

Задание 1 ТМЯ

Кодирование текстовой информации.

Алфавит - набор букв, цифр, знаков препинания используемых в тексте.

Задание 1 ТМЯ

Мощность алфавита - количество символов алфавита.

Например - как Вам известно, в русском алфавите 33 буквы, т.е. его мощность равна 33. Используются там буквы от А до Я, из них строятся все слова в русском языке. Двоичный алфавит, в нем есть только 0 и 1. Его мощность равна 2. Почему мы говорим о двоичном алфавите, потому что в компьютере информация как-раз и хранится в виде слов из этого алфавита, таких как 101, 1101 и т.д.

При алфавитном подходе к измерению информации считается, что каждый символ в алфавите имеет свой вес (т.е. занимает определенное место в памяти, выраженное в битах).

Обозначим мощность алфавита буквой N , и введем формулу $N = 2^b$, где b - вес одного символа в алфавите. Например: двоичный алфавит, мощность - 2, значит $2 = 2^b$, а $2 = 2$, только в первой степени, значит $b=1$. Один символ двоичного алфавита занимает в памяти ровно 1 бит.

Представим, что используется алфавит, где 16 символов, т.е $N=16$. Сколько будет весить один символ такого алфавита? (ответ 4 бита, т.к. $16 = 2^b$, 2 в какой степени равно 16?).

Для того, чтобы легко ориентироваться в таких задачах, полезно знать степени двойки.

Степень 2	Значение
0.	1
1.	2
2.	4
3.	8
4.	16
5.	32
6.	64
7.	128

8.	256
9.	512
10.	1024
11.	2048

Что делать в том случае, когда мощность алфавита не является степенью 2? Например, в алфавите 40 символов. Сколько весит 1 символ? В данном случае поступим следующим образом рассмотрим степени двойки. 32 - этого будет мало, у нас 40. Возьмем следующую - 64 - это гораздо больше чем 40, но для того чтобы "взвесить" символ и получить целое число бит, в самый раз. Таким образом, ограничиваем нашу мощность с двух сторон степенями двойки, и выбираем большую из них $32 < 40 < 64$, берем 64, а это 2 в 6 степени, поэтому ответ 6 бит.

Часто бывает, что нужно решить задачу наоборот, т.е. зная, сколько весит 1 символ, посчитать вес всего текста. Например - 1 символ алфавита весит 5 бит, сколько символов в алфавите? Подставим это в формулу. У нас есть вес одного символа, т.е. $b = 5$ бит $N = 2^5 = 32$. Если 1 символ весит 5 бит, то в алфавите 32 символа.

Ну, или такой пример - 1 символ кодируется 4 битами, определите размер следующего сообщения: **Привет, как дела?** В данном сообщении 17 символов (включая пробел и знаки препинания). 1 символ 4 бита, значит 17 символов = $17 * 4 = 64$ бита.

Если сообщение большое, то и размер его тоже большой, для удобства биты группируются в байты, килобайты и мегабайты следующим образом:

1 байт	8 бит	2^3 бит	
1 Килобайт	1024 байта	2^{10} байт	2^{13} бит
1 Мегабайт	1024 Килобайта	2^{10} Килобайт	2^{20} байт = 2^{23} бит

Таким образом 64 бита = $64/8$ байт = 8 байт

Или так $64 = 2^6$, $8 = 2^3$, $\frac{2^6}{2^3} = 2^3 = 8$ байт

Теперь разберем задачу из демо-версии ГИА 2016

1 В одной из кодировок Unicode каждый символ кодируется 16 битами. Определите размер следующего предложения в данной кодировке. Я к вам пишу – чего же боле? Что я могу ещё сказать?

- 1) 52 байт
- 2) 832 бит
- 3) 416 байт
- 4) 104 бит

Ответ:

По условию, каждый символ кодируется 16 битами, а 16 бит - это 2 байта. 1 символ - 2 байта. Посчитаем количество символов - 52 символа. $52 \cdot 2 = 104$ байта. У нас не указано в каких единицах измерения должен быть ответ, но ответа 104 байта у нас нет. Есть 104 бита, чтобы запутать нас, но нас не запутать. Мы переведем 104 байта в биты, для этого $104 \cdot 8 = 832$ бита. Такой ответ есть, поэтому ответ 2.

Теперь подсчитаем объем книги. В книге 30 страниц, на каждой странице по 32 строки, в каждой строке по 64 символа. Каждый символ кодируется 8 битами. Определите размер книги в Килобайтах. Если решать на прямую, то нужно посчитать количество символов в книге, для этого $32 \cdot 64 \cdot 30 = 61440$ символов в книге. Каждый кодируется 8 битами, или 1 байтом. Книга содержит 61440 байт информации, но нам нужно знать сколько здесь килобайт. Для этого байты переведем в килобайты $61440 / 1024 = 60$ Килобайт. Трудно в уме сразу рассчитывать такие числа, проще работать со степенями. $32 = 2^5$, $64 = 2^6$, $30 = 2 \cdot 15$

$$\frac{2^5 \cdot 2^6 \cdot 2^1 \cdot 15}{2^{10}} = \frac{2^{12} \cdot 15}{2^{10}} = 2^2 \cdot 15 = 4 \cdot 15 = 60$$