

**Информатика**

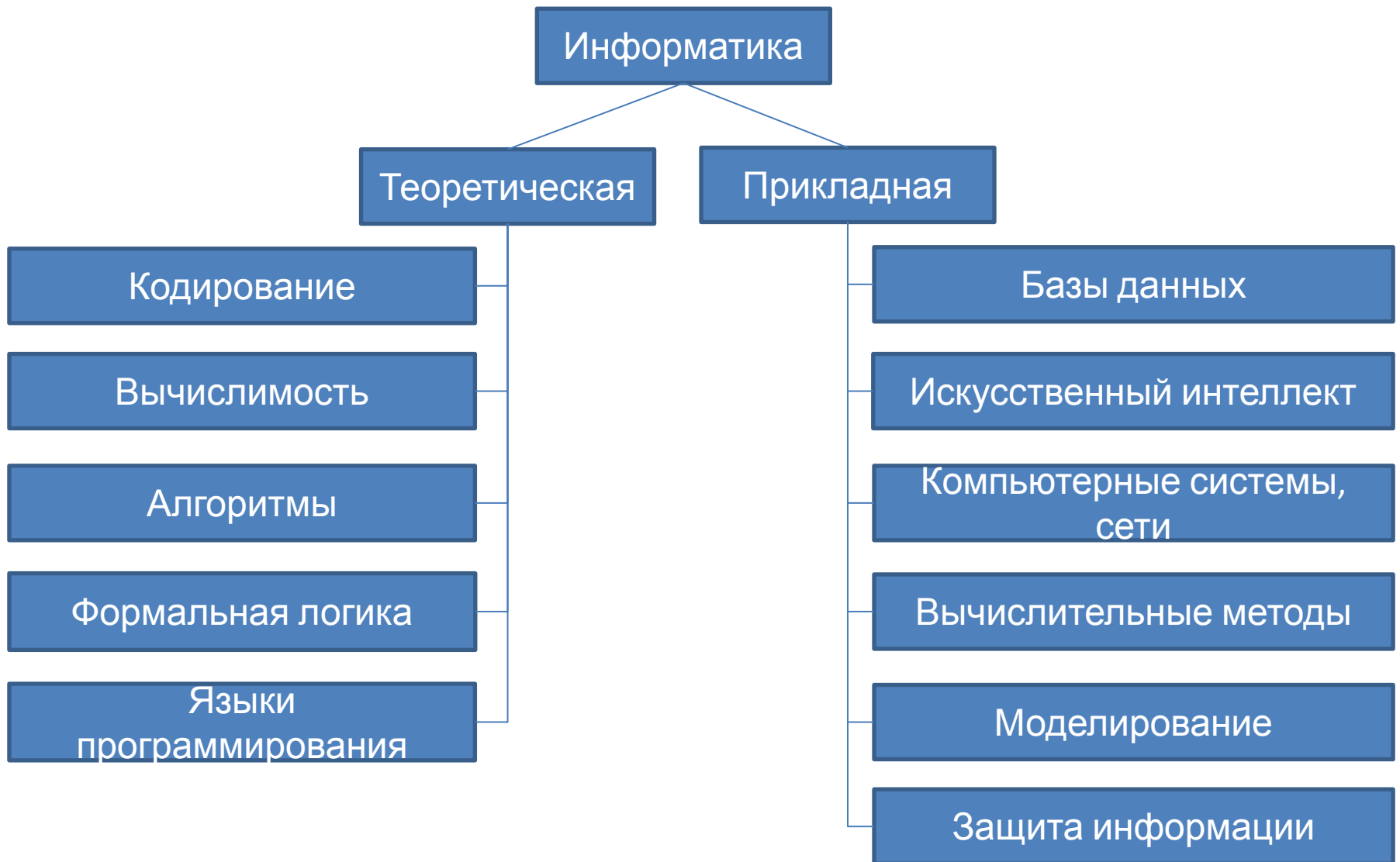
# ***Что такое информатика***

**Информатика – наука о методах работы с информацией:**

- Сбор**
- Хранение**
- Передача**
- Обработка (преобразование)**
- Анализ**

**Цель – обеспечение возможности принятия решения**

# Структура информатики



# ***Информационные технологии***

**ИТ** – совокупность процессов создания, развития и эксплуатации информационных систем (ИС).

**ИС** – система обработки информации, т.е. систематизированная (организованная) совокупность ресурсов (технических, финансовых, человеческих), предназначенная для обработки информации.

**ИВС** – информационно-вычислительная система, т.е. ИС, реализованная при помощи компьютерных технологий.

**САx** – Автоматизированная Система, ИВС специализированного инженерного назначения. Различают CAD/CAM/CAE/PDM, в русском объединенные аббревиатурой САПР (система автоматизированного проектирования).

