Основы математической обработки информации

План лекции

1. Эволюция ЭВМ
2. Представление текстовой и числовой информации в ПК
3. Обработка данных на компьютере
4. Математика как наука
5. Аксиоматический метод
6. Математическая обработка данных и моделирование
7. Математика в филологии и лингвистике

1. Эволюция ЭВМ. Человек и компьютер

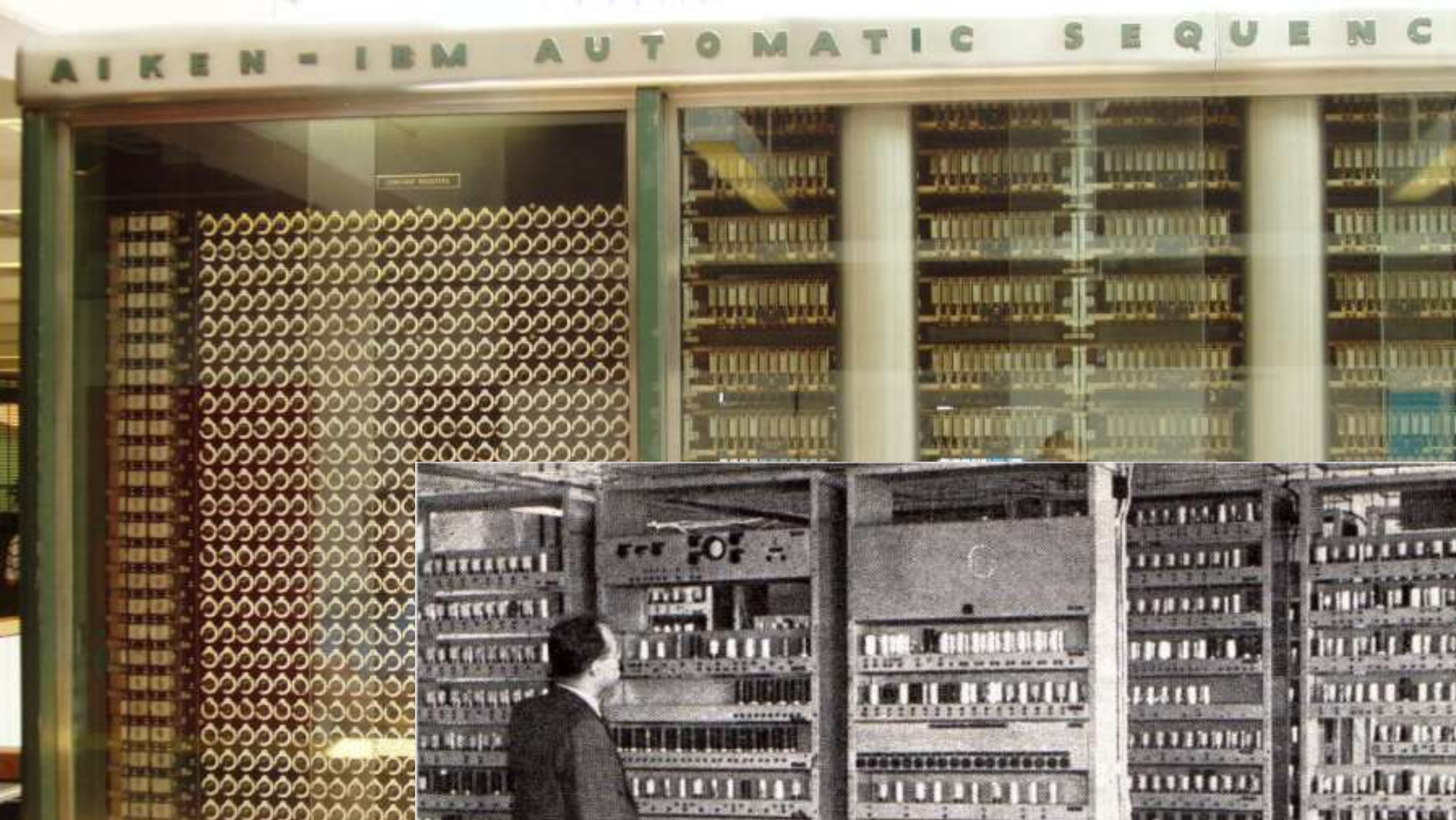
С давних времен люди стремились облегчить свой труд. С этой целью создавались орудия труда, машины и механизмы, усиливающие физические возможности человека.

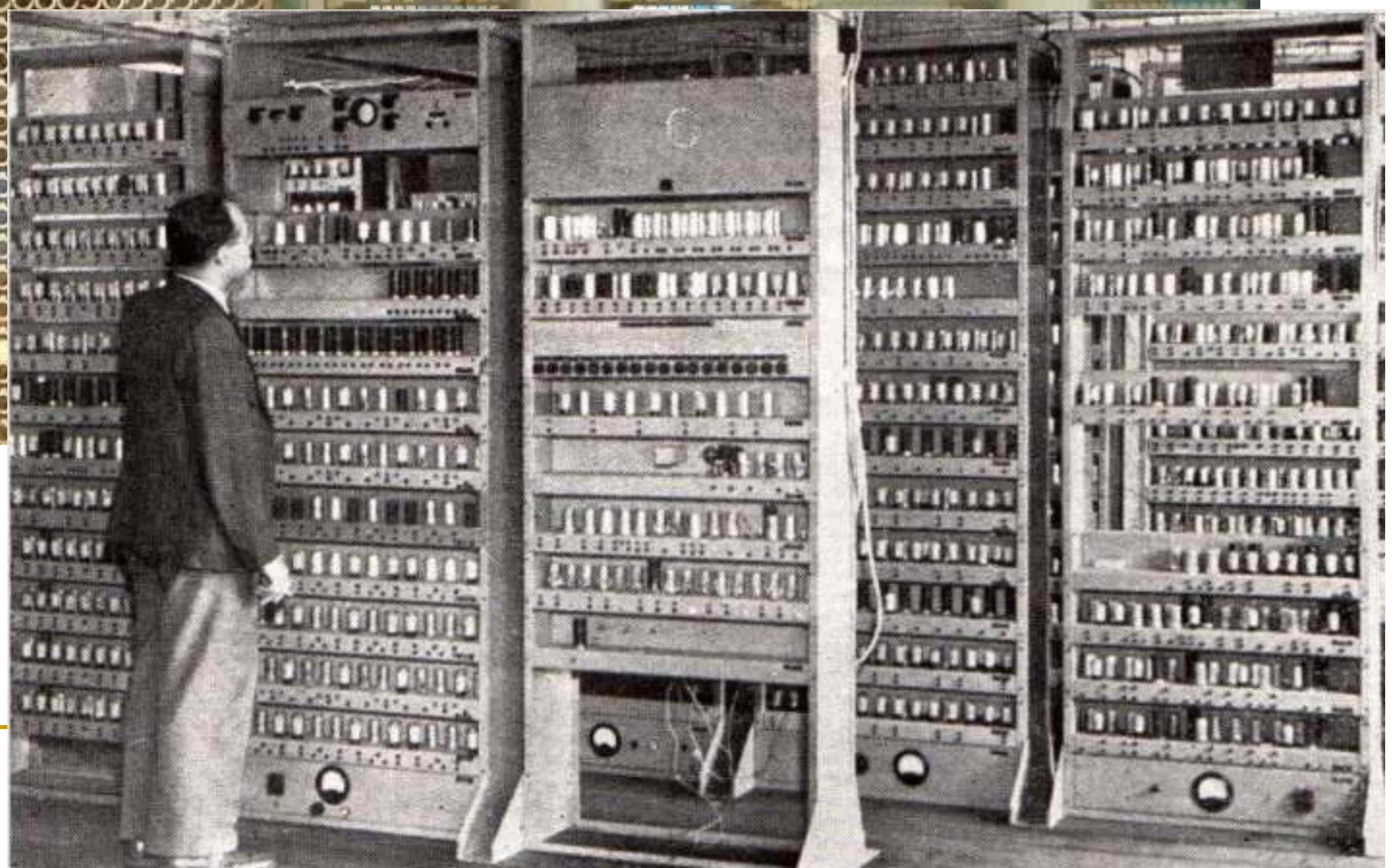
Компьютер был изобретен в середине XX века для усиления умственных возможностей человека, т. е. для работы с информацией.

- Первый компьютер создан **1943 г.** американцем Говардом Эйкеном **Марк-1 на электромеханических реле** на одном из предприятий фирмы IBM

ЭВМ была создана в **США в 1945 году**. Эта машина называлась ENIAC. Конструкторы Дж.Моучли и Дж.Эккерт. Весила она 27 тонн, скорость счета этой машины превосходила скорость «Марк-1» в тысячу раз. «Гигантский мозг» был использован для вычислений, связанных с созданием водородной бомбы.

В нашей стране первая ЭВМ была создана в 1951 году. Называлась она МЭСМ и была лучшей в мире, позднее появилась БЭСМ





Mark 1

ENIAK

Есть ли в природе прототип у компьютера? Да! Таким прототипом является сам человек. Только изобретатели стремились передать компьютеру не физические, а интеллектуальные возможности человека.

По своему назначению компьютер - универсальное техническое средство для работы человека с информацией.

По принципам устройства компьютер - это модель человека, работающего с информацией.

Эволюция ЭВМ и возможности компьютера в области обработки данных

В 50-60-е годы первые ЭВМ могли только вычислять. Процесс обработки информации состоял в операциях над числовыми данными.

В 70-е годы компьютер "научился" работать с текстом. Пользователь получил возможность редактировать и форматировать текстовые документы.

В 80-е годы появились компьютеры, способные работать с графикой. Сейчас компьютерная графика широко используется в бизнесе (построение диаграмм, графиков и так далее), в компьютерном моделировании, при подготовке презентаций, при создании Web-сайтов, в рекламе на телевидении, в анимационном кино и так далее.

В 90-е годы компьютеры начали обрабатывать звук. Любой пользователь современного ПК может воспользоваться стандартными приложениями Windows для прослушивания, записи и редактирования звуковых файлов.

Для того чтобы числовая, текстовая, графическая и звуковая информация могли обрабатываться на компьютере, они должны быть представлены в форме данных. Данные хранятся и обрабатываются в компьютере на машинном языке, то есть в виде последовательностей нулей и единиц.

Информация, представленная в компьютерной форме (на машинном языке) и обрабатываемая на компьютере, называется **данными**.

2.Представление информации в ЭВМ

В настоящее время во всех вычислительных машинах информация представляется с помощью электрических сигналов. При этом возможны две формы представления численного значения какой-либо переменной:

- в виде одного сигнала — например, электрического напряжения, которое сравнимо с величиной *X* (аналогично ей). Например, при *X* = 2003 единиц на вход вычислительного устройства можно подать напряжение 2,003 В
- в виде нескольких сигналов — нескольких импульсов напряжений, (например, при *X*, равном 195 единицам, на вход вычислительного устройства можно подать три импульса напряжением 1 В, 9 В и 5 В).

Первая форма представления информации (с помощью сходной величины - аналога) называется **аналоговой**, или непрерывной. Величины, представленные в такой форме, могут принимать любые значения в определенном диапазоне. Количество значений, которые может принимать такая величина, бесконечно велико. Отсюда названия — непрерывная величина и непрерывная информация.

Вторая форма представления информации называется **дискретной**. Такие величины, принимающие не все возможные, а лишь вполне определенные значения, называются дискретными (прерывистыми). В отличие от непрерывной величины количество значений дискретной величины всегда будет конечным.

Три класса вычислительных машин

В зависимости от формы представления информации вычислительные машины делятся на три класса:

• **цифровые вычислительные машины (ЦВМ),**обрабатывающие информацию, представленную в цифровой форме;

• **аналоговые вычислительные машины (АВМ),**обрабатывающие информацию, представленную в виде непрерывно меняющихся значений какой-либо физической величины (электрического напряжения, тока и т. д.);

• **гибридные вычислительные машины (ГВМ),**содержащие как аналоговые, так и цифровые вычислительные устройства.

АВМ предназначены в основном для решения задач, описываемых системами дифференциальных уравнений: исследования поведения подвижных объектов, моделирования процессов и систем, решения задач параметрической оптимизации и оптимального управления.

