает возможность делать достоверные выводы на бесконечном множестве натуральных чисел с помощью математической индукции, поэтому ее называют **аксиомой математической индукции**.

Действительно, если истинно некоторое высказывание $T(1)$, а из истинности $T(a)$ следует истинность $T(a + 1)$, причем $1 \in A$, $a \in A$, $a + 1 = a' \in A$, где A — подмножество натуральных чисел, для которых истинно $T(n)$, то множество A совпадает со всем множеством натуральных чисел $\mathbb{N}$ и $T(n)$ истинно для всех натуральных n.

Вошедшая в употребление во всем мире система аксиом натуральных чисел была предложена итальянским математиком и логиком, профессором Туринского университета Д. Пеано в статье «О понятии числа» в 1891 г.

5.6.4. Метод математической индукции

Метод полной индукции имеет весьма ограниченную область применения в математике, поскольку большинство рассматриваемых множеств бесконечны: множество натуральных чисел, множество простых чисел, множество многогранников и т.д. Проверить справедливость гипотезы напрямую при бесконечном множестве исследуемых объектов невозможно.

В таких случаях применяется метод рассуждений, заменяющий полный перебор всех вариантов, который также дает достоверный вывод. Этот метод носит название **метода математической индукции** (ММИ). С помощью ММИ определяются сложение и умножение натуральных чисел, свойства этих операций, вводятся отношения «больше» и «меньше» и их свойства, доказываются делимость и формула для n-й степени бинома (бином Ньютона). Так, с помощью ММИ можно доказать одно из важнейших свойств натуральных чисел — свойство полной упорядоченности (в любом непустом подмножестве множества натуральных чисел есть наименьшее число).

Смысл ММИ заключается в применении аксиомы Пеано в виде некоторого алгоритма.

1. Утверждение проверяется для некоторого начального элемента, например для $n = 1$.

2. Формулируется гипотеза о том, что утверждение справедливо для некоторого $k \in \mathbb{N}$.

3. Доказывается (устанавливается истинность утверждения), что если из того, что утверждение справедливо для произвольного $k \in \mathbb{N}$ *следует*, что оно справедливо и для $\forall k + 1 \in \mathbb{N}$, то оно справедливо для любого натурального числа: $\forall n \in \mathbb{N}$.

В теории функционально полных формальных систем мы упоминали, что система должна быть минимальной. Поэтому удобнее пользоваться одной аксиомой.

Аксиома натуральных чисел. Любое непустое подмножество натуральных чисел имеет минимальный элемент. Минимальный элемент самого множества $\mathbb{N}$ обозначается 1 и называется «единицей». Операцию «следования за» будем обозначать не штрихом, а привычно: + 1, например 3 = 2 + 1 (аналог 3 = 2′). При такой аксиоматике математическая индукция будет уже не аксиомой, а *теоремой математической индукции.*

Теорема. Если для множества $A \subset \mathbb{N}$ выполнены условия: $1 \in A$ и если из $k \in A$ следует $k + 1 \in A$, то $A = \mathbb{N}$.

Предлагаем читателю самому доказать эту теорему с помощью аксиомы. Легче всего это сделать методом от противного.

Полученный вывод замечателен тем, что в нем речь идет об абстрактных множествах. Фактически слова «утверждение справедливо» означают выполнение некоего свойства, а мы знаем, что задать характеристическое свойство все равно что задать множество. Поэтому-то две приведенные аксиоматики являются эквивалентными, так как из одной системы можно получить другую в качестве теоремы.

Метод математической индукции чаще всего применяется к *натуральным* числам и *счетным* множествам для доказательства формул, неравенств, делимости натуральных чисел и комбинаторных формул.

? Как быть, если по каким-то причинам утверждение не удается доказать подобным методом?

Даже если в результате применения ММИ получается заведомая ложь или упрощение не получается вовсе, то сначала нужно попытаться опровергнуть гипотезу логическим путем. Например, можно привести один факт, разоблачающий ее. Если это не получается, то нужно искать другие логически корректные методы доказательства утверждения, поскольку из ложной или неопределенной посылки может следовать все, что угодно.

Задача 27. Доказать, что $\forall n \in \mathbb{N}$ справедливо равенство:

$$1 + 4 + 7 + \ldots + (3n - 2) = \frac{n(3n-1)}{2}.$$

Решение. Проверим равенство при $n = 1$. Имеем $1 = \frac{1 \cdot (3 \cdot 1 - 1)}{2}$, $1 = 1$ — значит, формула верна.

Гипотеза: пусть формула справедлива для $n = k$.

$$1 + 4 + 7 + \ldots + (3k - 2) = \frac{k(3k-1)}{2}.$$

Доказываем, что формула верна для $n = k + 1$, т.е. имеет место выражение

$$1+4+7+\ldots+(3k-2)+(3(k+1)-2)=\frac{(k+1)(3(k+1)-1)}{2}.$$

Упростив выражение и воспользовавшись гипотезой, имеем

$$\frac{k(3k-1)}{2}+\frac{(3k+1)}{1}=\frac{(k+1)(3k+2)}{2},$$

что после приведения к общему знаменателю примет вид

$$\frac{3k^2+5k+2}{2}=\frac{3k^2+5k+2}{2}.$$

Видим, что формула справедлива для $n=k+1$ при условии ее выполнимости при $n=k$, следовательно, она справедлива для любого натурального числа $n\in\mathbb{N}$.

Схема метода математической индукции может быть записана кратко:

1. $S(1)$ — выполняется.
2. Γ: $S(k)$, $\forall k\in\mathbb{N}$.
3. $S(k)\Rightarrow S(k+1)$.

В математической литературе традиционно ММИ представляется в виде этапов:

1. Проверить справедливость $S(a_1)$.
2. Доказать, что если выполняется $S(k)$, то справедливо $S(k+1)$.

Однако предпочтительнее выделять второй этап в виде гипотезы, чтобы уже от нее переходить непосредственно к доказательству (этап 3).

Конечно, *доказать* формулу можно только в том случае, когда она действительно имеет место, т.е. гипотеза верна.

Задача 28. Проверить, является ли формула $S_n=\frac{n+1}{3n+1}$ суммой ряда $S_n=\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{n(n+1)}$.

Решение. Применим метод математической индукции.

При $n=1$ имеем $\frac{1}{1\cdot 2}=\frac{1+1}{3\cdot 1+1}$ или $\frac{1}{2}=\frac{1}{2}$, формула верна.

Гипотеза: при $n=k$ формула справедлива: $\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{n(n+1)}=\frac{n+1}{3n+1}$.

Убедимся в ее справедливости при $n=k+1$:

$$\left(\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{k(k+1)}\right)+\frac{1}{(k+1)(k+2)}=\frac{k+1+1}{3(k+1)+1}.$$

Используя гипотезу, имеем

$$\frac{k+1}{3k+1}+\frac{1}{(k+1)(k+2)}=\frac{k+2}{3k+4}\quad\text{или}\quad\frac{(k+1)^2(k+2)+(3k+1)}{(3k+1)(k+1)(k+2)}=\frac{k+2}{3k+4}.$$

После упрощения получим

$$\frac{k^3+4k^2+8k+3}{(3k+1)(k+1)(k+2)}=\frac{k+2}{3k+4}.$$

Левая и правая части оказались различными, т.е. гипотеза не подтвердилась. Но из того, что доказать не удалось, не следует неправильность гипотезы. Например, технический арсенал доказывающего может быть недостаточен. Чтобы себя проверить, посмотрим на каком-нибудь небольшом k, например $k = 4$. $S_4=\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\frac{1}{4\cdot 5}=\frac{4}{5}$, но $\frac{4+1}{3\cdot 5+1}=\frac{5}{16}$. А вот это уже достоверно гарантирует, что формула $S_n=\frac{n+1}{3n+1}$ не выполняется для любого $n \in \mathbb{N}$.

При применении метода математической индукции одинаково важны все этапы алгоритма:

первый этап — база индукции дает возможность определить нижнюю границу применения формулы или действия неравенства. В то же время мы проверяем справедливость этой формулы или неравенства для первого элемента множества.

Второй этап — шаг к обобщению, который формулируется в виде гипотезы, справедливой для всех $n = k$. Этот этап называют индуктивным переходом или индуктивной фазой, т.е. от одного частного случая мы перешли к обобщению для $n = k$.

На *третьем этапе* мы должны установить, насколько сильны индуктивные выводы: либо убедиться в их справедливости, либо их опровергнуть для значения, следующего за обобщенным, т.е. для $k' = k + 1$. Это так называемая фаза доказательства.

ММИ широко применяется для доказательства делимости чисел.

Задача 29. Доказать, что $6^{2n} - 1$ кратно 35.

Решение. Проверим справедливость утверждения при $n = 1$. Имеем $6^{2\cdot 1} - 1 = 36 - 1 = 35$, действительно делится на 35. Делимость нацело принято обозначать в виде $(6^{2\cdot 1} - 1) \vdots 35$.

Гипотеза: пусть при $n = k$ справедливо $(6^{2k} - 1) \vdots 35$.

Докажем тогда, что при $n = k + 1$ верно $(6^{2(k+1)} - 1) \vdots 35$. Имеем $6^{2(k+1)} - 1 = 6^{2k+2} - 1 = 6^{2k} \cdot 6^2 - 1 = 36 \cdot 6^{2k} - 1$. Чтобы доказать делимость, необходимо в выражении «увидеть» нашу гипотезу. Для этого к нему одновременно добавим и вычтем число 36. После группировки и вынесения общего множителя имеем $36 \cdot 6^{2k} - 1 = (36 \cdot 6^{2k} - 36) + 36 - 1 = 36(6^{2k} - 1) + 35$.

В полученном выражении каждое слагаемое делится на 35. Действительно, произведение $36 \cdot (6^{2k} - 1)$ кратно 35 по гипотезе. Так как второе слагаемое тоже делится на 35, то и вся сумма кратна 35, что и следовало доказать.

Методом математической индукции можно доказывать неравенства. При этом нужно особенно внимательно относиться к первому этапу алгоритма, так как в условии не всегда оговаривается его область определения.

Задача 30. Доказать справедливость неравенства $2^n > 2n + 1$.

Решение. Проверим неравенство для $n = 1$. Имеем $2 > 3$ ложно, для $n = 2$ имеем $2^2 > 5$ ложно, для $n = 3$ имеем $2^3 > 7$ истинно. Таким образом, неравенство может быть верным для всех $n \geq 3$, $n \in \mathbb{N}$.

Гипотеза: неравенство справедливо $\forall k \geq 3$, т. е. $2^k > 2k + 1$.

Докажем справедливость формулы для $n = k + 1$, т. е. $2^{k+1} > 2(k + 1) + 1$ или $2^{k+1} > 2k + 3$.

Действительно, $2^{k+1} = 2 \cdot 2^k > 2(2k + 1)$ согласно гипотезе. Тогда имеем $2(2k + 1) = (2k + 1) + (2k + 1) = (2k + 1 + 2) + (2k + 1 - 2) = (2k + 3) + (2k - 1) > 2k + 3$. Следовательно, $2^{k+1} > 2k + 3$, поэтому можно утверждать справедливость неравенства $2^n > 2n + 1$ для $\forall n \geq 3$, $n \in \mathbb{N}$, что и требовалось доказать.

? Дает ли метод математической индукции вывод индуктивного характера? Почему метод носит такое название?

Метод математической индукции применяется также и для доказательства теорем. Докажем методом математической индукции знаменитую теорему Эйлера о соотношении вершин, ребер и областей плоского графа (см. гл. 2).

Теорема Эйлера. Пусть G — связный планарный граф, у которого n вершин, m ребер, а его плоская реализация разбивает плоскость на r областей. Тогда справедливо соотношение $n - m + r = 2$.

Доказательство. Проведем индукцию по m.

База $m = 0$. Тогда граф имеет одну вершину $n = 1$ и не разбивает плоскость на области, т. е. $r = 1$. Подставив значения m, n и r в формулу Эйлера, имеем верное равенство $1 - 0 + 1 = 2 \Rightarrow 2 = 2$.

Пусть теорема верна для всех графов, имеющих $m = k$ ребер, и формула $n - k + r = 2$ истинна.

Докажем справедливость для $m' = k + 1$ ребра. Для этого добавим графу еще одно ребро, при этом возможны два варианта.

Если новое ребро инцидентно лишь имеющимся в графе ребрам, то новых вершин не добавилось, значит, $n' = n$. Новое ребро разбило одну из существующих областей на две части, т. е. $r' = r + 1$. Поэтому формула Эйлера имеет вид $n' - m' + r' = 2$ или $n - (k + 1) + (r + 1) = 2$, что эквивалентно $n - k + r = 2$, а это истинно по нашему допущению.

Если новое ребро соединяет существующие вершины с новой, то $n' = n + 1$, но новых областей не добавилось $r' = r$. Подставим эти значения в формулу Эйлера и получим $n + 1 - (k + 1) + r = 2$, т. е. $n - k + r = 2$, что подтверждает справедливость теоремы.

ММИ также дает возможность изучать свойства конечных множеств.

Задача 31. Найти число подмножеств (мощность булеана) конечного множества M.

Решение. Очевидно, что если множество M пусто, то оно содержит единственное подмножество — самого себя, $\varnothing$. Если $|M| = 1$, например $M = \{a_1\}$, то множество M содержит два подмножества: $\varnothing$ и $\{a_1\}$. Если $|M| = 2$, например множество содержит элементы a_1 и a_2, то в него входят четыре подмножества: $\varnothing$, $\{a_1\}$, $\{a_2\}$ и $\{a_1, a_2\}$. Аналогично множество, состоящее из элементов a_1, a_2 и a_3, содержит восемь подмножеств: $\varnothing$, $\{a_1\}$, $\{a_2\}$, $\{a_3\}$, $\{a_1, a_2\}$, $\{a_1, a_3\}$, $\{a_2, a_3\}$, $\{a_1, a_2, a_3\}$.

Пусть число подмножеств n-элементного множества M равно $S(n)$. Тогда имеем $S(0) = 1 = 2^0$, $S(1) = 2 = 2^1$, $S(2) = 4 = 2^2$, $S(3) = 8 = 2^3$... Напрашивается гипотеза о том, что число подмножеств n-элементного множества M равно 2^n.

Проведем доказательство методом математической индукции.

Поскольку база индукции уже проверена, причем начиная с $n = 0$, то сформулируем гипотезу: $\forall k$, $S(k) = 2^k$.

Докажем, что это утверждение справедливо для $n = k + 1$, т.е. в $(k + 1)$-элементном множестве число подмножеств равно 2^{k+1}. Для доказательства рассмотрим множество $M' = \{a_1, a_2, a_3, ..., a_k, a_{k+1}\}$, которое отличается от $M = \{a_1, a_2, a_3, ..., a_k\}$ добавлением к нему элемента a_{k+1}. Среди искомого числа подмножеств уже есть 2^k подмножеств M, и ни в одном из них a_{k+1} не содержится. Кроме того, подмножествами M' будут являться все те же 2^k подмножеств, но с добавленным в каждое элементом a_{k+1}, и все они будут различны. Очевидно, что других подмножеств нет. Тогда общее число подмножеств равно их сумме: $2^k + 2^k = 2^{k+1}$, что соответствует нашей гипотезе. Поэтому на основании принципа математической индукции делаем вывод о том, что действительно для $\forall n \in \mathbb{N}$ число подмножеств n-элементного множества равно 2^n , что и требовалось доказать.

Задача 32. Найти мощность декартова произведения конечных множеств.

Решение. Известно, что число элементов декартова произведения двух множеств $X \times Y$ при $|X| = m$, $|Y| = k$ равно $m \cdot k$, т.е. $|X \times Y| = |X| \cdot |Y| = m \cdot k$.

Пусть даны множества $X = \{x_1, x_2, ..., x_m\}$, $Y = \{y_1, y_2, ..., y_k\}$. Тогда $X \times Y$ состоит из пар вида (x_i, y_j), где $1 \le i \le m$, $1 \le j \le k$.

Доказательство проведем методом математической индукции.

Рассмотрим s множеств $X_1, X_2, ..., X_s$, из которых составлены кортежи длины s: $\alpha = \langle x_1, x_2, ..., x_s \rangle$, где $x_i \in X_i$, причем $1 \le i \le s$. Множество, состоящее из таких кортежей, называется декартовым (прямым) произведением s множеств $X_1, X_2, ..., X_s$ и обозначается $X_1 \times X_2 \times ... \times X_s$. Тогда число элементов такого декартова произведения должно вычисляться по формуле $|X_1 \times X_2 \times ... \times X_s| = |X_1| \cdot |X_2| \cdot ... \cdot |X_s|$.

База индукции уже проверена при $s = 2$. Пусть формула справедлива при $s = k$, т.е. для $\forall k \in \mathbb{N}$ выполняется равенство $|X_1 \times X_2 \times \times ... \times X_k| = |X_1| \cdot |X_2| \cdot ... \cdot |X_k|$.

Проверим, будет ли справедлива при этой гипотезе формула для $s = k + 1$, т.е. проверим, выполняется ли равенство $|X_1 \times X_2 \times ... \times X_k \times X_{k+1}| = |X_1| \cdot |X_2| \cdot ... \cdot |X_k| \cdot |X_{k+1}|$.

Действительно, рассмотрим произвольный элемент $<x_1, x_2, ..., x_k, x_{k+1}>$ множества $X_1 \times X_2 \times ... \times X_k \times X_{k+1}$. Обозначим $\alpha = <x_1, x_2, ..., x_k>$, тогда между множеством наборов вида $<x_1, x_2, ..., x_k, x_{k+1}>$ и множеством пар вида $<\alpha, x_{k+1}>$ установлено взаимно-однозначное соответствие, т.е. мы имеем дело с кортежами одной длины. Обозначим множество всех кортежей α через A, тогда мощность множества декартова произведения $X_1 \times X_2 \times ... \times X_k \times X_{k+1}$ равна мощности множества кортежей вида $<\alpha, x_{k+1}>$, т.е. $|A| \cdot |X_{k+1}|$ (база $s = 2$). Но согласно гипотезе индукции известно, что $|A| = |X_1 \times X_2 \times ... \times X_k| = |X_1| \cdot |X_2| \cdot ... \cdot |X_k|$. Отсюда следует справедливость равенства $|X_1 \times X_2 \times ... \times X_k \times X_{k+1}| = |X_1| \cdot |X_2| \cdot ... \cdot |X_k| \cdot |X_{k+1}|$ для $\forall s \geq 2$.

5.6.5. Статистические обобщения

Статистические обобщения — особый вид индуктивных умозаключений, связанный с анализом массовых событий. Примерами массовых событий могут служить транспортные перевозки пассажиров и грузов, рождаемость и смертность людей, рост народонаселения, распространение заболеваний и эпидемий, динамика преступлений, всевозможные прогнозы: погоды, выборов, продажи товаров, рынка труда и т.д.

Анализ массовых явлений позволяет установить устойчивое распределение интересующих признаков явления. Полученная цифровая информация имеет конкретное практическое значение для организации медицинского обслуживания и торговли, прогнозирования уровня экономической стабильности и борьбы с преступностью и т.д.

Для статистических обобщений необходим анализ индуктивных умозаключений. Но полной индукцией воспользоваться невозможно, поэтому применяется вид неполной индукции — **выборка**.

Итак, **статистические обобщения** — умозаключения неполной индукции, в которых установленная в посылках количественная информация о частоте признака в исследуемом множестве переносится на все множество.

Основными *задачами* статистических обобщений являются изучение количественных отношений причинной связи явлений и описание и измерение закономерностей общественного развития.

Статистические обобщения исследуют массовые объекты или явления, дифференцируя их по группам, подмножествам, определяя средние количественные характеристики различного вида.

При работе со статистическими обобщениями находит подтверждение один из законов философии, переход количественных изменений в качественные.

Обычно используются различные методы статистических наблюдений (перепись, выборочные обследования), методы обработки и анализа данных (балансы, вычисление средних величин и т.д.). Для обработки данных, для определения ошибки выборки, надежности результата применяются методы *математической статистики и теории вероятностей.*

С понятием «вероятность» люди сталкиваются довольно часто: и в повседневной жизни («невероятно, но факт!»), и при изучении математики (теория вероятностей) и логики (вероятностные умозаключения).

? Что же такое «вероятность» и сколько обличий она принимает?

Вероятность — это количественная мера возможности появления события A.

Количественная мера вероятности выражается либо словами, либо численно. Описание и условная шкала вероятностей показаны на рис. 5.12.

Известны четыре определения вероятности, но для анализа массовых явлений используется только статистическое определение, связанное с частотой появления событий.

Пусть m — число благоприятных событию A исходов, n — общее число исходов. Тогда **статистическая вероятность события** A находится по формуле $P(A) = m/n$, причем $0 \le P(A) \le 1$, $m \le n$.

Например, из ста изученных случаев получения осложнений после гриппа в 85 случаях больной переносил заболевание на ногах и не обращался к врачу. Значит, вероятность получить осложнение после гриппа при самолечении (событие A) равна $P(A) = 85/100 = 0{,}85$, или 85 %.

Статистическая вероятность характеризует не результат одного какого-то события, а весь класс, все множество аналогичных явлений.

Вероятность статистических обобщений можно увеличить с помощью анализа и отбора фактов. Так, изучая результаты выборки из некоторой генеральной совокупности, необходимо стремиться,

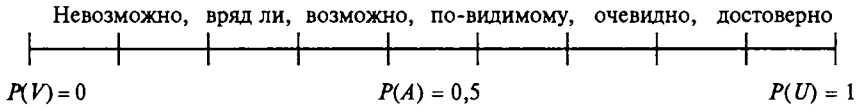


Рис. 5.12. Шкала вероятностей

чтобы *структура* выборки наиболее точно характеризовала структуру соответствующей генеральной совокупности.

Условия повышения вероятности статистических обобщений.

1. Количество исследуемых элементов выборки должно быть достаточным для выводов. Выборка должна быть представительна или *репрезентативна.* Например, минимальное количество деталей в партии, проверяемой на качество, устанавливается с помощью законов теории вероятностей и математической статистики; количество опрошенных свидетелей преступления должно быть достаточным для того, чтобы вывод не оказался поспешным, но их показания не должны быть противоречивы.

2. Элементы выборки должны быть разнообразны, взяты *случайно.* Должен соблюдаться принцип *рандомизации.* Например, психологи, изучая характерные повадки в поведении отдельных преступников, создали психологические обобщенные портреты, которые помогают в конкретных расследованиях; благодаря планомерно и грамотно проведенной выборке для составления прогноза о результатах выборов в президенты, в США и Литве были точно предсказаны результаты выборов.

3. Изучаемый признак должен быть *характерен*, типичен для всех элементов множества изучаемых объектов — генеральной совокупности.

4. Изучаемый признак является *существенным* для всех элементов данного класса.

Статистические обобщения дают возможность совершать обратные умозаключения по типу дедуктивных умозаключений: *от генеральной совокупности к выборке.* Схема рассуждений имеет вид

$$\frac{n\ \%\ \text{генеральной совокупности имеет признак } P}{\text{вероятно, } n\ \%\ \text{выборки будет иметь признак } P}.$$

5.7. Виды аналогии. Моделирование как метод

> Ошибка относится к истине, как сон к пробуждению. Пробуждаясь от ошибки, человек с новой силой обращается к истине.
>
> *И. В. Гёте*

Умозаключение **по аналогии** — это вывод о принадлежности определенного признака исследуемому единичному объекту на основании его сходства в существенных чертах с другими, уже известными единичными объектами.

По степени достоверности различают три вида аналогии (рис. 5.13).

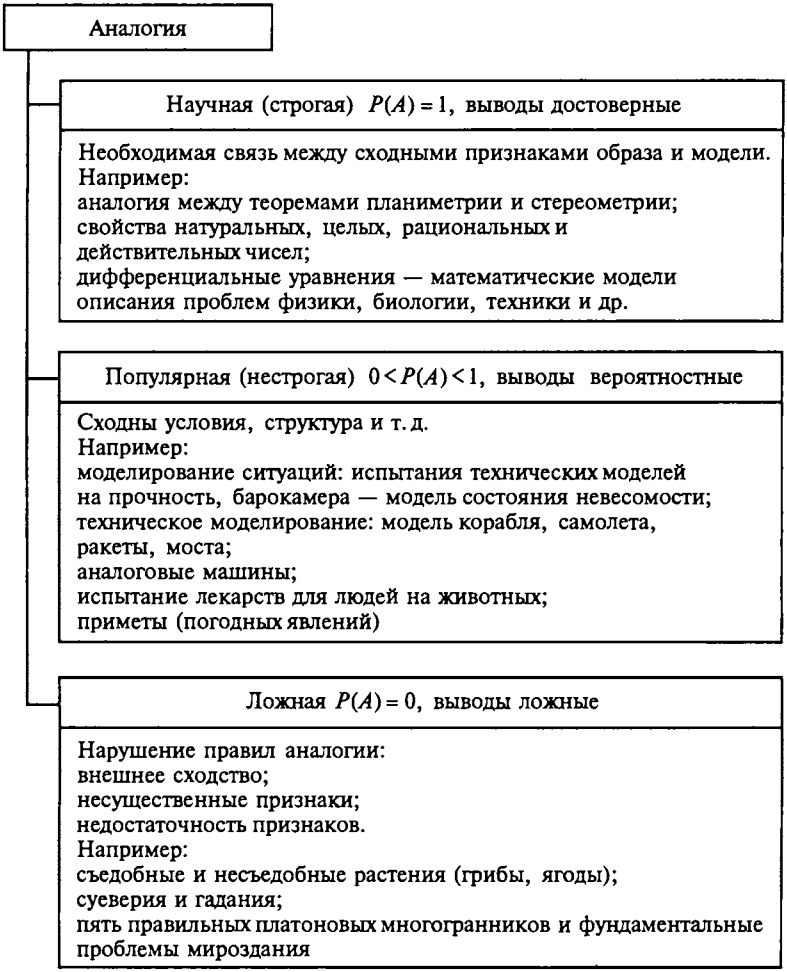


Рис. 5.13. Виды аналогии

Приведем примеры научной аналогии. По аналогии Солнца со свечой М. В. Ломоносов сделал открытие — вывод о том, что свет есть материя. По аналогии между падением яблока и движением небесных тел Ньютон открыл закон всемирного тяготения.

Аналогия необходима не только ученым для открытий новых законов. Умение проводить аналогии между различными, но в чем-то схожими явлениями, предметами и другими объектами помогает учащимся осваивать науки, использовать на практике правила и формулы. Так, одно и то же дифференциальное уравнение описывает разные законы физики, биологии, социологии и т. д. Например, с помощью уравнения показательного роста $y' = ky$,

имеющего решение $y = Ce^{kx}$, можно описать распад радиоактивного вещества, рост народонаселения и т.д.

Аналогия — это общеучебное умение переносить знания с одного предмета на другой в аналогичных заданиях. Например, в физике и математике необходимо решать задачи на движение, знать соотношения между тригонометрическими функциями.

? Почему математику называют и царицей, и служанкой всех наук (какая поэтичная, но противоречивая аналогия)?

Умозаключения по аналогии имеют широкий спектр применения: в физике и различных направлениях техники, в лингвистике и кибернетике, в истории и биологии. При этом достаточно просто решаются многие вопросы путем создания ситуаций, аналогичных реальным. Например, при конструировании самолета в аэродинамической трубе испытывается его точная копия. Проверяется устойчивость и модели, и реального самолета в полете. И если выводы будут неблагоприятными, то конструкторы будут продолжать доводку самолета. При этом промышленность экономит реальные деньги. В технике также давно вошло в практику предварительное изучение условий эксплуатации на модели. И только при получении положительных результатов проект воплощается в действительность: строится мост, плотина, ГЭС и т.д.

Моделирование как метод. Слово «модель» произошло от латинского *modulus* — мера. В науке под **моделью** понимают искусственно созданный объект, аналогичный своему прототипу. Главная задача модели: воспроизвести в упрощенном варианте взаимосвязи и отношения реального исследуемого объекта — прообраза — с действительностью с целью изучения этих отношений тогда, когда непосредственно это сделать сложно или невозможно.

Существуют различные типы моделей: физические, логико-математические, вещественно-математические.

Простейшим примером *физической* модели является макет. Природа *вещественно-математических* моделей полностью отличается от их реально существующих прототипов, но математическое описание модели аналогично прообразу. *Логико-математические* модели конструируются из знаков. Например, модель мыслительной операции может быть изображена чертежом или дедуктивной формулой.

Все три вида моделей могут быть взаимосвязаны. Первоначально выстроенный на бумаге вариант логико-математической модели может при необходимости дополниться вещественно-математической моделью или даже физическим макетом.

В настоящее время существует реальная возможность моделирования многих умственных процессов. Но разработки в области искусственного интеллекта имеют свои пределы, и на данном этапе моделируются лишь отдельные виды умственного труда. Создание

модели дает возможность изучать явление, изолированное от измененной информации, и потому порой в ходе такого анализа быстрее возникают перспективные гипотезы, догадки, ведущие к новым открытиям.

В настоящее время именно компьютер дает возможность моделировать многие процессы, что очень выгодно с экономической точки зрения. Так, дорогостоящие, например химические, реакции заменяют их компьютерной моделью, а их результат удается увидеть даже с сохранением цвета синтезируемого вещества.

Большое значение имеет компьютерное моделирование при изучении многих реальных экономических задач. В настоящее время имитация реальных процессов случайными числами осуществляется с помощью ЭВМ. Процесс отыскания возможных значений случайных чисел называется **моделированием или «разыгрыванием» случайной величины**.

Компьютерное моделирование дает возможность математикам применять метод статистических испытаний, метод Монте-Карло, моделируя процесс получения случайных чисел. Метод Монте-Карло полезен и эффективен для вероятностных расчетов в решении самых различных прикладных задач, которые трудно или невозможно решить аналитически. Метод статистических испытаний универсален, так как не ограничен рамками сложных формул и условий. Поэтому его используют для проверки степени точности построенных аналитических моделей, применяемых в конкретных ситуациях в различных предметных областях: теории массового обслуживания, теории игр, статистической физике, экономике и т.д.

? С какими математическими моделями вы познакомились в процессе изучения дискретной математики?

5.8. Гипотезы

> Проблема в том, чтобы найти проблему. Решить проблему — вопрос чисто технический.
>
> *Н. Векшин*

? Какой путь проходит мысль человека от незнания к знанию? Как рождаются на свет новые теоремы, положения теории и т.д.?

В познавательной цепочке от незнания к знанию решающим звеном является *гипотеза* (в переводе с греческого — основание, предположение).

Гипотеза — это предположение о свойствах, причинах, структуре, связях изучаемых объектов или явлений.

