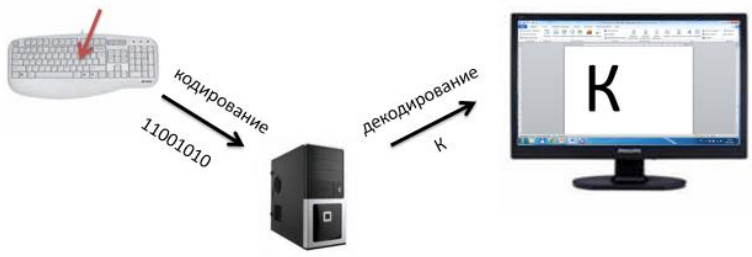
ИНФОРМАЦИЯ. ЕДИНИЦЫ ИЗМЕРЕНИЯ ИНФОРМАЦИИ.

Символ в компьютере – это любая буква, цифра, знак препинания, математический знак, специальный символ. В общем, все, что можно ввести с клавиатуры.

Но компьютер «не понимает» человеческий язык. Поэтому каждый символ кодируется.

ПК «понимает» только нули и единички – с помощью них и представляется информация в компьютере. **Эти «нули и единички» называются битом. Бит может принимать одно из двух значений – 0 или 1.**

При вводе в компьютер текстовой информации происходит её двоичное кодирование. А при выводе на экран декодирование. Для каждого символа, вводимого с клавиатуры, выделен свой код.



Эти коды хранятся в виде двоичных чисел одинаковой длины.

Для того чтобы однозначно закодировать 1 символ компьютерной клавиатуры из 256 символов нужно 8 таких ячеек. Значит информационный вес **1 символа = 8 бит** Величина равная 8 битам называется байтом. **8 бит = 1 байт** информационный вес **1 символа - 1 байт**

Давайте попробуем посчитать объем информации в словах. Сколько байт в слове «МАМА»? Правильно, 4 байта. Ведь 1 буква – это 1 символ, а 1 символ – 1 байт. А сколько байт в предложении «МАМА И ПАПА!»? Правильно, 12 байт. Пробел, т.е. разделитель между словами, тоже символ, только пустой или белый (от этого и название – пробел).

Широко используются также еще более крупные производные единицы информации:

Единицы измерения информации:
1байт = 8 бит
1Кб (килобайт) = 2^{10} байт = 1024 байт
1Мб (мегабайт) = 2^{10} Кб = 1024 Кб
1Гб (гигабайт) = 2^{10} Мб = 1024 Мб

1 Терабайт (Тбайт) = 1024 Гбайт,
1 Петабайт (Пбайт) = 1024 Тбайт.

В единицах измерения информации измеряют как саму информацию, так и носители информации. Например емкость CD диска - 700 Мб, DVD диска - 4,5ГБ, флешки бывают емкостью 1ГБ, 2 ГБ, 4 Гб, 8 Гб, 16 Гб, 32 Гб, 64 Гб, 128 Гб. жесткие диски сейчас выпускаются емкостью 500 гб, 1 Тб, 2 Тб,4 ТБ, 8 Тб

Давайте сравним их с объемом информации которую мы на них записываем
1 фотография среднего качества "весит" около 3,5 Мб
1 фильм с невысоким разрешением "весит" - около 700-800 Мб, более высокого качества -1,2 -1,4 Гб, высокого качества

Пример 1 Сколько фотографий размером 3,5 Мб можно разместить на CD и на DVD диске
700 / 3,5 = 200 шт
4,5 Гб =4,5*1024 = 4608 Мб 4608 / 3,5 ≈ 1316 шт

Пример 2 Книга содержит 100 страниц; на каждой странице -- 35 строк, в каждой строке -- 50 символов. Рассчитаем объем информации, содержащийся в книге.

Страница содержит 35 х 50 = 1750 байт информации. Объем всей информации в книге (в разных единицах): 1750 х 100 = 175000 байт. 175000 / 1024 = 170,8984 Кб 170,8984 / 1024 = 0,166893 Мб

Пример3 Определить, сколько времени будет передавать информацию страницы текста из 40 строк по 80 символов в строке модем, работающий со скоростью 1200 бит/сек.

объем сообщения : 40*80 =3200 байт 3200*8=25600 бит время передачи: 25600/1200=21,3 сек

Описанная выше кодировка 1 байтовая (или 8 битная) - это означает, что в этой кодировке для хранения 1 символа требуется 1 байт. Но бывают и другие вариант кодировок, например двухбайтовые (16 битные) - на 1 символ отводится 2 байта.

