

Е. В. МИХЕЕВА, О. И. ТИТОВА

# ИНФОРМАТИКА

**УЧЕБНИК**

*Рекомендовано  
Федеральным государственным учреждением  
«Федеральный институт развития образования»  
в качестве учебника для использования  
в учебном процессе образовательных учреждений,  
реализующих программы среднего профессионального образования*

*Регистрационный номер рецензии 340  
от 4 октября 2010 г. ФГУ «ФИРО»*

10-е издание, стереотипное



Москва  
Издательский центр «Академия»  
2014

УДК 681.518(075.32)  
ББК 32.81я723  
М695

Рецензенты:

зав. кафедрой «Информатизация банковской деятельности» Московского  
банковского института, канд. техн. наук *А. Н. Герасимов*;  
преподаватель информационных технологий и информационных систем  
Московского колледжа железнодорожного транспорта *Н. А. Федосеев*

**Михеева Е. В.**

М695 Информатика : учебник для студ. учреждений сред. проф.  
образования / Е. В. Михеева, О. И. Титова. — 10-е изд., стер. —  
М. : Издательский центр «Академия», 2014. — 352 с.  
ISBN 978-5-4468-0837-3

Учебник может быть использован при изучении дисциплины математического и общего естественно-научного цикла «Информатика» в соответствии с ФГОС СПО для всех специальностей.

Изложены основы базовых знаний по информатике: автоматизированная обработка информации, состав персонального компьютера и его программное обеспечение, работа с информацией и ее защита, сетевые технологии обработки и прикладные программные средства, а также автоматизированные информационные системы.

Для студентов учреждений среднего профессионального образования.

УДК 681.518(075.32)  
ББК 32.81я723

*Оригинал-макет данного издания является собственностью  
Издательского центра «Академия», и его воспроизведение любым способом  
без согласия правообладателя запрещается*

© Михеева Е. В., Титова О. И., 2007

© Образовательно-издательский центр «Академия», 2007

ISBN 978-5-4468-0837-3

© Оформление. Издательский центр «Академия», 2007

## ПРЕДИСЛОВИЕ

Характерной чертой нашего времени является рост производства, потребления и накопления информации во всех отраслях человеческой деятельности. Читает ли человек, разговаривает ли по телефону, смотрит ли телевизор — вся его жизнь неразрывно связана с получением, накоплением и обработкой информации.

Информатика как наука стала развиваться в 40-х гг. XX в., и это связано со стремительным развитием вычислительной техники (ВТ), когда технической базой вычислительной техники стала электроника, а затем микроэлектроника.

В условиях информационного общества одним из основных элементов рабочего места любого специалиста является персональный компьютер как инструмент сбора, обработки, хранения и передачи информации, поэтому очень важно для успешной деятельности обладать твердыми знаниями и четкими понятиями по информатике.

В данном учебнике рассмотрены различные аспекты применения достижений информатики и современных информационных технологий.

В главах 1—2 рассмотрены основные понятия информатики и технологии обработки информации, виды и свойства информации, единицы измерения информации и системы ее кодирования, информационные процессы и информационные технологии.

В главах 3—6 представлены материалы по аппаратному (Hardware) и программному (Software) обеспечению компьютера, рассмотрены операционная система Windows, файловая система организации данных, файловые менеджеры и архиваторы.

В главах 7 и 8 рассмотрены обработка и хранение информации, защита ее от несанкционированного доступа и антивирусные средства защиты.

В главе 9 представлены материалы по локальным и глобальным сетям, рассмотрены основные сервисы Интернета и современные технологии создания Web-сайтов.

В главах 10—13 рассмотрены информационные технологии работы с текстовой, табличной, графической информацией, а также с данными реляционных баз.

В главе 14 представлены классификация автоматизированных информационных систем (АИС) и принципы организации типовых профессиональных автоматизированных систем.

Авторами было приложено максимум усилий для создания учебника, дающего реальную пользу и устойчивые навыки работы с информацией и информационными технологиями.

Кроме того, для привития навыков практической работы с компьютером и программами можно использовать «Практикум по информатике» автора Е.В.Михеевой, изданный в Издательском центре «Академия».

## ВВЕДЕНИЕ

В середине XX в. потребность в обработке, передаче больших объемов информации и управлении сложными системами (в первую очередь, военно-стратегического назначения) послужили стимулом для бурного развития кибернетики и теории информации.

Основу активного развития вычислительной техники заложили исследования Норберта Винера и Клода Шеннона. Можно сказать, что они стояли у истоков научно-технической революции.

Технологии работы с информацией, обеспечивающие сбор, обработку, хранение, распределение и отображение информации, называются *информационными технологиями* (ИТ). В основе современных информационных технологий лежит использование новейших аппаратных средств (компьютеров и связанных с ними устройств) и программного обеспечения.

Первые компьютеры были созданы почти полвека назад. В то время они занимали целые залы, но их быстродействие сравнимо с быстродействием современных персональных компьютеров.

Миниатюризация и стремление сделать компьютеры достоянием масс привели к тому, что в XXI в. мы вошли с большим обилием красочных компьютерных журналов, написанных простым и понятным миллионам пользователей персональных компьютеров языком.

