



Количество информации

Количество информации

- ▶ Процесс познания окружающего мира приводит к накоплению информации в форме знаний (фактов, научных теорий и т.д.)
- ▶ Получение новой информации приводит к расширению знаний (уменьшению неопределенности знания).
- ▶ Если некоторое сообщение приводит к уменьшению неопределённости нашего знания, то можно говорить, что такое сообщение содержит информацию.

Примеры



После сдачи зачета вы мучаетесь неопределённостью, какую оценку вы получили. Учитель объявил результаты и вы получили одно из двух информационных сообщение: «зачёт» или «незачёт».

Информационное сообщение об оценке за зачёт приводит к уменьшению неопределённости вашего знания в два раза, так как получено одно из информационных сообщений.

После выполнения контрольной работы вы мучаетесь неопределённостью, какую оценку вы получили. Учитель объявил результаты и вы получили одно из четырёх информационных сообщение: «2», «3», «4», «5».

Информационное сообщение об оценке за контрольную работу приводит к уменьшению неопределённости вашего знания в четыре раза, так как получено одно из четырёх информационных сообщений.



Чем больше неопределенна первоначальная ситуация (чем больше количество информационных решений), тем больше мы получим новой информации при получении сообщения (тем в больше количества раз уменьшится неопределённость знания).

Количество информации можно рассматривать как меру уменьшения неопределённости знания при получении информационных сообщений.

БИТ



За единицу количества информации принимается такое количество информации, которое содержится в информационном сообщении, уменьшающем неопределённость знания в два раза. Такая единица измерения информации называется – **битом**.

Бит – минимальная единица измерения количества информации

Производные единицы измерения количества информации

Минимальной единицей измерения количества информации является бит, а следующей по величине единицей - байт, причем:

$$1 \text{ байт} = 8 \text{ битов} = 2^3 \text{ битов.}$$



В информатике система образования кратных единиц измерения несколько отличается от принятых в большинстве наук. Традиционные метрические системы единиц, например Международная система единиц СИ, в качестве множителей кратных единиц используют коэффициент 10^n , где $n = 3, 6, 9$ и т. д., что соответствует десятичным приставкам «Кило» ($10(3)$), «Мега» ($10(6)$), «Гига» ($10(9)$) и т. д.



В компьютере информация кодируется с помощью двоичной знаковой системы, и поэтому в кратных единицах измерения количества информации используется коэффициент 2^n .

Двоичная система

Сколько байт в 1 килобайте — $2(10)$ (два в десятой степени) или же те самые 1024 байта

А сколько байтов в 1 мегабайте — $2(20)$ (два в двадцатой) или же 1048576 байт (что эквивалентно 1024 умноженному на 1024)

А сколько байт в 1 гигабайте — $2(30)$ или 107374824 байт ($1024 \times 1024 \times 1024$)

Производные единицы измерения количества информации



БАЙТ	(1 байт = 8 битов)
КИЛОБАЙТ	(1 Кбайт = 1024 байт)
МЕГАБАЙТ	(1 Мбайт = 1024 Кбайт)
ГИГАБАЙТ	(1 Гбайт = 1024 Мбайт)

Алфавитный подход к измерению информации

При алфавитном подходе к определению количества информации отвлекаются от содержания информации и рассматривают информационное сообщение как последовательность знаков определенной знаковой системы.

Все множество используемых в языке символов называют **алфавитом**.

Обычно под алфавитом понимают только буквы, но поскольку в тексте могут встречаться знаки препинания, цифры, скобки, то их тоже нужно включить в алфавит. В алфавит также следует включить и пробел, т.е. пропуск между словами.

Полное количество символов алфавита принято называть мощностью алфавита. Обозначают эту величину буквой N.

При алфавитном подходе к измерению информации количество информации зависит не от содержания, а от размера текста и мощности алфавита.

т.е. количество информации, содержащееся в символьном сообщении из k символов, равно $k \cdot i$, где i - вес одного символа в битах.

Количество информации i , которое несет один символ в тексте и мощность алфавита N связаны формулой

$$N=2^i$$



В русском алфавите (не считая букву ё) 32 символа. Определим информационный вес одного символа.

Ответ:

$$N = 32 = 2 * 2 * 2 * 2 * 2 = 32 = 5 \text{ бит (информационный вес одного символа)}$$

Информационный объем сообщения

В русском алфавите (не считая букву ё) 32 символа. Определим информационный объем сообщения: «Информатика – наука об информации».

$$I = K * i$$

Решение:

$$N = 32$$

$$K = 33 \text{ (количество символов в сообщении)}$$

$$N = 32 = i = 5 \text{ бит}$$

$$I = 33 * 5 = 165 \text{ бит}$$

Решение задач:

Книга набранная на ПК содержит 150 страниц, на каждой странице 40 строк, в строке 60 символов. Какой информационный объем книги?

Решение:

$N=256$ (алфавит ПК)

$$N=2^i$$

$$256=2^i$$

$k=150 \times 40 \times 60$

найдем i , $i=8$ бит на символ или

1 байт

$V=?$

$k=150 \times 40 \times 60=360000$ символов

$V=k \cdot i=360000 \cdot 1=360000$ байт

переведем ответ в КБ

$360000/1024=351,6$ КБ

Упражнение:

1. Вычислите, какое количество информации в битах содержится в 1 Кб, 1 Мб, 1 Гб.
2. Была получена телеграмма: «Встречайте, вагон 7». Известно, что в составе поезда 16 вагонов. Какое количество информации было получено?

в 1 кб содержится 1024 байт = 8192 бит ($1024 * 8$)

$$16 = 2^4$$

в 1 мб содержится 1024 кб = 1048576 байт =
8388608 бит ($1024 * 1024 * 8$)

Смотрим степень двойки - 4.

в 1 гб содержится 1024 мб = 1048576 кб =
1073741824 байт = 8589934592 бит ($1024 * 1024 * 1024 * 8$)

Ответ: 4 бита.

Самостоятельная работа

1. Была получена телеграмма: «Встречайте, вагон 7». Известно, что в составе поезда 64 вагона. Какое количество информации было получено?
2. Была получена телеграмма: «Встречайте, вагон 7». Известно, что в составе поезда 32 вагона. Какое количество информации было получено?
3. Книга набранная на ПК содержит 180 страниц, на каждой странице 60 строк, в строке 50 символов. Какой информационный объем книги?

527,4 Кб



4. В русском алфавите (не считая букву ё) 32 символа. Определим информационный объем сообщения: «Количество информации».

5. В русском алфавите (не считая букву ё) 32 символа. Определим информационный объем сообщения: «Единицы измерения информации».