Аналоговый компьютер

Гибридный суперкомпьютер СКИФ МГУ

Цифровой компьютер

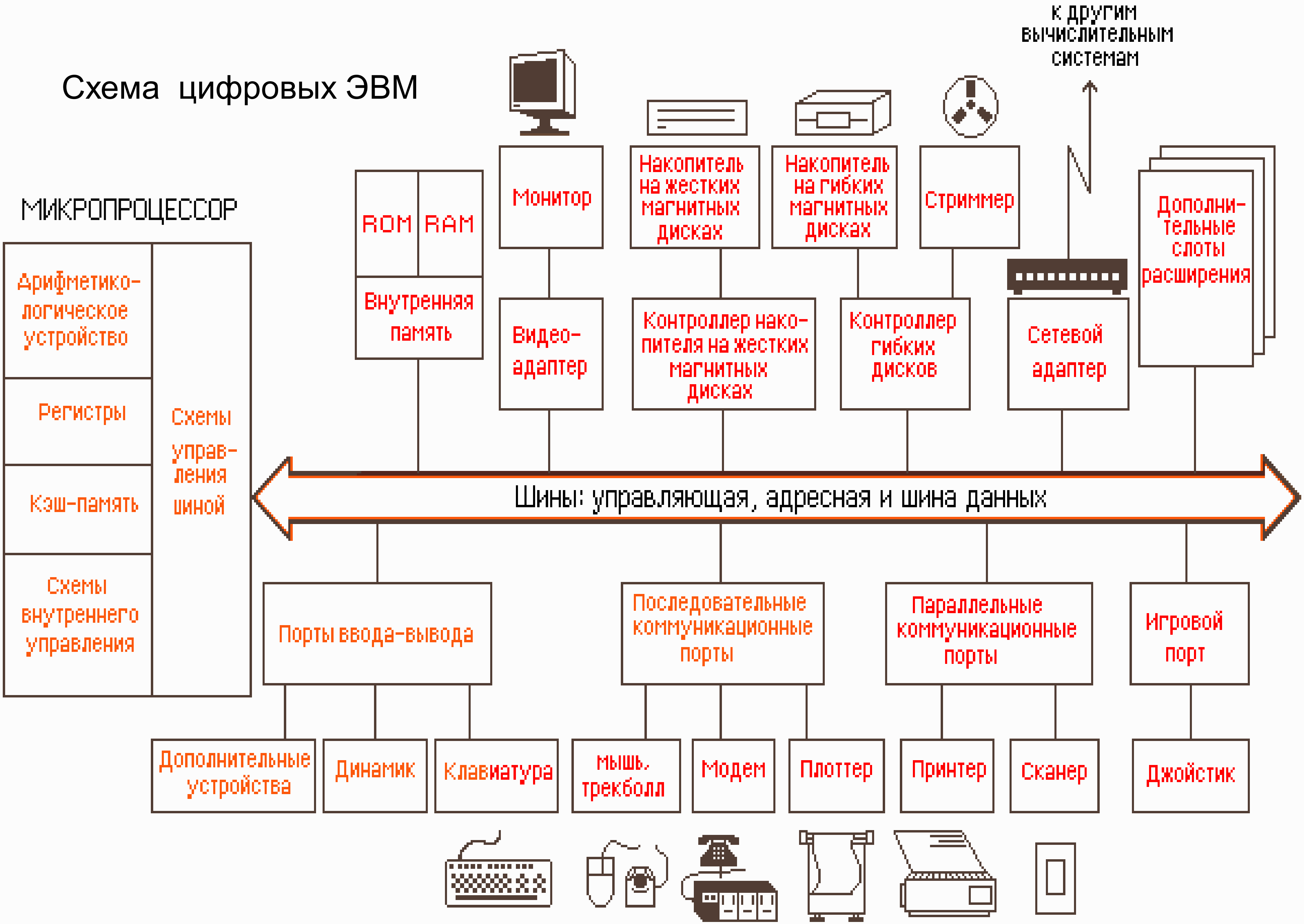
Аналоговые и гибридные ЭВМ

- В основу функционирования АВМ заложен принцип аналогии, заключающийся в том, что входной физической величине в машине ставится в соответствие величина другой физической природы, но меняющаяся по тому же физическому закону, что и входная. Так, при использовании в качестве модели (аналога) объекта электронных цепей каждой физической переменной величине (давление, температура, электромагнитное поле и т. д.) ставится в соответствие определенная переменная величина электронной цепи. В отличие от ЦВМ, точность которых определяется их разрядностью, точность вычислений на АВМ ограничена и характеризуется качеством изготовления элементной базы и основных узлов. В то же время, для целого класса задач скорость решения задач на АВМ может быть значительно выше, чем на ЦВМ. Это объясняется параллельным принципом решения задач на АВМ, при котором решение получается мгновенно и одновременно во всех точках модели. Данная особенность обусловливает использование АВМ в замкнутых системах автоматического регулирования и для решения задач в режиме реального времени.
- В гибридных вычислительных машинах есть и цифровые устройства для управления и выполнения логических операций, и аналоговые устройства — для решения дифференциальных уравнений.

Таблица 1 - Качественные параметры различных компьютеров

	Аналоговый	Цифровой	Гибридный
Форма сигналов	Аналоговая	Цифровая	Цифровая и аналоговая
Быстродействие	Высокое	Среднее	Среднее
Точность	Низкая	Высокая	Высокая
Технология программирования	Ручная	Автоматизирован ная	Автоматизирован ная
Совместимость с др. ЭВМ	Аналоговый	Цифровой	Цифровой и аналоговый

Схема цифровых ЭВМ



Цифровые компьютеры

- В настоящее время подавляющее большинство компьютеров являются цифровыми, далее слова *«компьютер»* или «ЭВМ» будем употреблять в значении *«цифровой компьютер»*. Для обработки аналоговой информации на таком компьютере ее сначала преобразуют в цифровую форму.
- В ЭВМ последовательность действий, составляющих задачу обработки информации, называют процессом.
- **Процесс** определяется программой, состоящей из машинных команд, набором данных, а также дескриптором процесса, который описывает текущее состояние любого выделенного процессу ресурса ЭВМ.
- Каждый сеанс пользователя с вычислительной системой, например ввод-вывод данных в ЭВМ, также является процессом. В вычислительной системе может одновременно существовать произвольное число процессов, поэтому между ними возможна конкуренция за обладание ресурсами, в первую очередь временем процессора — основного вычислительного устройства.
- Это приводит к необходимости управления процессами и их планирования. Для этого служат операционные системы (ОС)

2.1. Представление текста в ЭВМ

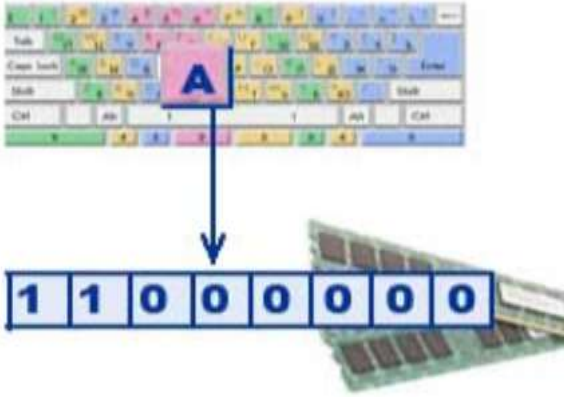
Представление информации в виде текста стало одним из первых доступных для обработки с помощью ЭВМ и до сих пор является наиболее универсальным.

Текстом считают такое представление информации, в котором она представлена в виде записи слов некоторого языка и доступна для чтения человеком.

Язык для текста описывается допустимым набором символов или алфавитом.

Поскольку компьютер работает только с двоичным кодом, то для записи и обработки требуется взаимно-однозначно сопоставить символы и двоичные коды.

Кодировка - правило сопоставления двоичного кода и символа –буквы алфавита.



ЕДИНИЦЫ ИЗМЕРЕНИЯ ИНФОРМАЦИИ

Бит (bit) - это самая наименьшая единица измерения количества информации: ноль или единица.
Байт (byte) - единица хранения и обработки цифровой информации; совокупность битов, обрабатываемая компьютером одномоментно. В современных вычислительных системах байт состоит из восьми битов и, соответственно, может принимать одно из 256 различных значений.

1 Бит **8 Бит или 1 Байт** **16 Бит или 2 Байта**

1 Килобайт	= 1024 байт	= 2^{10}	= 1024
1 Мегабайт	= 1024 Кбайт	= 2^{20}	= 1 048 576
1 Гигабайт	= 1024 Мбайт	= 2^{30}	= 1073 741 824
1 Терабайт	= 1024 Гбайт	= 2^{40}	= 1 099 511 627 776
1 Петабайт	= 1024 Тбайт	= 2^{50}	= 1 125 899 906 842 624
1 Эксабайт	= 1024 Пбайт	= 2^{60}	= 1 152 921 504 606 846 976
1 Зеттабайт	= 1024 Эбайт	= 2^{70}	= 1 180 591 620 717 411 303 422
1 Йоттабайт	= 1024 Збайт	= 2^{80}	= 1 208 925 819 614 629 174 706 176

Кодирование текста

Для хранения двоичного кода одного символа обычно выделяется 1 байт = 8 бит, иногда 2 байта или 16 бит.

Каждый бит принимает значение 0 или 1, количество их возможных сочетаний в байте равно $2^8 = 256$. Значит, с помощью 1 байта можно получить 256 разных двоичных кодовых комбинаций и отобразить с их помощью 256 различных символов.

С помощью 2 байт или 16 бит можно получить $2^{16}=65536$ разных двоичных кодов и отобразить с их помощью 65536 различных символов.

Такое количество символов вполне достаточно для представления текстовой информации, включая прописные и заглавные буквы русского и латинского алфавита, цифры, знаки, графические символы и т.д.

Кодирование заключается в том, что каждому символу ставится в соответствие уникальный десятичный код от 0 до 255 или соответствующий ему двоичный код от 00000000 до 11111111.

Таким образом, человек различает символы по их начертанию, а компьютер — по их коду.

Кодировка ASCII

Первый широко распространенный стандарт кодирования — таблица кодировки ASCII

American Standard Code for Information Interchange,

американский стандартный код для обмена информацией был разработан в 1963 году. В системе ASCII закреплены две таблицы кодирования — базовая и расширенная. Базовая таблица закрепляет значения кодов от 0 до 127, а расширенная относится к символам с номерами от 128 до 255.

Первые 33 кода (с 0 до 32) соответствуют **операциям** (перевод строки, ввод пробела и т. д.).

Коды с 33 по 127 являются интернациональными и соответствуют символам **латинского алфавита, цифрам, знакам** арифметических операций и знакам препинания.

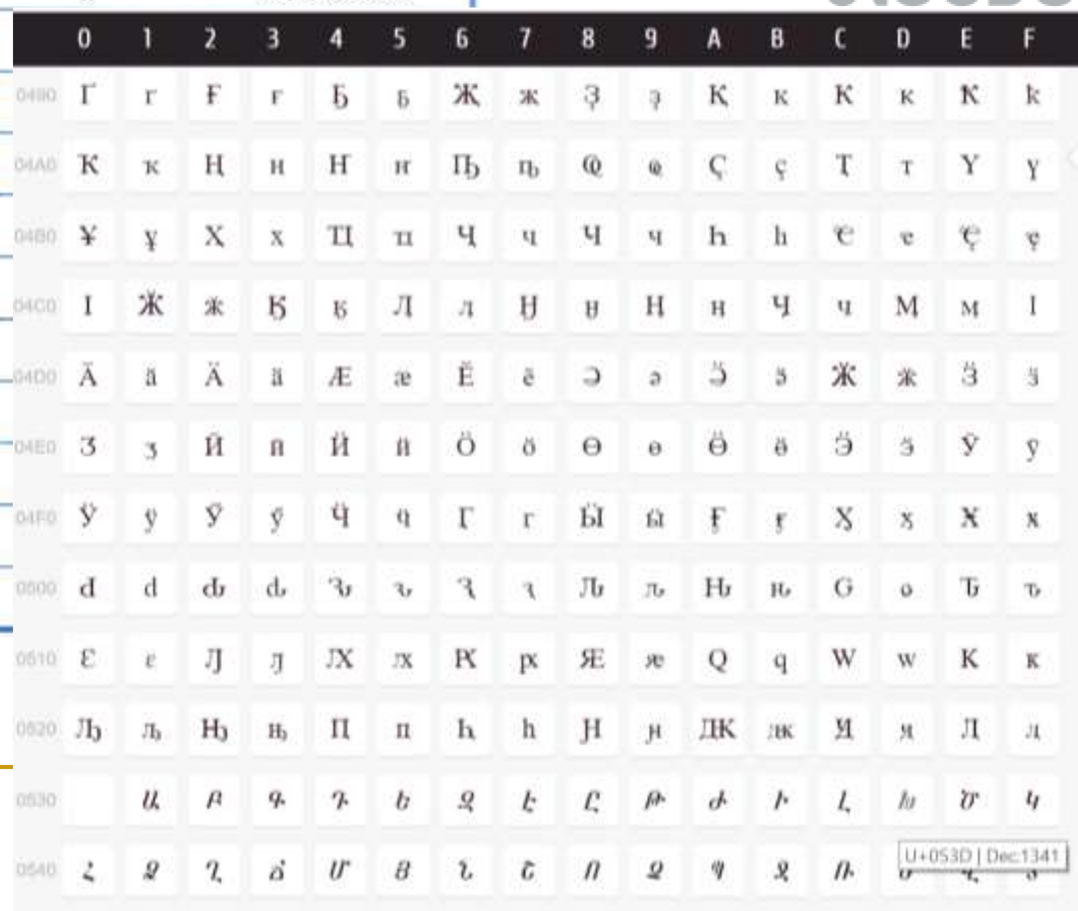
Коды с 128 по 255 служат для **национальных алфавитов.**

Каждый вариант этой второй половины (расширенной таблицы) исходной таблицы получил название “кодовой страницы” языка (code page).

