

- **символы** — связь между их формой и значением устанавливается по общепринятому соглашению (например, буквы, математические символы  $\int, \leq, \subseteq, \infty$ ; символы химических элементов).

Для представления информации используются знаковые системы, которые называются **языками**. Основу любого языка составляет **алфавит** — набор символов, из которых формируется сообщение, и набор правил выполнения операций над символами.

Языки делятся на:

- **естественные** (разговорные) — русский, английский, немецкий и др.;
- **формальные** — встречающиеся в специальных областях человеческой деятельности (например, язык алгебры, языки программирования, электрических схем и др.)

Системы счисления также можно рассматривать как формальные языки. Так, десятичная система счисления — это язык, алфавит которого состоит из десяти цифр 0..9, двоичная система счисления — язык, алфавит которого состоит из двух цифр — 0 и 1.

## Методы измерения количества информации: вероятностный и алфавитный

Единицей измерения количества информации является *бит*. **1 бит** — это количество информации, содержащейся в сообщении, которое вдвое уменьшает неопределенность знаний о чем-либо.

Связь между количеством возможных событий  $N$  и количеством информации  $I$  определяется **формулой Хартли**:

$$N = 2^I.$$

Например, пусть шарик находится в одной из четырех коробок. Таким образом, имеется четыре равновероятных события ( $N = 4$ ). Тогда по формуле Хартли  $4 = 2^I$ . Отсюда  $I = 2$ . То есть сообщение о том, в какой именно коробке находится шарик, содержит 2 бита информации.

### Алфавитный подход

При алфавитном подходе к определению количества информации отвлекаются от содержания (смысла) информации и рассматривают ее как последовательность знаков определенной знаковой системы. Набор символов языка (алфавит) можно рассматривать как различные возможные события. Тогда, если считать, что появление символов в сообщении равновероятно, по формуле Хартли можно рассчитать, какое количество информации несет каждый символ:

$$I = \log_2 N.$$

Например, в русском языке 32 буквы (буква ё обычно не используется), т. е. количество событий будет равно 32. Тогда информационный объем одного символа будет равен:

$$I = \log_2 32 = 5 \text{ битов.}$$

Если  $N$  не является целой степенью 2, то число  $\log_2 N$  не является целым числом, и для  $I$  надо выполнять округление в большую сторону. При решении задач в таком случае  $I$  можно найти как  $\log_2 N'$ , где  $N'$  — ближайшая к  $N$  степень двойки — такая, что  $N' > N$ .

Например, в английском языке 26 букв. Информационный объем одного символа можно найти так:

$$N = 26; N' = 32; I = \log_2 N' = \log_2(2^5) = 5 \text{ битов.}$$

Если количество символов алфавита равно  $N$ , а количество символов в записи сообщения равно  $M$ , то информационный объем данного сообщения вычисляется по формуле:

$$I = M \cdot \log_2 N.$$

### Примеры решения задач

**Пример 1.** Световое табло состоит из лампочек, каждая из которых может находиться в одном из двух состояний («включено» или «выключено»). Какое наименьшее количество лампочек должно находиться на табло, чтобы с его помощью можно было передать 50 различных сигналов?

*Решение.* С помощью  $n$  лампочек, каждая из которых может находиться в одном из двух состояний, можно закодировать  $2^n$  сигналов.  $2^5 < 50 < 2^6$ , поэтому пяти лампочек недостаточно, а шести хватит.

*Ответ:* 6.

**Пример 2.** Метеорологическая станция ведет наблюдения за влажностью воздуха. Результатом одного измерения является целое число от 0 до 100, которое записывается при помощи минимально возможного количества битов. Станция сделала 80 измерений. Определите информационный объем результатов наблюдений.

*Решение.* В данном случае алфавитом является множество целых чисел от 0 до 100. Всего таких значений 101. Поэтому информационный объем результатов одного измерения  $I = \log_2 101$ . Это значение не будет целочисленным. Заменим число 101 ближайшей к нему степенью двойки, большей 101. Это число  $128 = 2^7$ . Принимаем для одного измерения  $I = \log_2 128 = 7$  битов. Для 80 измерений общий информационный объем равен:

$$80 \cdot 7 = 560 \text{ битов} = 70 \text{ байтов.}$$

*Ответ:* 70 байтов.

### Вероятностный подход

Вероятностный подход к измерению количества информации применяют, когда возможные события имеют различные вероятности реализации. В этом случае количество информации определяют **по формуле Шеннона**:

$$I = - \sum_{i=1}^N p_i \log_2 p_i,$$

где  $I$  — количество информации;

$N$  — количество возможных событий;

$p_i$  — вероятность  $i$ -го события.