О суждении, являющемся гипотезой, в настоящее время нельзя сказать, истинно оно или ложно. Для ответа на этот вопрос необходимо найти подтверждения, обоснования, доказательства мысли, сформулированной в виде гипотезы.

Понятие «гипотеза» имеет несколько значений.

1. Гипотезой называют предположение о существовании явлений, о причинах их возникновения и характере развития.

2. Гипотезой называют процесс мышления, заключающийся в определении предположения и его доказательств.

3. Гипотезой называют теорию, в которой идея обоснована только до уровня предположения значительной степени вероятности.

По функциям в процессе познания характера объекта исследования гипотезы бывают *описательные* (Что представляет...? Какими свойствами обладает...?) и *объяснительные* (Почему...? Каковы...?).

Напомним, что в гуманитарных областях знаний — политологии, социологии, исторических исследованиях и судопроизводстве — гипотезы принято называть *версиями.*

Любая гипотеза, в том числе и ложная, стимулирует процесс познания, развивает науку. По степени общности гипотезы делятся на общие, частные и единичные. Отталкиваясь от рабочей гипотезы и находя ей подтверждения, можно получить прямой выход на научную теорию.

Гипотезы лежат в основе *гипотетико-дедуктивного* метода научного познания. Этот метод заключается в том, что поскольку в дедуктивных умозаключениях истинность посылок переносится на заключение, то посылка становится гипотезой, а заключение носит вероятностный характер.

Необходимость возникновения гипотез предопределена прогрессом науки. Новые гипотезы стимулируют открытие новых сведений, не имеющих объяснений через известные положения в той или иной предметной области.

Схема возникновения и становления гипотезы как формы развития теории может быть представлена таким образом.

1. Описание фактов:

средствами естественного языка;

статистическими методами (таблицы, ряды динамики и др.);

графически (графики, диаграммы, блок-схемы и др.).

2. Анализ известных теорий — объективных законов науки. Синтез выводов.

3. Догадки научного поиска (не исключая интуицию, заблуждения и т.д.).

4. Выдвижение рабочей гипотезы.

5. Развитие и проверка рабочей гипотезы (анализ следствий, полученных при допущении достоверности рабочей гипотезы, сопоставления следствий с данными опыта, эксперимента).

6. Отказ от отработанной гипотезы при наличии противоречий и выдвижение новой.

7. Развитие и проверка новой рабочей гипотезы.

8. Формулирование выводов — новых законов, теории и т.д.

Сформулируем условия состоятельности гипотезы.

Гипотеза должна быть:

внутренне *непротиворечивой*, а также согласовываться с исходными данными;

принципиально *проверяемой*, иметь очевидные подтверждения, проверяемой в эксперименте;

признанной *состоятельной* только после ее теоретического и экспериментального обоснования;

информативной, т.е. нести в себе *новую информацию*.

Превращение гипотезы в достоверное знание или в научную теорию. Приведем способы подтверждения гипотез.

1. Обнаружение предполагаемого объекта, свойства или явления.

2. Выведение следствий и их экспериментальная проверка.

3. Косвенный путь через опровержение всех ложных гипотез.

В способе подтверждения гипотез через выведение следствий практически невозможно установить все возможные следствия и доказать их истинность, поэтому заключение носит вероятностный характер.

Аналогично осуществляется опровержение гипотез, при этом схема рассуждений имеет вид $\dfrac{(\Gamma \to S)\overline{S}}{\overline{\Gamma}}$ (*modus tollens*), где Γ — гипотеза; S — следствия. Так как для опровержения достаточно убедиться в ложности даже одного из следствий, то вывод по законам дедукции имеет достоверный характер.

? Почему гипотеза не подтвердилась? Была сделана изначально ошибочная, ложная посылка?

Не всегда. Иногда это результат не совсем точных формулировок в следствиях или нарушение чистоты эксперимента во время проверки или гипотезы, или ее следствий.

В случае получения отрицательного результата не стоит сразу отказываться от гипотезы. Следует вновь и вновь перепроверить формулировки следствий, а также эксперименты, их подтверждающие или опровергающие, искать новые следствия.

Вообще, чем больше следствий, подтверждающих гипотезу опытным путем, тем ближе к истине эта гипотеза. Метод познания истины в том и заключается, что из частных гипотез, проверяемых опытом, вырастают общие абстрактные гипотезы.

Каждая из частных гипотез подтверждалась фактами, другими известными законами, опытом человечества. Поэтому общая

гипотеза переросла в научную теорию, имеющую достоверный характер.

Порой две альтернативные гипотезы носят противоречивый характер. Тогда ищутся факты, подтверждающие одну и опровергающие другую гипотезу.

Упражнения

5.1. В пропозиционных функциях сделайте подстановку переменной, чтобы в первом случае получилось истинное высказывание, а во втором ложное:

а) x — президент России; б) x делится на 2 и на y; в) x больше 7, но меньше y; г) неверно, что x старше y, но моложе z; д) x — первый космонавт России; е) x расположен между y и z.

5.2. Связывая переменные кванторами, превратите функции в высказывания:

а) x — автор романа y; б) город x стоит на берегу y; в) река x впадает в y; г) студент x учится на факультете y в учебном заведении z; д) x — число сторон, а y — число диагоналей для многоугольника z; е) словарь перевода с языка x на язык y.

5.3. Пусть $M = \langle N, S_3, P_3 \rangle$, где $S_3(x, y, z) = \text{И} \Leftrightarrow x + y = z$, а $P_3(x, y, z) = \text{И} \Leftrightarrow x \cdot y = z$ (И — истинность). Запишите формулу с одной свободной переменной x, истинную тогда и только тогда, когда

а) $x = 1$; в) $x = 2$; д) x — нечетное;
б) $x = 0$; г) x — четное; е) x — простое число.

5.4. Известно, что $P(x)$: «x — четное число», $K(x)$: «x — нечетное число», $T(x)$: «x — простое число», $Q(x, y)$: «y делится на x». Выразите в понятиях русского языка формулы логики предикатов:

а) $P(12)$;
б) $K(5) \wedge T(5)$;
в) $\forall x\,(Q(x, 2) \to P(x))$;
г) $\exists x\,(P(x) \wedge Q(x, 8))$;
д) $\forall x\,(P(x) \wedge \forall y\,(Q(x, y) \to \to P(y)))$;
е) $\forall x\,(T(x) \to \exists y\,(P(y) \wedge \wedge Q(x, y))$;
ж) $\forall x\ \overline{(P(x) \to Q(x, 8)}$;
з) $\forall x\,\overline{(K(x) \to \forall y\,(T(y) \to \to Q(x, y)))}$;
и) $\exists x\,(P(x) \wedge T(x) \wedge \overline{\exists x}\,((P(x) \wedge T(x)) \wedge \exists y\,(x \neq y) \wedge (P(x) \wedge T(y)))$.

5.5. Даны предикаты $P(X)$: $X < 3$, $Q(Y)$: $Y < 6$. Составьте предикаты:

а) $P(X) \wedge Q(Y)$; в) $P(X) \to Q(Y)$; д) $Q(X) \to P(X)$;
б) $P(X) \vee Q(Y)$; г) $Q(Y) \to P(X)$; е) $Q(X) \vee P(X)$.

5.6. Решите задачи.

1. Андрей, Ваня и Саша собрались в поход. Учитель, хорошо знавший этих ребят, высказал следующие предположения:

а) Андрей пойдет в поход только тогда, когда пойдут Ваня и Саша;

б) Андрей и Саша друзья, а это значит, что они пойдут вместе или же оба останутся дома;

в) чтобы Саша пошел в поход, необходимо, чтобы пошел Ваня.

Когда ребята пошли в поход, оказалось, что учитель немного ошибся: из трех его утверждений истинными оказались только два. Кто из названных ребят пошел в поход?

2. Петя решил поступить в МГУ и послал домой три сообщения:

а) если я сдам математику, то физику я сдам только при условии, что не завалю сочинение;

б) не может быть, чтобы я завалил сочинение и математику;

в) достаточное условие завала по физике — это двойка по сочинению.

После сдачи экзаменов оказалось, что из трех сообщений только одно было ложным. Как Петя сдал экзамены?

3. У разбойника, посаженного в тюрьму, было три сообщника, от каждого из которых он получил по одному письму.

1-е письмо: «Для того чтобы твой побег состоялся, достаточно, чтобы стража была подкуплена только тогда, когда будет закончено рытье подкопа».

2-е письмо: «Если стража будет подкуплена, то достаточное условие твоего побега будет состоять в своевременном рытье подкопа».

3-е письмо: «Если рытье подкопа будет закончено, то необходимо подкупить стражу. Но стражу подкупить не удается. Значит, побег невозможен».

Разбойник знал, что только один из его сообщников говорит правду, а остальные всегда обманывают. Значит, из полученных трех сообщений истинно только одно. Какую информацию получил разбойник?

4. Разбойник, посаженный в тюрьму, послал своим сообщникам две записки:

а) «Для побега достаточно, чтобы стража была подкуплена только тогда, когда вам удастся передать мне веревочную лестницу»;

б) «Для совершения побега достаточно, чтобы стража была подкуплена и мне была передана веревочная лестница».

На следующий день разбойник опять послал своим сообщникам две записки. Первая записка была похожа на предыдущие:

а′) «Если будет подкуплена стража, то для совершения побега достаточно передать мне веревочную лестницу».

Вторая записка была полна пессимизма:

б′) «Невозможно, чтобы стража была подкуплена, мне была передана лестница, и побег состоялся».

На третий день разбойник получил ответ: «Из каждой пары твоих высказываний истинно только одно». Какую информацию получил разбойник?

5. Совершено убийство. Подозреваются Браун, Джон и Смит. Один из них брат убитого, другой — сосед, а третий — случайный знакомый. Каждый из них сделал заявление.

Браун: Если ни я, ни Джон невиновны, то Смит тоже невиновен.

Смит: Чтобы обвинить меня и Брауна, достаточно признать Джона невиновным. Но Джон виновен. Значит, нельзя считать, что ни я, ни Браун невиновны.

Джон: Если меня сочтут виновным, то Смиту удастся оправдаться только тогда, когда оправдается Браун. Но виновен либо Смит, либо Браун. А я невиновен.

Следователь сообщил, что правду сказал только брат убитого, а остальные подозреваемые солгали. Кто убийца? Как фамилия убитого?

6. Коля пригласил свою сестру приехать к нему в гости. После этого он получил от нее три сообщения:

а) «Я поеду в гости, если только со мной поедет папа»;

б) «Чтобы я приехала, необходимо, чтобы меня сопровождала мама».

в) «Либо приедем мы с мамой, либо приедет только папа».

Когда приехали гости, оказалось, что из трех сообщений истинным было только одно. Кто приехал навестить Колю?

5.7. Придумайте примеры, иллюстрирующие действия разделительно-категорических силлогизмов, и проверьте истинность выводов. Дайте оценку характера вывода.

5.8. Сравните соотношения и на основе подмеченного сходства сформулируйте гипотезу об общем члене последовательности с помощью математических символов:

а) $3 + 7 = 10$, $3 + 17 = 20$, $13 + 17 = 30$, ...;

б) гипотеза Гольдтбаха (XVIII в.): $4 = 2 + 2$, $6 = 3 + 3$, $8 = 3 + 5$, $10 = 5 + 5 = 3 + 7$, $12 = 5 + 7$, $14 = 7 + 7 = 3 + 11$, $16 = 3 + 13 = 5 + 11$, ...;

в) $1 = 0 + 1$, $2 + 3 + 4 = 1 + 8$, $5 + 6 + 7 + 8 + 9 = 8 + 27$, $10 + 11 + 12 + 13 + 14 + 15 + 16 = 27 + 64$, ...;

г) 1, $1 + 3$, $1 + 3 + 5$, $1 + 3 + 5 + 7$, ...;

д) 1, $1 + 8$, $1 + 8 + 27$, $1 + 8 + 27 + 64$ + ...;

е) малая теорема Ферма: $(n^2 - n) \vdots 2$, $(n^3 - n) \vdots 3$, $(n^5 - n) \vdots 5$, ...

5.9. Каким методом установлена закономерность в упр. 5.11? Каков характер выводов? Как получить достоверный вывод? Докажите справедливость гипотезы либо опровергните ее.

5.10. Заполните табл. 5.11, установите закономерность, сделайте вывод и сформулируйте гипотезу о формуле общего члена последовательности, а также о виде формулы суммы ряда:

а) $S_n = 1 + 3 + 5 + 7 + \ldots$; б) $S_n = 1 + 2 + 4 + 8 + \ldots$;

в) $S_n = 1^2 - 2^2 + 3^2 - 4^2 + \ldots$; г) $S_n = -1 + 3 - 5 + 7 - 9 + \ldots$;

д) $S_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \ldots$; е) $S_n = \frac{1}{1 \cdot 3} + \frac{1}{3 \cdot 5} + \frac{1}{5 \cdot 7} + \ldots$

Таблица общего члена и суммы к упр. 5.10

	1	2	3	...	n
a_n					
S_n					

5.11. Каким методом установлена закономерность в упр. 5.10? Каков характер выводов? Как получить достоверный вывод? Докажите справедливость гипотезы либо опровергните ее.

5.12. Методом математической индукции докажите тождества:

а) $4 + 9 + 14 + \ldots + (5n - 1) = \dfrac{n(5n+3)}{2}$;

б) $2 + 6 + 10 + \ldots + 2(2n - 1) = 2n^2$;

в) $1^3 + 2^3 + \ldots + n^3 = \dfrac{n^2(n+1)^2}{4}$;

г) $3 + 9 + 27 + \ldots + 3^n = \dfrac{3}{2}(3^n - 1)$;

д) $1 \cdot 2 + 2 \cdot 5 + 3 \cdot 8 + \ldots + n(3n - 1) = n^2(n + 1)$;

е) $1^2 + 2^2 + \ldots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$.

5.13. Докажите, что при любых $n \in \mathbb{N}$ выражение:

а) $7^{2n} - 1$ делится на 24;
б) $13^n + 5$ делится на 6;
в) $6^n + 20n + 24$ делится на 25;
г) $5^n + 2 \cdot 3^n + 5$ делится на 8;
д) $4^n + 15n - 1$ кратно 9;
е) $3^{2n+2} + 8n - 9$ кратно 16.

5.14. Докажите справедливость формулы для суммы ряда:

а) $\dfrac{1}{1\cdot3} + \dfrac{1}{3\cdot5} + \dfrac{1}{5\cdot7} + \ldots + \dfrac{1}{(2n-1)(2n+1)} = \dfrac{n}{2n+1}$;

б) $\dfrac{1}{1\cdot4} + \dfrac{1}{4\cdot7} + \dfrac{1}{7\cdot10} + \ldots + \dfrac{1}{(3n-2)(3n+1)} = \dfrac{n}{3n+1}$;

в) $\dfrac{1}{2\cdot5} + \dfrac{1}{5\cdot8} + \dfrac{1}{8\cdot11} + \ldots + \dfrac{1}{(3n-1)(3n+2)} = \dfrac{n}{6n+4}$;

г) $\dfrac{1}{1\cdot5} + \dfrac{1}{5\cdot9} + \dfrac{1}{9\cdot13} + \ldots + \dfrac{1}{(4n-3)(4n+1)} = \dfrac{n}{4n+1}$;

д) $\dfrac{1^2}{1\cdot3} + \dfrac{2^2}{3\cdot5} + \dfrac{3^2}{5\cdot7} + \ldots + \dfrac{n^2}{(2n-1)(2n+1)} = \dfrac{n(n+1)}{2(2n+1)}$;

е) $\dfrac{1}{2!} + \dfrac{2}{3!} + \ldots + \dfrac{n-1}{n!} = 1 - \dfrac{1}{n!}$, $\forall n \geq 2$;

ж) $n! + \dfrac{(n+1)!}{1!} + \dfrac{(n+2)!}{2!} + \ldots + \dfrac{(n+m)!}{m!} = \dfrac{(n+m+1)!}{m!(n+1)}$.

5.15. Какое математическое предложение записано символами логики предикатов? Постройте отрицание определений:

а) $\forall a, \forall b, \exists c\ ((\log_a b = c) \wedge (a > 0) \wedge (b > 0) \wedge (a \neq 1)) \overset{def}{\Leftrightarrow} ((a^c = b))$;

б) $(\lim\limits_{x \to x_0} f(x) = A) \overset{def}{\Leftrightarrow} (\forall \varepsilon \in R_+)\ \exists \delta \exists x\ ((x \in R) \wedge (0 < |x - x_0| < \delta)) \Rightarrow$

$\Rightarrow |f(x) - A| < \varepsilon$;

в) $\forall \varepsilon > 0\ \exists \delta > 0 : \forall x_1, x_2 \in [a, b] |x_2 - x_1| < \delta \Rightarrow |f(x_2) - f(x_1)| < \varepsilon$;

г) $(a \parallel b) \Leftrightarrow (\exists \alpha((a \subset \alpha) \wedge (b \subset \alpha) \wedge (a \neq b)) \wedge (a \cap b = 0))$;

д) $(A \neq \lim\limits_{x \to x_0} f(x)) \Leftrightarrow (\exists \varepsilon \in R) \forall \delta\ \forall x\ ((x \in R) \wedge (0 < |x - x_0| < \delta)) \Rightarrow$

$\Rightarrow |f(x) - A| \geq \varepsilon$;

е) $\forall x_1 \in \mathbb{N}\ \exists x_2((x_2$ — простое$) \wedge (x_1 < x_2))$;

ж) самостоятельно запишите на языке логики предикатов любое математическое высказывание.

5.16. Представьте в виде блок-схемы этапы прохождения знания от гипотезы до научной теории.

5.17. Рассмотрите некоторую идею в виде ее поэтапного прохождения от гипотезы до научной теории.

5.18. Приведите примеры унарных и бинарных операций, которые можно задать на множествах: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$, B, множестве квадратов на плоскости, множестве груш, множестве голевых моментов в футболе.

5.19. Запишите кванторами, постройте отрицание определений:

а) натуральное число m называется делителем натурального числа n, если $\exists p \in \mathbb{N}$ такое, что $n = m \cdot p$. Говорят также, что n кратно m. Обозначение $n \vdots m$;

б) натуральное число m называется простым, если оно не имеет делителей, отличных от m и 1;

в) натуральные числа m и n называются взаимно простыми, если они не имеют общих (кроме 1) делителей;

г) рассмотрим отображение f: $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$, где $f(m, n) = \max\{p \mid n \vdots p,\ m \vdots p\}$. Число $f(m, n)$ называется наибольшим общим делителем натуральных чисел m и n;

д) рассмотрим отображение g: $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$, где $g(m, n) = \min\{p \mid p \vdots m,\ p \vdots n\}$. Число $g(m, n)$ называется наименьшим общим кратным натуральных чисел m и n;

е) докажите: m и n являются взаимно простыми $\Leftrightarrow f(m, n) = 1$;

ж) докажите тождество $f(m, n) \cdot g(m, n) = m \cdot n$.

Глава 6
ЭЛЕМЕНТЫ ТЕОРИИ И ПРАКТИКИ КОДИРОВАНИЯ

Для передачи, хранения и переработки информации в различных системах необходимы знания теории кодирования.

В этой главе мы познакомимся с историей кодирования от древности до наших дней, с формой представления информации и основными понятиями вероятностной теории информации, а также рассмотрим традиционные вопросы: обработку сообщений как кодирование, помехоустойчивое кодирование, методы эффективного кодирования и средства обеспечения контроля работы автомата.

6.1. История кодирования от древности до наших дней. Защита информации

Язык дан для того, чтобы скрывать свои мысли.

Народная мудрость

Теория кодирования — это раздел теории информации, изучающий способы отождествления сообщений с отображающими их сигналами. Задачей теории кодирования является согласование источника информации с каналом связи. Например, это может быть обеспечение заданных значений помехоустойчивости при заданных характеристиках помех в канале связи, достижение максимальной скорости переработки информации при выполнении некоторых арифметических действий.

Объектом кодирования служит как дискретная, так и непрерывная информация, которая поступает к потребителю через источник информации. Понятие **кодирование** означает преобразование информации в форму, удобную для передачи по определенному каналу связи.

Обратная операция — **декодирование** — заключается в восстановлении принятого сообщения из закодированного вида в общепринятый, доступный для потребителя.

В теории кодирования существует ряд направлений:
- статистическое или эффективное кодирование;
- помехоустойчивое кодирование;
- корректирующие коды;
- циклические коды;
- арифметические коды и др.

С глубокой древности люди искали эффективные способы передачи информации:
- движение факелов использовал древнегреческий историк Полибий (II в. до н. э.);
- оптический телеграф — семафор — впервые использовал Клод Шапп в 1791 г.;
- движение электромагнитной стрелки в электромагнитных телеграфных аппаратах впервые применили русский физик П. Л. Шиллинг (1832) и профессора Гёттингенского университета Вебер и Гаусс (1833);
- азбука и телеграфный аппарат Самюэла Морзе (1837);
- международный флажковый код для передачи информации оптическими сигналами впервые ввел капитан Фредерик Марьят в 1861 г. на основе свода корабельных сигналов;
- беспроволочный телеграф (радиопередатчик) был изобретен А. С. Поповым в 1895 г. и Маркони в 1897 г. независимо друг от друга;
- беспроволочный телефон, телевидение (1935), затем и ЭВМ — новые средства связи, появившиеся в XX в., с которыми связана новая эпоха в информатизации общества.

Одновременно с потребностью передавать информацию люди искали способы скрыть смысл передаваемых сообщений от посторонних любопытных глаз. Императоры, торговцы, политики и шпионы искали способы шифрования своих посланий. Образцы тайнописи можно встретить еще у Геродота (V в. до н. э.).

К тайнописи — *криптографии* (от греч. κρυπτος — тайна, γραφω — пишу) прибегал Гай Юлий Цезарь, заменяя в своих тайных записях одни буквы другими. Использовали шифрование не только древнегреческие жрецы, но и ученые Средневековья: математики итальянец Джероламо Кардано и француз Франсуа Виет, нидерландский гуманист, историк, юрист Гроций, выдающийся английский философ Фрэнсис Бэкон. Однако отцом криптографии считается архитектор Леон Баттиста Альберти (1404—1472), который ввел шифрующие коды и многоалфавитные подстановки.

В начале XX в. каждый германский офицер был обязан знать криптографию. Историки и лингвисты используют знания основ криптографии для расшифровки находок, дошедших до нас из глубин веков и выполненных, например, в виде клинописи или иероглифов. Так, в XIX в. немецкие филологи и историки Г. Ф. Гротефенд и Вестергорд расшифровали древнюю клинопись, а французский исследователь Ж. Шампольон — египетские иероглифы.

Первые манипуляции с символами в виде различных кодов возникли с потребностью шифровать информацию.

? Но какое отношение шифрование, коды и кодирование имеют к математике?

Оказывается, знания математики нужны для того, чтобы, во-первых, найти простую, но надежную систему кодирования, недоступную для расшифровки посторонним лицам, во-вторых, найти способы декодирования чужой системы тайнописи, чужих кодов.

Например, механическая замена одних букв или чисел другими — подстановки Цезаря — достаточно легко поддается дешифровке. Причем сам процесс декодирования аналогичен решению неопределенных уравнений со многими неизвестными. В отличие от любой шифровки, в основе принципа кодирования лежит замена исходной информации цифрами. Так, первый код, придуманный Полибием, состоял из 25 пар, включающих 25 букв и 5 цифр.

Сэр Фрэнсис Бэкон (1561—1626), английский писатель и философ, лорд-канцлер елизаветинской эпохи, автор двухлитерного кода, доказал в 1580 г., что для передачи информации достаточно двух знаков. Ф. Бэкон сформулировал три *требования* к шифру. Шифр должен быть:

несложен, прост в работе;

надежен, труден для дешифровки посторонним;

скрытен, по возможности не должен вызывать подозрений.

Шифры Бэкона — сочетание шифрованного текста с дезинформацией в виде нулей. Двузначные коды и шифры использовались задолго до появления ЭВМ.

Защита информации. В наше время потребность в кодировании информации не менее актуальна, чем в былые времена. Шифруется дипломатическая и экономическая корреспонденция, военные сообщения и медицинские диагнозы, сенсационные сообщения прессы и информация биржевых маклеров.

Кодирование имеет значение не только в конспиративных целях для шифровки информации. Так, в математике с помощью кодирования изучение одних объектов заменяют изучением других, более доступных или уже известных. Ярким примером кодирования в математике является метод координат, введенный Декартом, который дает возможность изучать геометрические объекты через их аналитическое выражение в виде чисел, букв и их комбинаций — формул.

Теория кодирования — довольно молодая наука. Автором первых циклических кодов был Ж. М. Э. Бодо (1889), одношаговых — Бодо и Грей (1953). Обеспечением надежности телеграфных кодов занимался Ван Дурен (1937). Код, исправляющий ошибки, изобрел Ричард Хемминг (1950).

Исследование надежности кодов получило новый импульс после создания в 1948 г. Клодом Эльвудом Шенноном *теории информации.*

В основе теории информации лежит гипотеза о статистическом характере источника сообщений. Случайная последовательность знаков не несет информации, так же как и ключ кода. А расшифровать код можно, используя знания о статистических закономерностях сообщения и кода. Теория количества информации Шеннона основана на известной со времен Аристотеля альтернативе выбора одного из двух знаков между 0 и 1.

Во все времена шифры являлись государственной тайной, требующей соответствующей защиты от любых посягательств.

В частности, в годы Второй мировой войны в группе дешифровки англичан успешно трудился английский математик Алан Матисон Тьюринг (1912—1954). Благодаря работе этой группы англичане владели способом дешифровки немецкой криптосистемы и своевременно узнавали о планах противника.

С появлением управляющих систем, в частности ЭВМ, роль кодирования существенно возросла и изменилась, так как без кодирования невозможна передача информации. В последнее время в связи с развитием телекоммуникационных систем и широким использованием вычислительной техники для обработки и хранения информации возникла новая область знаний — *информационная безопасность.*

На рубеже XX и XXI вв. потребность в защите информации возросла из-за широкого применения ЭВМ и появления Интернета. Доступность проникновения через Интернет в экономические, политические, военные системы любой страны выявила новую мировую проблему — борьбу с компьютерными взломщиками, хакерами, и компьютерными террористами. С одной стороны, стала актуальной разработка эффективной и надежной защиты, с другой, возросла ответственность за воспитание специалистов по работе с ЭВМ, обладающими моральными и нравственными принципами, не противоречащими общечеловеческим нормам.

Проблемами защиты информации занимается наука **криптология** (от греч. κρυπτος — тайный, λογος — наука), состоящая из криптографии и криптоанализа. Если **криптография** занимается поиском и исследованиями математических методов преобразования информации, то **криптоанализ** исследует принципы расшифровки сообщений без знаний ключа.

Современная криптография включает в себя разделы:

- симметричные криптоносители;
- криптосистемы с открытым ключом;
- системы электронной подписи;
- управление ключами.

Криптографические методы используются для передачи секретной информации по таким каналам связи, как, например, электронная почта, с целью установления истинности передаваемого сообщения, а также с целью хранения информации на носителях в зашифрованном виде.

Как было сказано, криптоаналитики часто пользуются математическими методами при работе с информацией. Так, методы декодирования включают в себя решение различных уравнений. Разгадывание занимательных заданий, арифметических ребусов, в которых одинаковые цифры заменены одинаковыми буквами, а

разные — разными, связано с решением уравнений со многими неизвестными и основано на разложении числа по степеням основания q.

Задача 33. Расшифруем один такой ребус (десятичная система счисления):

$$\begin{array}{r} \textit{КРОНА} \\ + \\ \textit{КРОНА} \\ + \\ \textit{КРОНА} \\ \hline \textit{ФРАНК} \end{array}$$

Решение. Во-первых, нужно учесть, что в реальной дешифровке результат должен быть единственным.

Представим зашифрованные числа в виде разложения в многочлен по степеням числа 10: $\textit{КРОНА} = (\text{К} \cdot 10^4 + \text{Р} \cdot 10^3 + \text{О} \cdot 10^2 + \text{Н} \cdot 10^1 + \text{А} \cdot 10^0)$. $3 \cdot (\text{К} \cdot 10^4 + \text{Р} \cdot 10^3 + \text{О} \cdot 10^2 + \text{Н} \cdot 10^1 + \text{А} \cdot 10^0) = \text{Ф} \cdot 10^4 + \text{Р} \cdot 10^3 + \text{А} \cdot 10^2 + \text{Н} \cdot 10^1 + \text{К} \cdot 10^0$. Ребус представился уравнением на множестве цифр {0, 1, 2, ..., 9}. Но это не единственная информация, которую можно выжать из перевода слов ребуса в цифры. Есть уравнение из последнего столбца «последняя цифра числа $3 \cdot$ А равна К», что на языке деления с остатком (см. подразд. 6.6) имеет лаконичную запись: $(3 \cdot \text{А}) \bmod 10 = \text{К}$. В следующем разряде будет: $\left(3\text{Н} + \dfrac{3\text{А} - \text{К}}{10}\right) \bmod 10 = \text{Н}$ и т.д. Но дешифровщик-перехватчик не знает ключей и кодов, ему приходится работать только с самим перехваченным сообщением и довольствоваться минимумом информации. Поэтому можно попробовать обойтись только самим алгебраическим уравнением. Перенесем неизвестные влево. После приведения подобных слагаемых получили одно уравнение с шестью неизвестными:

$$29\,999 \cdot \text{К} - 10\,000 \cdot \text{Ф} - 2000 \cdot \text{Р} + 300 \cdot \text{О} - 97 \cdot \text{А} + 20 \cdot \text{Н} = 0.$$

При решении учтем, что все неизвестные — цифры, символизирующие целые неотрицательные числа от 0 до 9, причем Ф, $\text{К} \neq 0$.

Воспользуемся методом решения числовых ребусов на чашечных весах, а результаты оформим в виде таблицы (табл. 6.1). Смысл метода заключается в последовательном подборе таких значений неизвестных, чтобы частичные суммы от последовательного сложения одночленов этого многочлена принимали минимальное по модулю значение. Расположив одночлены в порядке убывания абсолютной величины коэффициентов ($29\,999 > 10\,000 > 20\,000 > 300 > 97 > 20$), начнем с первой частичной суммы $29\,999 \cdot \text{К}$ и заполним графы табл. 6.1.