Кодовые таблицы

Кодовая таблица - таблица соответствий символов и их компьютерных кодов. В РФ распространены следующие кодировки: WIN1251 (Windows), KOI-8

65	A	01000001	78	N	01001110
66	B	01000010	79	O	01001111
67	C	01000011	80	P	01010000
68	D	01000100	81		
69	E	01000101	82		
70	F	01000110	83		
71	G	01000111	84		
72	H	01001000	85		
73	I	01001001	86		
74	J	01001010	87		
75	K	01001011	88		
			89		
			90		



Кодировки русского алфавита

Windows-1251 — в системах Windows;

Семейство кодовых страниц КОИ-8;

ISO-8859-5 (Unix)

MacCyrillic для ПК Макинтош

Unicode - уникальный код для любого символа, независимо от платформы, независимо от программы, независимо от языка

- Unicode) — стандарт кодирования символов, включающий в себя знаки почти всех письменных языков мира[3]. является доминирующим в Интернете.
- Стандарт предложен в 1991 году некоммерческой организацией «Консорциум Юникода». Применение этого стандарта позволяет закодировать очень большое число символов из разных систем письменности: в документах, закодированных по стандарту Юникод, могут соседствовать китайские иероглифы, математические символы, буквы греческого алфавита, латиницы и кириллицы, символы музыкальной нотной нотации, при этом становится ненужным переключение кодовых страниц[6].
- Стандарт состоит из двух основных частей: универсального набора символов (англ. Universal character set, **UCS**) и семейства кодировок (англ. Unicode transformation format, **UTF**). Универсальный набор символов перечисляет допустимые по стандарту Юникод символы и присваивает каждому символу код в виде неотрицательного целого числа, записываемого обычно в шестнадцатеричной форме с префиксом U+, например, U+040F. Семейство кодировок определяет способы преобразования кодов символов для передачи в потоке или в файле.

Зачем нужны кодировки текстов?

Символы на экране компьютера формируются на основе **векторных форм** (символов, которые находятся в файлах со шрифтами на компьютере) и **двоичного кода**, который позволяет выбрать из этого набора именно тот символ, который нужно отобразить.

За векторные формы отвечают шрифты, а за кодирование текста отвечает операционная система и другие программы.

Любой текст на компьютере представляет собой набор байтов, а в каждом байте закодирован один единственный символ этого самого текста.

Программа, отображающая этот текст на экране (текстовый редактор, браузер и т.п.), при нажатии клавиши считывает кодировку очередного знака и ищет соответствующую ему векторную форму в файле шрифта, который подключен для отображения данного текстового документа.

2.2.Представление числовой информации в компьютере

Системы счисления

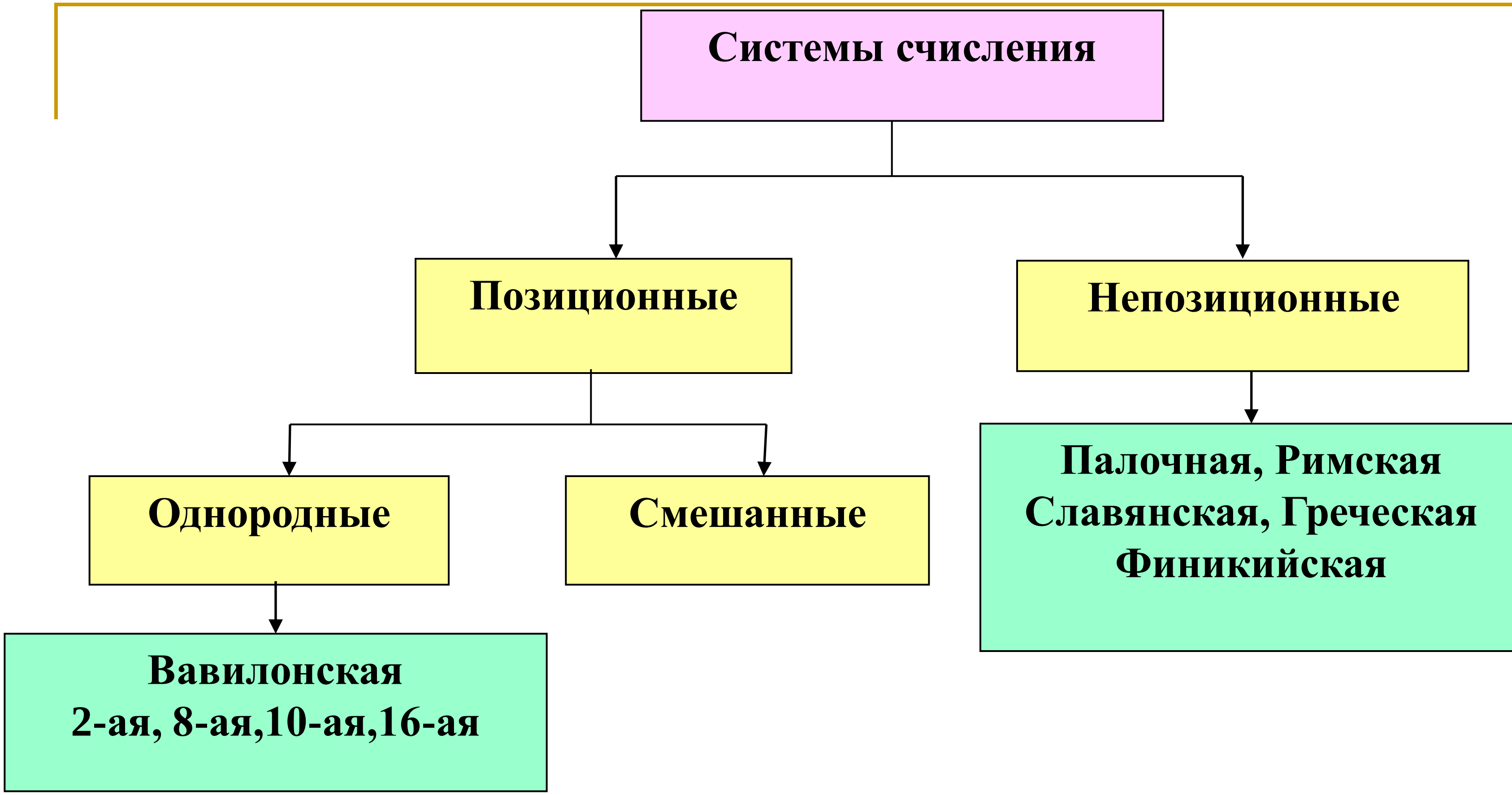
Система счисления – это способ представления чисел с помощью символов, имеющих определенные количественные значения.

Основание позиционной системы счисления — количество различных цифр, используемых для изображения чисел в данной системе счисления.

Первые системы записи чисел были непозиционными.

Существовали единичная система счисления – **палочная**, в которой любое число записывалось с помощью последовательности знаков - палочек. Чем больше было число, тем больше палочек нужно было нарисовать.

Непозиционными системами были также славянская, греческая и финикийская. В них цифры обозначались буквами алфавита.



Позиционной системой счисления называется такая система записи чисел, в которой значение числа зависит от позиции каждой цифры в нем.

Позиционными системами счисления являются двоичная, восьмеричная, десятичная, шестнадцатеричная системы.

В **непозиционной системе** счисления величина каждой цифры не зависит от ее местонахождения в числе.

Представление числовой информации в компьютере

Для записи чисел в ЭВМ используются системы счисления с основанием, являющимся целой степенью числа 2, а именно:

- **двоичная** (используются цифры 0, 1);
- **восьмеричная** (используются цифры 0, 1, ..., 7);
- **шестнадцатеричная** (для первых целых чисел от нуля до девяти используются цифры 0, 1, ..., 9, а для следующих чисел — от десяти до пятнадцати — в качестве цифр используются символы A, B, C, D, E, F).

Правило записи любого числа в позиционной системе счисления

Пусть N любое число, где цифры обозначены буквами $a_m - a_s$.

$$N = a_m\, a_{m-1}\, a_{m-2} \ldots a_1\ a_0\ a_{-1}\ a_{-2} \ldots a_s,$$

Пусть q – основание системы счисления, тогда формула перевода любого числа из одной позиционной системы счисления в другую:

$$\mathbf{N = a_m * q^m + a_{m-1} * q^{m-1} + a_{m-2} * q^{m-2} + a_1 * q^1 + a_0 * q^0 +}$$

$$\mathbf{+ a_{-1} * q^{-1} + a_{-2} * q^{-2} + \ldots + a_s * q^{-s}}$$

При переводе числа из любой системы счисления в десятичную систему счисления нужно каждый символ этого числа умножить на основание системы счисления, в которой записано это число, в степени соответствующей положению символа в записи числа и все произведения сложить.

2 1 0 -1

$$\mathbf{3\,7\,5,\,2_{10} = 3*10^2 + 7*10^1 + 5*10^0 + 2*10^{-1}}$$

Перевод чисел из любой системы счисления в десятичную систему счисления

1) переведём **101100, 10112** из двоичной системы в десятичную систему счисления:

$$101100,101_2 = 1*2^5 + 0*2^4 + 1*2^3 + 1*2^2 + 0*2^1 + 0*2^0 + 1*2^{-1} + 0*2^{-2} + 1*2^{-3} =$$

$$= 32 + 0 + 8 + 4 + 0 + 0 + 0{,}5 + 0 + 0{,}125 = 44{,}625_{10}$$

2) переведём **375, 6248** из восьмеричной системы в десятичную систему счисления:

$$375,624_8 = 3*8^2 + 7*8^1 + 5*8^0 + 6*8^{-1} + 2*8^{-2} + 4*8^{-3} =$$

$$= 192 + 56 + 5 + 0{,}75 + 0{,}03125 + 0{,}00781835938 = 253{,}78906835938_{10}$$

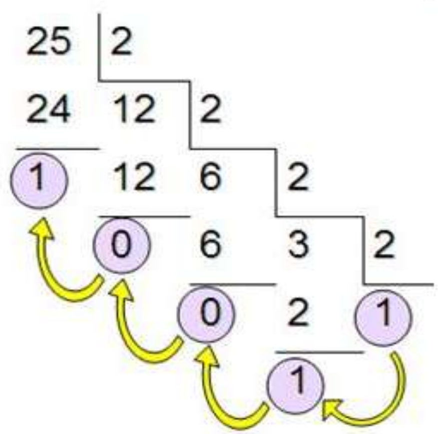
3) переведём **ACF, 5D16** из шестнадцатеричной системы в десятичную систему:

$$ACF,5D_{16} = 10*16^2 + 12*16^1 + 15*16^0 + 5*16^{-1} + 13*16^{-2} =$$

$$= 256 + 192 + 15 + 0{,}3125 + 0{,}050775 = 463{,}363275_{10}$$

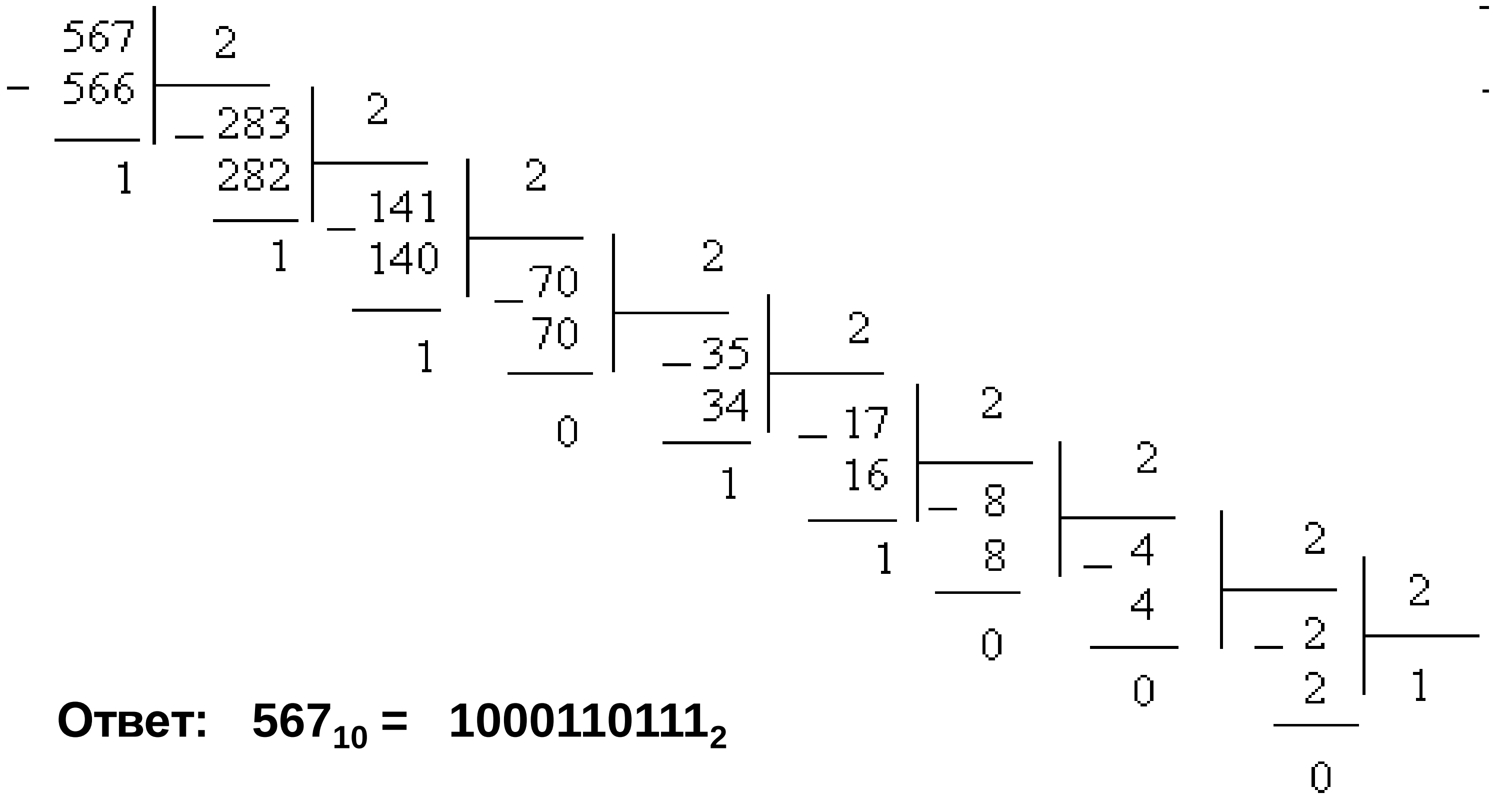
Перевод чисел из десятичной системы счисления в любую другую позиционную систему счисления

Правило перевода чисел: для перевода числа из одной позиционной системы в другую нужно делить число на основание новой системы счисления, затем частное от деления тоже делить на основание новой системы до тех пор, пока оно не станет меньше основания новой системы счисления. Выписав последнее частное и все остатки, начиная с последнего, получим представление этого числа в другой системе счисления.



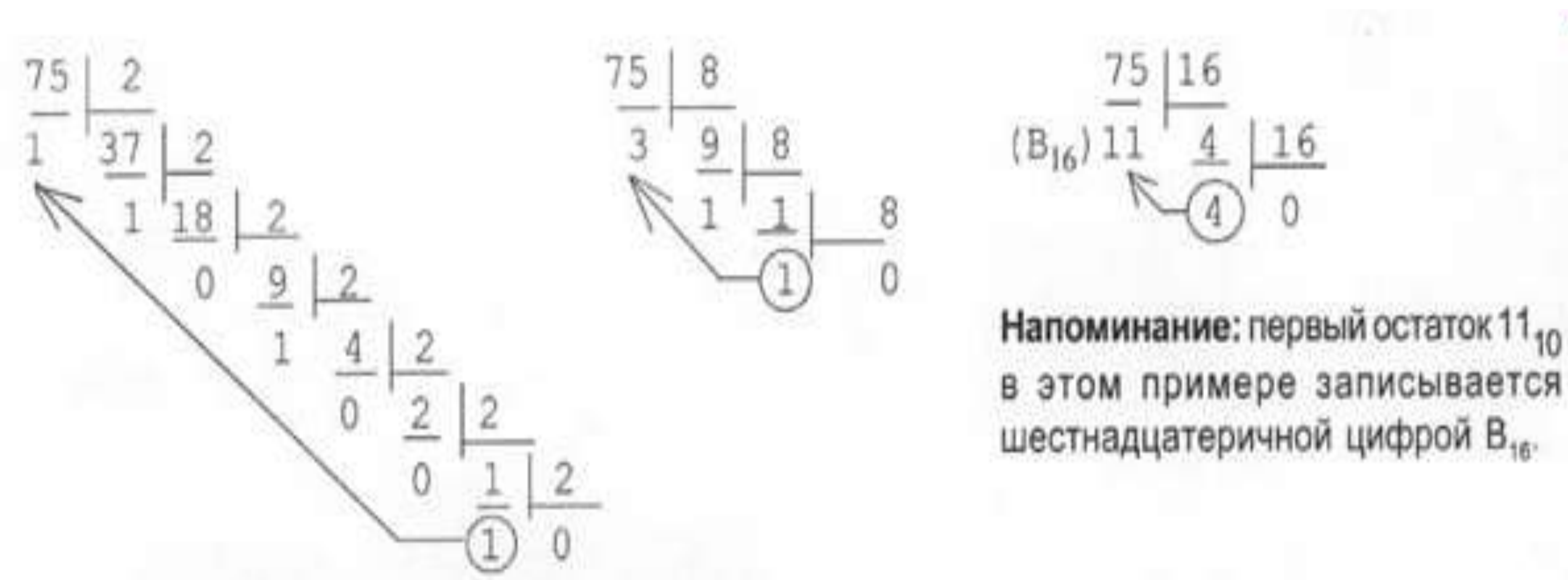
Ответ: $25_{10} = 11001_2$

Перевод десятичного числа в двоичную систему счисления



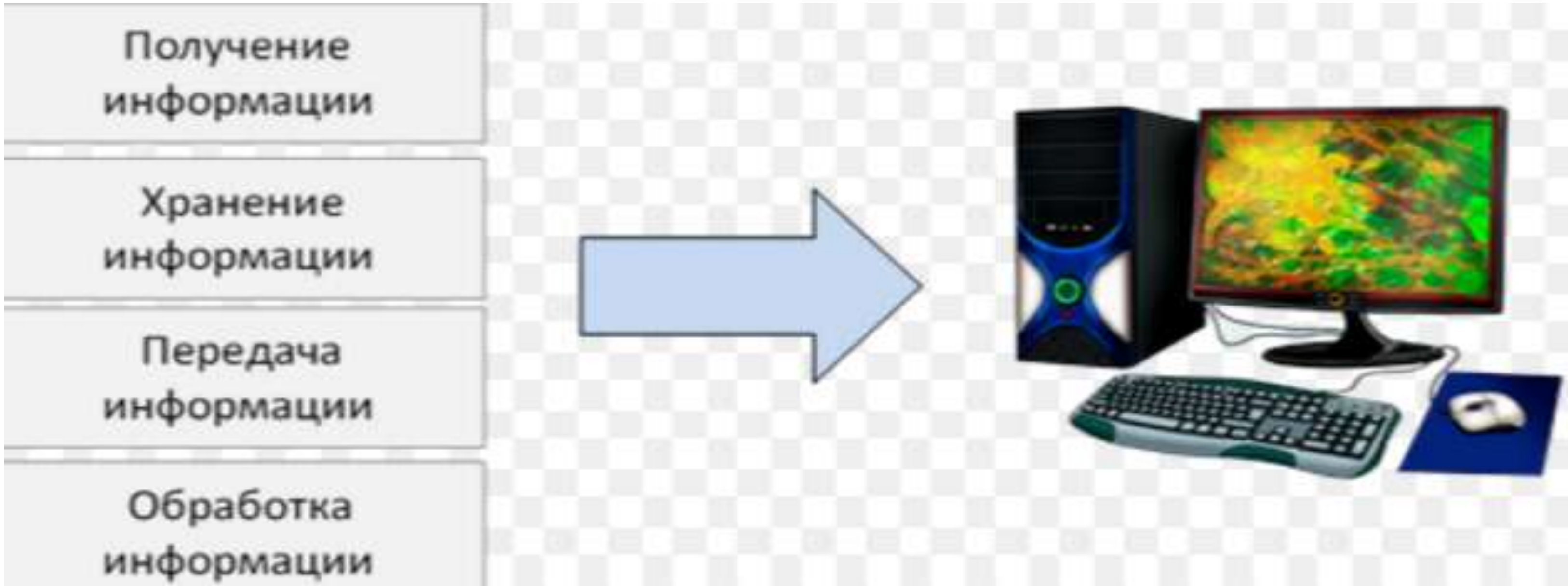
Ответ: 567_{10} = 1000110111_2

Пример перевода чисел из 10-ой системы в 2-ую, 8-ую, 16-ую системы счисления



Ответ: $75_{10} = 1\ 001\ 011_2 = 113_8 = 4B_{16}$.

3.Обработка информации на ПК



Компьютер –универсальное устройство обработки информации.

Под обработкой информации на компьютере понимают любые действия, которые преобразуют информацию из одного состояния в другое.

Компьютер имеет специальное устройство, называемое **процессором**, которое предназначено для быстрой обработки данных, со скоростями, доходящими до миллиардов операций в секунду. Данные для обработки процессор получает из **оперативной памяти** – от устройства, для временного хранения как входных, так и выходных данных. Процессор получает данные из оперативной памяти и записывает результаты в нее.

Обработка данных на компьютере

Тип информации	Человек	Компьютер
Числовая	1673	11010001001
Текстовая	М –русская буква	11101101
Графическая	●	00000000
Звуковая	Звук max громкости	11111111

11101001010010001010101010101010101

1000101010010100100010111101010101010101010001010101010101010110101010101

Обработка информации на ПК

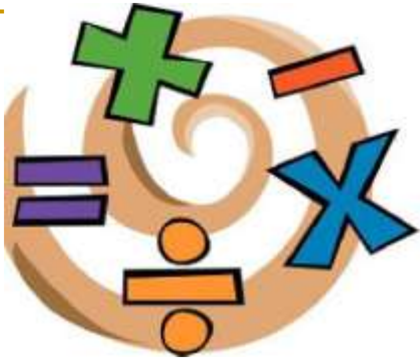
- Компьютер может производить вычисления, редактировать тексты, преобразовывать графические изображения и звуковые записи. Для того чтобы компьютер мог обрабатывать такие различные типы информации, она преобразуется в одинаковую цифровую форму – двоичный код.
- **Программы.** Для того чтобы компьютер «знал», что ему делать с данными, как их обрабатывать, он должен получить определенную команду (инструкцию). Например: «сложить два числа»; «заменить один символ в тексте на другой».
- Обычно решение задачи представляется в форме алгоритма, т. е. определенной последовательности команд. Такая последовательность команд (инструкций), записанная на «понятном» компьютеру языке, называется **программой.**
- **Программа** — это последовательность команд, которую выполняет компьютер в процессе обработки данных.

Основы математической обработки информации

Дисциплина, которая базируется на областях:

- Дискретная математика
- Линейная алгебра
- Логика
- Теория вероятностей
- Комбинаторика
- Статистика
- Информационные технологии

4.Математика



Математика – наука о количественных отношениях и пространственных формах действительного мира (Фридрих Энгельс).

Математика - наука о структурах, порядке и отношениях, сложившаяся на основе подсчёта, измерения и описания формы объектов.

Математические объекты создаются путём идеализации свойств реальных или других математических объектов и записи этих свойств на формальном языке.

Периоды развития математики

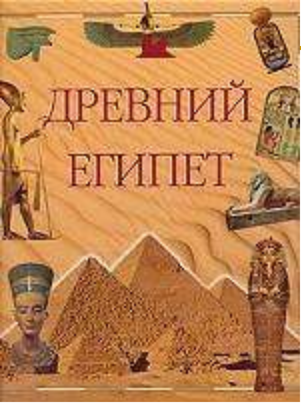
- **зарождение математики** – счет и мат. операции

  + - */ (VI-V вв. до н.э. Древний Египет, Вавилон);

- **элементарная математика** (5 в. до н.э. до 17 в. н.э.), Древняя Греция: Пифагор, Аристотель, Евклид, Архимед, Фалес, Демокрит, Птолемей и др., Индия, Китай, Древний Восток;

- **математика переменных величин** (XVII – начало XIX в.), Р.Декарт, И.Ньютон, Г.Лейбниц: функции, производные, дифференцирование, интегрирование, уравнения с неизвестными;

- **современная математика** (вторая половина XIX в. По наше время) – мат. логика, теория множеств, геометрия пространств.



