

Информация и её кодирование. Способы измерения информации.



Информатика – это наука об информации, способах ее представления, обработки и передачи.

Информация – это сведения, уменьшающие неопределённость нашего знания об окружающем нас мире.

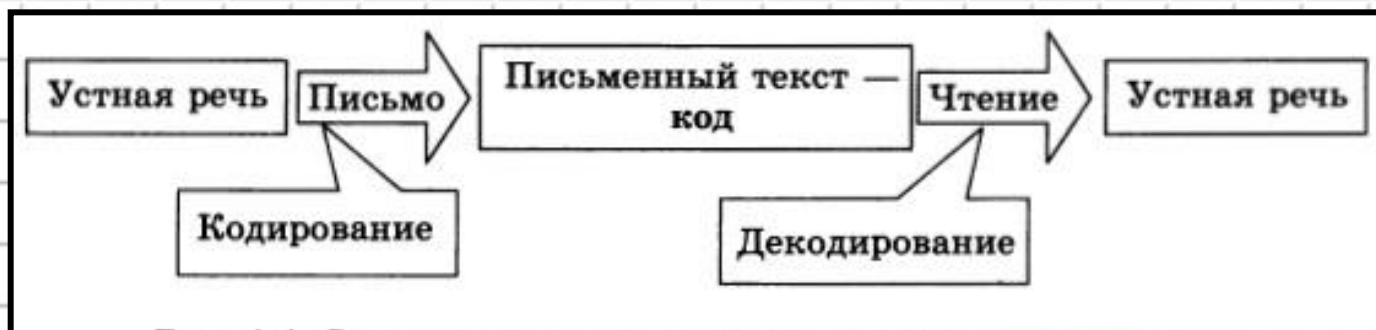
Язык – это знаковая форма представления информации.



Код – система условных знаков для представления информации.



Кодирование – это операция преобразования символов или группы символов одного кода в символы или группы символов другого кода.



Декодирование – процесс, обратный кодированию.



Длина кода – количество знаков в коде.

Текст:

- в России: *Привет, Вася!*
- Windows-1251:
CFF0E8E2E52C20C2E0F1FF21
- передача за рубеж (транслит): *Privet, Vasya!*
- стенография:
- шифрование: *Рсйгжу-!Гбта”*

Числа:

- для вычислений: *24*
- прописью: *двадцать четыре*
- римская система: *XXIV*



Код

Равномерный

Неравномерный

Код Бодо

Азбука Морзе

А	● ■	П	● ■ ■ ■ ●	Ь	■ ● ● ■
Б	■ ● ● ●	Р	● ■ ●	Ы	■ ● ■ ■
В	● ■ ■ ■	С	● ● ●	Й	● ■ ■ ■
Г	■ ■ ■ ●	Т	■	1	● ■ ■ ■ ■ ■
Д	■ ● ●	У	● ● ■	2	● ● ■ ■ ■ ■
Е	●	Ф	● ● ■ ■ ●	3	● ● ● ■ ■ ■
Ж	● ● ● ■	Х	● ● ● ●	4	● ● ● ● ■
З	■ ■ ■ ● ●	Ц	■ ● ■ ■ ●	5	● ● ● ● ●
И	● ●	Ч	■ ■ ■ ■ ■ ●	6	■ ● ● ● ●
К	■ ■ ● ■	Ш	■ ■ ■ ■ ■ ■	7	■ ■ ■ ● ● ●
Л	● ■ ● ●	Щ	■ ■ ■ ● ■	8	■ ■ ■ ■ ● ●
М	■ ■ ■	Э	● ● ■ ■ ● ●	9	■ ■ ■ ■ ■ ●
Н	■ ■ ●	Ю	● ● ■ ■ ■	0	■ ■ ■ ■ ■ ■
О	■ ■ ■ ■	Я	● ■ ● ■ ■		



Способы измерения информации

Содержательный подход

Алфавитный подход

Вероятностный подход



Единицы измерения информации

В двоичном коде один двоичный разряд несет одну единицу информации, которая называется 1 бит.

БИТ- это наименьшая единица измерения информации

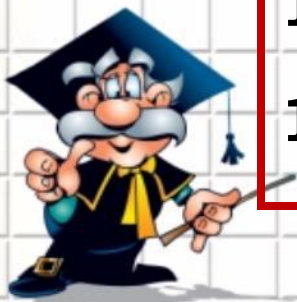
1 байт = 8 бит

1 Килобайт (Кбайт) = 1024 байт = 2^{10} байт

1 Мегабайт (Мбайт) = 1024 Кбайт = 2^{20} байт

1 Гигабайт (Гбайт) = 1024 Мбайт = 2^{30} байт

1 Терабайт (Тбайт) = 1024 Гбайт = 2^{40} байт



Алфавитный подход

$$N = 2^i$$

где, **i** - информационный вес одного символа

N - мощность алфавита

*Для записи сообщения использовали 64-символьный алфавит.
Каков информационный вес одного символа?*

$$64 = 2^i \longrightarrow 2^6 = 2^i \longrightarrow i = \mathbf{6 \text{ бит}}$$

*Сообщение, записанное буквами 64-символьного алфавита,
содержит 20 символов. Какой объем информации оно несет?*

$$I = i * K$$

I – информационный вес всего сообщения

K – количество символов в тексте

$$I = 6 \text{ бит} * 20 = \mathbf{120 \text{ бит}}$$



Содержательный подход

Сообщение, уменьшающее неопределенность знания в два раза, несет 1 бит информации.

$$N = 2^i$$

где, i - количество информации в сообщении об одном из N результатов.

N - количество возможных результатов события (неопределенность знания).

