

The background features a light gray grid pattern. Overlaid on this are several dynamic, wavy lines in shades of blue and red. In the bottom right corner, there are several small, blue, bubble-like shapes. The text is centered and presented in a clean, white, sans-serif font.

Измерение информации. Алфавитный подход

Объем информации

Актуальные вопросы:

- Достаточно ли места на носителе, чтобы разместить нужную информацию?
- Сколько времени потребуется, чтобы передать информацию по имеющемуся каналу связи.

Подходы к измерению информации

- Алфавитный подход
- Содержительный подход

Алфавитный подход

Позволяет измерять информационный объём текста на любом языке (естественном или формальном).

При использовании данного подхода объём информации не связывают с содержанием текста.

В данном случае, объём зависит от информационного веса символов.



Алфавит

Алфавит — весь набор букв, знаков препинания, цифр, скобок и других символов, используемых в тексте.

В алфавит включают и пробел (пропуск между словами).



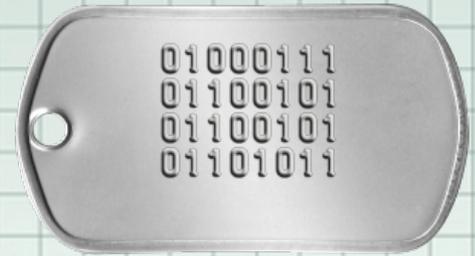
Мощность алфавита — полное число символов в алфавите.

Например:

Мощность алфавита русских букв и используемых символов равна 54:

33 буквы + 10 цифр + 11 знаков препинания, скобки, пробел.

