

Измерение информации



Алфавитный
подход

Единицы
измерения
информации

Содержательный
подход

Три кита на которых стоит мир



Вещество

Можно ли измерить
вещество?

Например
длина (мм, см, д
вес (мг, г, кг,
объем (мм^3 , см



Энергия

Можно ли измерить
энергию?

Н
тепло
электр



Информация

Как измерить
количество
информации?

Какие единицы измерения?

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



Понятное/ новое \Rightarrow
сообщение
информативно
(количество информации > 0)

Непонятное \Rightarrow сообщение
неинформативно
(количество информации $= 0$)

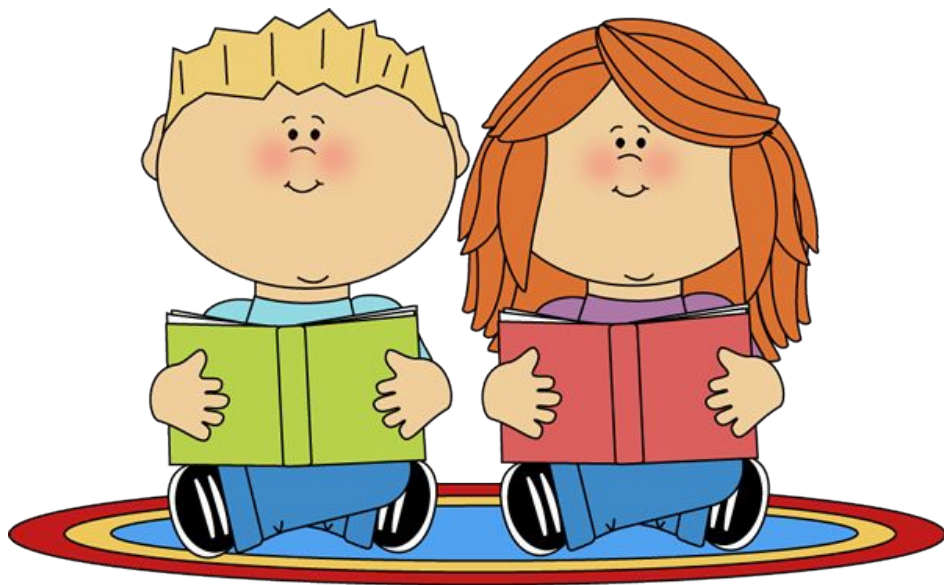
$$\text{Log}_a x + \text{Log}_a y = \text{log}_a(xy)$$

СООБЩЕНИЕ



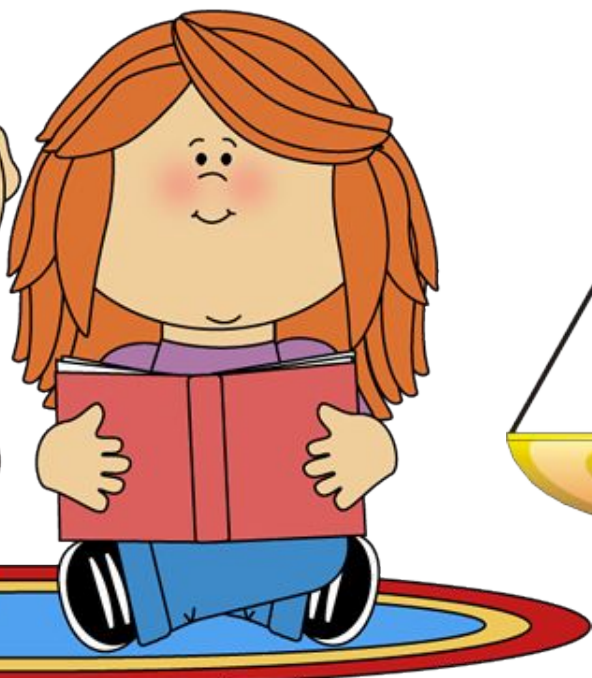
Вывод:

Сообщение несет информацию для человека **(информативно)**, если содержащиеся в нем сведения являются для него **новыми** и **понятными**.