# История информатики и ИТ

	Теория	Технологии
2700—2300 до н. э.	Позиционные системы счисления	Абак (Шумер)
150-100 до н. э.		Антикитерский механизм для расчета движения небесных тел
VII-XIII века	Десятичная система счёта, дроби, тригонометрические функции	Механические автоматы периода Исламского возрождения
XV век		Первый печатный станок
XVII век	Логарифмы Бинарная логика Математический анализ	Логарифмическая линейка Сумматор Паскаля Калькулятор Лейбница
XIX век	Первая программа Ады Лавлейс	Ткацкий станок, перфокарты Аналитическая машина Ч. Бэббиджа

# История информатики и ИТ

	Теория	Технологии
1931	Теоремы о неполноте	
1937	Машина Тьюринга	
1941		Первый тьюринг-полный компьютер Z3
1946	Архитектура фон Неймана	Первая ЭВМ «ЭНИАК»
1947		Первый транзистор
1948	Математическая теория связи Шеннона	
1957		FORTRAN
1962		Первые ИС
1969		ARPAnet
1971		Микропроцессор Intel 4040 Первая программа электронной почты
1981		ПК IBM 5150

# ***Информация***

**Информация – знания относительно фактов и понятий, которые в определённом контексте имеют конкретный смысл**

**Данные – форма представления информации**

**Информация – интерпретация данных**

**Кодирование – преобразование информации в форму, удобную для хранения и передачи. Чаще всего кодирование избыточно.**

# ***Измерение информации***

- **Определение требуемого размера хранилища данных**
- **Определение пропускной способности**
- **Сравнение и анализ**
- **Определение степени достоверности сообщения**



# *Измерение информации*

Три точки зрения на проблему определения количества информации:

- **Содержательная** – степень полезности информации
- **Алфавитная** – количество информации в сообщении, записанном при помощи алфавита
- **Вероятностная** – мера уменьшения неопределенности

# Количество информации

Измерение ← Сравнение ← Фиксация ← Изменение

- Для фиксации информации необходимо кодирование.  
Общий способ кодирования – составление сообщения из букв алфавита.

Пусть дан алфавит (множество букв)  $A$   
 $|A| = m$  – мощность алфавита (число букв)