Компьютеры заняли заслуженное первое место в процессе обработки информации, стали популярными и превратились из «религии одиночек» в орудие производства миллионов. Они с успехом решают поставленные перед ними задачи и удовлетворяют требованиям как ученых, так и топ-менеджеров крупных корпораций.

Историю развития информационных технологий можно разделить на несколько этапов.

Первый этап — это ручная технология сбора и обработки информации, господствовавшая до второй половины XIX в. Основными инструментами в то время были чернила, перо и простейшие приборы счета, а средства связи были представлены курьерской и почтовой связью.

Второй этап — электромеханический. В 1831 г. Джозеф Генри (США) и Сальваторе Дель Negro (Италия) создали электромагнитное реле, а в 1887 г. компостирование проездных билетов на железнодорожном транспорте натолкнула американского изобретателя и промышленника Германа Холлерита на изобретение элект-

тромеханического табулятора с вводом чисел с помощью перфокарт.

К 1930 г. общее число счетно-аналитических комплексов, установленных в США и других странах, достигло 6... 8 тыс. шт., что, естественно, потребовало развития индустрии для изготовления подобных устройств. В 1931 г. фирма IBM (США) начала выпуск табуляторов, приспособленных для выполнения операций умножения, а в 1934 г. — алфавитно-цифровых табуляторов.

В середине 1930-х гг. на основе табуляторов был создан прообраз первой локальной вычислительной сети. В Питсбурге (США) в универмаге была установлена система, состоящая из 250 терминалов, соединенных телефонными линиями с 20 табуляторами и 15 пишущими машинками для расчетов с покупателями. В 1934—1936 гг. немецкий инженер Конрад Цузе пришел к идее создания универсальной вычислительной машины с программным управлением и хранением информации в запоминающем устройстве. Он сконструировал машину «Z-3» — первую программно-управляемую вычислительную машину.

С появлением электронных пишущих машинок, диктофонов и копируемых машин в 40—60 гг. XX в. начался третий этап — с применением электронной технологии (с момента изобретения Т.Эдисоном первой электронной лампы — диода). Впоследствии Ли де Форест поместил третий электрод и появилась трехэлектродная лампа — триод. На основе триодов уже можно было создавать электронные быстродействующие реле и триггеры — основные компоненты ЭВМ.

Существенными недостатками ламповых ЭВМ были низкая экономичность (электронные лампы потребляли много энергии, выделяли много тепла и занимали большой объем) и, самое главное, ненадежность. Выход из строя всего одной из нескольких тысяч ламп мог полностью остановить работу ЭВМ.

Четвертый этап начался в 1947 г., когда У.Шоркли, Дж.Бардин и У.Бреттейн изобрели принципиально новое электронное устройство — транзистор. Это изобретение было лишено большинства недостатков электронных ламп и позволило сконструировать первую мини-ЭВМ. Новые типовые узлы и модули на порядок уменьшили размеры компьютеров.

Следующий решительный шаг был сделан в 1958 г., когда была создана интегральная микросхема. С ее созданием начался пятый этап микроэлектроники. Микросхема включала в себя все необходимые компоненты: транзисторы, резисторы, конденсаторы и соединяющие их проводники в одном кремниевом кристалле. Дальнейшее развитие информационных технологий было уже чисто технологическим — постоянная миниатюризация компонентов модуля, повышение надежности, увеличение количества узлов на единице площади или объема и т.д.

Для XX в. была характерна небывалая скорость развития науки, техники и новых технологий. Так, с момента изобретения книгопечатания в середине XV в. до создания радиоприемника в конце XIX в. прошло примерно 500 лет, между изобретением радио и телевидения прошло меньше 50 лет, а между изобретением транзистора и интегральной схемы — всего 5 лет.

Современные компьютеры сильно отличаются от первых вычислительных машин, хотя принципы обработки информации практически не изменились.

Информационные технологии дают людям возможность лучше проявить свои способности и сделать еще один шаг от чисто механического труда к творческому. XXI в. по праву считается веком информатики, компьютерных технологий и электронных телекоммуникаций, часто его называют цифровым веком.

Начиная с XVII в. объем научной информации увеличивался примерно в 10 раз каждые 100 лет. Сегодня лавинообразный поток информации является одной из важнейших проблем человечества.

Подсчитано, например, что специалист, наш современник, должен тратить больше половины своего рабочего времени для того, чтобы уследить за всеми новыми печатными трудами в его области деятельности.

Увеличение объема информации и растущий спрос на нее обусловили появление информатики — отрасли, связанной с автоматизацией обработки информации.

Информация является одной из основных потребностей современного человека. Она нужна для работы, принятия решений, приобретения товаров, путешествий, выполнения школьных заданий, заботы о здоровье, а также для других многочисленных видов деятельности.

Сегодня работа с большими объемами информации без применения современных средств автоматизации уже просто невыполнима.

