

Глава 1

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ

1.1. Информация в природе, обществе и технике

1.1.1. Информация и информационные процессы в неживой природе

В физике, которая изучает неживую природу, *информация является мерой упорядоченности системы по шкале «хаос — порядок»*. Один из основных законов классической физики утверждает, что замкнутые системы, в которых отсутствует обмен веществом и энергией с окружающей средой, стремятся с течением времени перейти из менее вероятного упорядоченного состояния в наиболее вероятное хаотическое состояние.

Например, если в одну половину замкнутого сосуда поместить газ, то через некоторое время в результате хаотического движения молекулы газа равномерно заполнят весь сосуд. Произойдет переход из менее вероятного упорядоченного состояния в более вероятное хаотическое состояние, и информация, которая является мерой упорядоченности системы, в этом случае уменьшится (рис. 1.1).

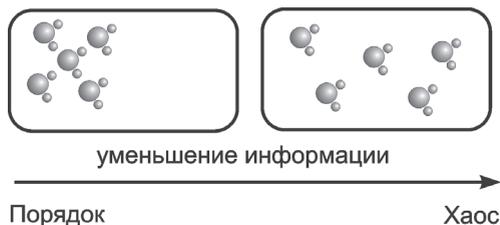


Рис. 1.1. Информация в неживой природе



В соответствии с такой точкой зрения физики в конце XIX века предсказывали, что нашу Вселенную ждет «тепловая смерть», т. е. молекулы и атомы со временем равномерно распределятся в пространстве и какие-либо изменения и развитие прекратятся.

Однако современная наука установила, что некоторые законы классической физики, справедливые для **макросел**, нельзя применять для **микро-** и **мегамира**. Согласно современным научным представлениям, наша Вселенная является динамически развивающейся системой, в которой постоянно происходят процессы усложнения структуры.

Таким образом, с одной стороны, в неживой природе в замкнутых системах идут процессы в направлении от порядка к хаосу (в них информация уменьшается). С другой стороны, в процессе эволюции Вселенной в микро- и мегамире возникают объекты со всё более сложной структурой, и, следовательно, информация, являющаяся мерой упорядоченности элементов системы, возрастает.

Мы живем в **макромире**, т. е. в мире, который состоит из объектов, по своим размерам сравнимых с человеком (рис. 1.2). Обычно макрообъекты разделяют на неживые (камень, льдина и т. д.), живые (растения, животные, сам человек) и искусственные (здания, средства транспорта, станки и механизмы, компьютеры и т. д.).

Макрообъекты состоят из молекул и атомов, которые, в свою очередь, состоят из элементарных частиц, размеры которых чрезвычайно малы. Этот мир называется **микромиром** (рис. 1.3).



Рис. 1.2. Макромир. Гулливер в стране лилипутов

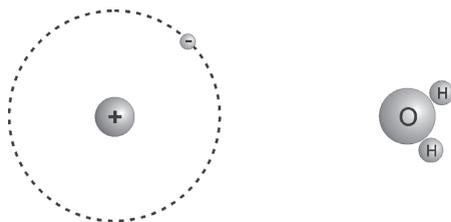


Рис. 1.3. Микромир. Атом водорода и молекула воды

Мы живем на планете Земля, которая входит в Солнечную систему, Солнце вместе с другими звездами образует нашу галактику Млечный Путь, а галактики образуют Вселенную. Все эти объекты имеют громадные размеры и образуют **мегамир** (рис. 1.4).

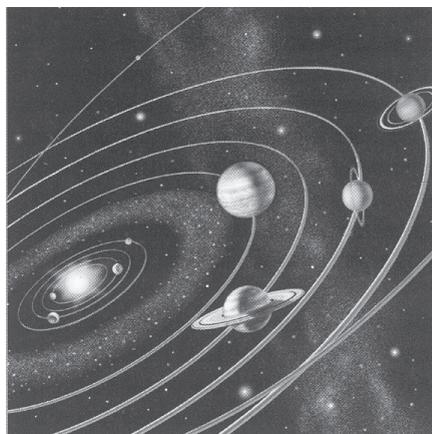


Рис. 1.4. Мегамир. Солнечная система

Контрольные вопросы

1. Приведите примеры перехода от хаоса к порядку (увеличения информации) в окружающем мире.
2. Приведите примеры перехода от порядка к хаосу (уменьшения информации) в окружающем мире.

1.1.2. Информация и информационные процессы в живой природе

Информация как мера увеличения сложности живых организмов. Примерно 3,5 миллиарда лет назад на Земле возникла жизнь. С тех пор идет саморазвитие, эволюция живой природы, т. е. повышение сложности и разнообразия живых организмов. Живые системы (одноклеточные, растения и животные) являются открытыми системами, так как потребляют из окружающей среды вещество и энергию и выбрасывают в нее продукты жизнедеятельности также в виде вещества и энергии.

Живые системы в процессе развития способны повышать сложность своей структуры, т. е. увеличивать информацию, понимаемую как меру упорядоченности элементов системы. Так, растения в процессе фотосинтеза потребляют энергию солнечного излучения и строят сложные органические молекулы из «простых» неорганических молекул.

