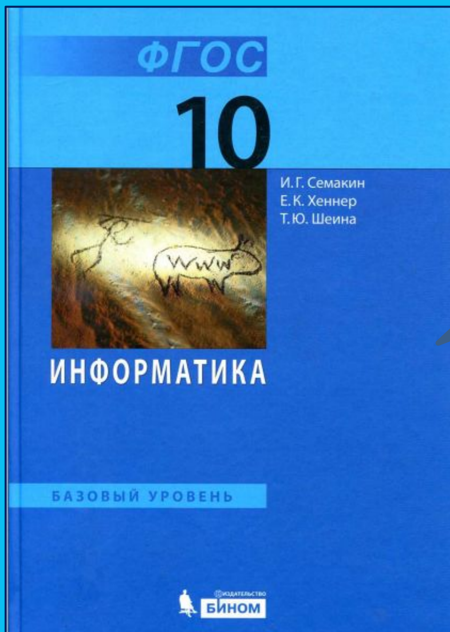
Урок 13-14

Готовимся к уроку

Домашнее задание



§ 4, *стр 26-34,*
вопросы 2-4, *стр. 35, устно*
вопросы 6,7,9, *стр. 35-36, письменно*

Уроки 4-5

Информатика

ТЕМА 3. ИЗМЕРЕНИЕ ИНФОРМАЦИИ (3 ЧАСА)

Количество информации как мера уменьшения неопределенности знаний.

содержательный подход;

алфавитный подход;

мощность алфавита;

информационный вес символа;

информационный объем текста;

единицы измерения информации;

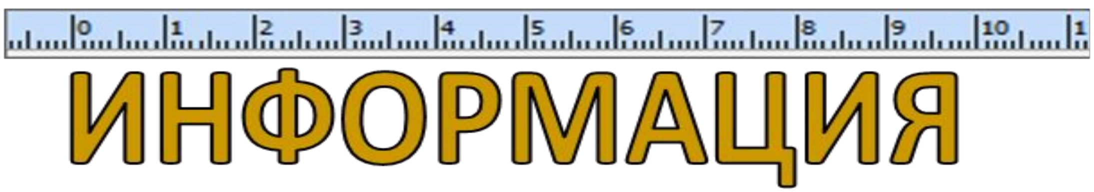
10 класс

ИЗДАТЕЛЬСТВО БИНОМ

Измерение информации.

Вопрос: «**Как измерить информацию?**»

Ответ на него зависит от того, что понимать под информацией. Но поскольку определять информацию можно по-разному, то и **способы измерения** тоже **могут быть разными**.

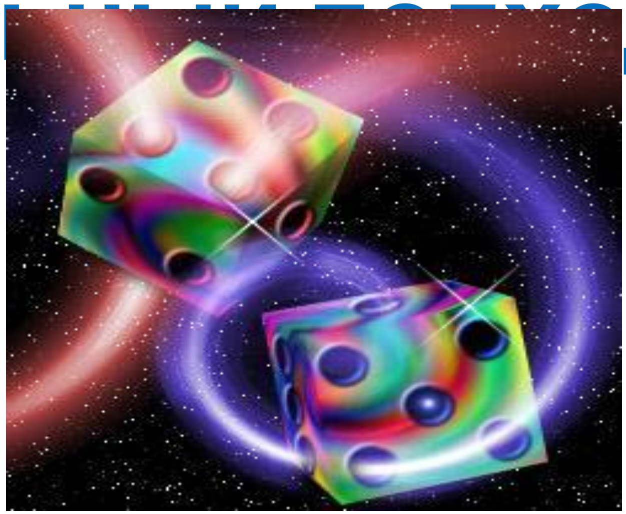


Токмакова Людмила Викторовна, учитель информатики МБОУ «СОШ №32» ЭМР саратовской области