Таблица 6.1

Табличный метод решения числовых ребусов

Число слагаемых в сумме	Вид суммы	Знак последующего слагаемого	Значение буквы	Значение суммы
1	$29999 \cdot К$	Не использован (–)	$К = 1$ (min, так как $К \neq 0$)	29999
2	$29999 - 10000 \cdot Ф$	–	$Ф = 3$	–1
3	$-1-2000 \cdot Р$	–	$Р = 0$	–1
4	$-1 + 300 \cdot О$	–	$О = 2$ (так как 1 уже есть: $К = 1$)	599
5	$599 - 97 \cdot А$	+ (так как min был бы при $А = 6$, но $17 + 20 \cdot Н = 0$ не решается)	$А = 7$	–80
6	$-80 + 20 \cdot Н$	–	$Н = 4$	0

Итак, закодированная информация имела вид $\begin{array}{r} 10247 \\ + \\ 10247 \\ + \\ 10247 \\ \hline 30741 \end{array}$.

Здесь дешифровщик уже знал, в каком алфавите зашифрованы сообщения. Малая мощность алфавита десятичной системы позволила бы и без метода чашечных весов путем обыкновенного перебора вариантов легко расшифровать ребус. Но на практике перебор всех вариантов осуществляется с помощью ЭВМ, и даже простейшая задача разложения натурального числа на простые множители (а это необходимо при некоторых способах шифрования и дешифрования) требует заведомо большого времени.

В реальных условиях при минимуме информации, имея только само перехваченное сообщение, необходимо в короткий срок расшифровать его и принять адекватное решение. Поэтому эффективность всей работы зависит от мастерства дешифровщика и программиста.

При этом дешифровщик несет всю ответственность за быстроту обработки информации и ее достоверность. Например, в сфере банковской деятельности неправильное прочтение сообщений может привести к финансовым потерям и другим негативным последствиям.

6.2. Системы счисления для представления информации в ЭВМ

Ей было тысяча сто лет,
Она в сто первый класс ходила,
В портфеле по сто книг носила.
Все это правда, а не бред.
Когда, пыля десятком ног,
Она шагала по дороге,
За ней всегда бежал щенок
С одним хвостом, зато стоногий.
Она ловила каждый звук
Своими десятью ушами,
И десять загорелых рук
Портфель и поводок держали.
И десять темно-синих глаз
Рассматривали мир привычно...
Но станет все совсем обычно,
Когда поймете мой рассказ.

А. Н. Стариков

В реальной жизни информация представляется как в непрерывном, так и в дискретном виде. В ЭВМ можно использовать только дискретную информацию, представленную числами в некоторой системе счисления.

При выборе системы счисления необходимо учитывать ряд условий:

- наличие взаимно-однозначного соответствия между некоторой информацией и ее закодированным аналогом, необходимого для ее кодирования и декодирования;
- возможность кодирования любой информации в выбранной системе счисления;
- простота применения кода для чисел в данной системе счисления.

Система счисления — это совокупность приемов и правил записи чисел с помощью символов. В нашей стране принята наиболее распространенная в мире десятичная система счисления (ДСС), в которой для записи любого числа используются десять цифр от 0 до 9.

Для того чтобы различить числа, написанные в разных системах счисления одинаковыми цифрами, пользуются индексом, равным основанию системы счисления. Например:

$101_{10} = 1 \cdot 10^2 + 0 \cdot 10^1 + 1 \cdot 10^0 = 101_{10}$;
$101_8 = 1 \cdot 8^2 + 0 \cdot 8^1 + 1 \cdot 8^0 = 65_{10}$;
$101_2 = 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 5_{10}$.

Система счисления называется **позиционной** (ПСС), если значение цифры зависит от ее положения в числе. Например, в системе счисления с основанием 10 число 9597 означает: $9597 = 9 \cdot 1000 + 5 \cdot 100 + 9 \cdot 10 + 7 \cdot 1 = 9 \cdot 10^3 + 5 \cdot 10^2 + 9 \cdot 10^1 + 7 \cdot 10^0$, т.е. «цифра 5» на третьем справа месте означает вклад в сумму, равный $5 \cdot 10^{3-1}$.

Любое действительное число A можно представить в виде суммы разложений по степеням основания системы счисления:

$$A_q = \sum_{i=-m}^{n} a_i q^i = a_n q^n + a_{n-1} q^{n-1} + \ldots + a_1 q^1 + a_0 q^0 + a_1 q^{-1} + \ldots + a_{-m} q^{-m}$$

с произвольным числом $q > 1$, называемым *основанием*. В общем случае $m \to \infty$, т.е. число A представимо в виде бесконечного ряда из степеней q. Понятно, что при заданном основании число A однозначно определяется всеми коэффициентами a_i с соблюдением порядка, т.е. кортежем длины $n + m + 1$. Общепринятым считается следующее представление:

$$A_q = \underbrace{a_n a_{n-1} \ldots a_1 a_0}_{\text{целые разряды}} \bullet \underbrace{a_{-1} \ldots a_{-m+1} a_{-m}}_{\text{дробные разряды}}.$$

плавающая точка

Для эффективности такой свертки важно, чтобы каждый из коэффициентов a_i содержал как можно меньше информации и представлялся только одним символом. Этого можно добиться, выбирая основание q натуральным. Тогда каждый коэффициент a_i будет также натуральным числом или нулем и будет называться *цифрой* в ПСС с основанием q, причем для соблюдения единственности представления цифра a_i должна быть меньше, чем q: $\forall i \in \{-m, \ldots, n\}\ 0 \le a_i < q$. Тогда, если $m < \infty$, то число A будет **рациональным**; если $m = 0$, то число A будет **целым**.

Теперь должен быть понятен смысл операции взятия целой части числа не только в абсолютном смысле, но и в виде представления в ПСС. Для неотрицательного числа нужно из записи числа в этой системе отбросить все знаки, стоящие справа от плавающей точки (иногда она изображается запятой), для отрицательного — отбросить дробные разряды (если они существуют), сохранить знак «минус» и вычесть единицу. Например, $[103{,}234_7] = 103_7$; $[-1003_6] = -1003_6$; $[-1003{,}005_6] = -1004_6$; $-788{,}12_9 = -800_9$.

В различных ПСС одно и то же число (в абсолютном смысле) имеет различный вид, и, наоборот, одинаковой записи в разных ПСС соответствуют разные числа. Например:

$123\,321_{10} = (1 \cdot 10^5 + 2 \cdot 10^4 + 3 \cdot 10^3 + 3 \cdot 10^2 + 2 \cdot 10^1 + 1 \cdot 10^0)_{10} = 123\,321_{\text{абс}}$;

$123\,321_8 = (1 \cdot 8^5 + 2 \cdot 8^4 + 3 \cdot 8^3 + 3 \cdot 8^2 + 2 \cdot 8^1 + 1 \cdot 8^0)_8 = 42\,705_{\text{абс}}$;

$123\,321_4 = (1 \cdot 4^5 + 2 \cdot 4^4 + 3 \cdot 4^3 + 3 \cdot 4^2 + 2 \cdot 4^1 + 1 \cdot 4^0)_4 = 1785_{\text{абс}}$,

где под символами 1, 2, 3 обозначены первая, вторая и третья (после 0) цифры соответствующих ПСС: $1_8 = 1_{\text{абс}} = 1_4$; $2_8 = 2_{\text{абс}} = 2_4$; $3_8 = 3_{\text{абс}} = 3_4$.

Вес разряда p_i в позиционной системе счисления есть отношение $p_i = q^i/q^0 = q^i$, где i — номер разряда при отсчете справа налево от плавающей точки. При накоплении в данном k-м разряде значения большего или равного единице разряда q_i накопленная единица переходит в старший разряд, осуществляется перенос, т.е. сложение этой дополнительной единицы с числом $k + 1$ разряда. При вычитании, если в k-м разряде недостает единиц, то производится заем необходимых единиц в $(k - 1)$-м разряде. Именно на этом и основано известное каждому сложение и вычитание столбиком.

Назовем **длиной числа** (ДЧ) количество разрядов в разложении этого числа по разрядам. В различных системах счисления одно и то же число имеет разную длину. Длина числа интересует специалистов с точки зрения записи этого числа в ЭВМ, так как эта длина не должна превышать длины разрядной сетки (ДРС). Если любое $n > 0$ — длина разрядной сетки ЭВМ, то максимальное и минимальное числа, которые можно уместить в этом компьютере, находятся по формулам: $A_{q\min} = -(q^n - 1)$, $A_{q\max} = q^n - 1$.

Но это же свойство конечности ДРС $(0 < n < \infty)$ делает невозможным точное представление дробных чисел, у которых полное число разрядов превышает ДРС, в частности иррациональных. Очевидно, что представление чисел возможно лишь с «шагом», большим $q^{-(n-1)}$.

Тогда *диапазон представления* чисел (ДП) в системе счисления с основанием q будет интервалом на числовой оси между минимальным и максимальным числом, т.е. $A_{q\min} \le \text{ДП} \le A_{q\max}$.

Поэтому выбор системы счисления определяется техническими возможностями ЭВМ. В настоящее время для кодирования информации приняты двоичная, восьмеричная и шестнадцатеричная системы счисления.

Двоичная система счисления. Для представления информации в ЭВМ используется двоичная система счисления, предложенная еще Фрэнсисом Бэконом.

Основание системы $q = 2$, а алфавит состоит, соответственно, из двух цифр: 0 и 1. Двоичный код имеет ряд достоинств:

- простота в обращении, так как использует всего два символа 0 и 1;
- соответствие этих символов техническим характеристикам цифровых схем, также имеющих два устойчивых состояния: рабочее (1) и нерабочее (0).

Многие явления природы и техники описываются также двумя состояниями: есть — нет, включен — выключен, проходит (ток, сигнал) — не проходит и др.

Эти же символы 0 и 1 Фрэнсис Бэкон ввел и для кодирования логических понятий «истина» (1) и «ложь» (0). В глобальном философском смысле это известные противоречивые понятия p и $\bar{p}$.

Таблица 6.2

Вес первых двенадцати разрядов

$q_1 = 10$	2048	1024	512	256	128	64	32	16	8	4	2	1
$q_2 = 2$	2^{11}	2^{10}	2^9	2^8	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0

В двоичной системе, как и в любой другой ПСС, каждый разряд имеет определенный вес, а именно 2^k для любого $k \in \mathbb{Z}$.

Вес первых 12 разрядов цифр двоичного числа можно представить в виде табл. 6.2.

В двоичной записи каждый разряд несет свою единицу информации и называется **битом** (от англ. *binary digit* — двоичное число). Самый младший двоичный разряд называется **наименьшим**, а самый старший — **наибольшим** значащим битом. Каждый из них имеет соответственно наименьший и наибольший вес. В двоичной записи числа это соответственно крайний слева и крайний справа биты. Для измерения большого количества информации используется еще одна единица измерения: 8 бит образуют 1 *байт*.

Еще одним достоинством двоичной системы счисления является тот факт, что $2^{10} = 1024_{10} \approx 1000_{10}$. Хотя достоинством это является лишь потому, что ДСС является наиболее употребительной, но, тем не менее, во-первых, в информатике для построения производных от основной единицы информации приходится использовать приставки, аналогичные физическим величинам. Поэтому приставка «кило-» обозначает множитель 2^{10}, а «мега-» — 2^{20} в отличие от физических величин. Во-вторых, для нахождения информации часто требуется логарифмировать по основанию 2. Поэтому можно без применения микрокалькуляторов вычислять приближенные значения логарифмов чисел, представленных в ДСС. Например, требуется посчитать $\log_2 123\,456\,789$. Имеем: $\log_2 123\,456\,789 \approx \log_2 123\,000\,000 = \log_2(123 \cdot 1000 \cdot 1000) = \log_2 123 + \log_2 1000 + \log_2 1000 \approx \log_2 123 + 10 + 10 \approx \log_2 128 + 20 = 27$. Точное значение 26,88. Таким образом, нужно просто на каждые три значащие цифры влево от плавающей точки прибавлять по 10 и брать логарифм оставшихся разрядов. А оставшееся число (оно всегда меньше тысячи) удобно сравнить с ближайшей к нему степенью двойки (см. табл. 6.2).

Техника перевода целых чисел из одной системы счисления в другую. Предлагаемые здесь действия можно проводить в любой ПСС, например в десятичной.

1. Разделить данное число на основание новой системы счисления q. Значение остатка на первом этапе деления назовем младшим значащим разрядом (МЗР).

2. Дальнейшее деление полученного частного на каждом следующем шаге дает следующие за младшим значащие разряды.

3. Процесс деления заканчивается, когда частное на отдельном шаге окажется равным нулю. Полученный на этом этапе остаток от деления всегда меньше основания q и является старшим значащим разрядом (СЗР), расположенным крайним слева в кодовой записи числа.

4. Полученная запись всех остатков от деления слева направо (от СЗР к МЗР) и есть запись числа в новой системе счисления.

Латинские и русские буквы на клавиатуре ЭВМ тоже имеют свой код и в десятичной, и в двоичной системах счисления (ASCII-таблица). Так, буква А соответствует 01000001; Б — 010000010 и т.д.

Несмотря на очевидные преимущества двоичной системы счисления, она имеет значительный недостаток — громоздкую форму записи двоичного слова. Поэтому в ЭВМ нашли применение и другие системы счисления.

Шестнадцатеричная система счисления. Система счисления с основанием $q = 16$ ($16 = 2^4$) дает самую компактную запись числа в ЭВМ. Для перевода двоичного числа в эту систему разбивают двоичную запись на четверки (тетры) и заменяют их соответствующим эквивалентом в шестнадцатеричной системе, пользуясь таблицей перевода (табл. 6.3).

Таблица 6.3

Перевод десятичного числа в системы счисления с основаниями 2, 8, 16

Десятичное число	Двоичное число	Восьмеричное число	Шестнадцатеричное число
0	0000	00	0
1	0001	01	1
2	0010	02	2
3	0011	03	3
4	0100	04	4
5	0101	05	5
6	0110	06	6
7	0111	07	7
8	1000	10	8
9	1001	11	9
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F
16	10000	20	10
17	10001	21	11

Пример. Перевести $F5_{16}$ в двоичную систему счисления, а число 10001110_2 в шестнадцатеричную систему счисления. По табл. 6.3 находим: $F_{16} = 1111_2$, $5_{16} = 0101_2$; $1000_2 = 8_{16}$, $1110_2 = E_{16}$. Тогда $F5_{16} = 11110101_2$, а $10001110_2 = 8E_{16}$.

При чтении чисел в иных, не десятичных, системах счисления принято называть отдельно каждую цифру. Сравним: 25_{10} — это двадцать пять, но 25_8 читается как два—пять. В ЭВМ самостоятельно осуществляется перевод чисел из одной системы счисления в другую.

Проблема отношения между основанием ПСС, т.е. мощностью алфавита, и длиной получившегося слова целиком лежит в плоскости закона обратной зависимости объема и содержания. Обратим внимание на следствие этого в лингвистике. В последнее время встала проблема поиска оптимального алфавита, обострившаяся в связи с попытками перевести письменность некоторых народов с кириллицы (33 литеры) на латинский алфавит (26 знаков). Ученые считают, что 32—33 знака являются оптимальным объемом алфавита и для запоминания, и для написания и произношения слов. По этой же причине системы счисления с основанием $q = 10—16$ являются оптимальными, а десятичная — наиболее употребляемой из них.

6.3. Основные понятия вероятностной теории информации

> Информация проникает во все поры жизни людей и обществ... Жизнь невозможна ни в вещественно-энергетическом, ни в информационном вакууме.
>
> *А. Берг*

Сигнал есть сообщение, передаваемое с помощью некоторого носителя. Если источник вырабатывает сигнал, параметр которого является непрерывной функцией времени, то он называется **непрерывным**. Непрерывна и информация, передаваемая такими сигналами.

Сигнал называется **дискретным**, если соответствующие ему числовые значения существуют в конечный момент времени и представляют собой конечную последовательность символов, например 0 и 1.

В дискретных сообщениях, конечных или бесконечных последовательностях знаков, можно проводить разбиения на конечные последовательности знаков — **слова**.

Пусть дано некоторое множество знаков M. Говорят, что задано **слово над множеством** M, если оно составлено из элементов

множества M. Для множества двоичных символов $B = \{0, 1\}$ двоичным словом длины n будет последовательность знаков, содержащая k нулей и l единиц, причем $k + l = n$. Например, кортеж 010111 есть слово длины 6 над B, двоичный код некоторого числа, содержащий $k = 2$ нулей и $l = 4$ единиц. Длина двоичного слова может быть неограниченной.

Процедура преобразования непрерывных сигналов в дискретные называется **дискретизацией**. Для ее проведения выбирают из алфавита конкретные символы, которые характеризуют все целостное сообщение.

К анализу информации возможны различные подходы.

Объемный подход связан с подсчетом количества информации в двоичной системе счисления с учетом количества символов.

Вероятностный подход учитывает частоту появления символа в тексте.

Однако при таких подходах упускается истинность (семантическая характеристика), а также полнота, актуальность, ценность текста.

К *свойствам информации* можно отнести:

- запоминаемость (возможность сохранения);
- передаваемость без искажений, с помощью различных каналов связи;
- воспроизводимость;
- преобразуемость, т.е. представление информации в разной форме, в том числе ее копирование;
- стираемость — преобразование, при котором количество информации в новой форме уменьшается до нуля.

Измерение информации. Математическое представление об информации связано с ее измерением. Существует несколько подходов к измерению количества информации.

Под **количеством информации** принято понимать числовую характеристику сигнала, которая не зависит от его формы и содержания и соответствует степени неопределенности, исчезающей после получения сообщения в виде этого сигнала.

Дело в том, что информация, переданная в виде последовательности сигналов (знаков), представляет собой определенный *порядок*, некоторую структурную определенность. Но порядок не существует сам по себе, а всегда сопоставим с *беспорядком* — отсутствием структуры.

Мерой беспорядка любой системы является **энтропия**. Понятие энтропии как меры беспорядка системы заимствовано из термодинамики и стало одним из основных не только в теории информации, но и в современном философском осмыслении открытых биосоциальных систем.

В замкнутых системах любые процессы протекают так, чтобы энтропия системы не убывала.

Реальная ценность понятия энтропии определяется в первую очередь тем, что выражаемая им «степень неопределенности» опытов оказывается во многих случаях именно той характеристикой, которая играет роль в разнообразных процессах, встречающихся в природе и технике и так или иначе связанных с передачей и хранением каких-либо сообщений.

Формула Хартли. Итак, в передачу сообщения входит процесс выбора, который осуществляет как отправитель, так и получатель сообщения. Этот процесс выбора можно измерить количественно.

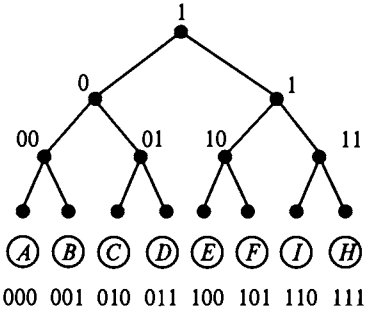
Пусть нам необходимо передать одно из восьми возможных сообщений: *A*, *B*, *C*, *D*, *E*, *F*, *I*, *H*. Условимся для удобства поставить в соответствие каждому сообщению двоичное число и представим результат в виде графа (рис. 6.1), где каждому сообщению соответствует свой набор сигналов.

Тогда один двоичный сигнал позволяет выбрать одно из двух сообщений: $2 = 2^1$. Два двоичных сигнала позволяют выбрать одно из четырех сообщений: $4 = 2^2$. Набор из трех сигналов — одно из восьми сообщений: $8 = 2^3$, из шести — одно из шестидесяти четырех сообщений: $64 = 2^6$, из m двоичных сигналов — одно из 2^m сообщений. И наоборот, для набора $N = 2^m$ сообщений нужно иметь не менее m сигналов. Важно знать, что одно из тридцати двух сообщений выбирается с помощью пяти сигналов. Поэтому для распознавания одного определенного сообщения из N возможных сообщений требуется $m = \log_2 N$ сигналов.

Так, при $N = 100$ требуется $m = \log_2 100 = \log_2 10^2 = 2\log_2 10 = 2 \cdot 3{,}32 = 6{,}64 \approx 7$ знаков. Следовательно, для выбора одного из N вариантов нужна информация в J бит: $J = \log_2 N$.

Эту формулу в 1928 г. предложил американец Хартли. В качестве единицы измерения используется бит. Таким образом, для выбора из 8 знаков потребуется 3 бита.

Формула Шеннона. Если сообщение появляется с одинаковой частотой, то полученные выводы единственно возможные. Если относительная частота сообщений одинакова, то для одного из



A	*B*	*C*	*D*	*E*	*F*	*I*	*H*
000	001	010	011	100	101	110	111

Рис. 6.1. Двоичное дерево для восьми символов

четырех сообщений требуется $J = \log_2 4 = 2$ бита. Если частота появления отдельных сообщений различна, то может быть использовано и меньшее количество сигналов. Пусть задана вероятность событий, соответствующая схеме (рис. 6.2): $P(A) = \frac{1}{2}$; $P(B) = \frac{1}{4}$; $P(C) = P(D) = \frac{1}{8}$, причем $P(A) + P(B) + P(C) + P(D) = 1$ как вероятность достоверного события.

Если закодировать буквы двоичными символами: A — 0 (1 двоичный знак), B — 10 (2 двоичных знака), C — 110 (3 двоичных знака), D — 111 (3 двоичных знака), то с учетом частоты появления каждого из двоичных знаков имеем $J = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 = 1\frac{3}{4}$ бит, что действительно меньше 2 бит.

В естественном языке, например русском, учет вероятности (частоты) появления слов различной длины заключается в том, что с большей вероятностью появляются короткие слова. В последнем рассмотренном примере для передачи сообщения A с вероятностью $P(A) = \frac{1}{2}$ необходимо $J = \log_2 \frac{1}{P(A)} = \log_2 \frac{1}{1/2} = \log_2 2 = 1$ бит, т.е. нужен один двоичный знак. Для передачи сообщения B с вероятностью $P(B) = \frac{1}{4}$ нужно $J = \log_2 \frac{1}{P(B)} = \log_2 \frac{1}{1/4} = \log_2 4 = 2$ бита, т.е. два двоичных знака. Для передачи сообщений C и D с вероятностью $P(C) = P(D) = \frac{1}{8}$ нужно $J = \log_2 \frac{1}{1/8} = \log_2 8 = 3$ бита, т.е. нужно три двоичных знака. Это значит, что если набор возможных сообщений, имеющихся у источника информации, и их вероятности известны, то для каждого отдельного сообщения количество информации равно логарифму величины, обратной частоте (статистической вероятности) появления этого сообщения.

Для расчета среднего количества информации суммируем произведения статистической вероятности на количество информации каждого сообщения, т.е. на логарифм величины, обратной этой вероятности, и обозначим эту сумму через H. Тогда

$$H = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} = -\sum_{i=1}^{n} p_i \log_2 p_i.$$

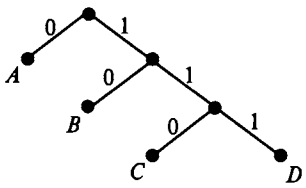


Рис. 6.2. Двоичное дерево для четырех символов

Это формула К. Шеннона, дающая среднее значение *количества информации*, приходящееся на один символ алфавита.

Функция H достигает максимума, когда частота для всех n сообщений одинакова и, следовательно, равна $p_i = 1/n$. В этом случае

$$H_{\max} = \sum_{i=1}^{n} \frac{1}{n} \log_2 n = n \frac{1}{n} \log_2 n = \log_2 n.$$

Таким образом, формула Р. Хартли является частным случаем формулы Шеннона, когда исходы равновероятны. В тех случаях, когда мы не обладаем всей информацией о статистической вероятности исходов, а условно считаем их равновероятными, мы также будем пользоваться для подсчетов формулой Хартли. Чтобы отличать такое ее применение от непосредственного, будем использовать понятие частной информации. Тем самым мы подчеркиваем тот частный случай, когда принимаем условие равновероятных исходов. Частную информацию, содержащуюся в сообщении X_i, будем находить по формуле $J_{X_i} = \log_2(1/p_i)$.

Отметим очевидные следствия формулы Хартли. Если возможно передать только один сигнал ($n = 1$), то $H = \log_2 1 = 0$. Видим, что подобные сообщения не содержат информации вообще. Например, если каждое утро восходит солнце, то, вставши в любой день и увидев светящее солнце, мы ничего нового не узнаем.

Если с равной вероятностью могут передаваться два сигнала ($n = 2$), то $H = \log_2 2 = 1$ бит. Это также понятно, поскольку формула Хартли как раз и строилась в предположении о двоичности бита, и мы лишь проверили ее действие на себя и тем самым непротиворечивость. Так, нетикающие часы свидетельствуют нам о том, что они остановились, тикающие — что идут.

С вероятностной точки зрения рассмотрим H_1 — энтропию до принятия сообщения и H_2 — энтропию после принятия сообщения. Тогда $J = H_1 - H_2$.

Пусть эксперимент состоит в бросании игральной кости, где J — информация о результате бросания. По окончании эксперимента известен результат бросания, неопределенность снята и $H_2 = 0$, поэтому информация равна энтропии системы $J = H$.

Задача 34. Найдем количество информации, имеющееся при получении сообщения на русском языке.

Решение. Так как в русском языке 33 буквы и 1 пробел между словами, то необходимо 34 места для символов. Тогда по формуле Хартли имеем $J = \log_2 34 = 5{,}09$ бит.

В словах русского языка буквы встречаются с разной частотой, поэтому необходим вероятностный подход, т. е. учет статистической вероятности появления буквы в тексте. Анализ больших текстов дает такие численные значения статистической вероятности появления букв в тексте.

Если учитывать частоту появления каждой буквы в тексте, то по формуле Шеннона $H = 4{,}35$ бит — для русского текста. В среднем $J = \log_2 27 = 4{,}76$ бит — для всех языков, основанных на латинском алфавите, в котором 26 букв.

Задача 35. На одной из клеток шахматной доски стоит фигура. Будем считать, что все положения этой фигуры на шахматной доске равновероятны.

1. Какое количество информации несет в себе сообщение о точном месте нахождения этой фигуры?

Решение. Неопределенность информации, т.е. энтропия этой системы A с N равновероятными состояниями, равна по формуле Хартли $\log_2 N$, поэтому информация $J = H(A) = \log_2 64 = \log_2 2^6 = 6$ бит. Сообщение о том, что фигура находится, например, на клетке D5, несет в себе 6 бит информации.

2. Какое количество информации несет в себе сообщение о том, что фигура находится на одной из угловых клеток доски?

Решение. Так как на доске четыре угловых клетки, а вероятность находиться на каждой из них 1/64, то вероятность попасть на угол $p = 4/64 = 1/16$. Тогда частная информация сообщения $J_A = \log_2 \frac{1}{p(A)} = \log_2 \frac{1}{1/16} = 4$ бит.

Задача 36. Какую частную информацию несет сообщение: «В субботу я иду в гости»?

Решение. В гости можно пойти в любую субботу года. Число суббот в году: $12 \cdot 4 = 48$. Таким образом, вероятность пойти в гости в субботу $P = 1/48$, а частная информация $J_{A_i} = \log_2 \frac{1}{P_i} = \log_2 \frac{1}{1/48} = \log_2 48 = 5{,}59 \approx 5{,}6$ бит.

Задача 37. Мой новый знакомый сообщил: «На этой неделе мой день рождения».

1. Какое количество информации заключено в сообщении?

Решение. Так как любой день недели равновероятен, то $P_1 = P_2 = \ldots = P_7 = 1/7$. Тогда частная информация $J_A = \log_2 \frac{1}{1/7} = \log_2 7 = 2{,}8 \approx 3$ бит.

2. Каким минимальным числом уточняющих вопросов, подразумевающих ответы «да» и «нет», можно узнать точный день рождения знакомого?

Решение. Так как информация заключена в сообщении $I_A = 3$, то необходимо минимум три вопроса для выяснения точного дня рождения. Представим все дни недели в виде отрезка с семью делениями (рис. 6.3).

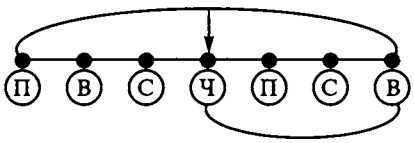


Рис. 6.3. Метод половинного деления

Принцип постановки вопроса основан на методе половинного деления. Пусть неизвестный день рождения в субботу, тогда вариант диалога будет таким.

Вопрос 1: «Твой день рождения до четверга?»

Ответ 1: «Нет».

Вопрос 2: «Твой день рождения после пятницы?»

Ответ 2: «Да».

Вопрос 3: «Твой день рождения в воскресенье?»

Ответ 3: «Нет».

Вывод: днем рождения знакомого будет суббота.

Из рассмотренных примеров и задач видно, что если о системе известны определенные начальные условия, то можно измерить количество информации, заключенной в сообщении об этой системе.

Итак, *теория информации* — один из разделов математики, появившийся сравнительно недавно. XXI в. называют веком информации из-за ее фундаментального значения для общества и всех объектов живой природы. Любое выдающееся общественное явление, значительное событие культуры, достижение современной техники по своей сути являются различными видами передачи, переработки и хранения информации. Обработка разнообразной информации — основной смысл функционирования ЭВМ.

Математическая теория информации появилась после того, как было установлено, что количеству информации можно поставить в соответствие число аналогично тому, как с помощью чисел выражаются различные физические величины: скорость, температура, масса и др.

Принцип получения формулы для количества информации можно проиллюстрировать на примере игры «Бар-Кохба». Смысл игры заключается в том, что для установления некоторой информации можно задавать уточняющие вопросы, на которые даются лишь два ответа: «Да» и «Нет» или на языке двузначной логики «1» и «0». Тогда последовательности вопросов будет соответствовать серия ответов в виде кодового слова над множеством $\{0, 1\}$.

Проведя аналогию между игрой и подсчетом количества информации, можно прийти к выводу, что для измерения количества некоторой информации необходимо закодировать ее в виде последовательности нулей и единиц. Тогда мерой количества информации будет служить длина кодового слова. При таком количественном подходе, естественно, содержание информации, ее ценность и значение не учитываются.

Любое сообщение, подразумевающее выбор одной из двух альтернатив, содержит 1 бит информации. Так, 1 бит информации несет в себе ответ на любой уточняющий вопрос.

Обычно, в зависимости от характера информации, выбирается наиболее целесообразный способ ее кодирования. Телевизионный канал связи кодирует изображение в виде совокупности некоторых «точек». Черно-белые телевизионные приемники получают два вида сигналов. Назовем их условно «точки черного и белого цвета», которые переданы при помощи электромагнитных волн по каналу связи. Мы получаем уже раскодированную информацию, которую телевизор представил вновь в виде изображения.

Наше письмо — тоже способ кодирования. Звукам мы ставим в соответствие буквы некоторого алфавита. Тогда процесс декодирования есть чтение написанного текста. В таком случае устройством, кодирующим и декодирующим озвученную информацию, является мозг человека, а каналом связи — написанный текст.