Животные подхватывают эстафету увеличения сложности живых систем, поедают растения и используют растительные органические молекулы в качестве строительного материала при создании еще более сложных молекул.

Биологи образно говорят, что «живое питается информацией», создавая, накапливая и активно используя информацию.

Информационные сигналы. Нормальное функционирование живых организмов невозможно без получения и использования информации об окружающей среде. Целесообразное поведение живых организмов строится на основе получения **информационных сигналов**. Информационные сигналы могут иметь различную физическую или химическую природу. Это звук, свет, запах и др.

Даже простейшие одноклеточные организмы (например, амeba) постоянно воспринимают и используют информацию, например, о температуре и химическом составе среды для выбора наиболее благоприятных условий существования.

Выживание популяций животных во многом базируется на обмене информационными сигналами между членами одной популяции. Информационный сигнал может быть выражен в различных формах: позах, звуках, запахах и даже вспышках света (ими обмениваются светлячки и некоторые глубоководные рыбы).



Генетическая информация. Одной из основных функций живых систем является размножение, т. е. создание организмов данного вида. Воспроизведение себе подобных обеспечивается наличием **генетической информации**, которая передается по наследству.

Генетическая информация представляет собой набор генов, каждый из которых «отвечает» за определенные особенности строения и функционирования организма. При этом дети не являются точными копиями своих родителей, так как каждый организм обладает уникальным набором генов, который определяет различия в строении и функциональных возможностях.



Контрольные вопросы



1. Приведите примеры информационных сигналов.



2. Для чего нужна генетическая информация? Подготовьте сообщение.

1.1.3. Человек: информация и информационные процессы

Примерно 40 тысяч лет назад в процессе эволюции живой природы появился *человек разумный* (лат. homo sapiens). Человек существует в «море» информации, он постоянно получает информацию из окружающего мира с помощью органов чувств, хранит ее в своей памяти, анализирует с помощью мышления и обменивается информацией с другими людьми.

Способы восприятия информации. Целесообразное поведение человека, так же как и животных, строится на основе анализа информационных сигналов, которые он получает с помощью органов чувств. Нервные окончания органов чувств (рецепторы) воспринимают воздействие (например, на глазном дне колбочки и палочки реагируют на воздействие световых лучей) и передают его по нервной системе в мозг.

Способы восприятия информации живыми организмами зависят от наличия у них тех или иных органов чувств. Человек может использовать разные способы восприятия информации с помощью различных органов чувств:

- зрения — с помощью глаз информация воспринимается в форме зрительных образов;
- слуха — с помощью ушей и органов слуха воспринимаются звуки (речь, музыка, шум и т. д.);
- обоняния — с помощью специальных рецепторов носа воспринимаются запахи;
- вкуса — рецепторы языка позволяют различить сладкое, соленое, кислое и горькое;
- осязания — рецепторы кожи (особенно кончиков пальцев) позволяют получить информацию о температуре объектов и типе их поверхности (гладкая, шершавая и т. д.);
- ориентации в пространстве — гравитационные рецепторы позволяют получить информацию о положении тела в пространстве.

Наибольшее количество информации (около 90%) человек получает с помощью зрения, около 9% — с помощью слуха и только 1% — с помощью других органов чувств (обоняния, осязания, вкуса и ориентации в пространстве).

Полученную информацию в форме зрительных, слуховых и других образов человек хранит в памяти, обрабатывает с помощью мышления и использует для управления своим поведени-

ем и достижения поставленных целей. Например, при переходе дороги человек видит сигналы светофора и движущиеся автомобили, анализирует полученную информацию и выбирает безопасный вариант перехода.

Живые организмы и растения обрабатывают информацию с помощью своих органов и систем. Компьютеры обрабатывают информацию путем выполнения некоторых алгоритмов. Если процесс обработки информации выполняет техническое средство, например компьютер, то вводятся понятия «данные» и «обработка данных».

Данными называют факты, сведения, представленные в формализованном виде (закодированные), занесенные на носители и допускающие обработку с помощью специальных технических средств (в первую очередь компьютером).

Обработка данных предполагает производство различных операций над ними (в первую очередь арифметических и логических) для получения новых данных.

Информация в форме сообщений. Человек не может жить вне общества. В процессе общения с другими людьми человек передает и получает информацию в форме **сообщений**. На заре человеческой истории для передачи информации использовался язык жестов, затем появилась устная речь. В настоящее время обмен сообщениями между людьми производится с помощью сотен естественных языков (русского, английского и других) и искусственных языков (эсперанто, Бейсик, С++ и других).

Сообщение — форма представления информации в виде совокупности знаков (символов) какого-либо алфавита естественного или искусственного языка, используемая для передачи. Каждый новый символ в сообщении увеличивает объем информации, представленной последовательностью символов конкретного алфавита.

Для того чтобы информация была **понятна**, язык должен быть известен всем людям, участвующим в общении. Чем большее количество языков вы знаете, тем шире круг вашего общения. Понятность — это одно из свойств информации.

Согласно библейской легенде о вавилонском столпотворении, строившаяся в древнем городе Вавилоне башня не была закончена и разрушилась, так как сотни строителей вдруг заговорили на различных языках и перестали понимать друг друга (рис. 1.5).