В современном обществе информация рассматривается как ресурс, который можно добывать, перерабатывать, использовать и распространять, как и традиционные ресурсы (труд, энергия, полезные ископаемые). Еще в 1994 г. на проходившем в Москве Третьем международном форуме по информатизации прозвучали слова о том, что раньше для производства нужны были три вещи: земля, орудия труда, капитал, а сегодня к этому перечню добавилась информация.

В наше время информация является общенаучным понятием, включающим в себя обмен сведениями между людьми и автоматами, обмен сигналами в растительном и животном мире, передачу признаков от организма к организму, от клетки к клетке. В технической сфере информация часто определяется на основе понятия обмена сообщениями как отражения внешнего мира с помощью знаков и сигналов.

Информатика — область человеческой деятельности, связанная с процессами преобразования информации с помощью компьютеров и взаимодействия со средой их применения.

Информатика появилась с появлением персональных компьютеров.



## 1.1. Понятия информатики и информации

**Понятия и предмет информатики.** Термин «информатика» (от фр. *information* — информация + *automatique* — автоматика) означает автоматическую обработку информации. Кроме Франции термин «информатика» используется в ряде стран Восточной Европы. В большинстве стран Западной Европы и США используется другой термин — *Computer Science* (наука о средствах вычислительной техники).

*Информатика* — это наука, изучающая структуру и общие свойства информации, а также закономерности и методы ее создания, хранения, поиска, передачи и преобразования с использованием компьютерных технологий.

Выделение информатики в качестве самостоятельной сферы человеческой деятельности связано с развитием компьютерной (микропроцессорной) техники.

Информатика базируется на компьютерной технике и немыслима без нее.

В качестве источников информатики обычно называют две науки: документалистику и кибернетику. Документалистика сформировалась в конце XIX в. в связи с бурным развитием производственных отношений. Ее целью являлось повышение эффективности документооборота.

Основы близкой к информатике технической науки кибернетики (от гр. *kyberneticos* — искусный в управлении) были заложены трудами по математической логике американского математика Норберта Винера, опубликованными в 1948 г.

Впервые термин «кибернетика» ввел французский физик Ампер в первой половине XIX в. Он занимался разработкой единой системы классификации всех наук и обозначил этим термином гипотетическую науку об управлении, которой в то время не существовало, но которая, по его мнению, должна была существовать.

Сегодня кибернетика изучает принципы построения и функционирования систем автоматического управления и методы моделирования процесса принятия решений техническими средствами. На практике кибернетика во многих случаях опирается на те же программные и аппаратные средства вычислительной техники, что и информатика, а информатика, в свою очередь, заимствует у кибернетики математическую и логическую базы для развития этих средств.

В информатике все ориентировано на эффективность. Вопрос, как сделать ту или иную операцию, для информатики является важным, но не основным. Основным же является вопрос, как сделать данную операцию эффективно.

В информатике используются следующие понятия:

- аппаратное обеспечение средств вычислительной техники;
- программное обеспечение средств вычислительной техники;
- средства взаимодействия аппаратного и программного обеспечения;
- средства взаимодействия человека с аппаратными и программными средствами.

Особое внимание в информатике уделяется вопросам взаимодействия. Существует специальное понятие — «интерфейс». *Пользовательским интерфейсом* называются методы и средства взаимодействия человека с аппаратными и программными средствами. Соответственно существуют аппаратные, программные и аппаратно-программные интерфейсы.

Основной задачей информатики является систематизация приемов и методов работы с аппаратными и программными средствами вычислительной техники. Цель систематизации состоит в выделении, внедрении и развитии передовых, наиболее эффективных, технологий, в автоматизации этапов работы с данными, а также в методическом обеспечении новых технологических исследований.

В составе основной задачи информатики можно выделить следующие направления для практических приложений: архитектура вычислительных систем, интерфейсы вычислительных систем, программирование, преобразование данных, защита информации, автоматизация и стандартизация.

**Понятие информации.** Ключевым понятием информатики является понятие информации, с которым мы сталкиваемся ежедневно, однако единого ее определения до сих пор не существует. Поэтому вместо определения обычно используют понятие об информации.

Первоначально под информацией (от *лат.* informatio — разъяснение, изложение, сообщение, осведомление) понимались сведения, передаваемые людьми различными способами: устно, с помощью сигналов или технических средств.

Федеральный закон от 27.07.2006 № 149-ФЗ «Об информации, информатизации, информационных технологиях и о защите информации» определяет информацию следующим образом.

*Информация* — сведения (сообщения, данные) независимо от формы их представления.

*Информатизация* — организационный социально-экономический и научно-технический процесс создания оптимальных условий для удовлетворения информационных потребностей и реализации прав граждан, органов государственной власти, органов местного самоуправления, организаций, общественных объединений на основе формирования и использования информационных ресурсов.

Основные понятия, определения и термины информатики также определяются ГОСТ 15971—90 «Системы обработки информации.

Термины и определения», согласно которому, *информация* — это сведения о фактах, концепциях, объектах, событиях и идеях, которые в данном контексте имеют вполне определенное значение. Отметим, что информация — это не просто сведения, а сведения нужные, имеющие значение для лица, обладающего ими. В этих определениях информация отражены основные важные свойства понятия информации.

