

входит в УМК по информатике для старшей школы (10–11 классы).

соответствует федеральному государственному образовательному стандарту
общего (полного) образования (2012 г.).

входит в федеральный перечень учебников, рекомендованных Министерством
образования и науки Российской Федерации.



Константин Юрьевич Поляков — доктор технических наук, профессор кафедры судовой автоматики и измерений Санкт-Петербургского государственного морского технического университета, учитель информатики школы № 163 Санкт-Петербурга. Победитель Всероссийского конкурса для педагогов по включению ресурсов Единой коллекции цифровых образовательных ресурсов в образовательный процесс. Лауреат профессиональной премии «Лучший учитель Санкт-Петербурга». Награжден знаком «Почетный работник общего образования РФ». Ведущий автор и редактор авторского коллектива по разработке комплекта учебников по информатике углубленного уровня.



Евгений Александрович Еремин — кандидат физико-математических наук. Начиная с 1974 года, работает с вычислительной техникой. Преподает связанные с информатикой курсы с момента появления этого предмета в школе в 1985 году. Хорошо известен по публикациям в газете «Информатика» издательства «Первое сентября» (около 50 статей). Всего имеет более 150 публикаций, в том числе несколько книг.

5-9963-3283-0



К.Ю. Поляков
Е.А. Еремин

ИНФОРМАТИКА 10 (1)



ФГОС

10



К.Ю. Поляков
Е.А. Еремин

ИНФОРМАТИКА

1

УГЛУБЛЕННЫЙ УРОВЕНЬ



ИЗДАТЕЛЬСТВО
БИНОМ

УДК 004.9
ББК 32.97
П54

Поляков К. Ю.

П54 Информатика. Углубленный уровень : учебник для 10 класса : в 2 ч. Ч. 1 / К. Ю. Поляков, Е. А. Еремин. — 6-е изд., стереотип. — М. : БИНОМ. Лаборатория знаний, 2017. — 344 с. : ил.

ISBN 978-5-9963-3283-0 (Ч. 1)

ISBN 978-5-9963-3285-4

Учебник предназначен для изучения информатики на углубленном уровне в 10 классах общеобразовательных организаций. Содержание учебника опирается на изученный в 7–9 классах курс информатики для основной школы.

Рассматриваются теоретические основы информатики, аппаратное и программное обеспечение компьютера, компьютерные сети, алгоритмизация и программирование, информационная безопасность.

Учебник входит в учебно-методический комплект (УМК), включающий также учебник для 11 класса и компьютерный практикум.

Предполагается широкое использование ресурсов портала Федерального центра электронных образовательных ресурсов (<http://fcior.edu.ru/>).

Соответствует федеральному государственному образовательному стандарту среднего общего образования (2012 г.).

УДК 004.9
ББК 32.97

Учебное издание

Поляков Константин Юрьевич
Еремин Евгений Александрович

ИНФОРМАТИКА. УГЛУБЛЕННЫЙ УРОВЕНЬ

Учебник для 10 класса
В 2 частях
Часть 1

Ведущий редактор О. Полежаева
Художник Н. Новак
Технический редактор Е. Денюкова
Корректор Е. Климина
Компьютерная верстка: Л. Катуркина

Подписано в печать 25.04.17. Формат 70х100/16. Усл. печ. л. 27,95.
Тираж 31 000 экз. Заказ № 40109.

ООО «БИНОМ. Лаборатория знаний»
127473, Москва, ул. Краснопролетарская, д. 16, стр. 1,
тел. (495)181-53-44, e-mail: binom@lbz.ru
<http://www.lbz.ru>, <http://metodist.lbz.ru>

Отпечатано в соответствии с качеством предоставленных издательством
электронных носителей в АО «Саратовский полиграфкомбинат».
410004, г. Саратов, ул. Чернышевского, 59. www.sarpk.ru

Оглавление

От авторов	5
Навигационные значки	8
Глава 1. Информация и информационные процессы	9
§ 1. Информатика и информация	9
§ 2. Что можно делать с информацией?	19
§ 3. Измерение информации	25
§ 4. Структура информации	31
Глава 2. Кодирование информации.	55
§ 5. Язык и алфавит	55
§ 6. Кодирование	60
§ 7. Дискретность.	76
§ 8. Алфавитный подход к измерению количества информации	84
§ 9. Системы счисления	88
§ 10. Позиционные системы счисления	91
§ 11. Двоичная система счисления	102
§ 12. Восьмеричная система счисления	109
§ 13. Шестнадцатеричная система счисления	114
§ 14. Другие системы счисления	118
§ 15. Кодирование символов	122
§ 16. Кодирование графической информации	127
§ 17. Кодирование звуковой и видеoinформации	146
Глава 3. Логические основы компьютеров	159
§ 18. Логика и компьютер	159

ISBN 978-5-9963-3283-0 (Ч. 1)
ISBN 978-5-9963-3285-4

© ООО «БИНОМ. Лаборатория знаний», 2013

§ 19. Логические операции	161
§ 20. Диаграммы Венна	180
§ 21. Упрощение логических выражений	185
§ 22. Синтез логических выражений.	192
§ 23. Предикаты и кванторы	196
§ 24. Логические элементы компьютера.	201
§ 25. Логические задачи	210
Глава 4. Компьютерная арифметика	221
§ 26. Особенности представления чисел в компьютере	221
§ 27. Хранение в памяти целых чисел	227
§ 28. Операции с целыми числами	235
§ 29. Хранение в памяти вещественных чисел	252
§ 30. Операции с вещественными числами	261
Глава 5. Как устроен компьютер.	266
§ 31. История развития вычислительной техники.	267
§ 32. Принципы устройства компьютера	283
§ 33. Магистрально-модульная организация компьютера	294
§ 34. Процессор	301
§ 35. Память.	309
§ 36. Устройства ввода	324
§ 37. Устройства вывода	333

От авторов

Вы держите в руках учебник информатики углублённого уровня для 10 класса. Если посмотреть на оглавление, может показаться, что многое из представленного в учебнике материала вам уже знакомо. Действительно, в 7–9 классах вы знакомились с понятиями «информация» и «информационный процесс», изучали кодирование данных, программное обеспечение и основы программирования.

Однако углублённый уровень изучения предполагает, что вы уже начинаете готовиться к освоению будущей профессии. Для этого необходимо более глубокое понимание всех этих вопросов, выход на следующий уровень владения материалом, когда человек может не только воспроизводить полученные знания, но и решать новые сложные задачи с их помощью. Цель этого учебника — дать такие знания, которые позволят вам грамотно решать задачи, не рассмотренные в самом учебнике.

Первые две главы учебника — базовые, в них содержится информация, необходимая для понимания всех последующих глав. Большинство остальных глав можно изучать в разном порядке, они относительно независимы друг от друга.

В углублённом курсе информатики самое серьёзное внимание уделяется разделу «Алгоритмизация и программирование». Авторы выбрали два языка программирования, которые изначально задумывались как учебные языки:

- алгоритмический язык свободно распространяемой системы КуМир (<http://lpm.org.ru/kumir2/>, <http://kpolyakov.spb.ru/school/kumir.htm>);
- язык Паскаль, который изучается во многих школах России; в качестве среды программирования можно использовать АЛГО (<http://petriv.ho.com.ua/algo/rus>), PascalABC.NET (<http://pascalabc.net>), Free Pascal (<http://www.freepascal.org>).

Вы можете использовать любой из этих языков, потому что практически все программы и примеры представлены в двух вариантах.

В конце каждого параграфа есть контрольные вопросы, которые помогут понять, хорошо ли вы разобрались в материале. В тексте нет прямых ответов на некоторые вопросы, но есть вся необходимая информация для ответа на них.

Задачи в конце параграфов помогут закрепить материал на практических работах. Самые сложные задачи (на взгляд авторов) отмечены звёздочкой (*).

Мы старались сделать так, чтобы содержание учебника как можно меньше зависело от программного обеспечения, установленного на ваших компьютерах. Весь курс можно успешно изучать, используя только свободное программное обеспечение (СПО) — операционную систему *Linux*, офисный пакет *OpenOffice.org* или его модификации (например, *LibreOffice*), компилятор *FreePascal* и др.

В заключение нам хочется поблагодарить наших коллег, которые взяли на себя труд прочитать предварительные версии отдельных глав учебника и высказать множество полезных замечаний, позволивших сделать учебник более точным, ясным и понятным:

- *А. П. Шестакова*, кандидата педагогических наук, заведующего кафедрой информатики и вычислительной техники Пермского государственного педагогического университета, который вдохновил авторов на написание этого учебника;
- *М. А. Ройтберга*, доктора физико-математических наук, заведующего лабораторией прикладной математики Института математических проблем биологии РАН, г. Пущино;
- *С. С. Михалковича*, кандидата физико-математических наук, доцента кафедры алгебры и дискретной математики Южного федерального университета, г. Ростов-на-Дону;
- *О. А. Тузову*, учителя информатики школы № 550, г. Санкт-Петербург;
- *А. Г. Тамаревскую*, учителя информатики лицея № 533, г. Санкт-Петербург;
- *Н. Д. Шумилину*, кандидата педагогических наук, учителя информатики Тверской гимназии № 6, г. Тверь;

- *Л. Б. Кулагину*, учителя информатики ФМЛ № 239, г. Санкт-Петербург;
- *В. Н. Разумова*, учителя информатики МОУ «Большеелховская средняя общеобразовательная школа», с. Большая Елховка, республика Мордовия;
- *Ю. М. Розенфарба*, учителя информатики МОУ Межозёрная СОШ, Челябинская область;
- *Т. А. Мисаренкова*, учителя информатики школы № 163, г. Санкт-Петербург;
- *В. В. Потопахина*, педагога дополнительного образования Краевого центра технического творчества, г. Хабаровск;
- *К. А. Малеванова*, руководителя направления телематики ООО «ОБИТ».

Навигационные значки

Уважаемые ученики!

В работе с книгой вам помогут навигационные значки:



— важное утверждение или определение.



— вопросы и задания к параграфу.



— дополнительное разъяснение.



— задания для подготовки к итоговой аттестации.



— к каждой главе учебника рекомендуются

- 1) электронные образовательные ресурсы (ЭОР) с сайта Федерального центра образовательных ресурсов (ФЦИОР): <http://fcior.edu.ru>

Доступ к ЭОР из каталога ФЦИОР:

<http://fcior.edu.ru/catalog/meta/4/mc/discipline%2000/mi/4.06/p/page.html>, где ресурсы размещены в алфавитном порядке, согласно названиям учебных тем;

- 2) электронное приложение к УМК:

- практические, самостоятельные и контрольные работы на сайте поддержки учебника <http://kpolyakov.spb.ru/school/probook/prakt.htm>;
- презентации для проведения уроков на сайте поддержки учебника <http://kpolyakov.spb.ru/school/probook/slides.htm>



— Проектное или исследовательское задание.

В ходе выполнения проекта (исследования) определите вид, в котором будут представлены его результаты. Можно выбрать один (или несколько) вариант:

- подготовить набор полезных ссылок с использованием веб-ресурсов;
- подготовить небольшое выступление с использованием презентации (5–7 мин);
- оформить доклад и поместить его на сайт школьной конференции;
- подтвердить полученные результаты расчётами или графиками (диаграммами);
- подготовить видеоролик;
- разместить материалы проекта (исследования) в кол-лекции обучающих модулей по предмету на сайте школы.

Глава 1 Информация и информационные процессы

§ 1 Информатика и информация

Информатика

Задачи, связанные с хранением, передачей и обработкой информации, человеку приходилось решать во все времена: требовалось передавать знания из поколения в поколения, искать нужные книги в хранилищах, шифровать секретную переписку. К концу XIX века количество документов в библиотеках стало настолько велико, что возникла необходимость применить научный подход к задачам хранения и поиска накопленной информации. В это время зародилось новое научное направление, в котором изучалась *документальная* информация, т. е. информация в виде документов (книг, журналов, статей и т. п.). В английском языке оно получило название *information science* (информационная наука, наука об информации).

Применение компьютерной техники значительно увеличило возможности людей в области работы с информацией, позволив автоматизировать рутинную работу. Считается, что слово «**информатика**»¹ в современном значении образовано в результате объединения двух слов: «информация» и «автоматика». Таким образом, получается «автоматическая работа с информацией». В английском языке существует близкое по значению выражение *computer science* (наука о компьютерах).

Современная информатика, которая стала самостоятельной наукой в 70-х годах XX века, изучает теорию и практику обработ-

¹ Впервые этот термин использовал немецкий ученый К. Штейнбух в 1957 г. (нем. *Informatik*). В 1962 г. Ф. Дрейфус ввёл слово *informatique* во французский язык, затем оно было переведено на английский (англ. *informatics*).

ки информации с помощью компьютерных систем. Обычно к информатике относят следующие научные направления:

- **теоретическую информатику** (теорию информации, теорию кодирования, математическую логику, теорию автоматов и др.);
- **вычислительную технику** (устройство компьютеров и компьютерных сетей);
- **алгоритмизацию и программирование** (создание алгоритмов и программ);
- **прикладную информатику** (персональные компьютеры, прикладные программы, информационные системы и т. д.);
- **искусственный интеллект** (распознавание образов, понимание речи, машинный перевод, логические выводы, алгоритмы самообучения).

Раньше эти вопросы частично рассматривались в других науках — математике, лингвистике (науке о языке), электронике и др. В теоретической информатике затрагиваются некоторые вопросы, относящиеся к кибернетике — науке об управлении, и к теории систем. С появлением компьютеров стало ясно, что все эти направления тесно связаны, и постепенно начала формироваться новая область научной деятельности. Информатика — это область науки в процессе становления, и круг её вопросов в будущем может измениться.

В нашем курсе мы также познакомимся с *информационными технологиями*, которые связаны с применением компьютеров во всех областях современной жизни: при оформлении документов; при подготовке книг и журналов к печати; для расчёта зарплаты; для продажи билетов на поезда и самолёты; для автоматизации производства; при проектировании зданий, кораблей, станков и т. д. Во всех этих сферах используется понятие «*информация*».

Что такое информация?

Латинское слово *informatio* переводится как «разъяснение», «сведения». В быту под информацией мы обычно понимаем любые сведения или данные об окружающем нас мире и о нас самих. Однако дать общее определение информации весьма непросто. Более того, в каждой области знаний слово «*информация*» имеет свой смысл.

Философы говорят о том, что информация, как зеркало, отражает мир (реальный или вымышленный). Биологи рассматривают

информационные процессы в живой природе. Социологи изучают ценность и полезность информации в человеческом обществе. Специалисты по компьютерной технике в первую очередь интересуется представление информации в виде знаков.

Попробуем посмотреть на информацию с разных сторон и попытаться выявить некоторые её свойства.

Прежде всего информация «бестелесна», или *нематериальна*, она не имеет формы, размеров, массы. С этой точки зрения информация — это то содержание, которое человек с помощью своего сознания «выделяет» из окружающей среды.

Информация нематериальна.

В то же время для того, чтобы её хранить, обрабатывать и передавать, используется материальный носитель.

Давайте сравним два изображения одинакового размера (рис. 1.1). На первом из них пусто, а на втором мы видим фотографию. Вряд ли кто-то способен долго разглядывать чистый лист, а на фотографию можно долго смотреть, открывая всё новые и новые детали. Почему так?

Первый рисунок разглядывать неинтересно, там всё одинаково — везде белый цвет. На втором рисунке есть *разнообразие*, он неоднороден. Поэтому можно сказать, что он содержит больше информации, чем первый.

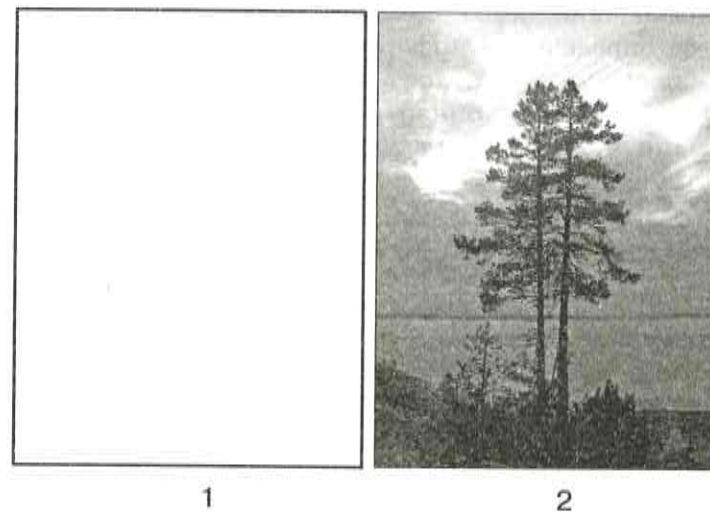


Рис. 1.1



Информация **характеризует разнообразие** (неоднородность) в окружающем мире.

Зачем вообще нам нужна информация? Дело в том, что наше знание всегда в чём-то неполно, в нём есть *неопределённость*. Например, вы стоите на остановке и не знаете, на каком именно автобусе вам нужно ехать в гости к другу (его адрес известен). Неопределённость мешает вам решить свою задачу. Нужный номер автобуса можно определить, например, по карте с маршрутами транспорта. Очевидно, что при этом вы получите новую информацию, которая увеличит знание и уменьшит неопределённость.



При получении новой информации **уменьшается неопределённость знания**.

Многие выдающиеся учёные XX века (Н. Винер, У. Эшби, К. Шеннон, А. Урсул, А. Моль, В. Глушков) давали своё определение информации, но ни одно из них не стало общепринятым. Дело в том, что слово «информация» используется в самых разных ситуациях для обозначения того общего, что есть в разговоре людей, обмене письмами, чтении книги, прослушивании музыки, передаче сообщения через компьютерную сеть и т. д. Поэтому дать строгое определение информации не удаётся, можно только объяснить значение этого слова на примерах и сравнить с другими понятиями. Норберт Винер, создатель *кибернетики* — науки об управлении и связи писал: «Информация есть информация, а не материя и не энергия».

Как получают информацию

Человек получает информацию через свои органы чувств: глаза, уши, рот, нос и кожу. Поэтому получаемую нами информацию можно разделить на следующие виды:

- *зрительная информация* (визуальная, от англ. *visual*) — поступает через глаза (по разным оценкам, это 80–90% всей получаемой нами информации);
- *звуковая информация* (аудиальная, от англ. *audio*) — поступает через уши;

- *вкусовая информация* — поступает через язык;
- *обонятельная информация* (запахи) — поступает через нос;
- *тактильная информация* — мы её получаем с помощью осязания (кожи), «на ощупь».

Ещё выделяют информацию, получаемую с помощью «мышечного чувства» (человеческий мозг получает импульсы от мышц и суставов при перемещении частей тела).

Некоторые животные чувствуют магнитное поле Земли и используют его для выбора направления движения.

Формы представления информации

Информация может быть представлена (зафиксирована, закодирована) в различных *формах*:

- *текстовая информация* — последовательность символов (букв, цифр, других знаков); в тексте важен порядок их расположения, например, КОТ и ТОК — два разных текста, хотя они состоят из одинаковых символов;
- *числовая информация*;
- *графическая информация* (рисунки, картины, чертежи, карты, схемы, фотографии и т. п.);
- *звуковая информация* (звучание голоса, мелодии, шум, стук, шорох и т. п.);
- *мультимедийная информация*, которая объединяет несколько форм представления информации (например, видеоинформация).

Обратим внимание, что одна и та же информация может быть представлена по-разному. Например, результаты измерения температуры в течение недели можно сохранить в виде текста, чисел, таблицы, графика, диаграммы, видеоролика и т. д.

Человек, информация, знания

Обо всех изменениях в окружающем мире человек узнает с помощью своих органов чувств: сигналы от них («первичная» информация) постоянно поступают в мозг. Чтобы понять эти сигналы, т. е. извлечь информацию, человек использует **знания** — свои представления о природе, обществе, самом себе. Знания позволяют человеку принимать решения, определяют его поведение и отношения с другими людьми.

Можно считать, что знания — это модель мира, которая есть у человека. Получив информацию («поняв» сигналы, поступившие от органов чувств), он корректирует эту модель, дополняет свои знания.

Всегда ли полученная информация увеличивает наши знания? Очевидно, что нет. Например, информация о том, что $2 \cdot 2 = 4$ вряд ли увеличит ваши знания, потому что вы это уже знаете, эта информация для вас не нова. Однако она будет новой для тех, кто изучает таблицу умножения. Это значит, что изменение знаний при получении сообщения зависит от того, что человек знал до этого момента. Если он знает всё, что было в полученном сообщении, знания не изменяются.

Вместе с тем сообщение «учёт вибронных взаимодействий континуализирует моделирование диссипативных структур» (или сообщение на неизвестном языке) также не увеличивает знания, потому что эта фраза, скорее всего, вам непонятна. Дело в том, что имеющихся знаний не хватает для того, чтобы воспринять новую информацию.

Эти идеи послужили основой *семантической* (смысловой) теории информации, предложенной в 60-х годах XX века советским математиком Ю. А. Шрейдером. На рисунке 1.2 показано, как зависит количество полученных знаний I от того, какая доля информации θ в сообщении уже известна получателю.

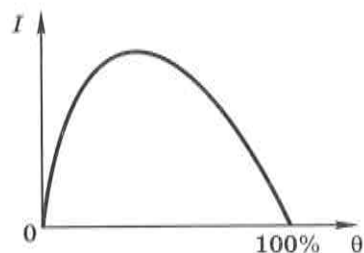


Рис. 1.2



Сообщение увеличивает знания человека, если оно понятно и содержит новые сведения.