Подходы к измерению информации

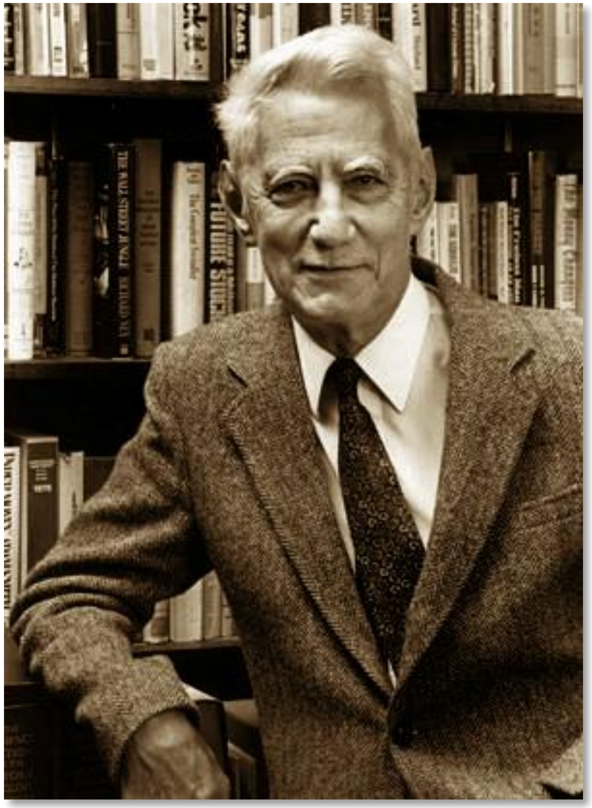
1. Алфавитный

2. Содержательный подход



Содержательный подход

Информация – это снятая неопределенность. Величина неопределённости некоторого события – это количество возможных результатов данного события.



Клод Элвуд Шеннон (1916-2001) – американский инженер и математик. Является основателем теории информации, нашедшей применение в современных высокотехнологических системах связи.

В 1948 году предложил использовать слово «*бит*» для обозначения наименьшей единицы информации.

УМЕНЬШЕНИЕ НЕОПРЕДЕЛЁННОСТИ ЗНАНИЯ

Пусть у нас имеется монета, которую мы бросаем на ровную поверхность.

С равной вероятностью произойдет одно из двух возможных событий – монета окажется в одном из двух положений: «орёл» или «решка».

Возможные события

Произошедшее событие

События равновероятны, если при возрастающем числе опытов количества выпадений «орла» и «решки» постепенно сближаются.

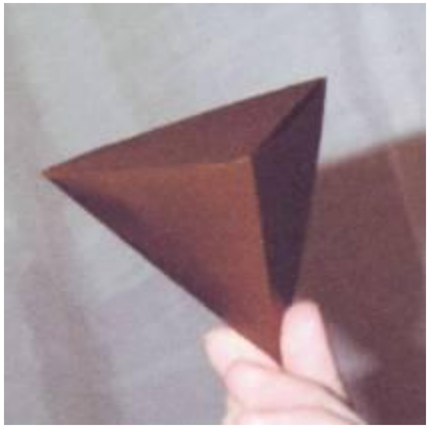
Перед броском существует неопределённость нашего знания (возможны два события), а после броска наступает полная определённость.

Неопределённость нашего знания уменьшается в два раза, так как из двух возможных равновероятностных событий реализовалось одно.

УМЕНЬШЕНИЕ НЕОПРЕДЕЛЁННОСТИ ЗНАНИЯ

Неопределенность знания о результате некоторого события — это число возможных результатов события.

УМЕНЬШЕНИЕ НЕОПРЕДЕЛЁННОСТИ ЗНАНИЯ



При бросании равносторонней четырехгранной пирамиды существуют 4 равновероятных события.

При бросании шестигранного игрального кубика существует 6 равновероятных событий.

Содержательный подход

Не знаю

- 8
- 4
- 4

Шар розовый?

1 вопрос

Знаю

i = **1 бит**

Сколько информации?

Содержательный подход

Не знаю

- 8
  - 4

Шар розовый?

- 4
  - 2

Шар синий?

2 вопроса

Знаю

i = 2 бита

Сколько информации?

Содержательный подход

Не знаю

- •8
- •4 Шар розовый?
- •4
- •2 Шар синий?
- •1
- •1
- •2 Шар зеленый?

3 вопроса

Знаю

i = **3 бита**

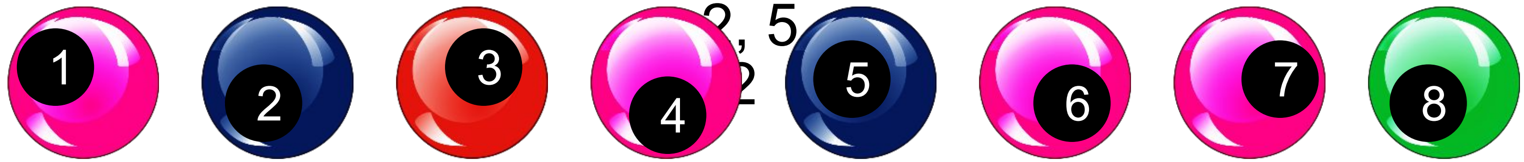
Сколько информации?

Метод половинного деления

•1-8

•2, 5

•2



•5

Исследуйте, сколько вопросов с ответами *Да/Нет* надо задать, чтобы определить цифру на шаре, если начать с вопроса: «*Шар синий?*»

По какому принципу необходимо выбирать вопросы, чтобы определить искомый ответ как можно быстрее?