Во-первых, информация не является материальным объектом, ее передают от одного человека к другому, при этом первый ее не утрачивает. В результате передачи оба эти человека будут владеть переданной информацией. Информация — единственный ресурс, который при передаче не уменьшается, а только увеличивается.

Во-вторых, для передачи информация должна быть представлена на каком-нибудь материальном носителе.

В-третьих, содержание информации должно быть неизменным при ее переносе с одного носителя информации на другой.

**Информационное сообщение.** С практической точки зрения информация всегда представляется в виде сообщения. Информационное сообщение связано с источником сообщения, получателем сообщений и каналом связи (рис. 1.1).

Сообщение от источника к приемнику передается в материально-энергетической форме (электрический, световой, в виде звуковых сигналов и т. д.). В зависимости от вида сигнала, определяемого свойствами передающего устройства, различают непрерывную (аналоговую) и дискретную (цифровую) информацию.

Источником аналоговой информации обычно являются различные природные объекты (например, температура, давление и влажность воздуха), объекты технологических производственных процессов (например, нейтронный поток в активной зоне, давление и температура теплоносителя в контурах ядерного реактора) и др.

Информационные сообщения, используемые человеком, имеют характер дискретных сообщений, например сигналы тревоги, передаваемые посредством световых сообщений, телеграфные сигналы, языковые сообщения, передаваемые в письменном виде или с помощью звуковых сигналов и др.

Человек воспринимает сообщения при помощи органов чувств, и, как правило, в основном это непрерывная информация, а вот логическое мышление человека имеет, скорее, дискретный характер.

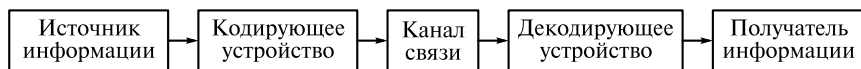


Рис. 1.1. Общая схема передачи информации

Приемники аналоговой информации обычно воспринимают сообщения с помощью различной измерительной и регистрирующей аппаратуры. Приемники цифровой информации, например компьютеры, воспринимают сообщения в виде чисел и обрабатывают информацию с помощью электрических сигналов.

**Классы вычислительных машин.** В зависимости от вида перерабатываемой информации вычислительные машины подразделяются на два основных класса: аналоговые и цифровые.

*Аналоговая вычислительная машина (АВМ)* — это машина, оперирующая информацией, представленной в виде непрерывных изменений некоторых физических величин. При этом в качестве физических переменных используются сила тока электрической цепи, угол поворота вала, скорость и ускорение движения тела и др. Используя тот факт, что многие явления в природе математически описываются одними и теми же уравнениями, АВМ позволяют с помощью одного физического процесса моделировать различные процессы. Это и определило сферу применения АВМ.

*Цифровая вычислительная машина (ЦВМ)* — это машина, оперирующая информацией, представленной в дискретном виде. В настоящее время разработаны методы численного решения многих видов уравнений, что дало возможность решать на ЦВМ различные уравнения и задачи с помощью набора простых арифметических и логических операций. Компьютеры — это ЦВМ.

## 1.2. Свойства и носители информации

**Свойства информации.** При работе с информацией и разработке информационных систем и технологий важно оценить свойства поступающей, хранимой и передаваемой информации.

Сформулируем следующие определения свойств информации.

*Адекватность* — свойство информации однозначно соответствовать отображаемому объекту или явлению.

*Достоверность* — свойство информации не иметь скрытых ошибок.

*Полнота* — свойство информации исчерпывающе характеризовать отображаемый объект или процесс.

*Доступность* — свойство информации, характеризующее возможность ее получения данным пользователем.

*Релевантность* — способность информации соответствовать запросам пользователя.

*Качество информации* — обобщенная положительная характеристика информации, отражающая степень ее полезности.

*Актуальность информации* — степень соответствия информации текущему моменту времени. Нередко с актуальностью, как и с полнотой, связывают коммерческую ценность информации.

Существуют еще и другие, менее существенные свойства информации.

*Краткость и четкость информации* — отсутствие в информации ненужных сведений.

*Ценность* — степень важности информации для решения задачи.

*Понятность* — выражение информации на языке, понятном тем, кому она предназначена.

*Своевременность* — актуальность информации и наличие в ней сведений, необходимых в данный момент для понимания и принятия решения.

В каждом конкретном случае нужно учитывать только те свойства информации, которые действительно оказывают влияние на рассматриваемую систему или технологию.

Информационные процессы растянуты во времени, поэтому достоверная и адекватная, но устаревшая информация может привести к ошибочным решениям. Необходимость поиска (или разработки) адекватного метода для работы с данными может привести к такой задержке получения информации, что она становится неактуальной и ненужной. На этом, в частности, основаны многие современные системы шифрования данных с открытым ключом. Лица, не владеющие ключом (методом) для чтения данных, могут заняться поиском ключа, поскольку алгоритм его работы доступен, но продолжительность этого поиска столь велика, что за время работы информация теряет актуальность и, естественно, связанную с ней практическую ценность. Одним из способов превращения информации в сообщение является запись ее на материальном носителе. Если мы применяем материальные носители, предназначенные для использования в компьютерной технике, то мы имеем дело с данными.