Рис. 1.5. Вавилонская башня. Картина Брейгеля-старшего

Информация в форме знаний. С самого начала человеческой истории возникла потребность накопления информации для ее передачи во времени из поколения в поколение и передачи в пространстве на большие расстояния. Процесс накопления информации начался с изобретения в IV тысячелетии до нашей эры письменности и первых носителей информации (шумерских глиняных табличек и древнеегипетских папирусов).

Для того чтобы человек мог правильно ориентироваться в окружающем мире, информация должна быть **полной** и **точной**. Полнота и точность — это еще два свойства информации. Задача получения полной и точной информации о природе, обществе и технике стоит перед наукой. Процесс систематического научного познания окружающего мира, в котором информация рассматривается как **знания**, начался с середины XV века после изобретения книгопечатания.

Для долговременного хранения знаний (передачи из поколения в поколение) и распространения их в обществе (тиражирования) необходимы носители информации. Материальная природа носителей информации может быть разной.

Вплоть до наших дней в качестве основного носителя информации используется бумага. В прошлом веке широкое распространение для хранения графической информации получила фото- и киноплёнка. В настоящее время для хранения информации широко применяются также магнитные носители, флеш-память и оптические носители (CD и DVD) (рис. 1.6).





Рис. 1.6. Первая печатная книга и современный оптический диск

Средства массовой информации. Широко известен термин «средства массовой информации» — СМИ (газеты, радио, телевидение), которые доводят информацию до каждого члена общества. Такая информация должна быть **достоверной, актуальной и полезной**. Это свойства информации, важные для СМИ. Недостоверная информация вводит членов общества в заблуждение и может быть причиной возникновения социальных потрясений. Неактуальная информация не имеет применения в настоящий момент времени, и поэтому никто, кроме историков, не читает прошлогодних газет. Бесполезная информация создает информационный шум, который затрудняет восприятие полезной информации.

Информация, ее формы, свойства, информационные процессы являются объектом исследования целого ряда научных дисциплин, таких как **теория информации** (математическая теория систем передачи информации), **кибернетика** (наука о связи и управлении в машинах и животных, а также в обществе и человеческих существах), **семиотика** (наука о знаках и знаковых системах), **теория массовой коммуникации** (исследование средств массовой информации и их влияния на общество), **информатика** (изучение процессов сбора, преобразования, хранения, защиты, поиска и передачи всех видов информации и средств их автоматизированной обработки), **информодинамика** (наука об открытых информационных системах), **информациология** (наука о получении, сохранении и передаче информации для различных множеств объектов), что и определяет общие закономерности протекания информационных процессов в системах различной природы.



Контрольные вопросы



1. Какие способы и органы чувств использует человек при восприятии информации?



2. Каковы должны быть свойства информации, представленной в форме сообщений?
3. Каковы должны быть свойства информации, представленной в форме знаний?
4. Каковы должны быть свойства информации, распространяемой средствами массовой информации?

1.1.4. Информация и информационные процессы в технике

Системы управления техническими устройствами. Функционирование систем управления техническими устройствами связано с информационными процессами, т. е. процессами приема, хранения, обработки и передачи информации. Системы управления могут выполнять различные функции. Например, такие системы могут поддерживать определенное состояние технической системы.

Так системы автоматической терморегуляции холодильника, утюга и кондиционера обеспечивают поддержание заданной температуры. В системе терморегуляции управляющее устройство получает информацию от температурных датчиков, обрабатывает ее (сравнивает реальную температуру с заданной) и передает команды нагревательному элементу (усилить или уменьшить нагрев) (рис. 1.7).

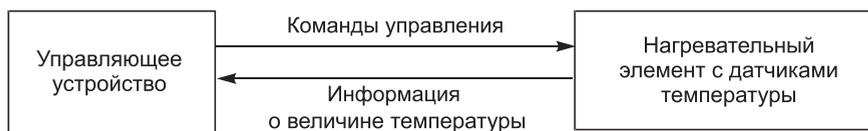


Рис. 1.7. Система управления, регулирующая температуру

Системы управления встроены практически во всю современную бытовую технику, станки с числовым программным управлением, транспортные средства и пр.

Системы управления могут обеспечивать функционирование технической системы по заданной программе. Например, системы программного управления обеспечивают стирку в стиральной машине в заданном режиме, обработку детали на станке с программным управлением.

В некоторых случаях главную роль в процессе управления выполняет человек, в других управление осуществляет встроенный в техническое устройство **микропроцессор** или подключенный компьютер. Например, управление полетом самолета может осуществлять летчик или в режиме автопилота борто-

вой компьютер. Они получают информацию о режиме полета от датчиков (скорости, высоты и пр.), обрабатывают ее и передают команды на исполнительные механизмы (закрылки, клапаны, регулирующие работу двигателей, и пр.), изменяющие режим полета.

Первый микропроцессор Intel 4004, «дедушка» современных процессоров, был разработан в 1971 году специально для использования в автоматизированных системах управления. Процессор включал 2300 электронных переключателей, обладал памятью объемом 640 байт и мог выполнять 100 тысяч операций в секунду (рис. 1.8).