•1, 3, 4, 6, 7, 8

•1, 4, 6, 7

•1, 7

•1

•7

•6, 4

•4

•6

Количество информации i, содержащееся в сообщении об одном из N равновероятных результатов некоторого события, определяется из решения уравнения

$$2^i = N.$$

4 вопроса

Содержательный подход

Стол телевизионной игры «Что? Где? Когда?» разбит на 13 равных секторов. Какое количество информации содержит сообщение ведущего: «*Волчок указывает на супер-блиц*».

•13

Не знаю

•7

$2^i = N$

•4

•2

$13 \le 16 = 2^4$

•2

•1

$i = 4$ **бита**

4 вопроса

Знаю

С позиции содержательного подхода к измерению информации решается вопрос о количестве информации в сообщении, получаемом человеком. Рассматривается следующая ситуация:

1) человек получает сообщение о некотором событии; при этом заранее известна ***неопределенность знания*** человека об ожидаемом событии. Неопределенность знания может быть выражена либо числом возможных вариантов события, либо вероятностью ожидаемых вариантов события;

2) в результате получения сообщения неопределенность знания снимается: из некоторого возможного количества вариантов оказался выбранным один;

3) по формуле вычисляется количество информации в полученном сообщении, выраженное **в битах.**

Формула, используемая для вычисления количества информации, зависит от ситуаций, которых может быть две:

1. Все возможные варианты события равновероятны. Их число конечно и равно N.

2. Вероятности (p) возможных вариантов события разные и они заранее известны:

$\{p_i\}$, $i = 1..N$.

Здесь по-прежнему N — число возможных вариантов события.

Равновероятные события

Если обозначить буквой i количество информации в сообщении о том, что произошло одно из N равновероятных событий, то величины i и N связаны между собой формулой Хартли:

$$2^i = N$$

1 бит — это количество информации в сообщении об одном из двух равновероятных событий.

Формула Хартли — это показательное уравнение. Если i — неизвестная величина, то решением данного уравнения будет:

$$i = \log_2 N$$

Данные формулы тождественны друг другу.

ПОКАЗАТЕЛЬНОЕ УРАВНЕНИЕ

$$2^{i} = N$$

N → i

i → N

*Количество **i** информации в сообщении о том, что произошло одно из **N** равновероят-ных событий.*

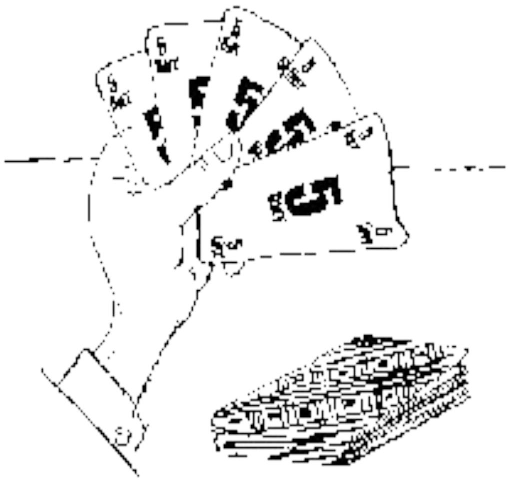
N	i	N	i	N	i	N	i
1	0.00000	17	4.08746	33	5.04439	49	5.61471
2	1.00000	18	4.16993	34	5.08746	50	5.64386
3	1.58496	19	4.24793	35	5.12928	51	5.67243
4	**2.00000**	20	4.32193	36	5.16993	52	5.70044
5	2.32193	21	4.39232	37	5.20945	53	5.72792
6	2.58496	22	4.45943	38	5.24793	54	5.75489
7	2.80735	23	4.52356	39	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**

Таблица представлена в приложении 1

Рассмотрим несколько примеров:

Пример 1. Сколько информации содержит сообщение о том, что из колоды карт достали даму пик?

Решение: В колоде 32 карты. В перемешанной колоде выпадение любой карты — равновероятные события. Если *i* — количество информации в сообщении о том, что выпала конкретная карта (например, дама пик), то из уравнения Хартли:

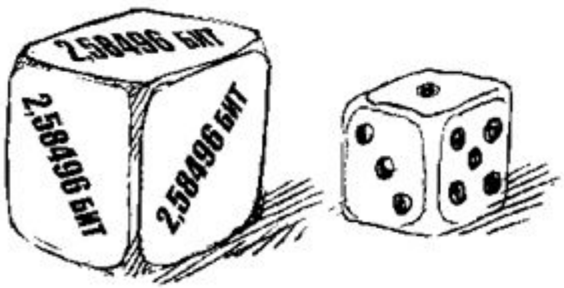


$$2^i = 32 = 2^5$$

Отсюда: i = 5 бит.