*Данные* — это информация, представленная в виде, пригодном для обработки ее автоматическими средствами при возможном участии человека.

**Операции с данными.** В ходе информационного процесса данные преобразуются из одного вида в другой. Обработка данных включает в себя следующие операции.

*Сбор данных* — накопление информации с целью обеспечения достаточной полноты для принятия решений.

*Формализация данных* — приведение данных, поступающих из разных источников, к одинаковой форме для того, чтобы сделать их сопоставимыми между собой.

*Сортировка данных* — упорядочивание данных по заданному признаку с целью удобства их использования; сортировка данных повышает доступность информации.

*Фильтрация данных* — отсеивание лишних данных, в которых нет необходимости для принятия решений; при этом достоверность и адекватность данных должны возрастать.

*Архивация данных* — организация хранения данных в компактной сжатой форме; архивация данных повышает общую надежность информационного процесса и используется для снижения затрат по хранению данных.

*Защита данных* — комплекс мер, направленных на предотвращение утраты, воспроизведения и изменения данных.

*Преобразование данных* — перевод данных из одной формы в другую; преобразование данных часто связано с изменением типа носителя.

*Прием и передача данных* — процессы, осуществляемые между участниками информационного процесса; при этом источник данных в информатике принято называть сервером, а потребителя — клиентом.

**Носители данных.** Самым распространенным носителем данных, хотя и не самым экономичным, является бумага. На бумаге данные регистрируются путем изменения оптических характеристик ее поверхности. Для хранения компьютерных данных используются специальные носители, на которых сохраняется кодированная информация.

В компьютере для хранения информации предназначена память, которая подразделяется на основную (энергозависимую), участвующую только в процессе обработки информации, и внешнюю (энергонезависимую).

Внешняя память в зависимости от характера носителя подразделяется на несколько типов:

- память на магнитных носителях — гибкие и жесткие магнитные диски (винчестеры), ziv-диски и магнитные ленты;
- память на оптических носителях — компакт-диски с однократной и многократной записью;
- энергонезависимая электронная память — флеш-память.

Внешняя память выполняет функции хранения исходных, промежуточных и окончательных данных в процессе обработки информации, архивного хранения данных и переноса информации с одного компьютера на другой. В каждом из этих случаев применяются свои виды внешней памяти.

Фактически единственным типом внешней памяти, используемым в процессе обработки информации, являются накопители на жестких магнитных дисках. Их основной характеристикой является объем вмещаемой информации.

### ***Это интересно***

Первый жесткий диск был выпущен компанией IBM в 1956 г. Он назывался RAMAC, имел объем памяти 5 Мбайт и был сконструирован из пятидесяти 24-дюймовых пластин. Ни о каком персональном применении тогда не было и речи.

Для архивного хранения данных долго использовались магнитные ленты и гибкие магнитные диски (дискеты). В последние годы дискеты постепенно утрачивают функцию архивного хранения, так как имеют небольшую емкость (1,44 Мбайт), но ими еще пользуются из-за низкой стоимости носителя. Оптические диски, стоимость которых сейчас резко снизилась, все больше используются для архивного хранения.

Средства переноса данных отличаются большим разнообразием. Дискеты по-прежнему сохраняют свое значение в качестве наиболее удобного средства переноса небольших объемов информации. Это связано с тем, что их можно использовать на любом компьютере, в отличие от других носителей, требующих наличия на компьютере дополнительных аппаратных или программных возможностей.

Удобным и дешевым средством переноса информации являются записываемые и перезаписываемые оптические диски, но они требуют наличия пишущего оптического дисковода и соответствующего программного обеспечения. Еще более удобным средством переноса информации являются устройства флэш-памяти с USB-интерфейсом (USB-брелки) и карты флэш-памяти с USB-переходником, однако их стоимость пока еще достаточно высока. Для переноса больших объемов информации, измеряемых десятками гигабайтов, используются внешние накопители на жестких магнитных дисках (портативные винчестеры и ziv-диски), оборудованные средствами для подключения к компьютеру через внешние разъемы.

От свойств носителя нередко зависят такие свойства информации, как полнота, доступность и достоверность. Задача преобразования данных с целью смены носителя является одной из основных задач информатики.

Перерабатываемая информация связана с различными материальными носителями, однако главную роль в информационных технологиях играет не сам носитель, а связанная с ним информация.

### **1.3. Виды информации и ее кодирование**

**Виды информации.** Обычно для классификации объектов одной природы используется то или иное свойство либо набор свойств объектов. Нас интересует классификация информации в плане автоматизации основных информационных процессов.

Первоначально вычислительные машины применялись только для обработки числовой информации, однако довольно быстро выяснилось, что их возможности не ограничиваются только работой с числами.

Далеко не вся информация окружающего нас мира может быть обработана компьютером, ведь пока не придумали такого компьютера, который мог бы чувствовать или наслаждаться произведениями искусства. Поэтому, говоря об информации, необходимо выделить те ее виды, которые компьютер воспримет и позволит человеку использовать свои ресурсы для обработки, хранения и передачи такой информации.