К сожалению, «измерить» смысл информации, оценить его числом, довольно сложно. Поэтому для оценки количества инфор-

мации используют другие подходы, о которых вы узнаете чуть позже.

Когда человек хочет поделиться с кем-то своим знанием, он может сказать: «Я знаю, что...» или «Я знаю, как...». Это говорит о том, что есть два разных вида знаний. В первом случае знания — это некоторый известный факт, например: «Я знаю, что Луна вращается вокруг Земли». Такие знания называются *декларативными*, человек выражает их словами (*декларирует*). Декларативные знания — это факты, законы, принципы.

Второй тип знаний («Я знаю, как...») называют *процедурными*. Они выражаются в том, что человек знает, как нужно действовать в той или иной ситуации. К процедурным знаниям относятся алгоритмы решения различных задач.

Для того чтобы сохранить знания и передать другим людям, нужно выразить их на каком-то языке (например, рассказать, записать, нарисовать и т. п.). Только после этого их можно хранить, обрабатывать, передавать, причём с этим может справиться и компьютер. В научной литературе информацию, зафиксированную (закодированную) в какой-то форме, называют *данными*, имея в виду, что компьютер может выполнять с ними какие-то операции, но не способен понимать смысл.

Для того чтобы данные стали *информацией*, их нужно понять и осмыслить, а на это способен только человек. Если человек, получающий сообщение, знает язык, на котором оно записано, он может понять смысл этого сообщения, т. е. получить информацию. Обработывая и упорядочивая информацию, человек выявляет закономерности — получает *знания*, которые он затем снова использует для извлечения информации из полученных сообщений.

Мы увидели, что в науке существуют достаточно тонкие различия между понятиями «данные», «информация», «знания». Тем не менее на практике чаще всего всё это называется общим термином «информация».

Свойства информации

В идеале информация должна быть:

- *объективной* (не зависящей от чьего-либо мнения);
- *понятной* для получателя;
- *полезной* (позволяющей получателю решать свои задачи);
- *достоверной* (полученной из надёжного источника);

- *актуальной* (значимой в данный момент);
- *полной* (достаточной для принятия решения).

Конечно, информация не всегда обладает всеми этими свойствами. Информация в сообщении «В стакане мало молока» не объективна (для пессимиста полстакана — это мало, а для оптимиста — много). Сообщение 私は散歩に行った. непонятно для незнающих японский язык (оно означает «Я пошёл гулять», только по-японски).

Полезность информации определяется для каждого человека в конкретной ситуации. Например, информация о том, как древние люди добывали огонь, для большинства городских жителей бесполезна, поскольку она никак не помогает им решать свои жизненные проблемы. Вместе с тем в экстремальной ситуации, когда человек оказывается один на один с природой, такие знания очень полезны, потому что сильно увеличивают шансы на выживание, т. е. помогают достичь цели.

Слухи, байки, искажённая информация (в том числе дезинформация) — это примеры недостоверной информации.

Сообщение «10 лет назад здесь был ларёк с мороженым» неактуально, эта информация устарела.

Информация в сообщении «Сегодня будет концерт» неполна, потому что не указано время и участники концерта, и из-за этого мы не можем принять решение (идти или не идти?).

Развитие глобальной сети Интернет, в которую ежеминутно вносится огромное количество самых разнообразных данных, во многом перевернуло привычные представления о работе с информацией. Например, основным источником для поиска учебных материалов теперь фактически является Интернет, а не библиотеки. Однако при использовании информации из Интернета необходимо относиться к ней критически, так как её достоверность никто не гарантирует.

Роль информации в человеческом обществе очень велика. Информация, получаемая нами из разных источников, позволяет принимать решения и во многом определяет всю нашу жизнь. Огромно влияние на общество средств массовой информации (СМИ) — газет, телевидения, изданий в Интернете.

В будущем ожидается переход к информационному обществу, где большая часть населения будет заниматься сбором, обработ-

кой и распространением информации, поэтому высказывание немецкого банкира Н. Ротшильда «Кто владеет информацией, тот владеет миром» становится актуальным как никогда.

Информация в технике

Практически все современные технические устройства (телевизоры, телефоны, стиральные машины, системы управления самолётами и судами и т. д.) строятся на **микропроцессорах**, которые обрабатывают информацию: анализируют сигналы с датчиков, выбирают нужный режим работы. Широко используются **системы программного управления**, например, станки, обрабатывающие детали по программе, заложенной в памяти. Эту программу очень легко поменять и настроить станок на изготовление другой детали.

Многие опасные, тяжёлые и утомительные работы за человека могут выполнить **роботы**, у которых датчики заменяют органы чувств. Например, человекоподобный робот (*андроид*) Asimo (рис. 1.3), разработанный фирмой Honda, умеет распознавать предметы, жесты, звуки, узнавать лица, разговаривать через домофон, передавать данные через Интернет.

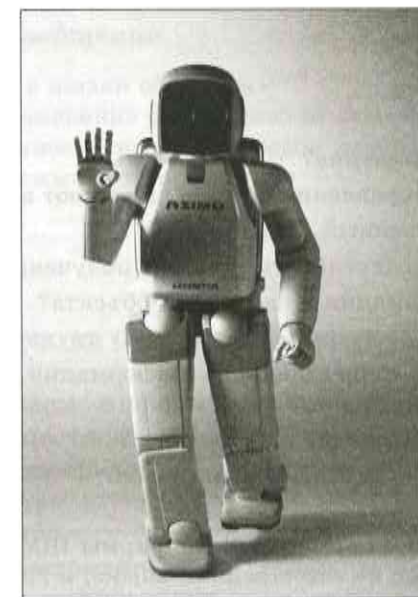


Рис. 1.3. Робот Asimo (www.robotonline.net)

Наиболее универсальным устройством для обработки информации можно считать **компьютер**. Хотя современные компьютеры пока не умеют работать с вкусовой и обонятельной информацией (запахами), работы в этом направлении ведутся. Уже существуют экспериментальные приборы, названные «электронный нос» и «электронный язык»; они построены на основе химических датчиков.

Сейчас в теоретической информатике считается, что компьютер может хранить и обрабатывать только *данные*, но не информацию. Многие учёные полагают, что машина принципиально не может научиться понимать смысл информации и делать выводы¹. Эту точку зрения подтверждает фактический провал проекта «компьютеров пятого поколения» (Япония, 1980-е гг.), в ходе которого планировалось создать машины, общающиеся с человеком на естественном языке. Тем не менее учёные уделяют этим проблемам огромное внимание. Например, возникло целое научное направление *data mining* («добыча данных»), в котором изучаются методы извлечения информации («смысла», закономерностей, связей, знаний) из огромных наборов данных. В некоторых случаях действительно удается использовать огромные вычислительные мощности компьютеров для того, чтобы найти неизвестные ранее закономерности, которые можно использовать на практике.



Вопросы и задания

1. Что изучает информатика?
2. Какие научные направления обычно включают в информатику?
3. Что такое искусственный интеллект?
4. Как связана неопределённость знания с получением информации?
5. Как связана информация и сложность объекта?
6. Объясните, почему термин «информация» трудно определить.
7. Согласны ли вы с «определением» информации, которое дал Н. Винер? Как вы его понимаете?
8. Как человек воспринимает информацию?
9. Чем отличается текст от набора символов?

¹ Тем не менее суперкомпьютер Watson фирмы IBM, умеющий отвечать на вопросы, заданные на естественном языке, в 2011 г. выиграл у лучших игроков в телевизионной викторине *Jeopardy!* (аналог телепередачи «Что? Где? Когда?»).

10. К какому виду информации относятся видеофильмы?
11. Что такое тактильная информация?
12. Всякая ли информация увеличивает знания? Почему?
13. На каких идеях основана семантическая теория информации?
14. Приведите примеры своих декларативных и процедурных знаний.
15. В чём, на ваш взгляд, разница между понятиями «данные», «информация», «знания»?
16. Почему считают, что компьютер может работать только с данными?
17. Какими свойствами должна обладать «идеальная» информация?
18. Приведите примеры необъективной, непонятной, бесполезной, недостоверной, неактуальной и неполной информации.
19. Может ли информация быть достоверной, но бесполезной? Достоверной, но необъективной? Объективной, но недостоверной? Актуальной, но непонятной?
20. Приведите примеры обработки информации в технических устройствах.
21. Что умеет робот Asimo? Какую информацию он обрабатывает?
22. Что такое «электронный нос» и «электронный язык»?
23. Как вы считаете, смогут ли компьютеры научиться понимать смысл данных?

Подготовьте сообщение

- а) «Информация в жизни общества»
- б) «Интернет и изменение уклада жизни людей»
- в) «Информационное общество: плюсы и минусы»
- г) «Как оценить смысл информации?»

§ 2

Что можно делать с информацией?

Как мы уже знаем, информация сама по себе нематериальна. Но она может существовать только тогда, когда связана с каким-то объектом или средой, т. е. с носителем.

Материальный носитель — это объект или среда, которые могут содержать информацию.



Изменения, происходящие с информацией (т. е. изменения свойств носителя), называются **информационными процессами**. Все эти процессы можно свести к двум основным:

- **передача информации** (данные передаются с одного носителя на другой);
- **обработка информации** (данные изменяются).

Часто информационными процессами называют также и многие другие операции с информацией (например, копирование, удаление и др.), но они, в конечном счёте, сводятся к двум названным процессам.

Для хранения информации тоже используется какой-то носитель. Однако при этом никаких изменений не происходит, поэтому хранение информации нельзя назвать процессом.

Передача информации

При **передаче информации** всегда есть два объекта — источник и приёмник информации. Эти роли могут меняться, например, во время диалога каждый из участников выступает то в роли источника, то в роли приёмника информации.

Информация проходит от источника к приёмнику через **канал связи**, в котором она должна быть связана с каким-то **материальным носителем** (рис. 1.4). Для передачи информации свойства этого носителя должны изменяться со временем. Например, если включать и выключать лампочку, то можно передавать сообщения с помощью азбуки Морзе.



Рис. 1.4

При разговоре людей носитель информации — это звуковые волны в воздухе. В компьютерах информация передается с помощью электрических сигналов или радиоволн (в беспроводных устройствах). Информация может передаваться с помощью света, лазерного луча, звуковых волн и др.

Информация поступает по каналу связи в виде сигналов, которые приёмник может обнаружить с помощью своих органов чувств (или датчиков) и «понять» (раскодировать).

Сигнал — это изменение свойств носителя, которое используется для передачи информации.

Примеры сигналов — это изменение частоты и громкости звука, вспышки света, изменение напряжения на контактах и т. п.

Человек может принимать сигналы только с помощью своих органов чувств. Чтобы передавать и принимать информацию, например, с помощью радиоволн, нужны вспомогательные устройства: радиопередатчик, преобразующий звук в радиоволны, и радиоприёмник, выполняющий обратное преобразование. Они позволяют расширить возможности человека.

С помощью одного сигнала (одного изменения) невозможно передать много информации. Поэтому чаще всего используется не одиночный сигнал, а *последовательность сигналов*, которая называется **сообщением**. Важно понимать, что сообщение — это только «оболочка» для передачи информации, а информация — это *содержание* сообщения. Приёмник должен сам «извлечь» (раскодировать) информацию из полученной последовательности сигналов. Можно принять сообщение, но не принять информацию, например, услышав речь на незнакомом языке или перехватив шифровку.

Одна и та же информация может быть передана с помощью сообщений, имеющих разные физические носители (например, через устную речь, с помощью записки или с помощью флажного семафора, который используется на флоте). В то же время одно и то же сообщение может нести разную информацию для разных приёмников. Так фраза «В Сантьяго идёт дождь», переданная в 1973 г. на военных радиочастотах, для сторонников генерала Пиночета послужила сигналом к началу государственного переворота в Чили.

К сожалению, в реальном канале связи всегда действуют помехи: посторонние звуки при разговоре, шумы радиоэфира, электрические и магнитные поля. Помехи могут полностью или частично исказить сообщение, вплоть до полной потери информации (например, телефонные разговоры при перегрузке сети).

Чтобы содержание сообщения, искажённого помехами, можно было восстановить, оно должно быть *избыточным*, т. е. в нём должны быть «лишние» элементы, без которых смысл всё равно восстанавливается. Например, в сообщении «Влг впдт в Кспск мр» многие угадают фразу «Волга впадает в Каспийское море», из которой убрали все гласные. Этот пример говорит о том, что естественные языки содержат много «лишнего», их избыточность оценивается в 60–80% (если удалить 60–80% текста, его смысл всё равно удастся восстановить).

В курсе информатики мы будем рассматривать передачу информации именно как передачу сообщений между компьютерными системами, отвлекаясь от смысла сообщений.

Обработка информации

Обработка — это изменение информации: её формы или содержания. Среди важнейших видов обработки можно назвать:

- *создание новой информации*, например, решение задачи с помощью вычислений или логических рассуждений;
- *кодирование* — запись информации с помощью некоторой системы знаков для передачи и хранения; один из вариантов кодирования — шифрование, цель которого — скрыть смысл (содержание) информации от посторонних;
- *поиск информации*, например, в книге, в библиотечном каталоге, на схеме или в Интернете;
- *структурирование информации*, т. е. выделение важных элементов в сообщениях и установление связей между ними;
- *сортировка* — расстановка элементов списка в заданном порядке, например, расстановка чисел по возрастанию или убыванию, расстановка слов по алфавиту; задача сортировки — облегчить поиск и анализ информации.

Для обработки информации человек использует, в первую очередь, свой мозг. *Нейроны* (нервные клетки) коры головного мозга «переключаются» примерно 200 раз в секунду — значительно медленнее, чем элементы памяти компьютеров. Однако человек практически безошибочно отличает собаку от кошки, а для компьютеров эта задача пока не разрешима. Дело, по-видимому, в том, что мозг решает такие задачи не «в лоб», не путем сложных вычислений, а как-то иначе (как — пока никто до конца не знает).

Компьютер позволяет «усилить» возможности человека в тех задачах обработки информации, решение которых требует дли-

тельных расчётов по известным алгоритмам. Однако, в отличие от человека, для компьютера недоступны фантазия, размышления, творчество.

Хранение информации

Для хранения информации человек, прежде всего, использует свою память. Можно считать, что мозг — это одно из самых совершенных хранилищ информации, во многом превосходящее компьютерные средства.

К сожалению, человек многое забывает. Кроме того, необходимо передавать знания другим людям, в том числе и следующим поколениям. Поэтому в древности люди записывали информацию на камне, папирусе, берёсте, пергаменте, затем — на бумаге. В XX веке появились новые средства хранения информации: перфокарты и перфоленты, магнитные ленты и магнитные диски, оптические диски, флэш-память.

В любом случае информация хранится на каком-то носителе, который обладает «памятью», т. е. может находиться в разных состояниях, переходить из одного состояния в другое при каком-то внешнем воздействии, и сохранять свое состояние.

При записи информации свойства носителя меняются: на бумагу наносятся текст и рисунки; на магнитных дисках и лентах намагничиваются отдельные участки; на оптических дисках образуются области, по-разному отражающие свет. При хранении эти свойства остаются неизменными, что позволяет потом читать (получать) записанную информацию.

Отметим, что процессы записи и чтения — это процессы передачи информации.

Вопросы и задания

1. Кто (что) может быть источником (приёмником) информации? Приведите примеры.
2. Что такое сигнал? Приведите примеры сигналов.
3. Что такое сообщение? Чем отличается получение информации от получения сообщения?
4. Приведите примеры, когда приём сообщения не означает приём информации.
5. Приведите примеры, когда одна и та же информация может быть передана с помощью разных сообщений.

6. Приведите примеры, когда одно и то же сообщение несёт разную информацию для разных людей.
7. Расскажите, как помехи влияют на передачу информации. Приведите примеры.
8. Что такое избыточность? Почему она полезна при передаче информации?
9. Представьте, что придумали язык, в котором нет избыточности. В чем будет его недостаток?
10. Как вы думаете, какой вариант русского языка обладает наибольшей избыточностью: разговорный, литературный, юридический, язык авиадиспетчеров?
11. В каком из перечисленных выше языков наиболее важна помехоустойчивость? За счёт чего она достигается?
12. Какие виды обработки информации вы знаете?
13. При каких видах обработки информации меняется её содержание?
14. При каких видах обработки информации меняется только форма её представления?
15. К какому виду обработки можно отнести шифрование? Почему?
16. Работники удалённой метеостанции каждые 3 часа измеряют температуру и влажность воздуха, и передают данные по радию в районный метеоцентр. Там эти данные сводят в таблицу и отправляют по электронной почте в Гидрометцентр, где мощные компьютеры составляют прогноз погоды. Выделите здесь процессы, связанные с обработкой и передачей информации.
17. Ученик нашёл в старой книге сведения о населении Москвы в XIX веке, составил таблицу по этим данным, построил диаграмму и сделал доклад на школьной конференции. Выделите здесь процессы, связанные с обработкой и передачей информации.
18. Зачем человек записывает информацию?
19. В чём преимущества и недостатки человеческой памяти по сравнению с компьютерной?

Подготовьте сообщение

- а) «Компьютер и человек: кто сильнее?»
- б) «Носители информации: вчера, сегодня, завтра»

§ 3

Измерение информации

Любая наука рано или поздно приходит к необходимости как-то измерять то, что она изучает. Измерение информации — это одна из важнейших задач теоретической информатики.

Для человека информация — это, прежде всего, смысл, заключённый в сигналах и данных. Как измерить смысл? На этот вопрос пока нет однозначного ответа.

Вспомним, что компьютеры не могут обрабатывать смысл, они работают только с данными (а не с информацией). При этом возникают чисто практические задачи: определить, сколько места займёт на диске текст, рисунок или видеофильм; сколько времени потребуется на передачу файла по компьютерной сети и т. п. При решении этих задач используется **объёмный подход** к измерению информации. Он заключается в том, что количество информации оценивается просто по числу символов, используемых для её кодирования. С этой точки зрения стихотворение А. С. Пушкина и случайный набор букв могут содержать одинаковое количество информации. Конечно, такой подход не универсален, но он позволяет успешно решать практические задачи, связанные с компьютерной обработкой и хранением данных.

Что такое бит?

Рассмотрим электрическую лампочку, которая может находиться в двух состояниях: «горит» и «не горит». Тогда на вопрос «Горит ли сейчас лампочка» есть два возможных варианта ответа, которые можно обозначить цифрами 1 («горит») и 0 («не горит») (рис. 1.5). Поэтому ответ на этот вопрос (полученная информация) может быть записан как 0 или 1¹.

Цифры 0 и 1 называют **двоичными**, и с этим связано название единицы измерения количества информации — **бит**. Английское слово *bit* — это сокращение от выражения *binary digit* — «двоичная цифра». Впервые слово «бит» в этом значении использовал американский инженер и математик Клод Шеннон в 1948 г.

Бит — это количество информации, которую можно записать (закодировать) с помощью одной двоичной цифры.

¹ Конечно, вместо 0 и 1 можно использовать два любых знака.

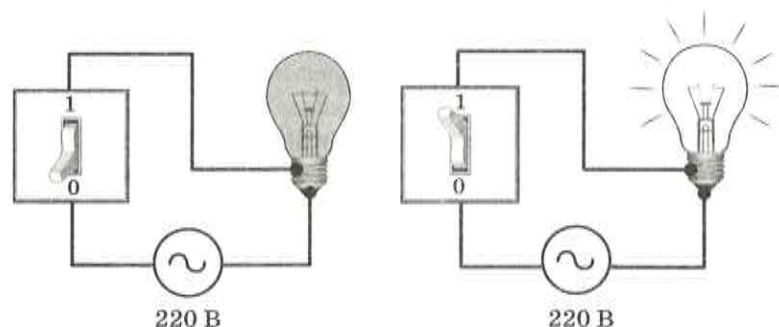


Рис. 1.5

Конечно, нужно договориться, что означают 0 и 1 (1 — это «горит» или «не горит?»), но для измерения количества информации это не важно.

Как вы уже поняли, 1 бит — это количество информации, которое позволяет выбрать один из двух возможных вариантов. Например, в сообщении «подброшенная монета упала гербом» содержится 1 бит информации, потому что монета могла упасть гербом (обозначим это через 0) или «решкой» (1). Сообщение «Дверь открыта» тоже содержит 1 бит, если считать, что дверь может быть в двух состояниях: открыта (0) или закрыта (1). Вот ещё пример диалога, в котором получена информация в 1 бит:

- Вы будете чай или кофе?
- Кофе, пожалуйста.

2 бита, 3 бита...

А если возможных вариантов не два, а больше? Понятно, что в этом случае количество информации будет больше, чем 1 бит. Представим себе, что на вокзале стоят 4 одинаковых поезда (рис. 1.6), причём только один из них проследует в Москву. Сколько битов понадобится для того, чтобы записать информацию о номере платформы, где стоит поезд на Москву?

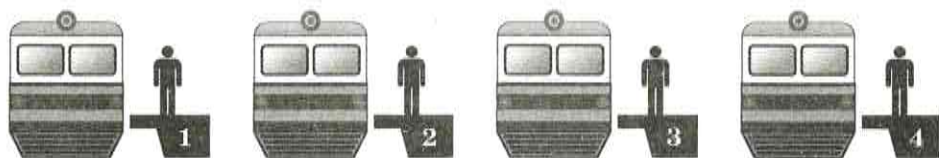


Рис. 1.6

Очевидно, что одного бита недостаточно, так как с помощью одной двоичной цифры можно закодировать только два варианта — коды 0 и 1. А вот два бита дают как раз 4 разных сообщения: 00, 01, 10 и 11. Теперь нужно сопоставить эти коды номерам платформ, например, так: 1 — 00, 2 — 01, 3 — 10, 4 — 11. Тогда сообщение 10 говорит о том, что поезд на Москву стоит на платформе № 3. Это сообщение несёт 2 бита информации.