В индейских племенах существовал своеобразный канал связи в виде узелков на веревке. Такие «узелки на память» могли передавать различные сообщения, т. е. служить источником информации.

В горах альпинисты, терпящие бедствие, передают сигнал SOS свистками, а мореплаватели на судах — световыми сигналами.

Студия звукозаписи тоже осуществляет кодирование музыкальных звуков и речи и их запись на магнитную кассету. А композитор кодирует ту же музыку с помощью семи нот.

Задача 38. Какими предварительными знаниями необходимо обладать для того, чтобы подсчитать, сколько бит информации содержит номер моего паспорта?

Решение. Так как в стране 145 миллионов россиян, то надо найти длину кодового слова, соответствующего одному из двух вариантов, для того чтобы определить, мой это паспорт или нет. Тогда $2^J = 145\,000\,000$ или $J = \log_2 145\,000\,000$, т. е. $J \approx 27$.

Формула Хартли $J = \log N$ позволяет установить важное свойство **аддитивности информации** для определения количества информации, соответствующего двум сообщениям. Если $J_1 = \log_2 N_1$, а $J_2 = \log_2 N_2$, то $J_1 + J_2 = \log_2 N_1 + \log_2 N_2 = \log_2 (N_1 N_2)$.

Хорошо известное из курса математики свойство суммы логарифмов можно также объяснить с помощью игры «Бар-Кохба». Пусть нам необходимо отгадать информацию X_1 как элемент множества H_1, содержащий N_1 элементов, и X_2 — элемент множества H_2, содержащего N_2 элементов. Это значит, что необходимо установить информацию о некоторой паре (x_1, x_2), которая взята из множества $H_1 \times H_2$, содержащего $N_1 \cdot N_2$ элементов. Таким образом, чтобы установить пару (x_1, x_2), необходимо по формуле Хартли задать $\log_2 (N_1 N_2)$ уточняющих вопросов и, следовательно, получить $\log_2 (N_1 N_2)$ бит информации..

Ту же самую информацию можно получить, отгадав каждый из элементов X_1 или X_2 по отдельности. Для отгадывания X_1 потребуется $\log_2 N_1$ уточняющих вопросов, а для X_2 соответственно $\log_2 N_2$

вопросов. Поэтому количество информации, необходимое для определения (x_1, x_2), равно $\log_2 N_1 + \log_2 N_2$. Тождественность разных подходов к установлению количества информации, таким образом, математически записывается в виде свойства *аддитивности*: $\log_2(N_1 N_2) = \log_2 N_1 + \log_2 N_2$.

? Какие синонимы понятия «информация» вы знаете?

Знание формулы Хартли для подсчета количества информации помогает решать различные задачи в одном из разделов теории информации — **теории поиска.** Задачи этого раздела аналогичны задачам поиска фальшивой монеты из некоторого множества, решение которых было рассмотрено в теории графов. Фальшивой может оказаться любая из имеющихся 27 монет, поэтому можно воспользоваться формулой Хартли. Минимальное количество информации о фальшивой (тяжелой либо легкой) монете равно $J = \log_2 27 = \log_2 3^3 = 3\log_2 3$.

С другой стороны, любое взвешивание на чашечных весах даст нам три варианта ответа: либо тяжелее правая чашечка, либо левая, либо весы уравновешены. Поэтому при одном взвешивании становится известным количество информации в $\log_2 3$ бит. В таком случае при m взвешиваниях будут выявлены $m\log_2 3$ бит — максимальное количество информации. Поэтому $m\log_2 3 \geq 3\log_2 3$. Так как $\log_2 3 > 1$, то, разделив обе части неравенства на $3\log_2 3$, имеем $m \geq 3$, т.е. для выявления фальшивой монеты необходимо не менее трех взвешиваний, что и было показано в гл. 2.

? Зависит ли количество информации от степени точности сформулированных вопросов?

Откуда произошла игра «Бар-Кохба» и что означает это название?

Согласно легенде Бар-Кохба («сын звезды») — предводитель восстания против владычества римлян — послал в лагерь римлян разведчика, которого римляне схватили, жестоко пытали и отрезали ему язык. Сбежав от врагов и вернувшись в осажденную крепость, несчастный лазутчик не имел возможности рассказать о том, что разведал в стане римлян. Тогда ему стали задавать лишь уточняющие вопросы, на которые можно было отвечать кивком головы — положительным, в случае согласия, и отрицательным в противоположном случае.

Задача 39. Найти количество информации, которое можно получить при ответе на вопрос: при бросании игрального кубика у вас выпала цифра 5?

Решение. Вероятность выпадения числа 5 на кубике $p = \frac{1}{6}$, а выпадения другой цифры $\bar{p} = \frac{5}{6}$. Тогда по формуле Шеннона имеем $J = -\frac{1}{6}\log_2 \frac{1}{6} - \frac{5}{6}\log_2 \frac{5}{6} = 0{,}65$ бита.

6.4. Обработка сообщений как кодирование

Ты должен идти не той дорогой, какой идут другие, но той, которой следует идти.

Сенека

Основные понятия теории кодирования. Пусть заданы два алфавита A и B, состоящие из конечного числа символов: $A = \{a_1, a_2, ..., a_N\}$ и $B = \{b_1, b_2, ..., b_M\}$. Кортеж в алфавите A назовем **словом** $\alpha = a_1, a_2, ..., a_i, ..., a_n$, где $i \in [1, n]$, число n показывает длину слова α и обозначается $n = l(\alpha)$.

Пустое слово обозначается Λ: $l(\Lambda) = |\Lambda| = 0$. Для слова $\alpha = \alpha_1, ..., \alpha_2$ буква α_1 называется началом, или **префиксом**, слова α, а буква α_2 — окончанием, или **постфиксом**, слова α. Слова, как и кортежи вообще, можно соединять. Для этого префикс второго слова должен следовать сразу за постфиксом первого, при этом в новом слове они, естественно, утрачивают свой статус, если только одно из слов не было пустым. Соединение слов α_1 и α_2 обозначается $\alpha_1\alpha_2$, соединение n одинаковых слов α обозначается α^n, причем $\alpha^0 = \Lambda$.

Множество всех непустых слов алфавита A обозначим A^*: $A^* = \{\alpha \mid l(\alpha) > 0\}$. Множество A называют **алфавитом** сообщений, а множество B — **кодирующим алфавитом**. Множество слов, составленных в алфавите B, обозначим B^*.

Обозначим через F отображение слов алфавита A в алфавит B. Тогда слово $\beta = F(\alpha)$ назовем **кодом** слова α (от лат. *codex*). **Кодированием** называют универсальный способ отображения информации при ее хранении, передаче и обработке в виде системы соответствий между элементами сообщений и сигналами, при помощи которых эти элементы можно зафиксировать. Таким образом, код — правило однозначного преобразования (т.е. функция) сообщения из одной символической формы представления (исходного алфавита A) в другую (объектный алфавит B), обычно без каких-либо потерь информации. Процесс преобразования $F: A^* \to B^*$ слов исходного алфавита A в алфавит B называется **кодированием** информации.

Процесс обратного преобразования слова $\beta \in B^*$ в слово $\alpha \in A^*$ называется **декодированием.** Очевидно, для этого должно быть $\beta \in F(A)$. Таким образом, декодирование — функция, обратная F (рис. 6.4).

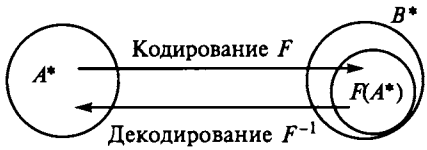


Рис. 6.4. Схема декодирования

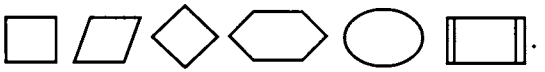
Правила обработки сообщений могут быть различными: поэлементное кодирование, с помощью алгоритма, кодирование с помощью различных

видов графики и др. Так как для любого кодирования должна выполняться операция декодирования, то отображение должно быть обратимым (биекция).

Если все кодовые слова имеют одинаковую длину, то код называется **равномерным**, или **блочным.**

Под **абстрактным** алфавитом будем понимать упорядоченное дискретное множество символов. Например:

- алфавит букв русского языка А, Б, В, ..., Я, мощность которого $n_1 = 33$;
- алфавит арабских цифр 0, 1, 2, ..., 9, $n_2 = 10$;
- алфавит знаков шестигранной кости (кубика) ·, ⁚, ∴, ⁝, ⁘, ⁙, $n_3 = 6$;
- двоичный алфавит $B = \{0, 1\}$, $n_4 = 2$;
- двоичный алфавит {«точка», «тире»}, $n_5 = 2$;
- алфавит римской системы счисления I, V, X, L, C, D, M...;
- алфавит языка блок-схем для изображения алгоритмов:

.

Особенно важное значение имеют алфавиты, содержащие два элемента — *двоичный,* или *бинарный,* набор символов. Это такие пары символов, как цифры {0, 1}, семантические значения «истина» и «ложь» {*и, л*}, пара ответов {да, нет}, пара состояний «включено» и «выключено», знаки азбуки Морзе {·, —} («точка», «тире»).

Алфавитное кодирование. Пусть заданы конечный алфавит $A = \{a_1, ..., a_n\}$ и множество слов над алфавитом A — A^*. Напомним, что для слова $\alpha = \langle a_1 a_2 ... a_n \rangle \in A^*$ количество используемых букв называется длиной слова и обозначается $l(\alpha) = |\alpha| = |a_1 ... a_n| = n$.

Назовем элементарной частью сообщения ту его минимальную часть, которой ставится в соответствие символ из множества B. Будем рассматривать простейший случай, когда элементарной частью сообщения является одна буква алфавита A.

Алфавитное, т.е. побуквенное, кодирование можно задать таблицей кодов. Фактически кодом, кодирующей функцией, будет служить некоторая подстановка σ. Тогда σ: $\langle \alpha_1 \to \sigma(\alpha_1), ..., \alpha_n \to \sigma(\alpha_n) \rangle = \langle \beta_{i_1} ... \beta_{i_n} \rangle$ или $\sigma = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \\ \beta_1 & \dots & \beta_n \end{pmatrix}$, где $\alpha_i \in A$, $\beta_i \in B^*$. Такое побуквенное кодирование обозначается $K^{\alpha_i}_{\beta_i}$. Множество кодов букв $\{\beta_i\}$ называется **множеством элементарных кодов**. Алфавитное кодирование можно использовать для любого множества сообщений. Таким образом, алфавитное кодирование является самым простым, и его всегда можно ввести на непустых алфавитах.

Пусть, например, заданы алфавиты A и B, состоящие соответственно из {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} и {0, 1}. Тогда таблица кодирования может быть подстановкой:

$$\sigma = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0000 & 0001 & 0010 & 0011 & 0100 & 0101 & 0110 & 0111 & 1000 & 1001 \end{pmatrix}.$$

Это *двоично-десятичное кодирование*, оно является взаимно-однозначным и потому допускает декодирование.

Однако схема $\sigma = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 10 & 11 & 100 & 101 & 110 & 111 & 1000 & 1001 \end{pmatrix}$ не является взаимно-однозначной. Например, кортеж из шести единиц 111111 может соответствовать как слову 333, так и 77, 111111, а также 137, 3311 или 7111 плюс любая перестановка.

Кодирование с минимальной избыточностью. *Оптимальным* является кодирование, при котором сообщение имеет наименьшую из возможных длину. Но для выявления наилучшего из вариантов необходима дополнительная информация. Например, для сообщений, представленных на естественном языке, такой дополнительной информацией может быть распределение вероятности появления букв в некотором тексте. Тогда задача построения оптимального кода приобретает точную математическую формулировку и строгое решение.

Пусть задана некоторая схема алфавитного кодирования $\sigma = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \\ \beta_1 & \dots & \beta_n \end{pmatrix}$. Тогда любая схема $\sigma' = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \\ \beta'_1 & \dots & \beta'_n \end{pmatrix}$, где кортеж $\langle \beta'_1 \dots \beta'_n \rangle$ есть перестановка кортежа $\langle \beta_1 \dots \beta_n \rangle$, также будет задавать некое кодирование. В таком случае, если длины элементарных кодов равны, то их перестановка в схеме не влияет на длину закодированного сообщения. В том случае, если длины элементарных кодов различны, то длина кода сообщения напрямую зависит и от того, какие элементарные коды каким буквам поставлены в соответствие, и от того, каков состав букв в сообщении.

Поэтому для определения алгоритма назначения элементарных кодов, дающих минимальную длину фиксированного сообщения S при фиксированной схеме σ, *необходимо:*

отсортировать буквы в порядке убывания количества вхождений;
отсортировать элементарные коды в порядке возрастания длины;
поставить коды в соответствие буквам в установленном порядке.

Назовем схему σ **префиксной,** если элементарный код одной буквы не является префиксом кода другой буквы.

Схема σ называется **разделимой**, если любое слово, составленное из элементарных кодов, разлагается на элементарные коды единственным образом, т.е. $\forall \beta = \langle \beta_1 \beta_2 \dots \beta_n \rangle$, $\beta_i \in B$ $\exists!$ разбиение $\beta = \langle \underbrace{b_1^1 \dots b_1^{n_1}}_{b_1} \dots \underbrace{b_k^1 \dots b_k^{n_k}}_{b_k} \rangle = \langle b_1 b_2 \dots b_k \rangle$, где $b_j \in B^*$, а $b_i^j \in B$, причем $\forall b_j \in B^*$ $\exists!$ $\alpha_j \in A^*$: $\alpha_j = \sigma^{-1}(b_j)$. Полученное слово можно разбить на k подслов длины n_i: $b_1^1 = \beta_1$, ..., $b_1^{n_1} = \beta_{n_1}$, $b_2^1 = \beta_{n_1+1}$, ..., $b_k^{n_k} = \beta_n$,

каждое из которых допускает дешифровку, и это разбиение единственно. Очевидно, $n_1 + ... + n_k = n$.

Алфавитное кодирование с разделимой схемой допускает декодирование. Можно доказать, что префиксная схема является разделимой. Пусть задан алфавит $A = \{a_1, ..., a_n\}$ и вероятности появления букв в сообщении: $P = \{p_1, ..., p_n\}$, где $p_i > 0$ — вероятность появления буквы α_i, причем буквы с нулевой вероятностью появления в сообщении исключены.

Для разделимой схемы $\sigma = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \\ \beta_1 & \dots & \beta_n \end{pmatrix}$ алфавитного кодирования существует так называемая средняя цена, или **длина кодирования**, — это математическое ожидание длины закодированного сообщения, которая обозначается I_σ и определяется как $I_\sigma = \sum_{i=1}^{n} p_i l(\beta_i)$. Также эта величина служит мерой среднего удлинения слова при кодировании. Учитывая, что кодируются буквы, длина которых 1, имеем для изменения средней длины $\Delta I_\sigma =$ $= \sum_{i=1}^{n} p_i (l(\beta_i) - 1) = I_\sigma - \sum_{i=1}^{n} p_i = I_\sigma - 1$.

Рассмотрим пример. Пусть заданы алфавиты $A = \{x, y\}$, $B = \{0, 1\}$ и разделимая схема $\sigma = \begin{pmatrix} x & y \\ 0 & 01 \end{pmatrix}$. При распределении вероятностей (0,5; 0,5) цена кодирования равна $I_\sigma = 0{,}5 \cdot 1 + 0{,}5 \cdot 2 = 1{,}5$, а при распределении вероятностей (0,8; 0,2) соответственно $I_\sigma = 0{,}8 \cdot 1 + 0{,}2 \cdot 2 = 1{,}2$.

6.5. Кодирование информации как средство обеспечения контроля работы автомата

Scientia vinces — наукой победишь.

Латинская формула

Система контроля. Рассмотренные методы кодирования информации обеспечивают надежное кодирование только в том случае, если ЭВМ функционирует без каких-либо нарушений. Если же появились отклонения в работе ЭВМ, повлекшие за собой ошибки при передаче и хранении информации, то пользователь ЭВМ даже не узнает об этом.

Совокупность определенных методов и средств, обеспечивающих необходимое качество работы как всей ЭВМ, так и отдельных ее элементов, а также автоматическое исправление возникающих ошибок называется **системой контроля**.

В задачи системы контроля надежности работы ЭВМ входят:

• профилактический контроль, т. е. предупреждение появления возможных ошибок;

• оперативный контроль — проверка качества выполнения машинных операций.

Однако принципиально важно не только установить наличие ошибки, но и устранить ее с помощью ЭВМ. Поэтому разработаны различные методы кодирования, позволяющие решать поставленные задачи.

Несмотря на то что надежность электронной аппаратуры постоянно возрастает, в работе ЭВМ неизбежны систематические и случайные ошибки. Искажение сигнала за счет помех канала связи может быть вызвано повреждением поверхности магнитного носителя и потерей контакта.

Будем рассматривать канал связи с помехами как модель различных ошибок. Пусть заданы алфавиты $A = B = \{0, 1\}$ и Λ — пустое слово, а кодирование выполняется без дополнительных помех. Возможны три типа ошибок:

• $0 \to 1$, $1 \to 0$ — ошибки замещения разряда;

• $0 \to \Lambda$, $1 \to \Lambda$ — ошибки выпадения разряда (длина слова уменьшается на 1);

• $\Lambda \to 1$, $\Lambda \to 0$ — ошибки вставки разряда (длина слова увеличивается на 1).

Общая характеристика ошибок канала, т. е. их количество и тип, обозначается Q. Так, канал с характеристикой $Q = (1, 1, 0)$ означает, что в канале возможны по одной ошибке замещения и выпадения разряда при передаче сообщения.

В процессе кодирования исходный текст, называемый *открытым*, заменяют зашифрованным. Тогда **дешифрование** — процесс, благодаря которому информационный текст с помощью ключа преобразуется в исходный (рис. 6.5).

Под **ключом** мы понимаем информацию, необходимую для беспрепятственного шифрования и дешифрования текстов.

К основным видам криптосистем относят симметричные и системы с открытым ключом. Симметричные криптосистемы используют единый ключ для шифрования и дешифрования информации.

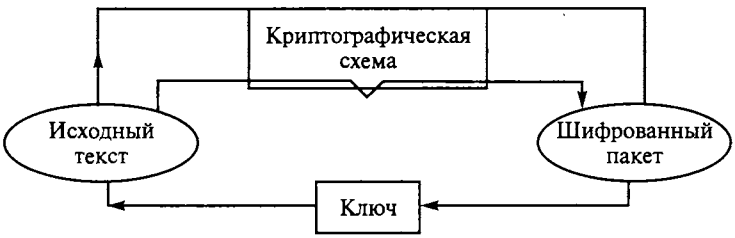


Рис. 6.5. Схема процесса дешифрования

Системы с открытым ключом подразумевают также и наличие закрытого ключа, доступного лишь получателю сообщения.

Криптостойкостью ключа называют его способность противостоять криптоанализу. На современные криптографические системы защиты информации накладываются общепринятые требования, среди которых важнейшими являются следующие:

- зашифрованное сообщение поддается прочтению только с помощью ключа;
- любые изменения ключа влекут за собой значительные изменения зашифрованного сообщения;
- структура алгоритма, информация — постоянны;
- длина зашифрованного текста равна длине исходного;
- алгоритм должен допускать различные виды реализации (программный и (или) аппаратный).

Задачей теории кодирования является проблема наилучшего согласования источника информации и канала связи. Симметричные криптосистемы включают в себя в числе других подстановки, перестановки и блочные шифры. При подстановке происходит замена символов исходного текста буквами того же алфавита. При этом шифруется каждый символ исходного текста.

Самый древний из известных видов шифрования — «подстановки Цезаря» (I в. до н. э.). Этот код назван в честь знаменитого римского императора Гая Юлия Цезаря, поручившего Цицерону составить послания с использованием пятибуквенного алфавита. Подстановка Цезаря заключалась в замене символов (букв) исходного текста алфавитом, смещенным на три буквы вправо. Очевидно, сопоставление алфавитов делает код Цезаря легко поддающимся вскрытию.

Эффективное кодирование. Закодируем сообщение, записанное с помощью алфавита, содержащего n символов, например с помощью русского алфавита, содержащего 33 буквы и знаки препинания. Чтобы закодировать это сообщение, необходимо указать правило, которое устанавливает взаимно-однозначное соответствие между самим сообщением и теми кортежами из m символов, которыми оно записано в этом алфавите. Как это осуществить наиболее эффективно?

Будем называть закодированное сообщение **эффективным**, если для его передачи по некоторому каналу связи использовалось минимальное количество сигналов. Тогда с учетом скорости передачи каждого сигнала на передачу всего сообщения будет затрачено минимальное количество времени. Такой «выгодный» код позволяет получить реальный экономический эффект благодаря уменьшению времени эксплуатации канала связи.

Как известно, возможность эффективного использования лишь двух символов 0 и 1 для кодирования указал еще Фрэнсис Бэкон. Он первым применил в XVI в. принцип двоичного кодирования для маскировки тай-

нописи. Так, Бэкон, передавая сообщения с помощью букв *A* и *B*, закодировал пятерками, состоящими из этих букв, весь латинский алфавит. Как известно, Лейбниц использовал принцип двоичного кодирования не только в своей двоичной арифметике, но также и в так называемой двухзначной логике.

Кодирование информации преследует разные цели и поэтому использует различные методы. В качестве целей могут служить:

- экономичность, т. е. уменьшение избыточности сообщений;
- обеспечение максимальной пропускной способности канала связи;
- надежность, т. е. защита сообщений от случайных искажений, помехоустойчивое кодирование;
- сохранность, т. е. защита информации от постороннего доступа к ней;
- повышение скорости передачи информации;
- удобство физической реализации через соответствующие средства связи;
- удобство восприятия сообщений.

Иногда эти цели вступают в противоречие друг с другом. Так, надежность кода требует избыточности информации, а это неэкономично. Такую избыточную информацию содержат, например, многие финансовые документы, где денежную сумму записывают и числом, и прописью.

Разработаны методы эффективного кодирования информации, позволяющие учитывать одновременно группу целей, решать комплекс задач.

Методы эффективного кодирования информации. Как мы знаем, для выявления ошибок в процессе передачи закодированной информации необходимо ввести информационную избыточность.

Метод Шеннона — Фано. Один из самых эффективных путей кодирования основан на применении основной теоремы Шеннона для каналов без шума. Согласно этой теореме сообщение, составленное из букв некоторого алфавита, можно закодировать так, что среднее число двоичных символов на букву будет сколь угодно близко к *энтропии* источника этих сообщений, но не меньше этой величины. Это означает, что при выборе каждого символа кодовой комбинации необходимо добиваться, чтобы знак нес максимальную информацию, т. е. принятие каждым символом значений 0 и 1 должно быть равновероятным, и этот выбор не должен зависеть от предыдущих значений символов. По имени авторов первых эффективных методов кодирования этот код получил название *метода Шеннона — Фано.*

Для построения кода буквам алфавита ставят соответствие их частоты в порядке убывания вероятности. Затем их разделяют на две группы, уравнивая суммарные вероятности каждой группы. Процесс разделения на подгруппы продолжают до тех пор, пока

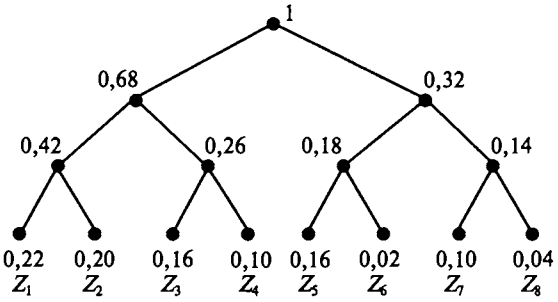


Рис. 6.6. Кодовое дерево для восьми символов с учетом вероятностей

в каждой подгруппе не окажется по одной букве (см. рис. 6.6). Такой способ кодирования не всегда приводит к однозначному построению кода, так как разбиение на группы произвольно. Эту процедуру можно представить в виде двоичного дерева.

Рассмотрим пример. На рис. 6.6 показан алфавит кодирования из восьми символов, который представлен в табл. 6.4. Алфавит можно разместить в таблице без учета статистических характеристик и с учетом вероятности. В этом примере энтропия набора букв $H = 2{,}88 < 3$, но разброс длины кода от 2 до 5 ведет к неоднозначному кодированию.

В решении этой задачи отсутствие помех служит ограничением при реальном кодировании.

Двоичное кодирование непосредственно связано с задачей поиска ответа при помощи уточняющих вопросов (см. подразд. 6.3).

Таблица 6.4

Кодовая таблица

Буквы	Вероятности	Кодовые комбинации	
		без учета вероятности	с учетом вероятности
Z_1	0,22	11	11
Z_2	0,20	101	10
Z_3	0,16	100	011
Z_4	0,10	01	010
Z_5	0,16	001	001
Z_6	0,02	0001	0001
Z_7	0,10	00001	00001
Z_8	0,04	00000	00000

Пусть имеется неизвестное кодовое слово, двоичными символами которого закодировано сообщение в исходном алфавите. При определении загаданного слова задаем уточняющие вопросы. Например, на вопрос: «В кодовом слове первая цифра 1?» — возможный ответ «Да» или «Нет» однозначно определяет первую цифру кодового слова. Аналогично определяются и другие символы этого кодового слова, соответствующего неизвестному сообщению. Поэтому каждый двоичный код в n-буквенном алфавите соответствует определенному методу установления неизвестных символов в этом алфавите с помощью уточняющих вопросов.

Итак, выбор кодов зависит от различных обстоятельств:

- удобство процесса передачи информации;
- удобство восприятия информации;
- необходимость обеспечения максимальной пропускной способности канала связи;
- качественность характеристик кодирования (простота или однозначность кодирования и др.);
- потребность в помехоустойчивом кодировании.

Теория кодирования исследует проблемы оптимального сочетания вопросов экономичности и надежности процесса передачи информации. Так, защита информации от постороннего вмешательства ведет к увеличению ее объема за счет ввода специальных шифров, паролей, ключей и т.д., а увеличение объема замедляет работу с информацией.

Помехоустойчивое кодирование. Кодирование само по себе не гарантирует 100 %-ной защиты при передаче информации. Искажение информации может происходить по различным причинам от помех или шумов, которые могут быть вызваны различными обстоятельствами, такими, как неисправность ЭВМ, отклонение от стандартов напряжения сети или влажности помещения, внешние помехи и др.

Поэтому одной из важнейших проблем теорий кодирования является проблема *обнаружения и исправления ошибок*, возникающих при передаче информации. Различают три вида ошибок, возникающих по различным причинам: погрешности в данных, методические погрешности и неисправности в работе ЭВМ. Ошибки первого и второго вида система контроля не сможет исправить самостоятельно, их надо учесть заранее.

Помехами в канале связи называется потеря информации между входом (X_{inp}) и выходом (X_{out}), связанная с самим каналом. Пусть одним множеством (A) является вход (X_{inp}), его мощность $H(X_{inp})$ — информация на входе, а другим (B) — выход (X_{out}) (см. рис. 1.17, *а*). Тогда их пересечение $X_{inp} \cap X_{out}$ можно отождествить со взаимной информацией $I(X_{inp}, X_{out})$. Чем оно ближе к самим множествам X_{inp} и X_{out}, тем лучше передался сигнал. Один из остатков — множество $X_{inp} \setminus X_{out}$ — отождествляется с потерей информации за

счет канала. Тогда разность $\Delta I_1 = H(X_{inp}) - I(X_{inp}, X_{out})$ и будет мерой потери информации. Оставшийся «полумесяц» — множество $X_{out} \backslash X_{inp}$ — это то, чего не было на входе, но появилось на выходе, т.е. **шум**, а изменение информации, связанное с этим, равно $\Delta I_2 = H(X_{out}) - I(X_{inp}, X_{out})$.

В задачи помехоустойчивого кодирования входит обнаружение и исправление ошибок третьего вида. Эта цель достигается введением *избыточной информации.* Избыточность информации можно получить аппаратными (схемными), логическими и информационными средствами.

Конечно, для обнаружения ошибки можно перепроверить результат другим способом, например применив другой математический метод. Но эта перепроверка сможет только установить наличие ошибки.

Другой способ обнаружения ошибки заключается в удвоении каждого символа. Например, слово «март» можно закодировать в слово «ммаарртт». Тогда, в случае искажения одного из символов, ошибку можно обнаружить по второму, неискаженному символу. Так, получив искаженное слово «ммаарртс», мы должны проверить оба варианта гипотез: «марс» и «март».

При кодировании неизбежно возникает избыточная информация (шум), поэтому ставится задача сведения этой избыточности к минимуму.

Теоремы Шеннона. Существует несколько десятков разнообразных методов распознавания и коррекции ошибок кодирования, предназначенных для каналов связи с различными характеристиками помех.

Выдающимся достижением в теории кодирования являются теоремы, сформулированные и доказанные Клодом Шенноном, о том, что при любых помехах и шумах возможна передача информации с минимумом потерь. Оговоримся, что здесь мы дадим популярную интерпретацию теорем Шеннона, так как их фактическая формулировка оперирует с такими понятиями теории вероятностей, как «условная вероятность», «условное распределение», и выходит далеко за рамки вузовской программы. Строгое доказательство теорем основано на неравенстве Фано.

Согласно первой теореме Шеннона, если канал передачи не содержит собственных помех, то возможно так закодировать сообщение, чтобы среднее число элементов кода (двоичных символов), приходящихся на один элемент кодируемого алфавита, было минимальным. Это так называемое *эффективное статистическое кодирование.*

Первая теорема Шеннона утверждает, что вероятность ошибок при передаче информации сколь угодно мала, если энтропия S множества передаваемых сообщений меньше пропускной способности канала связи C, определяемой как наибольшее число бит,

которые возможно передавать по этому каналу связи за единицу времени.

Для каналов связи с помехами справедлива *вторая теорема Шеннона* (*прямая теорема кодирования*), согласно которой всегда существует способ кодирования, такой, что информация может быть передана с какой угодно высокой достоверностью при большой длине передаваемых слов, если скорость передачи не выше пропускной способности канала связи. Пусть кодом являются N слов $q_1, q_2, \ldots, q_N$ длины n. Пусть $p(q_i)$ — вероятность неправильной передачи сообщения q_i. Тогда величина $\frac{1}{N}\sum_{i=1}^{N} p_i$ будет средней вероятностью ошибки. Минимальную среднюю ошибку среди всевозможных кодов обозначим $p(n, M)$. Тогда скоростью передачи будет $R = \log_2 N / n$. Теорема Шеннона утверждает, что при $R < C$ $p(n, N) \xrightarrow[n\to\infty]{} 0$.

Таким образом, вторая теорема Шеннона доказывает принципиальную возможность *помехоустойчивого кодирования*. Обоснованная Шенноном, схема передачи данных представлена в общем виде на рис. 6.7.