**Зависит от
человека
(субъективно)**

Для того, чтобы измерить
информацию объективно,
нельзя опираться на её содержание



Алфавитный (объёмный) подход к измерению информации



Алфавитный подход

- ✓ Алфавитный подход позволяет измерять количество информации в тексте (символьном сообщении), составленном из символов некоторого алфавита.
- ✓ Информационный объём сообщения не зависит от содержания (объективный подход)
- ✓ Используется в технике



Алфавит - ?

Алфавит – набор знаков, используемых при кодировании информации с помощью некоторого языка.

Примеры:

АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ N=33

ABCDEFGHIJKLMNOPQRSTUVWXYZ N=26

- + N=2

0123456789 N=10

Мощность алфавита – количество символов в алфавите (N).



Алфавитный подход

✓ Каждый символ имеет свой информационный вес

текст записан с помощью символов
некоторого алфавита (различаются своей
мощностью) \Rightarrow **вес символа зависит от
мощности алфавита**



Какова минимальная мощность алфавита, с помощью которого можно записывать (кодировать) информацию?



Двоичный алфавит

Информационный вес символа
двоичного алфавита принят за
единицу информации – **1 бит**

0

1



Информация, записанная на машинном языке, весит:

011 100 1 7 бит

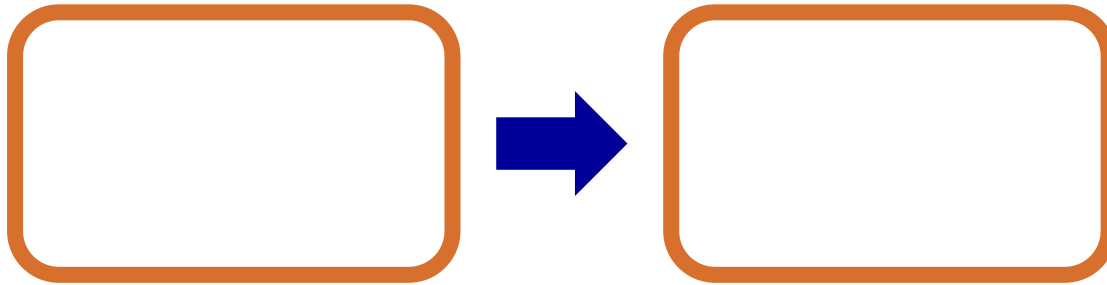
010 011

6 бит

101 3 бита

011 001 101 111 01

14 бит



**Информационный вес символа
= длине двоичного кода**



Например:

11101000111011011111010011101110111100001
110110011100000111101101110100011111111

содержит 40 бит информации

Сколько символов можно закодировать двумя битами?

1) 0 0

2) 0 1

3) 1 0

4) 1 1

Значит в алфавите мощностью **4 символа** информационный вес каждого символа - **2 бита**.

Сколько символов можно закодировать тремя битами?

1) 0 0 0

2) 0 0 1

3) 0 1 0

4) 0 1 1

5) 1 0 0

6) 1 0 1

7) 1 1 0

8) 1 1 1

Значит в алфавите мощностью **8 символов** информационный вес каждого символа - **3 бита**.

Единицы измерения информации

1 бит – минимальная единица измерения информации

КОМПЬЮТЕРНЫЙ АЛФАВИТ



- *русские буквы*
- *латинские буквы*
- *цифры (1, 2, 3, 4, 5, 6, 7, 8, 9, 0)*
- *математические знаки (+, -, *, /, ^, =)*
- *прочие символы («», №, %, <, >, :, ;, #, &)*

Компьютерный алфавит содержит 256 символов.

$256 = 2^8 \Rightarrow i=8$ битов

1 байт = 8 битов

1 байт - информационный вес символа алфавита мощностью 256.