Три бита дают уже 8 вариантов: 000, 001, 010, 011, 100, 101, 110 и 111. Таким образом, каждый бит, добавленный в сообщение, увеличивает количество вариантов в 2 раза (табл. 1.1).

Таблица 1.1

I , битов	1	2	3	4	5	6	7	8	9	10
N , вариантов	2	4	8	16	32	64	128	256	512	1024

Наверно, вы заметили, что все числа в нижней строчке таблицы — это степени числа 2: $N = 2^I$.

Осталось выяснить, чему равно количество информации, если выбор делается, скажем, из 5 возможных вариантов (или из любого количества, не являющегося степенью числа 2). С точки зрения приведённого выше рассуждения случаи выбора из 5, 6, 7 и 8 вариантов не различаются — для кодирования двух двоичных цифр мало, а трёх — достаточно. Поэтому использование трёх битов для кодирования одного из 5 возможных вариантов *избыточно*, ведь три бита позволяют закодировать целых 8 вариантов! Значит, выбор из 5 вариантов даёт меньше трёх битов информации.

Чтобы количественно измерить разницу между, скажем, 5-ю и 8-ю вариантами, придется допустить, что количество информации в битах может быть дробным числом. При этом информация, полученная при выборе из 5 вариантов, больше, чем 2 бита, но меньше, чем 3 бита. Точную формулу для ее вычисления получил в 1928 г. американский инженер Ральф Хартли. Эта формула использует понятие логарифма, поэтому мы познакомимся с ней в 11 классе.



Тем не менее уже сейчас вы можете оценить количество информации при выборе из 5 вариантов. Допустим, на завтрак в лагере отдыха каждый день дают кашу одного из 5 видов. Чтобы закодировать вид каши, которую дают в понедельник, нужно, как мы знаем, 3 бита. Но меню на два дня может быть составлено 25 разными способами ($5 \cdot 5$), поэтому для его кодирования достаточно 5 битов, ведь $25 < 32 = 2^5$. И так, количество информации о меню за 2 дня меньше 5 битов, поэтому на каждый день (выбор из 5 вариантов) приходится меньше, чем $5/2 = 2,5$ бита. Но и эту оценку можно уточнить. Для трёх дней получаем $5 \cdot 5 \cdot 5 = 125$ вариантов. Так как $125 < 128 = 2^7$, количество информации при выборе из 5 вариантов составляет не больше, чем $7/3 = 2,33$ бита. И так далее. Попробуйте самостоятельно показать, что при выборе из 5 вариантов количество информации больше 2,25 бита. Верно ли, что при выборе из 6 вариантов количество информации менее 2,5 бита?

Другие единицы

Считать большие объёмы информации в битах неудобно, хотя бы потому, что придется работать с очень большими числами (миллиардами, триллионами и т. д.). Поэтому стоит ввести более крупные единицы.

1 байт = 8 битов.

Сразу возникает вопрос: а почему не 10 битов? Дело в том, что слово «байт¹» (англ. *byte*) имеет второе значение — так называют наименьший блок (ячейку) памяти, который процессор компьютера может считать за один раз. Для современных компьютеров он состоит из 8 элементов, каждый из которых хранит 1 бит данных. Это связано с тем, что до недавнего времени при обработке текста использовался набор из 256 символов, так что для кодирования каждого символа было нужно 8 битов.

Объёмы данных, с которыми работают компьютеры, нередко измеряются миллионами и миллиардами байтов. В таких случаях используют единицы, образованные с помощью приставок:

1 Кбайт (килобайт) = 1024 байта = 2^{10} байта = 2^{13} битов.
 1 Мбайт (мегабайт) = 1024 Кбайт = 2^{10} Кбайт = 2^{20} байтов = 2^{23} битов.
 1 Гбайт (гигабайт) = 1024 Мбайт.
 1 Тбайт (терабайт) = 1024 Гбайт.

¹ Впервые его использовал американский инженер В. Бухгольд в 1956 г.

Так сложилось исторически, что при измерении количества информации приставка «кило-» обозначает, в отличие от международной системы единиц СИ, увеличение не в 1000 раз, а в $1024 = 2^{10}$ раз. Аналогично «мега-» — это увеличение в $1024^2 = 2^{20} = 1\,048\,576$ раз, а не в 1 млн = 1000^2 раз.

Строго говоря, нужно называть такие кило- (мега-, гига-, ...) байты *двоичными*, поскольку множитель 1024 — это 2^{10} . Стандарт Международной электротехнической комиссии (МЭК) предлагает называть их «кибибайт», «мебибайт», «гибибайт» и «тебибайт», но эти названия на практике не прижились.

Для перевода количества информации из одних единиц в другие нужно использовать приведённые выше соотношения. При переводе из крупных единиц в мелкие числа умножают на соотношение между единицами. Например:

$$2 \text{ Кбайт} = 2 \cdot (1 \text{ Кбайт}) = 2 \cdot 1024 \text{ байтов} = 2048 \text{ байтов} = \\ = 2048 \cdot (1 \text{ байт}) = 2048 \cdot 8 \text{ битов} = 16\,384 \text{ бита}.$$

$$2 \text{ Кбайт} = 2 \cdot 2^{10} \text{ байтов} = 2^{11} \text{ байтов} = 2^{11} \cdot 2^3 \text{ битов} = 2^{14} \text{ битов}.$$

В последней строке все расчёты сделаны через степени числа 2, очень часто так бывает проще.

При переводе количества информации из мелких единиц в крупные нужно делить на соотношение между единицами. Например:

$$8192 \text{ бита} = 8192 \cdot (1/8 \text{ байта}) = 8192 : 8 \text{ байтов} = 1024 \text{ байта} = \\ = 1024 \cdot (1/1024 \text{ Кбайт}) = 1024 : 1024 \text{ Кбайт} = 1 \text{ Кбайт}.$$

$$8192 \text{ бита} = 2^{13} \text{ битов} = 2^{13} \cdot (1/2^3 \text{ байта}) = 2^{10} \text{ байтов} = \\ = 2^{10} \cdot (1/2^{10} \text{ Кбайт}) = 1 \text{ Кбайт}.$$

Вопросы и задания

1. Дайте определение минимальной единицы измерения количества информации.
2. Приведите примеры сообщений, количество информации в которых равно 1 биту.
3. Что такое двоичные цифры?
4. Объясните, почему все числа во второй строке табл. 1.1 — это степени числа 2.
5. Какие единицы используют для измерения больших объёмов информации?
6. Что означают приставки «кило-», «мега-», «гига-» и «тера-» при измерении количества информации?



7. Какие приставки рекомендует МЭК для обозначения двоичных килобайта и мегабайта? Как вы думаете, почему они редко используются?



Подготовьте сообщение

- а) «Бит и байт: как возникли термины?»
б) «Стандарт МЭК и единицы измерения количества информации»



Задачи

1. Пассажир не знает, какой (только один!) из 8 поездов, стоящих на вокзале, проследует в Санкт-Петербург. В справочном бюро он задает 8 вопросов: «Поезд на 1-й платформе проследует в Санкт-Петербург?», «Поезд на 2-й платформе проследует в Санкт-Петербург?» и т. д. На первые 7 вопросов он получает ответ «нет», а на последний — «да». Пассажир считает, что он получил 8 битов информации. Прав он или нет? Почему?
- *2. В зоопарке содержится 10 обезьян, причём одна из них выступает в цирке. Обезьяны сидят в двух вольерах, в первом — 8 животных, а во втором — два. Посетитель зоопарка считает, что сообщение «Обезьяна, выступающая в цирке, сидит во втором вольере» содержит 1 бит информации. Прав он или нет? Рассмотрите разные варианты уточнения постановки задачи.
3. В горах, рядом с которыми живёт племя Тумба-Юмба, есть 4 пещеры. В каждой из них может быть (а может не быть) клад. Можно ли закодировать сведения о том, где есть клад, используя 3 бита? 4 бита? 5 битов?
4. Известно, что ровно в двух пещерах из четырёх есть клад. Сколько битов нужно, чтобы закодировать информацию о расположении кладов?
- *5. Известно, что дверь с двумя замками открывается двумя из четырёх имеющихся ключей. Оцените количество информации в сообщении «Дверь открывается ключами № 2 и № 4». Закодируйте его, используя наименьшее количество двоичных цифр.
- *6. Известно, что дверь открывается двумя из пяти имеющихся ключей. Оцените количество информации в сообщении «Верхний замок открывается ключом № 1, а нижний — ключом № 4». Закодируйте его, используя наименьшее количество двоичных цифр.
7. Вася задумал число от 1 до 100. Нужно отгадать это число за наименьшее число попыток, задавая Васе вопросы, на которые он отвечает только «да» и «нет». За сколько вопросов вы берётесь угадать

число? Как нужно задавать вопросы, чтобы их число было минимальным даже в худшем случае?

8. Даниил задумал число от 20 до 83. Сколько битов информации содержится в сообщении «Даниил задумал число 77»? Закодируйте это сообщение, используя наименьшее количество двоичных цифр.
9. Двое играют в «крестики-нолики» на поле размером 4×4 клетки. Какое количество информации получил второй игрок, узнав первый ход соперника?
10. На вокзале посёлка Сосново три платформы, у каждой из них стоит поезд. Девушка в справочном окне отвечает на все вопросы только «да» и «нет». За какое минимальное число вопросов можно гарантированно узнать, в каком порядке отходят поезда?
11. Переведите 1 Мбайт во все изученные единицы измерения количества информации.
12. Переведите 2^{26} битов во все изученные единицы измерения количества информации.
13. Сколько килобайтов содержится в 32 768 битах?
14. Сколько битов в 8 Кбайтах?
15. Сколько битов содержит 1/16 Кбайт?
16. Сколько битов содержит 1/512 Мбайт?

§ 4

Структура информации

Зачем структурировать информацию?

Давайте сравним четыре сообщения.

Первое:

«Для того чтобы добраться из Москвы до села Васино, нужно сначала долететь на самолёте до города Ивановска. Затем на электричке доехать до Ореховска. Там на пароме переправиться через реку Слоновую в посёлок Ольховка, и оттуда ехать в село Васино на попутной машине».

Второе:

«Как ехать в Васино:

1. На самолёте из Москвы до г. Ивановска.
2. На электричке из г. Ивановска до г. Ореховска.
3. На пароме из г. Ореховска через р. Слоновую в пос. Ольховка.
4. На попутной машине из пос. Ольховка до с. Васино».

Третье:

Откуда	Куда	Транспорт
Москва	г. Ивановск	Самолёт
г. Ивановск	г. Ореховск	Электричка
г. Ореховск	пос. Ольховка	Паром (р. Слоновая)
пос. Ольховка	с. Васино	Попутная машина

Четвёртое (рис. 1.7):



Рис. 1.7

Можно считать, что все эти (такие разные по форме!) сообщения содержат одну и ту же информацию. Какие из них проще воспринимать? Очевидно, что человеку «вытащить» полезную информацию из сплошного текста (первое сообщение) сложнее всего. Во втором случае мы сразу видим все этапы поездки и понимаем, в каком порядке они следуют друг за другом. Третье сообщение (таблицу) и четвёртое (схему) можно понять сразу, с первого взгляда. Второй, третий и четвёртый варианты воспринимаются лучше и быстрее первого, потому что в них выделена **структура информации**, в которой самое главное — этапы поездки в Васино.

Зачем книгу разбивают на главы и разделы, а не пишут сплошной текст? Зачем в тексте выделяют абзацы? Прежде всего для того, чтобы подчеркнуть основные мысли — каждая глава, раздел, абзац содержат определённую идею. Благодаря особенностям человеческого восприятия, при таком выделении структуры улучшается передача информации от автора к читателю.

При автоматической (компьютерной) обработке правильно выбранная структура данных облегчает доступ к ним, позволяет быстро найти нужные данные.

Средства, облегчающие поиск информации, знакомы вам по работе с книгами. Самый простой (но очень утомительный!) способ найти в книге то, что нужно, — перелистывать страницу за страницей. Однако в большинстве книг есть оглавление, которое

позволяет сразу найти нужный раздел, и это значительно ускоряет поиск.

В словарях слова всегда расставлены в алфавитном порядке (представьте себе, что было бы, если бы они были расположены произвольно!). Поэтому открыв словарь в любом месте, мы можем сразу определить, куда дальше листать страницы для поиска нужного слова — вперёд или назад.

В больших книгах используют **индексы** (рис. 1.8) — списки основных терминов с указанием страниц, на которых они встречаются.

Оглавление:	Словарь:	Индекс:
1. Информация 5	автомат — <i>automaton</i>	А
1.1. Что такое информация? 6	автор — <i>author</i>	аксиома 45
1.2. Виды информации 8	адрес — <i>address</i>	алгоритм 30, 78
1.3. Информация в технике 11	алгебра — <i>algebra</i>	архиватор 125
2. Измерение информации 12	алгоритм — <i>algorithm</i>	Б
2.1. Что такое бит? 13	архив — <i>archive</i>	бит 5, 15, 25, 43
2.2. Байт и другие единицы 14	архитектура — <i>architecture</i>	брандмауэр 112
	асимметрия — <i>asymmetry</i>	браузер 322

Рис. 1.8

Структурирование — это выделение важных элементов в информационных сообщениях и установление связей между ними.

Цели структурирования для человека — облегчение восприятия и поиска информации, выявление закономерностей. При компьютерной обработке структурирование ускоряет поиск нужных данных.

Знакомые структуры данных

С некоторыми структурами данных вы уже знакомы. Например, на уроках математики вы изучали **множество** — некоторый набор элементов. Чтобы определить множество, мы должны перечислить все его элементы (например, множество, состоящее из Васи, Пети и Коли) или определить характерный признак, по ко-

торому элементы включаются в это множество (например, множество драконов с пятью зелёными хвостами или множество точек, в которых функция принимает положительные значения).

Множество может состоять из конечного числа элементов (множество букв русского алфавита), бесконечного числа элементов (множество натуральных чисел) или вообще быть пустым (множество слонов, живущих на Северном полюсе). Множества, с которыми работает компьютер, не могут быть бесконечными, потому что его память конечна.

В документах множество часто оформляют в виде маркированного списка, например:

- процессор;
- память;
- устройства ввода;
- устройства вывода.

В таком списке порядок элементов не важен, от перестановки элементов множество не меняется (рис. 1.9).



Рис. 1.9

Линейный список состоит из конечного числа элементов, которые должны быть расположены в строго определённом порядке. В отличие от множества, элементы в списке могут повторяться. Список обычно упорядочен (отсортирован) по какому-то правилу, например, по алфавиту, по важности, по последовательности действий и т. д. В тексте он часто оформляется как нумерованный список, например:

- 1) надеть носки;
- 2) надеть ботинки;
- 3) выйти из дома.

Переставить местами элементы такого списка нельзя (это будет уже другой список). Список можно задать перечислением элементов, с первого до последнего:

(надеть носки, надеть ботинки, выйти из дома),

а также представить в виде цепочки связанных элементов (рис. 1.10).

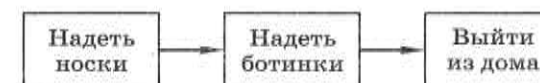


Рис. 1.10

Ещё одна знакомая вам структура — **таблица**. С помощью таблиц устанавливается связь между несколькими элементами. Например, в табл. 1.2 элементы в каждой строке связаны между собой — это свойства некоторого объекта (человека).

Таблица 1.2

Фамилия	Имя	Рост, см	Вес, кг	Год рождения
Иванов	Иван	175	67	1996
Петров	Пётр	164	70	1998
Сидоров	Сидор	168	63	2000

Именно так хранится информация в базах данных: строка таблицы, содержащая информацию об одном объекте, называется *записью*, а столбец (название свойства) — *полем*.

Возможен и другой вариант таблицы, когда роли строк и столбцов меняются. В первом столбце записываются названия свойств, а данные в каждом из следующих столбцов описывают свойства какого-то объекта. Например, табл. 1.3 содержит характеристики разных марок автомашин.

Таблица 1.3

Марка	Лада Приора	Лада Калина	ВАЗ 2110	ВАЗ 21099
Мощность двигателя, л. с.	98	89	79	70
Максимальная скорость, км/ч	183	165	165	156
Время разгона до 100 км/ч, с	11,5	12,5	14	15

Иерархия (дерево)

Линейных списков и таблиц иногда недостаточно для того, чтобы представить все связи между элементами. Например, в некоторой фирме есть директор, ему подчиняются главный инженер и главный бухгалтер, у каждого из них есть свои подчинённые. Если мы захотим нарисовать схему управления этой фирмы, она получится *многоуровневой* (рис. 1.11, а).

Такая структура, в которой одни элементы «подчиняются» другим, называется **иерархией** (от древнегреческого *ἱεραρχία* — священное правление). В информатике иерархическую структуру называют **деревом**. Дело в том, что если перевернуть схему на рис. 1.11 вверх ногами, она становится похожа на дерево (точнее, на куст, см. рис. 1.11, б).



Рис. 1.11

Дерево состоит из узлов и связей между ними (они называются **дугами**). Самый первый узел, расположенный на верхнем уровне (в него не входит ни одна стрелка-дуга) — это **корень дерева**. Конечные узлы, из которых не выходит ни одна дуга, называются **листьями**. Все остальные узлы, кроме корня и листьев — промежуточные.

Из двух связанных узлов тот, который находится на более высоком уровне, называется **родителем**, а другой — **сыном**. Корень — это единственный узел, у которого нет родителя; у листьев нет сыновей.

Используются также понятия **предок** и **потомок**. Потомок какого-то узла — это узел, в который можно перейти по стрелкам от узла-предка. Соответственно, предок какого-то узла — это узел, из которого можно перейти по стрелкам в данный узел.

В дереве на рис. 1.12 родитель узла *E* — это узел *B*, а предки узла *E* — это узлы *A* и *B*, для которых узел *E* — потомок. Потомками узла *A* (корня) являются все остальные узлы.

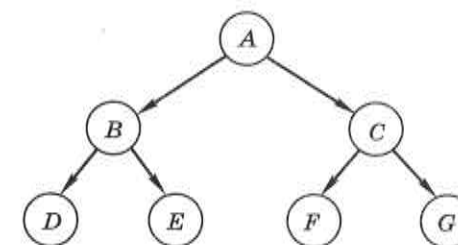


Рис. 1.12

Типичный пример иерархии — различные *классификации* (животных, растений, минералов, химических соединений). Например, отряд *Хищные* делится на два подотряда: *Псообразные* и *Кошкообразные*. В каждом из них выделяют несколько семейств (рис. 1.13).



Рис. 1.13

Конечно, на рис. 1.13 показаны не все семейства, остальные обозначены многоточиями.

В текстах иерархию часто представляют в виде многоуровневого списка. Например, оглавление книги о хищниках может выглядеть так:

- Глава 1. Псообразные
- 1.1. Псовые
- 1.2. Енотовые
- 1.3. Медвежьи

...

Глава 2. Кошкообразные

2.1. Кошачьи

2.2. Гиеновые

2.3. Мангустовые

...

Работая с файлами и каталогами, мы тоже встречаемся с иерархией: классическая файловая система имеет древовидную структуру¹. Вход в каталог — это переход на следующий (более низкий) уровень иерархии (рис. 1.14).



Рис. 1.14

Алгоритм вычисления арифметического выражения тоже может быть представлен в виде дерева (рис. 1.15).

Здесь листья — это числа и переменные, тогда как корень и промежуточные вершины — знаки операций. Вычисления идут «снизу вверх», от листьев — к корню. Показанное дерево можно записать так:

$$(- (* (+ (a, 3), 5), * (2, b)))$$

¹ В современных файловых системах (NTFS, ext3) файл может «принадлежать» нескольким каталогам одновременно. При этом древовидная структура, строго говоря, нарушается.

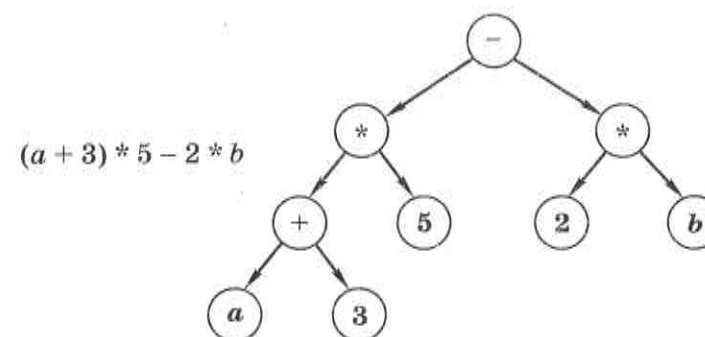


Рис. 1.15

Самое интересное, что скобки здесь необязательны; если их убрать, то выражение все равно может быть однозначно вычислено:

$$- * + a 3 5 * 2 b$$

Такая запись, которая называется *префиксной* (операция записывается перед данными), просматривается с конца. Как только встретится знак операции, эта операция выполняется с двумя значениями, записанными справа. В рассмотренном выражении сначала выполняется умножение:

$$- * + a 3 5 (2 * b)$$

затем — сложение:

$$- * (a + 3) 5 (2 * b)$$

и ещё одно умножение

$$- (a + 3) * 5 (2 * b)$$

и, наконец, вычитание:

$$(a + 3) * 5 - (2 * b)$$

Для получения префиксной записи мы обходили все узлы дерева в порядке «корень — левое поддерево — правое поддерево». Действительно, сначала записана метка корня («-»), затем все метки левого поддерева, а затем — все метки правого поддерева. Для поддеревьев используется тот же порядок обхода. Если же обойти дерево в порядке «левое поддерево — правое поддерево —

корень», получается *постфиксная* форма (операция *после* данных). Например, рассмотренное выше выражение может быть записано в виде

$$a \ 3 + 5 * 2 \ b \ -$$

Для вычисления такого выражения скобки также не нужны, и это очень удобно для автоматических расчётов. Когда программа на языке программирования высокого уровня переводится в машинные коды, часто все выражения записываются в бесскобочной постфиксной форме и именно так и вычисляются. Постфиксная форма для компьютера удобнее, чем префиксная, потому что, когда программа доходит до знака операции, все данные для выполнения этой операции уже готовы.

