

Измерение информации

Алфавитный подход к измерению информации

Алфавитный подход позволяет измерить информационный объём сообщения, представленного на некотором языке (естественном или формальном), независимо от его содержания.

При алфавитном подходе считается, что каждый символ некоторого сообщения имеет определённый информационный вес — несёт фиксированное количество информации. Все символы одного алфавита имеют один и тот же вес, зависящий от мощности алфавита. Информационный вес символа двоичного алфавита принят за минимальную единицу измерения информации и называется 1 бит.

Информационный вес символа произвольного алфавита

Ранее мы выяснили, что алфавит любого естественного или формального языка можно заменить двоичным алфавитом. При этом мощность исходного алфавита N связана с разрядностью двоичного кода i , требуемой для кодирования всех символов исходного алфавита, соотношением: $N = 2^i$.

Разрядность двоичного кода принято считать информационным весом символа алфавита. Информационный вес символа алфавита выражается в битах.

Информационный вес символа алфавита i и мощность алфавита N связаны между собой соотношением: $N = 2^i$.

Задача 1. Алфавит племени Пульти содержит 8 символов. Каков информационный вес символа этого алфавита?

Решение. Составим краткую запись условия задачи.

$$\frac{N = 8}{i = ?}$$

Известно соотношение, связывающее величины i и N : $N = 2^i$.

С учётом исходных данных: $8 = 2^i$. Отсюда: $i = 3$.

Полная запись решения в тетради может выглядеть так:

$$\frac{N = 8}{i = ?}$$

Информационный объём сообщения

Информационный объём сообщения (количество информации в сообщении), представленного символами естественного или формального языка, складывается из информационных весов составляющих его символов.

Информационный объём сообщения I равен произведению количества символов в сообщении K на информационный вес символа алфавита i ; $I = K \cdot i$.

Задача 2. Сообщение, записанное буквами 32-символьного алфавита, содержит 140 символов. Какое количество информации оно несёт?

Решение.

$$\begin{array}{l|l} N = 32 & I = K \cdot i, \quad N = 2^i \\ K = 140 & 32 = 2^i, \quad i = 5, \quad I = 140 \cdot 5 = 700 \text{ (битов)} \\ I = ? & \end{array}$$

Ответ: 700 битов.

Единицы измерения информации

В наше время подготовка текстов в основном осуществляется с помощью компьютеров. Можно говорить о «компьютерном алфавите», включающем следующие символы: строчные и прописные русские и латинские буквы, цифры, знаки препинания, знаки арифметических операций, скобки и др. Такой алфавит содержит 256 символов. Поскольку $256 = 2^8$, информационный вес каждого символа этого алфавита равен 8 битам. Величина, равная восьми битам, называется байтом. 1 байт — информационный вес символа алфавита мощностью 256.

$$1 \text{ байт} = 8 \text{ битов}$$

Бит и байт — «мелкие» единицы измерения. На практике для измерения информационных объёмов используются более крупные единицы:

$$1 \text{ килобайт} = 1 \text{ Кб} = 1024 \text{ байта} = 2^{10} \text{ байтов}$$

$$1 \text{ мегабайт} = 1 \text{ Мб} = 1024 \text{ Кб} = 2^{10} \text{ Кб} = 2^{20} \text{ байтов}$$

1 гигабайт = 1 Гб = 1024 Мб = 2^{10} Мб = 2^{20} Кб = 2^{30} байтов
 1 терабайт = 1 Тб = 1024 Гб = 2^{10} Гб = 2^{20} Мб = 2^{30} Кб = 2^{40} байтов

Задача 4. Информационное сообщение объёмом 4 Кбайта состоит из 4096 символов. Каков информационный вес символа используемого алфавита? Сколько символов содержит алфавит, с помощью которого записано это сообщение?

Решение.

$I = 4 \text{ Кб}$ $K = 4096$ $i = ? \quad N = ?$	$I = 4 \text{ Кб} = 4 \cdot 1024 \cdot 8 \text{ битов}$	$N = 2^i,$ $I = K \cdot i, \quad i = I/K$
---	---	--

Ответ: 8 битов, 256 символов.