Например, пусть при бросании несимметричной четырехгранной пирамидки вероятности отдельных событий будут равны:

$$p_1 = \frac{1}{2}, p_2 = \frac{1}{4}, p_3 = \frac{1}{8}, p_4 = \frac{1}{8}.$$

Тогда количество информации, которое будет получено после реализации одного из них, можно вычислить по формуле Шеннона:

$$I = - \left( \frac{1}{2} \cdot \log_2 \frac{1}{2} + \frac{1}{4} \cdot \log_2 \frac{1}{4} + \frac{1}{8} \cdot \log_2 \frac{1}{8} + \frac{1}{8} \cdot \log_2 \frac{1}{8} \right) = \frac{14}{8} \text{ битов} = 1.75 \text{ бита.}$$

### Единицы измерения количества информации

Наименьшей единицей информации является *бит* (англ. *binary digit (bit)* — двоичная единица информации).

**Бит** — это количество информации, необходимое для однозначного определения одного из двух равновероятных событий. Например, один бит информации получает человек, когда он узнает, опаздывает с прибытием нужный ему поезд или нет, был ночью мороз или нет, присутствует на лекции студент Иванов или нет и т. д.

В информатике принято рассматривать последовательности длиной 8 битов. Такая последовательность называется **байтом**.

**Производные единицы измерения количества информации:**

1 байт = 8 битов

1 килобайт (Кб) = 1024 байта =  $2^{10}$  байтов

1 мегабайт (Мб) = 1024 килобайта =  $2^{20}$  байтов

1 гигабайт (Гб) = 1024 мегабайта =  $2^{30}$  байтов

1 терабайт (Тб) = 1024 гигабайта =  $2^{40}$  байтов

## Процесс передачи информации. Виды и свойства источников и приемников информации. Сигнал, кодирование и декодирование, причины искажения информации при передаче

Информация передается в виде сообщений от некоторого *источника* информации к ее *приемнику* посредством *канала связи* между ними.

В качестве источника информации может выступать живое существо или техническое устройство. Источник посылает передаваемое сообщение, которое кодируется в передаваемый *сигнал*.

**Сигнал** — это материально-энергетическая форма представления информации. Другими словами, **сигнал** — это переносчик информации, один или несколько параметров которого, изменяясь, отображают сообщение. Сигналы могут быть *аналоговыми* (непрерывными) или *дискретными* (импульсными).

Сигнал посылается по каналу связи. В результате в приемнике появляется принимаемый сигнал, который декодируется и становится принимаемым сообщением.

Передача информации по каналам связи часто сопровождается воздействием помех, вызывающих искажение и потерю информации.

### Примеры решения задач

**Пример 1.** Для кодирования букв А, З, Р, О используются двухразрядные двоичные числа 00, 01, 10, 11 соответственно. Этим способом закодировали слово РОЗА и результат записали шестнадцатеричным кодом. Указать полученное число.

*Решение.* Запишем последовательность кодов для каждого символа слова РОЗА: 10 11 01 00. Если рассматривать полученную последовательность как двоичное число, то в шестнадцатеричном коде оно будет равно:  $1011\ 0100_2 = B4_{16}$ .

*Ответ:*  $B4_{16}$ .

## Скорость передачи информации и пропускная способность канала связи

Прием/передача информации может происходить с разной скоростью. Количество информации, передаваемое за единицу времени, есть **скорость передачи информации**, или **скорость информационного потока**.

Скорость выражается в битах в секунду (бит/с) и кратных им Кбит/с и Мбит/с, а также в байтах в секунду (байт/с) и кратных им Кбайт/с и Мбайт/с.

Максимальная скорость передачи информации по каналу связи называется **пропускной способностью канала**.

### Примеры решения задач

**Пример 1.** Скорость передачи данных через ADSL-соединение равна 256000 бит/с. Передача файла через данное соединение заняла 3 мин. Определите размер файла в килобайтах.

*Решение.* Размер файла можно вычислить, если умножить скорость передачи информации на время передачи. Выразим время в секундах: 3 мин =  $3 \cdot 60 = 180$  с. Выразим скорость в килобайтах в секунду:  $256000 \text{ бит/с} = 256000 : 8 : 1024 \text{ Кбайт/с}$ . При вычислении размера файла для упрощения расчетов выделим степени двойки:

Размер файла =  $(256000 : 8 : 1024) \cdot (3 \cdot 60) = (2^8 \cdot 10^3 : 2^3 : 2^{10}) \cdot (3 \cdot 15 \cdot 2^2) = (2^8 \cdot 125 \cdot 2^3 : 2^3 : 2^{10}) \cdot (3 \cdot 15 \cdot 2^2) = 125 \cdot 45 = 5625 \text{ Кбайт}$ .

*Ответ:* 5625 Кбайт.