Графы

Подумайте, как можно структурировать такую информацию:

«От посёлка Васюки три дороги идут в посёлки Солнцево, Грибное и Ягодное. Между Солнцевым и Грибным и между Грибным и Ягодным также есть дороги. Кроме того, есть дорога, которая идет из Грибного в лес и возвращается обратно в Грибное».

Можно, например, нарисовать схему дорог (рис. 1.16, а). На рисунке 1.16, б населённые пункты для краткости обозначены латинскими буквами.

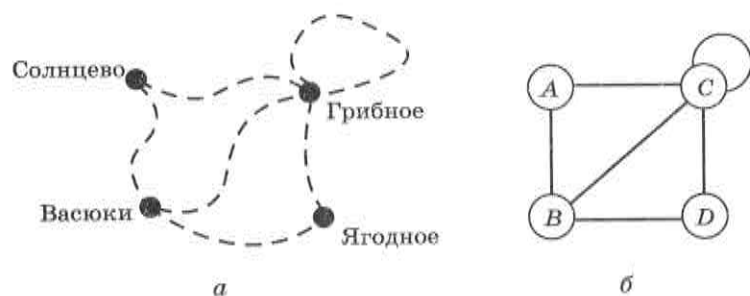


Рис. 1.16

Для исследования таких схем используют графы.

Граф — это набор вершин и связей между ними (рёбер).

Для хранения информации о вершинах и связях графа, соответствующего схеме на рис. 1.16, можно использовать таблицу (матрицу), показанную на рис. 1.17.

	A	B	C	D
A	0	1	1	0
B	1	0	1	1
C	1	1	1	1
D	0	1	1	0

Рис. 1.17

Единица на пересечении строки A и столбца B означает, что между вершинами A и B есть связь. Ноль в ячейке таблицы указывает на то, что связи между соответствующими вершинами нет. Такая таблица называется **матрицей смежности**. Она симметрична относительно главной диагонали (серые клетки в таблице).

На пересечении строки C и столбца C стоит единица, которая говорит о том, что в графе есть **петля** — ребро, которое начинается и заканчивается в одной и той же вершине.

Можно поступить иначе: для каждой вершины перечислить все вершины, с которыми связана данная вершина. В этом случае мы получим **список смежности**. Для рассмотренного графа список смежности выглядит так:

(A (B, C), B (A, C, D), C (A, B, C, D), D (B, C))

Строго говоря, граф — это математический объект, а не рисунок. Конечно, его можно нарисовать на плоскости (например, как на рис. 1.16, б), но матрица смежности и список смежности не дают никакой информации о том, как именно следует располагать вершины друг относительно друга. Для таблицы на рис. 1.17 возможны, например, варианты, показанные на рис. 1.18.

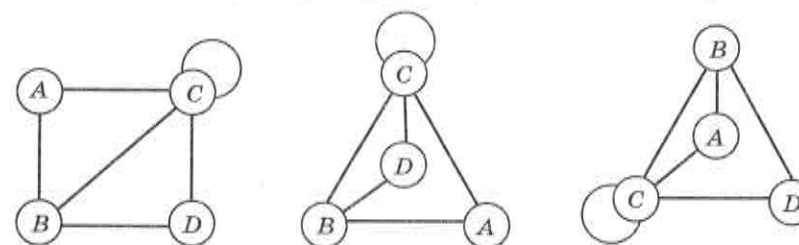


Рис. 1.18

В рассмотренном примере все вершины связаны, т. е. между любой парой вершин существует **путь** — последовательность рёбер, по которым можно перейти из одного узла в другой. Такой граф называется **связным**.



Связный граф — это граф, между любыми вершинами которого существует путь.

Теперь представьте себе, что дороги Васюки — Солнцево, Васюки — Грибное и Грибное — Ягодное завалило снегом (или размыло дождём) так, что по ним не пройти и не проехать (рис. 1.19).

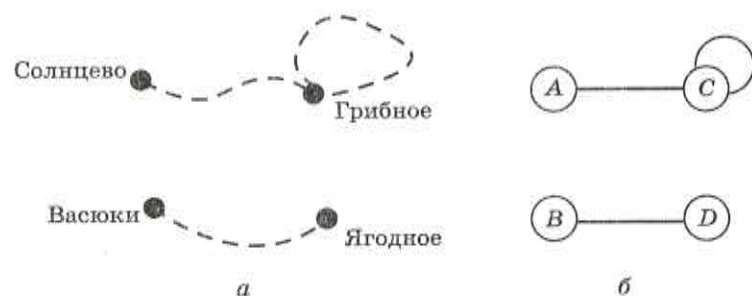


Рис. 1.19

Схему на рис. 1.19, а тоже можно считать графом (она подходит под определение), но в таком графе есть две несвязанные части, каждая из которых — связный граф. Такие части называют **компонентами связности**.

Вспоминая материал предыдущего пункта, можно сделать вывод, что дерево — это частный случай связного графа. Но у дерева есть одно важное свойство — в нём нет замкнутых путей (**циклов**). Граф на рис. 1.17 не является деревом, потому что в нём есть циклы: $ABCA$, $BCDB$, $ABDCA$.



Дерево — это связный граф, в котором нет циклов.

Если в первом примере с дорогами нас интересуют ещё и расстояния между поселками, каждой связи нужно сопоставить число (**вес**) (рис. 1.20).

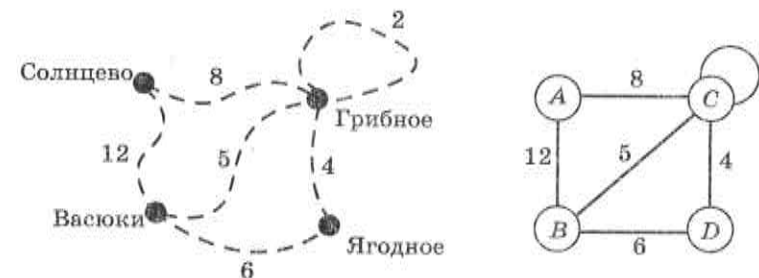


Рис. 1.20

Такой граф называется **взвешенным**, поскольку каждое ребро имеет свой вес. Весом может быть не только расстояние, но и, например, стоимость проезда или другая величина.

Как хранить информацию о таком графе? Ответ напрашивается сам собой — нужно в таблицу записывать не 1 или 0, а вес ребра. Если связи между двумя вершинами нет, на бумаге можно оставить ячейку таблицы пустой, а при хранении в памяти компьютера записывать в неё условный код, например -1 или очень большое число. Такая таблица называется **весовой матрицей**, потому что содержит веса рёбер. В данном случае она выглядит так, как показано на рис. 1.21.

	A	B	C	D
A		12	8	
B	12		5	6
C	8	5		4
D		6	4	

Рис. 1.21

Так же как и матрица смежности, весовая матрица симметрична относительно главной диагонали. Нижняя ячейка в столбце A и верхняя в столбце D говорят о том, что между вершинами A и D нет ребра.

Если в графе немного вершин, весовая матрица позволяет легко определить наилучший маршрут из одной вершины в другую простым перебором вариантов. Рассмотрим граф, заданный весовой матрицей на рис. 1.22 (числа определяют стоимость поездки между соседними пунктами).

	A	B	C	D	E
A			3	1	
B			4	5	1
C	3	4			2
D	1	5			1
E		1	2	1	

Рис. 1.22

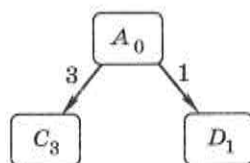


Рис. 1.23

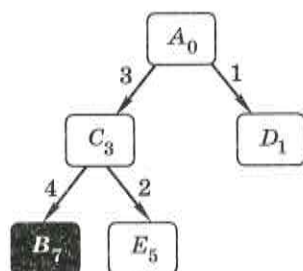


Рис. 1.24

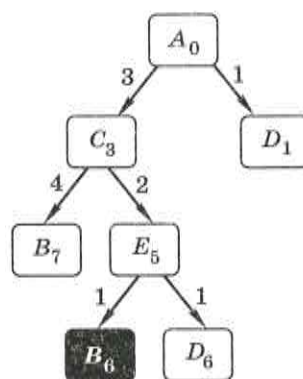


Рис. 1.25

Найдём наилучший путь из A в B — такой, при котором общая стоимость поездки минимальная. Сначала видим, что из пункта A напрямую в B ехать нельзя, а можно ехать только в C и D . Изобразим это на схеме (рис. 1.23).

Числа около рёбер показывают стоимость поездки по этому участку, а индексы у названий вершин показывают общую стоимость проезда в данную вершину из вершины A .

Теперь разберём варианты дальнейшего движения из вершины C (рис. 1.24). Вершину A уже не нужно рассматривать, так как мы из неё пришли.

Видим, что из C сразу можно попасть в B , стоимость проезда в этом случае равна 7. Но, возможно, это не самый лучший вариант, и нужно проверить ещё путь через вершину E . Действительно, оказывается, что можно сократить стоимость до 6 (рис. 1.25).

Исследовать дальше маршрут, содержащий цепочку $ACED$, нет смысла, потому что его стоимость явно будет больше 6. Аналогично строим вторую часть схемы (вы догадались, что схема представляет собой дерево!) (рис. 1.26).

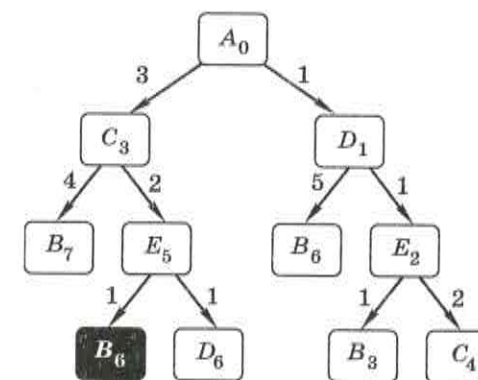


Рис. 1.26

Таким образом, оптимальный (наилучший) маршрут — $ADEB$, его стоимость — 3. Маршрут $ADEC$, не дошедший до вершины B , далее проверять не нужно, он не улучшит результат.

Конечно, для более сложных графов метод перебора работает очень долго, поэтому используются более совершенные (но более сложные) методы, о которых мы поговорим в 11 классе.

Наверное, вы заметили, что при изображении деревьев, которые описывают иерархию (подчинение), мы ставили стрелки от верхних уровней к нижним. Это означает, что для каждого ребра указывается направление, и двигаться можно только по стрелкам, но не наоборот. Такой граф называется **ориентированным** (или коротко **орграфом**). Он может служить, например, моделью системы дорог с односторонним движением. Матрица смежности и весовая матрица для орграфа уже не обязательно будут симметричными.

На схеме на рис. 1.27 всего две дороги с двусторонним движением, по остальным можно ехать только в одну сторону.

Рёбра в орграфе называют дугами. Дуга, в отличие от ребра, имеет начало и конец.

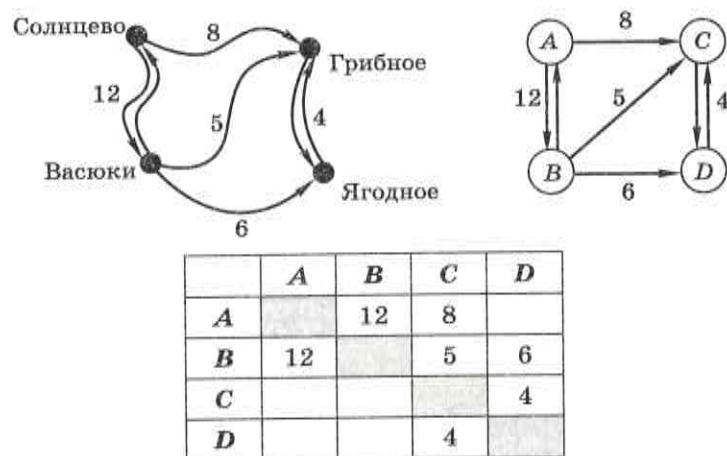


Рис. 1.27

Рассмотрим следующую задачу: определить количество возможных путей из вершины A в вершину K для ориентированного графа, показанного на рис. 1.28.

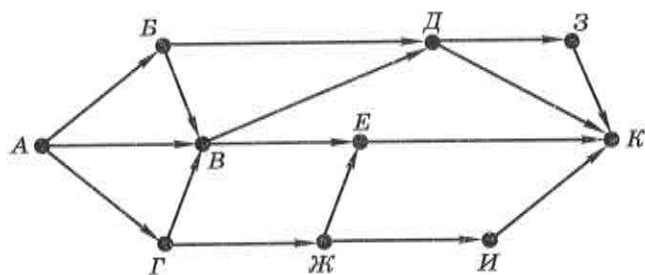


Рис. 1.28

Так как граф ориентированный, переходить в другую вершину можно только по стрелкам.

Будем двигаться по стрелкам от начальной вершины A. Около каждой вершины запишем количество возможных путей из вершины A в эту вершину.

Найдём все вершины, в которые можно попасть только из A: это вершины B и Г, записываем около них количество путей 1 (рис. 1.29).

В саму вершину A можно попасть (из A) единственным способом — никуда не ехать, поэтому отмечаем её числом 1.

Теперь ищем вершины, в которые можно попасть только из уже отмеченных вершин. Например, в вершину B есть один путь из A

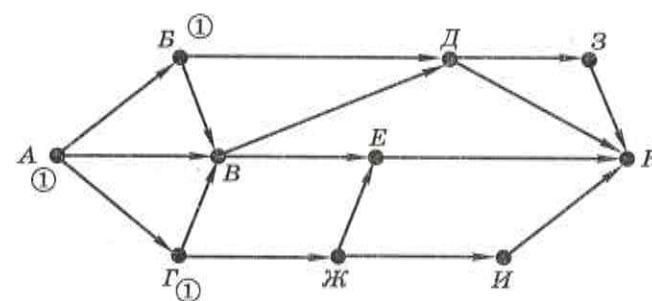


Рис. 1.29

напрямую, а также по одному пути через B и Г (так как эти вершины отмечены числом 1). Общее количество путей из A в B вычисляется как сумма отметок предыдущих вершин. В данном случае получаем всего три маршрута из A в B. В вершину Ж можно попасть только из Г, поэтому число путей в Г и Ж совпадает (рис. 1.30).

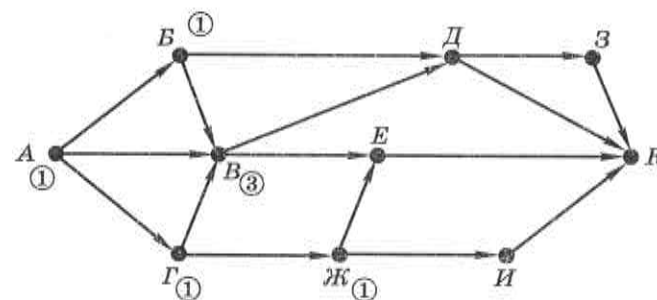


Рис. 1.30

В вершину D идёт один путь через B и три пути через B, поэтому общее число путей — четыре. Аналогично получаем четыре пути в вершину E: три пути через B и один — через Ж (рис. 1.31).

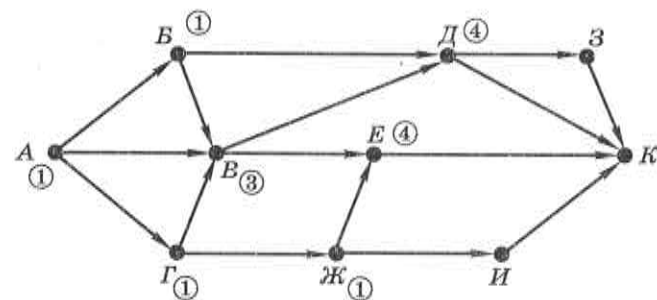


Рис. 1.31

Далее находим один путь из A в I (только через $Ж$) и четыре пути из A в $З$ (все через $Д$). В конечную вершину K можно приехать через вершины $Д$ (четыре пути), $З$ (четыре пути), E (четыре пути) и I (один путь), таким образом, общее количество различных путей равно 13 (рис. 1.32). Это и есть ответ задачи.

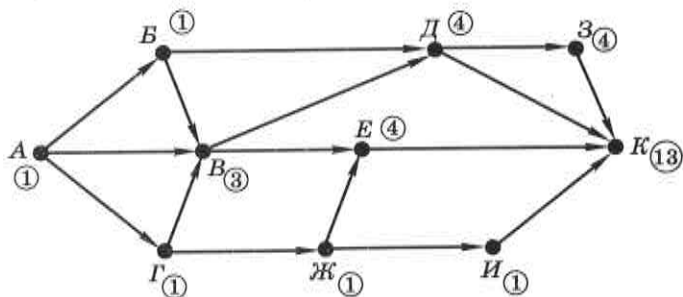


Рис. 1.32



Вопросы и задания

1. Что такое структурирование информации? Зачем оно нужно?
2. Что такое алфавитный порядок? Как поступают, если начальные символы слов совпали?
3. Расскажите, как используются оглавление, словарь и индекс для быстрого поиска нужной информации. Чем эти средства отличаются друг от друга?
4. Какими способами можно задать множество? Что такое пустое множество?
5. Приведите примеры множеств.
6. Чем отличаются множество и линейный список?
7. Что такое матрица?
8. Как можно записать табличные данные в виде списка?
9. Что такое иерархия? Приведите примеры.
10. Вспомните известные вам классификации, которые вы изучали на других предметах.
11. Как называется иерархическая структура в информатике?
12. Что такое корень, лист, родитель, сын, предок, потомок в структуре «дерево»?
13. В дереве 4 потомка, и все они являются листьями. Нарисуйте это дерево. Сколько в нём вершин?
14. В чём разница между понятиями «ребро» и «дуга»?
15. В чём различие понятий «дерево», «граф»?
16. Какой граф называется связным?
17. Что такое компонента связности?
18. Что такое петля? Как по матрице смежности определить, есть ли петли в графе?

19. Что такое орграф? Когда для представления данных используются орграфы? Приведите примеры.

20. Выберите наиболее подходящий способ структурирования информации для хранения:

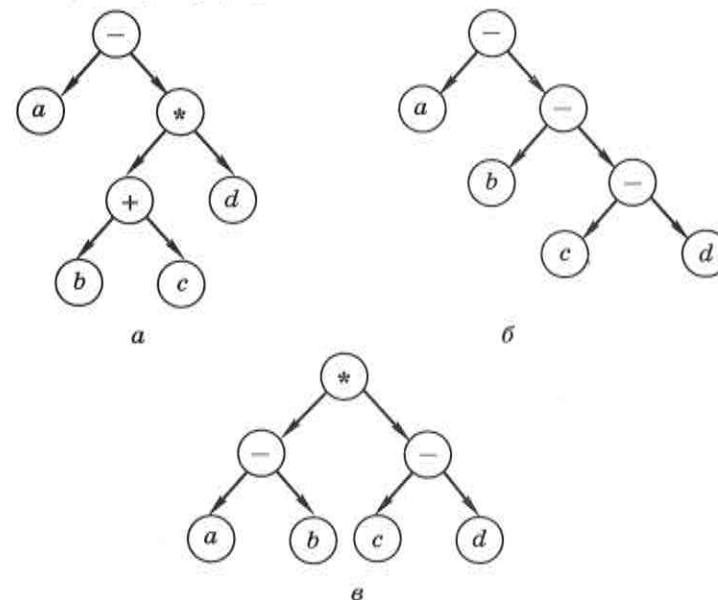
- а) данных о крупнейших озерах мира;
- б) рецепта приготовления шашлыка;
- в) схемы железных дорог;
- г) схемы размещения файлов на флэш-диске.

Подготовьте сообщение:

- а) «Как вычисляются арифметические выражения?»
- б) «Постфиксная и префиксная формы записи выражений»
- в) «Графы в практических задачах»
- г) «Диаграммы связей (mind maps)»
- д) «Системы классификации книг (ДДК, УДК)»

Задачи

1. Определите выражения, соответствующие каждому из деревьев, в «нормальном» виде со скобками (эту форму называют *инфиксной* — операция записывается между данными). Постройте для каждого из них постфиксную форму.



2. Постройте деревья, соответствующие следующим арифметическим выражениям. Запишите эти выражения в префиксной и постфиксной формах:

- а) $(a+b)*(c+2*d)$
- б) $(2*a-3*d)*c+2*b$

- в) $(a+b+2*c)*d$
- г) $3*a-(2*b+c)*d$

3. Вычислите выражения, записанные в постфиксной форме:

а) $12\ 6\ +\ 7\ 3\ -\ 1\ -\ * \ 12\ +$

б) $12\ 10\ -\ 5\ 7\ +\ * \ 7\ -\ 2\ *$

в) $5\ 6\ 7\ 8\ 9\ +\ -\ +\ -$

г) $5\ 4\ 3\ 2\ 1\ -\ -\ -\ -$

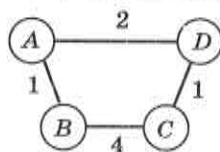
Запишите каждое из них в инфиксной и в префиксной формах, и постройте соответствующее дерево. Единственно ли такое дерево? В этом дереве назовите корень, листья и промежуточные вершины.