Кодирующее устройство преобразует некоторое сообщение m_1 в код m_2. В канале связи за счет помех (шумов) сообщение может быть искажено: $m_2 \to m_3$. Декодирующее устройство преобразует сообщение m_3 в m_2^1. Затем m_2^1 преобразуется в m_1^1.

При кодировании кодер канала обеспечивает заданную достоверность сообщения в процессе ее передачи и хранения. Однако из-за помех в канале связи при декодировании возможны искажения. Поэтому возможно преобразование сообщения m_3 не в m_2, а в m_2^1 (искаженное m_2), а затем в m_1^1 (искаженное m_1).

Сформулируем три глобальные проблемы кодирования:

- создание шифра кодирования;

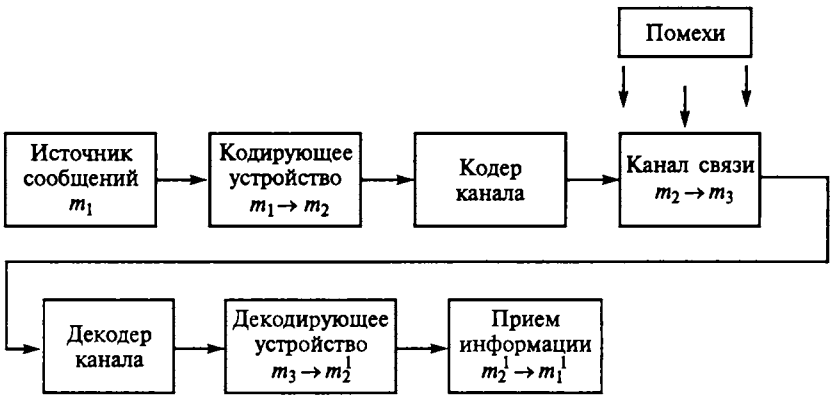


Рис. 6.7. Обобщенная схема передачи данных

• создание специальных корректирующих кодов для поиска и исправления ошибок, возникающих в результате передачи и хранения информации;

• минимизация избыточной информации для успешной коррекции и сокращения потерь в скорости передачи сообщения.

Для повышения помехоустойчивости каналов передачи информации разработаны специальные методы, сокращающие влияние шумов. Несмотря на то что каналы связи весьма разнообразны, а помехи специфичны, существуют общие способы борьбы с ними.

Необходимо максимально снизить уровень помех и максимально повысить уровень полезного сигнала, т. е. увеличить отношение сигнал/шум. Увеличение отношения сигнал/шум может достигаться за счет представления информации в виде импульсов, четко выделяемых на фоне помех.

В целях борьбы с помехами используется фильтрация информации, принимаемой на выходе канала связи.

Корректирующие коды. Борьба с помехами уменьшает, но не исключает вероятность появления ошибок. Поэтому рассмотрим подробнее корректирующие коды.

Корректирующие коды — класс кодов, обладающих свойством обнаруживать с заданной точностью и исправлять возникающие ошибки с целью повышения помехоустойчивости информационных систем при передаче, хранении и обработке дискретной информации.

Согласно теореме Шеннона для каналов связи с помехами, при помощи подходящих кодов можно передавать информацию с любой скоростью, не превосходящей пропускную способность канала связи, так, что вероятность ошибки при декодировании будет сколь угодно мала. Для организации работы корректирующих кодов используется введение принципа избыточности информации, за счет чего основная информация приобретает помехоустойчивые свойства. Но для расположения такой избыточной информации необходимо ввести дополнительные разряды, которые заполняются соответствующим кодом.

Корректирующие коды бывают разных видов. Так, в *блоковых* кодах используют последовательности из n символов, причем кодовые слова составляют лишь часть всех возможных n-символьных последовательностей. Эта часть слов, составляющих код, подбирается соответственно требуемой корректирующей способности.

По принципу образования кодовых слов блоковые коды делятся на разделяемые и неразделяемые. В *разделяемых* кодах кодовые позиции разделяются на информационные, содержащие исходную кодируемую информацию, и проверочные, которые содержат избыточную информацию, необходимую для коррекции ошибок.

В *неразделяемых* кодах избыточность появляется при перекодировании всей исходной информации.

В процессе декодирования принимается решение относительно истинного значения закодированного слова с учетом статистического характера самой гипотезы. Поэтому при конструировании кода очень важен выбор модели канала с помехами. Наиболее распространенными считаются симметричные каналы с равновероятными ошибками различных типов, для которых наиболее полно разработана математическая модель.

Систематический код. Рассмотрим математическую сторону разделяемого *систематического* кода.

Кроме информационных он содержит контрольные разряды, в которые записывается информация о слове, т. е. та самая избыточность, необходимая для диагностики ошибок.

Пусть k — количество контрольных разрядов, m — количество информационных разрядов, n — число разрядов в информационном слове. Представим это на отрезке в виде схемы (рис. 6.8).

Корректирующая способность кода численно равна вероятности обнаружения и исправления ошибок.

Пусть вероятность искажения одного разряда кодового слова равна p. Если искажены все k символов, причем каждый с вероятностью p, то по правилу произведения вероятностью их искажения будет p^k. Тогда остальные $n - k$ символов сохранены с вероятностью $1 - p$, что по правилу произведения даст $(1 - p)^{n-k}$. Так как в кодовом слове учитывается и вероятность искажения, и вероятность сохранения символов, то по правилу произведения их вероятность равна $p^k(1 - p)^{n-k}$. Для определения количества комбинаций применим формулу сочетаний, так как число кодовых комбинаций из всего множества символов n по k искажениям не зависит от порядка. Для вычисления вероятности искажения информации надо учесть число комбинаций сохраненных и искаженных символов в кодовом слове, а также вероятность их сохранения в $n-k$ разрядах и вероятность их искажения в k разрядах. Тогда по правилу произведения

корректирующая способность кода равна $P = \sum_{i=1}^{k} \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}$.

Было установлено, что наибольший вес в сумме вероятностей имеет вероятность искажения одного символа $p \in [10^{-4}; 10^{-3}]$. Это подтверждает тот факт, что легче обнаружить и исправить одну ошибку.

Введем ряд определений. Назовем **весом** кодовой комбинации $V(A)$ количество единиц в кодовой комбинации.

Например, пусть $A = 01001101$, $B = 11011001$. Вес кодовых комбинаций составляет $V(A) = 4$, $V(B) = 5$.

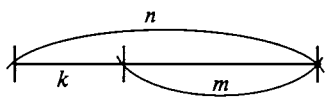


Рис. 6.8. Схема для представления систематического кода

Назовем **кодовым**, или **хемминговым**, **расстоянием** $d(A, B)$ для кодовых ком-

бинаций A и B вес третьей кодовой комбинации, полученной при сложении исходных по модулю два.

В нашем примере $C = A \oplus B = 10101000$, так как $\begin{array}{r} 01001101 \\ +\,01011011 \\ \hline 10101000 \end{array}$.

Тогда $V(C) = d(A, B) = \sum_{i=1}^{8} C_i = 3$.

В теории кодирования известно, что систематический код обнаруживает ошибки, когда минимальное кодовое расстояние $d_{\min} = 2t$, где t — кратность обнаруженных ошибок. Так, при единичных ошибках $t = 1$.

Для кодов, корректирующих независимые ошибки, необходимо учитывать расстояние Хемминга, которое определяет корректирующие возможности кода по отношению к независимым ошибкам.

Пусть код содержит t информационных разрядов и s контрольных разрядов. После проведения s проверок обнаруженной ошибке поставим в соответствие число 1, а отсутствию ошибки — число 0. При этом полученный кортеж двоичных символов является контрольным числом, указывающим номер позиции, где произошла ошибка. Поэтому кортежу из одних нулей соответствует безошибочная передача сообщения. Итак, число $(t + s + 1)$ соответствует всем возможным событиям, поэтому для того, чтобы код обнаружил все комбинации из t ошибок, необходимо и достаточно, чтобы кодовое расстояние было $(t + s + 1)$.

Код Хемминга обозначают (n, k), где n — длина кодового слова; k — число информационных позиций.

Итак, любое слово можно записать в виде кортежа из 0 и 1. Так как в русском языке 33 буквы, а нам необходимо закодировать все, то для кодирования понадобится $2^5 = 32$ разряда. Поскольку $32 < 33$, примем за один символ пары Е и Ё и И и Й. Тогда, имея в запасе пробел, сможем записать в виде кортежей 0 и 1 не только слова, но и предложения.

В двоичных словах помехой является замена 0 на 1 или 1 на 0, т.е. инверсия (отрицание). Учитывая тождество $1 \oplus x_i = \overline{x_i}$, видим, что математически помеха описывается сложением (в смысле операции М2) 1 с исходным сообщением. Отсюда видна важная роль суммы по модулю два. Обозначим передаваемый символ x, полученный символ — y. Тогда, рассматривая $x \oplus y$, получим:

$0 \oplus 0 = 0$ — передавали 0, получили 0, ошибки нет;

$0 \oplus 1 = 1$ — передавали 0, получили 1, ошибка есть;

$1 \oplus 0 = 1$ — передавали 1, получили 0, ошибка есть;

$1 \oplus 1 = 0$ — передавали 1, получили 1, ошибки нет.

Таким образом, наличие ошибки в данной позиции означает получение 1 в результате суммы по модулю два исходного и полу-

ченного сигналов (т.е. код ошибки) или прибавление 1 к исходному. Обозначим код ошибки E. Тогда $x \oplus y = E$. Прибавив ($\oplus$) к обеим частям равенства x, получим $y = x \oplus E$, что уже было описано. Прибавим к последнему равенству E и получим $y \oplus E = (x \oplus E) \oplus E = x$. Видим, что если выходное сообщение еще раз пропустить через тот же самый канал связи, то восстановится исходное.

В технике механизм действия помех на кодовые слова из 0 и 1 практически не связан с самим передаваемым сигналом, а связан в основном с внешними воздействиями на передающую технику. Например, при передаче сообщения произошла ошибка в определенном разряде. Тогда при передаче другого сообщения с большой вероятностью помеха произойдет в том же разряде, что и в первый раз.

Пусть передавали двоичное слово 101001, а получили 110001, т.е. ошибка произошла на втором и третьем разряде. Если такая ошибка произойдет повторно, то восстановится исходное сообщение, поскольку двойная инверсия дает исходный аргумент: $\overline{\overline{x}} = x$. Действительно,

в первый раз:			*во второй раз*:	
передавали	101001		передавали	110001
ошибка	011000	;	ошибка	011000
получили	110001		получили	101001

Напомним, что здесь использовалась операция М2 в каждом разряде. Так как совпали первое слово при передаче в первом сообщении и последнее слово при передаче во втором сообщении, то первоначальное сообщение восстановлено: две одинаковые ошибки взаимно уничтожили друг друга. Поэтому для коррекции нужно полученное сообщение еще раз прогнать через канал связи.

Эти идеи развил Вернан для построения *абсолютно надежного шифра* — одноразового блокнота. Пусть сообщение П зашифровано в виде кортежа из 0 и 1. Нужно придумать случайный ключ ошибки К и сложить его с помощью операции М2 с начальным сообщением П.

Полученное сообщение Ш вторично сложить с ключом ошибки К. Тогда в первом действии П $\oplus$ К = Ш, во втором действии Ш $\oplus$ К = П $\oplus$ К $\oplus$ К = П $\oplus$ 0 = П. Две взаимно-обратные ошибки уничтожили друг друга, и передаваемое письмо П пришло к нам без изменений. Клод Шеннон доказал, что такой шифр является абсолютно надежным.

Например, пусть наш алфавит содержит лишь буквы *а*, *м*, *р*, *т*. Поставим в соответствие *а* — 00, *м* — 01, *р* — 10, *т* — 11.

Пусть передается буква *М* (П = 01) с ошибкой (ключ ошибки К = 11). Тогда принимаемое сообщение имеет вид Ш = П $\oplus$ К = = 01 $\oplus$ 11 = 10, а число 10 соответствует букве *р*. Угадаем передан-

ную букву, если ключ ошибки не известен. Тогда необходимо проверить все возможные равновероятные варианты:

если К = 00, то П = Ш ⊕ К = 10 ⊕ 00 = 10, т.е. передавали *р*;

если К = 01, то 10 ⊕ 01 = 11, передавали *т*;

если К = 10, то 10 ⊕ 10 = 00, передавали *а*;

если К = 11, то 10 ⊕ 11 = 01, передавали *м*.

Можно сформулировать основные требования к ключу:

- длина ключа равна длине кодового слова;
- полная равновозможность ключей;
- для нового слова создается свой кодовый ключ.

Но как передать отправителю *А* для получателя *В* информацию о секретном ключе ошибок К? Для этого можно использовать достижения новой науки квантовой криптографии. Квантовая криптография берет за основу действие единичных фотонов.

Как известно, свет состоит из набора частиц — фотонов. Ученые создали приборы, генерирующие единичные фотоны, которые условно назовем *зеленый нуль*, *зеленая единица*, *красный нуль*, *красная единица*. Эти единичные фотоны обладают особыми свойствами:

- при воздействии зеленого прибора на зеленый фотон последний не изменяется, что дает возможность установить его характер — зеленый нуль или зеленая единица;
- при воздействии на красный нуль зеленым прибором он с одинаковой вероятностью может превратиться в зеленый нуль и в зеленую единицу.

Будет известно, каким стал фотон после воздействия, но мы не узнаем о том, что произошло превращение.

Пусть отправитель *А* случайным образом посылает один из четырех типов фотона получателю *В*. Получатель *В* случайным образом определяет цвет прибора, который будет воздействовать на фотон. Пусть опыт проводился, например, 1024 раза. Примерно половину фотонов получатель *В* испортит, посмотрев на них с помощью прибора другого цвета. Пусть отправитель *А* сообщит получателю *В* последовательность цветов, сохранив в тайне их численные значения. Тогда *В* сможет определить, какие фотоны приняты верно, а в каких допущена ошибка. Приблизительно 512 фотонов, о которых известны и цвет, и численные значения, необходимо проверить и установить, не превратил ли посторонний агент красный цвет в зеленый и наоборот. Для установления этого отправитель *А* сообщает точные численные значения этих 512 фотонов с тем, чтобы получатель сравнил их и сопоставил с имеющимися. Если все фотоны совпали, то остальные 256 можно использовать в качестве ключа, в противном случае ясно, что корреспонденцию вскрывали.

Этот способ был реализован на практике в 1989 г. в виде квантовых каналов связи, но ввиду того что он очень дорогой, им

пользуются лишь в крайних случаях, например руководители государств. Более подробно современные методы шифрования рассмотрены в статье И. Б. Писаренко «Что такое абсолютно надежный шифр?».

Контроль по четности. Рассмотрим более распространенные методы помехоустойчивого кодирования.

Несмотря на простоту в использовании двоичных кодов, они обладают серьезным недостатком за счет значительной разницы в наборах цифр между двумя соседними значениями. Так, система управления углом поворота вала или линейных перемещений некоторого объекта соответствует двум соседним кодовым состояниям датчика положений, имеющего определенные выходные двоичные сигналы. Если они выражены последовательными числами 7 и 8 в десятичной системе исчисления, то соответствующие им двоичные кодовые комбинации 0111 и 1000 отличаются во всех четырех позициях. Возможность ошибки увеличивается, и значительно усложняется процесс установления первоначального значения кодового слова. В результате система будет подавать неверный управляющий сигнал.

Для уменьшения вероятности ошибки в двоичных кодах используют код Грея, в котором каждые две позиции отличаются только одним разрядом, т. е. на 1 бит. Поэтому выходной сигнал может быть представлен лишь одним из двух состояний — истинным или ложным.

Для обнаружения однократной ошибки и ее исправления применяется **контроль по четности**, который заключается в том, что сумма двоичных единиц в машинном слове, включая контрольный разряд, должна иметь определенную четность, т. е. быть либо всегда четной, либо всегда нечетной.

Пусть необходимо сохранить информацию, содержащую 4 бит, например 1011. Расположим значения битов машинного слова во внутреннюю область кругов Эйлера так, чтобы они располагались в пересекающихся областях соответственно с номерами 1, 2, 3, 4 (рис. 6.9, *а*). Оставшиеся внутренние области с номерами 5, 6, 7 заполняются «битами четности» так, чтобы сумма единиц в каждом круге была, например, четной (рис. 6.9, *б*). Тогда области с номерами 5, 6, 7 приобретут соответственно числа 0, 1, 0 (рис. 6.9, *в*). При заполнении этих областей необходимо помнить, что сумма всех единиц внутри каждого круга должна быть четной. Тогда изменение одного разряда кодового слова, например 1010 (рис. 6.9, *г*), можно обнаружить, проверив на четность каждый из кругов Эйлера. В данном случае проверка показала, что изменение четности произошло в кругах *B* и *C*, на пересечении которых находится область 4 (рис. 6.9, *д*). Значит, ошибка произошла в четвертом разряде. Для исправления достаточно изменить ее значение на противоположное. Каждая ошибка однозначно определяет

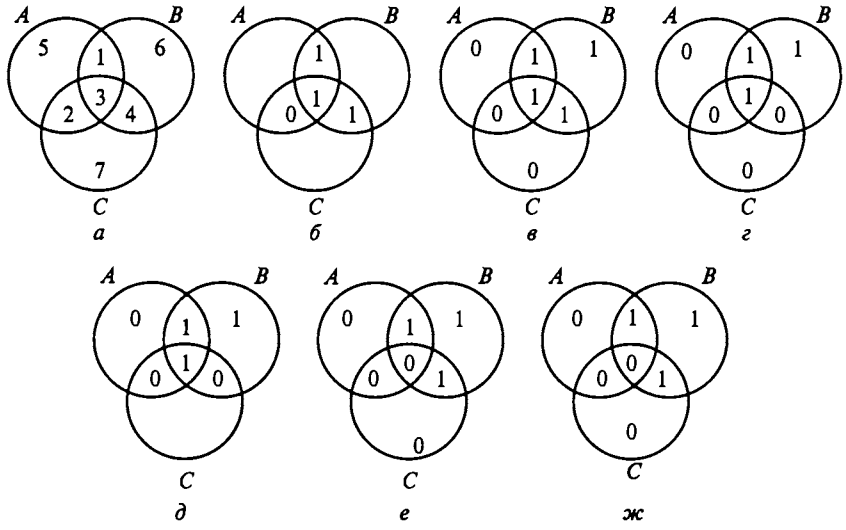


Рис. 6.9. Контроль по четности:

а — в — представление исходной информации в виде кругов Эйлера; *г* — изменение исходной информации в 4-м разряде; *д, ж* — проверка на четность; *е* — изменение исходной информации в 3-м разряде

область на кругах Эйлера и, следовательно, возможность ее устранения. Пусть, например, ошибка произошла в третьем разряде, т.е. получено кодовое слово 1001 (рис. 6.9, *е*). Проверка кругов Эйлера на четность показывает, что нечетны все три круга *A*, *B* и *C*. Но на пересечении всех трех кругов лежит единственная область 3 (рис. 6.9, *ж*), что и соответствует фактической ошибке.

В случае возникновения большого числа ошибок исправление одной из них может повлечь за собой появление новых. Для того чтобы уменьшить такие последствия, вводится еще одна — внешняя область, которая также заполняется битом четности. Значение этого бита выбирается так, чтобы полное число единиц на диаграмме было четным. Код Хемминга в таком случае дает возможность лишь обнаружить, но не устранить ошибки. Код Хемминга и его модификации легко реализуются в ЭВМ.

Поиски надежных шифров, как это ни странно, привели к использованию одного из известнейших и глубоко изученных вопросов математики — делению с остатком.

Этот раздел математики и теории чисел был детально разработан много лет назад применительно к другим проблемам. В настоящее время он широко используется при разработке новых методов кодирования сообщений.

Далее мы даем подробное изложение этого вопроса в связи с его актуальностью в новых криптографических исследованиях.

6.6. Основы алгебры вычетов и их приложение к простейшим криптографическим шифрам

Все проблемы, в конечном счете,
являются научными проблемами.

Б. Шоу

? Подумайте, может ли квадрат некоторого целого числа равняться 1 278 075?

Квадрат какого числа может оканчиваться цифрами 75?

Деление с остатком. *Разделить натуральное число a на натуральное число b с остатком означает записать его в виде $a = qb + r$,* где q и r — неотрицательные числа, причем $r < b$. При этом q называется частным, а r — остатком от деления a на b. Такое деление выполняется уголком.

Например, $187 : 12 = 15$(ост. 7).

В такой записи сразу видно, что для делимого 187 и делителя 12 частным является число 15, а остаток от деления равен 7. Тогда $187 = 15 \cdot 12 + 7$, т.е. $a = 187$, $b = 12$, $q = 15$, $r = 7$. Это представление не требует, чтобы выполнялось условие $a > b$. Если $a < b$, то $a = 0 \cdot b + r$, т.е. $q = 0$, $r = a$. Например, разделить 4 на 9 — значит представить его в виде $4 = 0 \cdot 9 + 4$, где $a = r = 4$, $b = 9$, $q = 0$.

Можно обобщить понятие «деление с остатком» на целые числа: разделить целое a на натуральное b с остатком — значит представить a в виде $a = qb + r$, где $q \in \mathbb{Z}$, а $0 \le r < b$. Отметим, что это целиком соответствует операции «взятие целой части». Тогда, если $a = qb + r$, то $\left[\frac{a}{b}\right] = q$. Остаток r от деления a на b с остатком будем записывать аналитически: $r = a \bmod b$.

Например, для $a = -12$, $b = 8$ имеем $-12 = (-2) \cdot 8 + 4$, для $a = -324$, $b = -15$ имеем $-324 = 22 \cdot (-15) + 6$.

Число a делится нацело на b, если остаток $r = 0$: $a = qb$.

Следствие. Если a и b делятся на c, то при любых целых k и l сумма $ka + lb$ делится на c: т.е. $(a \vdots c,\ b \vdots c,\ \forall k, l, a, b \in \mathbb{Z}) \Rightarrow ((ka + lb) \vdots c)$. Действительно, если $a = q_1 c$, $b = q_2 c$, тогда $ka + lb = kq_1 c + lq_2 c = c(kq_1 + lq_2)$.

Задача 40. Найти такое число $d \neq 1$, на которое делятся числа $6n + 5$ и $9n + 2$. Найти n.

Решение. Если сами числа $6n + 5$ и $9n + 2$ делятся на d, то на d делится и разность $3(6n + 5) - 2(9n + 2) = 15 - 4 = 11$ — простое число, т.е. $d = 11$. Итак, $6n + 5 = 11k$. Введем $m = n - 1$, $l = k - 1$, тогда $6m = 11l$. Поскольку 6 и 11 — взаимно простые натуральные числа, общее решение: $m = 11p$, $l = 6p$. Подставляя, получим $n = m + 1 = 11p + 1$. Так как 9 и 11 так же взаимно просты, аналогично

получим то же решение для уравнения $(9n + 2) \vdots 11$. Итак, $n = 11p + 1$ для любого целого p: $n = ...-21, -10, 1, 12, 23...$

Сравнение по модулю. Рассмотрим множество натуральных чисел.

? На какую цифру оканчивается число 2^{2003}?

Как известно, натуральные степени числа 2 оканчиваются цифрами {2, 4, 8, 6}. Действительно: 2, 4, 8, 16, 32, 64 и т.д. Определим функцию f: $\mathbb{N} \to \{2, 4, 8, 6\}$, которая ставит в соответствие каждому натуральному числу n последнюю цифру числа 2^n. Эта функция периодична с периодом 4, т.е. $f(n) = f(n + 4) = f(n + 4k)$. И обратно: $f(n) = f(n - 4) = f(n - 4k)$. Для любого n нужно найти минимальное натуральное m, такое, что $f(m) = f(m + 4k) = f(n)$. Но это задача на остаток: k — частное деления n на 4, m — остаток. Поэтому последняя цифра числа 2^n зависит от остатка деления показателя на 4: $f(n) = \begin{cases} f(n \bmod 4), & \text{если } n \bmod 4 \neq 0; \\ f(4) = 6, & \text{если } n \bmod 4 = 0 \quad (n \vdots 4). \end{cases}$

В таком случае, так как $2003 = 500 \cdot 4 + 3$, то $f(2003) = f(3) = 8$, и число 2^{2003} оканчивается цифрой 8. Таким образом, множество всех возможных показателей состоит из четырех подмножеств $\forall n \in \mathbb{N}$: $4k$, $4k + 1$, $4k + 2$, $4k + 3$.

Тогда $\forall n \in \mathbb{N}$ все целые числа разбиваются на p подмножеств, которые соответствуют числу, полученному в остатке при делении на p. Такие подмножества, зависящие от остатка деления на p, назовем:

0 если $a \bmod p = 0$, т.е. $a \vdots p$;
1 если $a \bmod p = 1$, т.е. $(a - 1) \vdots p$ (или $\exists k$: $a = kp + 1$);
2 если $a \bmod p = 2$, т.е. $(a - 2) \vdots p$ (или $\exists k$: $a = kp + 2$);
... .
$p - 1$ если $a \bmod p = p - 1$, т.е. $(a + 1) \vdots p$ (или $\exists k$: $a = kp + (p - 1)$).

Очевидно, что любое число $a \in \mathbb{Z}$ принадлежит одному из этих p подмножеств. Причем разность любых двух чисел одного подмножества делится на p, а разность чисел из разных множеств не должна делиться на p.

Два целых числа, разность которых кратна натуральному числу $p \geq 2$ $(a - b \vdots p)$, называются **сравнимыми по модулю** p. Обозначение $a \equiv b \pmod p$, если $\exists k \in \mathbb{Z}$: $a - b = kp$.

Это значит, что числа a и b сравнимы по модулю p тогда и только тогда, когда они принадлежат одному подмножеству, т.е. дают одинаковые остатки при делении на p. Например, $36 \equiv 16 \pmod{10}$; $-24 \equiv 6 \pmod{30}$.

Отметим, что в записи $a \equiv b \pmod p$ выражение $b \pmod p$ означает не «остаток от деления b на p, который сравним с a», а если расставить логические скобки, то $a \equiv b \pmod p$ следует понимать

как $(a \equiv b)(\mathrm{mod}\, p)$. В том случае, когда перед $b\,(\mathrm{mod}\, p)$ стоит знак = или знак арифметической операции, то $b\,(\mathrm{mod}\, p)$ означает остаток. Для доказательства неэквивалентности выражений $a \equiv b\,(\mathrm{mod}\, p)$ (сравнение) и $a = b\,(\mathrm{mod}\, p)$ (равенство) заметим, что в сравнении $a \equiv b\,(\mathrm{mod}\, p)$ a и b симметричны. Если их поменять местами, то сравнение $b \equiv a(\mathrm{mod}\, p)$ будет иметь место и будет полностью эквивалентно исходному: $a \equiv b\,(\mathrm{mod}\, p) \Leftrightarrow b \equiv a(\mathrm{mod}\, p)$. Этого свойства нет в выражении $a = b\,(\mathrm{mod}\, p)$. Например, $4 = 12(\mathrm{mod}\, 8)$, $12 \neq 4(\mathrm{mod}\, 8)$.

Отношение сравнимости $\equiv$ рефлективно, симметрично и транзитивно, следовательно, является *отношением эквивалентности.*

Отдельные классы эквивалентности для отношения сравнимости (по модулю p) называют **вычетами по модулю** p и обозначают Z_p, а раздел математики, изучающий вычеты по модулю, — *алгеброй вычетов*, *теорией вычетов* или *модулярной арифметикой.*

Свойства сравнений.

1. Два числа, сравнимые с третьим по одному модулю, сравнимы между собой: $(a \equiv b\,(\mathrm{mod}\, p),\ b \equiv c(\mathrm{mod}\, p)) \Rightarrow a \equiv c(\mathrm{mod}\, p)$.

2. Сравнения можно почленно складывать и вычитать: $(a \equiv b(\mathrm{mod}\, p);\ c \equiv d(\mathrm{mod}\, p)) \Rightarrow (a \pm c) \equiv (b \pm d)(\mathrm{mod}\, p)$.

Действительно, если $a \equiv b\,(\mathrm{mod}\, p)$ и $c \equiv d(\mathrm{mod}\, p)$, то по определению $(a-b) \vdots p$ и $(c-d) \vdots p$. Тогда разность $(a+c)-(b+d) = (a-b) + (c-d)$ равна сумме двух слагаемых, каждое из которых делится на p. Значит, $(a \pm c) \equiv (b \pm d)(\mathrm{mod}\, p)$.

Поэтому слагаемые можно переносить из одной части сравнения в другую с противоположным знаком.

3. Сравнения можно почленно перемножать: $(a \equiv b(\mathrm{mod}\, p),\ c \equiv d(\mathrm{mod}\, p)) \Rightarrow (ac \equiv bd(\mathrm{mod}\, p))$.

4. Обе части сравнения можно умножить на одно и то же число $k \in \mathbb{Z}$:

$(a \equiv b(\mathrm{mod}\, p)) \Rightarrow (ak \equiv bk(\mathrm{mod}\, p))$. Действительно, $ak - bk = (a-b)k \vdots p$.

5. Обе части сравнения и модуль можно умножить на одно и то же число: $(a \equiv b\,(\mathrm{mod}\, p)) \Rightarrow (ak \equiv bk(\mathrm{mod}\, pk))$.

Действительно, если $a \equiv b\,(\mathrm{mod}\, p)$, то $\exists k$: $a = b + kp$, умножим на число m, тогда $am = bm + kmp$, т.е. $am \equiv bm\,(\mathrm{mod}\, pm)$.

6. Обе части сравнения можно возвести в степень (следствие свойства 3): $(a \equiv b\,(\mathrm{mod}\, p)) \Rightarrow (a^n \equiv b^n(\mathrm{mod}\, p))$.

Впервые понятие сравнения ввел К. Ф. Гаусс в работе «Арифметические исследования» (1802). Алгебра вычетов возникает в тех случаях, когда рассматриваются некоторые циклически повторяющиеся события, например время в течение дня, повторяющееся каждые 24 часа, углы по окружности, повторяющиеся через период 2π, и т.д.

Алгебра вычетов — один из тех удивительных разделов математики, которые рождались в мыслях ученых как некоторые фор-

мальные рассуждения и только спустя годы нашли свое практическое применение.

Контроль по модулю. Методы, основанные на свойствах сравнений, дают возможность осуществлять операцию контроля кода. Различают числовой и цифровой методы получения контрольного кода.

Сущность *числового* метода заключается в том, что код заданного числа определяется как остаток деления числа на выбранный модуль p: $r_A = A \bmod p = A - [A/p] \cdot p$, где в скобках заключена целая часть от деления числа; A — контролируемое число.