Пример 2. Сколько информации содержит сообщение о выпадении грани с числом 3 на шестигранном игральном кубике?

Решение: Считая выпадение любой грани событием равновероятным, запишем формулу Хартли:



$$2^i = 6.$$

Отсюда: $i = \log_2 6 = 2{,}58496$ бит.

Неравновероятные события (вероятностный подход). Если вероятность некоторого события равна *p*, а *i* (бит) — это количество информации в сообщении о том, что произошло это событие, то данные величины связаны между собой формулой:

$$2^i = 1/p$$

Решая данное показательное уравнение относительно *i*, получаем:

$$i = \log_2(1/p)$$

формула Шеннона

Качественный подход

Информация — *это знания людей, получаемые ими из различных сообщений.*

Сообщение — *это информационный поток (поток данных), который в процессе передачи информации поступает к принимающему его субъекту.*

Сообщение

Информативные **сообщение**, *которое пополняет знания человека, т.е. несет для него информацию.*

Неинформативные *сведения “старые”, т.е. человек это уже знает, или содержание сообщения непонятно человеку*

Количественный подход в приближении равновероятности

События равновероятны, *если ни одно из них не имеет преимущества перед другими.*

Рассмотрим на примере.
«Сколько информации несет сообщение о результате бросания шестигранного кубика?»

Из уравнения Хартли: $2^i = 6$.
Поскольку $2^2 < 6 < 2^3$, следовательно, $2 < i < 3$.

Затем определяем более точное значение (с точностью до пяти знаков после запятой), что i = 2,58496 бит.

Отметить, что при данном подходе количество информации может быть выражено дробной величиной.

Вероятностный подход к измерению информации

Вероятность некоторого события — это величина, которая может принимать значения от нуля до единицы.

Вероятность ***невозможного*** *события равна* ***нулю*** (например: «завтра Солнце не взойдет над горизонтом»)

Вероятность ***достоверного*** *события равна* ***единице*** (например: «Завтра солнце взойдет над горизонтом»).

Вероятность ***некоторого*** *события определяется путем многократных наблюдений (измерений, испытаний). Такие измерения называют статистическими. И чем большее количество измерений выполнено, тем точнее определяется вероятность события.*

Рассмотрим несколько примеров:

Пример 3. На автобусной остановке останавливаются два маршрута автобусов: № 5 и № 7.

Ученику дано задание: определить, сколько информации содержит сообщение о том, что к остановке подошел автобус № 5, и сколько информации в сообщении о том, что подошел автобус № 7.

Решение: Ученик провел исследование. В течение всего рабочего дня он подсчитал, что к остановке автобусы подходили 100 раз.

Из них — 25 раз подходил автобус № 5 и 75 раз подходил автобус № 7.

Сделав предположение, что с такой же частотой автобусы ходят и в другие дни, ученик вычислил вероятность появления на остановке автобуса № 5:

$$p_5 = 25/100 = 1/4,$$

и вероятность появления автобуса № 7:

$$p_7 = 75/100 = 3/4.$$

Отсюда, количество информации в сообщении об автобусе № 5 равно:

$$i_5 = \log_2 4 = 2 \text{ бита}.$$

Количество информации в сообщении об автобусе № 7 равно:

$$i_7 = \log_2(4/3) = \log_2 4 - \log_2 3 = 2 - 1{,}58496 = 0{,}41504 \text{ бита}.$$

Пример 4. Рассмотрим другой вариант задачи об автобусах. На остановке останавливаются автобусы № 5 и № 7. Сообщение о том, что к остановке подошел автобус № 5, несет 4 бита информации. Вероятность появления на остановке автобуса с № 7 в два раза меньше, чем вероятность появления автобуса № 5. Сколько бит информации несет сообщение о появлении на остановке автобуса № 7?

Решение: Запишем условие задачи в следующем виде:

i_5 = 4 бита, $p_5 = 2 \cdot p_7$

Вспомним связь между вероятностью и количеством информации: $2^i = 1/p$

Отсюда: $p = 2^{-i}$

Подставляя в равенство из условия задачи, получим:

$$2^{-i_5} = 2 \times 2^{-i_7}; \qquad 2^{-4} = 2 \times 2^{-i_7} = 2^{1-i_7};$$

Отсюда: $i_7 - 1 = 4;$ $\qquad i_7 = 5$ бит

Из полученного результата следует вывод: уменьшение вероятности события в 2 раза увеличивает информативность сообщения о нем на 1 бит.

Очевидно и обратное правило: увеличение вероятности события в 2 раза уменьшает информативность сообщения о нем на 1 бит. Зная эти правила, предыдущую задачу можно было решить «в уме».