

§ 2. Представление информации в компьютере

Кодирование информации использовалось с древности. Широко известен шифр Юлия Цезаря, применявшийся для записи секретных сообщений. Каждый символ в тексте заменялся символом, находящимся на некотором постоянном расстоянии левее или правее его в алфавите.



Например, при кодировании информации с помощью букв русского алфавита путем сдвига вправо на 3 буква «А» была бы заменена на «Г», буква «Б» станет «Д» и т. д.

Пример 2.1. Сегодня широко применяются штрих-коды на различных товарах. Перед вами штрих-код сгущенного молока:



2.1. Кодирование информации

Для общения люди используют естественный язык, например белорусский или русский. В основе естественного языка лежит алфавит — система графических знаков для передачи звуков устной речи. Алфавит естественного языка является универсальным кодом любой письменной культуры.

Кроме естественных, человек использует искусственно созданные языки со своими особыми кодами: язык математических или химических формул, ноты и др.

Код — совокупность условных знаков, каждому из которых присваивается определенное значение (примеры 2.1 и 2.2).

Процесс записи или преобразования информации в соответствии с правилами, заданными некоторым кодом, называют **кодированием**. Процесс, обратный кодированию, называют **декодированием**.

Кодировать и передавать информацию можно различными способами: устно, письменно, жестами и др. Компьютер может обрабатывать числовую, текстовую, графическую и звуковую информацию только в цифровом формате, который в компьютере представлен в виде двоичного кода.

Двоичный код — способ кодирования, в котором каждый **разряд** принимает одно из двух возможных значений, обычно обозначаемых цифрами 0 и 1. **Разряд** в этом случае называется **двоичным разрядом**.

Такой способ кодирования связан с тем, что проще всего реализуются технические устройства, обладающие двумя устойчивыми состояниями: включено/выключено, соединено/разъединено и др.

Для кодирования числовой информации в компьютере вместо десятичной системы счисления используется двоичная, основанная на двоичном коде.

Кодирование текстовой информации в компьютере выполняется при помощи специальных кодовых таблиц, где каждому символу ставится в соответствие определенная последовательность из нулей и единиц (пример 2.3).

Пример 2.2. С появлением смартфонов начали распространяться QR-коды. Они позволяют быстро заносить в телефон текстовую информацию, добавлять контакты в адресную книгу, переходить по web-ссылкам, отправлять SMS-сообщения и т. д. Вот, например, QR-код со ссылкой на статью в Wikipedia о QR-кодах:

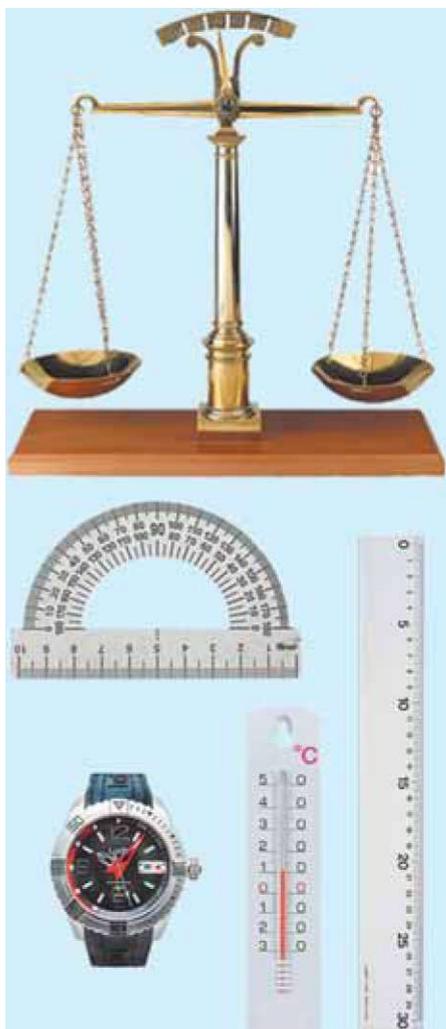


Пример 2.3. Кодирование некоторых букв английского алфавита на компьютере.

Буква	Двоичный код
<i>A</i>	01000001
<i>B</i>	01000010
<i>C</i>	01000011
<i>D</i>	01000100
<i>E</i>	01000101
<i>F</i>	01000110
<i>G</i>	01000111
<i>H</i>	01001000
<i>I</i>	01001001
<i>J</i>	01001010
<i>K</i>	01001011
<i>L</i>	01001100
<i>M</i>	01001101
<i>N</i>	01001110



Пример 2.4.



Пример 2.5. Соотношение между битом и байтом.

Бит

0 0 1 1 0 1 0 1

Байт (8 бит)

2.2. Единицы измерения объема информации

Человек применяет различные единицы измерения. Так, для измерения времени используются секунды, минуты, часы, для измерения расстояния — метры, километры и др. Измерения проводят с помощью измерительных приборов (пример 2.4).

Для определения количества информации есть свои единицы измерения. Минимальное количество информации, для кодирования которой достаточно одного двоичного разряда, называют **битом** (bit).

Слово «бит» произошло от английских слов *binary* (двоичный) и *digit* (знак). Бит — минимальная единица, выражающая количество информации. Он может принимать одно из двух значений — 0 или 1. Для удобства введена более крупная единица измерения информации — байт.

Байт — единица измерения количества информации, состоящая из восьми последовательных и взаимосвязанных битов.

1 байт = 2^3 бит = 8 бит (пример 2.5).

Правообладатель Народная асвета

Для обозначения большего объема информации используются другие единицы измерения:

$$1 \text{ Кбайт (килобайт)} = 1024 \text{ байта};$$

$$1 \text{ Мбайт (мегабайт)} = 1\,048\,576 \text{ байт};$$

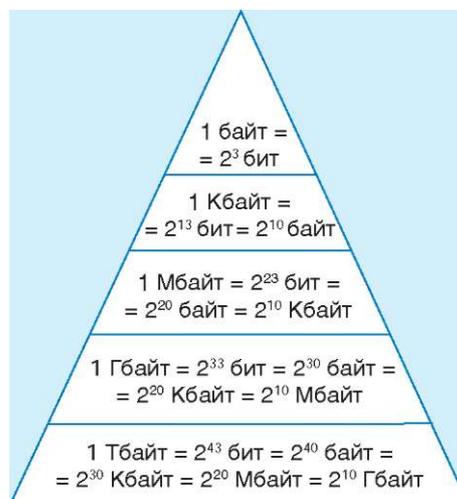
$$1 \text{ Гбайт (гигабайт)} = 1\,073\,741\,824 \text{ байта};$$

$$1 \text{ Тбайт (терабайт)} = 1\,099\,511\,627\,776 \text{ байт}.$$

Значения данных единиц измерения информации для удобства кодирования связаны со степенью числа 2 (пример 2.6).

В этих единицах измеряются количество (объем) оперативной или внешней памяти компьютера, размеры файлов. В примере 2.7 показано, каким образом выполняется перевод одних единиц измерения информации в другие.

Пример 2.6. Соотношение единиц измерения информации.



Пример 2.7.

Переведем 2368 Мбайт в килобайты и гигабайты:

$$\begin{aligned} 2368 \text{ Мбайт} &= \\ &= (2368 \cdot 2^{10}) \text{ Кбайт} = \\ &= 2\,424\,832 \text{ Кбайт}; \end{aligned}$$

$$\begin{aligned} 2368 \text{ Мбайт} &= \\ &= (2368 / 2^{10}) \text{ Гбайт} - \\ &- 2,3 \text{ Гбайт}. \end{aligned}$$