В мешке лежат 16 красных яблок. Сколько информации содержит сообщение, что достали красное яблоко?

$$16 = 2^i \longrightarrow 2^6 = 2^i \longrightarrow i = \mathbf{4 \text{ бита}}$$



Содержательный подход

Задача 1.

Какое количество информации содержится в неинформационном сообщении?

Дано:

$$\underline{N=0}$$

i -?

Решение:

$$N=0 \Rightarrow 2^i=0 \Rightarrow i=\text{«пустое множество»}$$

Ответ: «пустое множество»

Задача 2.

Найти количество информации в однозначном сообщении.

Дано:

$$\underline{N=1}$$

i -?

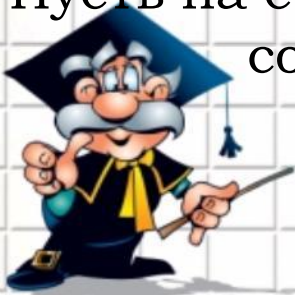
Решение:

$$N=1 \Rightarrow 2^i=1 \Rightarrow i=0 \text{ бит}$$

Ответ: 0 бит

Задача 3.

Пусть на стеллаже 16 полок. Сколько информации содержится в сообщении о месте нахождения книги?



Дано:

$$\underline{N=16}$$

i -?

Решение:

$$N=16 \Rightarrow 2^i=16 \Rightarrow 16=2^4 \Rightarrow i=4 \text{ бита}$$

Ответ: 4 бита

Задача 4.

Какое количество информации потребуется для кодирования одного шахматного поля?

Дано:

$$N=8*8$$

$$i=?$$

Решение:

$$N=8*8=64 \Rightarrow 2^i=64 \Rightarrow i=6 \text{ бит}$$

Ответ: 6 бит

Задача 5.

Получено сообщение, объемом 10 бит. Какое количество сообщений возможно составить из полученных данных?

Дано:

$$i=10 \text{ бит}$$

$$N=?$$

Решение:

$$i=10 \Rightarrow 2^{10}=1024 \Rightarrow N=1024 \text{ сообщения}$$

Ответ: 1024 сообщения

Задача 6.

Какое количество слов получится из фразы в 8 бит?

Дано:

$$i=8 \text{ бит}$$

$$N=?$$

Решение:

$$i=8 \Rightarrow 2^8=256 \Rightarrow N=256 \text{ слов}$$

Ответ: 256 слов



Задача 7.

Сообщение о том, что ваш друг живет на 6 этаже несет 4 бита информации. Сколько этажей в доме?

Дано:

$$\frac{i=4 \text{ бита}}{N=?}$$

Решение:

$$i=4 \Rightarrow 2^4=16 \Rightarrow N=16 \text{ этажей}$$

Ответ: 16 этажей

Задача 8.

В школьной библиотеке 16 стеллажей с книгами. На каждом стеллаже 8 полок. Библиотекарь сообщил Пете, что нужная ему книга находится на пятом стеллаже на третьей сверху полке. Какое количество информации библиотекарь передал Пете?

Дано:

$$\frac{N=16*8 \text{ бит}}{i=?}$$

Решение:

$$N=16*8 \text{ вариантов местонахождения книги}$$

$$\Rightarrow N=128 \Rightarrow 2^i=128 \Rightarrow i=7 \text{ бит}$$

Ответ: 7 бит





Алфавитный подход

Задача 1.

Книга, набранная на компьютере, имеет 150 страниц, на странице 40 строк, в каждой строке 60 символов. Каков полный объем информации?

Дано:

$N=256$

I-?

Решение:

Мощность компьютерного алфавита равна

$$2^i=256 \Rightarrow 2^8=256 \Rightarrow i=8 \text{ бит} \Rightarrow i = 1 \text{ байт}$$

Объем всей информации: $150 * 40 * 60 = 360\ 000$ байт

Ответ: 360 000 байт



Задача 2.

Объем сообщения, содержащего 2048 символов, составил 1/512 часть Мбайта. Определить мощность алфавита.

Дано:

$$I = 1/512 \text{ Мбайта}$$
$$k = 2048$$

N-?

Решение:

Переведем в биты информационный объем сообщения.

$$I = 1/512 * 1024 * 1024 * 8 = 16384 \text{ бит}$$

$$i = I / K = 16384 / 1024 = 16 \text{ бит} - \text{приходится на}$$

один символ алфавита.

$2^{16} = 65536$ СИМВОЛОВ - МОЩНОСТЬ
ИСПОЛЬЗОВАННОГО АЛФАВИТА.

Именно такой алфавит используется в кодировке Unicode

Ответ: 65536 СИМВОЛОВ



Задача 3.

Два текста содержат одинаковое количество символов. Первый текст записан на русском языке, а второй на языке племени нагури, алфавит которого состоит из 16 символов. Чей текст несет большее количество информации?

Дано:

$$N_1 = 32$$

$$N_2 = 16$$

$$k_1 = k_2$$

I-?

Решение:

$I = i * k$ т.к. оба текста имеют одинаковое число символов (k), то разница зависит от информативности одного символа алфавита (i).

$$N_1 = 2^{i_1} \Rightarrow 2^{i_1} = 32 \Rightarrow i = 5 \text{ бит}$$

$$N_2 = 2^{i_2} \Rightarrow 2^{i_2} = 16 \Rightarrow i = 4 \text{ бита}$$

$$I_1 = 5 \text{ бит} * k$$

$$I_2 = 4 \text{ бита} * k$$

$$I = I_1 / I_2$$

Текст, записанный на русском языке в 5/4 раза несет больше информации

Ответ: Текст на русском языке