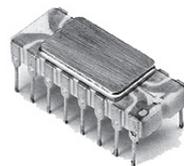


Рис. 1.8. Первый микропроцессор Intel 4004

Роботы. Роботами называются автоматические устройства, предназначенные для осуществления производственных, научных и других работ. Роботы могут иметь различные внешний вид и размеры, но все они выполняют действия на основании заложенной в них программы обработки информации.

Промышленные роботы обычно заменяют человека в тех отраслях производства, где требуется проведение утомительных и однообразных работ (например, конвейерная сборка автомобилей и электронных устройств), опасных технических работ (например, работа с радиоактивными материалами), а также работ, где присутствие человека физически невозможно (например, автоматические космические и глубоководные аппараты).

В последние годы появились роботы, оснащенные органами чувств, аналогичными органам чувств человека (зрение, слух, тактильные ощущения), имеющие память и способные обрабатывать полученную информацию и осуществлять целенаправленные действия. Такие роботы могут работать дома (уже производится робот-пылесос), в больнице (экспериментальные образцы разносят больным лекарства), на других планетах (лунноходы и марсоходы путешествуют по поверхностям небесных тел) и т. д.

Большой интерес всегда вызывают роботы, подобные человеку или животным по внешнему виду и действиям. Они могут ходить, преодолевать препятствия, реагировать на внешние раздражители и даже разговаривать (рис. 1.9), что характерно для систем искусственного интеллекта. Искусственный интеллект (*англ.* Artificial intelligence, AI) — это наука и ряд технологий создания интеллекту-

альных («думающих») машин, особенно интеллектуальных компьютерных программ. Искусственный интеллект связан со сходной задачей использования компьютеров для понимания человеческого интеллекта..



Рис. 1.9. Собакообразный и человекообразный роботы

Контрольные вопросы

1. Приведите примеры систем управления техническими устройствами.
2. Подготовьте доклад об использовании роботов в различных сферах деятельности.



1.2. Кодирование информации с помощью знаковых систем

1.2.1. Знаки: форма и значение

С древних времен знаки используются человеком для долговременного хранения информации и ее передачи на большие расстояния.

Форма знаков. В соответствии со способом восприятия знаки можно разделить на зрительные, слуховые, осязательные, обонятельные и вкусовые, причем в человеческом общении используются знаки первых трех типов.

К зрительным знакам, воспринимаемым с помощью зрения, относятся буквы и цифры (они используются в письменной речи), знаки химических элементов, музыкальные ноты, дорожные знаки и т. д.

К слуховым знакам, воспринимаемым с помощью слуха, относятся звуки (они используются в устной речи), а также звуковые сигналы, которые производятся с помощью звонка, колокола, свистка, гудка, сирены и т. д.

Для слепых разработана азбука Брайля, которая использует осязательный способ восприятия текстовой информации. К ряду осязательных знаков принадлежат также жесты-касания: рукопожатия, похлопывания по плечу и др.

В коммуникации многих видов животных особую роль играют обонятельные знаки. Например, медведи и другие дикие животные помечают место обитания клочьями шерсти, сохраняющей запах, чтобы отпугнуть чужака и показать, что данная территория уже занята.

Для длительного хранения знаки записываются на носители информации.

Для передачи информации на большие расстояния используются знаки в форме **сигналов**. Всем известны световые сигналы светофора, звуковые сигналы школьного звонка оповещают о начале или конце урока, электрические сигналы передают информацию по телефонным и компьютерным сетям, электромагнитные волны передают сигналы радио и телевидения.

Значение знаков. Знаки отображают объекты окружающего мира или понятия, т. е. имеют определенное значение (смысл).

Знаки различаются по способу связи между их формой и значением. **Иконические знаки** позволяют догадаться об их смысле, так как имеют форму, похожую на отображаемый объект. Примером таких знаков являются значки на *Рабочем столе* операционной системы компьютера, например значок *Компьютер*.

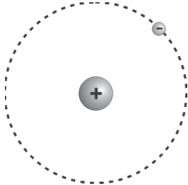
Символами называются знаки, для которых связь между формой и значением устанавливается по общепринятому соглашению. Примером таких знаков являются символы химических элементов, отображающие атомы химических веществ (табл. 1.1).

Если неизвестно соглашение о связи формы и значения символов, то ничего нельзя сказать о смысле информации, записанной такими знаками. Существуют найденные археологами и до сих пор не расшифрованные тексты на древних языках, так как неизвестно значение знаков, которыми они записаны.

В современном мире широко применяется шифрование, которое использует секретный ключ в качестве соглашения о связи формы символов с их значениями. Если секретный ключ неизвестен, то без применения специальных методов дешифровки содержание передаваемого текста понять невозможно.

Таблица 1.1

Иконические знаки и символы

Тип знака	Форма знака	Отображаемый объект
Иконический знак		
Символ	Н	

Контрольные вопросы

1. Приведите примеры зрительных, слуховых, осязательных, обонятельных и вкусовых знаков. Какие типы знаков применяются в человеческом общении?
2. Приведите примеры знаков в форме сигналов.
3. В чем состоит различие между иконическими знаками и символами?
4. Приведите примеры символов, которые могут иметь несколько значений.

