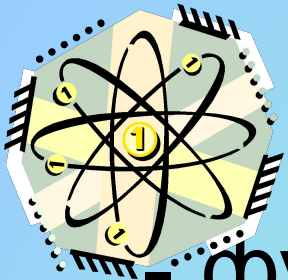


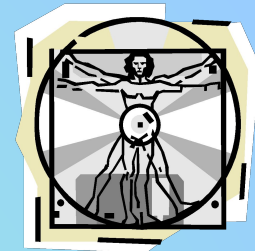
**Подходы к понятию
информации и измерению
информации.**

**Информационные объекты
различных видов.**

**Универсальность
дискретного (цифрового)
представления информации**

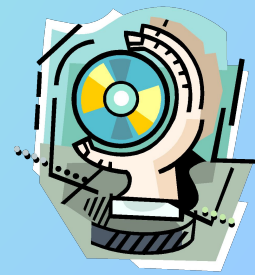
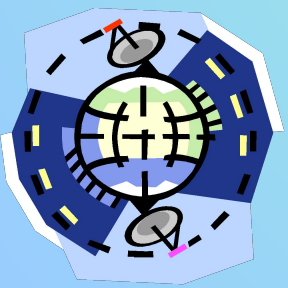


ИНФОРМАЦИЯ



— фундаментальное понятие науки,
поэтому определить его
исчерпывающим образом через
какие-то более простые понятия
НЕВОЗМОЖНО

С позиции человека информация – это содержание разных сообщений, это самые разнообразные сведения, которые человек получает из окружающего мира через свои органы чувств.



Подходы к понятию информации

Теория информации	Информация – содержание, заложенное в знаковые (сигнальные последовательности)
Кибернетика	Информация – содержание сигналов, передаваемых по каналам связи в системах управления
Нейрофизиология	Информация – содержание сигналов электрохимической природы, передающихся по нервным волокнам организма
Генетика	Информация – содержание генетического кода – структуры молекул ДНК, входящих в состав клетки живого организма
Философия	<i>Атрибутивная концепция:</i> Информация – всеобщее свойство (атрибут) материи
	<i>Функциональная концепция:</i> Информация и информационные процессы присущи только живой природе, являются ее функцией
	<i>Антропоцентрическая концепция:</i> Информация и информационные процессы присущи только человеку

Существует два подхода к измерению информации:

- **содержательный (вероятностный);**
- **объемный (алфавитный).**

Содержательный (вероятностный) подход к измерению информации

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

Главная формула информатики

**связывает между собой
количество возможных
информационных сообщений
N и количество информации I,
которое несет полученное
сообщение:**

$$N = 2^I$$

За **единицу количества информации** принимается такое количество информации, которое содержится в информационном сообщении, уменьшающем неопределенность знания в два раза.

Такая единица названа **бит**.

Бит – наименьшая единица измерения информации.

С помощью набора битов можно представить любой знак и любое число. Знаки представляются восьмизрядными комбинациями битов – байтами.

1 байт = 8 битов = 2^3 битов

Байт – это 8 битов, рассматриваемые как единое целое, основная единица компьютерных данных.

Рассмотрим, каково количество комбинаций битов в байте.

- Если у нас **две** двоичные цифры (бита), то число возможных комбинаций из них:

$2^2=4$: 00, 01, 10, 11

- Если **четыре** двоичные цифры (бита), то число возможных комбинаций:

$2^4=16$: 0000, 0001, 0010, 0011,
 0100, 0101, 0110, 0111,
 1000, 1001, 1010, 1011,
 1100, 1101, 1110, 1111

Так как в байте- **8 бит** (двоичных цифр),
то число возможных комбинаций битов в

байте:

$$2^8=256$$

Т.о., байт может принимать одно из 256 значений или комбинаций битов.

Для измерения информации
используются более крупные
единицы:

*килобайты, мегабайты,
гигабайты, терабайты и т.д.*

1 Кбайт = 1 024 байт

1 Мбайт = 1 024 Кбайт

1 Гбайт = 1 024 Мбайт

1 Тбайт = 1 024 Гбайт

**Проведем аналогию с единицами
длины:**

если 1 бит «соответствует» 1 мм,

то:

1 байт – 10 мм = 1 см;

1 Кбайт – 1000 см = 10 м;

1 Мбайт – 10 000 м = 10 км;

**1 Гбайт – 10 000 км (расстояние от
Москвы до Владивостока).**

**Страница учебника содержит
приблизительно 3 Кбайта информации;
1 газета – 150 Кбайт.**

Объемный (алфавитный подход)

к измерению информации



Алфавитный подход
позволяет измерить
количество
информации

в тексте, составленном
из символов
некоторого алфавита.