Число различных сообщений  
 $N = m^n$ , где  $n$  – количество букв в сообщении

Количество информации  
 $I = n = \log_m N$

Сообщения несут информацию при  
 $m > 1$

Если  $m = 1$ , то  $N = 1^n = 1$ , следовательно  $I = \log 1 = 0$

# ***Количество информации***

- Минимальный алфавит имеет мощность 2  
 $|A| = m = 2$

**Например**

**{Да, Нет}      {0, 1}**

**Число различных сообщений**

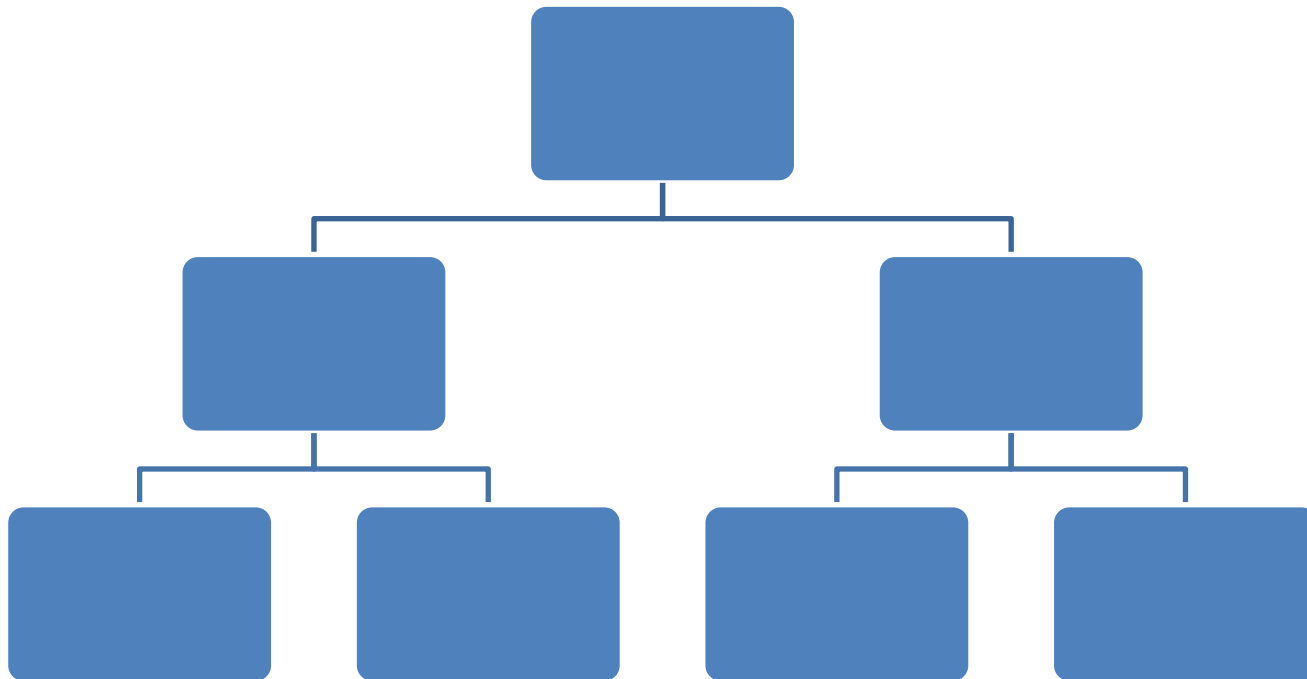
$$N = m^n = 2^n$$

**Количество информации, бит**

$$I = n = \log_2 N \quad (\text{формула Хартли})$$

# *Количество информации*

**С каждым полученным двоичным символом или при наступлении очередного события типа «да/нет» информационная неопределенность уменьшается вдвое.**



# ***Количество информации***

**Пусть имеется множество  $X$  независимых событий  $x_i$   
Если возможные события имеют различные вероятности  $p_i$ ,  
то для отдельного события**

$$2^I = \frac{1}{p_i}$$
$$I = \log_2 \frac{1}{p_i} = -\log_2 p_i$$

**Сумма вероятностей всех исходов (вероятность любого  
исхода)**

$$\sum_{i=1}^N p_i = 1$$

# Формула Шеннона

$$H(X) = - \sum_{i=1}^N p_i \log_2 p_i$$

*H* – среднее количество информации (информационная двоичная энтропия)

**Формула Хартли**  
**для равновероятных событий**  
**(появлений символа в сообщении)**

$$I = \log_2 N$$

# ***Единицы измерения количества информации***

1 бит - количество информации, содержащееся в сообщении, уменьшающем неопределенность знаний в два раза {0;1}

1 байт = 8 бит

1 Килобайт (Кб) [кибибайт, KiB] =  $2^{10}$  байт = 1024 байт

1 Мегабайт (Мб) [мебибайт, MiB] =  $2^{10}$  Кб =  $2^{20}$  байт

1 Гигабайт (Гб) [гибибайт, GiB] =  $2^{10}$  Мб =  $2^{20}$  Кб =  $2^{30}$  байт

1 Терабайт (Тб) [тебибайт, TiB] =  $2^{10}$  Гб =  $2^{20}$  Мб =  $2^{30}$  Кб =  $2^{40}$  байт

1 Петабайт (Пб) [пебибайт, PiB] =  $2^{50}$  байт

1 Эксабайт (Эб) [эксбибайт, EiB] =  $2^{60}$  байт



# ***Кодирование текста***

- наборы **ASCII** кодов – 1 байт на символ  
ANSI X3.4, ISO-8859  
кириллические кодовые страницы:  
windows-1251, CP866, KOI8-R, ISO-8859
- **Юникод** (Unicode) – 1-4 байта на символ  
UCS - *universal character set*  
UCS-2, UCS-4  
UTF - *unicode transformation format*  
UTF-8, UTF-16, UTF-32

# Пример 1

Последовательное уточнение параметров требуемого товара при покупке в магазине, снижает неопределенность знаний продавца пропорционально числу вариантов каждого параметра.

Допустим, происходит покупка телевизора Samsung, 42", UltraHD.

Пусть имеется 4 вида товаров: холодильники, телевизоры, стиральные машины, свч-печи.

В наличии телевизоры 4х компаний Sony, Samsung, LG, Panasonic.

Телевизоры имеют диагонали 27, 35, 42 и 49 дюймов.

Телевизоры с диагональю 42" имеют разрешение HD и UltraHD.

Общее количество информации, которое необходимо сообщить продавцу

$$I = \log_2 4 + \log_2 4 + \log_2 4 + \log_2 2 = 2 + 2 + 2 + 1 = 7 \text{ бит}$$

# Пример 2

Дан текст длиной 1200 символов.

Известно, что в тексте встречаются следующие символы:

Символ	Количество в тексте
Пробел	10%
А	20%
Б	30%
В	40%

Среднее количество информации в одном символе

$$H = -(0.1 \log_2 0.1 + 0.2 \log_2 0.2 + 0.3 \log_2 0.3 + 0.4 \log_2 0.4) \approx 1.8464 \text{ бит}$$

## ***Пример 3***

Текст книги набран на компьютере в текстовом редакторе в ASCII кодировке (1 байт на символ). Размер текста – 150 страниц, 40 строк на странице, 60 символов в строке.

Размер файла, содержащего книгу, составляет

$$1 \cdot 40 \cdot 60 \cdot 150 = 360\,000 \text{ байт} \approx 351.56 \text{ Кб}$$