**1.2.2. Знаковые системы**

В основе знаковой системы лежит набор знаков, называемый **алфавитом**. Эти знаки имеют определенную физическую природу. С некоторыми знаковыми системами вы хорошо знакомы и постоянно ими пользуетесь (языки, числа, дорожные знаки), с другими познакомитесь в этом параграфе.

Каждая **знаковая система** строится на основе определенного алфавита (набора знаков) и правил выполнения операций над знаками.



Естественные языки. Человек широко использует для представления информации знаковые системы, которые называются **языками**. Естественные языки начали формироваться еще в древнейшие времена в целях обеспечения обмена информацией между людьми. В настоящее время существуют сотни естественных языков (русский, английский, китайский и др.).

В устной речи, которая используется как средство коммуникации при непосредственном общении людей, в качестве знаков языка используются различные звуки (**фонемы**).

В основе письменной речи лежит алфавит, т. е. набор знаков (букв), которые человек различает по их рисунку. В большинстве современных языков буквы соответствуют определенным звукам устной речи. Алфавит русского языка называется **кириллицей** и содержит 33 знака, английский язык использует **латиницу** и содержит 26 знаков.

На основе алфавита по правилам **грамматики** образуются основные объекты языка — слова. Правила, согласно которым из слов данного языка строятся предложения, называются **синтаксисом**. Необходимо отметить, что в естественных языках грамматика и синтаксис языка формулируются с помощью большого количества правил, из которых существуют исключения, так как такие правила складывались исторически. Кроме того, естественные языки допускают много вариантов передачи знаками смысла сообщений. Одну и ту же информацию можно передать разными предложениями.

Формальные языки. В процессе развития науки были разработаны **формальные языки** (системы счисления, язык алгебры, языки программирования и др.), отличие которых от естественных языков состоит в существовании ограниченного количества строгих правил грамматики и синтаксиса и в однозначной записи знаками смысла сообщения.

Например, десятичную систему счисления можно рассматривать как формальный язык, имеющий алфавит (цифры) и позволяющий именовать и записывать объекты (числа) и выполнять над ними арифметические операции по строго определенным правилам.

Существуют формальные языки, в которых в качестве знаков используют не буквы и цифры, а другие символы, например обозначения химических элементов, музыкальных нот, изображения элементов электрических или логических схем, дорожные знаки, точки и тире (код азбуки Морзе).

Физическая реализация знаков в естественных и формальных языках может быть различной. Например, текст и числа могут быть напечатаны на бумаге, высвечены на экране монитора компьютера, записаны на магнитном или оптическом диске.



Генетический алфавит. Он используется живыми организмами для строительства единой системы хранения и передачи наследственной информации.

Генетическая информация хранится в клетках живых организмов в специальных молекулах. Эти молекулы состоят из двух длинных скрученных друг с другом в спираль цепей, построенных из молекулярных фрагментов четырех различных типов. Эти фрагменты образуют «генетический алфавит» и обычно обозначаются латинскими прописными буквами {A, G, C, T} (рис. 1.10).

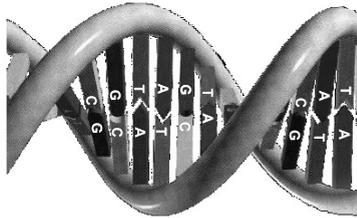


Рис. 1.10. Модель молекулы ДНК, содержащей генетический код

Как слова в языках записывают с помощью букв, так и гены состоят из знаков генетического алфавита. В процессе эволюции от простейших организмов до человека количество генов постоянно возрастало, так как было необходимо закодировать все более сложное строение и функциональные возможности живых организмов.

Двоичная знаковая система. В процессах хранения, обработки и передачи информации в компьютере используется двоичная знаковая система, алфавит которой состоит всего из двух знаков {0, 1}. Физически знаки реализуются в форме электрических импульсов (нет импульса — 0, есть импульс — 1), а также состояний ячеек оперативной памяти и участков поверхностей носителей информации (одно состояние — 0, другое состояние — 1).

Именно двоичная знаковая система используется в компьютере, так как существующие технические устройства могут надежно сохранять и распознавать только два различных состояния (знака).



В 60-е годы XX века в СССР учеными Московского государственного университета им. М. В. Ломоносова была разработана и запущена в производство ЭВМ «Сетунь» (всего было произведено 50 экземпляров) (рис. 1.11). «Сетунь» использовала троичное кодирование информации и, соответственно, состояла из устройств, способных находиться в одном из трех возможных состояний.

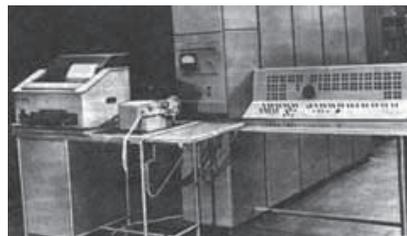


Рис. 1.11. ЭВМ «Сетунь»



Контрольные вопросы

1. Приведите примеры знаковых систем. Какой может быть физическая природа знаков?
2. В чем состоит различие между естественными и формальными языками?
3. Обладают ли генетическим кодом растения? Животные? Человек?
4. Почему в компьютерах используется двоичная знаковая система для кодирования информации?



