

Дополнение к главе I

$$2 \times 2 = 4$$



1.1. Неопределенность знания и количество информации

Основные темы параграфа:

- *содержательный подход к измерению информации;*
- *неопределенность знания;*
- *сообщение, несущее 1 бит информации;*
- *количество информации в сообщении об одном из N равновероятных событий.*

Содержательный подход к измерению информации

В § 4 приведен алфавитный подход к измерению информации. Здесь мы рассмотрим другой подход, который назовем **содержательным**. В этом случае количество информации связывается с содержанием (смыслом) полученного человеком сообщения. 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.

Как же с этой точки зрения определяется единица измерения информации? Вы уже знаете, что эта единица называется битом. Единица измерения информации была определена в науке, которая называется **теорией информации**. Ее определение звучит так:



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

В этом определении есть понятия, которые требуют пояснения.

Неопределенность знания

Что такое **неопределенность знания**? Лучше всего это объяснить на примерах. Допустим, вы бросаете монету, загадывая, что выпадет: орел или решка. Есть всего два варианта возможного результата бросания монеты. Причем ни один из этих вариантов не имеет преимущества перед другим. В таком случае говорят, что они **равновероятны***.

Так вот в этом случае перед подбрасыванием монеты неопределенность знания о результате равна двум. Игральный 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. Пусть имеется 100 участников соревнований. Тогда неопределенность знания спортсменом своего номера до жеребьевки равна 100.

Следовательно, можно сказать так: неопределенность знания о некотором событии — это количество возможных результатов события (бросания монеты, кубика; вытаскивания жребия и пр.).

Сообщение, несущее 1 бит информации

Вернемся теперь к примеру с монетой. После того как вы бросили монету, она упала, и вы посмотрели на нее, вы получили зрительное сообщение, что выпал, например, орел. Произошло одно из двух возможных событий. В этом случае считают, что неопределенность знания уменьшилась в два раза: было два варианта, остался один. Значит, *узнав результат бросания монеты, вы получили 1 бит информации.*



Сообщение о том, что произошло одно событие из двух равновероятных, несет 1 бит информации.



Количество информации в сообщении об одном из N равновероятных событий

А теперь такая задача: студент на экзамене может получить одну из четырех оценок: 5 — «отлично», 4 — «хорошо», 3 — «удовлетворительно», 2 — «неудовлетворительно». Представьте себе, что ваш

* Более строгое определение равновероятности: если увеличивать количество бросаний монеты (100, 1000, 10 000 и т. д.), то число выпадений орла и число выпадений решки будут все более близкими.

товарищ пошел сдавать экзамен. Причем учиться он очень неровно и может с одинаковой вероятностью получить любую оценку от 2 до 5. Вы волнуетесь за него, ждете результата экзамена. Наконец, он пришел и на ваш вопрос: «Ну, что получил?» ответил: «Четверку!».

Вопрос: сколько битов информации содержится в его ответе?

Если сразу сложно ответить на этот вопрос, то давайте подойдем к ответу постепенно. Будем отгадывать оценку, задавая вопросы, на которые можно ответить только «да» или «нет».

Вопросы будем ставить так, чтобы каждый ответ уменьшал количество вариантов в два раза и, следовательно, приносил 1 бит информации.

Первый вопрос:

— Оценка выше тройки?

— Да.

После этого ответа число вариантов уменьшилось в два раза. Остались только 4 и 5. Получен 1 бит информации.

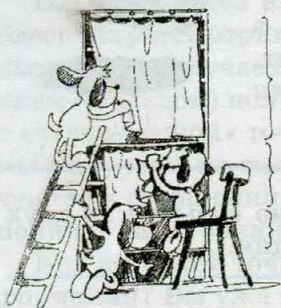
Второй вопрос:

— Ты получил пятерку?

— Нет.

Выбран один вариант из двух оставшихся: оценка — четверка. Получен еще 1 бит информации. В сумме имеем 2 бита.

Сообщение о том, что произошло одно из четырех равновероятных событий, несет 2 бита информации.



Разберем еще одну частную задачу, а потом получим общее правило.

На книжном стеллаже восемь полок. Книга может стоять на любой из них. Сколько информации содержит сообщение о том, где находится книга?

Будем действовать таким же способом, как в предыдущей задаче. Метод поиска, на каждом шаге которого отбрасывается половина вариантов, называется **методом половинного деления**. Применим метод половинного деления к задаче со стеллажом.