Единицы измерения информационного объёма

1 байт = 2^3 бит = 8 бит

1 Кбайт = 2^{10} байт = 1024 байт

1 Мбайт = 2^{10} Кбайт = 1024 Кбайт

1 Гбайт = 2^{10} Мбайт = 1024 Мбайт

1 Тбайт = 2^{10} Гбайт = 1024 Гбайт

Единицы измерения информации

1 байт (*byte*) = 8 бит

1 Кбайт (килобайт) = **1024** байт = 2^{10} байт

1 Мбайт (мегабайт) = **1024** Кбайт

1 Гбайт (гигабайт) = **1024** Мбайт

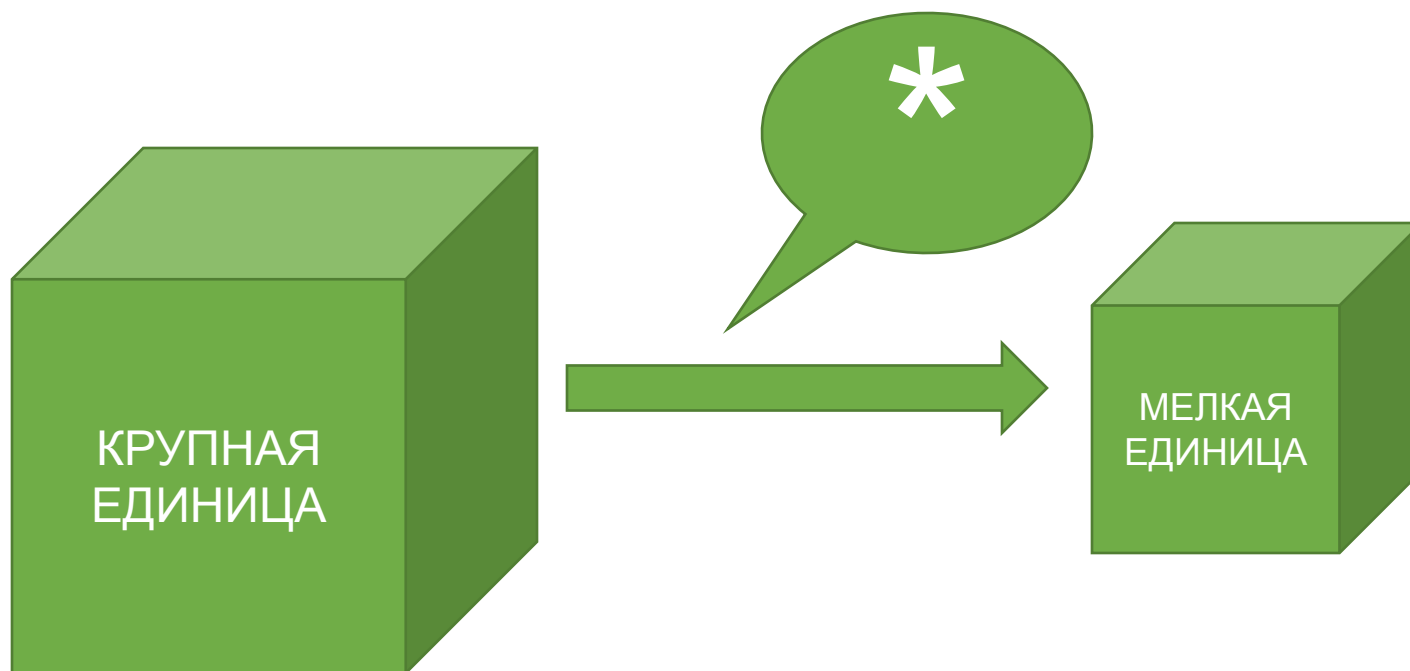
1 Тбайт (терабайт) = **1024** Гбайт

1 Пбайт (петабайт) = **1024** Тбайт

У
М
Н
О
Ж
Е
Н
И
Е

Д
Е
Л
Е
Н
И
Е

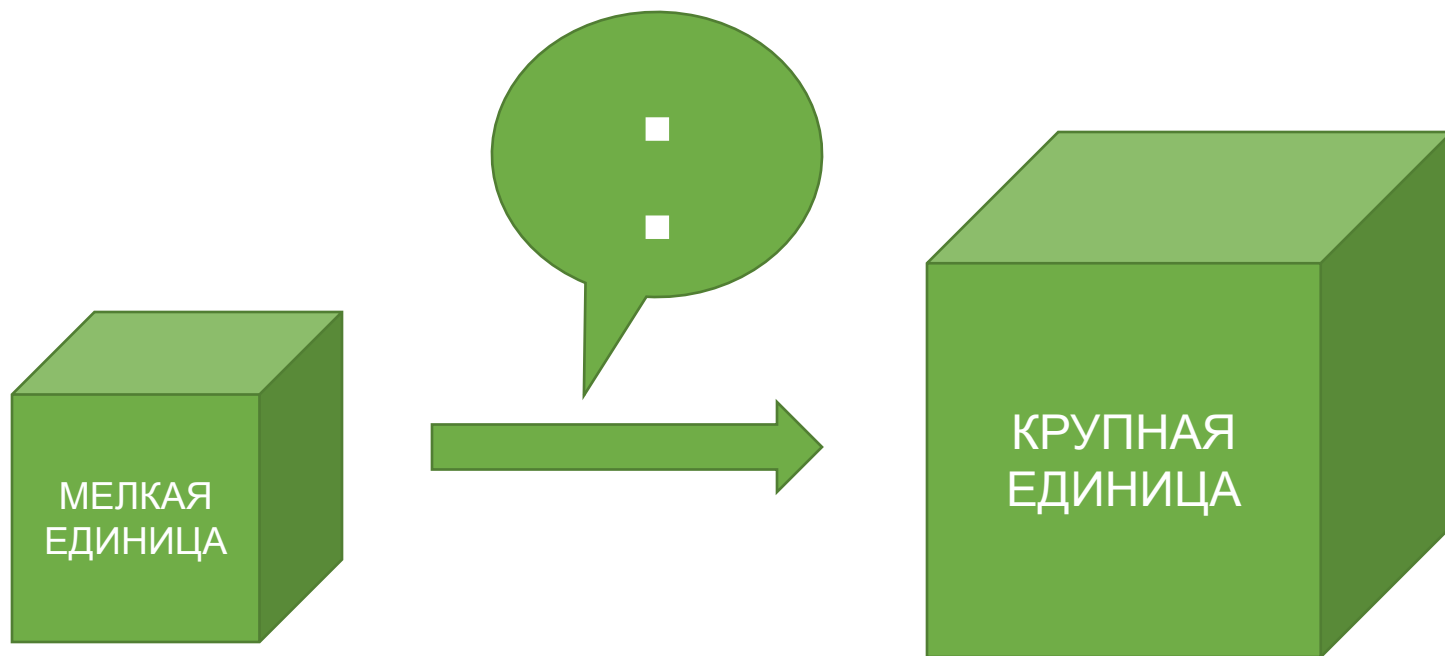
Перевод из крупных единиц в мелкие



Переведите в биты:

$$10 \text{ байт} = 80 \text{ бит}$$

Перевод из мелких единиц в крупные



Переведите в байты:

$$32 \text{ бита} = 4 \text{ байта}$$

Перевод в другие единицы

3 Кбайта = ~~3 байт~~ · 1024 байт = 3072 байта

15 байт = ~~15 бит~~ · 8 бит = 120 бит

2048 Кбайт = ~~2048 байт~~ · 1024 Мбайт = 2 Мб

1024 Мбайт = ~~1024~~ : 1024 Гб = 1 Гб

1 Кбайт = ~~1 бит~~ · 1024 · 8 бит = 8192 бита

байт



Зависимость между мощностью алфавита N и информационным весом символа i

N	2	4	8	16
i	1 бит	2 бита	3 бита	4 бита

Заметим, что $2 = 2^1$, $4 = 2^2$, $8 = 2^3$, $16 = 2^4$.