I период

II период

Пифагор Архимед Евклид Фалес Эратосфен

III период

Лейбниц

Бернулли Ньютон Декарт

5.Аксиоматический метод в математике

В основе построения математической теории лежит аксиоматический метод. Базой научной теории являются некоторые исходные положения, называемые аксиомами, а все остальные положения теории получаются, как логические следствия аксиом.

- Основные понятия
- Определения
- Аксиомы
- Теоремы и их доказательство.

Понятие

Понятие – это отражение в сознании человека отличительных особенностей предметов и явлений, их общих и специфических признаков, выраженное словом или группой слов.

Понятие представляет собой высший уровень обобщения, присущий только словесно-логическому виду мышления. Понятия бывают конкретные и абстрактные.

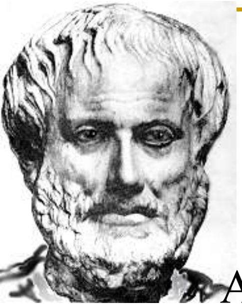
Конкретные понятия отражают предметы, явления, события окружающего мира, абстрактные отражают отвлеченные идеи. Например, «человек», «осень», «праздник» – **конкретные понятия;** «истина», «красота», «добро» – **понятия абстрактные.**

Содержание понятий раскрывается в **суждениях,** которые также всегда имеют словесную форму.

Аксиоматический метод в математике

Аксиома – утверждение, принимаемое без доказательств.

- ✓ Какова бы не была прямая, существуют точки, принадлежащие этой прямой, и точки, не принадлежащие ей.
- ✓ Через любые две точки можно провести прямую, и только одну.
- ✓ Из трех точек на прямой одна и только одна лежит между двумя другими.
- ✓ Каждый отрезок имеет определенную длину, б**о**льшую нуля.



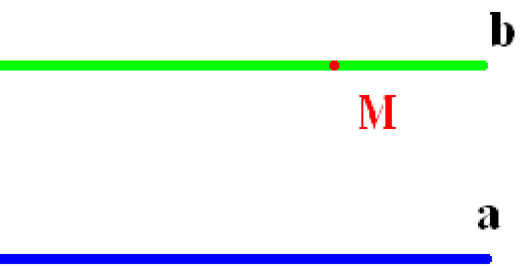
Аксиомы

Аксиома – утверждение, принимаемое без доказательств.

Аксиоматический метод в математике берет начало по меньшей мере от **Евклида**, хотя термин «аксиома» часто встречается и у **Аристотеля**: «… Ибо невозможны доказательства для всего: ведь доказательство должно даваться исходя из чего-то относительно чего-то и для обоснования чего-то. Таким образом, выходит, что все, что доказывается, должно принадлежать к одному роду, ибо все доказывающие науки одинаково пользуются аксиомами.

«… Аксиома обладает наивысшей степенью общности и суть начала всего. «Началами доказательства я называю общепринятые положения, на основании которых все строят свои доказательства.

«…О началах знания не нужно спрашивать «почему», а каждое из этих начал само по себе должно быть достоверным. Правдоподобно то, что кажется правильным всем или большинству людей или мудрым – всем или большинству из них или самым известным и славным». (Аристотель).



Аксиома Евклида: В плоскости через точку, не лежащую на данной прямой, можно провести одну и только одну прямую, параллельную данной.

Суть аксиоматического метода

Таким образом, аксиоматический метод сводится к следующему:

1)**выбирается множество принимаемых без доказательств аксиом**, причем входящие в аксиомы понятия явно не определяются в теории;

2) **фиксируются правила определения** и **правила вывода данной теории**, позволяющие логически выводить одни предположения из других;

3) **все теоремы выводятся из аксиом** на основе этих **правил**.

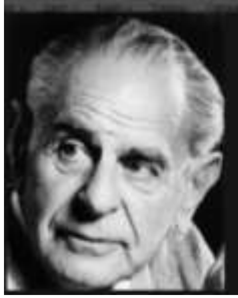
Таким методом в настоящее время построены различные разделы

математики (геометрия, теория вероятностей, алгебра и др.),

физики (механика, термодинамика);

делаются попытки аксиоматизации ***химии*** и ***биологии***.

Требования к аксиомам



Карл Раймунд Поппер – философ и социолог 1902 -1994**:**

«**Теоретическую систему можно назвать аксиоматизированной**, если множество **аксиом**, удовлетворяет четырем фундаментальным требованиям:

1) система аксиом должна быть **непротиворечивой** (то есть в ней не должно быть ни самопротиворечивых аксиом, ни противоречий между аксиомами). Набор аксиом называется непротиворечивым, если нельзя доказать одновременно и некое утверждение, и его отрицание.

2) аксиомы данной системы должны быть **независимыми**, то есть система не должна содержать аксиом, выводимых из остальных аксиом.

Эти два условия относятся к самой системе аксиом.

Что же касается отношения системы аксиом к основной части теории, то аксиомы должны быть:

3) **достаточными для дедукции всех высказываний**, принадлежащих к аксиоматизируемой теории, и

4) **необходимыми** в том смысле, что система не должна содержать излишних предположений. <…> Я считаю допустимыми две различные интепретации любой системы аксиом. Аксиомы можно рассматривать либо (1) как *конвенции*, либо (2) как эмпирические, или научные *гипотезы*»

5.Аксиоматический метод в математике

Куртом Гёделем доказана **невозможность полной аксиоматизации достаточно развитых научных теорий** (например, арифметики натуральных чисел), откуда следует невозможность полной формализации научного знания. При исследовании естественнонаучного знания аксиоматический метод выступает в форме *гипотетико-дедуктивного метода*.

Гедель К.Ф. - австрийский логик, математик и философ математики, наиболее известный сформулированной и доказанной им теоремой о неполноте.

Неполнота означает наличие высказываний, которые нельзя ни доказать, ни опровергнуть, исходя из аксиом этой теории.

Противоречивость — возможность доказать любое высказывание: как истинное, так и ложное.

Теорема

Теорема – утверждение, нуждающееся в доказательстве.

- *О параллельных прямых*: если накрест лежащие углы равны, то прямые параллельны. И наоборот, если прямые параллельны, то накрест лежащие углы равны.

- *Теорема Пифагора*: В прямоугольном треугольнике квадрат длины гипотенузы равен сумме квадратов длин катетов $\mathbf{c^2=a^2+b^2}$

- *Теорема Виета:* Для приведенного квадратного уравнения: $\mathbf{x^2+px+q=0}$ сумма корней равна коэффициенту **p**, взятому с обратным знаком, а произведение корней равно свободному члену **q**: $\mathbf{x1 + x2 = -p;\quad x1 \cdot x2 = q}$

Доказательство – это рассуждение, которое показывает, что истинность утверждения логически вытекает из истинности предыдущих теорем или аксиом.

Математические суждения

Каждая наука по существу представляет собой определенную систему **суждений об объектах, являющихся предметом ее изучения.**

Каждое из суждений оформляется в виде некоторого предложения, выраженного в терминах и символах, присущих этой науке.

Математика также представляет собой определенную систему суждений, выраженных в математических предложениях посредством математических или логических терминов или соответствующих им символов.

Математические термины обозначают те понятия, которые составляют содержание математической теории, логические термины обозначают логические операции, с помощью которых из одних математических предложений строятся другие математические предложения, из одних суждений образуются другие суждения, вся совокупность которых и составляет математику как науку.

Истинность суждений

В мышлении понятия не выступают разрозненно, они определенным способом связываются между собой. Формой связи понятий друг с другом является суждение.

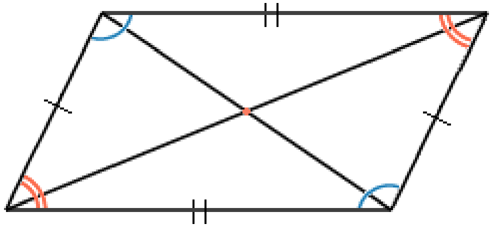
Суждение — форма мышления, в которой что-либо утверждается или отрицается о предмете, его свойствах или отношениях между предметами.

Если суждения правильно отображают эти объективно существующие зависимости между предметами, то мы такие суждения считаются истинными, в противном случае суждения являются ложными.

Так, например,

- **суждение "всякий ромб является параллелограммом" - истинное суждение;**
- **суждение "всякий параллелограмм является ромбом" - ложное суждение.**

Суждение имеет свою языковую оболочку - предложение, однако не всякое предложение является суждением.



Умозаключения

Умозаключением называется процесс получения нового суждения вывода из одного или нескольких данных суждений.

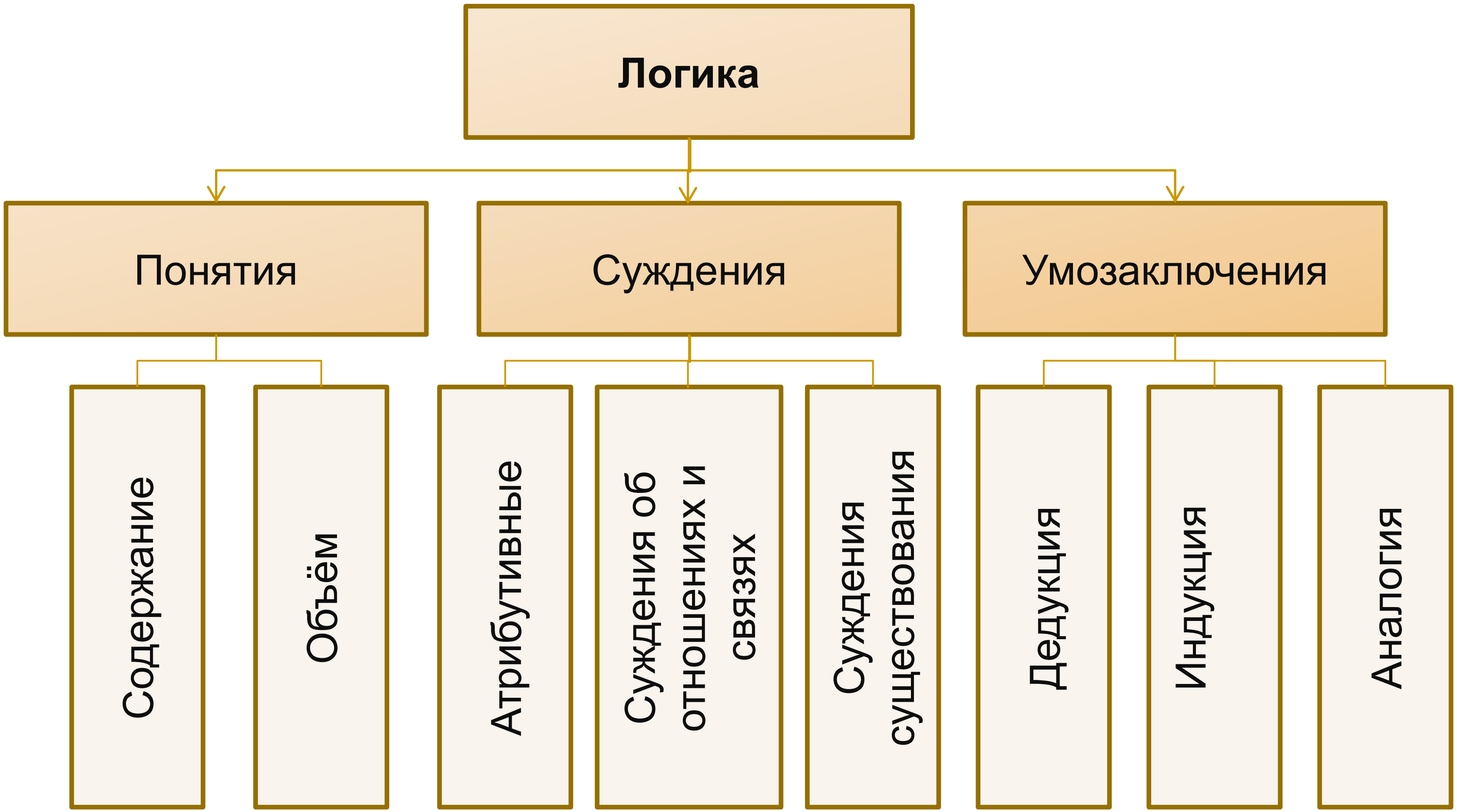
Первое суждение: Диагональ параллелограмма делит его на два равных треугольника .