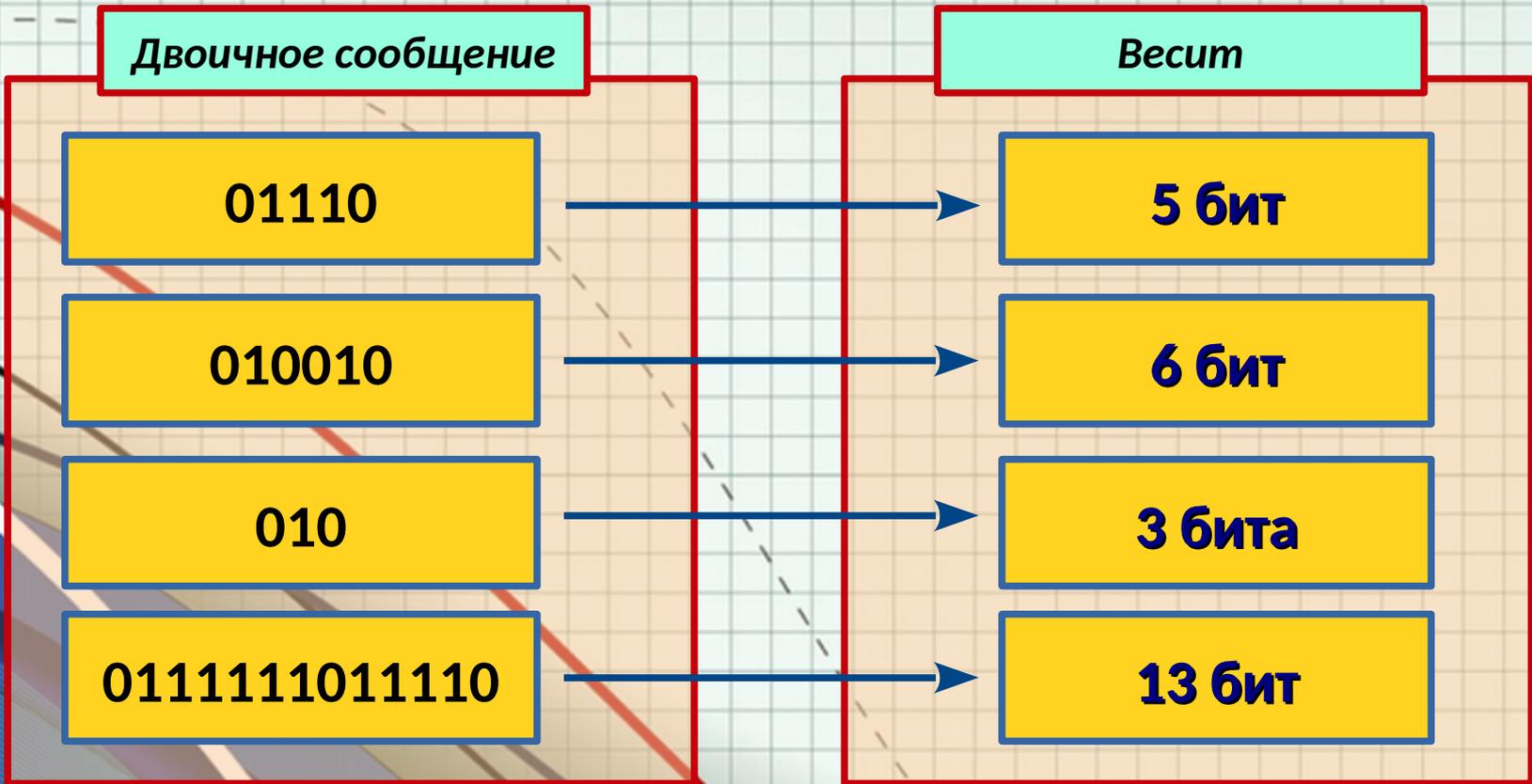
Двоичный алфавит



Наименьшую мощность имеет алфавит, используемый в компьютере (машинный язык), его называют двоичным алфавитом, т.к. он содержит только два знака «0», «1».

Информационный вес символа двоичного алфавита принят за единицу измерения информации и называется **1 бит.**

Примеры:



Алфавитный подход

При алфавитном подходе считают, что каждый символ текста, имеет информационный вес.

Информационный вес символа зависит от мощности алфавита.

С увеличением мощности алфавита, увеличивается информационный вес символа.



Информационный вес символа

Если мощность произвольного алфавита равна 4, то каждый символ может быть закодирован кодом, состоящим из 2-х знаков:

Четырёхзначный алфавит	Символ	↙	Ω	♂	∞
	Номер	1	2	3	4
Двухзначный код		00	01	10	11

Следовательно, каждый символ четырёхзначного алфавита весит **2 бита**.

Информационный вес символа

Если мощность произвольного алфавита равна 8, то каждый символ может быть закодирован кодом, состоящим из 3-х знаков:

Восьмизначный алфавит	Символ	◁	Ω	♂	∞	¿	≠	∑	ℚ
	Номер	1	2	3	4	5	6	7	8
Двузначный код		000	001	010	011	100	101	110	111

Следовательно, каждый символ четырёхзначного алфавита весит **3 бита**.

Информационный вес символа

- Алфавит из шестнадцати символов можно закодировать с помощью четырёхзначного двоичного кода.
- Какой объём информации содержат 3 символа 16-символьного алфавита?

Решение:

Так как каждый символ алфавита мощностью 16 знаков можно закодировать с помощью четырёхзначного двоичного кода, каждый символ исходного алфавита весит 4 бита.

Так как всего использовали 3 символа алфавита мощностью 16 символов, следовательно: $4 \text{ бит} \cdot 3 = 12 \text{ бит}$

Информационный вес символа

Запишем таблицу соответствия мощности алфавита (N) и количеством знаков в коде (i) - разрядностью двоичного кода.

Мощность алфавита (N)	2	4	8	16
Количество знаков в коде (i)	1	2	3	4

Заметим, что $2 = 2^1$, $4 = 2^2$, $8 = 2^3$, $16 = 2^4$.

Следовательно:

Информационный вес каждого символа, выраженный в битах (i), и мощность алфавита (N) связаны между собой формулой: $N = 2^i$

Мощность алфавита

Воспользуемся формулой $N = 2^i$ и узнаем мощность алфавита, в зависимости от объёма информации, содержащегося в одном символе данного алфавита.

Информационный вес одного символа (i)	1	2	3	4	5	6	7	8	9	10
Мощность алфавита (N)	2	4	8	16	32	64	128	256	512	1024

Мощность алфавита

- Алфавит из которого составляется на компьютере текст (документ) состоит из 256 символов.
- Этот алфавит содержит символы: строчные и прописные латинские и русские буквы, цифры, знаки арифметических операций, всевозможные скобки, знаки препинания...
- Из формулы $N = 2^i$ следует $256 = 2^8$
- Значит, каждый символ алфавита используемого в компьютере для печати документов весит 8 бит.

Эту величину приняли так же за единицу измерения информации и дали название байт.

8 бит = 1 байт

Пример

Статья содержит 30 страниц, на каждой странице - 40 строк, в каждой строке 50 символов. Какой объём информации содержит статья?

Решение:

- 1) На каждой странице $50 \cdot 40 = 2000$ символов;
- 2) Во всей статье $2000 \cdot 30 = 60000$ символов;
- 3) Т. к. вес каждого символа равен 1 байту, следовательно, информационный объём всей статьи $60000 \cdot 1 = 60000$ байт
или $60000 \cdot 8 = 480000$ бит

Как видно из задачи байт «мелкая» единица измерения информационного объёма текста, поэтому для измерения больших объёмов информации используются более крупные единицы.