Информационный вес каждого символа, выраженный в битах (i), и мощность алфавита (N) связаны между собой формулой:

$$N = 2^i$$



Задачи

$$N = 2^i$$

№1

$$N=2$$

Найти i

Решение:

$$2=2^i$$

$$\underline{i=1 \text{ бит}}$$

№2

$$N=8$$

Найти i

Решение:

$$8=2^i$$

$$\underline{i=3 \text{ бита}}$$

№3

$$N=32$$

Найти i

Решение:

$$\underline{i=5 \text{ бит}}$$

№4

$$i=7 \text{ битов}$$

Найти N

Решение:

$$N=2^7$$

$$\underline{N=128 \text{ СИМВОЛОВ}}$$

Информационный объём сообщения

Информационный объём сообщения

Количество символов (знаков) в тексте

$$I = K \times i$$

Информационный вес одного символа (знака)



Алгоритм вычисления информационного объема сообщения

- 1) Определяем мощность алфавита N ;
- 2) Вычисляем информационный вес символа i (*бит*) :

$$N = 2^i$$

- 3) Вычисляем информационный объем сообщения:

$$I = K \cdot i$$

где K – количество символов в сообщении.

Информационный объём текста

Сообщение, записанное буквами из 64-символьного алфавита, содержит 2000 символов. Какой объём информации оно несёт?

Дано: $K=2000$, $N=64$

Найти: I - ?

Решение:

$$N=2^i$$

$$64=2^i$$

$i=6$ (бит) – информационный вес одного символа.

$$I=K*i$$

$$I=2000*6=12000 \text{ (бит)}$$

Ответ: 12000 бит.



Задача №1

Сколько килобайтов составит сообщение из **2048** символов **16-ти** символьного алфавита

Дано:

$$K=2048, N=16$$

Найти: I - Кб?

Решение:

$$N=2^i$$

$$16=2^i$$

$i=4$ (бита) – информационный вес одного символа.

$$I=K*i$$

$$I = \frac{2048 \cdot 4}{8 \cdot 1024} = 1 \text{ Кбайт}$$

Ответ: 1 Кбайт.

Задача №2

Информационный объём сообщения равен 3 Кб.
Информационный вес символа – 32 бита. Сколько символов содержит сообщение?

Дано: $I=3$ Кб,
 $i=32$ бита

Найти: K ?

Решение:

$$I=K*i$$

$$K=I/i$$

$$K = \frac{3 \cdot 1024 \cdot 8}{32} = \frac{3 \cdot 2^{10} \cdot 2^3}{2^5} = 3 \cdot 2^8 = 3 \cdot 256 = 768 (\text{символов})$$

Ответ: 768 символов

Основные понятия

! **Алфавитный подход** позволяет измерить объём информации не зависимо от её содержания.

! Каждый символ несёт некоторое количество информации и имеет **информационный вес (i)**.

! Минимальная единица измерения информации – **1 бит**.

! Мощность алфавита и информационный вес символа связаны отношением: **$N = 2^i$** .

! Информационный объём сообщения вычисляется по формуле: **$I = K * i$** .

! **1 байт = 8 бит**



! **Байты, килобайты (КБ), мегабайты (МБ), гигабайты (ГБ), терабайты (ТБ)** – единицы измерения информации. Каждая в **1024** раза больше предыдущей.