Качество контроля во многом зависит от величины модуля p. Если $p = q$, где q — основание системы счисления, в которой выражено число, то контроль осуществляется только над младшим разрядом числа. Если $p = q^m$, то контролируются все разряды числа, но не фиксируются ошибки в m старших разрядах.

Применение метода числового контроля требует получения остатка с помощью операции деления, следовательно, сопровождается большими затратами машинного времени.

Остаток от деления числа на модуль p сравним с самим числом по модулю p, т.е. если $r_A = A \bmod p$, то $A \equiv r_A (\bmod p)$.

Тогда $A + B \equiv (r_A + r_B)(\bmod p)$ и $r_{A+B} = (r_A + r_B)(\bmod p)$.

Аналогично $r_{A-B} = (r_A - r_B)(\bmod p)$ и $r_{AB} = r_A r_B (\bmod p)$.

Задача 41. Пусть заданы числа $A = 312$ и $B = 98$. Определить контрольные коды этих чисел, их суммы и разности, если модуль $p = 15$.

Решение. Контрольные коды чисел найдем по формуле $r_A = A \bmod p$. Тогда $r_A = 312 \bmod 15 = 12$, $r_B = 98 \bmod 15 = 8$.

Аналогично найдем контрольные коды суммы и разности чисел A и B:

для $A + B = 312 + 98 = 410$ $r_{A+B} = 410 \bmod 15 = 5 = (r_A + r_B) \bmod 15$;

для $A - B = 312 - 98 = 214$ $r_{A-B} = 214 \bmod 15 = 4 = (r_A - r_B) \bmod 15$.

Цифровой метод контроля. Пусть натуральное число A задано в некоторой ПСС: $A = \langle a_1, a_2, \ldots, a_n \rangle$. Для получения контрольного кода при цифровом методе контроля необходимо разделить сумму цифр числа на модуль p, т.е. $r' = \left(\sum_i a_i \right) (\bmod p)$, где a_i — цифры кодируемого числа.

Например, найдем контрольные коды тех же чисел $A = 312$ и $B = 98$ цифровым методом.

Найдем суммы цифр чисел A и B:

$$\sum_i a_i = 3 + 1 + 2 = 6; \quad \sum_i b_i = 9 + 8 = 17.$$

Найдем контрольные коды этих чисел:

$r'_A = 6 \bmod 15 = 6$, $r'_B = 17 \bmod 15 = 2$.

Найдем суммы цифр для суммы и разности чисел A и B:

$C = A + B = 410$, $\sum_i c_i = 4 + 1 = 0 = 5$;

$D = A - B = 214$, $\sum_i d_i = 2 + 1 + 4 = 7$.

Найдем контрольные коды суммы и разности чисел C и D:

$r'_C = 5 \bmod 15 = 5$; $r'_D = 7 \bmod 15 = 7$.

Сравним результаты: $(r'_A + r'_B) \bmod 15 = 8$, $r'_{A+B} = 5$, $(r'_A - r'_B) \bmod 15 = 4$, $r'_{A+B} = 7$.

Видим, что ответы, полученные различными способами, не совпали. Но результат контроля должен быть определен однозначно.

? В чем причина плохого результата?

Оказывается, цифровой метод контроля не всегда дает точный результат. Это связано с тем, что нарушаются свойства сравнений из-за переносов единиц из одного разряда на другой при выполнении арифметических действий над числами. Во-вторых, сумма цифр зависит от системы счисления, и поэтому необходимо корректировать контрольный код.

Так, пусть заданы числа A и B и их контрольные коды $r'_A \equiv \sum_i a_i (\bmod p)$, $r'_B \equiv \sum_i b_i (\bmod p)$. Найдем контрольные коды $r_C = r_{A+B}$, где $C = A + B$.

Зная цифры a_i и b_i и количество переносов в каждом разряде, можно найти сумму цифр c_i. В результате каждого переноса q единиц (например, $q = 100$ при заеме из второго разряда в третий операции сложения столбиком) уходит в следующий разряд и одновременно добавляется одна единица из младшего разряда. Поэтому сумма цифр уменьшается на $q - 1$ при каждом переносе.

Тогда $(\sum c_i)' = \sum a_i + \sum b_i - l(q-1)$, где l — количество переносов в больший разряд, возникших в процессе сложения $\sum a_i + \sum b_i$. Подставим это выражение вместо суммы цифр. Имеем $r''_C = (r'_A + r'_B - l(q-1)) (\bmod p)$.

Для разности аналогичная формула имеет вид $r''_D = (r'_A - r'_B + S(q-1)) (\bmod p)$, где S — количество заемов. Очевидно, что r''_C и r''_D являются требуемыми скорректированными кодами для r'_{A+B} и r'_{A-B} соответственно. Если появляется отрицательный остаток, то к сравнению надо добавить модуль p столько раз, сколько потребуется для получения первого положительного остатка.

Задача 42. В предыдущем примере контрольные коды не совпали. Проверим их по формулам для случая подсчета контрольного кода с помощью суммирования цифр по модулю.

Решение. Имеем $r'_A = 6$, $r'_B = 2$. Расчет ведем по формуле $r'_{A+B} = (r'_A + r'_B - l(q-1) (\bmod p)$. Так как $\sum a_i + \sum b_i = 23$, то имеем $q = 10$,

$l = [23/10] = 2$, $r'_{A+B} = (6 + 2 - 2 \cdot 9)(\mathrm{mod}\,15) = -10(\mathrm{mod}\,15) = 5$. Аналогично для разности: $r'_{A-B} = (r'_A - r'_B + S(q - 1)(\mathrm{mod}\,p)$, $\sum a_i - \sum b_i = -11$, $q = 10$, $S = -[-11/10] = 2$. Получим $r'_{A+B} = (6 - 2 + 2 \cdot 9)(\mathrm{mod}\,15) = -22(\mathrm{mod}\,15) = 7$.

Полученные таким образом результаты для r'_{A+B} и r'_{A-B} полностью совпадают с вычисленным ранее, т.е. с помощью новых формул ошибки цифрового метода удалось избежать.

Выбор модуля для контроля. Числовой метод контроля имеет весомое преимущество над иными методами. Он использует свойства сравнений, имеющие достоверный, а не вероятностный характер. Однозначность полученных ответов облегчает осуществление контроля выполнения арифметических операций, сокращает затраты времени.

Для того чтобы эти достоинства числового метода были использованы в цифровом коде, необходимо выполнить условие $r_A = r'_A$: т.е. если $r_A = A(\mathrm{mod}\,p)$; $r'_A = \sum a_i(\mathrm{mod}\,p)$, то $\sum_{i=1}^{n} a_i q^i \equiv \equiv \sum_{i=1}^{n} a_i(\mathrm{mod}\,p)$.

Это требование выполняется, в частности, если или $a_i q^i \equiv a_i(\mathrm{mod}\,p)$, или $q^i \equiv 1(\mathrm{mod}\,p)$.

Но $q^i \equiv 1(\mathrm{mod}\,p)$ означает, что $\exists m \in \mathbb{Z}$: $q^i = mp + 1$ или $p = \dfrac{q^i - 1}{m}$.

Замечая, что свойство должно выполняться для любого $i \geq 1$, в том числе и для $i = 1$, а также вспоминая тождество $q^i = 1 = (q - 1)(q^{i-1} + q^{i-2} + \ldots + 1)$, нетрудно догадаться, что $p = q - 1$ является наилучшим кандидатом на роль модуля при данном основании системы счисления.

Для выбора системы счисления необходимо учесть требования, накладываемые на величину модуля p:

- величина модуля p должна быть небольшой, так как рост числа контролирующих операций усложняет процесс;
- при появлении любой арифметической или логической ошибки изменять сравнимость контрольных кодов;
- получение контрольного кода осуществлять предельно упрощенными средствами.

Компромиссным вариантом для выбора системы счисления служат системы с основанием $q = 2^S$. Так, восьмеричная (2^3) и шестнадцатеричная (2^4) системы счисления легко переводятся в двоичную. Для того чтобы осуществить переход из двоичной системы счисления в $q = 2^S$, необходимо разбить двоичное слово справа на кортежи длины S, а затем суммировать результат по модулю $p = 2^S - 1$.

Таким образом, при $S = 2$ вся информация разбивается на пары — диады, при $S = 3$ — на триады, при $S = 4$ — на тетрады и т.д.

Процесс разбиения кодовой информации на кортежи длины S и получения контрольного кода называется *свертыванием.*

Такие свертки или свернутые коды получаются в процессе суммирования образовавшихся кортежей длины S по модулю p.

Частные случаи образования сверток:

1. Контроль по модулю 3, т.е. $p = 3$ (при $S = 2$).
2. Контроль по модулю 7, т.е. $p = 7$ (при $S = 3$).

Цифровая подпись. Для организации многосторонней секретной связи используется шифр с открытым ключом. Кодирование сообщения A заключается в преобразовании $F: A \to A^d (\mathrm{mod}\, p)$, где пара (d, p) называется **ключом**. Получатель сигнала декодирует его таким же преобразованием с помощью ключа (l, p). Очевидно, что получатель принципиально сможет получить исходное A только если $A < p$, поэтому если надо закодировать много информации (большое слово), его надо разбить на кортежи длиной, меньшей p.

Очевидно, операции кодирования и декодирования информации, по сути, тождественны и отличаются друг от друга лишь показателями степени, поэтому для них выполняется переместительный закон:

$$A \to (A^l)^d (\mathrm{mod}\, p) = A^{ld} (\mathrm{mod}\, p) = A^{dl} (\mathrm{mod}\, p) = (A^d)^l (\mathrm{mod}\, p) \leftarrow A.$$

В практике кодирования используются различные приемы, объединенные названием *цифровая* или *электронная подпись.*

Отправитель кодирует сообщение A закрытым ключом $C = A^d (\mathrm{mod}\, p)$ и посылает получателю информацию, т.е. пару (d, p) в виде подписанного сообщения. Получатель, получив это сообщение, декодирует подпись сообщения открытым ключом (l, p), т.е. находит $A' = C^l (\mathrm{mod}\, p)$.

Если $A = A'$, то письмо дошло правильно и без помех или оно было отправлено в нешифрованном виде. Если $A \neq A'$, то сообщение при передаче было искажено, т.е. произошла потеря информации. В теории вычетов доказывается, что при отсутствии помех и выполнения некоторых условий (взаимной простоты чисел d, l, p) результат $A = A'$ достигается всегда.

Рассмотрим пример. Зашифруем сообщение $A = \langle 4, 3, 2 \rangle$ ключом (5, 7) и пошлем его получателю: $4 \to 4^5 (\mathrm{mod}\, 7) = 1024 (\mathrm{mod}\, 7) = 2$,

$3 \to 3^5 (\mathrm{mod}\, 7) = 243 (\mathrm{mod}\, 7) = 5$,

$2 \to 2^5 (\mathrm{mod}\, 7) = 32 (\mathrm{mod}\, 7) = 4$. Итого $C = F(A) = \langle 2,5,4 \rangle$.

Получатель без труда раскодирует его ключом (11, 7):

$2 \to 2^{11} (\mathrm{mod}\, 7) = 2048 (\mathrm{mod}\, 7) = 4$,

$5 \to 5^{11} (\mathrm{mod}\, 7) = 48\,828\,125 (\mathrm{mod}\, 7) = 3$,

$4 \to 4^{11} (\mathrm{mod}\, 7) = 4\,194\,304 (\mathrm{mod}\, 7) = 2$. Получили $A' = F^{-1}(C) = \langle 4, 3, 2 \rangle = A$.

Сообщение дошло до получателя без искажений.

Однако кроме математических проблем могут возникнуть нравственные, например перехват сообщения и замена ключа, а также

подмена сообщений неким злоумышленником. Цифровая подпись лишь констатирует факт, что сообщение пришло от того же отправителя, что и открытый ключ.

Шифрование с открытым ключом (СОК). Вообще-то это словосочетание в сокращенном виде должно быть записано как ШОК. Построение аббревиатуры (от лат. *brevis* — короткий) есть не что иное, как преобразование информации с ее потерей. Чтобы не шокировать читателя, мы воспользовались побуквенным переводом соответствующей английской аббревиатуры. Внимательный читатель конечно же сразу определит, что это алфавитное кодирование, причем в обе стороны. В современных информационных системах стало популярным шифрование с открытым ключом, которое осуществляется на основе математических знаний, например, таких разделов, как разложения чисел на простые множители, вычисление логарифмов чисел, решение алгебраических уравнений.

На основании теоремы Рабина доказано, что разложение на простые множители двух больших чисел эквивалентно раскрытию ключа для шифра RSA и практически невозможно в реальном времени с учетом возможностей современных ЭВМ.

Шифры с открытым ключом достаточно просты в обращении, практичны и обладают высокой криптостойкостью. И хотя сравнительно просто найти пару больших взаимно простых чисел, к настоящему времени не разработаны эффективные алгоритмы разложения чисел на простые множители. Так, разложение на множители числа в 200 и более цифр займет сотни лет работы компьютера. А так как при употреблении шифра с открытым ключом используются очень большие простые числа, содержащие сотни цифр в десятичной системе счисления, то вскрыть такие шифры весьма сложно.

Поэтому поиск простых чисел и их общей формулы в настоящее время представляет не только теоретический, но и практический интерес.

Получив сообщение, получатель сначала расшифровывает его закрытым ключом, а затем проверяет его подлинность. Для этого он сравнивает дешифрованный текст с тем, который был получен с помощью открытого ключа.

Алгоритмы кодирования и декодирования на самом деле весьма сложны. Основные идеи специальных кодов изложены в соответствующей литературе и защищены от злоумышленников на различных уровнях, включая юридическую защиту.

Рассмотрим один из таких алгоритмов, алгоритм RSA для некоторого пользователя A_i.

1. Пользователь выбирает пару различных простых чисел p_i и g_i.

2. Находит произведение $r_i = p_i \cdot g_i$ и **функцию Эйлера** φ от r_i — число взаимно-простых с r_i натуральных чисел, меньших r_i, включая 0 и 1: $\varphi(r_i)$.

3. Находит открытый ключ c_i — взаимно-простое с $\varphi(r_i)$ и меньшее его.

4. Находит закрытый ключ d_i — произвольное решение сравнения $d_i \cdot c_i \equiv 1(\mathrm{mod}\,\varphi(r_i))$ и меньшее $\varphi(r_i)$.

5. Публикует открытый ключ — пару $\{c_i, r_i\}$, которые доступны для любого пользователя.

6. Для отправления сообщения w абоненту A_i его нужно зашифровать открытым ключом $\{c_i, r_i\}$: $w' = w^{c_i}(\mathrm{mod}\,r_i) < r_i$.

7. Получив сообщение, A_i дешифрует его своим секретным ключом $\{d_i, r_i\}$: $w'' = w'^{d_i}(\mathrm{mod}\,r_i)$. Очевидно, что для полностью правильной дешифровки ($w = w''$) нужно, чтобы и исходное сообщение было меньше r_i. Для этого нужно *предварительно* зашифровать сообщение в форму, мощность алфавита которой меньше r_i, или разбить сообщение на несколько.

Задача 43. Зашифруем сообщение «САВА». Выберем для простоты небольшие простые числа (в отличие от тех больших, которые выбирают при реальном кодировании).

Пусть $p_1 = 2$, $g_1 = 11$.

Тогда $r_1 = 2 \cdot 11 = 22$, $\varphi(22) = 10$.

Выберем в качестве c_1 взаимно-простое с 10, например $c_1 = 3$.

Выберем d_1 из сравнения $(3 \cdot d_1) \equiv 1(\mathrm{mod}\,10)$. Таким (минимальным) числом является $d_1 = 7$.

Кодируем сообщение отображением множества букв на множество простых чисел с помощью подстановки $\sigma = \begin{pmatrix} A & B & C \\ 1 & 2 & 3 \end{pmatrix}$.

Поэтому сообщение имеет вид $\langle 3, 1, 2, 1\rangle$.

Зашифруем сообщение с помощью открытого ключа $\{3, 22\}$ для того, чтобы послать его абоненту A_1. Имеем

$3 \to (3^3)(\mathrm{mod}\,22) = 27(\mathrm{mod}\,22) = 5$,

$1 \to (1^3)(\mathrm{mod}\,22) = 1$,

$2 \to (2^3)(\mathrm{mod}\,22) = 8(\mathrm{mod}\,22) = 8$.

Получим сообщение $w' = \langle 5, 1, 8, 1\rangle$.

A_1 расшифрует это сообщение $\langle 5, 1, 8, 1\rangle$ с помощью закрытого ключа $\{7, 22\}$:

$5 \to 5^7(\mathrm{mod}\,22) = ((5^2)^3 \cdot 5)(\mathrm{mod}\,22) = ((5^2(\mathrm{mod}\,22))^3 \cdot 5)(\mathrm{mod}\,22) =$
$= (3^3 \cdot 5)(\mathrm{mod}\,22) = (27(\mathrm{mod}\,22) \cdot 5)(\mathrm{mod}\,22) = (5 \cdot 5)(\mathrm{mod}\,22) = 3$,

$1 \to (1^7)(\mathrm{mod}\,22) = 1$,

$8 \to (8^7)(\mathrm{mod}\,22) = (2^{21})(\mathrm{mod}\,22) = ((2^5)^4 \cdot 2)(\mathrm{mod}\,22) =$
$= ((2^5(\mathrm{mod}\,22))^4 \cdot 2)(\mathrm{mod}\,22) = (10^4 \cdot 2)(\mathrm{mod}\,22) = (100 \cdot 100 \cdot 2)(\mathrm{mod}\,22) =$
$= (12 \cdot 12 \cdot 2)(\mathrm{mod}\,22) = (12 \cdot 2)(\mathrm{mod}\,22) = 2$.

Таким образом, расшифрованное сообщение $w'' = \langle 3, 1, 2, 1\rangle$ соответствует исходному САВА. Полная схема, соответствующая рис. 6.4—6.5, имеет вид

$$\text{САВА} \xrightarrow{\sigma} 3123 \xrightarrow{F=\{3,22\}} 5181 \xrightarrow{F^{-1}=\{7,22\}}$$
$$\xrightarrow{F^{-1}=\{7,22\}} 3121 \xrightarrow{\sigma^{-1}} \text{САВА}.$$

Таблица 6.5

Трудоемкость разложения простых чисел

Длина	Число операций	Примечание
50	$1{,}4 \cdot 10^{10}$	Раскрываем на суперкомпьютерах
100	$2{,}3 \cdot 10^{15}$	На пределе современных технологий
200	$1{,}2 \cdot 10^{23}$	За пределами современных технологий
400	$2{,}7 \cdot 10^{34}$	Требует существенных изменений в технологии
800	$1{,}3 \cdot 10^{51}$	Требует существенных изменений в технологии (не раскрываем)

С точки зрения практической реализации такой шифр обладает высокой криптостойкостью, так как в его основе лежит выбор простых чисел. В настоящее время, как известно, не найдена математическая формула или алгоритм установления простого числа. Генерация простых чисел достаточно трудоемкая задача. Еще большую трудность вызывает определение ключей, также являющихся простыми числами.

Известны данные о трудоемкости разложения простых чисел различной длины, установленные Шроппелем (табл. 6.5).

Авторы алгоритма RSA рекомендуют использовать размеры модуля n для: частных лиц 768 бит, коммерческой информации 1024 бит.

На практике встречаются и иные криптоустойчивые системы, использующие как теорию вычетов, так и поиск логарифма числа — не менее трудоемкую математическую задачу.

Упражнения

6.1. Переведите в двоичную, восьмеричную и шестнадцатеричную системы счисления числа:

а) 53; б) 62; в) 71; г) 84; д) 96; е) 47.

6.2. Переведите двоичное число в десятичную, восьмеричную и шестнадцатеричную системы счисления:

а) 010011; б) 100111; в) 110100; г) 010101; д) 101010; е) 010010.

6.3. Найдите сумму, разность и произведение чисел A и B в двоичной системе счисления и сделайте проверку результата в десятичной системе счисления:

а) $A = 100110$, $B = 001001$; б) $A = 010111$, $B = 101000$; в) $A = 011011$, $B = 010010$; г) $A = 101100$, $B = 101000$; д) $A = 111010$, $B = 110000$; е) $A = 110011$, $B = 001001$.

6.4. Выполните сложение и вычитание чисел в восьмеричной системе счисления и проверьте результат, сложив эти же числа, переведенные в десятичную систему счисления:

а)	б)	в)	г)
$A_8 = 53412$,	$A_8 = 72047$,	$A_8 = 45236$,	$A_8 = 63405$,
$B_8 = 25706$;	$B_8 = 43576$;	$B_8 = 25743$;	$B_8 = 47345$;

д)	е)
$A_8 = 32657$,	$A_8 = 267354$;
$B_8 = 12476$;	$B_8 = 154757$.

6.5. Выполните сложение и вычитание в шестнадцатеричной системе счисления и сделайте проверку в десятичной:

а)	б)	в)	г)
A_{16} = C3896,	A_{16} = F79832,	A_{16} = 87FDC,	A_{16} = 987F4,
B_{16} = 845FA;	B_{16} = C9A53;	B_{16} = 5AB63;	B_{16} = 5ADF7;

д)	е)
A_{16} = B4599,	A_{16} = D5B6C,
B_{16} = A74F5;	B_{16} = B7AF8.

6.6. Составляя тексты для последующей компьютерной обработки, необходимо излагать их в сжатом, компактном виде. Изложите в 10 предложениях самую важную информацию, содержащуюся в:

а) главе 1; б) главе 2; в) главе 3; г) главе 4; д) главе 5; е) главе 6.

6.7. Систематизируйте информацию в табличном виде, содержащуюся в:

а) главе 1; б) главе 2; в) главе 3; г) главе 4; д) главе 5; е) главе 6.

6.8. Составьте кроссворд, в который бы входили основные понятия:

а) главы 1; б) главы 2; в) главы 3; г) главы 4; д) главы 5; е) главы 6.

6.9. Расшифруйте текст и составьте алфавит кодирования.

1. Федя договорился с бывшим одноклассником, живущим в деревне, что проведет у него отпуск. Его товарищ собирался ехать в тренировочный лагерь, и весь его дом должен был остаться в распоряжении Феди. Федя на целые сутки задержался в городе, и, когда он с вещами наконец выбрался в деревню, товарища уже не было, а дверь была на замке. Однако в щели Федя нашел адресованное ему письмо, содержание которого передано на рис. 6.10.

Федя сразу узнал тайнопись, которой они пользовались в школьные годы, но, к несчастью, он успел за прошедшие годы забыть шифр. Он помнил лишь то, что каждой рожице соответствовала определенная буква, причем буквы Е и Ё не различались. Однако после некоторого раздумья Феде все же удалось расшифровать тайнопись, найти ключ от двери и попасть в дом. Как звали школьного товарища Феди и что он писал в этом письме? Найдите ошибку в тексте письма.

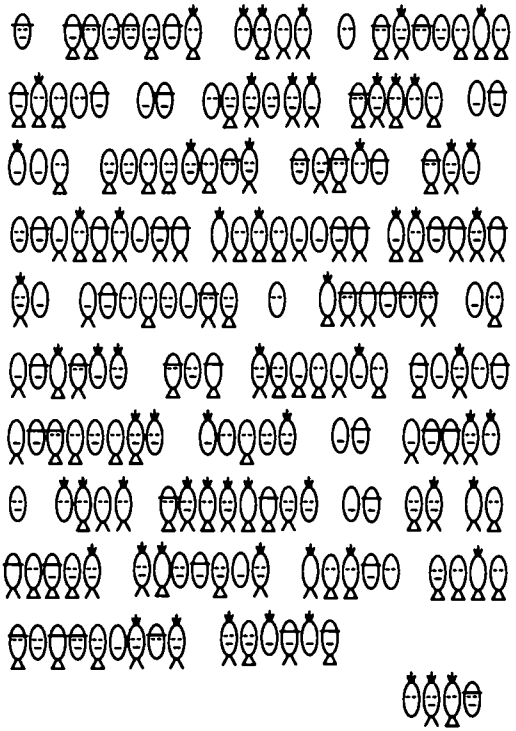


Рис. 6.10. Зашифрованное письмо

2. Расшифруйте отрывок из книги Л. Кэрролла «Алиса в Зазеркалье»:

«— БЕРПИ Э ЙДЕМГОКВЭЫ БИБЕО-ЖАКЙПЧ ЗВЕЛЕ, — ЗБИСИВ ФИВМИУ-КЕВМИУ ПЕЛЕВЧЖЕ ДГОСГАМОВЧЖЕ, — ЕЖЕ ЕСЖИЬИОМ МЕВЧБЕ МЕ, ЬМЕ Э ЦЕЬЙ, ЬМЕКЮ ЕЖЕ ЕСЖИЬИВЕ, — ЖА КЕВЧФО, ЖА ТОЖЧФО».

Текст зашифрован так: десять букв («а», «е», «и», «й», «о», «у», «ы», «э», «ю», «я») разбиты на пары, и каждая из этих букв в тексте заменена второй из пары. Все остальные буквы точно так же разбиты на пары.

3. Найдите ключ к «тарабарской грамоте» — тайнописи, применявшейся ранее в России для дипломатической переписки:

«Пайцике тсюг т “камащамлтой чмароке” — кайпонили, имирепяшвейля мапее ш Молии цся цинсоракигелтой неменилти».

4. Как-то раз Таня ехала в поезде. Чтобы не скучать, она стала зашифровывать названия городов, заменяя буквы их порядковыми номерами в алфавите. Когда Таня зашифровала пункты прибытия и отправления поезда, то с удивлением обнаружила, что они записываются с помощью всего лишь двух цифр: 21221 — 211221. Откуда и куда шел поезд?

5. Расшифруйте знаменитое послание Цезаря, которое зашифровал в I в. до н.э. по его поручению знаменитый римский политический деятель, оратор и философ Марк Туллий Цицерон:

УЮЫОЛХИВРСЕЮИВНЦНГКРЛБ (?)

6. Придумайте свой код и предложите закодированный текст с ключом к его расшифровке.

6.10. Латинский вариант двоичного кода Фрэнсиса Бэкона имел вид:

a	*b*	*c*	*d*	*e*	*f*
AAAAA	*AAAAB*	*AAABA*	*AAABB*	*AABAA*	*AABAB*
g	*h*	*i*	*k*	*l*	*m*
AABBA	*AABBB*	*ABAAA*	*ABAAB*	*ABABA*	*ABABB*
n	*o*	*p*	*q*	*r*	*s*
ABBAA	*ABBAB*	*ABBBA*	*ABBBB*	*BAAAA*	*BAAAB*
t	*v*	*w*	*x*	*y*	*z*
BAABA	*BAABB*	*BABAA*	*BABAB*	*BABBA*	*BABBB*

Зашифруйте кодом Бэкона латинское крылатое выражение:

а) «*Ab altero expectes, alteri quod feceris*» — «Жди от другого того, что сам ты сделал другому»;

б) «*Ad cogitandum etagendum homo natusest*» — «Человек рожден для мышления и деятельности»;

в) «*Audendo virtus crescit, tardando timor*» — «Мужество возрастает благодаря смелым поступкам, а страх — от медлительности»;

г) «*Beate vivere est honeste vivere*» — «Счастливо жить — значит честно жить»;

д) «*Bono ingenio me esse ornatum, quam multo auro mavolo*» — «Лучше, чтобы моим украшением был приятный характер, чем гора золота»;

е) «*Bonorum vita vacua est metu*» — «Жизнь честных (людей) свободна от страха».

6.11. Набор знаков азбуки Морзе из двух символов «точка—тире» имеет вид:

Буквы		*Ф* · · — ·	*Цифры*
А · —	*К* — · —	*Х* · · · ·	1 · — — — —
Б — · · ·	*Л* · — · ·	*Ц* — · — ·	2 · · — — —
В · — —	*М* — —	*Ч* — — — ·	3 · · · — —
Г — —	*Н* — ·	*Ш* — — — —	4 · · · —
Д — · ·	*О* — — —	*Щ* — — · —	5 · · · · ·
Е ·	*П* · — — ·	*Ь, Ъ* — · · —	6 — · · · ·
Ж · · · —	*Р* · — ·	*Ы* — · — —	7 — — · · ·
З — — ·	*С* · · ·	*Э* · · — · ·	8 — — — · ·
И · ·	*Т* —	*Ю* · · — —	9 — — — — ·
Й · — — —	*У* · · —	*Я* · — · —	0 — — — — —

Зашифруйте с помощью азбуки Морзе крылатые латинские изречения:

а) «Короткий век человека достаточно продолжительный для того, чтобы жить хорошо и честно»;

б) «С болтуном не спорь: речь дается всем, разум — немногим»;

в) «Каждому человеку свойственно ошибаться, упорствовать в заблуждениях свойственно только глупцу»;

г) «Бедным людям многого не хватает, жадным — всего»;

д) «Для ученого и образованного человека жить — значит мыслить»;

е) «Плоды науки сладкие, корни — горькие».

6.12. В какой системе счисления x возможно равенство:

а) $4_{10} = (10)_x$; б) $8_{10} = (11)_x$; в) $9_{10} = (100)_x$;

г) $26_{10} = (31)_x$; д) $23_{10} = (17)_x$; е) $42_{10} = (52)_x$?

6.13. Расшифруйте арифметический ребус:

а)
$$\begin{array}{r} \text{сорок} \\ + \\ \text{один} \\ \hline \text{триста} \end{array}$$

б)
$$\begin{array}{r} \text{квант} \\ + \\ \text{квант} \\ \hline \text{наука} \end{array}$$

в)
$$\begin{array}{r} \text{лето} \\ + \\ \text{лето} \\ \hline \text{полет} \end{array}$$

г)
$$\begin{array}{r} \text{муха} \\ + \\ \text{муха} \\ \hline \text{слон} \end{array}$$

д)
$$\begin{array}{r} \text{елка} \\ + \\ \text{елка} \\ + \\ \text{елка} \\ \hline \text{лесок} \end{array}$$

е)
$$\begin{array}{r} \text{ахинея} \\ + \\ \text{ахинея} \\ \hline \text{чепуха} \end{array}$$

Одинаковым буквам соответствуют одинаковые цифры, разным — разные.

6.14. Определите, какую информацию несут в себе сообщения:

а) из каждых семи выпускников колледжа пятеро продолжают образование;

б) на этой неделе необходимо сдать зачет по математической логике;

в) в четверг наша группа идет в театр;

г) из 24 студентов нашей группы 16 занимаются в творческих объединениях и спортивных секциях колледжа;

д) из 18 участников научно-исследовательского студенческого общества 8 студентов выступили с сообщениями на конференции;

е) из 12 участников конференции 8 студентов посещали занятия научно-исследовательского студенческого общества.