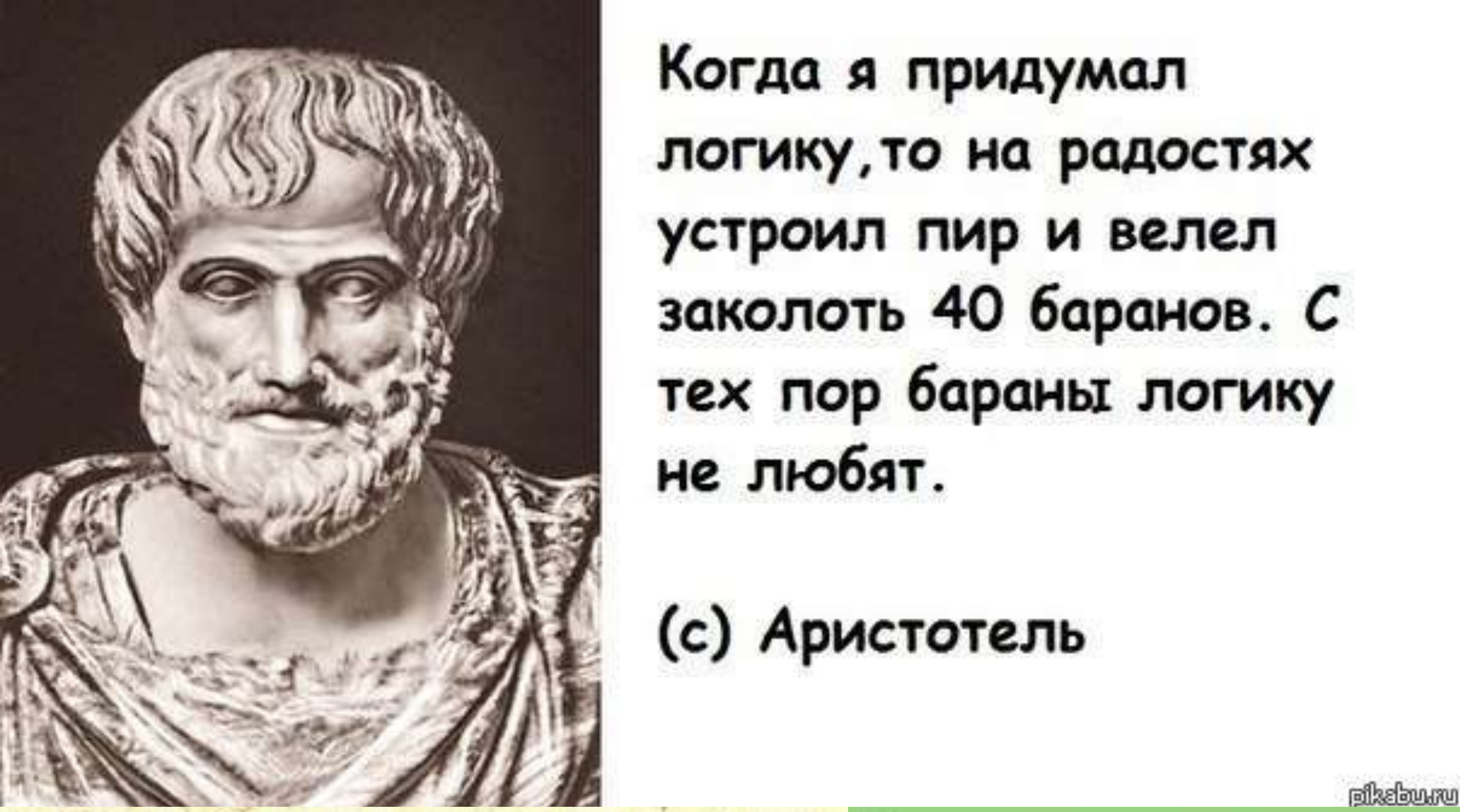
Второе суждение: Сумма внутренних углов треугольника равна 180°

Новое суждение-вывод: Сумма внутренних углов параллелограмма 360°

Познавательное значение математических умозаключений чрезвычайно велико. Они «расширяют» границы наших знаний об объектах и явлениях реального мира в силу того, **что большая часть математических предложений является выводом из сравнительно небольшого числа основных суждений**, которые получены, как правило, путем непосредственного опыта и в которых отражены наши наиболее простые и общие знания об его объектах.

Любая научная область основана на логике




Когда я придумал логику, то на радостях устроил пир и велел заколоть 40 баранов. С тех пор бараны логику не любят.
(с) Аристотель
pikabu.ru


Мужская логика - правильная.
Но женская интереснее!


Законы логики: да или нет.
Законы женской логики:
да, нет, да нет наверное,
да нет, я не уверена,
ой, всё, бесишь.

6. Математическая обработка данных и моделирование

Применение мат. обработки данных в научном познании:

1) Теоретическое направление
2) Экспериментальное направление
3) Вычислительное направление

6.1.Теоретическое направление

- Выдвижение гипотезы и построение математической модели (в виде уравнений или неравенств)
- Исследование математической модели (решение математической задачи)
- Модификация модели

Основа теоретического подхода –

математическое моделирование

6.2.Экспериментальное направление

Наблюдение → Эксперимент → Экспериментальные данные

Математическая обработка результатов эксперимента (экспериментальных данных):

- определение истинных значений измеряемых величин
- определение вида функциональной зависимости исследуемых величин – построение эмпирических зависимостей
- определение количественных характеристик (параметров) функциональных зависимостей

6.3.Вычислительное направление

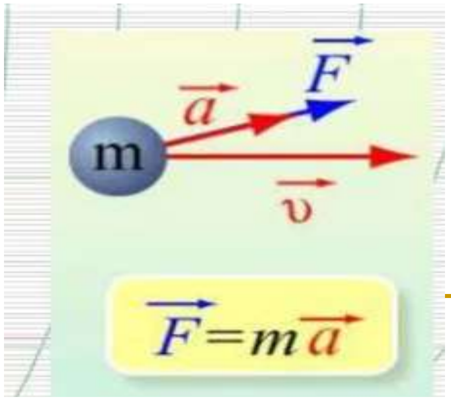
- Выбор или создание математической модели
- Разработка алгоритма решения задачи
- Составление компьютерной программы
- Проведение вычислений с помощью компьютера
- Анализ результатов и их проверка

Модель + Алгоритм + Программа – основа вычислительного эксперимента

6.4.Математическое моделирование

- **Модель** – формализованное представление объекта исследования, отражающее наиболее существенные свойства этого объекта. Может быть идеальной или физической.
- **Математическая модель** — это представление изучаемого объекта его свойств, структуры, связей с помощью математического языка.

Второй закон Ньютона



Структура математических моделей

Примеры математических моделей

Примеры математических моделей:

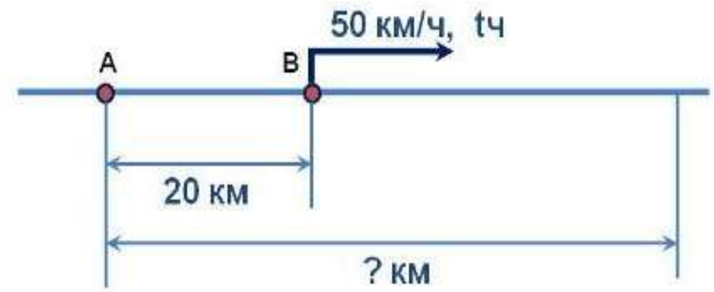
- модель прямолинейного перемещения тела

$$x = x_0 + v_x t + \frac{a_x t^2}{2}$$

- математическая модель периода колебаний пружинного маятника

$$T = 2\pi\sqrt{\frac{m}{k}}$$

Математическая модель



$$S = 50t + 20$$

$$F = 2x_1 - 3x_2 - 2x_3 + x_4 \to \min,$$

$$\begin{cases} -2x_1 + 4x_2 + x_3 - x_4 \le 2 \\ x_1 - 2x_2 + 3x_3 - 4x_4 = 5 \\ 3x_1 - 5x_2 - x_3 \ge 1 \\ x_2 \ge 0; x_4 \ge 0 \end{cases}$$

Математическая модель движения снаряда

Снаряд пущен с земли с начальной скоростью v_0 под углом $\alpha°$ к ее поверхности. Требуется найти траекторию его движения и дальность полета.

Математическая модель:

Дано:

v_0 =70скорость Высота подъема снаряда: $H=V_0*\sin\alpha - g*t^2/2$

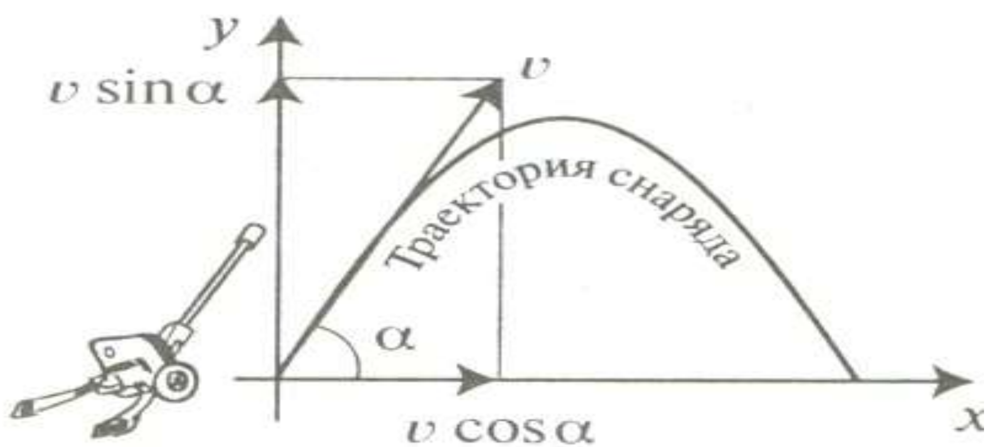
α =45°угол Дальность полета снаряда: $L=V_0*\cos\alpha*t$

$g=9.8$м/с^2

Найти:

траекторию полета

H и L -?



Модели солнечной системы

Развитие астрономии и физики давало рождение новым моделям. Сначала приблизительным и в чем-то ошибочным, затем все более точным..

- Геоцентрическая модель (Клавдий Птолемей –Земля в центре солнечной системы и неподвижна)
- Гелиоцентрическая модель (Николай Коперник)
- Модель Кеплера (законы Кеплера).
- Динамическая модель (Исаак Ньютон, закон всемирного тяготения)
- Теоретическое предсказание местоположения планет Нептун и Плутон

Геоцентрическая модель солнечной системы Клавдий Птолемей, II век, до н.э.

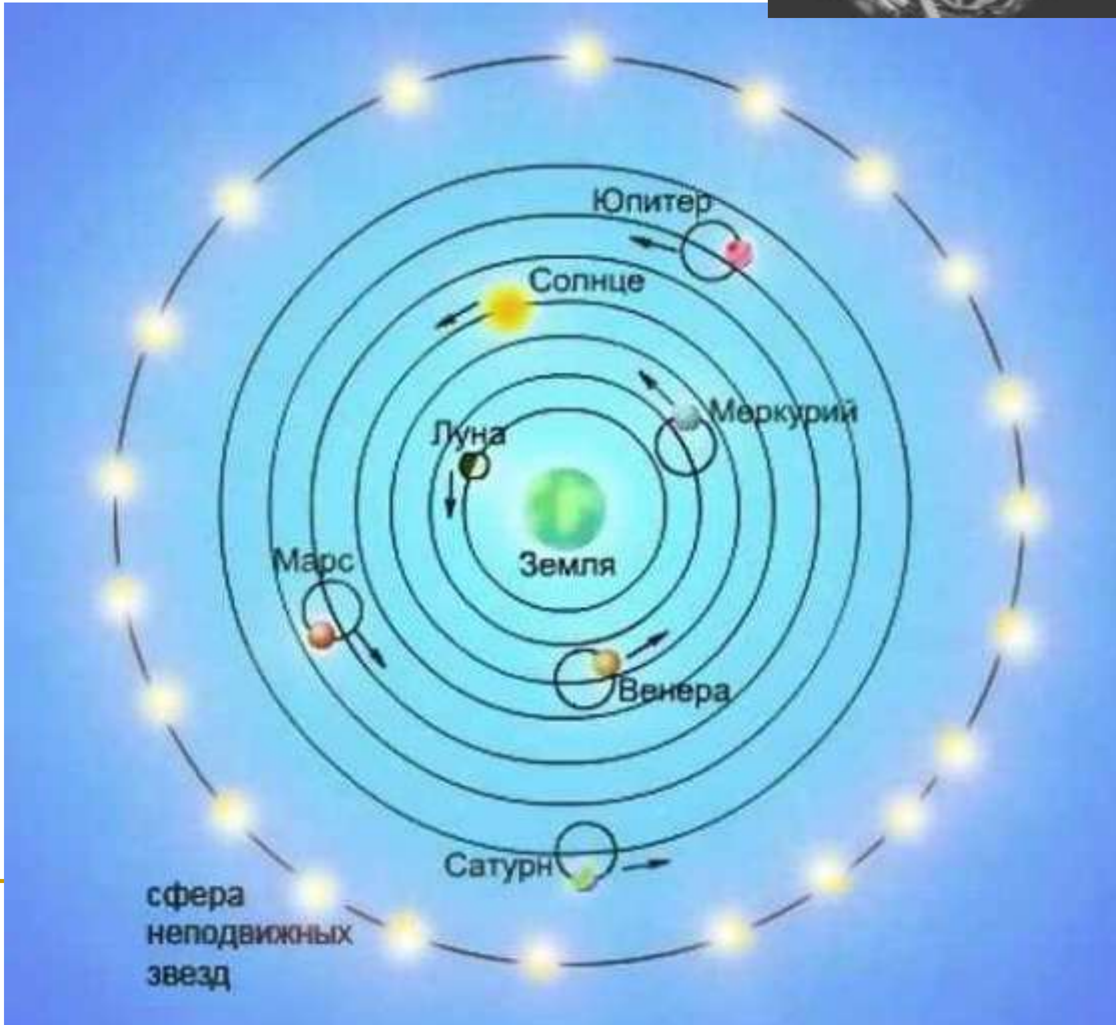
В основе системы мира Птолемея лежат четыре постулата:

I. Земля находится в центре Вселенной.

II. Земля неподвижна.

III. Все небесные тела движутся вокруг Земли.

IV. Движение небесных тел происходит по окружностям с постоянной скоростью



Модель солнечной системы

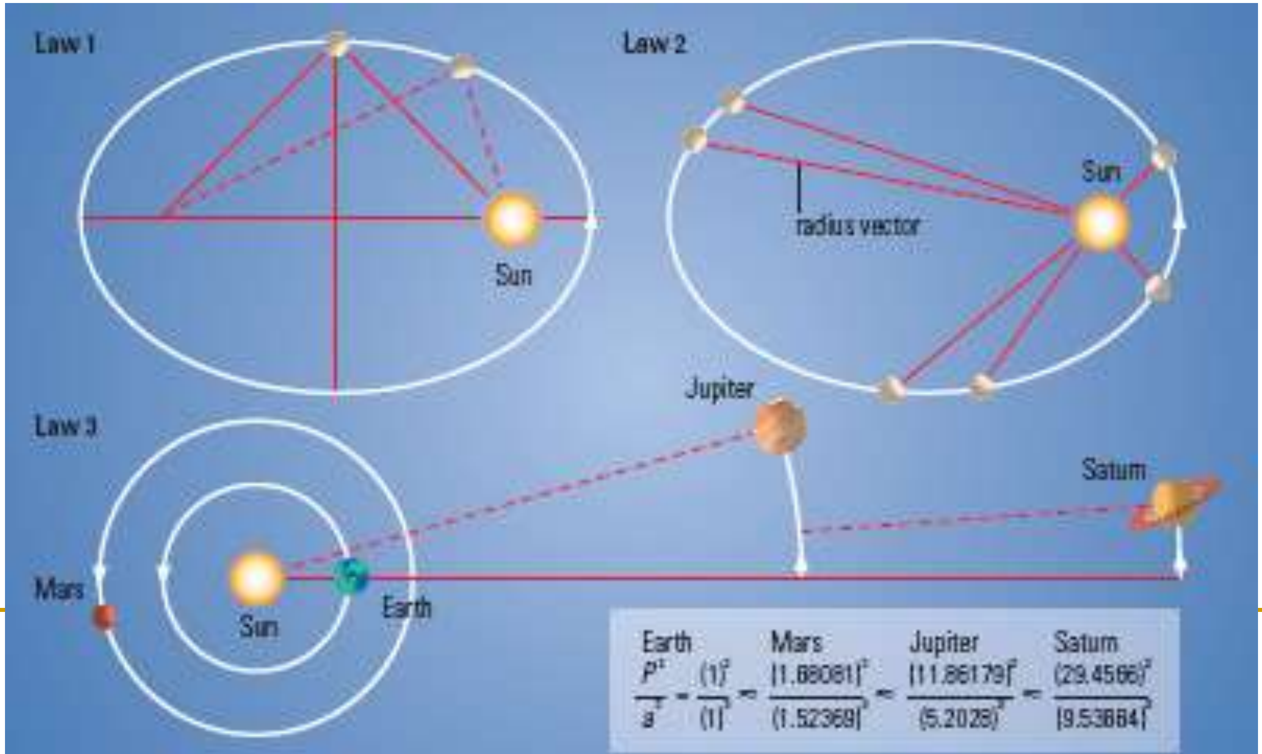
Гелиоцентрическая модель Н. Коперник, 1514 г.

Модель солнечной системы И.Кеплера (1609 г.)

1. Все планеты движутся по эллипсам, в одном из фокусов которых Солнце.

2. Радиус-вектор планеты в равные промежутки времени описывает равновеликие площади.

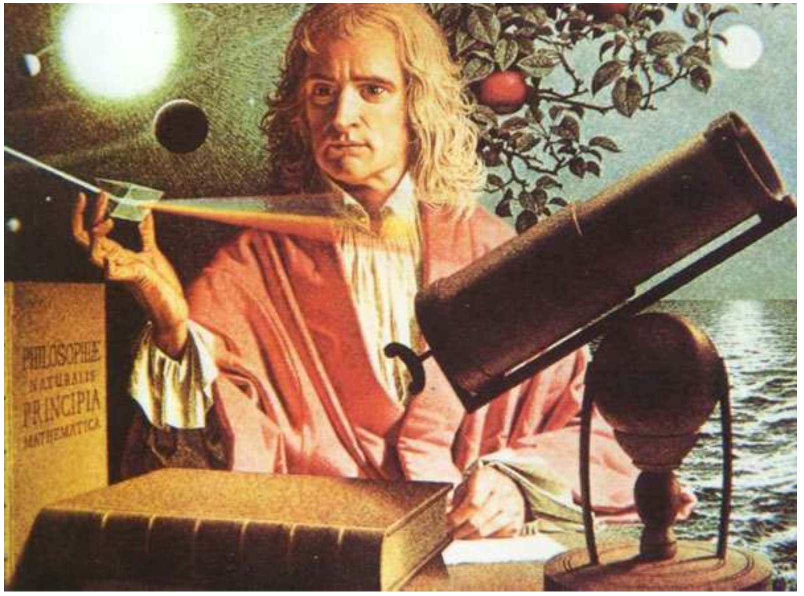
3. Квадраты сидерических периодов обращений планет вокруг Солнца пропорциональны кубам больших полуосей их эллиптических орбит.



Закон всемирного тяготения

Динамическая модель на основе закона всемирного тяготения Исаака Ньютона 1666 г

$$F = G\frac{Mm}{r^2}$$

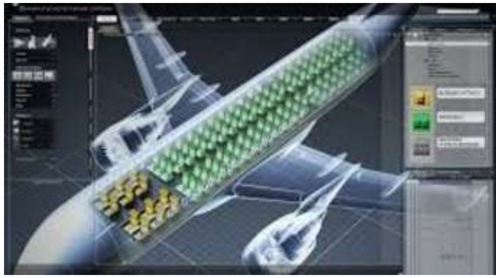


- Теоретическое предсказание положения планеты Нептун по возмущениям орбиты планеты Уран (Урбен Леверье и Джон Адамс)
- Теоретическое предсказание местоположения Плутона

Компьютерное моделирование

Компьютерная модель — это численная модель, представляющая объект в форме алгоритма, включающего данные о свойствах объекта и динамику их изменения со временем.

Компьютерные модели бывают



1) структурно-функциональные, которые представляют собой **условный образ объекта**, описанный с помощью компьютерных технологий;

2) **программа**, описывающая основные свойства и закономерности объекта моделирования - математическую модель, по которой можно получить определенные результаты.

7.МАТЕМАТИКА В ФИЛОЛОГИИ И ЛИНГВИСТИКЕ

ФИЛОЛОГИЯ - *совокупность наук, изучающих культуру народа, выраженную в языке и литературном творчестве*

«Бывают среднестатистические тексты, а бывают такие, в которых чего-то неожиданно много или чего-то неожиданно мало. Что это за неожиданности, которые встречаются в текстах, и как из измерить?

Математическая лингвистика

- Исследовать литературу цифровыми методами, для извлечения из большого количества текстов каких-либо общих тенденций и закономерностей можно с помощью Национального корпуса русского языка (НКРЯ) – платформы, которая является самым главным инструментом современных лингвистов.
- **Корпус русского языка** — это информационно-справочная система, основанная на собрании русских текстов в электронной форме.
- «Точки соприкосновения лингвистики и математики – самые неожиданные. И тем более я не думал раньше, что литературу можно выразить математически. Это интересный эксперимент: когда буквы уже не буквы, а рандомные номерные знаки, встречающиеся в тексте с определенной частотой»
- *Проникновение в лингвистику математических методов и «математического духа» способствовало развитию лингвистики в сторону точности и объективности. Однако на пути ее дальнейшего развития в этом направлении стоят серьезные препятствия*

Национальный корпус русского ×

Не защищено | www.ruscorpora.ru/index.html

НАЦИОНАЛЬНЫЙ КОРПУС
РУССКОГО
ЯЗЫКА

главная
архив новостей

поиск в корпусе

что такое корпус?
состав и структура
статистика
графики
частоты
морфология
обороты
синтаксис
семантика
параметры текстов

studiorum
форум

Национальный корпус русского языка

На этом сайте помещен корпус современного русского языка общим
система, основанная на собрании русских текстов в электронной фо

Корпус предназначен для всех, кто интересуется самыми разными в
языка, школьников и студентов, иностранцев, изучающих русский яз

Развитие подкорпусов НКРЯ (основного, поэтического, параллельно
проекты № 15-04-12018 «Развитие специализированных модулей НК
Пополнение и разметка».

Как пользоваться Корпусом (инструкция в формате PDF)

Подробнее о корпусе

Новости проекта

28 ноября 2018 года
28 ноября с 19 до 20 часов по техническим причинам поиск будет не

3 апреля 2018 года
Объём латышско-русского и русско-латышского параллельного корп
русско-бурятского параллельного корпуса вырос более чем вдвое и

Национальный корпус русско ×

Не защищено | search.ruscorpora.ru/search.

НАЦИОНАЛЬНЫЙ КОРПУС
РУССКОГО
ЯЗЫКА

Результаты поиска

красная

№	Вхождения	Документы	Фрагмент
1	786	320	красная армия
2	323	74	красная шапочка
3	277	246	и красная
4	205	128	красная звезда
5	146	136	красная и
6	129	110	красная площадь
7	112	107	как красная
8	107	87	что красная
9	103	63	красная девица
10	102	92	вся красная
11	86	75	красная как
12	84	80	была красная

Математическая лингвистика

Л. С. Выготский писал в 1934 году в книге «Мышление и речь»: «Первым, кто увидел в математике мышление, происходящее из языка, но преодолевающее его, был, по-видимому, Декарт» и продолжал: «Наш обычный разговорный язык из-за присущих ему колебаний и несоответствий грамматического и психологического находится в состоянии подвижного равновесия между **идеалами математической** и фантастической гармонии и в непрестанном движении, которое мы называем эволюцией».

В Древней Греции учение о грамматических категориях представляло собой описание важнейших аспектов строения языка с помощью абстрактных моделей, близких по стилю к тем моделям, которые были созданы древнегреческими математиками для описания пространственных форм; только привычность таких понятий, как падеж, род и т. п., ставших, как писал Х. Штейнталь, «нашей второй натурой», мешает нам понять, какого высокого уровня абстрактного мышления потребовало их создание. Первые попытки использовать для описания языкового «идеала математической гармонии» настоящие математические средства были предприняты лишь в середине XX столетия.

Математическая лингвистика

в XIX веке математики начали строить неколичественные абстрактные модели, отличавшиеся от античных более высоким уровнем абстракции, а также — что для нашей темы особенно важно — тем, что они могли использоваться для описания значительно более широкого круга явлений, чем пространственные формы; нередко такие модели оказывались удобным и даже необходимым средством для изучения явлений, о которых строившие их математики вовсе не думали и даже не знали об их существовании. Среди этих моделей были и те, которые впоследствии получили применение в лингвистике; особенно интенсивное развитие математических дисциплин, содержанием которых было их построение, пришлось на первую половину XX столетия.