Содержательный подход к измерению информации



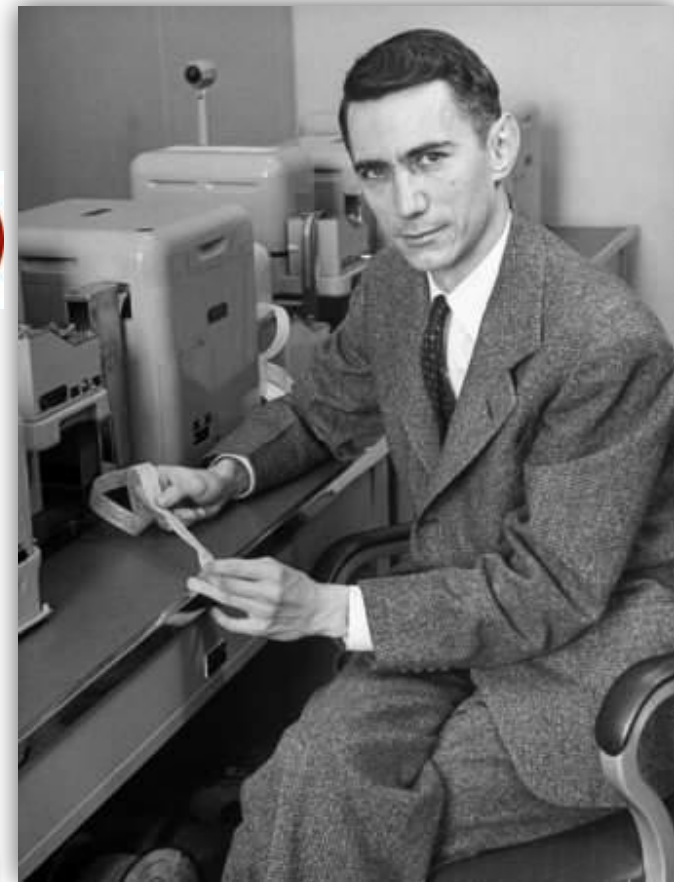
Содержательный подход к измерению информации

Информация — уменьшение неопределенности наших знаний.

Неопределенность знания некоторого события — это число возможных вариантов результата .



Равновероятные события — ни одно из них не имеет преимущества перед другими.



**Клод Элвуд
Шеннон**

Единица измерения информации

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



При бросании монеты неопределенность равна $2 \Rightarrow$ сообщение о том, что выпал «Орёл» несет **1 бит информации**

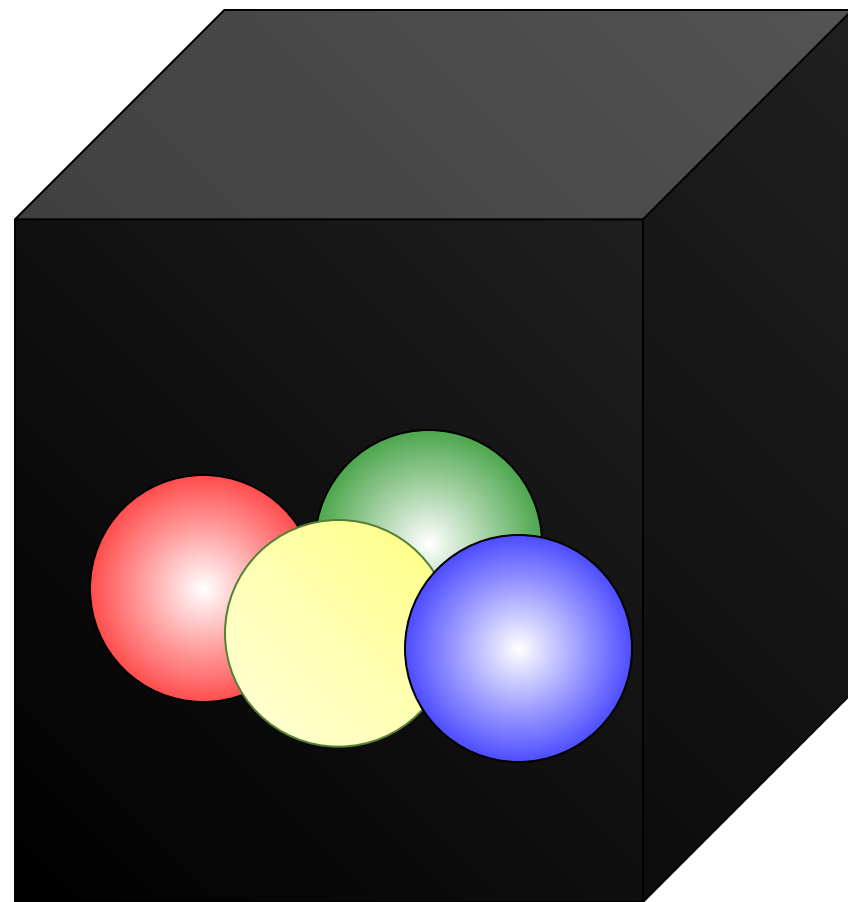


Книга лежит на одной из двух полок \Rightarrow неопределенность знаний равна 2



Сообщение о том, что **книга лежит на верхней полке** уменьшает неопределенность в два раза \Rightarrow данное сообщение несет **1 бит информации**