6.15. Какую информацию несет в себе сообщение о том, что при обнаружении фальшивой монеты из n имеющихся на чашечки весов положили k_i монет? Какое минимальное число взвешиваний при этом необходимо?

а) $n = 26$, $k_1 = k_2 = 9$, $k_3 = 8$; г) $n = 29$, $k_1 = k_2 = 10$, $k_3 = 9$;

б) $n = 25$, $k_1 = k_2 = 8$, $k_3 = 9$; д) $n = 31$, $k_1 = k_2 = 10$, $k_3 = 11$;

в) $n = 28$, $k_1 = k_2 = 9$, $k_3 = 10$; е) $n = 32$, $k_1 = k_2 = 10$, $k_3 = 12$.

Глава 7
КОНЕЧНЫЕ АВТОМАТЫ

Алгебраическая теория автоматов является разделом теоретической кибернетики, который изучает дискретные автоматы с абстрактной алгебраической точки зрения.

Общая теория автоматов содержит различные подразделы. В зависимости от предмета изучения она делится на абстрактную теорию автоматов и структурную теорию автоматов.

Абстрактная теория автоматов изучает переходы, совершаемые автоматом, на который воздействуют входные сигналы, а также выходные сигналы как результат этих переходов.

Предметом изучения *структурной* теории автоматов является структура автомата, а также структура входных и выходных сигналов, например, способы кодирования входных и выходных сигналов и др.

7.1. Определение конечных автоматов

> В принципе, автоматы могут обладать всеми основными свойствами самых сложных систем, и даже человеческого мозга.
>
> *А. Н. Колмогоров*

Под **автоматом** (от греч. *αυτοματος* — самодействующий) принято понимать абстрактную модель устройства, функционирующего в дискретном времени, которая перерабатывает последовательность входных сигналов (стимулов) и превращает их в последовательность выходных сигналов (реакций). В процессе работы автомата происходит последовательная смена его внутренних состояний, причем состояние автомата в определенный момент времени однозначно определяется входным и выходным сигналами. Такие автоматы представляют собой основу всей современной вычислительной техники и всевозможных дискретных систем автоматического контроля и управления.

С определенной долей абстрагирования можно утверждать, что попытки описания информационных моделей поведения биологических систем средствами математики также приводят к понятию автомата. Так, фон Нейман рассматривал автоматы как метаязык для описания кибернетических систем. Советский ученый М.Л. Цетлин (1924—1966) совместно с учениками использовал такой подход при исследовании процессов целесообразного поведения взаимодействующих объектов любой природы, в том числе живых существ.

? Когда появились первые автоматы?
Кого можно считать автором первых автоматов?

История не сохранила ответа на эти вопросы. Дело в том, что под определение автомата подходят любые устройства, в том числе те, которыми первобытные люди охотились или метали камни, защищая свое жилище от неприятеля. Характерными признаками древних автоматов были автономный источник энергии и некоторый исполнительный механизм. Энергия, переходя от источника к механизму, заставляет его совершать заданные действия. Для управления такой системой к механизму должны поступать определенные указания в виде некоторых сигналов. Причем источником сигналов, средствами их передачи, накопления и преобразования являются элементы самой системы.

Такой автомат выполняет некоторые конкретные команды по заданному алгоритму. Под **алгоритмом** мы понимаем точное формальное предписание, однозначно определяющее содержание и последовательность операций, переводящих заданную совокупность исходных данных в искомый результат.

Считается, что первым программным устройством, созданным человеком, были часы. Часовые механизмы с помощью пружины, приводящей в действие шестеренки и кулачковые механизмы, зубчатые колеса и рычаги, осуществляют ряд определенных действий. Примером такого часового механизма могут быть знаменитые часы на Центральном театре кукол в Москве, где он приводит в действие двенадцать сказочных героев, расположенных на циферблате.

Первые башенные часы были установлены в 1352 г. на соборе в городе Страсбурге. В течение дня во время боя часов на их циферблате появлялись изображения людей и животных.

Сейчас все чаще употребляется термин «интеллектуальный робот», создано целое научное направление — «искусственный интеллект». Понятие «робот» вошло в историю с легкой руки чешского писателя Карела Чапека, который в 1920 г. написал пьесу «R.U.R.» («Россумовские универсальные роботы»).

Перелистаем некоторые страницы истории создания автоматов.

Тридцать лет, с 1216 по 1246 г., создавал своего железного слугу знаменитый немецкий философ и алхимик Альберт Великий. Его автомат выполнял в доме обязанности привратника. Механический слу-

га-швейцар с поклоном открывал двери гостям в доме Альбертуса Магнуса.

Механический орел знаменитого немецкого астронома Иоганна Региамонтана (1436—1476) приветствовал наклоном головы и движением крыльев въезд в Нюрнберг императора священной Римской империи Максимилиана II.

Автором первого в мире автоматического ткацкого станка был французский механик Жак де Вакансон (1709—1782). Его механическая утка — точная копия своего живого двойника — плавала, чистила перья, глотала с ладони зерна и т.д. Утка, а также механический флейтист, исполнявший одиннадцать музыкальных пьес, продолжали путешествовать по свету и участвовать в различных выставках даже после смерти Вакансона.

Целый механический кабинет талантливого русского изобретателя XIX в. А.М.Гамулецкого — настоящий музей, в котором было множество сконструированных им автоматов, от говорящей головы чародея до амура, играющего на арфе, которые поражали воображение современников.

Два мальчика, пишущий и рисующий, и их сестра-музыкант — автоматические куклы, осуществлявшие все действия, аналогичные действиям живых людей (повороты головы, движение глаз, пальцев рук и т.д.), а также результаты их творчества (написанный текст, рисунки, игра на фисгармонии), были виртуозно выполнены двумя швейцарцами — часовых дел мастерами отцом и сыном Дро во второй половине XVIII в. Железные люди Пьера и Анри Дро вместе со своими создателями пережили массу приключений: попадали в кораблекрушение, после чего так называемые «андроиды» ржавели, их создатели подвергались нападкам инквизиции и т.д.

Первый примитивный арифмометр сконструировал в 1641 г. Блез Паскаль. Толчком для открытия были мучения его отца — налогового инспектора, который днями и ночами работал с большими вычислениями. Изобретя арифмометр, восемнадцатилетний сын избавил отца от сложных вычислений, а миру подарил первый калькулятор, производящий сложение и вычитание чисел.

Первый шахматный автомат был построен в 1890 г. испанским инженером Торресом Кеведо. Такой автомат мог разыграть лишь ладейный эндшпиль (король и ладья против короля).

Первую вычислительную машину с автоматическим управлением создал Чарльз Баббедж в 1822 г. Он спроектировал сверхарифмометр, который имел запоминающие и арифметические устройства, аналогичные современным ЭВМ.

Виды автоматов. Напомним, что **автоматами** принято называть устройства, выполняющие процессы приема, преобразования и передачи энергии, материалов или информации в соответствии с заложенной в них программой, но без непосредственного участия человека.

По виду деятельности автоматы делятся на три класса: информационные, управляющие и вычислительные.

К **информационным** автоматам мы относим разнообразные справочные таблицы, например, на вокзалах, информационные табло на стадионах, различные устройства аварийной сигнализации. К **управляющим** автоматам принято относить устройства для управления лифтом, конвейером, различные типы управляемых станков, автоматические шлагбаумы. К **вычислительным** автоматам мы будем относить микрокалькуляторы (МК) и процессоры в ЭВМ. Однако, строго говоря, эти сложные системы являются одновременно и вычислительными, и управляющими, и информационными.

В информатике **конечным автоматом** называют дискретный преобразователь информации, т.е. преобразователь, в котором множество входных и выходных сигналов, а также множество внутренних состояний являются *конечными множествами.*

Автомат называют **цифровым,** если он преобразует *цифровую* информацию. В таком автомате входные сигналы задаются в виде конечного множества мгновенных символов: их длительность настолько мала, что ею можно пренебречь. За фиксированное время t происходит преобразование входных символов, а на выходе происходит скачкообразный переход из состояния $q_i(t)$ в состояние $q_{i+1}(t)$ (рис. 7.1).

Таким образом, абстрактный автомат отображает множество слов входного алфавита X во множество слов выходного алфавита Y. Входной и выходной алфавиты, а также множество состояний автомата называют его **базовыми множествами.**

Характерной особенностью конечного автомата является наличие *памяти*, которая определяет состояние автомата в зависимости от времени. Внешним проявлением различных состояний автомата является его реакция на однотипные воздействия (сигналы).

В работе конечных цифровых автоматов важное значение имеет *время.* В зависимости от того, одновременно или последовательно принимаются входной сигнал и сигнал смены состояний, автоматы делятся на *синхронные* и *асинхронные.*

В **синхронных** автоматах продолжительность входных сигналов и время переходов согласованы. Они используются, например, в вычислительных комплексах, АСУ и т.д.

В **асинхронных** автоматах продолжительность входных сигналов и время переходов не согласованы. Они зависят от внешних источников — различных событий, а *интервал дискретности* является переменным (например, в кодовых замках). В асинхронных автоматах очередное изменение значений входных сигналов может произойти только при условии, что закончился переходный процесс, вызванный предыдущим изменением этих сигналов.

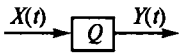


Рис. 7.1. Схема автомата

Классификацию автоматов можно проводить и по другим основаниям.

Если в основании классификации лежит *объем памяти*, то различие заключается в том, имеет ли автомат конечное или бесконечное число внутренних состояний.

Под **бесконечным** автоматом обычно понимают определенную математическую идеализацию представлений об автомате, имеющую бесконечное число состояний. Память такого автомата потенциально может неограниченно возрастать. Например, известные абстрактные автоматы Поста и Тьюринга являются бесконечными автоматами, но сама ЭВМ или ее отдельные части — конечными автоматами.

Если в основании классификации лежит *механизм случайного выбора*, то различают детерминированные и вероятностные (стохастические) автоматы. В **детерминированных** автоматах поведение и структура в каждый момент времени однозначно определены текущей входной информацией и состоянием самого автомата в предшествующий момент времени.

В вероятностных автоматах эта зависимость связана еще и с некоторым случайным выбором. **Вероятностный** автомат — это дискретный преобразователь информации, функционирование которого в каждый момент времени зависит только от состояний памяти и описывается статистическими законами.

В теории автоматов доказано, что для выполнения различных преобразований информации достаточно построить *универсальный* автомат с помощью программы и соответствующего кодирования, способный решать любые задачи.

Пусть имеется цифровой автомат с одним входом, математическая модель которого задается пятью объектами:

X — конечное множество входных символов, входной алфавит: $X = \{x_1(t), x_2(t), ..., x_n(t)\}$; множество слов над X – X^*;

Y — конечное множество выходных символов, выходной алфавит: $Y = \{y_1(t), y_2(t), ..., y_k(t)\}$;

Q — конечное множество состояний автомата: $Q = \{q_0(t), q_1(t), ..., q_s(t)\}$, где q_0 — его начальное состояние;

$\delta(q, x)$ — функция перехода автомата из одного состояния в другое: $Q \times X \to Q$;

$\lambda(q, x)$ — функция выхода автомата: $Q \times X \to Y$.

Таким образом, конечный автомат $Z = \langle X, Q, Y, \delta, \lambda \rangle$ определяется рекуррентными соотношениями

$q(0) = q_0$, $q(t + 1) = \delta(q(t), x(t))$, $y(t) = \lambda(q(t), x(t))$.

Здесь t — дискретизированный момент времени, т. е. образ монотонной функции t: $T \to \mathbb{N}$, где T — обычное непрерывное время, $\mathbb{N}$ — множество натуральных чисел. Таким образом, все время работы T разбивается на конечное число интервалов, на границе которых происходит изменение состояния автомата. При таком

подходе $t(T_0)$ — число таких изменений, произошедших до момента времени T_0. Примером дискретизации служит обычный кинематограф: время разбито на интервалы длительностью 1/(24 с), и человеческий глаз воспринимает следование дискретных кадров как непрерывное движение.

В зависимости от *способа организации функции выхода* синхронные автоматы делятся на автоматы Мили и автоматы Мура.

В **автоматах Мили** (автоматы I рода) выходной сигнал $y(t)$ однозначно определяется входным сигналом $x(t)$ и состоянием $q(t-1)$ автомата в предшествующий момент времени $(t-1)$. Математической моделью таких автоматов служит система уравнений:

$$\begin{cases} q(t) = \delta(q(t-1),\ x(t)), \\ y(t) = \lambda(q(t-1),\ x(t)), \end{cases} \text{ где } t = 1,\ 2,\ \ldots$$

В **автоматах Мура** (автоматы II рода) выходной сигнал $y(t)$ однозначно определяется входным сигналом $x(t)$ и состоянием $q(t)$ в данный момент времени t. Математической моделью таких автоматов является система: $\begin{cases} q(t) = \delta(q(t-1),\ x(t)), \\ y(t) = \lambda(q(t)). \end{cases}$

В таких автоматах функция выхода зависит только от состояний автомата в данный момент времени и не зависит от входного сигнала. Таким образом, входная строка такого автомата однократно считывается слева направо, осуществляя поочередный просмотр символов. В определенный момент времени конечный автомат находится в некотором внутреннем состоянии, которое изменяется после считывания очередного символа. Новое состояние можно охарактеризовать считанным символом и текущим состоянием.

Комбинационными называются автоматы, в которых выходной символ не зависит от его состояния и определяется лишь текущими входными символами, т.е. в этом автомате все состояния эквивалентны. В таком автомате вырождена функция перехода, она принципиально не важна и в процессе функционирования неизменна. Поэтому минимальный комбинационный автомат имеет лишь одно состояние.

Кроме комбинационных, существуют также и **логические** автоматы, у которых входной алфавит состоит из 2^m двоичных наборов длины m, а выходной — из 2^n двоичных наборов длины n. Для *логических комбинационных* автоматов функция выхода имеет вид системы n логических функций от m переменных.

Представление событий в автомате. Обозначим через X^* множество входных слов автомата Z, составленное из множества букв входного алфавита X. Множества слов входного алфавита называют также **событием** или **языком.**

Для некоторого автомата Z определим функции перехода δ_z и выхода λ_z *индуктивно*:

• функция перехода $\delta(q_i, x_j)$ задается автоматной таблицей Z, т.е. известна функция перехода при входных буквах;

• для любого слова $\alpha \in X^*$ и любой буквы $x_j \in X$ справедливо $\delta(q_i, \alpha x_j) = \delta(\delta(q_i, \alpha), x_j)$.

Фактически это значит, что, опираясь на значения функции перехода при входных буквах, мы определяем ее значения при входных словах, состоящих из большего количества букв. Последняя формула означает, что перед нажатием последней буквы x_j автомат после ввода слова α перешел во внутреннее состояние $q = \delta(q_i, \alpha)$. Если же $\alpha = \Lambda$, то на пустое слово автомат не должен реагировать: это задает начальное условие $\delta(q_i, \Lambda) = \delta(q_i)$.

Такая расширенная функция перехода дает возможность определить и расширенную функцию выхода, для которой справедливо $\lambda(q_i, \alpha x_j) = \lambda(\delta(q_i, \alpha), x_j)$.

Событие $E \subseteq X^*$ представимо в автомате $Z = \langle X, Q, Y, \delta, \lambda \rangle$ тогда и только тогда, когда $\forall \alpha \in E\ \exists q_i \in Q$, такое, что определено $\delta(q_i, \alpha)$. На графе автомата события изображаются в виде подмножества всех путей, исходящих из вершины q_i. Событие называют **представимым** в конечном автомате, если существует конечный автомат, в котором оно представимо.

Итак, когда событие, представимое автоматом, получено непосредственно из входного алфавита, т.е. без переходов в новые состояния, соответствующее слово является пустым. Пустое слово аналогично 1: для любого α справедливо $\Lambda\alpha = \alpha\Lambda = \alpha$. Однако нельзя путать пустое слово с пустым событием, которое состоит в том, что ни одно из его заключительных состояний не может быть достижимо из начального.

Доказано, что любое конечное множество слов представимо в автомате. Однако существуют события, непредставимые в автомате. Так, конечные автоматы не распознают непериодические бесконечные последовательности слов, так как множество таких событий несчетно, а множество состояний автомата счетно.

7.2. Способы задания конечных автоматов

Принципы — не исходный путь исследования, а его конечный результат.

Ф. Энгельс

Существуют три способа задания конечных автоматов:

• табличный (матрицы переходов и выходов);

• графический (с помощью графов);

• аналитический (с помощью формул).

Аналитический способ. Как отмечалось в предыдущем подразделе, аналитически автомат можно задать системой уравнений:

$$\begin{cases} q(t) = \delta(q(t-1),\ x(t)), \text{ где } x(t) \in X, \\ y(t) = \lambda(q(t-1),\ x(t)), \text{ где } y(t) \in Y, \\ q(0) = q_0, \text{ где } q(t) \in Q. \end{cases}$$

Табличный способ. Функцию перехода δ и выхода λ можно представить в виде таблицы, столбцы которой соответствуют входным символам (элементы входного алфавита X), а строки — состояниям (элементы конечного множества Q). Поэтому пересечение i-й строки и j-го столбца — клетка (i, j) — соответствует аргументу функций δ и λ автомата в момент, когда он находится в состоянии q_i, на его входе — слово x_j, а в самой соответствующей клетке запишем значения функций δ и λ. Таким образом, вся таблица соответствует множеству $Q \times X$.

Для автомата Мили она имеет вид табл. 7.1.

Таблица 7.1

Таблица переходов автомата Мили

	x_1	x_2	...	x_k
q_1	$\delta(q_1, x_1) \mid \lambda(q_1, x_1)$	$\delta(q_1, x_2) \mid \lambda(q_1, x_1)$	...	$\delta(q_1, x_k) \mid \lambda(q_1, x_k)$
q_2	$\delta(q_2, x_1) \mid \lambda(q_1, x_1)$	$\delta(q_2, x_2) \mid \lambda(q_1, x_1)$	...	$\delta(q_2, x_k) \mid \lambda(q_2, x_k)$
...	...	...	...	...
q_n	$\delta(q_n, x_1) \mid \lambda(q_n, x_1)$	$\delta(q_n, x_2) \mid \lambda(q_n, x_2)$	...	$\delta(q_n, x_k) \mid \lambda(q_n, x_k)$

В этой таблице δ: $Q \times X \to Q$, λ: $Q \times X \to Y$.

В таблице переходов для автомата Мура каждая строка определяет состояние q_i и выходной сигнал $y_j = \lambda(q_i)$, соответствующий этому состоянию (табл. 7.2).

Таблица 7.2

Таблица переходов автомата Мура

	x_1	x_2	...	x_m	$\lambda(q)$
q_1	$\delta(q_1, x_1)$	$\delta(q_1, x_2)$	...	$\delta(q_1, x_m)$	y_1
q_2	$\delta(q_2, x_1)$	$\delta(q_2, x_2)$	...	$\delta(q_2, x_m)$	y_2
...	...	...	...	...	...
q_n	$\delta(q_n, x_1)$	$\delta(q_n, x_2)$	...	$\delta(q_n, x_m)$	y_n

В этой таблице δ: $Q \times X \to Q$, λ: $Q \to Y$.

При заполнении таблицы переходов каждая клеточка однозначно определяется парой символов: символом следующего состояния и символом выходного сигнала. Таким образом, рассматрива-

ются все возможные ситуации, встречающиеся при работе этого автомата. Однако при большом числе состояний и входных сигналов таблица переходов становится громоздкой и неудобной в употреблении.

Графический способ. Задание с помощью ориентированного графа — более удобная и компактная форма описания автомата. Граф автомата содержит вершины, соответствующие состоянию $q_i \in Q$, а также дуги, соединяющие вершины, — переходы автомата из одного состояния в другое. На дугах принято указывать пары входных и выходных сигналов — сигналов переходов.

Если автомат переходит из состояния q_i в состояние q_j под воздействием нескольких входных сигналов, то на соответствующей дуге графа этот вариант будет представлен через дизъюнкцию. Для представления автомата используют двухполюсные графы с выделенными начальным и конечным состояниями.

Задача 44. Опишем работу кодового замка с пятью кнопками. Назовем их A, B, C, D, E. Замок открывается при нажатии одной кнопки и остается открытым 5 с, причем никакие две кнопки одновременно не набираются.

Решение. Множество входных символов содержит 6 элементов: сигналы A, B, C, D, E — сигналы нажатия соответствующей буквы, и сигнал * — сигнал паузы (кнопка не нажата): $X = \{A, B, C, D, E, *\}$. Множество выходных сигналов Y содержит два элемента: 0 — означает, что дверь закрыта (замок не открыт), и 1 — замок открыт. Замок открывается при нажатии одной кнопки (например, кнопки B), тогда выходной сигнал зависит только от текущего входного сигнала. Таким образом, любое кодовое слово, имеющее B только в конце, например $D^*C^*A^*E^*B$, открывает замок. У такого замка может быть два состояния: основное q_0 — он закрыт и на любые нажатия (кроме B) не реагирует, и состояние, когда он среагировал на B и из q_0 перешел в q_1, а дверь открылась. Далее, удерживая B, автомат остается в состоянии q_1: $\delta(q_1, B) = q_1$ и $\lambda(q_1, B) = 1$.

Любое нажатие влечет за собой закрытие двери и переход автомата обратно в q_0. Учитывая, что две клавиши одновременно не набираются, клавишу B рано или поздно придется отпустить, тогда будет подаваться *, и дверь закроется.

Такой кодовый замок задается табл. 7.3.

Таблица 7.3

Задание автомата — кодового замка

	A	B	C	D	E	*
q_0	$q_0 \mid 0$	$q_1 \mid 1$	$q_0 \mid 0$	$q_0 \mid 0$	$q_0 \mid 0$	$q_0 \mid 0$
q_1	$q_0 \mid 0$	$q_1 \mid 1$	$q_0 \mid 0$	$q_0 \mid 0$	$q_0 \mid 0$	$q_0 \mid 0$

Автомат является комбинационным.

Задача 45. Описать работу кодового замка, состоящего из пяти последовательно нажимаемых кнопок, который открывается при нажатии двух кнопок.

Решение. Пусть кодовый замок с пятью кнопками открывается при последовательном нажатии двух кнопок, например кнопок B и D. Таким образом, для открытия замка необходим набор символов B, *, D, причем, пока кнопка D нажата, замок открыт. Для задания состояний такого кодового замка-автомата составим таблицу переходов и выходов (табл. 7.4).

Эта таблица показывает, что на пересечении q_i строки и x_j столбца наш автомат находится в состоянии $\delta(q_i, a_j)$ и $\lambda(q_i, a_j)$, что можно задать аналитически. Перечислим эти состояния:

$\delta(q_2, B) = q_2$ и $\lambda(q_2, B) = 0$;

$\delta(q_3, *) = q_3$ и $\lambda(q_3, *) = 0$;

$\delta(q_4, D) = q_4$ и $\lambda(q_4, D) = 1$ — т.е. замок открылся.

Для всех других входных сигналов A, C и E выполняется $\delta(q_1, A) = q_1$ и $\lambda(q_1, A) = 0$, $\delta(q_1, C) = q_1$ и $\lambda(q_1, C) = 0$, $\delta(q_1, E) = q_1$ и $\lambda(q_1, E) = 0$.

Этот автомат можно представить в виде диаграммы — ориентированного графа. Граф состояний такого автомата содержит 4 вершины — состояния q_i и дуги, соответствующие переходу автомата из состояния q в состояние q_j. На дугах такого ориентированного графа указывают входной сигнал, вызывающий этот переход, и выходной сигнал как результаты, которые отделяются друг от друга вертикальной чертой (рис. 7.2).

Такой ориентированный взвешенный граф наглядно демонстрирует путь для отпирания замка, т.е. необходимую последовательность символов $\{B, *, D\}$. Заметим, что если считать уже нажатую клавишу поступающим входным сигналом, то, во-первых, кодом для открытия будут слова вида $B...B*...*D...D$, во-вторых, в таблице перехода появятся $\delta(q_2, B) = q_2$ и т.д., в-третьих, аналогичные петли появятся и в графе на рис. 7.2. В таком случае, даже

Таблица 7.4

Таблица переходов автомата — кодового замка

Состояния	Входы					
	A	B	C	D	E	*
q_1	$q_1 \mid 0$	$q_2 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$
q_2	$q_1 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$	$q_3 \mid 0$
q_3	$q_1 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$	$q_4 \mid 1$	$q_1 \mid 0$	$q_1 \mid 0$
q_4	$q_1 \mid 0$	$q_1 \mid 0$	$q_1 \mid 0$	$q_4 \mid 1$	$q_1 \mid 0$	$q_1 \mid 0$

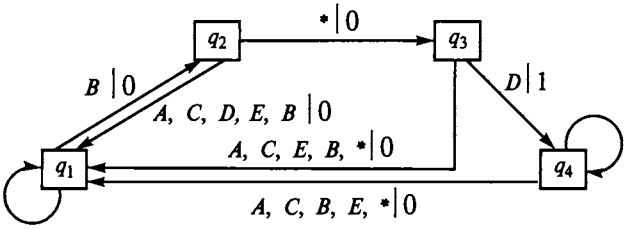


Рис. 7.2. Схема работы кодового замка

если никто не пытается взломать замок, автомат продолжает работать — подается сигнал *.

Рассмотренный кодовый замок будет являться *асинхронным* автоматом. Действительно, как известно из определения, у асинхронных автоматов время действия входного сигнала не совпадает со временем смены состояний, а следующий входной сигнал автомат получает после перехода в новое состояние. Для того чтобы неопределенное время действия входного сигнала не отражалось на работе такого автомата, после получения повторного сигнала он не меняет своего состояния. Так, при повторном нажатии кнопки B автомат остается в состоянии q_2, а при повторной паузе * — в q_3.

Описанный процесс называют этапом *абстрактного синтеза* автоматов или этапом **синтеза** абстрактных автоматов.

Итак, абстрактный автомат есть математическая модель, алгоритм действия некоторого преобразования кодовых последовательностей, закон преобразования входного алфавита в выходной.

7.3. Общие задачи теории автоматов

Sapienti sat (Plautus) — для понимающего достаточно.

Латинская формула

Абстрактный автомат рассматривается только как преобразование входных последовательностей в выходные, без учета его структуры. Но не всякое преобразование последовательностей может быть представлено конечным автоматом. Одной из главных проблем абстрактной теории автоматов является выяснение того, какие возможны преобразования, как их описать и каким образом математическая сущность этих преобразований связана с числом состояний, а значит, и со степенью сложности автоматов.

Принято выделять **три основные задачи** теории автоматов.

Задача синтеза заключается в том, что необходимо построить автомат Z, реализующий отображение множества слов над алфавитом X во множество слов над алфавитом Y.

В рамках общей задачи синтеза конечных автоматов возникают отдельные частные проблемы.

Проблема существования: существует ли автомат данного типа, соответствующий заданным условиям?

Проблема единственности: единственным ли образом можно синтезировать заданный автомат?

Проблема минимизации: по заданным начальным условиям с помощью эквивалентных преобразований привести автомат к изоморфному, удовлетворяющему определенным критериям оптимальности.

При выборе языка предпочтение отдается языкам, основанным на использовании логики предикатов.

Обратная задача — **задача анализа** — заключается в поиске по заданному графу конечного автомата Z отображения слов входного алфавита X во множество слов его выходного алфавита Y.

Для конечных детерминированных автоматов разработано достаточное число алгоритмов их синтеза и анализа.

Анализ можно рассматривать как вычисление предиката. Так как каждый предикат полностью характеризуется своим множеством истинности, то задача анализа автоматов может быть сведена к поиску множества истинности соответствующего предиката. В таком случае говорят, что это множество распознается автоматом. Причем для многих классов автоматов классы распознаваемых ими множеств хорошо известны.

Однако далеко не всегда по заданным автомату и множеству удается установить, насколько точно автомат распознал это множество. Так, не существует алгоритмов, которые смогли бы выполнить эту задачу, она считается алгоритмически неразрешимой.

Решение задач анализа и синтеза автоматов связано с понятием *автоматного времени*. Так, в синхронных автоматах моменты времени, в которые фиксируются изменения их состояний, задают специальные устройства — *генераторы синхросигналов*. Они выдают импульсы через равные промежутки времени, задавая так называемый постоянный *интервал дискретности*.

Задача декомпозиции заключается в поиске таких эквивалентных преобразований автомата, которые при минимальном числе состояний имеют те же свойства, что и при заданном. То есть необходимо определить, как можно получить заданный автомат Z, соединив минимальное число более простых автоматов, может быть, за счет усложнения комбинационных схем. Но при этом поведение автомата характеризуется одинаковыми внешними проявлениями. Так как понятие композиции имеет разное значение, то задачу декомпозиции автоматов можно ставить по-разному.

Простейшим примером является рассмотрение автоматов в виде суммы, произведения или последовательного соединения. Но наибольший интерес представляют композиции более простые, чем

исходный автомат, т.е. те, у которых либо меньшее число состояний, либо меньшее число входных каналов. Поэтому задача декомпозиции решается неоднозначно.

Тождественными принято считать такие автоматы, которые одинаково ведут себя при различном числе состояний. Основная задача декомпозиции автоматов состоит в поиске эффективных процедур, позволяющих по заданному автомату находить композицию, моделирующую исходный. Так, с практической точки зрения представляют значительный интерес замены сложной системы на более простые и экономически более выгодные системы, выполняющие те же функции, что и исходная.

Композицией автоматов называют операции, используемые для создания новых автоматов из других, исходных. Автоматы, полученные в результате таких операций, также называют **композицией** автоматов. Основная задача композиции — минимизация числа состояний автомата, т.е. построение по произвольно заданному конечному автомату нового с наименьшим числом возможных состояний и обладающего теми же свойствами, что и исходный.

На практике задача синтеза автоматов может встретиться в следующей ситуации: заказчик придумал конкретный оператор, реализуемый некоторым автоматом, но его базовая подготовка не дает возможности описать и построить задуманный автомат. Тогда исполнитель при помощи заказчика описывает автомат, что дает возможность его синтезировать.

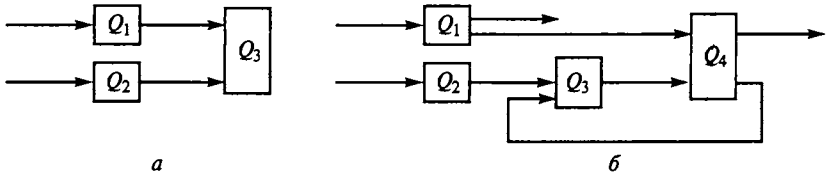
Проблемы синтеза автоматов аналогичны проблемам, возникающим в современной алгебре при рассмотрении представления некоторой общей системы в виде отдельных более простых систем того же вида. На практике большое значение имеют задачи, связанные с необходимостью получить информацию о свойствах и состоянии автомата на данный момент времени, не «вскрывая» его, а только сравнивая результаты входного и выходного алфавита. Эта проблема составляет основу поиска методов технической диагностики конечного автомата, т.е. поиска его неисправностей. Порой, только с помощью прочтения или расшифровки «черного ящика» удается узнать о причинах гибели транспортных средств в экстремальных ситуациях.