Компьютер может работать с текстовой, числовой, табличной, графической информацией, а также со звуковой, анимационной и видеoinформацией. Также компьютер воспринимает специальную двоичную информацию.

В настоящее время практически все компьютерные технологии ограничиваются обработкой перечисленных видов информации. С развитием компьютерной техники увеличиваются объемы перерабатываемой информации. И хотя современные компьютеры могут делать очень много, все же их возможности не безграничны. Наибольший эффект от применения компьютера будет там, где оправдано его применение.

**Понятие кодирования.** *Кодирование информации* — это преобразование одной последовательности сигналов в другую. Под *кодированием данных* понимается выражение данных одного типа через данные другого типа. Для автоматизации работы с данными, относящимися к различным типам, очень важно унифицировать их форму представления, поэтому обычно используется прием кодирования.

*Человеческий язык* — это система кодирования понятий для выражения мыслей посредством речи. *Азбука* — это система кодирования компонентов языка с помощью графических символов.

Своя система существует и в вычислительной технике. Она называется двоичным кодированием и основана на представлении данных последовательностью всего двух цифр: 0 и 1.

Для представления дискретной информации в компьютере применяется алфавитный способ, основанный на использовании фиксированного конечного набора символов (алфавита). Примерами алфавитов могут служить алфавиты естественных человеческих языков, совокупность десятичных цифр, любая другая упорядоченность знаков, предназначенная для образования и передачи сообщений. Символы из набора алфавита называются *буквами*, а любая конечная последовательность букв — *словом* в этом алфавите. При этом не требуется, чтобы слово обязательно имело языковое смысловое значение.

Процесс преобразования информации часто требует представлять буквы одного алфавита средствами (буквами, словами) другого алфавита. Такое представление и называется кодированием. Процесс обратного преобразования информации относительно ранее выполненного кодирования называется *декодированием*.



**Предыстория кодирования информации.** Люди общаются в основном с помощью сказанных или написанных слов. Эта система нормально работает, когда все участники находятся поблизости друг от друга (в пределах слышимости или видимости). А если мы хотим связаться с удаленным собеседником? С древних времен до XIX в. для этой цели использовались курьеры с устными или письменными сообщениями. Такая связь работала неплохо, хотя часто слишком медленно; к тому же сообщение или курьер до адресата порой не доходили.

Шло время, развивались технологии, и люди изобретали различные коммуникационные приспособления. В доиндустриальную эпоху для передачи сообщений на большие расстояния использовали устройства наподобие маяков. Индейцы Северной Америки применяли дымовые сигналы, в армиях для передачи сообщений использовали флаги и зеркала. Создавались и хитроумные механизмы для передачи сообщений на все увеличивающиеся расстояния.

Техническая революция сопровождалась распространением электричества и телеграфа, позволявшего мгновенно передавать сообщения на очень большие расстояния по одному проводу. Теперь уже не нужно было видеть человека на другом конце провода или посылать к нему посредника-почтальона. Телеграф и дымовые сигналы имеют одно общее свойство — им требуется некоторый код, чтобы перевести человеческий язык в форму, которую мог бы передать механизм или телеграфный аппарат. На принимающем конце этот код необходимо перевести обратно на человеческий язык. Уже в ранних коммуникационных устройствах сформировались две идеи, которые легли в основу современных компьютеров:

- 1) цифровой (digital), т. е. дискретный, код, основанный на двух состояниях (включено — выключено, или 0 и 1);
- 2) специализированный машинный язык (обычно цифровой), используемый машиной для обработки данных.

Телеграф и первые радиостанции применяли для передачи сообщений специальный код — азбуку Морзе, названную по имени ее создателя Сэмюэла Ф. Б. Морзе. В ней с каждой буквой алфавита сопоставлена комбинация точек (коротких импульсов) и тире (длинных импульсов). Импульсы передаются по проводам в определенной последовательности, которую оператор на принимающем устройстве переводит обратно в буквы и слова. Как правило, оператор использует справочник по кодам, но опытные операторы знают код настолько хорошо, что могут расшифровывать каждый символ по памяти.

Современные компьютеры похожи на ранний телеграф, ведь они передают информацию по проводам в цифровой форме, используя специальный код. Но если основная задача телеграфа —

передавать информацию на далекие расстояния, то компьютер передает данные внутри себя. При этом компьютер использует другой кодовый язык и несколько проводов, а не один, как телеграф.

**Кодирование данных двоичным кодом.** На современном языке телеграф можно назвать устройством для цифровой последовательной связи. Связь является цифровой, потому что в ней используется дискретный (включено—выключено) код; последовательной, потому что элементы языка (точки и тире) отправляются последовательно один за другим. Если мы разработаем код, в котором каждая буква алфавита будет представлена комбинацией из восьми элементов (0 или 1), и будем отправлять их один за другим, то мы создадим цифровое последовательное устройство. При наличии единственного провода такой способ связи работает прекрасно, но медленно (ведь нам приходится посылать по очереди восемь единиц информации, чтобы передать одну букву). А если вместо одного у нас было бы восемь проводов? Тогда мы могли бы передать все восемь элементов сразу, или параллельно. Именно так данные передаются в компьютере.