1

3

4

6

7

Σ

9

5

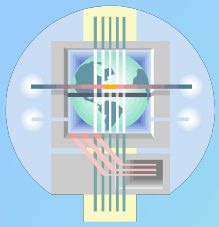
8

X

F

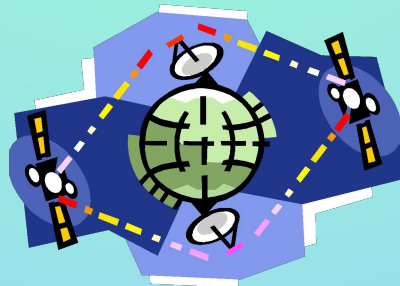
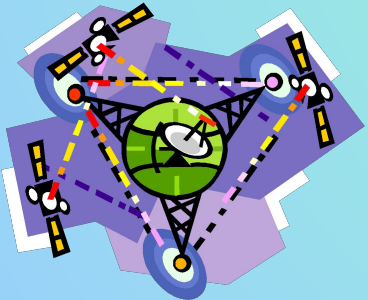
A

B



Алфавитный подход к измерению информации

Это объективный,
количественный метод для
измерения информации,
циркулирующей в
информационной технике.



Алфавит- множество символов,
используемых для представления
информации.

Мощность алфавита – число
символов в алфавите (его размер)
N.



Например, алфавит десятичной системы счисления – множество цифр- 0,1,2,3,4,5,6,7,8,9.

Мощность этого алфавита – 10.

Компьютерный алфавит, используемый для представления текстов в компьютере, использует 256 символов.

Алфавит двоичной системы кодирования информации имеет всего два символа- 0 и 1.

Алфавиты русского и английского языков имеют различное число букв, их мощности – различны.

Информационный вес символа
(количество информации в одном
символе), выраженный в битах (i), и
мощность алфавита (N) связаны между
собой формулой:

$$N = 2^i$$

где N – это количество знаков в алфавите знаковой системы или мощность

Тогда информационный вес символа:

$$i = \log_2 N$$

Информационная емкость знаков зависит от их количества в алфавите. Так, информационная емкость буквы в русском алфавите, если не использовать букву «ё», составляет:

$$32 = 2^I,$$

$$I = \ln 32 / \ln 2 = 3.46 / 0.69 = 5$$

$$I = 5 \text{ битов}$$

В латинском алфавите 26 букв. Информационная емкость буквы латинского алфавита также 5 битов.

Количество информации в сообщении или информационный объём текста- I_c , равен количеству информации, которое несет один символ- i , умноженное на количество символов K в сообщении:

$$I_c = K * i \quad \underline{\text{БИТ}}$$

Например, в слове «информатика» 11 знаков ($K=11$), каждый знак в русском алфавите несет информацию 5 битов ($I=5$), тогда количество информации в слове «информатика» $I_c=5 \times 11=55$ (битов).

С помощью формулы $N = 2^I$ можно определить количество информации, которое несет знак в двоичной знаковой системе: $N=2 \Rightarrow 2=2^I \Rightarrow \underline{2^1=2^I \Rightarrow I=1 \text{ бит}}$

Таким образом, *в двоичной знаковой системе 1 знак несет 1 бит информации. При двоичном кодировании объем информации равен длине двоичного кода.*

Чем большее количество знаков содержит алфавит знаковой системы, тем большее количество информации несет один знак.

Информационные объекты различных видов

Информационный объект – обобщающее понятие, описывающее различные виды объектов; это предметы, процессы, явления материального или нематериального свойства, рассматриваемые с точки зрения их информационных свойств.

Простые информационные объекты:

звук, изображение, текст, число.

Комплексные (структурированные) информационные объекты:

элемент, база данных, таблица, гипертекст, гипермедиа.

Информационный объект:

- обладает определенными потребительскими качествами (т.е. он нужен пользователю);
- допускает хранение на цифровых носителях;
- допускает выполнение над ним определенных действий путем использования аппаратных и программных средств компьютера.