4. Нарисуйте граф, в котором 5 вершин и три компонента связности. Постройте его матрицу смежности. Сколько существует разных графов, удовлетворяющих условию задачи?

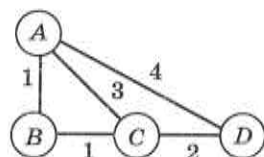
5. Структурируйте следующую информацию разными способами: «Между посёлками Верхний и Нижний есть просёлочная дорога длиной 10 км. Село Сергеево соединяется двумя асфальтовыми шоссе с Нижним (22 км) и Верхним (16 км). В село Солнечное можно доехать только из Сергеева по грунтовой дороге (5 км)». Можно ли сказать точно, как расположены эти пункты?

6. Для графа, полученного в предыдущей задаче, постройте матрицу смежности, список смежности, весовую матрицу. Является ли этот граф деревом?

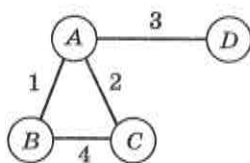
7. Постройте матрицы смежности и весовые матрицы графов:



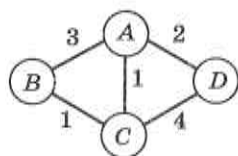
а



б



в



г

8. Постройте графы, соответствующие матрицам смежности.

а)

	A	B	C	D	E
A		0	1	1	0
B	0		1	0	1
C	1	1		0	1
D	1	0	0		0
E	0	1	1	0	

б)

	A	B	C	D	E
A		0	1	1	1
B	0		1	0	0
C	1	1		0	1
D	1	0	0		0
E	1	0	1	0	

в)

	A	B	C	D	E
A		0	1	1	1
B	0		1	0	1
C	1	1		0	1
D	1	0	0		0
E	1	1	1	0	

г)

	A	B	C	D	E
A		0	0	1	0
B	0		1	0	1
C	0	1		1	1
D	1	0	1		0
E	0	1	1	0	

9. Постройте графы, соответствующие весовым матрицам.

а)

	A	B	C	D	E
A		4	3		7
B	4			2	
C	3			6	
D		2	6		1
E	7			1	

б)

	A	B	C	D	E
A		2	5		6
B	2			3	
C	5				
D		3			1
E	6			1	

в)

	A	B	C	D	E
A			2	2	6
B				2	
C	2			2	
D	2	2	2		
E	6				

г)

	A	B	C	D	E
A		5	2		6
B	5			5	
C	2			2	
D		5	2		3
E	6			3	

10. Стоимость перевозок между пунктами, которые для краткости обозначены буквами A, B, C, D и E, задается таблицей (весовой матрицей графа). Нужно перевезти груз из пункта A в пункт B. Для каждого из четырех вариантов определите оптимальный маршрут и полную стоимость перевозки.

а)

	A	B	C	D	E
A			3	1	
B			4		2
C	3	4			2
D	1				
E		2	2		

б)

	A	B	C	D	E
A			3	1	1
B			4		
C	3	4			2
D	1				
E	1		2		

в)

	A	B	C	D	E
A			3	1	4
B			4		2
C	3	4			2
D	1				
E	4	2	2		

г)

	A	B	C	D	E
A				1	
B			4		1
C		4		4	2
D	1		4		
E		1	2		

11. Постройте орграфы, соответствующие весовым матрицам.

а)

	A	B	C	D	E
A			3	1	
B	2		4		2
C	3				
D	1				
E			2		

б)

	A	B	C	D	E
A			5	1	1
B			6	4	
C	3	4			2
D		2			
E			3		

в)

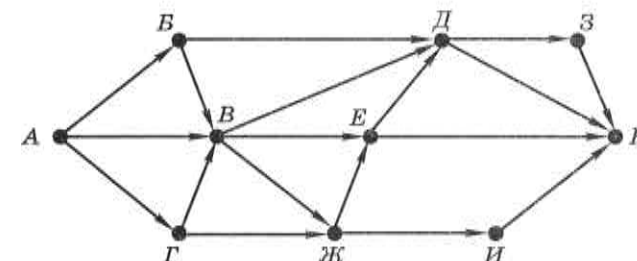
	A	B	C	D	E
A			3	1	4
B			4		2
C		4			2
D					
E	4		2		

г)

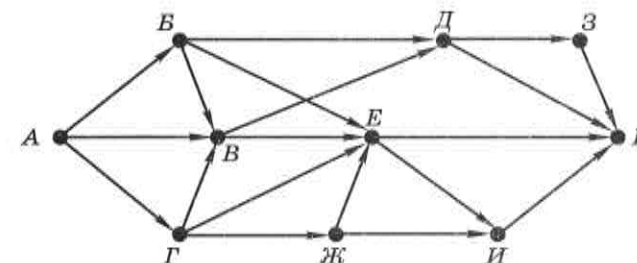
	A	B	C	D	E
A				1	
B			4		1
C	3	4		4	2
D	1	2	4		
E	1	1	2		

12. Для каждого из орграфов найдите количество различных маршрутов из вершины A во все остальные вершины.

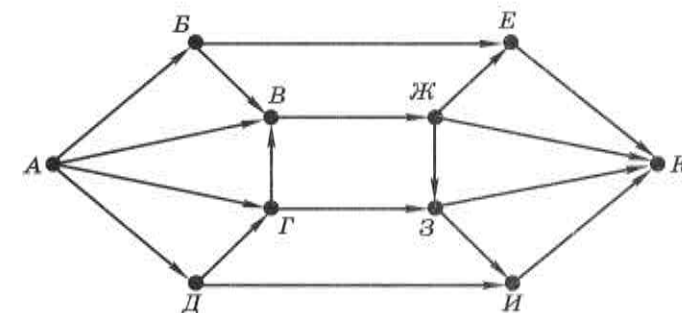
а)

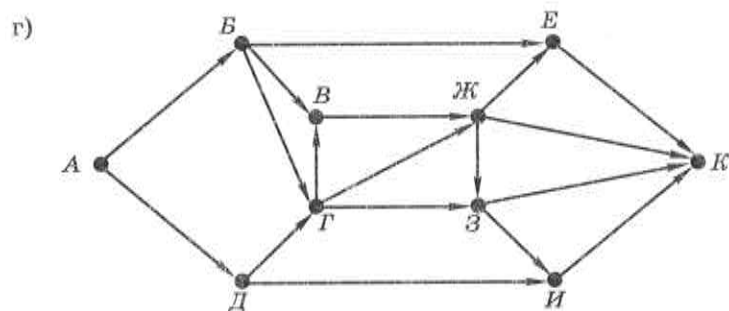


б)



в)





Практические работы к главе 1

- Работа № 1 «Оформление документа»
 Работа № 2 «Структуризация информации (таблица, списки)»
 Работа № 3 «Структуризация информации (деревья)»
 Работа № 4 «Графы»

ЗОР к главе 1 на сайте ФЦИОР (<http://fcior.edu.ru>)

- Виды и свойства информации
- Информация и информационные процессы
- Информация, информационные процессы в обществе, природе и технике
- Что изучает «Информатика»
- Классификация информационных процессов
- Единицы измерения информации

Самое важное в главе 1

- Человек получает информацию об окружающем мире с помощью органов чувств. Для расширения их возможностей можно использовать различные приборы.
- Зафиксированная (закодированная) информация — это данные. Компьютеры работают только с данными. Данные сохраняются на носителях.
- Информатика изучает широкий круг вопросов, связанных с автоматической обработкой данных.
- Основные информационные процессы — это передача и обработка информации (данных).
- Минимальная единица измерения количества информации — это бит. 1 бит — это количество информации, которое можно закодировать с помощью одной двоичной цифры (0 или 1).
- Для того чтобы облегчить работу с информацией, её нужно упорядочить (структурировать). Структурирование данных обязательно для компьютерной обработки.

Глава 2 Кодирование информации

§ 5 Язык и алфавит

Как записать информацию?

Для того чтобы хранить и передавать информацию, её необходимо как-то зафиксировать, например, записать с помощью символов (знаков) на каком-то языке.

Язык — это система знаков, используемая для хранения, передачи и обработки информации.

Естественные языки (русский, английский и др.) сформировались в результате развития человеческого общества и используются для общения людей.

Сначала древние люди овладели устной речью. Поскольку человек может издавать и различать на слух не так много звуков, он стал комбинировать их, составляя слова, каждому из которых приписывался некоторый смысл.

Затем люди стали записывать информацию, например, для передачи потомкам. В первое время жизненный опыт пытались зафиксировать в виде рисунков животных и предметов, затем пиктограмм (схематических изображений), иероглифов (рис. 2.1).

В большинстве современных языков используется **алфавитное письмо**, где каждый знак (или сочетание знаков) обозначает некоторый звук, так что с помощью небольшого набора знаков (алфавита) можно записать любые слова устной речи.

Алфавит — это набор знаков, который используется в языке.

Чаще всего подразумевается, что символы в алфавите расположены в определённом порядке.

Египетское письмо		Иероглифы (Китай)	
	Рука	日	Солнце
	Дом	月	Луна
	Кобра	雨	Дождь
	Лев	山	Гора
	Вода	马	Лошадь
	Рот	鱼	Рыба
	Мужчина	人	Человек
	Женщина	女	Женщина

Рис. 2.1

В алфавите русского языка 33 буквы, в английском алфавите — 26. К алфавиту языка, вообще говоря, нужно отнести пробел (пропуск между словами), цифры (знаки для записи чисел), знаки препинания, скобки.



Мощность алфавита — это количество знаков в алфавите.

Например, алфавит, состоящий из 33 русских букв, 10 цифр, пробела и 12 знаков препинания (точка, запятая, точка с запятой, вопросительный и восклицательный знаки, тире, двоеточие, мно-

готочие, кавычки, круглые скобки) имеет мощность 56 (а если различать прописные и строчные буквы, то 89).

Слово — это последовательность символов алфавита, которая используется как самостоятельная единица и имеет определённое значение.



Из слов составляются **предложения**, каждое из которых выражает определённую законченную мысль (сообщение, порцию информации). В языке определяются правила построения слов (**грамматика**), правила построения предложений (**синтаксис**) и правила расстановки знаков препинания (**пунктуация**).

С точки зрения теории информации, сообщение — это любой набор знаков некоторого алфавита. Определим, сколько различных сообщений можно построить с помощью заданного количества знаков. Пусть, например, алфавит состоит из четырёх знаков: @ # \$ %. С его помощью можно записать 4 разных сообщения из одного символа: @, #, \$ и %.

Теперь рассмотрим сообщения из двух знаков. Первый знак можно выбрать четырьмя способами, и для каждого из них есть 4 варианта выбора второго знака. Поэтому сообщений, состоящих из двух знаков, будет $4^2 = 16$:

@@	#@	\$@	%@
@#	##	\$#	%#
@\$	#\$	\$\$	%\$
@%	#%	\$%	%%

Рассуждая аналогично, получим, что трёхсимвольных сообщений будет $4^3 = 64$, а четырёхсимвольных — $4^4 = 256$ и т. д.

Если алфавит языка состоит из N символов (имеет мощность N), количество различных сообщений длиной L знаков вычисляется как

$$M = N^L$$



Естественные и формальные языки

Вы знаете, что в естественных языках кроме правил есть и исключения. Кроме того, одно и то же слово может иметь различный смысл в зависимости от *контекста*, т. е. фрагмента текста, в котором оно употребляется. Например, в книге по географии слово «рукав», скорее всего, будет означать ответвление русла реки, а не деталь одежды. Существует некоторая «свобода понимания»: слова и выражения могут пониматься немного по-разному в зависимости от множества условий: опыта человека, его культуры, уровня образованности, настроения и т. п. Так кулинарные рецепты часто содержат совет «добавить соль по вкусу», который каждый выполняет по-своему.

В то же время, например, в научных публикациях, такая ситуация недопустима, потому что смысл текста должен быть понят однозначно. В таких случаях используют языки специального типа, в которых каждое слово и словосочетание имеет чётко определенное единственное значение, и нет никаких исключений.



Формальный язык — это язык, в котором однозначно определяется значение каждого слова, а также правила построения предложений и придания им смысла.

Вот некоторые примеры формальных языков:

- математические формулы: $y = 3 \sin x + 1$;
- химические формулы и правила записи реакций:
 $2H_2 + O_2 = 2H_2O$;
- системы счисления (правила записи чисел с помощью специальных знаков — цифр): 12345, XXI;

Allegro animato

- нотная запись:



- язык записи шахматных партий: 1. e2–e4 e7–e5...;

- язык программирования Паскаль:

```
program qq;
begin
  write('Привет, Вася!')
end.
```

Все формальные языки — **искусственные**. В отличие от естественных языков, которые формировались в течение многих веков и неотделимы от истории каждого народа и его культуры, формальные языки разрабатываются людьми для обмена информацией в специальных областях знаний. Например, нотная запись позволяет сохранить и передать музыкальное произведение. К формальным языкам часто относят эсперанто — искусственный язык, разработанный в конце XIX века польским врачом Л. М. Заменгофом для международного общения.

В таблице 2.1 проведено сравнение естественных и формальных языков.

Таблица 2.1

Естественные языки	Формальные языки
Сформировались в результате развития общества	Созданы людьми целенаправленно
Используются для общения в быту	Используются для обмена информацией в специальных областях знаний
Часто встречаются слова с неточным и неясным содержанием	Не допускаются слова с неточным и неясным содержанием
Значения отдельных слов и предложений зависят не только от них самих, но и от их окружения (контекста)	Значения отдельных слов и предложений не зависят от контекста
Встречаются синонимы (разные слова имеют одинаковый смысл)	Как правило, синонимов нет
Встречаются омонимы (одно слово может иметь несколько значений)	Омонимов нет
Нет строгих правил образования предложений	Правила образования предложений строго определены
Для многих правил существуют исключения	Нет исключений из правил



Вопросы и задания

1. Что такое язык?
2. Зачем нужны языки?
3. Какие языки называются естественными?
4. Что такое алфавит языка?
5. Как вы думаете, почему алфавиты большинства современных языков содержат небольшое число знаков?
6. Чем отличается алфавитное письмо от использования иероглифов?
7. Какие правила существуют в языке? Как они называются?
8. В каких областях требуется использование формальных языков?
9. Чем отличается формальный язык от естественного?
10. Что такое контекст? Почему меню, которое появляется при щелчке правой кнопкой мыши на объекте, называют контекстным?
11. Приведите примеры формальных языков, о которых не упоминалось в тексте учебника.
12. Объясните, почему любой язык программирования — это формальный язык.
13. Как вы думаете, почему люди не отказываются от естественных языков и не переходят на формальные во всех областях?
14. Как вы думаете, почему любой формальный язык не является универсальным и хорошо подходит для записи информации только в определённой области?
15. Выберите из табл. 2.1 те свойства естественных языков, которые затрудняют и не позволяют полностью автоматизировать перевод с одного языка на другой. Приведите примеры.



Подготовьте сообщение:

- а) «Что такое алфавит?»
- б) «Зачем нужны формальные языки?»
- в) «Язык эсперанто»

§ 6

Кодирование

Как показано в главе 1 и в § 5, для хранения и передачи информации нужно записать её, зафиксировать на некотором языке (с помощью какого-то алфавита), т. е. закодировать. Это особенно важно в наше время, когда данные в компьютерных системах передаются, хранятся и обрабатываются в закодированном виде.



Кодирование — это представление информации в форме, удобной для её хранения, передачи и обработки. Правило такого преобразования называется **кодом**. Кодом называют также набор знаков закодированного сообщения.

В зависимости от конкретной задачи информация может кодироваться разными способами. Например, фраза «Привет, Вася!» может быть закодирована *транслитом* (так сокращённо называют транслитерацию — русский текст, записанный латинскими буквами): «Privet, Vasya!». Такой метод используют в электронных письмах, когда у одного собеседника (или у обоих) на компьютере нет поддержки русского языка. То же самое сообщение можно просто перевести на английский (или какой-то другой) язык, если собеседник не знает русского языка. А можно даже зашифровать: «Рсйгжу-!Гбта». Шифрование — это один из способов кодирования, при котором нужно скрыть смысл сообщения от посторонних¹.

Для кодирования числовой информации в разных ситуациях тоже используют разные способы. Например, число 21 можно записать как XXI (в римской системе счисления) или «двадцать один» (в финансовых документах).

Код Морзе

Долгое время для передачи сообщений по телеграфу и радио применялся код Морзе² (азбука Морзе), предложенный американским художником и изобретателем Самюэлем Морзе. В этом коде все буквы и цифры кодируются в виде различных последовательностей точек и тире (рис. 2.2).



Самюэль Морзе
(1791–1872)

¹ Это сообщение зашифровано с помощью шифра Цезаря. Попробуйте разгадать этот шифр и сформулировать правила кодирования и декодирования.

² Код Морзе применялся в британском флоте с 1865 г. для передачи сообщений с помощью флажков (днем) и фонарей (ночью). Для этой же цели использовали прожектора, у которых закрывали и открывали специальные жалюзи, а также сирены (для звуковой связи). С начала XX века код Морзе начали применять в радиосвязи.

Код Морзе для русских букв и цифр

А	•—	О	— — —	Э	• • — • •
Б	— • • •	П	• — — •	Ю	• • — —
В	• — —	Р	• — •	Я	• — • —
Г	— — •	С	• • •		
Д	— • •	Т	—	1	• — — — —
Е	•	У	• • —	2	• • — — —
Ж	• • • —	Ф	• • — •	3	• • • — —
З	— — • •	Х	• • • •	4	• • • • —
Й	• •	Ц	— • — •	5	• • • • •
И	• — — —	Ч	— — — •	6	— • • • •
К	— • —	Ш	— — — —	7	— — • • •
Л	• — • •	Щ	— — • —	8	— — — • •
М	— —	Ъ	— • • —	9	— — — — •
Н	— •	Ы	— • — —	0	— — — — —

Рис. 2.2

Код Морзе — *неравномерный*, т. е. коды символов могут быть разной длины. Для сокращения общего времени передачи буквы, которые встречаются чаще, имеют более короткие коды. Чтобы узнать, как часто встречается каждая буква в текстах, Морзе посетил типографию и подсчитал количество используемых литер с изображениями разных букв. Поэтому английская буква «Е», которая встречается в текстах чаще всего, получила код •. Коды Морзе для русских букв совпадают с кодами похожих по звучанию английских букв, например коды букв «Л» и «L» одинаковы¹.

Чтобы отделить последовательности (коды букв) друг от друга, вводят еще один символ — пробел (пауза). Например, имя «Вася», закодированное с помощью кода Морзе, выглядит так:

• — — • — • • • • — • —

Если бы не было разбивки на буквы, текст перестал бы расшифровываться однозначно. Например, сообщение • — — • — можно было бы прочитать как ВА, АК, ПТ или даже ЕМЕТ.

¹ Поэтому код Морзе для русских букв менее эффективен.

Двоичное кодирование

Для передачи информации обязательно нужно, чтобы свойства носителя как-то изменялись. Самый простой используемый код должен содержать, по крайней мере, два разных знака. Такое кодирование называют **двоичным** (от слова «два»), оно используется практически во всех современных компьютерах.

Двоичное кодирование — это кодирование с помощью двух знаков.

Например, сообщение АБАГДАВ может быть закодировано с помощью кодовой таблицы

А	Б	В	Г	Д
000	001	010	010	111

следующим образом: 000001000011111000010.

Кодирование чисел с помощью нулей и единиц впервые применил в своей (механической) вычислительной машине немецкий мыслитель Готфрид Вильгельм Лейбниц в конце XVII века. Затем, уже в середине XX века, двоичное кодирование информации стало повсеместно применяться для электронных компьютеров.

Чаще всего используется равномерный код, когда все символы исходного сообщения кодируются с помощью одинакового количества двоичных знаков. Каждый знак соответствует выбору одного из двух вариантов (0 или 1), поэтому несет 1 бит информации.

Длина кода определяется количеством вариантов, которые нужно закодировать. Поскольку алфавит двоичного кода содержит 2 символа, применяя общую формулу (см. § 5), получаем количество различных сообщений длиной I битов:

$$N = 2^I.$$



Готфрид Вильгельм
Лейбниц
(1646–1716)

Если заданное количество вариантов не равно степени числа 2, выбирают длину кода с запасом. Например, для кодирования номера спортсмена в интервале от 1 до 200 нужно использовать не меньше, чем 8 битов, поскольку

$$2^7 = 128 < 200 \leq 256 = 2^8.$$

Если нужно передать информацию о номерах первых 20 спортсменов, пришедших к финишу, информационный объём такого сообщения будет равен $8 \cdot 20 = 160$ битов = 20 байтов.



Вопросы и задания

1. Что такое кодирование?
2. Зачем кодируют информацию?
3. Что такое код?
4. Какой алфавит используется в коде Морзе?
5. Какие буквы в коде Морзе имеют самые короткие коды? Почему?
6. Предложите, как можно изменить азбуку Морзе, чтобы сообщения на русском языке стали более короткими.
7. Запишите своё имя с помощью кода Морзе.
8. Почему в коде Морзе необходим символ-разделитель (пауза)?
9. В каком случае применяется транслитерация?
10. Где сейчас используются числа, записанные в римской системе счисления?
11. Как вы думаете, зачем в финансовых документах денежные суммы пишут прописью?
12. Какое кодирование называют двоичным?
13. Можно ли при двоичном кодировании использовать не 0 и 1, а другие знаки (например, буквы А и Б)?
14. Объясните, как при двоичном кодировании связаны длина сообщения и количество информации в нём.