Задания для самостоятельного выполнения

- 1.1. *Задание с развернутым ответом.* Заполните таблицу: введите алфавит и перечислите возможную физическую природу знаков для различных знаковых систем.

Знаковая система	Алфавит	Физическая природа знаков
Русский язык (письменный)		
Русский язык (устный)		
Английский язык (письменный)		
Десятичная система счисления		
Генетический алфавит		
Двоичный компьютерный код		

1.2.3. Кодирование информации

В процессах восприятия, передачи и хранения информации живыми организмами, человеком и техническими устройствами происходит ее кодирование.

Код. Длина кода. В процессе представления информации с помощью знаковой системы производится ее **кодирование**. Результатом кодирования является последовательность знаков данной знаковой системы, т. е. **информационный код**. Отдельная комбинация таких знаков – **кодовое слово**.

Примерами кодов являются последовательности букв в тексте, цифр в числе, генетический код, двоичный компьютерный код и т. д.

Код состоит из определенного количества знаков (например, текстовое сообщение состоит из определенного количества букв, число — из определенного количества цифр и т. д.), т. е. имеет определенную длину.

Количество знаков в коде называется длиной кода.

Так, длина кода текста данного учебника составляет около 300 тысяч знаков, а генетический код человека в 10 тысяч раз длиннее, так как состоит из 3 миллиардов знаков генетического алфавита.

Перекодирование информации из одной знаковой системы в другую. В процессе обмена информацией между людьми часто приходится переходить от одной формы представления информации к другой. Так, в процессе чтения вслух производится переход от письменной формы представления информации к устной и, наоборот, в процессе выполнения диктанта или записи объяснения учителя происходит переход от устной формы к письменной. В процессе преобразования информации из одной формы представления (знаковой системы) в другую происходит **перекодирование информации**.

Перекодирование — это операция преобразования знаков или групп знаков одной знаковой системы в знаки или группы знаков другой знаковой системы.

Средством перекодирования служит таблица соответствия знаковых систем (таблица перекодировки), которая устанавливает взаимно однозначное соответствие между знаками или группами знаков двух различных знаковых систем. Таблица 1.2 устанавливает соответствие между гласными буквами русского алфавита и фонемами.



В русской письменной речи только шесть гласных букв могут быть озвучены в устной речи соответствующими звуками. Для озвучивания остальных четырех гласных букв используются составные звуки, которые начинаются со звука [j].

Таблица 1.2

Соответствие букв и звуков

Буквы	Звуки (фонемы)
а	[а]
о	[о]
у	[у]
и	[и]
ы	[ы]
э	[э]
е	[j] + [э]
ё	[j] + [о]
ю	[j] + [у]
я	[j] + [а]



Контрольные вопросы



1. Приведите примеры кодов и определите их длины.
2. Приведите примеры перекодирования информации из одной знаковой системы в другую. Какие в этих случаях используются таблицы перекодировки?



Задания для самостоятельного выполнения

- 1.2. *Задание с кратким ответом.* Перекодируйте с русского письменного языка на русский устный имя Юля.

1.3. Количество информации

1.3.1. Количество информации как мера уменьшения неопределенности знания

Процесс познания окружающего мира приводит к накоплению информации в форме знаний (фактов, научных теорий и т. д.). Получение новой информации приводит к расширению знания или, как иногда говорят, к уменьшению неопределенности зна-

ния. Если некоторое сообщение приводит к уменьшению неопределенности нашего знания, то можно говорить, что такое сообщение содержит информацию.

Например, после сдачи зачета или выполнения контрольной работы вы мучаетесь неопределенностью, вы не знаете, какую оценку получили. Наконец, учитель объявляет результаты, и вы получаете одно из двух информационных сообщений: «зачет» или «незачет», а после контрольной работы одно из четырех информационных сообщений: «2», «3», «4» или «5».

Информационное сообщение об оценке за зачет приводит к уменьшению неопределенности вашего знания в два раза, так как получено одно из двух возможных информационных сообщений. Информационное сообщение об оценке за контрольную работу приводит к уменьшению неопределенности знания в четыре раза, так как получено одно из четырех возможных информационных сообщений.

Ясно, что чем более неопределенна первоначальная ситуация (возможно большее количество информационных сообщений), тем больше мы получим новой информации при получении информационного сообщения (тем в большее количество раз уменьшится неопределенность знания).

Рассмотренный подход к информации как мере уменьшения неопределенности знания позволяет количественно измерять информацию.

Количество информации можно рассматривать как меру уменьшения неопределенности знания при получении информационного сообщения.

Существует формула, которая связывает между собой количество возможных информационных сообщений N и количество информации I , которое несет полученное сообщение:

$$N = 2^I. \quad (1.1)$$

Бит. Для количественного выражения любой величины необходимо сначала определить единицу измерения. Так, для измерения длины в качестве единицы выбран метр, для измерения массы — килограмм и т. д. Аналогично, для определения количества информации необходимо ввести единицу измерения.