Известны два способа структуризации автомата, его конкретизации:

• если абстрактные автоматы могут быть соединены в сети подключением выходов одних автоматов ко входам других, то решаются две задачи:

задача композиции — построение автомата, эквивалентного некоторой сети из исходных автоматов, например на рис. 7.3;

задача декомпозиции — представление автомата в виде сети из исходных по заданным свойствам или имеющим заданную **структуру**;



а *б*

Рис. 7.3. Схема автоматов:
а — простая; *б* — сложная

• если абстрактные автоматы имеют несколько каналов входа и выхода, но число возможных сигналов на каждом из них ограничено, то рассматривается структура входных и выходных сигналов.

Так как в основе изучаемых автоматов лежит двоичный код, то автоматы называют *двоичными* или *логическими*. Тогда и входной алфавит X, и выходной — Y представлены множеством {0, 1}.

В задаче 45 с кодовым замком существует пять двоичных входов ($X \in \{A, B, C, D, E\}$) и шесть возможных выходов: $Y = \{00000, 00001, 00010, 00100, 01000, 10000\}$.

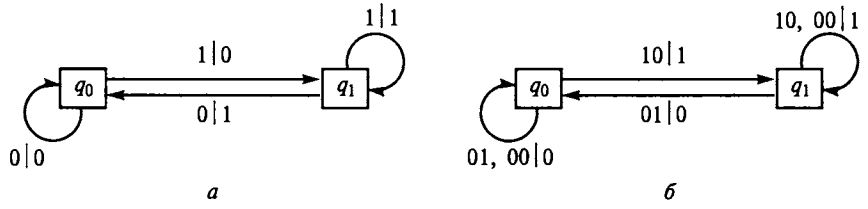
В результате такой структуризации получают *логическую схему* — сеть элементарных автоматов в виде двоичных автоматов с ограниченным числом входов, выходов и состояний.

При этом используются известные логические элементы с двумя входами: «логическое И», т.е. $z = x_1 \cdot x_2$; «логическое ИЛИ», т.е. $z = x_1 \vee x_2$; «2И—НЕ», т.е. $z = \overline{x_1 \cdot x_2}$, а также элементы с четырьмя входами: $z = \overline{x_1 x_2 x_3 x_4}$, $z = \overline{x_1 x_2 \vee x_3 x_4}$.

Кроме логических элементов в такой схеме используются *элементы памяти*: элемент задержки и двухвходовый триггер.

Элемент задержки представляет собой автомат, имеющий один двоичный вход и один двоичный выход, через который с определенной задержкой времени проходит сигнал, поступающий на вход. Его граф имеет вид, представленный на рис. 7.4, *а*.

Время схемы состояний такого элемента соответствует величине задержки времени.



а *б*

Рис. 7.4. Графы элементов памяти:
а — элемент задержки; *б* — двухвходовый триггер

Двухвходовым триггером называют автомат, имеющий два входа x_1 и x_2, который может также находиться в двух состояниях Q_0 и Q_1. Его граф имеет вид, изображенный на рис. 7.4, *б*.

Автомат, получив на входе x_1 сигнал 1, перешел из состояния Q_0 в состояние Q_1 (запись) и на выходе выдает сигнал 1. Получив сигнал 1 на входе x_2, автомат переходит из состояния Q_1 в состояние Q_0 (сброс), а на выходе при этом выдается сигнал 0, причем одновременная подача единиц на оба входа x_1 и x_2 невозможна.

Имея минимальный набор элементов, например И, ИЛИ, НЕ, а также элементы памяти — задержку и триггеры, можно построить любую логическую схему, а значит, и любой конечный автомат.

Проектируя автомат как техническое устройство, мы осуществляем процесс перехода от описания его свойств к синтезу логической схемы, соответствующей заданным свойствам. Причем в теории автоматов разработаны специальные алгоритмические методы и программы для автоматизации процесса синтеза логических схем.

Такой способ синтеза логических схем называют **аппаратным**, так как его результатом является создание конкретной электронной аппаратуры.

Существует также **программный** способ синтеза логических схем, по которому создается не сама логическая схема, а программа, ей соответствующая.

Очевидно, конечный автомат является частным случаем алгоритма, в частности алгоритма с конечной памятью, что соответствует конечному числу состояний автомата.

? Какие перспективы и применения имеет теория конечных автоматов?

Благодаря исследованиям в области теории автоматов математика пополнилась новым направлением, связанным с понятием сложности некоторой системы.

Выдающийся американский математик XX в., венгр по происхождению, Джон фон Нейман использовал язык автоматов для описания основных законов взаимодействия сложных систем, т. е. как *метаязык* кибернетики. Впоследствии М. Л. Цетлин, исследуя задачи целесообразности поведения взаимодействующих объектов, рассматривал их как задачи коллективного поведения автоматов. В своих исследованиях для описания поведения любых объектов, в том числе и живых, он использовал термины языка конечных автоматов.

Новые научные направления «теория коллективного поведения автоматов» и «поведение автоматов в случайных средах» опираются на тактику поведения обычных конечных автоматов. Изучены условия, характеризующие их способность к адаптации в стационарных и нестационарных средах.

Причем оказалось, что отдельный вид автоматов — автоматы с переменной структурой в нестационарных средах — более приспособлены к функционированию. Автоматы такого типа используют в системах автоматического управления стационарными и нестационарными объектами, когда информация о функционировании объекта известна частично или вообще не известна.

Благодаря широкому развитию вычислительной техники и глобальной информатизации современного общества особую актуальность приобрела проблема машинного интеллекта. Например, исследования в этом направлении английского математика А. М. Тьюринга показали, что автомат может решать любые логические правильно сформулированные задачи, которые может решать человек.

? Может ли автомат заменить человека?

Очевидны преимущества автомата в скорости, точности и надежности выполнения различных механических операций. С каждым годом мы узнаем о прорывах в области искусственного интеллекта. Писатели-фантасты давно предсказали глобальные войны между машинным интеллектом и его родителем — человеческим сообществом.

? Насколько реальны предсказания фантастов о грядущем вытеснении человека автоматами?

Ведь не секрет, что многие предсказания писателей-фантастов сбываются. Так, известны данные о результативности писательских прогнозов на конец XX в. (табл. 7.5).

? Какие предсказания научной фантастики, связанные с вычислительными машинами, искусственным интеллектом и другими современными проблемами века информации, оказались решенными к настоящему времени?

Таблица 7.5

Судьба фантастических идей

Авторы научно-фантастических произведений	Общее количество независимых фантастических идей (на период написания)	Судьба фантастических идей					
		Сбылись или сбудутся в ближайшее время		Подтвердилась принципиальная осуществимость		Оказались ошибочными или неосуществимыми	
		Количество	%	Количество	%	Количество	%
Жюль Верн	108	64	59	34	32	10	9
Герберт Уэллс	86	57	66	20	23	9	11
Александр Беляев	50	21	42	26	52	3	6

Упражнения

7.1. Опишите работу кодового замка, состоящего из пяти последовательно нажимаемых кнопок, который открывается при нажатии двух кнопок:

а) *D* и *E*; б) *A* и *C*; в) *C* и *E*; г) *A* и *B*; д) *B* и *C*; е) *A* и *D*.

7.2. Опишите работу кодового замка из задачи 45, если считать уже нажатую клавишу поступающим входным сигналом. Составьте таблицу переходов и нарисуйте соответствующий граф.

7.3. Для ответа на вопросы о перспективах искусственного интеллекта сделайте сравнительный анализ возможностей человека и автомата современного поколения по направлениям:

- надежность конструкции:
  - температура среды обитания;
  - чувствительность к радиоактивному излучению;
  - приспособленность органов чувств к обнаружению механических или электромагнитных колебаний;
  - подверженность усталости;
- обеспечение функциональной надежности (как решается проблема повреждений системы или выхода из строя отдельных ее элементов):
  - скорость восприятия реакции;
  - возможности обеспечения информацией;
  - проблема распознания образов;
- возможности правильной идентификации информации в связи с возможным выходом «из образа»;
- сравнение интеллектуальных функций:
  - обработка информации;
  - рефлексия и другие психические реакции;
  - творческая деятельность.

ЗАКЛЮЧЕНИЕ

Нам тайны нераскрытые раскрыть пора, —
Лежат без пользы тайны, как в копилке.
Мы тайны эти с корнем вырвем у ядра,
На волю пустим джина из бутылки!

В. Высоцкий

Подходит к концу наше знакомство с этой удивительной наукой — дискретной математикой, наукой старой и совсем молодой, такой непохожей на классическую математику. Если зрелость науки определять годами ее развития, то дискретная математика окажется одной из самых юных. Но ведь это означает, что она продолжает развиваться, и рано еще писать заключение и ставить точку. Кто знает, может, кто-то из читателей нашей книги успеет внести свой вклад в развитие этой науки.

Еще столетие назад, когда Д. Гильберт выступил со своим знаменитым докладом «Математические проблемы» на II Международном математическом конгрессе в Париже (1900), ни один из его участников не знал такой науки. Удивительные вопросы прозвучали в этом выступлении.

1. Кто из нас не хотел бы приоткрыть завесу, за которой наше будущее, чтобы хоть одним взглядом проникнуть в предстоящие успехи нашего знания и тайны его развития в ближайшие столетия?

2. Какие новые методы и новые факты будут открыты в новом столетии на широком и богатом поле математической мысли?

Для того «чтобы представить себе возможный характер развития математического знания в ближайшем будущем», Гильберт сформулировал наиболее актуальные, на его взгляд, проблемы, которые необходимо было решить в ближайшее время, т.е. в XX в. Он взял на себя смелость выдвинуть гипотезы дальнейшего развития математики, сформулировал двадцать три проблемы этой науки.

? Насколько предвидение Д. Гильберта оказалось выполненным?

Познакомимся с некоторыми проблемами, сформулированными на том памятном конгрессе, с тем, как они решались в XX в. и какое отношение они имеют к дискретной математике. В скобках будут указаны соответствующие разделы учебника.

1. В этом списке первой стояла знаменитая проблема континуума. Как мы знаем из теории множеств, еще Г. Кантор пытался ее разрешить. Только в 1963 г. математик П. Коэн доказал, что континуум-гипотеза есть утверждение, не зависимое от системы аксиом теории множеств. Обратите внимание: на такой исход указал в 1940 г. К. Гёдель (подразд. 1.4, пп. 4.8.3, 5.6.3).

2. Шестая проблема была связана с аксиоматизацией физики и теории вероятностей. Аксиоматизация теории вероятностей была осуществлена А. Н. Колмогоровым в 1933 г., что послужило толчком для ее бурного развития и применения во второй половине XX в. (п. 5.6.5).

3. Восьмая проблема, связанная с задачами теории простых чисел, включая проблему Гольдбаха, до настоящего времени не решена (подразд. 5.5, п. 5.6.4, упр. к гл. 5).

4. Тринадцатая проблема, связанная с гипотезой о невозможности решения алгебраического уравнения седьмой степени в общем случае с помощью суперпозиции непрерывных функций только двух переменных, опровергнута нашими соотечественниками А. Н. Колмогоровым и В. И. Арнольдом в 1956—1957 гг. (подразд. 1.3).

5. Последняя, двадцать третья проблема касалась развития методов вариационного исчисления и в XX в. сформировалась в отдельную, имеющую большое практическое применение в современной жизни, математическую дисциплину.

Не все проблемы, поставленные Гильбертом, оказались актуальными в XX в., не все из них удалось решить или опровергнуть. Но для нас важно, что часть проблем была связана с наукой «дискретная математика», которая четко обозначилась лишь к концу XX в.

? В настоящее время, в начале не только столетия, а тысячелетия, математики ставят перед собой те же вопросы, что и Гильберт 100 лет назад.

Могут ли они увидеть векторы развития науки в новом тысячелетии?

Смогут ли они найти решения еще не решенных старых и новых проблем?

Могут ли математики прогнозировать будущее развитие науки, как это зачастую делают писатели-фантасты?

Вы, наверное, обратили внимание, что в нашей книге постоянно переплетаются три ветви: математическая, литературная и историческая. Мы убеждены, что разделение на математический и гуманитарный склад ума хоть и объяснимо, но нежелательно. Оно порождает интеллектуально бедных людей. Помните, как у К. Пруткова: «Специалист подобен флюсу: полнота его одностороння». Ведь не случайно во времена декабристов — глубоко образованных, всесторонне развитых интеллектуалов XIX в. — была своего рода мода на математику. И хотя не каждому из них давалась эта наука, считалось признаком хорошего тона разбираться в ней.

Наиболее ярким примером может служить А. С. Пушкин, который с детства не имел способностей к математике, но уже будучи известным поэтом и издателем литературного журнала «Современник», в одном из первых его номеров опубликовал статью князя П. Б. Козловского «Разбор Парижского математического еженедельника», а в третьем номере — его статью по теории вероятности. Этот удивительный исторический факт объясняется словами

из стихотворения А. С. Пушкина «Чаадаеву» (1821), где высказано желание «в просвещении стать с веком наравне».

И, наверное, теперь уже не очень удивит такое хронологическое совпадение. Известно, что в 1827 г. А. С. Пушкин высказал мысль о том, что в геометрии воображение нужно так же, как и в поэзии. И приблизительно в тот же период, 24 февраля 1826 г., великий и не признанный современниками математик Н. И. Лобачевский сделал свой знаменитый доклад о «воображаемой геометрии», которую впоследствии назвали его именем. По-видимому, и вам при изучении такой абстрактной науки — дискретной математики — не раз приходилось подключать свое воображение.

Итак, математическая культура — та грань культуры, без которой невозможно стать образованным человеком. Но это всего лишь одна из граней общечеловеческой культуры. А среди других граней — знание истории цивилизации, истории открытий, по которым, как по ступенькам, человечество передвигается в будущее.

Обычно мы узнаем о мире из книг, например из учебников по отдельным предметам, и смотрим на мир как через отдельные окошки. И наверное, поэтому удается увидеть в этих окошках не все грани культуры в их многообразии, а лишь отдельные из них. При таком знакомстве теряется целостность восприятия мира. В учебнике была предпринята попытка показать смешанные вопросы, пересекающие проблемы математики, истории математики и одновременно культуры в целом, и поэтому назван учебник междисциплинарным.

И не важно, гуманитарии или математики читали его. Надеемся, каждый смог найти в этом учебнике и то, что ему было понятно и интересно, и то, к чему пока, быть может, интерес явно не проявился.

Возможно, после знакомства с такой математикой одним она станет ближе, а другие заинтересуются историей или литературой. Главное, чтобы желание задавать вопросы и узнавать новое, такое естественное в детском возрасте, не угасло с годами. А способность задавать вопросы и искать ответы на них характерна для любознательного, интеллектуального общества, от которого зависит будущее всех нас, будущее нашей страны, будущее цивилизации.

? И, наконец, зададим себе последние в этой книге вопросы.

Какие новые открытия ждут нас в будущем?

Какова роль наших соотечественников в продвижении цивилизации, в ее развитии и совершенствовании?

Какое место вы отводите себе в историческом стремлении к прогрессу?

Что лично для вас означает фраза «Если не я, то кто же?»

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Автоматы 341
- асинхронные 344
- вероятностные 345
- комбинационные 346
- Мили 346
- Мура 346
- синхронные 344

Аксиома 211, 258, 268
- независимости 211
- непротиворечивости 213
- полноты 214

Алгебра
- булева 131, 153
- высказываний 140, 153, 219
- вычетов 329

Алгоритм 342
Алфавит 35, 300, 309
- формальной теории 210
- логики высказываний 219

Аналогия 259

Базис 196, 214
- Жегалкина 199
- индукции 115

Биекция 23, 45
Булеан 31
Буква 35
Булева функция 133

Вершина (узел) графа 70
- висячая 71
- изолированная 71
- нечетная 71
- четная 71

Вес разряда 297
Вывод 211
Выводимость 211, 220, 221
- непосредственная 211

Выборка 276, 278
Высказывание 132, 219, 226
- обратное 148
- противоположное 148, 239

Вычет по модулю 329

Генеральная совокупность 278
Гипотеза 211, 281
Граф 69
- взвешенный (см. Сеть) 89
- гамильтонов 77
- несвязный 73
- ориентированный 72
- планарный 75
- плоский 75
- полный 71
- связный 73
- эйлеров 77

Двойственность 194
Дедукция 243—245, 258
Декартово (прямое) произведение 36
Декодирование 289
Деление
- нацело 288, 327
- понятия 120
- с остатком 327

Дерево 80
- бинарное 82
- ориентированное 92

Дешифрование 289, 313
Диапазон представления чисел 297
Дизъюнкция 137, 222, 228
- элементарная 171

Дискретный (сигнал) 300, 345
Длина
- маршрута 72
- разрядной сетки 297
- числа 297

Дополнение
- графа 71
- множества 18

Закон
- склеивания 181
- Де Моргана 152

– достаточного основания 163
– идемпотентности 152
– инверсии (отрицания) 142, 152
– исключенного третьего 161
– противоречия 159
– тождества 157

Измерение информации 301
Изоморфизм 38
– графов 75, 84
Импликация 138, 228
Инверсия 58
Индукция 262
Интерпретация 212
– исчисления предикатов 224
– формул 212
– формальной теории 212
Инцидентность 70, 85
Инъекция 23
Истинность (семантическая характеристика) 105
Исчисление
– высказываний 219
– предикатов 224

Канал 319
Карты Карно 180
Квантификация 233
Кванторы 16, 231
– двойственные 232
– общности 231
– существования 231
Класс
– линейных функций 194
– монотонных функций 195
– самодвойственных функций 194
– функций, сохраняющих константу 1 194
– функций, сохраняющих константу 0 194
– эквивалентности 42
Классификация 123
Ключ 313, 333
– открытый 334
– закрытый 314
Код 309
– Хемминга 322
– сообщения 309
Кодерево 82
Кодирование 289
– алфавитное 309, 310
– двоичное 311
– оптимальное 311
– помехоустойчивое 317, 319
– с минимальной избыточностью 311, 318
– самокорректирующееся 320
– эффективное 314
Количество информации 301
Комбинаторика 45
Компонента связности 74, 80
Константа 225
Конъюнкция 137, 142, 222, 227
– элементарная 171
Корень дерева 80
Кортеж 34
Кратность ребра 70
Криптография 292

Лес 82
Лист дерева 82
Логическая
– аксиома 268
– связка 141, 227
– схема 175
– эквивалентность 220
Логическое следствие 220

Маршрут 72
Матрица
– инцидентности 85
– смежности 85
Метатеория 212
Метауровень 209
Метаязык 212
Метод резолюций 219, 282
Множество 14
– бесконечное 29
– истинности 226
– конечное 29
– несчетное 29
– пустое 15
– счетное 29
– универсальное 16
– упорядоченное 34
Модель 280
– формальной теории 211
Мост 74
Мощность множества 17

Непротиворечивость 213
Нуль-граф 71
Нормальные формы 170
– дизъюнктивные (ДНФ) 171

– совершенные 171
– конъюнктивные (КНФ) 171

Область
– действия квантора 231
– значений функции 22
– определения функции 22
Обобщение понятия 107, 110
Образ 20
Объединение
– графов 79
– множеств 18
Объем понятия 108
Ограничение понятия 109
Операции
– исчисления высказываний 219
– предикатов 227
– над множествами 19
– высказываниями 141
– обращения 247
– превращения 247
– противопоставления предикату 247
Остов 82
Отношение 38
– бинарное 39
– на множестве 38
– обратное 39
– порядка 43
– рефлективное 40
– симметричное 40
– толерантности 42
– транзитивное 40
– эквивалентности 40, 42
Отображение множеств (функция) 22
– единичное (тождественное) 27

Парадокс
– теории множеств 32
Переменная
– высказывательная 132
– пропозиционная 226
– свободная 231
– связанная 231
– существенная 139
– фиктивная 139
Пересечение множеств 18, 19
Перестановка 34
Петля 70
Подграф 80
Поддерево 82
Помехи 317
Подмножество 16
Подстановка 55
Позиционная система счисления 296
Поиск бинарный 94
Полином Жегалкина 192
Полнота 214
– системы функций 214
– формальной теории 214
Понятие 105
Порядок
– частичный 44
Постфикс 309
Посылка 211, 244
Потомок узла 82
Правило
– ∃-введения (*eg*-правило) 236
– ∀-введения (*ug*-правило) 236
– вывода 211, 268
– подстановки 221
– ∃-удаления (*es*-правило) 237
– ∀-удаления (*us*-правило) 237
– *modus ponens* 219
Предикат 225
Представление 38
Префикс 309
Произведение
– логическое (конъюнкция) 137
– множеств (см. Декартово) 36
– функций (см. Суперпозиция) 25
– чисел 27, 269
Путь
– гамильтонов 77
– эйлеров 77

Равенство
– множеств 17
– кортежей 34
– функций 23, 134
Равномощность (эквивалентность) множеств 24, 29
Разбиение на классы 20
Разложение функций по переменным 170, 175
Размещения 48
Разность множеств 18
Разрешимость формальной теории 218
Разряд 296
Расстояние
– между вершинами графа 73
– кодовое 321
– Хемминга 321
Ребро (дуга)

– графа 70
– кратное 70
Рефлективность 40

Связность 40
– вершин 73
– графа 74
Связка логическая 141
Свойство
– информации 301
– характеристическое 15
Сеть 89
Сигнал 300
Силлогизм 249
Система
– счисления 295
– двоичная 297
– восьмеричная 297
– десятичная 295
– шестнадцатеричная 299
Симметрическая разность 18
Симметричность 40
Слово 300
– над алфавитом (множеством) 35, 210
– пустое 309
Смежность
– вершин 70
– ребер 70
Содержание понятия 108
Соединение графов (кольцевая сумма) 80
Сообщение 309
– шифрованное 291
Соответствие множеств 20
– взаимно-однозначное 23
– обратное 21
Сочетания 48
Способы задания
– конечных автоматов 347
Сравнение по модулю 328
Степень вершины 70
– входа 72
– выхода 72
Стрелка Пирса 137
Суждения 131
– вероятностные 244
– достоверные 244
Сумма по модулю два 137, 187
Суперпозиция (композиция) функций 25
Схема
– аксиом 269
– логическая 175
– правил вывода 269
– комбинационная 176
Сын узла 82
Сюръекция 23

Таблица истинности 137
Тавтология 155
Тайнопись 290
Теорема 211
– Поста—Яблонского 196
– Эйлера 274
Теория 212
– семантически непротиворечивая 213
– формальная 211
– формальной арифметики 267
Терм 268
Транзитивность 41
Триггер 355

Умозаключение 244
– дедуктивное 244
– индуктивное 244
– традуктивное (аналогия) 259
Уровень узла 82
Условие
– необходимое 147
– достаточное 147
– эйлеровости графа 77

Фаза доказательства 273
Форма нормальная 170
Формула 135, 210, 268
– двойственная 194
– истинная 155
– ложная 155
– общезначимая 213, 241
– невыполнимая 213
– формальной теории 211
– Хартли 302
– Шеннона 302
Фрейм 90
Функциональная замкнутость 193
Функциональная полнота 196
Функция 45
– булева 133
– выхода автомата 346
– двойственная 194
– монотонная 195
– обратная 23

– перехода автомата 346
– самодвойственная 194

Цепь 73
– простая 73
– эйлерова 77
Цикл 73
– простой 73
– эйлеров 77
Цифра 296
Цифровая подпись 333

Числа взаимно простые 288, 333
Число
– инверсий 58
– перестановок 47
– простое 334
– размещений 48
– сочетаний 48

Шифр
– Бэкона 291
– надежный 291
– с открытым ключом 334, 339
Шифрование 289
Штрих Шеффера 138
Шум 317

Эквиваленция 137, 142, 150, 229
Энтимема 250
Энтропия 301
Электронная подпись 333
Элемент
– множества 14

Язык
– логики 131
– высказываний 219
– предикатов 225
– теории графов 69
– теории множеств 14
– формальной теории 211
Ярус дерева 80

СПИСОК ЛИТЕРАТУРЫ

1. *Бауэр Ф.Л., Гооз Г.* Информатика. — М.: Мир, 1990.

2. *Белов В.В., Воробьев Е.М., Шаталов В.Е.* Теория графов. — М.: Высш. шк., 1976.

3. *Ганчев И., Чимев К., Стоянов Й.* Математический фольклор. — М.: Знание, 1987.

4. *Козлова Е.Г.* Сказки и подсказки. — М.: Мирос, 1994.

5. *Кузнецов О.П., Адельсон-Вельский Г.М.* Дискретная математика для инженеров. — М.: Энергия, 1980.

6. *Кэрролл Л.* Алиса в Стране чудес. — М.: Лорис, 1993.

7. *Мадер В.В.* Школьнику об алгебре логики. — М.: Просвещение, 1993.

8. *Мендельсон Э.* Введение в математическую логику. — М.: Наука, 1984.

9. *Новиков Ф.А.* Дискретная математика для программистов. — СПб., 2001.

10. *Оре О.* Теория графов. — М.: Наука, 1968.

11. *Пахомов Ю.В.* Логика естествознания. — М.: Космополис, 1994.

12. *Писаренко И.Б.* Что такое абсолютно надежный шифр? — М.: Квант, 1992. — № 6.

13. *Пона Д.* Математика и правдоподобные рассуждения. — М.: Наука, 1975.

14. *Свифт Дж.* Путешествие Гулливера. — М.: Правда, 1987.

15. *Фрейденталь Х.* Язык логики. — М.: Наука, 1969.

16. *Черкасов О.Ю.* Задачи по математике. — М.: Моск. лицей, 1994.

17. *Эрдниев М.П., Эрдниев Б.П.* Укрупнение дидактических единиц как новая технология обучения математике. — М.: Просвещение, 1986.

18. *Яблонский С.В.* Введение в дискретную математику. — М.: Высш. шк., 2002.

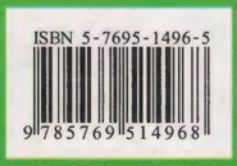
ОГЛАВЛЕНИЕ

Предисловие ... 3
Перечень математических символов и сокращений ... 6
Введение ... 8
Глава 1. **Множества** ... 14
1.1. Общие понятия теории множеств ... 14
1.2. Основные операции над множествами ... 17
1.3. Соответствия между множествами. Отображения ... 20
1.4. Классификация множеств. Мощность множества ... 28
1.5. Кортежи. Декартовы произведения ... 34
1.6 Отношения. Бинарные отношения и их свойства ... 38
1.7. Элементы комбинаторики ... 45
1.8. Подстановки ... 55
Упражнения ... 61
Глава 2. **Графы** ... 69
2.1. Основные понятия и определения графа и его элементов ... 69
2.2. Операции над графами ... 79
2.3. Деревья. Лес. Бинарные деревья ... 80
2.4. Способы задания графа. Изоморфные графы ... 84
2.5. Сети. Сетевые модели представления информации ... 89
2.6. Применение графов и сетей ... 91
Упражнения ... 96
Глава 3. **Понятия** ... 104
3.1. Понятие как форма мышления ... 104
3.2. Логические операции над понятиями: обобщение и ограничение понятий ... 109
3.3. Отношения между понятиями ... 111
3.4. Операции над понятиями. Определение понятий ... 114
3.5. Деление понятий. Классификация ... 120
Упражнения ... 127
Глава 4. **Математическая логика** ... 131
4.1. Суждения как форма мышления. Простые высказывания ... 131
4.2. Булевы функции ... 133
4.3. Сложные высказывания ... 140
4.3.1. Операции над сложными высказываниями ... 141
4.3.2. Необходимое и достаточное условия импликации ... 145
4.3.3. Формулы алгебры логики ... 152
4.4. Законы правильного мышления ... 156
4.5. Логика вопросов и ответов ... 166
4.6. Минимизация булевых функций ... 170
4.6.1. Разложение функций по переменным. Нормальные формы ... 170
4.6.2. Логические схемы ... 175
4.6.3. Карты Карно ... 180

4.7. Сумма по модулю два ... 187
4.8. Полином Жегалкина. Функционально замкнутые классы ... 192
4.8.1. Канонический полином Жегалкина ... 192
4.8.2. Функциональная замкнутость ... 193
4.8.3. Функционально полные системы функций ... 196
Упражнения ... 199
Глава 5. **Формальные системы и умозаключения. Логика предикатов** ... 208
5.1. Формальные системы ... 208
5.2. Исчисление высказываний ... 219
5.3. Логика предикатов ... 224
5.4. Умозаключения как форма мышления. Дедуктивные умозаключения и их виды ... 243
5.4.1. Непосредственные умозаключения по логическому квадрату ... 245
5.4.2. Простые категорические силлогизмы ... 249
5.4.3. Энтимемы ... 250
5.4.4. Умозаключения из сложных суждений ... 251
5.4.5. Применение аппарата алгебры высказываний для работы с умозаключениями ... 254
5.5. Методы научного познания ... 258
5.6. Индуктивные умозаключения и их виды ... 262
5.6.1. Виды индукции ... 263
5.6.2. Методы установления причинных связей ... 265
5.6.3. Формальная аксиоматическая теория для арифметики натуральных чисел ... 267
5.6.4. Метод математической индукции ... 270
5.6.5. Статистические обобщения ... 276
5.7. Виды аналогии. Моделирование как метод ... 278
5.8. Гипотезы ... 281
Упражнения ... 284
Глава 6. **Элементы теории и практики кодирования** ... 289
6.1. История кодирования от древности до наших дней. Защита информации ... 289
6.2. Системы счисления для представления информации в ЭВМ ... 295
6.3. Основные понятия вероятностной теории информации ... 300
6.4. Обработка сообщений как кодирование ... 309
6.5. Кодирование информации как средство обеспечения контроля работы автомата ... 312
6.6. Основы алгебры вычетов и их приложение к простейшим криптографическим шифрам ... 327
Упражнения ... 336
Глава 7. **Конечные автоматы** ... 341
7.1. Определение конечных автоматов ... 341
7.2. Способы задания конечных автоматов ... 347
7.3. Общие задачи теории автоматов ... 351
Упражнения ... 357
Заключение ... 358
Предметный указатель ... 361
Список литературы ... 366

ДИСКРЕТНАЯ МАТЕМАТИКА

ISBN 5-7695-1496-5



Издательский центр
«Академия»