Одним из результатов этой встречи было возникновение новой математической дисциплин — **математической лингвистики,** предметом которой является разработка математического аппарата для лингвистических исследований. Центральное место в математической лингвистике занимает теория формальных грамматик, по характеру используемого в ней аппарата родственная математической логике и в особенности теории алгоритмов. Она доставляет формальные методы описания правильных языковых единиц различных уровней, а также что

Математическая лингвистика

- В математической лингвистике разрабатываются также аналитические модели языка, в которых на основе тех или иных данных о «правильных текстах» производятся формальные построения, результатом которых является описание каких-то «составных частей» механизма языка. На этом пути можно получить формальное описание некоторых традиционных грамматических понятий. Сюда же следует отнести описание смысла предложения с помощью аппарата интенсиональной логики («семантику Монтегю»).
- С помощью математического аппарата можно описать только один из двух идеалов языка, о которых говорил Выготский; поэтому часто раздающиеся возражения против использования той или иной математической модели на том основании, что такие-то и такие-то частные случаи она не охватывает, не имеют смысла: для описания присущих языку «колебаний и несоответствий» нужны совсем другие, не математические средства, и как раз четкое описание «математического идеала» могло бы помочь их находить, поскольку оно позволило бы ясно отграничивать в языке «фантастическое» от «математического». Но это пока что дело будущего.

Математическая лингвистика

- Не меньшее, а может быть и большее значение, чем возникновение математической лингвистики, имело непосредственное проникновение в лингвистику фундаментальных математических идей и понятий — таких, как множество, функция, изоморфизм. В современной лингвистической семантике важную роль играют пришедшие из математической логики понятия предиката и квантора. (Первое из них возникло в логике еще тогда, когда она не отграничивалась от лингвистики, и теперь вернулось в лингвистику в обобщенном и математически обработанном виде.)
- Очень большое значение имеет уточнение языка лингвистических исследований, происходящее благодаря проникновению в лингвистику «математического духа» не только в тех ее областях, где возможно использование математических идей и методов. Все это можно коротко резюмировать так: ***лингвистика становится все более точной и более объективной наукой — не переставая, само собой, быть наукой гуманитарной.***

Математическая лингвистика

- Модель «Смысл - текст» — лингвистическая концепция, созданная

И. А. Мельчуком и представляющая **язык как многоуровневую модель преобразований смысла в текст и обратно** (модель «Смысл ⇔ Текст»).

Отличительной особенностью этой теории является также использование синтаксиса зависимостей.

Значительна роль, отводимая лексическому компоненту модели — Толково-комбинаторному словарю.

Задачи компьютерной лингвистики

В сфере обработки текстов на сегодняшний день сформировалось два подхода:

1) **на основе моделей языка и правил, составленных экспертами;**
2) **на базе машинного обучения.**

Первый позволяет достичь лучших результатов, однако составление моделей и правил настолько трудоемкий процесс, что уступающие по качеству методы машинного обучения практически его вытеснили.

Задачи компьютерной лингвистики

- **Анализ и градация мнений.** Соотнесение текста, написанного от первого лица, с дискретной шкалой оценок: плохо, хорошо, очень хорошо и т. д. Используется для анализа отзывов в интернет-магазинах и высказываний в социальных сетях.
- **Анализ тональности высказываний.** Выявление позитивного или негативного отношения к обсуждаемому предмету. Используется для анализа отзывов, генерации диалога и т. д.
- **Классификация текстов по темам.** Отнесение текста к той или иной тематике. Используется во многих приложениях для рубрикации текстов в онлайн-библиотеках и для организации новостных потоков.
- **Генерация речи.** Используется в робототехнике, смартфонах, навигаторах.
- **Ведение диалога.** Анализ реплик собеседника и формирование на их основе ответов. Используется в робототехнике, экспертных системах, Call-центрах. **Проверка правописания** в текстовых редакторах, поисковых системах.
- **Извлечение смысла из текста.** Выделение ключевых слов и словосочетаний, трендов, суммаризация. Применяется в новостных системах для агрегирования новостей, базах знаний для организации хранения знаний и вывода новых фактов.
- **Поиск ответов на вопросы.** Подборка по вопросу и, возможно, контексту наиболее релевантного ответа. Применяется в поисковых и экспертных системах.
- **Машинный перевод.**

Функции языка

- Язык - сложная знаковая система, естественно или искусственно созданная и соотносящая понятийное содержание и типовое звучание (написание).

Термин «язык», понимаемый в широком смысле, может применяться к произвольным знаковым системам.

- Язык в широком смысле – это словарь, грамматика, рассказы, повести, написанные на этом языке.

Функции языка:

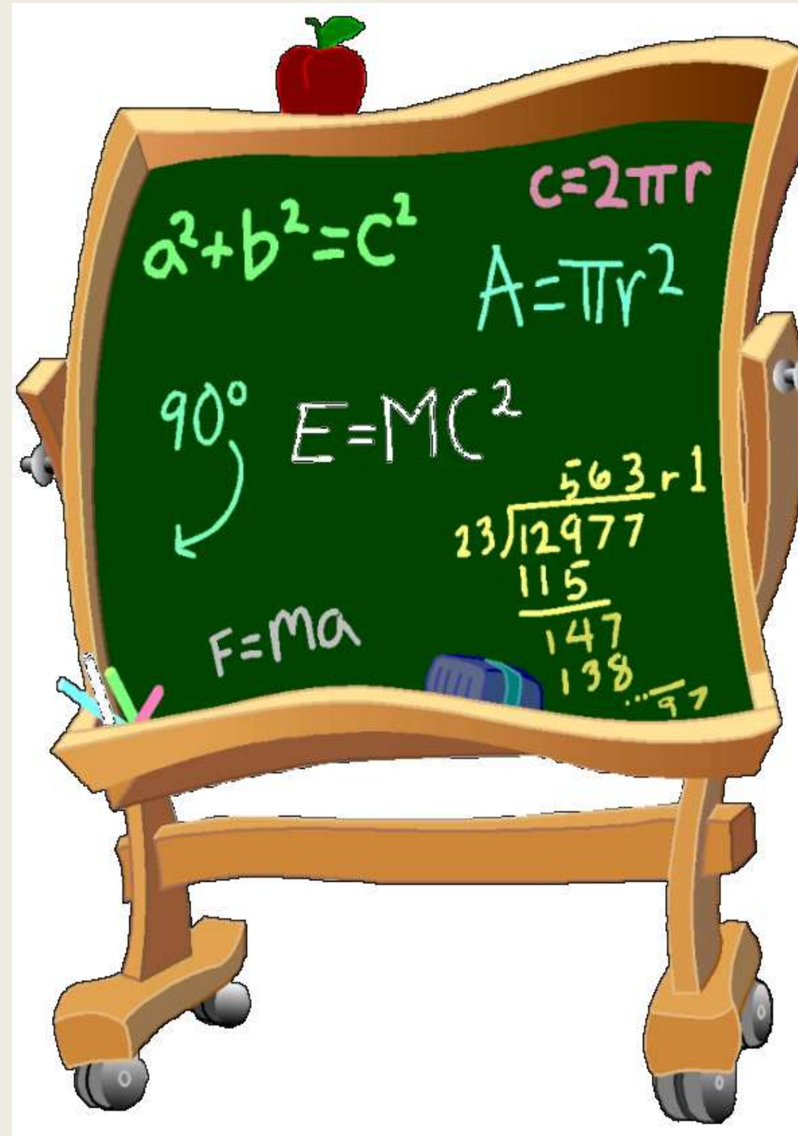
1) коммуникативная (функция общения) — использование языка для передачи информации;
2) когнитивная (гносеологическая) — накопление и сохранение информации, её передача;
3) аккумулятивная (накопительная функция) — накопление и сохранение знания

Математический язык

В математическом языке:

словарь и грамматика – *математическая операционная система*

рассказы, повести и прочее – *математические модели*



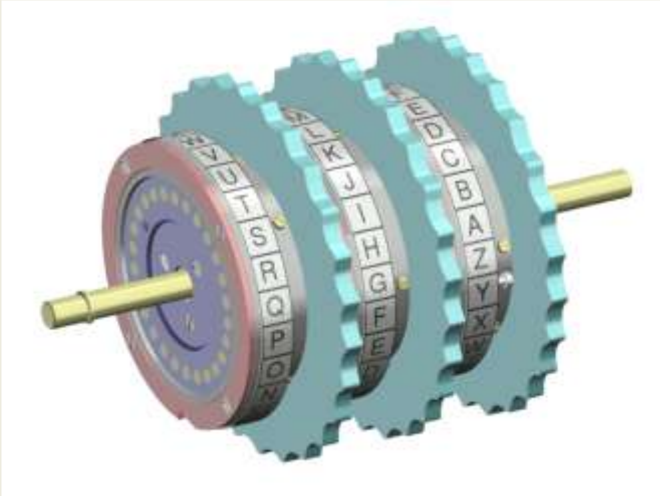
Математический язык

- Математика: мышление, чувствование и язык.
- Язык математики как и любой другой язык состоит из совокупности высказываний (предложений).
- Математические высказывания это математические символы, объединенные формулой.

Математика – язык символов и формул

Математика в филологии

- Криптография и расшифровка древних текстов
- Обоснование авторства текстов
- Перевод текстов с одного языка на другой
- Лексический анализ текстов
- Другие задачи…



Статистические методы анализа лексики

Лексика представляет собой статистически организованную структуру:

Вероятностные характеристики слова проявляются в неодинаковой частотности их в речи, в многообразных видах лексических связей

Установлено, например, что самые частотные слова в естественном языке, как правило, являются наиболее краткими, наиболее древними, наиболее простыми по морфологической структуре, наиболее многозначными

Статистические методы используются для изучения характера семантических связей между словами.

Так, например, установлено, что слова, часто встречающиеся вместе в определенном отрезке текста, теснее связаны между собой по смыслу, чем слова, реже появляющиеся рядом в этом же отрезке текста.

Математическая лингвистика

Лингвистика — наука, изучающая языки, наука о естественном человеческом языке вообще и обо всех языках мира как индивидуальных его представителях. В широком смысле слова, **лингвистика** подразделяется на научную и практическую.

Математическая лингвистика - математическая дисциплина, разрабатывающая формальный аппарат для описания строения естественных и некоторых искусственных языков. Возникла в 50-х годах 20 века и является направлением развития искусственного интеллекта. Базируется на методах алгебры, теории алгоритмов и теории автоматов.


Информатика
Теоретическая информатика
Математическая логика
Вычислительная математика
Теория информации
Системный анализ
Теория принятия решений
Искусственный интеллект
Психолингвистика
Когнитивная психология
Робототехника
Операционные системы
Экспертные системы
Программирование
Системное
Трансляторы
Прикладное
Языки программирования
Прикладная информатика
САПР
АСНИ
АСУ
АИС
АОС
Вычислительная техника
Кибернетика
Автоматическое управление
Бионика
Распознавание образов
Гомеостатика
Математическая лингвистика

Математическая лингвистика

Изучение способов математического описания текстов

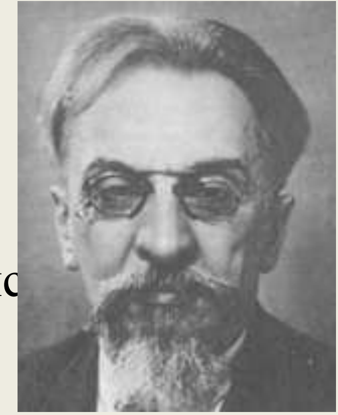
Для описания строения **предложения** можно либо выделить в нём "составляющие" — группы слов, функционирующие как цельные синтаксические единицы, либо указать для каждого слова те слова, которые от него непосредственно зависят. Математические объекты, возникающие при таком описании структуры предложения, называются деревом составляющих (1-й способ) и деревом синтаксического подчинения (2-й способ).

Теория формальных грамматик (Н.Хомский)

- Изучает способы описания закономерностей, которые характеризуют не отдельный текст, а всю совокупность текстов того или иного языка. Эти закономерности описываются путём построения "формальной грамматики" — абстрактного "механизма", позволяющего с помощью единообразной процедуры получать правильные тексты данного языка вместе с описаниями их структуры.
- Используется при разработке и описании искусственных языков (языков программирования)

Построение аналитических моделей языка, в которых на основе тех или иных данных о речи, считающихся известными (например, множества правильных предложений), производятся формальные построения, дающие некоторые сведения о структуре языка.

Лингвистика и алгебра



Пример был предложен академиком Л. В. Щербой в 1930-е годы и ис
на вводных лекциях к курсу «Основы языкознания».
Широкую известность эта фраза приобрела после публикации научно-популярной книги Льва Успенского «Слово о словах».

«Глокая куздра штеко будланула бокра и кудрячит бокрёнка»

$$y = x + a$$

Искусственная фраза на основе русского языка, в которой все корневые морфемы заменены на бессмысленные сочетания звуков.

Несмотря на это, общий смысл фразы понятен: некоторая сущность женского рода что-то сделала определённым образом с другим существом мужского пола, а затем начала (и продолжает до настоящего момента) делать что-то другое с его детёнышем (или более мелким представителем того же вида).

Фраза создана для иллюстрации того, что многие семантические признаки слова можно понять (не зная значений слов) из его морфологии.

Проблемы распознавания смысла