



Измерение объёма текстовой информации

АЛФАВИТ

Алфавит – это совокупность символов используемых в языках для представления информации.

Мощность алфавита – число символов в нём.



1 1 1 0 1 0 0 1 0 1

Двоичный алфавит – алфавит из двух символов.

Величину, способную принимать лишь два различных значения (0 и 1), называют бит.

Бит – минимальная единица количества информации.

1 байт = 8 бит

ИНФОРМАЦИОННЫЙ ВЕС

СИМВОЛА



Зависит
от
МОЩНОСТИ алфавита

$$2^i = N$$

ГДЕ:

i – информационный вес символа
(Измеряется в БИТАХ)

N – мощность алфавита

Например:

Мощность алфавита $N=16$ символов → $i=4$ бит информации

Мощность алфавита $N=32$ символов → $i=5$ бит информации

Мощность алфавита $N=64$ символов → $i=6$ бит информации

Мощность алфавита $N=128$ символов → $i=7$ бит информации

Мощность алфавита $N=256$ символов → $i=8$ бит информации

Информационный вес символа



АЛФАВИТНЫЙ ПОДХОД К ИЗМЕРЕНИЮ ИНФОРМАЦИИ

При выяснения информационного объёма текста содержание текста не учитывается.



$$\begin{array}{ccc} \mathbf{X} & & \mathbf{K} \\ \text{Информационный} & = & \text{Количество} \\ \text{объём} & & \text{символов в} \\ \text{текста} & & \text{тексте} \\ & & \mathbf{X} \\ & & \text{Информационный} \\ & & \text{вес} \\ & & \text{символов} \end{array}$$

РЕШИМ ЗАДАЧУ:

Первобытное племя отправило сообщение для всех соплеменников:

МИРУ – МИР!!



Выяснить какой объём сообщения в байтах, известно, что сообщение создано с помощью алфавита в котором 16 символов.

АЛФАВИТНЫЙ ПОДХОД К ИЗМЕРЕНИЮ ИНФОРМАЦИИ

РЕШЕНИЕ ЗАДАЧИ:

1. Посчитаем количество символов в сообщении:

МИРУ – МИР!!

- 12 символов (включая пробелы)

2. Выясним информационный вес символа.

Мощность алфавита $N=16$ символов \rightarrow информационный вес символа $i=4$ бит информации.

3. $X=K \cdot i \rightarrow 12 \cdot 4=48$ бит, $48:8=6$ байт

(в 1байт=8бит)

Ответ: объём сообщения = 6 байт.

Решаем задачи

Решите задачу.

Страница детской книжки содержит 20 строк, в каждой строке 60 символов.

Книжка содержит 10 страниц.

Для набора текста в этой книжке использовали алфавит мощностью 256 символов.

Выяснить информационный объём книжки в байтах, в Кбайтах.

Информация для проверки

Ответ: 12000 байт, 11.71Кб

Решаем задачи

Решите задачу.

Сообщение располагается на двух страницах и занимает объём 1/16 Кбайта информации.

На каждой странице записано 256 символов, какова мощность использованного алфавита?

Информация для проверки

Решение

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

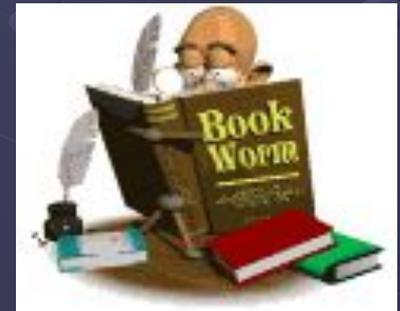
1024:16=64 байт на двух страницах, 32 байта – на одной странице,

32*8=256 бит – на одной странице.

256 бит: 256 символов = 1 бит на 1 символ.

$2^i=N$; $2^1=2$;

Ответ: 2 символа мощность алфавита.



Решаем задачи

Решите задачи самостоятельно в интерактивном режиме
Щелкните указателем мыши по картинке – перейдёте в интерактивную среду LearningApps

The screenshot shows a web browser window with the URL `learningapps.org/view84403`. The page contains a math problem and several input fields. The problem is: "Найти объем информации сообщения, содержащего 8 символов и записанного на языке с мощностью алфавита 16". The solution options are: 32, 128, 48, 2, 1024, and 256. The correct answer is 256. There are also two additional problem boxes on the page: "Сколько символов содержит сообщение, информационный объем которого составляет 1 Кбайт, а мощность алфавита с помощью которого записано это сообщение 256" and "3 страницы информационного сообщения, каждая страница которого состоит из 15 строк по 10 символов в строке, содержат 450 байт информации. Найдите мощность алфавита с помощью которого записано это сообщение".

learningapps.org/view84403

Перепроверить решение

32

128

48

2

1024

256

Найти объем информации сообщения, содержащего 8 символов и записанного на языке с мощностью алфавита 16

Сколько символов содержит сообщение, информационный объем которого составляет 1 Кбайт, а мощность алфавита с помощью которого записано это сообщение 256

3 страницы информационного сообщения, каждая страница которого состоит из 15 строк по 10 символов в строке, содержат 450 байт информации. Найдите мощность алфавита с помощью которого записано это сообщение