Подготовьте сообщение

- а) «Код Морзе»
- б) «Код Бодо»
- в) «История двоичного кодирования»
- г) «Код Грея»
- д) «Шрифт Брайля»



Задачи

1. Сколько существует в коде Морзе различных последовательностей из точек и тире, длина которых равна 4 символа? 6 символов?
2. Сколько различных пятизначных чисел можно записать с помощью цифр 4 и 2?

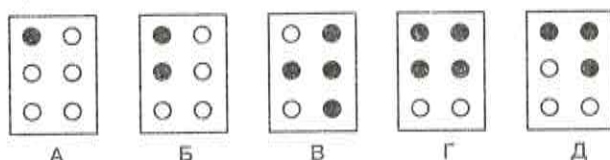
3. В алфавите языка племени «тамба-амба» две буквы: Й и Ы. Сколько различных 11-буквенных слов можно образовать в этом языке?
4. Алфавит языка «амба-карамба» состоит из 5 букв. Сколько различных четырехбуквенных слов можно образовать в этом языке?
5. В языке племени «тумба-юмба» разрешены только четырехбуквенные слова, которые можно образовывать из букв алфавита в любых комбинациях. Известно, что словарный запас языка составляет 81 слово. Какова мощность алфавита?
6. Некоторый язык содержит только трёхбуквенные слова, которые можно образовывать из букв его алфавита в любых комбинациях. Известно, что словарный запас языка составляет 216 слов. Какова мощность алфавита?
7. Какое наименьшее число символов должно быть в алфавите, чтобы с помощью всевозможных трёхбуквенных слов, состоящих из символов данного алфавита, можно было передать не менее 9 различных сообщений?
8. Световое табло состоит из лампочек. Каждая лампочка может находиться в одном из трех состояний («включено», «выключено» или «мигает»). Какое наименьшее количество лампочек должно находиться на табло, чтобы с его помощью можно было передать 18 различных сообщений?
9. Некоторое сигнальное устройство за одну секунду передает один из трёх сигналов. Сколько различных сообщений длиной в четыре секунды можно передать с помощью этого устройства?
10. Световое табло состоит из светящихся элементов, каждый из которых может гореть одним из двух различных цветов (или не гореть вообще). Сколько различных сообщений можно передать с помощью табло, состоящего из пяти таких элементов?
11. Для передачи сигналов на флоте используются специальные сигнальные флаги, вывешиваемые в одну линию (последовательность важна). Какое количество различных сообщений может передать корабль с помощью пяти сигнальных флагов, если на корабле имеются флаги четырёх различных видов (флагов каждого вида неограниченное количество)?
12. Вася и Петя передают друг другу сообщения, используя синий, красный и зелёный фонарики. Это они делают, включая по одному фонарику на одинаковое короткое время в некоторой последовательности. Количество вспышек в одном сообщении — 3 или 4, между сообщениями — паузы. Сколько различных сообщений могут передавать мальчики?
13. Для кодирования 300 различных сообщений используются 5 последовательных цветовых вспышек. Вспышки одинаковой длительности, для каждой вспышки используется одна лампочка определённого цвета. Лампочки скольких цветов должны использоваться при передаче (укажите минимально возможное количество)?

- *14. Некоторый алфавит содержит 4 различных символа. Сколько трёхбуквенных слов можно составить из символов этого алфавита, если символы в слове не могут повторяться?
- *15. В текстовом процессоре есть 5 кнопок, с помощью которых можно включать и выключать следующие режимы: жирный шрифт, курсив, подчёркивание, верхний индекс, нижний индекс. Сколько различных стилей оформления текста можно использовать?
16. Используя кодовую таблицу

А	Б	В	Г
00	01	10	11

закодируйте сообщение ГАВВАБ.

17. Шрифт Брайля — это специальный шрифт, с помощью которого незрячие люди могут читать. Для кодирования используются 6 точек, расположенных в два столбца. В каждой из них может быть выпуклость, которую человек воспринимает на ощупь. Коды Брайля первых букв русского алфавита (чёрная точка обозначает выпуклость):



Сколько различных символов можно закодировать с помощью кода Брайля?

18. Предложите какой-нибудь способ перехода от шрифта Брайля к двоичному кодированию.
19. В чём преимущества использования двоичного кодирования информации в современных компьютерах?
20. Сколько существует различных последовательностей из символов «плюс» и «минус» длиной ровно в пять символов?
21. На хранение целого числа отвели 12 битов. Сколько различных чисел можно закодировать таким образом?
22. Разведчик кодирует секретные сообщения, расставляя крестики и нолики в ячейки таблицы. Всего он может закодировать 512 сообщений. Сколько ячеек в таблице у разведчика?
23. Шахматная доска состоит из 8 столбцов и 8 строк. Какое минимальное количество битов потребуется для кодирования координат одной шахматной фигуры?
24. Какое минимальное количество битов потребуется для кодирования одного из натуральных чисел, меньших 60?

25. Для кодирования значений температуры воздуха (целое число в интервале от -50 до 40) используется двоичный код. Какова минимальная длина двоичного кода, необходимого для кодирования одного измерения?
26. В сельскохозяйственном институте изучают всхожесть семян растений. Результатом одного измерения является целое число от 0 до 100% , которое записывается с помощью минимально возможного количества битов. Всего исследовано 60 партий семян. Определите информационный объём результатов наблюдений.
27. Обычный дорожный светофор без дополнительных секций подаёт шесть видов сигналов (непрерывные красный, жёлтый и зелёный, мигающие жёлтый и зелёный, мигающие красный и жёлтый одновременно). Электронное устройство управления светофором последовательно воспроизводит записанные сигналы. Подряд записано 100 сигналов светофора. Определите информационный объём этого сообщения.
28. В некоторой стране автомобильный номер длиной 6 символов составляется из заглавных букв (всего используется 12 букв) и десятичных цифр в любом порядке. Каждый символ кодируется одинаковым и минимально возможным количеством битов, а каждый номер — одинаковым и минимально возможным количеством байтов. Определите объём памяти, необходимый для хранения 32 автомобильных номеров.
29. В базе данных хранятся записи, содержащие информацию о датах. Каждая запись содержит три поля: год (число от 1 до 2100), номер месяца (число от 1 до 12) и номер дня в месяце (число от 1 до 31). Каждое поле записывается отдельно от других полей с помощью минимально возможного числа битов. Определите минимальное количество битов, необходимое для кодирования одной записи.
30. В соревнованиях по ориентированию участвуют 768 спортсменов. Специальное устройство регистрирует финиш каждого из участников, записывая его номер с использованием минимально возможного количества битов, одинакового для каждого спортсмена. Каков будет информационный объём сообщения (в байтах), записанного устройством, после того как финишируют 200 спортсменов?

Декодирование

Декодирование — это восстановление сообщения из последовательности кодов.

Сообщения, записанные с помощью равномерного кода, всегда декодируются однозначно. Для этого достаточно разбить сообще-

ние на группы битов известной длины и восстановить исходный текст, используя кодовую таблицу.

В случае неравномерного кода декодирование также однозначно, если используется разделитель между символами. Например, закодированное сообщение

• — — • • — • — • — — — • — —

можно восстановить, используя код Морзе «в обратную сторону»: в этой строке закодирована фамилия «Петров».

В некоторых случаях даже при использовании неравномерного кода не требуется вводить символ-разделитель. Для этого достаточно выполнение условия Фано: ни одно кодовое слово не совпадает с началом другого кодового слова. Такой код называют **префиксным**.

Пример 1. Пусть для кодирования первых 5 букв русского алфавита используется таблица:

А	Б	В	Г	Д
000	10	01	110	001

Это неравномерный код, поскольку в нём есть двух- и трёхсимвольные кодовые слова. Построим для этой кодовой таблицы дерево, в котором от каждого узла (кроме листьев) отходят два ребра, помеченные цифрами 0 и 1. Чтобы найти код символа, нужно пройти по стрелкам от корня дерева к нужному листу, выписывая метки стрелок, по которым мы переходим (рис. 2.3).

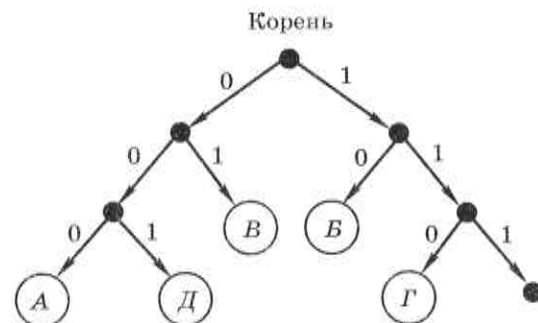


Рис. 2.3

Заметим, что ни один символ не лежит на пути от корня к другому символу. Это значит, что условие Фано выполняется, и любую правильную кодовую последовательность можно однозначно декодировать с начала сообщения. Например, рассмотрим цепочку 1100000100110. Букв с кодами 1 и 11 в таблице нет, поэтому сообщение начинается с буквы Г — она имеет код 110:

Г
110 | 0000100110

Следующий (единственно возможный) код — 000, это буква А:

Г А
110 | 000 | 0100110

Аналогично декодируем всё сообщение:

Г А В Д Б
110 | 000 | 01 | 001 | 10

Пример 2. Рассмотрим другую кодовую таблицу:

А	Б	В	Г	Д
000	01	10	011	100

Здесь условие Фано не выполняется, поскольку код буквы Б (01) является началом кода буквы Г (011), а код буквы Д (100) начинается с кода буквы В (10). Дерево для этой кодовой таблицы выглядит так (рис. 2.4).

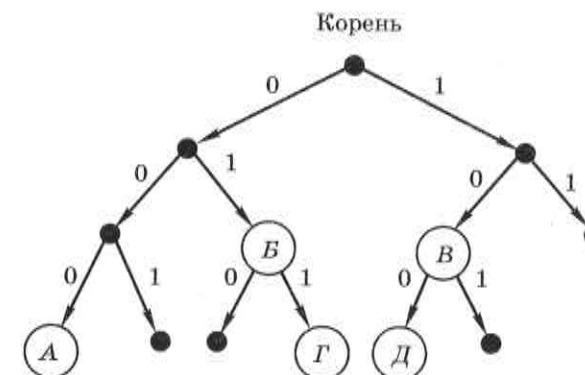


Рис. 2.4

Тем не менее можно заметить, что выполнено «обратное» условие Фано: ни одно кодовое слово не совпадает с окончанием другого кодового слова (такой код называют **постфиксным**). Поэтому закодированное сообщение можно однозначно декодировать с конца. Например, рассмотрим цепочку 011000110110. Последней буквой в этом сообщении может быть только В (код 10):

	В
0110001101	10

Вторая буква с конца — В (код 01):

	В	В
01100011	01	10

И так далее:

В	Д	Г	В	В
01	100	011	01	10

В общем случае (если код не является ни префиксным, ни постфиксным) декодировать сообщение удастся только перебором вариантов.

Пример 3. Декодируем сообщение 010100111101, закодированное с помощью кодовой таблицы:

А	Б	В	Г	Д
01	010	011	11	101

Здесь не выполняется ни «прямое», ни «обратное» условие Фано, поэтому декодировать сообщение однозначно, возможно, не удастся. На первом месте может быть, буква А или буква Б. Сначала предположим, что это буква А:

А0100111101

Тогда второй буквой также может быть буква А:

АА00111101.

Дальше декодировать не получается, потому что в таблице нет кодов 0, 00 и 001. Поэтому проверяем второй вариант: вторая буква — Б:

АБ0111101.

Третьей буквой может быть А:

АБА11101.

Тогда четвертая и пятая буквы определяются однозначно — это буквы Г и Д. Таким образом, один из подходящих вариантов — АБАГД.

Посмотрим, есть ли другие варианты. После сочетания АБ может стоять буква В:

АВВ1101.

тогда оставшиеся буквы — это ГА, а полное сообщение — АВВГА. Этот вариант тоже подходит.

Кроме того, на первом месте может стоять буква Б:

Б100111101.

но дальше декодировать не удастся, потому что в таблице нет кодов 1, 10 и 100. Таким образом, сообщение может быть декодировано двумя способами: АБАГД и АВВГА.

Пример 3 показывает, что неоднозначное декодирование возможно тогда, когда начало кода одной буквы совпадает с концом кода другой и можно переместить границу между кодами букв в сообщении. Например, последовательность 01011 может быть декодирована как АВ (01011) и БГ (01011). Следовательно, нужно обратить внимание на те цепочки, которые встречаются как в начале, так и в конце кодовых слов.

Покажем, как найти сообщения, которые декодируются неоднозначно. Для таблицы из примера 3 построим граф Ал. А. Маркова следующим образом.

1. Определим все последовательности, которые совпадают с началом какого-то кодового слова и одновременно с концом какого-то кодового слова; в данном случае это три последовательности:

0 (начало кода буквы А и конец кода буквы Б)

1 (начало кода буквы Г и конец кода буквы Д).

10 (начало кода буквы Д и конец кода буквы В).

Последовательности 01 и 11 не учитываем, потому что они совпадают с кодами букв А и Г.

2. Добавим к этому множеству $\{0, 1, 10\}$ пустую строку, которую обычно обозначают буквой Λ (прописная греческая буква «лямбда»); элементы полученного множества $\{\Lambda, 0, 1, 10\}$ становятся вершинами графа (рис. 2.5).
3. Соединим вершины дугами (направленными рёбрами) по такому правилу: две вершины X и Y соединяются дугой, если последовательная запись кода вершины X , кода некоторой буквы (или нескольких букв) и кода вершины Y даёт код ещё одной буквы (рис. 2.6).

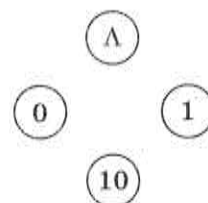


Рис. 2.5

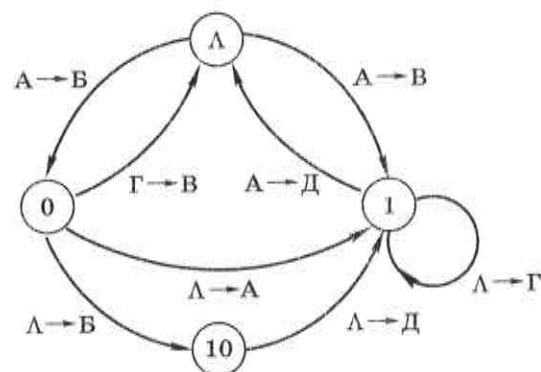


Рис. 2.6

Например, последовательная запись пустой строки (Λ), кода буквы A (01) и цепочки 0 даёт цепочку 010 , которая совпадает с кодом буквы B ; поэтому рисуем дугу из вершины Λ в вершину 0 ; у этой дуги пишем $A \rightarrow B$, и т. д. Поскольку код буквы Γ можно записать как $11 = 1\Lambda 1$, у вершины 1 появляется петля $\Lambda \rightarrow \Gamma$.

Любое сообщение декодируется однозначно тогда и только тогда, когда в полученном таким образом графе нет циклов, включающих вершину « Λ ».

В нашем графе есть несколько таких циклов, например:

- цикл $\Lambda 0 \Lambda$, соответствующий сообщению $\Lambda A 0 \Lambda = 01011$; это сообщение может быть расшифровано как AB и как $B\Gamma$;
- цикл $\Lambda 1 \Lambda$, соответствующий сообщению $\Lambda A 1 \Lambda = 01101$; это сообщение может быть расшифровано как AD и как BA ;

- цикл $\Lambda 0 1 \Lambda$, соответствующий сообщению $\Lambda A 0 1 \Lambda = 010101$; это сообщение может быть расшифровано как AAA и как BD ;
- цикл $\Lambda 0 1 0 1 \Lambda$, соответствующий сообщению $\Lambda A 0 1 0 1 \Lambda = 01010101$; это сообщение может быть расшифровано как ABD и как BDA .

Кроме того, из-за петли у вершины 1 неоднозначно декодируется любая последовательность вида $01\dots 101$, где многоточие обозначает любое количество единиц. Например, сообщение 0111101 может быть декодировано как AGD или BGA (см. пример 3).

Пример 4. Существуют коды, для которых условия Фано не выполняются, но все сообщения однозначно декодируются. В кодовой таблице

Λ	B	V
0	11	010

код буквы A совпадает как с началом, так и с окончанием кода буквы B , т. е. этот код не является ни префиксным, ни постфиксным.

Проверим, можно ли однозначно декодировать сообщения, построенные с помощью такого кода. Множество последовательностей, которые совпадают с началом и концом кодовых слов, состоит из пустой строки и единицы: $\{\Lambda, 1\}$. Граф, построенный с помощью приведённого выше алгоритма, содержит две вершины и одну петлю (рис. 2.7).

В этом графе нет цикла, содержащего вершину Λ , поэтому любое сообщение, записанное с помощью такого кода, декодируется однозначно. Это можно показать и простыми рассуждениями:

- 1) все цепочки 11 в сообщении — это коды букв B , иначе они не могут образоваться;
- 2) все цепочки 010 — это коды букв V ;
- 3) остальные символы сообщения могут быть только нулями — это коды букв A .

Иногда при кодировании и декодировании происходит искажение сообщения. Например, известно, что перевод художественных текстов (особенно стихов) на другой язык и затем обратный перевод может изменить их до неузнаваемости.

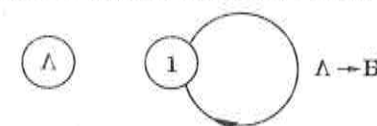


Рис. 2.7



Вопросы и задания

1. Что такое декодирование?
2. Всегда ли удаётся однозначно декодировать сообщение? В каких случаях это может быть не так?
3. Перечислите достаточные условия, при которых можно однозначно декодировать сообщение с неравномерным кодом.
4. В каких случаях для декодирования приходится использовать перебор вариантов?



Задачи

1. Расшифруйте сообщение, записанное с помощью кода Морзе, которое используется как международный сигнал бедствия:

... — — — ...

2. Покажите с помощью дерева, что кодовая таблица из примера 2 удовлетворяет «обратному» условию Фано.
3. Для кодирования сообщения используется таблица

А	Б	В	Г	Д
10	11	001	010	011

Найдите все способы декодирования сообщения 1111001011.

4. Для кодирования сообщения используется таблица

А	Б	В	Г	Д
01	11	100	010	110

Найдите все способы декодирования сообщения 1111001001100.

5. Для кодирования сообщения используется таблица

А	Б	В	Г	Д
0	11	101	110	111

Найдите все способы декодирования сообщения 1111001010.

6. Для кодирования сообщения используется таблица

А	Б	В	Г	Д
0	10	1	110	111

Найдите все способы декодирования сообщения 01110011.

7. Для кодирования сообщения используется таблица

А	В	С	Д	Е
000	01	100	10	011

Декодировать сообщение 0110100011000.

8. Для кодирования сообщения, состоящего только из букв А, В, С, Д и Е, используется неравномерный двоичный код:

А	В	С	Д	Е
000	11	01	001	10

Какие из сообщений могли быть закодированы с помощью этого кода?

- 1) 110000010011110
- 2) 110000011011110
- 3) 110001001001110
- 4) 110000001011110

- *9. Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный код: А = 0, Б = 10, В = 110. Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?
- *10. Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный код: А = 0, Б = 100, В = 101. Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?
- *11. Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный код: А = 01, Б = 1, В = 001. Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?

- *12. Для передачи по каналу связи сообщения, состоящего только из букв А, В, Г, решили использовать неравномерный код: А = 0, В = 100, Г = 110. Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?
- *13. Для передачи по каналу связи сообщения, состоящего только из букв А, В, Г, решили использовать неравномерный код: А = 00, В = 11, Г = 100 и Г = 10. Определите, допускает ли такой код однозначное декодирование сообщения. Выполняется ли для него условие Фано?

§ 7

Дискретность

Аналоговые и дискретные сигналы

Как вы уже знаете, информация передаётся в закодированном виде с помощью сигналов. Согласно определению из § 2, сигнал — это изменение свойств носителя, которое используется для передачи информации. Изменение выбранного свойства (например, силы тока, напряжения, освещённости) во времени можно описать в виде функции. Далее такую функцию тоже будем называть сигналом, как это принято в электронике и вычислительной технике.

В любой компьютерной системе очень важно обеспечить надёжный обмен данными в условиях, когда на сигнал действуют помехи. Поэтому желательно выбрать такой способ кодирования информации, который позволяет лучше всего решить эту задачу.

Элементы электронных устройств, как правило, обмениваются данными с помощью электрических сигналов; для получения информации приёмник должен измерить этот сигнал (чаще всего — напряжение на контактах). В таких устройствах, как радиоприёмник и микрофон, изменение электрического сигнала может произойти в любой момент и быть любым (в пределах допустимого диапазона). Такие сигналы называют **аналоговыми**.



Аналоговый сигнал — это сигнал, который в любой момент времени может принимать любые значения в заданном диапазоне.

Органы чувств человека воспринимают информацию в аналоговой форме: свет, звуковые волны, вкус, запах и т. п. Поэтому

раньше большинство технических устройств для работы с информацией (телефоны, магнитофоны, фотоаппараты) тоже были аналоговыми.

В 60-х годах XX века были широко распространены *аналоговые компьютеры*, которые выполняли вычисления с аналоговыми сигналами (сложение, вычитание, умножение, извлечение квадратного корня). Однако они решали достаточно узкий круг задач (моделирование законов движения), и их точность была невысока.

Дело в том, что при передаче сигнала всегда есть помехи, которые искажают его значения. В большинстве случаев эти искажения — случайные ошибки, не поддающиеся учёту. Фактически приёмник получает не исходный сигнал, посланный источником (сплошная линия на рис. 2.8), а искажённый (штриховая линия).