Кодирование может производиться без потери и с потерями информации. Так, преобразование принципиально различных видов информации — непрерывной в дискретную (аналого-цифровое преобразование (АЦП)) и дискретной в непрерывную (цифроаналоговое преобразование (ЦАП)) — возможно только с потерей информации.

К кодированию можно отнести и сжатие (архивацию) информации. *Сжатие* — это устранение избыточности информации, например за счет упрощения кодов путем исключения из них постоянных битов.

Другой разновидностью кодирования является введение избыточной информации, что широко применяется в криптографии. Примерами такого кодирования могут служить электронный сертификат, цифровая подпись и шифрование.

## 1.4. Измерение информации

**Системы счисления.** Для записи чисел люди используют различные системы счисления. Система счисления показывает, по каким правилам записываются числа и как выполняются арифметические действия над ними.

Обычно мы используем десятичную систему записи чисел, при которой число записывается с помощью 10 цифр (0, 1, ..., 9). Для счета времени в часах используется двенадцатеричная система счисления, в минутах и секундах — шестидесятеричная.

И десятичная, и двоичная системы счисления относятся к позиционным, т. е. значение цифры зависит от ее расположения в

записи числа. Место цифры в записи числа называется *разрядом*, а количество цифр в числе — *разрядностью числа*. Разряды нумеруются справа налево; каждому разряду соответствует степень основания системы счисления.

В компьютере для записи чисел используется двоичная система счисления, т.е. любое число записывается в виде сочетания двух цифр: 0 и 1, которые называются двоичными цифрами (binary digit, или сокращенно bit).

**Единицы информации.** *Бит* (bit) — это наименьшая единица информации, распознаваемая компьютером. Это нечто вроде лампы, которая может быть либо включена, либо выключена. Биты используются при передаче информации по однопроводной телеграфной системе.

### **Это интересно**

В 1946 г. математик из Принстонского университета Джон Таки впервые использовал в одной из своих статей термин «bit» (бит).

Одним битом могут быть выражены два понятия: 0 или 1 (да или нет, истина или ложь, черное или белое и т.д.). Если количество битов увеличить до двух, то уже можно выразить четыре различных понятия. Тремя битами можно закодировать восемь различных значений.

*Байтом* (byte) называется группа из восьми битов, необходимая для представления одного символа информации. Нажатие одной клавиши на клавиатуре эквивалентно отправке одного байта информации центральному процессору компьютера. Байт — это стандартная единица измерения памяти в компьютере. Обычно ее объем выражается в килобайтах (Кбайт) или мегабайтах (Мбайт).

Единицы компьютерной памяти и их значения:

- бит — наименьшая единица информации, сокращение для binary digit (двоичной цифры);
- тетрада — 4 бит (половина байта);
- байт — 8 бит (один символ равен 8 бит);
- слово — 16 бит (на больших компьютерах длина слова достигает 64 бит);
- килобайт (Кбайт) — 1 024 байт.

Более крупные единицы обозначаются добавлением префиксов мега (М), гига (Г), тера (Т), но в них пока нет практической надобности:

1 Мбайт = 1 048 576 байт (приблизительно миллион байтов, или 1 024 Кбайт, или  $2^{10}$  Кбайт);

1 Гбайт = 1 073 741 824 байт (приблизительно миллиард байтов, или 1 024 Мбайт, или  $2^{10}$  Мбайт);

1 Тбайт = 1 024 Гбайт (или  $2^{10}$  Гбайт).

**Двоичный код.** Как уже отмечалось ранее, бит существует в двух состояниях: «включено» или «выключено». Для наглядного представления байтов используются цифры: 1 — включено; 0 — выключено.

Далее представлен один байт информации, в котором все восемь битов имеют нулевые значения. В двоичной системе счисления это соответствует нулю.

0 0 0 0 0 0 0 0

Двоичная система счисления похожа на десятичную (decimal), которую мы используем в повседневной жизни. Приставка «dec» означает 10, т. е. эта система счисления основана на числе 10. Двоичная (binary) система основывается на 2 («bi» означает 2, как в слове «бинокль»).

Первая цифра справа в представлении байта — это столбец единиц; значения в этом столбце равны 1 или 0. Второй столбец представляет двойки и принимает значение 1 или 0. Следующие столбцы соответствуют 4, 8, 16 и т. д. Значение каждого столбца равно удвоенному значению столбца справа; 2 — это основа двоичной системы.

Максимальное число, которое можно представить одним байтом, — 256.

Компьютеры — это машины, и для связи им требуется собственный машинный язык. Компьютерный язык называется двоичным (binary); структурные элементы, на которых он основан, могут находиться в двух состояниях: «включено» или «выключено».

Компьютеры должны работать с разнообразной информацией, поэтому им нужен код, преобразующий человеческий язык в машинный. Таким кодом является код ASCII (American Standard Code for Information Interchange — стандартный код информационного обмена США).

