



ПОДХОДЫ К  
ИЗМЕРЕНИЮ  
ИНФОРМАЦИИ

# Алфавитный (объёмный) подход к измерению информации

С помощью алфавитного подхода можно определять количество информации в текстовом сообщении, которое состоит из символов определённого алфавита.



Алфавит — это набор символов, которые используются в некотором языке с целью представления информации.

В качестве символов могут быть использованы буквы, цифры, скобки, специальные знаки.



Мощность алфавита — это количество символов в алфавите, которое вычисляется по формуле:

$N = 2^i$ , где  $i$  — информационный вес символа.

# Алфавитный (объёмный) подход к измерению информации

Например, мощность алфавита, состоящего из 26 латинских букв и дополнительных символов (скобки, пробел, знаки препинания (11 шт.), 10 цифр), — 47.

Информационный вес 1 символа, $i$ (бит)	Мощность алфавита, $N$
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024

# Алфавитный (объёмный) подход к измерению информации

*Пример:*

*определи информационный объём сообщения (в Кбайтах), состоящего из 8192 символов, если используется алфавит из 256 символов.*

*Составим краткую запись условия задачи.*

*Дано:*

$$N = 256;$$

$$K = 8192.$$

*Найти:  $I$ .*

*Решение:*

*1. определим, какое количество бит необходимо для кодировки одного символа. Так как мощность используемого алфавита  $N = 256$ , то  $i = 8$  (использовали формулу  $N = 2^i$ ).*

*2. Вычислим информационный объём сообщения, которое состоит из 8192 символов, используя формулу  $I = K \cdot i$ :*

$$I = 8192 \cdot 8 = 65536 \text{ бит.}$$

*3. Переведём биты в Кбайты.*

*Поскольку 1 байт = 8 бит, 1 Кбайт = 1024 байт, получим:*

$$65536 \text{ бит} = \frac{65536}{8} \text{ байт} = 8192 \text{ байт} = \frac{8192}{1024} \text{ Кбайт} = 8 \text{ Кбайт.}$$

*Ответ: 8 Кбайт.*

# Алфавитный (объёмный) подход к измерению информации

## Единицы измерения информации

Любая компьютерная техника работает в двоичном коде, понимая только значения **0** — «сигнал есть» и **1** — «сигнала нет». Эти значения хранятся в бите — наименьшей единице измерения информации. Однако удобнее использовать более крупные единицы измерения информации, которые приведены в таблице.

<b>1</b> байт	8 бит = $2^3$ бит
<b>1</b> Кбайт (килобайт)	$2^{10}$ байт
<b>1</b> Мбайт (мегабайт)	$2^{10}$ Кбайт
<b>1</b> Гбайт (гигабайт)	$2^{10}$ Мбайт
<b>1</b> Тбайт (терабайт)	$2^{10}$ Гбайт

# Алфавитный (объёмный) подход к измерению информации

Пример:

1) определить, сколько Мбайт информации содержится в 512 битах. Ответ дай в виде степени числа 2.

Будем работать со степенями, учитывая, что  $512 = 2^9$ .

$$2^9 : 2^3 : 2^{10} : 2^{10} = 2^{-14} \text{ Мбайт.}$$

Ответ:  $2^{-14}$  Мбайт.

2) Какое количество бит содержится в  $\frac{1}{256}$  Гбайт памяти? Ответ дай в виде степени числа 2.

$$\frac{1}{256} \text{ Гбайт} = 2^{-8} \text{ Гбайт.}$$

$$2^{-8} \cdot 2^{10} \cdot 2^{10} \cdot 2^{10} \cdot 2^3 = 2^{25} \text{ бит.}$$

Ответ:  $2^{25}$  бит.

# Содержательный подход к измерению информации



**Информация** — это знания, которые получает человек.

С позиции содержательного подхода сообщение, которое информирует об исходе некоторого события, снимает неопределённость знания человека об этом событии. Сообщение несёт больше информации, в случае если изначально была большая неопределённость знания.



**Неопределённость знания** — это количество возможных вариаций ответа на поставленный вопрос.

Подбрасывая монету, мы не знаем, что выпадет: «орёл» или «решка» — это равновероятные события. После броска нам известен исход события — имеем полную определённость. Неопределённость знания уменьшается в **2** раза.



**Информация** — это снятая неопределённость знания человека об исходе некоторого события (Клод Шеннон).

Сообщение, которое уменьшает неопределённость знания в **2** раза, несёт **1 бит** информации.

# Содержательный подход к измерению информации



«Главная формула» информатики:

$2^i = N$ , где  $N$  — неопределённость знания (количество возможных исходов какого-то события);  $i$  — количество информации в сообщении об одном из  $N$  исходов.

Следовательно, для нахождения количества информации  $i$ , которое содержится в сообщении об одном из  $N$  равновероятных исходов какого-то события, нужно воспользоваться формулой:  $i = \log_2 N$ .



# Содержательный подход к измерению информации

*Пример:*

1) в доме **16** этажей, на каждом этаже по **4** квартиры. Какое количество информации несёт сообщение о том, что Игорь живёт на **7**-м этаже в квартире № **27**?

*Решение:*

всего в доме  $16 \cdot 4 = 64$  квартиры, т. е.  $N = 64$ .

Используем формулу  $i = \log_2 N$ :

$$i = \log_2 64 = 6 \text{ бит.}$$

*Ответ:* **6** бит.

2) Загадали некоторое целое число в определённом диапазоне. Угадывая это число, получили **7** бит информации. Сколько чисел содержит диапазон?

*Решение:*

$$i = 7, 2^i = N, \text{ значит, } N = 2^7 = 128.$$

*Ответ:* **128** чисел.

# Вероятностный подход к измерению информации

В жизни чаще всего случаются события, частота появления которых различна. Речь идёт о неравновероятных событиях. Например, летом чаще солнечно и ясно, чем пасмурно.



**Вероятность некоторого события** — это величина  $p$  ( $0 \leq p \leq 1$ ), которая показывает частоту появления этого события в ряде однотипных испытаний.

Если событие невозможно, то его вероятность:  $p = 0$ .

Если событие достоверно, то его вероятность:  $p = 1$ .

Если событие может наступить или не наступить с одинаковой вероятностью, то его вероятность:  $p = 0,5$ .

# Вероятностный подход к измерению информации

## Формула Шеннона

Американский учёный Клод Шеннон был одним из основоположников теории информации и криптографии. Им была выведена в 1948 году формула для вычисления количества информации равновероятных событий:

$$I = - \sum_{i=1}^N p_i \cdot \log_2 p_i$$
, где  $i$  — количество информации;  $N$  — количество возможных событий;  $p_i$  — вероятность  $i$ -го события.

# Вероятностный подход к измерению информации

*Пример:*

*пусть вероятность выпадения осадков в виде дождя равна 0,5, ветра — 0,25, грозы и молнии — 0,125. Определим, какое количество информации получим при реализации одного из них.*

*Решение.*

*Согласно формуле Шеннона получим:*

$$\begin{aligned} I &= -(0,5 \cdot \log_2 0,5 + 0,25 \cdot \log_2 0,25 + 0,125 \cdot \log_2 0,125 + 0,125 \cdot \log_2 0,125) = \\ &= -(-0,5 - 0,5 - 0,375 - 0,375) = 1,75. \end{aligned}$$

*Ответ: 1,75 бит.*