В ящике лежат 4 шара.
Какое количество
информации несет
сообщение
«Достали красный шар»





Формула Хартли (для равновероятных событий)

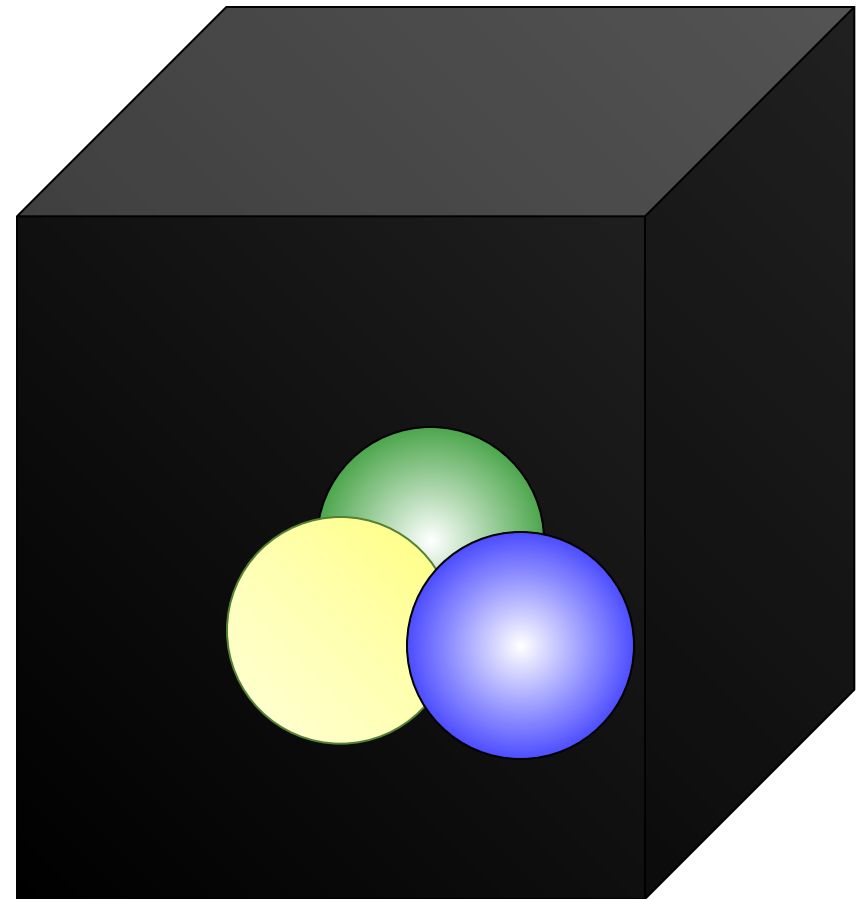
$$2^i = N$$



N – неопределённость знаний, т.е. число возможных результатов (вариантов сообщения)

i – количество информации в сообщении о том, что произошло одно из N равновероятных событий

$N=4$
 $2^i=4$
 $i=2$ (бита)



Задачи

1. В классе 32 ученика. Какое количество информации содержится в сообщении о том, что к доске пойдет Коля Сидоров?

$$N = 32$$

$$N = 2^i$$

$$32 = 2^i$$

$$i - ?$$

$$i = 5 \text{ бит}$$

2. Сообщение о том, что ваш друг живет на 5 этаже, несет 4 бита информации. Сколько этажей в доме?

$$i = 4 \text{ бита}$$

$$N = 2^i$$

$$N = 2^4$$

$$N - ?$$

$$N = 16 \text{ этажей}$$

Основные понятия

❗ **Содержательный подход** позволяет измерить объём информации в сообщении о том, что произошло одно из N событий.

❗ Сообщение о том, что произошло одно из двух равновероятных событий 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: **$N = 2^i$** .