Программы	Информационные объекты
Текстовые редакторы и процессоры	Текстовые документы
Графические редакторы и пакеты компьютерной графики	Графические объекты: чертежи, рисунки, фотографии
Табличные процессоры	Электронные таблицы
Пакеты мультимедийных презентаций	Компьютерные презентации
СУБД – системы управления базами данных	Базы данных
Клиент-программа электронной почты	Электронные письма, архивы, адресные списки
Программа-обозреватель Интернета (браузер)	Web-страницы, файлы из архивов Интернета

**Универсальность дискретного
(цифрового) представления
информации.**

Текстовая информация дискретна – состоит из отдельных знаков

Для обработки текстовой информации на компьютере необходимо представить ее в двоичной знаковой системе. Каждому знаку необходимо поставить в соответствие уникальный 8-битовый двоичный код, значения которого находятся в интервале от 00000000 до 11111111 (в десятичном коде от 0 до 255).

Дискретное (цифровое) представление графической информации

- **Изображение на экране монитора дискретно. Оно состоит из отдельных точек- пикселей.**
- **Пиксель — минимальный участок изображения, которому независимо можно задать цвет.**

В процессе дискретизации могут использоваться различные палитры цветов. Каждый цвет можно рассматривать как возможное состояние точки.

Количество цветов N в палитре и количество информации I , необходимое для кодирования цвета каждой точки, вычисляется по формуле:

$$N = 2^I$$

Пример

Наиболее распространенными значениями глубины цвета при кодировании цветных изображений являются 4, 8, 16 или 24 бита на точку.

Можно определить количество цветов в 24-битовой палитре: $N = 2^l = 2^{24} = 16\,777\,216$ бит.

Дискретное (цифровое) представление звуковой информации

Частота дискретизации звука — это количество измерений громкости звука за одну секунду.

Глубина кодирования звука — это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука.

Если известна глубина кодирования, то количество уровней громкости цифрового звука можно рассчитать по формуле

$$N = 2^l$$

Дискретное (цифровое) представление видеоинформации

ВИДЕОИНФОРМАЦИЯ -это сочетание звуковой и графической информации. Кроме того, для создания на экране эффекта движения используется дискретная технология быстрой смены статических картинок.

Способ уменьшения объема видео: первый кадр запоминается целиком (ключевой), а в следующих сохраняются только отличия от начального кадра (разностные кадры).

РЕШЕНИЕ задачи1

- Надо найти мощность алфавита N.

По условию задачи

$$I = 1,5 \text{ Кб} = 1.5 * 1024 * 8 = 12\ 288 \text{ бит}$$

$$I = i * k \text{ Значит, } i = I / k = 12\ 288 / 3072 = 4 \text{ бита}$$

Так как $N = 2^i$, то $N = 2^4 = 16$ символов.

ОТВЕТ: 16 СИМВОЛОВ



**Остальные задачи
попробуйте решить
самостоятельно!!!**

УСПЕХА!!!!



ЗАДАЧИ на дом

- 1. Информационное сообщение объёмом 1,5 Кбайта содержит 3072 символа. Сколько символов содержит алфавит, при помощи которого было записано это сообщение?
- 2. Сообщение занимает 2 страницы и содержит $1/16$ Кбайта информации. На каждой странице записано 256 символов. Какова мощность используемого алфавита?
- 3. Сколько килобайтов составляет сообщение, содержащее 12288 битов?

ЗАДАЧИ

4. Алфавит племени Мульти состоит из 8 букв. Какое количество информации несёт одна буква этого алфавита?
5. Сообщение, записанное буквами 64-х символьного алфавита, содержит 20 символов. Какой информационный объём оно несёт?

ЗАДАЧИ

6. Племя Мульти имеет 32-х символьный алфавит. Племя Пульти использует 64-х символьный алфавит. Вожди племён обменялись письмами. Письмо племени Мульти содержало 80 символов, а письмо племени Пульти – 70 символов. Сравните объёмы информации, содержащейся в письмах.

ЗАДАЧИ



7. Задача про марсиан!!!

Приветствие участникам олимпиады от марсиан записано с помощью всех символов марсианского алфавита:

ТЕВИРП!КИ!

Сколько информации оно несет?

ЗАДАЧИ

8. ДНК человека (генетический код) можно представить себе как некоторое слово в четырёхбуквенном алфавите, где каждой буквой помечается звено цепи ДНК, или нуклеотид.

Сколько информации (в битах) содержит ДНК человека, содержащий примерно $1,5 \cdot 10^{23}$ нуклеотидов?

