

# Измерение информации: алфавитный подход

---

**Информация и информационные процессы**

# Как измерить информацию?

Вопрос: «**Как измерить информацию?**»  
очень непростой.

Ответ на него зависит от того, что понимать под информацией. Но поскольку определять информацию можно по-разному, то и **способы измерения** тоже **могут быть разными**.



***ИНФОРМАЦИЯ***

# Алфавитный подход к измерению информации

**Способ измерения информации, который не связывает количество информации с содержанием сообщения, называется алфавитным подходом.**

При алфавитном подходе рассматривают информационное сообщение как **последовательность знаков** определенной знаковой системы.

# Алфавит и его мощность

Все множество используемых в языке символов будем называть **алфавитом**.

Полное количество символов алфавита принято называть **мощностью алфавита**.

Мощность алфавита будем обозначать  $N$ .

Например, мощность алфавита из заглавных русских букв и отмеченных дополнительных символов равна 54.

**АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЭЮЯ0123456789().,!?«»:-;  
(пробел)**

# Сколько информации несет один символ в русском языке

Представьте себе, что текст к вам поступает последовательно, по одному знаку, словно бумажная ленточка, выползающая из телеграфного аппарата. Предположим, что каждый появляющийся на ленте символ с одинаковой вероятностью может быть любым символом алфавита.

В каждой очередной позиции текста может появиться любой из  $N$  символов.

Тогда, согласно известной нам формуле  $2^I = N$ , каждый такой символ несет  $I$  бит информации, которое можно определить из решения уравнения:  $2^I = 54$ .

Получаем:  $I = 5.755$  бит.

Вот сколько информации несет один символ в русском тексте!



ПРИВЕТ! КАК Д

# Количество информации в тексте

А теперь для того, чтобы найти количество информации во всем тексте, нужно посчитать число символов в нем и умножить на  $I$ .

*Посчитаем количество информации на одной странице книги.*



Пусть страница содержит 50 строк. В каждой строке — 60 символов. Значит, на странице уместается  $50 \times 60 = 3000$  знаков. Тогда объем информации будет равен:  $5,755 \times 3000 = 17265$  бит.

***При алфавитном подходе к измерению информации количество информации зависит не от содержания, а от размера текста и мощности алфавита.***

# Задание 1:

- Определите информационный объем страницы книги, если для записи текста использовались только заглавные буквы русского алфавита, кроме буквы Ё.

Решение:

1.  $N = 32$

2.  $2^I = N$

3.  $2^I = 32$

4.  $I = 5$

5. На странице 3000 знаков, тогда объем информации =  $3000 * 5 = 15000$  бит.



# Двоичный алфавит

А что если алфавит состоит только из двух символов 0 и 1?

В этом случае:  $N = 2$ ;  $2^I = N$ ;  $2^I = 2$ ;  $I = 1$ !

При использовании двоичной системы (алфавит состоит из двух знаков: 0 и 1) каждый двоичный знак несет **1 бит** информации.

Интересно, что сама единица измерения информации «бит» получила свое название от английского сочетания «**binary digit**» - «двоичная цифра».



# Достаточный алфавит

Ограничения на максимальный размер алфавита теоретически не существует. Однако есть алфавит, который можно назвать **достаточным**. С ним мы скоро встретимся при работе с компьютером. Это алфавит **мощностью 256 символов**. В алфавит такого размера можно поместить все практически необходимые символы: латинские и русские буквы, цифры, знаки арифметических операций, всевозможные скобки, знаки препинания....

Поскольку  $256 = 2^8$ , то один символ этого алфавита «весит» 8 бит. Причем 8 бит информации — это настолько характерная величина, что ей даже присвоили свое название — байт.

**1 байт = 8 бит**

# Количество информации в тексте

Сегодня очень многие люди для подготовки писем, документов, статей, книг и пр. используют компьютерные текстовые редакторы. *Компьютерные редакторы, в основном, работают с алфавитом размером 256 символов.*

В этом случае легко подсчитать объем информации в тексте. Если 1 символ алфавита несет 1 байт информации, то надо просто сосчитать количество символов; полученное число даст информационный объем текста в байтах.



Пусть небольшая книжка, сделанная с помощью компьютера, содержит 150 страниц; на каждой странице — 40 строк, в каждой строке — 60 символов.

Значит страница содержит  $40 \times 60 = 2400$  байт информации.

Объем всей информации в книге:  $2400 \times 150 = 360\,000$  байт.

# Более крупные единицы информации

Название	Условное обозначение	Соотношение с другими единицами
Килобит	Кбит	1 Кбит = 1024 бит = $2^{10}$ бит $\approx$ 1000 бит
Мегабит	Мбит	1 Мбит = 1024 Кбит = $2^{20}$ бит $\approx$ 1 000 000 бит
Гигабит	Гбит	1 Гбит = 1024 Мбит = $2^{30}$ бит $\approx$ 1 000 000 000 бит
Килобайт	Кбайт (КБ)	1 Кбайт = 1024 байт = $2^{10}$ байт $\approx$ 1000 байт
Мегабайт	Мбайт (МБ)	1 Мбайт = 1024 Кбайт = $2^{20}$ байт $\approx$ 1 000 000 байт
Гигабайт	Гбайт (ГБ)	1 Гбайт = 1024 Мбайт = $2^{30}$ байт $\approx$ 1 000 000 000 байт

# Скорость передачи информации

Прием-передача информации могут происходить с разной скоростью.

Количество информации, передаваемое за единицу времени, есть скорость передачи информации или скорость информационного потока.

Очевидно, эта скорость выражается в таких единицах, как бит в секунду (бит/с), байт в секунду (байт/с), килобайт в секунду (Кбайт/с) и т. д.

# Вопросы:

- Что такое «алфавит»? Что такое «мощность алфавита»?
- Как определяется количество информации в сообщении с алфавитной точки зрения?
- Что больше 1 Кбайт или 1000 байт?
- Расположите единицы измерения информации в порядке возрастания:
  - **Гигабайт; Байт; Мегабайт; Килобайт.**
- Сколько информации содержится в сообщении, если для кодирования одного символа использовать 1 байт:
  - **«Компьютер – универсальный прибор.»**
- Два текста содержат одинаковое количество символов. Первый текст составлен в алфавите мощностью 32 символа, второй – мощностью 64 символа. Во сколько раз отличается количество информации в этих текстах?

## Задание 2:

Племя Мумбу-Юмбу использует алфавит из букв:  $\alpha\beta\gamma\delta\epsilon\zeta\eta\theta\lambda\mu\xi\sigma\psi$ , точки и для разделения слов используется пробел.

Сколько информации несет свод законов племени, если в нем 12 строк и в каждой строке по 20 символов?

## Задание 3:

Вычислите какова мощность алфавита, с помощью которого записано сообщение, содержащее 2048 символов, если его объем составляет 1.25 Кбайта.