Алфавитный подход к определению количества информации.

Кодирование текстовой информации.

Термины:

Знаковая система - это набор знаков определенного типа (алфавита) и правил выполнения операций над знаками.

Алфавит - это набор всех допустимых символов знаковой системы.

Термины:

$$N=2^i$$

Где **N** – **мощность алфавита** (количество знаков в алфавите);

i – **количество информации**, которое несет каждый символ алфавита.

Пример:

Русский алфавит

N = 32

$$32 = 2^i$$

i = 5 БИТ

Ответ: 5 бит информации несет 1 русская буква.

Пример:

Двоичный код

0 и 1

$$N = 2$$

$$2 = 2^i$$

$$i = 1$$

Ответ: 1 бит информации несет 1 символ.

Термины:

$$I=i*k$$

Где

I – количество информации в сообщении;

k – количество символов в сообщении;

i – количество информации, каждого символа.

Задача 1:

Какое количество информации передается в простом арифметическом выражении. Алфавит состоит из 16 знаков.

Решение:

$$N=16 \rightarrow 16=2^i \rightarrow i=4$$

Ответ: 4 бит

Задача 2:

Подсчитать количество информации в тексте, если текст состоит из 800 символов, а мощность используемого алфавита 128 символов.

Решение:

$$N=128 \rightarrow 128=2^i \rightarrow i=7$$

$$I=7*800=5600$$

Ответ: 5600 бит

Термины:

Текстовая информация

256 символов

N=256 **$256=2^i$** **I = 8**

Ответ: 8 бит

Термины:

Представление текстовой информации в компьютере:

От 00000000 до 11111111

Кодовая таблица – это соответствие символов и их кодов.

Пример 1:

Таблица 3.1. **Кодировки знаков**

Двоичный код	Десятичный код	КОИ-8	Windows	MS-DOS	Mac	ISO
00000000	0					
...						
00001000	8	удаление последнего символа (клавиша {Backspace})				
...						
00001101	13	перевод строки (клавиша {Enter})				
...						
00100000	32	клавиша {Пробел}				
00100001	33	!				
...						
01011010	90	Z				
...						
01111111	127	▯				
10000000	128	—	Ъ	А	А	к
...						
11000010	194	б	В	—	—	Т
...						
11001100	204	л	М	╬	╬	ь
...						
11011101	221	щ	Э	_	Ё	н
...						
11111111	255	ь	я	нераз. пробел	нераз. пробел	п

Пример 2:

Таблица 3.2. **Десятичные коды некоторых символов в различных кодировках**

Символ	Windows	MS-DOS	КОИ-8	Mac	ISO	Unicode
А	192	128	225	128	176	1040
В	194	130	247	130	178	1042
М	204	140	237	140	188	1052
Э	221	157	252	157	205	1069
я	255	239	241	223	239	1103

Unicode

$$N=2^i$$

$$2^i=2^{16}$$

65536 символов

Давайте вспомним!