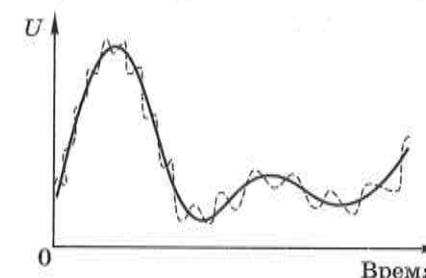


Рис. 2.8

Вспомним, что аналоговый сигнал может принимать любые значения в некотором диапазоне. «Очистить» его от помех в общем случае нельзя, потому что невозможно понять, искажён он или на самом деле имеет такое значение. Кроме того, дополнительные ошибки (погрешности) вносятся самими аналоговыми устройствами при измерении, передаче и обработке сигнала.

Если использовать аналоговые компьютеры, мы будем при каждом расчёте с одинаковыми исходными данными получать несколько отличающиеся результаты. Кроме того, при копировании аналоговая информация искажается (например, при каждом копировании звукозаписи на магнитной ленте качество копии ухудшается).

Эта ситуация не устраивала инженеров, разрабатывающих компьютеры, и они нашли интересное решение: если не удаётся точно измерить сигнал, нужно вообще отказаться от его измерения, а просто через некоторый интервал времени определять,

в каком из двух состояний находится сигнал (эти состояния можно обозначить как 1 и 0)¹. При использовании такого подхода мы получаем огромное преимущество: при небольших помехах искажение сигнала не влияет на передачу данных: если напряжение выше некоторого порога U_1 , считается, что сигнал равен 1, а все сигналы, меньшие другого порога U_0 , считаются равными нулю (рис. 2.9).

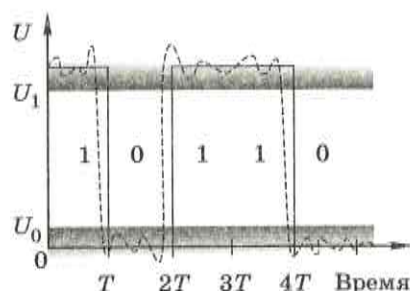


Рис. 2.9

Сигналы, с которыми работает компьютер, называются **дискретными** или **цифровыми**. Они обладают двумя важными свойствами:

- изменяются только в отдельные моменты времени (*дискретность по времени*);
- принимают только несколько возможных значений (*дискретность по уровню*).

Дискретный (цифровой) сигнал — это последовательность значений, каждое из которых принадлежит некоторому конечному множеству.

Обратите внимание на важный момент: мы естественным образом пришли к необходимости использования дискретных сигналов, когда стало необходимо точно и однозначно воспринимать передаваемую информацию с учётом неизбежных помех.

Так как каждому значению дискретного сигнала всегда можно поставить в соответствие определённый знак, такой сигнал

¹ Может быть и наоборот: 0 обозначает, что сигнал есть, а 1 — что сигнала нет.

можно рассматривать как сообщение, записанное с помощью конечного набора знаков (алфавита).

Этот принцип применим не только к компьютерам. Переход от наскальных рисунков к алфавитному письму, где каждый знак имеет чётко определённое значение, — это тоже переход от аналоговых сигналов к дискретным, цель которого — максимально исключить неоднозначное понимание смысла. Код Морзе и двоичный код — это тоже дискретные коды.

Дискретизация

Поскольку данные в компьютерах передаются с помощью дискретных сигналов, компьютеры могут хранить и обрабатывать только **дискретную информацию**, т. е. такую, которая может быть записана с помощью конечного количества знаков некоторого алфавита. Поэтому для ввода любых данных в компьютер их нужно перевести в дискретный код.

Дискретность означает, что мы представляем нечто целое (непрерывное) в виде набора отдельных элементов. Например, картина художника — это аналоговая (непрерывная) информация, а мозаика, сделанная на её основе (рисунок из кусочков разноцветного стекла), — дискретная. Множество вещественных чисел непрерывно (между любыми двумя различными числами есть ещё бесконечно много других), а множество целых чисел — дискретно.

Дискретизация — это представление единого объекта в виде множества отдельных элементов.

Всем известно иррациональное число π содержит бесконечное количество знаков в дробной части. Если мы хотим записать, чему равно π , необходимо остановиться на каком-то знаке, отбросив остальные, например: $\pi \approx 3,14$. Таким образом, мы перешли к дискретной информации, потому что рассматриваем только числа с шагом 0,01 — точки на числовой оси (рис. 2.10).

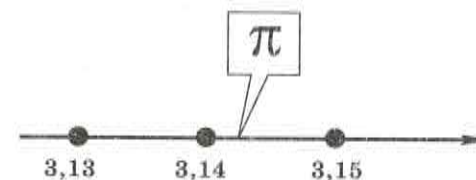


Рис. 2.10

Изменение высоты столбика термометра — это аналоговая информация, а *записанная* температура, округлённая до десятых долей градуса (например, $36,6^\circ$), — дискретная (рис. 2.11).

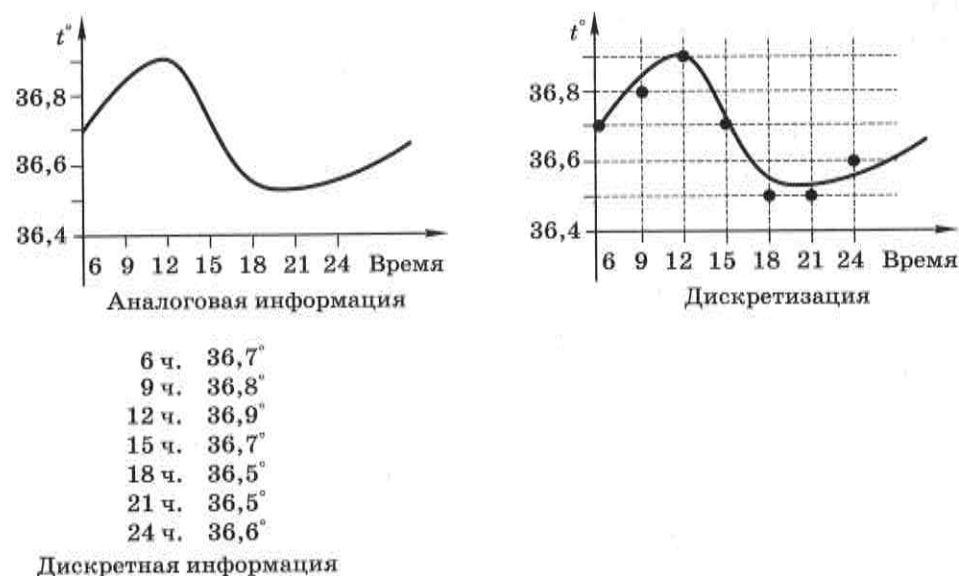


Рис. 2.11

Дискретность состоит в том, что записанные значения температуры изменяются скачкообразно (через $0,1^\circ$) — это *дискретизация по уровню*, или *квантование*. Кроме того, обычно температуру больного измеряют не непрерывно, а несколько раз в день — появляется *дискретизация по времени*.

Заметим, что при дискретизации, как правило, происходит *потеря информации*. В данном случае мы, во-первых, потеряли информацию об изменении температуры между моментами измерений и, во-вторых, исказили измеренные значения, округлив их до десятых (каждая дискретизация, и по времени, и по уровню, вносит свою ошибку). Чтобы уменьшить ошибки, нужно уменьшить шаг дискретизации — измерять температуру чаще, записывать показания термометра до тысячных долей градуса. Однако в любой практической задаче есть некоторый предел, после которого увеличение точности уже никак не влияет на конечный результат.

Из приведённого примера понятно, что непрерывность или дискретность — это не свойство самой информации, а свойство её *представления*. В данном случае информация — это сведения об изменении температуры человека в течение дня. Если бы температура измерялась постоянно и записывалась самописцем (в виде графика), можно было бы говорить о том, что информация представлена в аналоговой (непрерывной) форме.

Ещё один пример — аналоговые («стрелочные») и цифровые вольтметры, которые измеряют одну и ту же величину, но выводят результат измерения в разном виде (рис. 2.12).

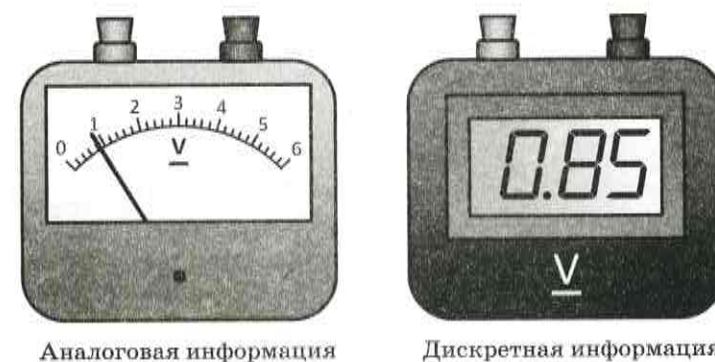


Рис. 2.12

Теперь подумаем, как записать аналоговую величину, которая может принимать бесконечное множество значений. Вы уже знаете, что с помощью знаков алфавита, состоящего из N символов, можно закодировать $M = N^L$ разных сообщений длины L . Поэтому теоретически для записи аналоговой величины придется использовать бесконечное число знаков.

Итак, когда мы хотим записать (зафиксировать) информацию с помощью какого-то алфавита, нужно переходить к дискретному представлению. С одной стороны, это делает более надёжной передачу данных (если обе стороны одинаково понимают используемые знаки). С другой стороны, при дискретизации часть информации теряется.

Хотя аналоговую информацию невозможно точно представить в дискретном виде, при увеличении точности дискретизации свойства непрерывной и дискретной информации практически совпадают. Например, для точной записи числа π требуется бесконеч-

ное количество цифр, но в расчётах чаще всего достаточно знать это значение с точностью не более 10 знаков.

Идеальная непрерывность существует только в теории. Мы считаем дерево, пластмассу, металл непрерывными, но на самом деле они состоят из отдельных молекул, расположенных на некотором расстоянии друг от друга — это значит, что вещество дискретно. Иллюстрация в книге кажется нам сплошной, но при сильном увеличении видно, что она строится из отдельных точек (имеет «растр») (рис. 2.13).

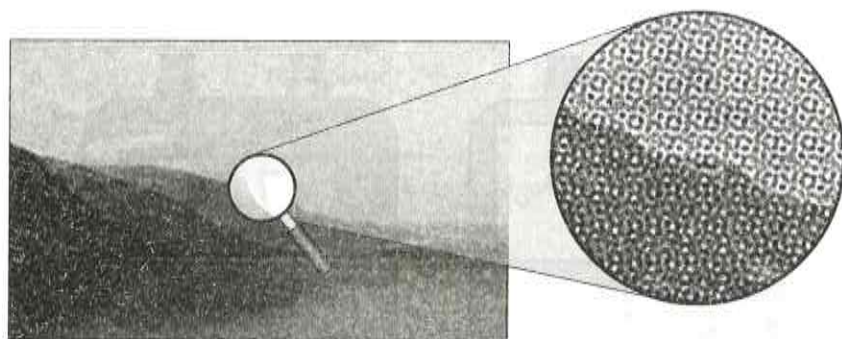


Рис. 2.13

«Плёночная» фотография считается аналоговой, но при увеличении снимка с фотоплёнки нельзя бесконечно получать все новые и новые детали — предел «уточнения» определяется величиной зерна светочувствительного материала.

Мы часто воспринимаем дискретные объекты как непрерывные, потому что наши органы чувств не позволяют различить отдельные элементы. Например, разрешающая способность глаза составляет около одной угловой минуты ($1' = 1/60$ часть градуса), это значение определяется размером элементов сетчатки глаза. Поэтому человек не может различить два объекта, если направления на них различаются меньше, чем на $1'$. Для того чтобы повысить разрешающую способность при наблюдении, применяют специальные приборы (например, бинокли и микроскопы).

Вопросы и задания

1. Что такое аналоговый сигнал?
2. Какие бытовые устройства работают с аналоговыми сигналами?
3. Какие системы связи используют аналоговые сигналы, а какие — дискретные?
4. Что такое аналоговые компьютеры? Почему они вышли из употребления?
5. Почему при использовании аналоговой техники передача информации всегда происходит с искажениями?
- *6. Как вы думаете, почему аналоговый сигнал нельзя полностью «очистить» от помех? Какую дополнительную информации нужно иметь, чтобы эта задача могла быть частично решена?
7. Что такое дискретный сигнал?
8. Почему с помощью дискретного сигнала можно передавать информацию практически без искажений? Искажается ли такой сигнал под воздействием помех?
9. Сигнал изменяется в моменты времени, кратные 1 секунде, и может принимать одно из 16 возможных значений. Является ли такой сигнал дискретным?
10. Выясните, какие музыкальные инструменты позволяют извлекать только дискретные звуки (заранее определённые ноты), а какие — звуки любой частоты.
11. Что такое дискретизация? Приведите примеры.
12. Что такое дискретизация по времени и дискретизация по уровню?
13. Приведите пример дискретизации сигнала только по уровню, только по времени. Нарисуйте графики таких процессов.
14. Объясните связь между дискретностью сигнала и алфавитным способом записи информации.
15. Почему при дискретизации, как правило, происходит потеря информации? В каких случаях потери информации не будет?
16. Как можно уменьшить потери информации при дискретизации?
17. Почему не стоит стремиться максимально уменьшить эти потери?
18. Приведите примеры, когда одна и та же информация может быть представлена в аналоговой и в дискретной форме.
19. Объясните фразу: «При увеличении точности дискретизации свойства непрерывной и дискретной информации практически совпадают».

Подготовьте сообщение:

- а) «Аналоговые вычислительные машины»
- б) «Аналоговые и дискретные измерительные устройства»
- в) «Непрерывность и дискретность в математике»
- г) «Непрерывность и дискретность в природе»

§ 8

Алфавитный подход к измерению количества информации

Представьте себе, что вы много раз бросаете монету и записываете результат очередного броска как 1 (если монета упала гербом) или 0 (если она упала «решкой»). В результате получится некоторое *сообщение* — цепочка нулей и единиц, например 0101001101001110. Вы наверняка поняли, что здесь используется *двоичное* кодирование — это сообщение написано на языке, *алфавит* которого состоит из двух знаков (символов): 0 и 1. Как вы знаете из § 1, каждая двоичная цифра несёт 1 бит информации, поэтому полная информация в сообщении 0101001101001110 равна 16 битов.

Теперь представим себе, что нужно закодировать программу для робота, который умеет выполнять команды «вперёд», «назад», «влево» и «вправо». Для этого можно использовать алфавит, состоящий из 4 символов: $\uparrow \downarrow \rightarrow \leftarrow$. Сколько информации содержится в сообщении $\uparrow \leftarrow \uparrow \uparrow \rightarrow \downarrow \downarrow \downarrow \rightarrow \leftarrow$? Каждый полученный символ может быть любым из 4-х символов алфавита, а для кодирования одного из 4-х вариантов требуется уже два бита. Поэтому полное сообщение из 11 символов содержит 22 бита информации.

Алфавитный подход к измерению количества информации состоит в следующем:

- 1) определяем мощность алфавита N (количество символов в алфавите);
- 2) по таблице степеней числа 2 определяем количество битов информации i , приходящихся на каждый символ сообщения — *информационную ёмкость (объём) символа*:

N , символов	2	4	8	16	32	64	128	256	512	1024
i , битов	1	2	3	4	5	6	7	8	9	10

- 3) умножаем i на число символов в сообщении L , это и есть полное количество информации:

$$I = L \cdot i.$$

Алфавитный подход к измерению информации

Обратим внимание на две важные особенности алфавитного подхода.

При использовании алфавитного подхода не учитывается, что некоторые символы могут встречаться в сообщении чаще других. Считается, что каждый символ несёт одинаковое количество информации.

Алфавитный подход не учитывает также частоты появления *сочетаний символов* (например, после гласных букв никогда не встречается мягкий знак).

Кроме того, никак не учитывается смысл сообщения, оно представляет собой просто набор знаков, которые приёмник, возможно, даже не понимает.

При использовании алфавитного подхода смысл сообщения не учитывается. Количество информации определяется только длиной сообщения и мощностью алфавита.

Во многих задачах такой подход очень удобен. Например, для устройств, передающих информацию по сети, её содержание не имеет никакого значения, важен только объём. Почтальону всё равно, что написано в письмах, важно только их количество, которое влияет на вес сумки. Для компьютера все данные — это последовательности нулей и единиц, смысла данных он не понимает.

Для вычисления информационного объёма текста чаще всего применяют именно алфавитный подход. Например, пусть требуется оценить количество информации в 10 страницах текста (на каждой странице 32 строки по 64 символа) при использовании алфавита из 256 символов. Задача решается так:

- 1) определяем информационную ёмкость одного символа: так как $256 = 2^8$, один символ несёт $i = 8$ битов, или 1 байт информации;
- 2) считаем количество символов на одной странице, в данном случае удобно использовать степени числа 2 ($32 = 2^5$, $64 = 2^6$): $2^5 \cdot 2^6 = 2^{11}$ символов на странице;

- 3) находим общее количество символов на 10 страницах:

$$L = 10 \cdot 2^{11} \text{ символов};$$

- 4) определяем информационный объём всего текста:

$$I = L \cdot i = 10 \cdot 2^{11} \cdot 1 \text{ байтов} = 10 \cdot 2^{11} \text{ байтов} = \\ = 10 \cdot 2^{11} \cdot (1/2^{10} \text{ Кбайт}) = 20 \text{ Кбайт}.$$



Вопросы и задания

1. В чём состоит алфавитный подход к измерению количества информации?
2. Приведите примеры ситуаций, когда смысл информации значения не имеет, а важен только её объём.
3. Учитывается ли при алфавитном подходе частота встречаемости символов в тексте?
4. Выберите все утверждения, справедливые для алфавитного подхода:
 - а) количество информации зависит от длины сообщения;
 - б) количество информации зависит от мощности алфавита;
 - в) чем больше мощность алфавита, тем больше количество информации;
 - г) важен смысл сообщения;
 - д) сообщение должно быть понятно для приёмника;
 - е) разные символы могут нести разное количество информации.
5. Технический документ перевели с одного языка на другой (считаем, что это было сделано максимально близко к тексту). Изменился ли смысл документа? Изменился ли его объём? Обоснуйте ответ.
6. Как вы думаете, почему компьютеру легко извлечь несколько предложений с конкретных страниц документа, но трудно составить аннотацию к документу?



Задачи

1. Сообщение состоит из 100 символов, используется алфавит, состоящий из 64 символов. Каков информационный объём этого сообщения?
2. Дан текст из 600 символов. Известно, что символы берутся из таблицы размером 16×32 , в которой все ячейки заполнены разными символами. Определите информационный объём текста в битах.
3. Для записи текста использовался алфавит, состоящий из 32 символов. Каждая страница текста содержит 32 строки. Информационный объём сообщения, состоящего из 5 страниц, составил 6400 байтов. Сколько символов в каждой строке текста?
4. Страница текста содержит 30 строк по 60 символов в каждой. Сообщение, состоящее из 4 страниц текста, имеет информационный объём 6300 байтов. Какова мощность алфавита?

5. Мощность алфавита равна 256. Сколько Кбайт памяти потребуется для сохранения 160 страниц текста, содержащего в среднем 192 символа на каждой странице?
6. Мощность алфавита равна 64. Сколько Кбайт памяти потребуется, чтобы сохранить 128 страниц текста, содержащего в среднем 256 символов на каждой странице?
7. Секретарь может набирать текст со скоростью 256 символов в минуту. Сколько Кбайт информации он сможет ввести в компьютер за 10 минут, если используется алфавит из 256 символов?
8. Для кодирования секретного сообщения используются 12 специальных знаков. При этом символы кодируются одним и тем же минимально возможным количеством битов. Чему равен информационный объём сообщения длиной в 256 символов?
9. Для кодирования нотной записи используются 7 знаков-нот. Каждая нота кодируется одним и тем же минимально возможным количеством битов. Чему равен информационный объём (в битах) сообщения, состоящего из 180 нот?
10. Объём сообщения равен 7,5 Кбайт. Известно, что данное сообщение содержит 7680 символов. Какова мощность алфавита?
11. Объём сообщения равен 12 Кбайт. Сообщение содержит 16 384 символа. Какова мощность алфавита?
12. Объём сообщения, содержащего 4096 символов, равен $1/512$ мегабайта. Какова мощность алфавита, с помощью которого записано это сообщение?
13. Два текста содержат одинаковое количество символов. Первый текст составлен в алфавите мощностью 16 символов, а второй текст — в алфавите из 256 символов. Во сколько раз количество информации во втором тексте больше, чем в первом?
14. Алфавит языка первого племени содержит 8 знаков, а алфавит языка второго племени — 16 символов. Племена обменивались сообщениями, состоящими из одинакового количества символов. Известно, что сообщение второго племени содержало 128 байтов информации. Каков информационный объём сообщения первого племени?
- *15. Два текста содержат одинаковое количество символов, но информационный объём второго текста в 1,5 раза больше, чем информационный объём первого. Определите мощности алфавитов, если известно, что в обоих алфавитах число символов меньше 10, и на каждый символ приходится целое число битов.
- *16. Два текста имеют одинаковый информационный объём, но количество символов во втором тексте в 3,5 раза больше, чем в первом. Определите мощности алфавитов, если известно, что в обоих текстах число символов меньше 200, и на каждый символ приходится целое число битов.

§ 9

Системы счисления

Далее в этой главе мы рассмотрим методы кодирования различных видов информации в компьютерной технике. Начнём с числовой информации, для обработки которой и были первоначально разработаны компьютеры. Кроме того, все остальные виды информации в компьютере хранятся как последовательности числовых кодов, поэтому важно разобраться, как можно записывать и обрабатывать числа.