За единицу количества информации принимается количество информации, содержащееся в информационном сообщении, уменьшающем неопределенность знания в два раза. Такая единица названа **бит**.

Если вернуться к рассмотренному выше получению информационного сообщения о результатах зачета, то здесь неопределенность как раз уменьшается в два раза и, следовательно, количество информации, которое несет сообщение, равно 1 биту.

Производные единицы измерения количества информации. Минимальной единицей измерения количества информации является **бит**, а следующей по величине единицей — **байт**, причем:

$$1 \text{ байт} = 8 \text{ бит} = 2^3 \text{ бит.}$$

В информатике система образования кратных единиц измерения несколько отличается от принятых в большинстве наук. Традиционные метрические системы единиц, например Международная система единиц СИ, в качестве множителей кратных единиц используют коэффициент 10^n , где $n = 3, 6, 9$ и т. д., что соответствует десятичным приставкам «кило» (10^3), «мега» (10^6), «гига» (10^9) и т. д.

В компьютере информация кодируется с помощью двоичной знаковой системы, поэтому в кратных единицах измерения количества информации используется коэффициент 2^n .

Так, кратные байту единицы измерения количества информации вводятся следующим образом:

$$1 \text{ килобайт (Кбайт)} = 2^{10} \text{ байтов} = 1024 \text{ байтов};$$

$$1 \text{ мегабайт (Мбайт)} = 2^{10} \text{ Кбайт} = 1024 \text{ Кбайт};$$

$$1 \text{ гигабайт (Гбайт)} = 2^{10} \text{ Мбайт} = 1024 \text{ Мбайт};$$

$$1 \text{ терабайт (Тбайт)} = 2^{10} \text{ Гбайт} = 1024 \text{ Гбайт};$$

$$1 \text{ петабайт (Пбайт)} = 2^{10} \text{ Тбайт} = 1024 \text{ Тбайт};$$

$$1 \text{ эксабайт (Эбайт)} = 2^{10} \text{ Пбайт} = 1024 \text{ Пбайт.}$$



Контрольные вопросы

1. Приведите примеры информационных сообщений, которые приводят к уменьшению неопределенности знания.
2. Приведите примеры информационных сообщений, которые несут 1 бит информации.

Задания для самостоятельного выполнения

- 1.3. *Задание с выборочным ответом.* За минимальную единицу измерения количества информации примут:
- 1) 1 бод 2) 1 пиксель 3) 1 байт 4) 1 бит
- 1.4. *Задание с кратким ответом.* Вычислите, какое количество информации в битах содержится в 1 килобайте, 1 мегабайте и 1 гигабайте.

1.3.2. Определение количества информации

Определение количества информационных сообщений. По формуле (1.1) можно легко определить количество возможных информационных сообщений, если известно количество информации. Например, пусть на экзамене студент берет экзаменационный билет, и учитель сообщает ему, что зрительное информационное сообщение о номере билета несет 5 бит информации. Если вы хотите определить количество экзаменационных билетов, то достаточно определить количество возможных информационных сообщений об их номерах по формуле (1.1):

$$N = 2^5 = 32.$$

Таким образом, количество экзаменационных билетов равно 32.

Определение количества информации. Наоборот, если известно возможное количество информационных сообщений N , то для определения количества информации, которое несет сообщение, необходимо решить уравнение (1.1) относительно I .

Представьте себе, что вы управляете движением робота и можете задавать направление его движения с помощью информационных сообщений: «север», «северо-восток», «восток», «юго-восток», «юг», «юго-запад», «запад» и «северо-запад» (рис. 1.12). Какое количество информации будет получать робот после каждого сообщения?

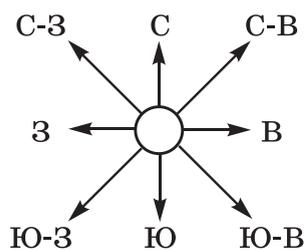


Рис. 1.12. Управление роботом с использованием информационных сообщений

Всего возможных информационных сообщений 8, поэтому формула (1.1) принимает вид уравнения относительно I :

$$8 = 2^I.$$

Разложим стоящее в левой части уравнения число 8 на множители, равные 2, и представим его в степенной форме:

$$8 = 2 \cdot 2 \cdot 2 = 2^3.$$

Получим степенное уравнение, из которого необходимо найти неизвестное I :

$$2^3 = 2^I.$$

Равенство левой и правой частей уравнения справедливо, если равны показатели степени числа 2. Таким образом, $I = 3$ бита, т. е. количество информации, которое несет роботу каждое информационное сообщение, равно 3 битам.



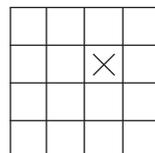
Задания для самостоятельного выполнения

1.5. Задание с выборочным ответом. Производится бросание симметричной четырехгранной пирамидки. Какое количество информации мы получаем в зрительном сообщении о ее падении на одну из граней?

- 1) 1 бит 2) 2 бита 3) 4 бита 4) 1 байт

1.6. Задание с кратким ответом. Из непрозрачного мешочка вынимают шарики с номерами, и известно, что информационное сообщение о номере шарика несет 5 битов информации. Определите количество шариков в мешочке.