Задаем вопросы:

— Книга лежит выше четвертой полки?

— Нет.

— Книга лежит ниже третьей полки?

— Да.

— Книга на второй полке?

— Нет.

— Ну теперь все ясно! Книга лежит на первой полке!

Каждый ответ уменьшал неопределенность знания в два раза. Всего было задано три вопроса. Значит, набрано 3 бита информации. И если бы сразу было сказано, что книга лежит на первой полке, то этим сообщением были бы переданы те же 3 бита информации.

А сейчас попробуем получить формулу, по которой вычисляется количество информации, содержащейся в сообщении о том, что произошло одно из множества равновероятных событий.

Обозначим буквой N количество возможных событий. Буквой i будем обозначать количество информации в сообщении о том, что произошло одно из N событий.

В примере с монетой: $N = 2, i = 1$.

В примере с оценками: $N = 4, i = 2$.

В примере со стеллажом: $N = 8, i = 3$.

Нетрудно заметить, что связь между этими величинами выражается такой формулой

$$2^i = N.$$

Действительно: $2^1 = 2$; $2^2 = 4$; $2^3 = 8$.

Если величина N известна, а i неизвестно, то данная формула становится уравнением для определения i . В математике оно называется *показательным уравнением*.

Например, пусть на стеллаже не 8, а 16 полок. Чтобы ответить на вопрос, сколько информации содержится в сообщении о том, где лежит книга, нужно решить уравнение

$$2^i = 16.$$

Поскольку $16 = 2^4$, то $i = 4$.



Количество информации i , содержащееся в сообщении о том, что произошло одно из N равновероятных событий, определяется из решения показательного уравнения $2^i = N$.



Если значение N равно целой степени двойки (4, 8, 16, 32, 64 и т. д.), то такое уравнение решается просто: i будет целым числом. А чему, например, равно количество информации в сообщении о результате бросания игральной кости, у которой имеется шесть граней и, следовательно, $N = 6$? Решение уравнения

$$2^i = 6$$

будет дробным числом, лежащим между 2 и 3, поскольку $2^2 = 4 < 6$, а $2^3 = 8 > 6$. С точностью до пяти знаков после запятой решение такое: 2,58496. Ниже приведена табл. 1.1, из которой можно определить i для различных значений N в диапазоне от 1 до 64.

Таблица 1.1. Количество информации в сообщении об одном из N равновероятных событий

N	i	N	i	N	i	N	i
1	0,00000	17	4,08746	33	5,04439	49	5,61471
2	1,00000	18	4,16993	34	5,08746	50	5,64386
3	1,58496	19	4,24793	35	5,12928	51	5,67243
4	2,00000	20	4,32193	36	5,16993	52	5,70044
5	2,32193	21	4,39232	37	5,20945	53	5,72792
6	2,58496	22	4,45943	38	5,24793	54	5,75489
7	2,80735	23	4,52356	39	5,28540	55	5,78136
8	3,00000	24	4,58496	40	5,32193	56	5,80735
9	3,16993	25	4,64386	41	5,35755	57	5,83289
10	3,32193	26	4,70044	42	5,39232	58	5,85798
11	3,45943	27	4,75489	43	5,42626	59	5,88264
12	3,58496	28	4,80735	44	5,45943	60	5,90689
13	3,70044	29	4,85798	45	5,49185	61	5,93074
14	3,80735	30	4,90689	46	5,52356	62	5,95420
15	3,90689	31	4,95420	47	5,55459	63	5,97728
16	4,00000	32	5,00000	48	5,58496	64	6,00000

Коротко о главном

Неопределенность знания о некотором событии — это количество возможных результатов события.

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

Для определения количества информации i , содержащейся в сообщении о том, что произошло одно из N равновероятных событий, нужно решить показательное уравнение $2^i = N$.

Вопросы и задания

1. Что такое неопределенность знания о результате какого-либо события? Приведите примеры.
2. Как с точки зрения содержательного подхода к измерению информации определяется единица измерения количества информации?
3. По какой формуле можно вычислить количество информации, содержащейся в сообщении?
4. Сколько битов информации несет сообщение о том, что из колоды в 32 карты достали «даму пик»?
5. Проводятся две лотереи: «4 из 32» и «5 из 64». Сообщение о результатах какой из лотерей несет больше информации?

ЕК ЦОР: Часть 1, дополнение к главе 1. ЦОР № 1–5.



www