Система счисления — это правила записи чисел с помощью специальных знаков — цифр, а также соответствующие правила выполнения операций с этими числами.

Первоначально люди считали на пальцах — это самый простой способ, который используется и сейчас. Один загнутый (или отогнутый) палец обозначал единицу (один день, одного человека, одного барана и т. п.). Такая система счисления называется **унарной** (от лат. *unus* — один). В качестве цифр унарной системы можно использовать камешки, узелки, счётные палочки (как в начальной школе), зарубки на дереве (как делал Робинзон Крузо) или на кости, чёрточки на бумаге, точки и другие одинаковые знаки или предметы.

С помощью унарной системы можно записывать только натуральные числа, причём запись больших чисел получается очень длинной (представьте себе, как записать миллион). Цифра в любой позиции числа, записанного в унарной системе, всегда обозначает единицу, поэтому это одна из **непозиционных систем счисления**.



Непозиционная система счисления — это такая система счисления, в которой значение цифры не зависит от её места в записи числа.

К непозиционным относится и десятичная египетская система счисления. Египтяне ввели 7 знаков-иероглифов, которые об-

означали степени числа 10 (чёрточка, хомут, верёвка, лотос, палец, лягушка, человек) — рис. 2.14.



Рис. 2.14

В этой системе, например, число 235 записывалось как

eeennnnlllll

В римской системе счисления (она также считается непозиционной) в качестве цифр используются латинские буквы: I обозначает 1, V — 5, X — 10, L — 50, C — 100, D — 500, M — 1000. Единицы, десятки, сотни и тысячи кодировались отдельными группами, например:

$$2368 = 2000 + 300 + 60 + 8 = (1000 + 1000) + (100 + 100 + 100) + (50 + 10) + (5 + 1 + 1 + 1) = \text{MMCCCLXVIII}.$$

Больше трёх одинаковых цифр подряд не ставили, поэтому число 4 записывали как IV. В такой записи меньшая цифра (I) стоит перед большей (V), поэтому она вычитается из неё. То есть:

$$\text{IV} = 5 - 1 = 4.$$

Аналогично записывались числа 9, 40, 90, 400 и 900:

$$\begin{aligned} \text{IX} &= 10 - 1 = 9, \text{XL} = 50 - 10 = 40, \text{XC} = 100 - 10 = 90, \\ \text{CD} &= 500 - 100 = 400, \text{CM} = 1000 - 100 = 900. \end{aligned}$$

Из-за этой особенности римскую систему нельзя считать полностью непозиционной, потому что значение меньшей цифры, стоящей слева от большей, меняется на отрицательное.

У римской системы есть несколько серьёзных недостатков:

- можно записывать только натуральные числа (что делать с дробными и отрицательными?);
- чтобы записывать большие числа, необходимо вводить всё новые и новые цифры (иногда использовались цифры с подчёркиванием или чертой снизу, что обозначало увеличение в 1000 раз: $\underline{\text{V}}$ — 5000, $\underline{\text{X}}$ — 10 000 и т. д.);
- сложно выполнять арифметические действия.

Сейчас римская система применяется для нумерации веков (XXI век), месяцев, глав в книгах, на циферблатах часов (например, на Спасской башне Московского Кремля).

В славянской системе счисления в качестве цифр использовались буквы алфавита, над которыми ставился знак Z («титло») (рис. 2.15).

A^{Z}	B^{Z}	Г^{Z}	Д^{Z}	Е^{Z}	С^{Z}	З^{Z}	И^{Z}	Ѡ^{Z}
1	2	3	4	5	6	7	8	9
I^{Z}	K^{Z}	L^{Z}	M^{Z}	Н^{Z}	O^{Z}	З^{Z}	П^{Z}	Ч^{Z}
10	20	30	40	50	60	70	80	90
P^{Z}	C^{Z}	T^{Z}	V^{Z}	Ф^{Z}	X^{Z}	Ψ^{Z}	W^{Z}	Ц^{Z}
100	200	300	400	500	600	700	800	900

Рис. 2.15

Если в ряду стояло несколько цифр, знак «титло» ставился только у предпоследней. Старшие цифры записывались *справа* от младших, например, число 11 записывалось как AI^{Z} . Славянская система счисления используется на циферблате часов Суздальского Кремля.



Вопросы и задания

1. Как можно закончить фразу: «Система счисления — это...»?
2. Что такое унарная система счисления? Приведите примеры её использования.
3. Какие недостатки имеет унарная система счисления?
4. Что такое непозиционная система счисления?
5. Какие системы счисления относятся к непозиционным?
6. Какие цифры используются в римской системе? Что они означают?
7. Можно ли называть римскую систему полностью непозиционной? Обоснуйте ответ.
8. Какое наибольшее число можно записать в классической римской системе счисления?
9. Где сейчас используется римская система?
10. Перечислите недостатки римской системы счисления. Как вы думаете, почему её не используют в компьютерах?

Подготовьте сообщение

- а) «Где применяется римская система счисления?»
- б) «Славянская система счисления»
- в) «Системы счисления разных народов»
- г) «Система остаточных классов»

Задачи

1. Переведите в римскую систему числа: 12, 345, 2999, 2444, 2888, 3777.
2. Переведите в десятичную систему числа: MCDXCIX, MMDCCXLVII, MDCXCIX.
3. Запишите в славянской системе числа: 15, 25, 38, 137, 596.

§ 10

Позиционные системы счисления

Основные понятия

Позиционная система счисления — это такая система счисления, в которой значение цифры («вес») полностью определяется её местом (позицией) в записи числа.

Пример позиционной системы счисления — привычная нам десятичная система. В числе 6375 цифра 6 обозначает тысячи (т. е. 6000), цифра 3 — сотни (300), цифра 7 — десятки (70), а цифра 5 — единицы:

$$6375 = 6 \cdot 1000 + 3 \cdot 100 + 7 \cdot 10 + 5 \cdot 1.$$

Алфавит системы счисления — это используемый в ней набор цифр.

Основание системы счисления — это количество цифр в алфавите (мощность алфавита).

В десятичной системе основание — 10, алфавит состоит из 10 цифр: 0, 1, 2, 3, 4, 5, 6, 7, 8 и 9. Число 10, вероятно, было выбрано потому, что люди сначала использовали для счета свои 10 пальцев на руках.



Разряд — это позиция цифры в записи числа. Разряды в записи целых чисел нумеруются с нуля справа налево.

В числе 6375 цифра 6 стоит в третьем разряде (тысячи, 10^3), 3 — во втором разряде (сотни, 10^2), 7 — в первом (десятки, 10^1), а 5 — в нулевом (единицы, 10^0). Не забывайте, что любое число (кроме нуля!) в нулевой степени равно 1. Поэтому

$$\begin{array}{l} \text{разряды} \rightarrow 3 \ 2 \ 1 \ 0 \\ 6375 = 6 \cdot 10^3 + 3 \cdot 10^2 + 7 \cdot 10^1 + 5 \cdot 10^0. \end{array}$$

Это так называемая **развёрнутая форма записи числа**. Из этой записи видно, что последняя цифра 5 — это остаток от деления числа на 10 (все остальные слагаемые делятся на 10); число, составленное из двух последних цифр (75) — это остаток от деления исходного числа на $100 = 10^2$ и т. д. Поэтому все числа, делящиеся на 100 без остатка, оканчиваются на два нуля.



Чтобы определить число, записанное в позиционной системе счисления, нужно значение каждой цифры умножить на основание системы счисления в степени, равной разряду этой цифры, и сложить полученные величины.

Число 6375 можно представить в другой форме — по *схеме Горнера*:

$$6375 = ((6 \cdot 10 + 3) \cdot 10 + 7) \cdot 10 + 5.$$

Эта форма позволяет найти число, используя только умножение и сложение (без возведения в степень).

Кроме десятичной системы, на практике используются ещё несколько позиционных систем:

- двоичная, восьмеричная и шестнадцатеричная в компьютерной технике;
- двенадцатеричная английская система мер (1 фут = 12 дюймов, 1 шиллинг = 12 пенсов);
- шестидесятеричная система измерения времени (1 час = 60 минут, 1 минута = 60 секунд).

Целые числа

Теперь можно записать аналогичные выражения для системы счисления с любым натуральным основанием $p > 1$. Её алфавит состоит из p цифр¹ от 0 до $p - 1$, т. е. «старшая» (наибольшая) цифра в позиционной системе счисления на единицу меньше, чем основание.

Рассмотрим четырёхзначное число $a_3a_2a_1a_0$, записанное в системе счисления с основанием p . Здесь a_3 , a_2 , a_1 и a_0 — отдельные цифры, стоящие соответственно в третьем, втором, первом и нулевом разрядах. Это число может быть записано в развёрнутой форме:

$$\begin{array}{l} \text{разряды} \rightarrow 3 \ 2 \ 1 \ 0 \\ a_3a_2a_1a_0 = a_3 \cdot p^3 + a_2 \cdot p^2 + a_1 \cdot p^1 + a_0 \cdot p^0 \end{array}$$

или с помощью схемы Горнера:

$$a_3a_2a_1a_0 = ((a_3 \cdot p + a_2) \cdot p + a_1) \cdot p + a_0.$$

Оба способа можно использовать для перевода числа из любой позиционной системы в десятичную систему. Например, пусть число 1234_5 записано в пятеричной системе счисления (с основанием 5). Нижний индекс 5 в записи 1234_5 обозначает основание системы счисления (для десятичной системы основание не указывают). Тогда:

$$1234_5 = 1 \cdot 5^3 + 2 \cdot 5^2 + 3 \cdot 5^1 + 4 \cdot 5^0 = 125 + 2 \cdot 25 + 3 \cdot 5 + 4 = 194,$$

$$1234_5 = ((1 \cdot 5 + 2) \cdot 5 + 3) \cdot 5 + 4 = (7 \cdot 5 + 3) \cdot 5 + 4 = 38 \cdot 5 + 4 = 194.$$

Схема Горнера очень удобна для обработки данных при вводе чисел с клавиатуры, когда цифры числа вводятся последовательно, начиная с первой, и их количество заранее неизвестно.

Развёрнутую запись числа можно использовать для обратного перехода, от десятичной системы к системе с основанием p . Действительно из формулы

$$a_3a_2a_1a_0 = a_3 \cdot p^3 + a_2 \cdot p^2 + a_1 \cdot p + a_0$$

следует, что a_0 — это остаток от деления исходного числа на основание p . Если мы разделим исходное число на p и отбросим остаток, мы получим:

$$a_3a_2a_1 = a_3 \cdot p^2 + a_2 \cdot p + a_1.$$

Теперь легко найти a_1 — это последняя цифра получившегося числа, которая, как мы знаем, равна остатку от его деления на p .

¹ При $p > 10$ используются также и латинские буквы, но об этом далее.

Разделив новое получившееся число на p и отбросив остаток, получим число

$$a_3 a_2 = a_3 \cdot p + a_2,$$

из которого найдём a_2 как остаток от деления на p . Разделив на p ещё раз, получаем последнюю цифру a_3 .

Переведём, например, число 194 в пятеричную систему счисления ($p = 5$). Найдём остаток от деления на 5:

$$194 = 38 \cdot 5 + 4.$$

Таким образом, мы нашли последнюю цифру — 4. Частное равно 38, повторяем ту же операцию:

$$38 = 7 \cdot 5 + 3.$$

Следующая (с конца) цифра числа — 3. Дальше получаем:

$$7 = 1 \cdot 5 + 2,$$

третья с конца цифра — 2, а четвёртая — 1 (единица уже не делится на 5). Обратим внимание, что с помощью этого способа мы находим цифры числа, начиная с последней. Поэтому полученные остатки нужно выписать в обратном порядке:

$$\begin{array}{r|l} 194 & 5 \\ \hline 190 & 38 \quad 5 \\ \hline \textcircled{4} & 35 \quad 7 \quad 5 \\ & \textcircled{3} \quad 5 \quad 1 \quad 5 \\ & \quad \textcircled{2} \quad 1 \quad 5 \\ & \quad \quad \textcircled{1} \quad 0 \quad 0 \end{array}$$

Ответ: 1234_5 .



Для перевода числа из десятичной системы в систему счисления с основанием p нужно делить число на p , отбрасывая остаток на каждом шаге, пока не получится 0. Затем надо выписать найденные остатки в обратном порядке.

Можно было заметить, что такой алгоритм фактически использует схему Горнера, «раскручивая» её в обратном порядке. При каждом делении частное и остаток определяются однозначно, поэтому представление числа в любой позиционной системе единственно.

Рассмотренные приёмы позволяют записать любое неотрицательное число в заданной позиционной системе счисления. Признаком отрицательного числа служит знак «-», после которого по тем же правилам записывается модуль числа.

Пример 1. Зная десятичное число и его запись в некоторой позиционной системе счисления, можно найти основание этой системы. Пусть, например, число 71 в некоторой системе с основанием x записывается как 56_x . Представим это число в развёрнутой форме:

$$71 = 56_x = 5 \cdot x^1 + 6 \cdot x^0 = 5 \cdot x + 6.$$

Решая уравнение $71 = 5 \cdot x + 6$ относительно неизвестного x , получаем: $x = 13$. Значит, искомое основание системы — 13.

Пример 2. В более сложных случаях может получиться алгебраическое уравнение второй (или ещё более высокой) степени. Например, пусть то же число 71 в некоторой системе с основанием x записывается как 155_x . Представим это число в развёрнутой форме:

$$71 = 155_x = 1 \cdot x^2 + 5 \cdot x^1 + 5 \cdot x^0 = x^2 + 5 \cdot x + 5.$$

Решая уравнение $71 = x^2 + 5x + 5$ относительно неизвестного x , получаем два решения: $x_1 = -11$ и $x_2 = 6$. Искомое основание положительно, поэтому правильный ответ — 6.

Пример 3. Если запись числа в системе счисления задана не полностью, решений может быть несколько. Например, найдём все основания систем счисления, в которых запись десятичного числа 24 оканчивается на 3. Здесь удобно использовать схему Горнера, из которой сразу следует

$$24 = k \cdot x + 3,$$

где x — неизвестное основание системы счисления, k — некоторое натуральное число или 0. Отсюда сразу получаем $21 = k \cdot x$, т. е. все интересующие нас основания являются делителями числа 21. Это могут быть 3, 7 и 21. Поскольку последняя цифра числа — 3, основание не может быть равно 3 (в троичной системе нет цифры 3), поэтому условию задачи удовлетворяют только основания 7 и 21.

Пример 4. Найдём все десятичные числа, не превосходящие 40, запись которых в системе счисления с основанием 4 оканчивается на 11. Используя схему Горнера, находим, что все интересные нас числа имеют вид

$$N = k \cdot 4^2 + 1 \cdot 4 + 1 = k \cdot 16 + 5,$$

где k — некоторое натуральное число или 0. Подставляя $k = 0, 1, 2, 3, \dots$, находим соответствующие числа $N = 5, 21, 37, 53, \dots$. Из них только 5, 21 и 37 удовлетворяют условию (не больше 40).

Пример 5. Все 5-буквенные слова, составленные из букв А, О, У, записаны в алфавитном порядке. Вот начало списка:

1. ААААА
2. ААААО
3. ААААУ
4. АААОА
- ...

Найдём слово, которое стоит на 140-м месте от начала списка.

Как ни странно, эта задача прямо связана с позиционными системами счисления. В словах используется набор из трёх разных символов, для которых задан порядок (алфавитный). Заменяя буквы А, О и У соответственно на цифры 0, 1 и 2, выпишем начало списка:

1. 00000
2. 00001
3. 00002
4. 00010

Это числа, записанные в троичной системе счисления в порядке возрастания. Тогда легко понять, что 140-м месте от начала списка стоит десятичное число 139, записанное в троичной системе счисления:

$$139 = 12011_3.$$

Заменяя обратно цифры на буквы, получаем ответ:
 $12011 \rightarrow ОУАОО$.

Вопросы и задания

1. Какие системы счисления называют позиционными?
2. Каким термином называется количество цифр в алфавите позиционной системы счисления?
3. Что такое разряд? Как нумеруются разряды?
4. Как связан в позиционной системе «вес» цифры и разряд, в котором она стоит?
5. Как вы думаете, почему в быту мы чаще всего используем десятичную систему?
6. Какие позиционные системы счисления используются сейчас на практике?
7. Чем хороша схема Горнера с точки зрения вычислений?
8. Как перевести число из любой позиционной системы в десятичную?
9. Какие цифры входят в алфавит девятеричной системы?
10. Как вы думаете, можно ли использовать систему счисления с основанием 1 000 000? В чём могут быть проблемы?
11. Сформулируйте алгоритм перевода числа из семеричной системы в десятичную.
12. Сформулируйте алгоритм перевода числа из десятичной системы в семеричную.
13. Как по записи числа в пятеричной системе сразу увидеть, делится ли оно на 5? на 25? на 125?

Задачи

1. Запишите число 12 345 в развёрнутой форме и в виде схемы Горнера.
2. Какое минимальное основание должно быть у системы счисления, чтобы в ней существовали числа 123, 463, 153 и 455? Ответ обоснуйте.
3. Выберите наибольшее число: $11_2, 11_7, 11, 11_{12}, 11_{16}, 11_{25}$.
4. Переведите числа $345_6, 345_7, 345_8$ и 345_9 в десятичную систему.
5. Переведите число 194 в троичную, шестеричную, семеричную и восьмеричную системы счисления.
6. Какие из чисел $1230_7, 124_7, 600_7, 530_7$ делятся на 7? На 49?
7. Десятичное число, переведенное в восьмеричную и в девятеричную систему, в обоих случаях заканчивается на цифру 0. Какое минимальное десятичное число удовлетворяет этому условию?
8. Сколько всего раз встречается цифра 2 в записи чисел 10, 11, 12, ..., 17 в системе счисления с основанием 5?
9. Сколько всего раз встречается цифра 3 в записи чисел 19, 20, 21, ..., 33 в системе счисления с основанием 6?

10. В системе счисления с некоторым основанием N число 12 записывается в виде 110_N . Найдите это основание.
11. Найдите все основания систем счисления, в которых запись числа 29 оканчивается на 5.
12. В системе счисления с некоторым основанием N десятичное число 129 записывается как 1004_N . Найдите это основание.
13. Запись числа 30 в системе счисления с основанием N выглядит так: 110_N . Укажите основание N этой системы счисления.
14. Запись числа 23 в системе счисления с основанием N выглядит так: 212_N . Укажите основание N этой системы счисления.
15. Запись числа 210_5 в системе счисления с основанием N выглядит так: 313_N . Укажите основание N этой системы счисления.
16. Запись числа 65_8 в системе счисления с основанием N выглядит так: 311_N . Укажите основание N этой системы счисления.
17. Найдите все десятичные числа, не превосходящие 25, запись которых в двоичной системе счисления оканчивается на 101.
18. Запись числа 67 в системе счисления с основанием N оканчивается на 1 и содержит 4 цифры. Укажите основание N этой системы счисления.
19. Найдите все основания систем счисления, в которых запись числа 30 оканчивается на 8.
20. Найдите все основания систем счисления, в которых запись числа 31 оканчивается на 11.
21. Найдите все основания систем счисления, в которых запись числа 63 оканчивается на 23.
22. Найдите наименьшее основание системы счисления, в которой запись числа 30 трёхзначна.
23. Найдите наименьшее основание системы счисления, в которой запись числа 70 трёхзначна.
24. Найдите все десятичные числа, не превосходящие 26, запись которых в троичной системе счисления оканчивается на 22.
25. Найдите все десятичные числа, не превосходящие 30, запись которых в четверичной системе счисления оканчивается на 31.
- *26. Найдите все десятичные числа, не превосходящие 25, запись которых в системе счисления с основанием 6 начинается на 4.
- *27. Найдите все десятичные числа, не превосходящие 30, запись которых в системе счисления с основанием 5 начинается на 3.
- *28. Найдите основание системы счисления x , для которого выполняется равенство

а) $32_x + 64_x = 106_x$;	б) $45_x + 55_x = 122_x$;
в) $42_x + 41_x = 133_x$;	г) $91_x + 93_x = 154_x$.

29. Все 5-буквенные слова, составленные из букв А, О, У, записаны в алфавитном порядке. Вот начало списка:

1. ААААА
2. ААААО
3. ААААУ
4. АААОА

Выполните следующие задания:

- а) определите, сколько всего слов в списке;
- б) укажите слова, которые стоят на 101-м, 125-м, 170-м и 210-м местах;
- в) укажите порядковые номера слов ОАОАО, УАУАУ, АОУОА, УОАОУ;
- г) укажите номера первого и последнего слов, которые начинаются с буквы О.

30. Все 5-буквенные слова, составленные из букв А, К, Р, У, записаны в алфавитном порядке. Вот начало списка:

1. ААААА
2. ААААК
3. ААААР
4. ААААУ
5. АААКА

Выполните следующие задания:

- а) определите, сколько всего слов в списке;
- б) укажите слова, которые стоят на 150-м, 250-м, 350-м и 450-м местах;
- в) укажите порядковые номера слов АКУРА, КАРАУ, РУКАА, УКАРА, УРАКА;
- г) укажите номера первого и последнего слов, которые начинаются с буквы Р.

Дробные числа

Дробные числа сначала рассмотрим на примере десятичной системы. Число 0,6375 можно представить в виде:

$$0,6375 = 6 \cdot 0,1 + 3 \cdot 0,01 + 7 \cdot 0,001 + 5 \cdot 0,0001.$$

Все множители, на которые умножаются значения цифр, представляют собой отрицательные степени числа 10 — основа-