Единицы

измерения информации

1 килобайт = 1 Кб = 2^{10} байт = 1024 байт

1 мегабайт = 1 Мб = 2^{10} Кб = 1024 Кб

1 гигабайт = 1 Гб = 2^{10} Мб = 1024 Мб

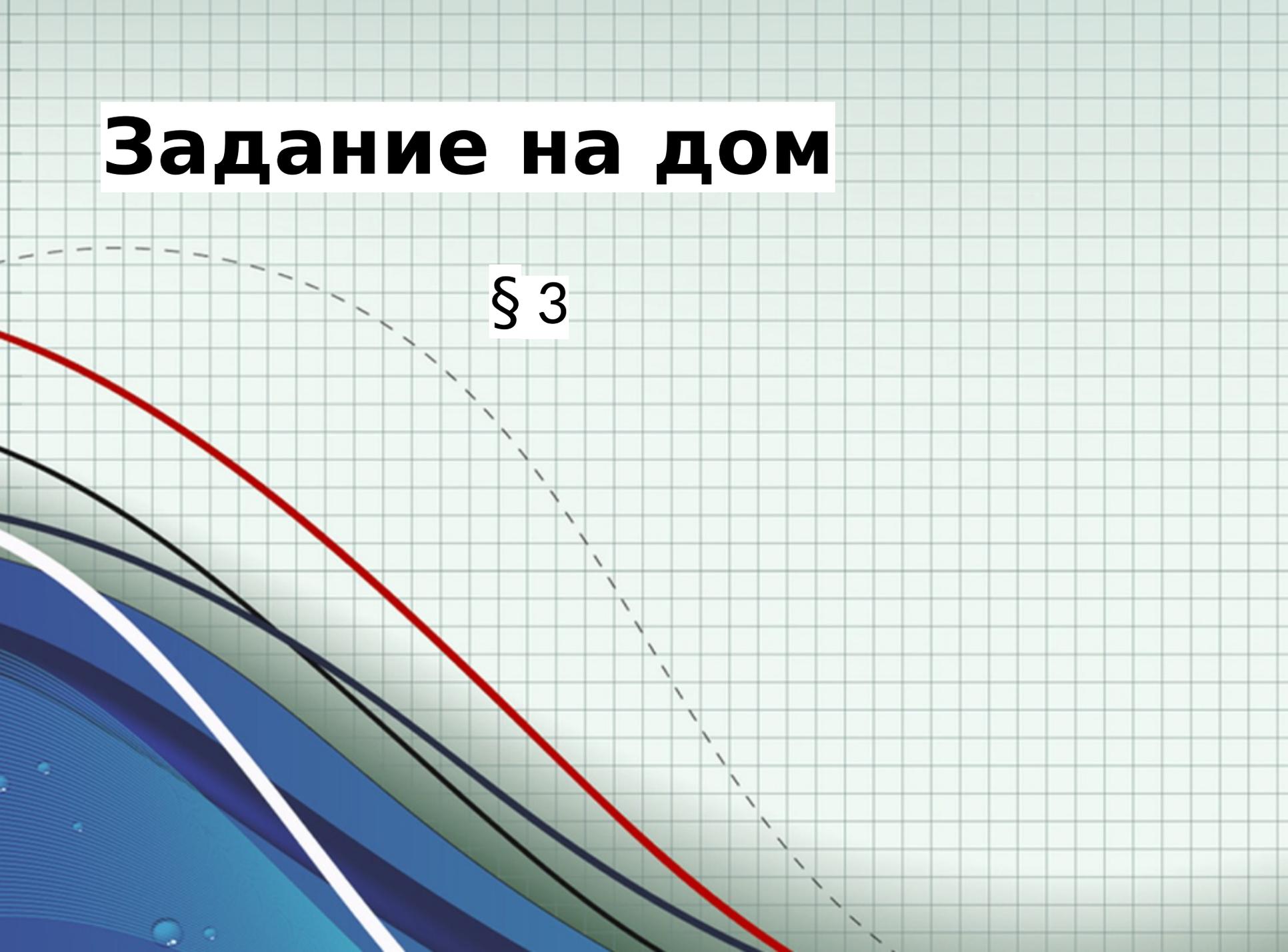
Значит, если перевести результат предыдущей задачи в более крупные единицы измерения получим:

60000 байт \approx 58,59375 Кб

60000 байт \approx 0,057 Мб

Задание на дом

§ 3

The background features a light gray grid. In the bottom-left corner, there are several overlapping, curved, abstract shapes in shades of blue and white, resembling a stylized landscape or water. A prominent red curve starts from the left edge and curves downwards towards the bottom right. A dashed gray line also starts from the left edge and curves downwards, following a similar path to the red curve but positioned slightly higher.

Задание

Вопросы и задания: 3, 8-10, стр 25