

Содержательный подход к измерению информации

В **содержательном подходе** количество информации, заключенное в сообщении, определяется объемом знаний, который это сообщение несет получающему его человеку.

Вспомним, что с «человеческой» точки зрения **информация** — это знания, которые мы получаем из внешнего мира. Количество информации, заключенное в сообщении, должно быть тем больше, чем больше оно пополняет наши знания.

Вы уже знаете, что за единицу измерения информации принимается **1 бит**.



1 бит — минимальная единица измерения количества информации.

Проблема измерения информации исследована в теории информации, основатель которой — *Клод Шеннон*.

В теории информации для бита дается следующее определение:



Сообщение, уменьшающее неопределенность знания в два раза, несет **1 бит** информации.

Что такое неопределенность знания, поясним на примерах.

Допустим, вы бросаете монету, загадывая, что выпадет: орел или решка. Есть всего два возможных результата бросания монеты. Причем ни один из этих результатов не имеет преимущества перед другим. В таком случае говорят, что они **равновероятны**.

В случае с монетой перед ее подбрасыванием неопределенность знания о результате равна двум.

Игральный же кубик с шестью гранями может с равной вероятностью упасть на любую из них. Значит, неопределенность знания о результате бросания кубика равна шести.

Еще пример: спортсмены-лыжники перед забегом путем жеребьевки определяют свои порядковые номера на старте. Допустим, что имеется **100** участников соревнований, тогда неопределенность знания спортсмена о своем номере до жеребьевки равна **100**.

Следовательно, можно сказать так:



Неопределенность знания о результате некоторого события (бросание монеты или игрального кубика, вытаскивание жребия и др.) — это количество возможных результатов.

Вернемся к примеру с монетой. После того как вы бросили монету и посмотрели на нее, вы получили зрительное сообщение, что выпал, например, орел. Определился один из двух возможных результатов. Неопределенность знания уменьшилась в два раза: было два варианта, остался один. Значит, узнав результат бросания монеты, вы получили 1 бит информации.

Сообщение об одном из двух равновероятных результатов некоторого события несет 1 бит информации.

Пусть в некотором сообщении содержатся сведения о том, что произошло одно из N равновероятных событий.

Тогда количество информации i , содержащееся в сообщении о том, что произошло одно из N равновероятных событий, можно определить из формулы Хартли:

$$N = 2^i.$$

Данная формула является показательным уравнением относительно неизвестного i .

Из математики известно, что решение такого уравнения имеет вид:

$$i = \log_2 N - \text{логарифм } N \text{ по основанию } 2.$$

Если N равно целой степени двойки (2, 4, 8, 16 и т. д.), то такое уравнение можно решить «в уме».

Пример:

Шахматная доска состоит из 64 полей: 8 столбцов на 8 строк.

Какое количество бит несет сообщение о выборе одного шахматного поля?

Решение.

Поскольку выбор любой из 64 клеток равновероятен, то количество бит находится из формулы:

$$2^i = 64,$$

$$i = \log_2 64 = 6, \text{ так как } 2^6 = 64.$$

Следовательно, $i = 6$ бит.

В противном случае количество информации становится нецелой величиной, и для решения задачи придется воспользоваться таблицей двоичных логарифмов.

Также, если N не является целой степенью 2, то можно выполнить округление i в большую сторону. При решении задач в таком случае i можно найти как $\log_2 K$, где K — ближайшая к N степень двойки, такая, что $K > N$.

Пример:

При игре в кости используется кубик с шестью гранями.

Сколько битов информации получает игрок при каждом бросании кубика?

Решение.

Выпадение каждой грани кубика равновероятно. Поэтому количество информации от одного результата бросания находится из уравнения: $2^i = 6$.

Решение этого уравнения: $i = \log_2 6$.

Из таблицы двоичных логарифмов следует (с точностью до 3-х знаков после запятой):

$i = 2,585$ бита.

Данную задачу также можно решить округлением i в большую сторону: $2^i = 6 < 8 = 2^3$, $i = 3$ бита.

Задача 1

Специальное устройство на автостоянке таксопарка регистрирует заезд на территорию автомобилей фирмы, записывая их индивидуальные номера с использованием минимально возможного количества бит, одинакового для каждого автомобиля. Каков информационный объём сообщения, записанного устройством, если на территорию за рассматриваемый промежуток времени заехало 16 из 27 машин таксопарка?

Задача 2

Специальное устройство на автостоянке таксопарка регистрирует заезд на территорию автомобилей фирмы, записывая их индивидуальные номера с использованием минимально возможного количества бит, одинакового для каждого автомобиля. Каков информационный объем сообщения, записанного устройством, если на территорию за рассматриваемый промежуток времени заехало 24 из 28 машин таксопарка? (Ответ дайте в байтах.)

Задача 3

При регистрации в компьютерной системе каждому пользователю выдаётся пароль, состоящий из 11 символов и содержащий только символы А, Б, В, Г, Д, Е. Каждый такой пароль в компьютерной программе записывается минимально возможным и одинаковым целым количеством байт, при этом используют посимвольное кодирование и все символы кодируются одинаковым и минимально возможным количеством бит. Определите, сколько байт необходимо для хранения 20 паролей.

Задача 4

При регистрации в компьютерной системе каждому пользователю выдаётся пароль, состоящий из 23 символов и содержащий только символы A, F, G, Y, S, L (таким образом, используется 6 различных символов). Каждый такой пароль в компьютерной программе записывается минимально возможным и одинаковым целым количеством байт (при этом используют посимвольное кодирование и все символы кодируются одинаковым и минимально возможным количеством бит). Определите объём памяти, отводимый этой программой для записи 50 паролей.

Задачи для закрепления

1. В велокроссе участвуют 28 спортсменов. Специальное устройство регистрирует прохождение каждым из участников промежуточного финиша, записывая его номер с использованием минимально возможного количества бит, одинакового для каждого спортсмена. Какой объём памяти будет использован устройством, когда все спортсмены прошли промежуточный финиш? (Ответ дайте в битах.)
2. Автомобильный номер состоит из 6 символов. Допустимыми символами считаются 10 цифр и 8 заглавных букв: А, В, С, Е, Н, К, М и Р. Для хранения каждого из 18 допустимых символов используется одинаковое и наименьшее возможное количество бит. Для хранения каждого номера используется одинаковое и минимально возможное количество байт. Сколько байт памяти потребуется для хранения 400 автомобильных номеров? Номера хранятся без разделителей.
3. При регистрации в компьютерной системе каждому пользователю выдается пароль, состоящий из 15 символов и содержащий только символы Е, Г, Э, 2, 0, 1, 3. Каждый такой пароль в компьютерной программе записывается минимально возможным и одинаковым целым количеством байт (при этом используют посимвольное кодирование и все символы кодируются одинаковым и минимально возможным количеством бит). Определите объём памяти, отводимый этой программой для записи 30 паролей. (Ответ дайте в байтах.)