1.7. Задание с развернутым ответом. Какое количество информации при игре в крестики-нолики на поле размером 4×4 клетки получит второй игрок после первого хода первого игрока?



1.3.3. Алфавитный подход к определению количества информации

При алфавитном подходе к определению количества информации отвлекаются от содержания (смысла) информации и рассматривают информационное сообщение как последовательность знаков определенной знаковой системы.

Информационная емкость знака. Представим себе, что необходимо передать информационное сообщение по каналу передачи информации от отправителя к получателю (рис. 1.13). Пусть сообщение кодируется с помощью знаковой системы, алфавит которой состоит из N знаков $\{1, \dots, N\}$. В простейшем случае, когда длина кода сообщения составляет один знак, отправитель может послать одно из N возможных сообщений: $1, 2, \dots, N$, которое будет нести количество информации I .

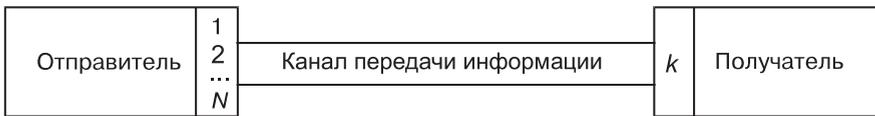


Рис. 1.13. Передача информации

Формула (1.1) связывает между собой количество возможных информационных сообщений N и количество информации I , которое несет полученное сообщение. В рассматриваемой ситуации N — это количество знаков в алфавите знаковой системы, а I — количество информации, которое несет каждый знак:

$$N = 2^I.$$

С помощью этой формулы можно, например, определить количество информации, которое несет один знак в двоичной знаковой системе:

$$N = 2 \Rightarrow 2 = 2^I \Rightarrow 2^1 = 2^I \Rightarrow I = 1 \text{ бит.}$$

Таким образом, в двоичной знаковой системе знак несет 1 бит информации. Интересно, что сама единица измерения количества информации «бит» (bit) получила свое название от английского словосочетания «Binary digiT» — «двоичная цифра».

Информационная емкость знака двоичной знаковой системы составляет **1 бит**.



Чем большее количество знаков содержит алфавит знаковой системы, тем большее количество информации несет один знак. В качестве примера определим количество информации, которое несет буква русского алфавита. В русский алфавит входят 33 буквы, однако на практике часто для передачи сообщений используются только 32 буквы (исключается буква «ё»). Алфавит — конечное множество символов (N).

С помощью формулы (1.1) определим количество информации, которое несет одна буква русского алфавита:

$$N = 32 \Rightarrow 32 = 2^I \Rightarrow 2^5 = 2^I \Rightarrow I = 5 \text{ бит.}$$

Таким образом, буква русского алфавита несет 5 бит информации.



Количество информации, которое несет знак, зависит от вероятности его получения. Если получатель заранее точно знает, какой знак придет, то полученное количество информации будет равно 0. Наоборот, чем менее вероятно получение знака, тем больше его информационная емкость.

В русской письменной речи частота использования букв в тексте различна. Так, в среднем на 1000 знаков осмысленного текста приходится 200 букв «а» и в 100 раз меньшее количество буквы «ф» (всего 2). Таким образом, с точки зрения теории информации, информационная емкость знаков русского алфавита различна (у буквы «а» она наименьшая, а у буквы «ф» — наибольшая).

Количество информации в сообщении. Сообщение состоит из последовательности знаков, каждый из которых несет определенное количество информации.

Если знаки несут одинаковое количество информации, то количество информации I_c в сообщении можно подсчитать, умножив количество информации I , которое несет один знак, на длину кода (количество знаков в сообщении) K :

$$I_c = I \cdot K. \tag{1.2}$$



Так, каждая цифра двоичного компьютерного кода несет 1 бит информации. Следовательно, две цифры несут 2 бита информации, три цифры — 3 бита и т. д. Количество информации в битах равно количеству цифр двоичного компьютерного кода (табл. 1.3).

Таблица 1.3

Количество информации, которое несет двоичный компьютерный код

Двоичный компьютерный код	1	0	1	0	1
Количество информации	1 бит				

Объем данных. В технике, где информацией считается любая хранящаяся, обрабатываемая или передаваемая последовательность знаков, сигналов, часто используют простой способ определения количества информации, который может быть назван **объемным**. Он основан на подсчете числа символов в сообщении, т. е. связан только с длиной сообщения и не учитывает его содержания.

Длина сообщения зависит от числа знаков, употребляемых для записи сообщения. Например, слово «мир» в русском алфавите записывается тремя знаками, в английском — пятью (pease), а в кодовой таблице КОИ-8 — двадцатью четырьмя битами (11101101 11101001 11110010).

Задания для самостоятельного выполнения

- 1.8. *Задание с выборочным ответом.* Какое количество информации содержит один разряд двоичного числа?
1) 1 байт 2) 3 бита 3) 4 бита 4) 1 бит
- 1.9. *Задание с кратким ответом.* Какое количество информации несет двоичный код 10101010?
- 1.10. *Задание с кратким ответом.* Какова информационная емкость знака генетического кода?