

Алфавитный
подход к
определению
информации



Алфавитный подход является объективным, т.е. он не зависит от субъекта (человека), воспринимающего текст.

При определении количества информации на основе уменьшения неопределенности наших знаний мы рассматриваем информацию с точки зрения содержания, ее понятности и новизны для человека.

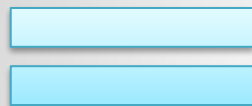


Набор символов знаковой системы (алфавит) можно рассматривать как различные возможные состояния (события).

Количество информации

Количество информации

И, которое содержит сообщение, закодированное с помощью знаковой системы



Количество информации

И, которое несет один знак



Количество знаков



Правило для измерения информации с точки зрения алфавитного подхода

1. Найти мощность алфавита – \underline{N}
2. Найти информационный объем одного символа – $\underline{i = \log_2 N}$
3. Найти количество символов в сообщении – \underline{K}
4. Найти информационный объем всего сообщения – $\underline{K * i}$



Пример №1

Найти объем информации, содержащейся в тексте из 3000 символов, написанном русскими буквами (33 буквы +21 специальных символов).

Решение:

1) Найдем мощность алфавита:

N=33 русских прописных буквы + **33** русских строчных буквы + **21** специальный знак = **87** СИМВОЛОВ.

2) Найдем информационный объем одного символа:

$i = \log_2 87 = 6,4$ бит.

3) Найдем объем всего текста:

$6,4 \text{ бит} * 3000 = 19140 \text{ бит} = 2392,5 \text{ байт} = 2,3 \text{ Кбайт}$



Пример №2

Найти количество информации, содержащейся в немецком тексте с таким же количеством символов (26 букв).

Решение:

1) Найдем мощность немецкого алфавита:

N=26 немецких прописных буквы + **26** немецких строчных буквы + **21** специальный знак = **73** символа.

2) Найдем информационный объем одного символа:

$i = \log_2 73 = 6,1$ бит.

3) Найдем объем всего текста:

$6,1 \text{ бит} * 3000 = 18300 \text{ бит} = 2287,5 \text{ байт} = 2,2 \text{ Кбайт}.$



Пример №3

Найти объем текста, записанного на языке, алфавит которого содержит 128 символов и 2000 символов в сообщении.



Дано: $K=2000$, $N=128$.

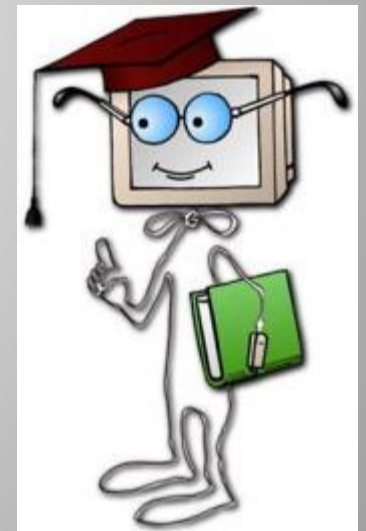
Найти: V_T - ?

Решение:

1) $i = \log_2 N = \log_2 128 = 7$ бит – объем одного символа.

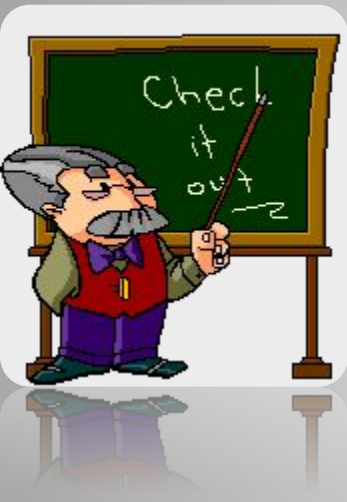
2) $V_T = i * K = 7 * 2000 = 14000$ бит = 1750 байт = 1,7 Кбайт – объем сообщения.

Ответ: 1,7 Кбайт .



Пример №4

Книга, набранная с компьютера, содержит 150 страниц: на каждой странице – 40 строк, в каждой строке – 60 символов. Каков объем информации в книге?



Дано: $K=150*40*60$, $N=256$.

Найти: V_T - ?

Решение:

1) $i = \log_2 N = \log_2 256 = 8$ бит – **объем одного символа.**

2) $V_T = i * K = 8 * 150 * 40 * 60 = 2880000$ бит = 360000 байт = 351,6 Кбайт – **объем сообщения.**

Ответ: 351,6 Кбайт.



Пример №5

Система оптического распознавания символов позволяет преобразовывать отсканированные изображения страниц документа в текстовый формат со скоростью 4 страницы в минуту и использует алфавит из 65 536 символов. Какое количество информации будет нести текст, каждая страница которого содержит 40 строк по 50 символов, после 10 минут работы приложения?



Дано: $N=65536$, $K=40*50$,

$V = 4$ стр/мин, $t= 10$ мин.

Найти: $V_T - ?$

Решение:



1) $i = \log_2 N = \log_2 65536 = 16$ бит – **объем одного символа.**

2) $V_{T \text{ стр}} = 40 * 50 * 16$ бит = 32000 бит – **объем сообщения в одной странице.**

3) $V_{T \text{ общ}} = 4 \text{ стр/мин} * 10 \text{ мин} * 32000 \text{ бит} = 1280000$ бит = 160000 байт = 156,25 Кбайт – **объем всего сообщения**

Ответ: 156,25 Кбайт.

Домашнее задание

- Стр. 111 задание 2.2