## 1.5. Системы кодирования данных

**Кодирование целых и действительных чисел.** Целые числа кодируются двоичным кодом достаточно просто: необходимо взять целое число и делить его пополам до тех пор, пока частное не будет равно единице. Совокупность остатков от каждого деления, записанная справа налево вместе с последним частным, и образует двоичный аналог десятичного числа.

Для кодирования целых чисел от 0 до 255 достаточно иметь восемь разрядов двоичного кода (8 бит). 16 бит позволяют закодировать целые числа от 0 до 65 535, а 24 бит — более 16,5 млн различных значений.

Для кодирования действительных чисел используют 80-разрядное кодирование. При этом число предварительно преобразовывают в нормализованную форму:

$$5,12345678 = 0,512345678 \cdot 10^1;$$

$$500\,000 = 0,5 \cdot 10^6.$$

Первая часть числа называется *мантиссой*, а вторая — *характеристикой*. Большую часть из 80 бит отводят для хранения мантиссы (вместе со знаком) и некоторое фиксированное количество разрядов отводят для хранения характеристики.

**Универсальная система кодирования (код ASCII).** С помощью двоичного кода можно кодировать текстовую информацию, если каждому символу алфавита сопоставить определенное целое число. Восемью двоичных разрядов достаточно для кодирования 256 различных символов. Этого хватает, чтобы выразить различными комбинациями восьми битов все символы английского и русского языков, как строчные, так и прописные, а также знаки препинания, символы основных арифметических действий и некоторые общепринятые специальные символы.

Для того чтобы весь мир одинаково кодировал текстовые данные, нужны единые таблицы кодирования, а это пока невозможно из-за противоречий между символами национальных алфавитов.

Институт стандартизации США ввел в действие систему кодирования ASCII, в которой закреплены две таблицы кодирования: базовая и расширенная. Базовая таблица закрепляет значения кодов от 0 до 127, а расширенная таблица относится к символам с номерами от 128 до 255.

Базовая таблица системы ASCII содержит 128 кодов. Первые 32 кода базовой таблицы, начиная с нулевого, отданы производителям аппаратных средств. В этой области размещаются управляющие коды, которым не соответствуют никакие символы языков. С 32-го по 127-й код размещены коды символов английского алфавита, знаков препинания, арифметических действий и некоторых вспомогательных символов.

Кодировка символов русского языка, известная как кодировка Windows-1251, была введена компанией Microsoft. Учитывая широкое распространение операционных систем и других продуктов этой компании в России, она глубоко закрепились и нашла широкое распространение. Кодировка символов русского языка закреплена в расширенной таблице кодирования системы ASCII с 192-го по 255-й код.

Большинство систем распознают 256 кодов: 128 стандартных и 128 дополнительных из расширенного набора символов.

Поскольку одному байту соответствует один символ, для представления строки из четырех символов необходимо 4 байта. Вот как выглядит, например, группа символов A12B, состоящая из букв и цифр, в кодировке ASCII:

А	1	2	В
01000001	00110001	00110010	01000010

А вот как выглядит двоичное представление шести символов слова «binary»:

В	И	Н	А	Р	У
01000010	01001001	01001110	01000001	01010010	01011001

В компьютерном тексте, в отличие от текста, напечатанного на пишущей машинке, «пробел» — это значащий символ и, как любой другой символ, он имеет соответствующее двоичное представление. При автоматизированной обработке информации отсутствие или наличие пробела играет важную роль, иногда приводя к путанице и сбивая с толку пользователей-новичков.

Буквам верхнего и нижнего регистров соответствуют разные коды ASCII. Например, прописной букве D соответствует код 68, а строчной d — 100.

Для кодирования букв русского алфавита чаще всего на практике используется кодировка Windows-1251, однако существуют и другие системы кодировки. Распространенной является кодировка КОИ-8 (код обмена информацией восьмизначный). Ее происхождение относится к временам действия Совета экономической взаимопомощи государств Восточной Европы. Сегодня кодировка КОИ-8 широко распространена в компьютерных сетях на территории России.

Международный стандарт, в котором предусмотрена кодировка символов русского языка, называется ISO (International Standard Organization — Международный институт стандартизации). На практике данная кодировка используется редко.

Следует всегда помнить, что компьютеры — это только машины, они не понимают единиц и нулей, зато они способны интерпретировать электрическое напряжение, воспринимая его наличие как 1, а отсутствие — как 0. Эта технология и позволяет компьютерам обрабатывать информацию.

**Кодирование графических данных.** Напечатанное на бумаге черно-белое графическое изображение состоит из мельчайших точек — пикселов (picture element — элемент изображения), образующих характерный узор, называемый растром.

Растровое кодирование позволяет использовать двоичный код для представления графических данных, поскольку линейные координаты и индивидуальные свойства каждой точки (яркость) можно выразить с помощью целых чисел. Общепринятым на сегодняшний день считается представление черно-белых иллюстраций в виде комбинации точек с 256 градациями серого цвета. Следовательно, для кодирования яркости любой точки обычно достаточно 8-разрядного двоичного числа.