Задание 1.
В одном из вариантов кодировки Unicode на каждый символ отводится два байта. Определите информационный объём сообщения из двадцати символов в этой кодировке. Ответ выразите в битах.

1 символ - 2 байта 20 символов: 20х2=40 байт 40*8=320 бит

Задание 2.
В кодировке КОИ-8 каждый символ кодируется одним байтом. Определите количество символов в сообщении, если информационный объём сообщения в этой кодировке равен 160 бит.

1 символ - 1 байт = 8 бит 160:8=20 символов

А может быть вопрос задан и обратным образом, например так:

Задание 3.
Информационное сообщение объемом 4 Кбайта состоит из 4096 символов. Каков информационный вес символа используемого алфавита?

объем сообщения: 4Х1024=4096 байт
4096:4096 =1 байт - информационный вес 1 символа, те 1 символ кодируется 1 байтом

Задание 4. Информационное сообщение объемом 720 битов состоит из 180 символов. Каким количеством бит кодируется каждый символ этого сообщения?

720:180 = 4 бита используются для кодировки одного символа

Почему такая разница в весе одного символа в различных кодировках? Это зависит от мощности используемого алфавита - те от количества в нем символов

Например, если в алфавите 2 символа то для его кодировки 1 символа достаточно 1 бит 1 0
если в алфавите 4 символа, то для кодирования каждого из них требуется 2 бита
так как в двухбитном коде кол-во возможных кодовых комбинаций 4 : 00 01 10 11
если в алфавите 8 символов, то для кодирования каждого из них требуется 3 бита

так как в трехбитном коде кол-во возможных кодовых комбинаций 8:
000 001 011 111 110 100 010 101
обратите внимание, что $4=2^2$, $8=2^3$, и тд

это означает, что если **мощность алфавита (кол-во кодируемых символов) $N=2^i$**, то для однозначного кодирования каждого из **N** символов требуется **i** бит

давайте ответим на вопрос, какова мощность алфавита в задании 4

1 символ = 4 бита
кол-во символов алфавита $N= 2^4=2x2x2x2=16$ символа

Часто для кодирования символов используются таблицы КОИ-8, Windows-1251, ISO, MAC, состоящие из 256 различных символов.

Таким образом, $N=2^i$, $256=2^i$, $2^8=2^i$, i=8 бит

Сейчас в качестве международной принята кодовая таблица Unicod, которая позволяет закодировать 65536 различных символов. Таким образом, N=2i, 65536 =2i, 216=2i, i=16 бит

Информационный вес одного символа:

В кодировках КОИ-8, Windows-1251, ISO, DOS, MAC ***i= 8 бит= 1 байт***

В кодировке Unicode ***i= 16 бит= 2 байт***

Информационный объём текста:

$I = K \cdot i$,

где *I*- количество информации в тексте,

K- количество символов в тексте,

i – информационный вес одного символа.

Задача 5. Информационное сообщение объемом 650 битов состоит из 130 символов. Какова мощность алфавита, с помощью которого это сообщение записано?

650:130=5 бит - 1 символ
$2^5=2x2x2x2x2=32$ символа - мощность алфавита

Это означает, что если мы используем 1 байтный код (8 битный), то мы можем закодировать $2^8=256$ символов, а если 2 байтный (16 битный), то $2^{16}=65536$ символов

То что мы с вами рассмотрели, называется **алфавитным подходом** к кодированию информации.

Задача 6. Автоматическое устройство осуществило перекодировку информационного сообщения на русском языке длиной в 25 символов, первоначально записанного в 16-и битной кодировке Unicode в 8-и битную кодировку КОИ-8. На сколько бит уменьшился информационный объём сообщения?
Решение:
Длина сообщения 25 символов
в 16-и битной кодировке Unicode I1=16*25= 400 бит
в 8-и битной кодировке КОИ-8 I2=8*25= 200 бит
I1-I2=400-200=200 бит
Ответ: на 200 бит

Задача 7. Автоматическое устройство осуществило перекодировку информационного сообщения на русском языке, первоначально записанного в 8-и битной кодировке КОИ-8 в 2-байтный код Unicode. При этом информационное сообщение увеличилось на 320 бит. Какова длина сообщения в символах?

Решение: Обозначим за x количество символов в сообщении.
Количество бит, которое было первоначально I1=8x бит
Количество бит, после перекодировки I2=2x байт=16x бит
Составим уравнение 16x-8x= 320. 8x=320; x=40 Ответ: 40 символов