Лекции по информатике  
(оригиналы любезно предоставлены  
Романом Кузнецовым и Алисой Потолоковой,  
оцифровано и оформлено Фарманяном Арташесом)

# Понятие информации и природа её возникновения

**Информатика** – дисциплина, основанная на использовании компьютерной техники, изучающая *структуру* и *общие свойства* информации, а также *закономерности* и *методы её создания, хранения, поиска, преобразований и применений* в различных сферах человеческой деятельности.

В 1978 году международный конгресс официально закрепил за понятием информатики области, связанные с обслуживанием **С**истем **О**бработки **И**нформации.

## Средства преобразования

Технические средства **(Hardware)**Программное обеспечение **(Software)**Алгоритмы  
**(Brainware)**

## Информатика

**Отрасли производстваФундаментальнаяПрикладная**

***Данные*** *+ обработка =* ***информация******Информация*** *+ обработка =* ***знания***

**Информация** – *сведения* об окружающем мире, *фактах* человеческой деятельности, которая является объектом *передачи*, *накопления* и *обработки*.

## Информационные процессы

**Хранение**Хранилище данных

Внутреннее Внешнее

**Передача**Источник – Канал –Получатель  
  
Тех. Средства Органы Чувств

**Обработка**

* С Тех. Сред.
* Без Тех. Сред.

# Информационные технологии

**Информационные технологии –** совокупность методов и способов получения, обработки и представления информации, направленных на изменение её состояния, свойств, формы, содержания и осуществляемых в интересах пользователя.

## Уровни информационных технологий

ТеоретическийИсследовательскийПрикладной

**Информационный ресурс** – это *концентрация* имеющихся *фактов*, *документов* и *знаний*, *отражающих* реальное изменяющееся во времени состояние общества и *используемых* в научных исследованиях и материальном производстве.

**Информационная система** – упорядоченная совокупность документированной информации и информационных технологий.

## Виды информационных систем

Ручные  
(поддержка целиком осуществляется человеком)

Автоматические  
(обработка происходит автоматически, человеческое участие всё ещё необходимо)Автоматизированные  
(обработка происходит с минимальным человеческим участием)

**Делимость** и **целостность** – свойства информационных систем.

**Информационная среда** – совокупность систематизированных и организованных специальным образом данных и знаний.

**Эффективность применения** и **качество функционирования** любых систем в значительной степени определяется **качеством** информации.

**Качество** информации – совокупность свойств информации, характеризующих степень её соответствия потребностям пользователя.

## Качества информации

Содержательность (внутреннее)

* Значимость
* Полнота
* Идентичность
* Конфиденциальность
* Доступность

Защищённость (внешнее)

* Кумулятивность
* Избирательность
* Гомоморфизм
* Сохранность
* Достоверность
* Скрытность
* Имитостойкость

## Внутренние свойства

**Содержательность** – совокупность сведений о конкретном объекте или процессе, содержащаяся в сообщении и воспринимаемая получателем.

**Значимость** – свойство информации сохранять ценность для потребителя с течением времени.

**Полнота** – свойство содержательной информации, характеризумое мерой её достаточности для решения определённых задач

**Идентичность** – свойство, заключающееся в соответствии содержательной информации состоянию объекта.

**Кумулятивность** – свойство содержательной информации, заключенной в массиве небольшого объёма, достаточно полно отображать действительность.

**Гомоморфизм** – свойство содержательной информации, связанное с достаточно полным отображением действительности, представленной информационным массивом большого объёма, с помощью малого числа информационных единиц на основе соответствующих моделей агрегирования.

**Избирательность -**  свойство содержательной информации, заключающееся в достаточно полном отображении действительности, представленной информационными массивами большого объема, с помощью малого числа информационных единиц на основе учёта конкретного потребителя.

## Внешние свойства

**Защищенность** отражает внешнее качество информации, определяемое совокупностью свойств информации, обеспечиваемых системой контроля и защиты информации (КЗИ) в конкретной информационной системе.

**Имитостойкость** определятся степенью её защиты от внедрения информационных массивов, имитирующих зарегистрированные массивы и заключается в способности недопустить навязывание дезинформации и нарушение нормального функционирования информационной системы.

**Сохранность** информации – свойство информации, характеризующееся степенью готовности определённых информационных массивов к целеовму приведению и определяемая способностью контроля защиты информции обеспечить постоянное наличие и своевременное представление информационных массивов. Основными показателями сохранности являются целостность и готовность. (???)

## Потребительские показатели качества информации

1. **Репрезентативность** связана с адекватным отображаением свойств объекта
2. **Содержательность**
3. **Достаточность**
4. **Доступность**
5. **Актуальность**
6. **Своевременность**
7. **Точность**
8. **Достоверность**
9. **Устойчивость**

## Уровни проблем передачи информации

1. Синтаксический (порядок символов)
2. Семантический (смысл)
3. Прагматический (восприятие получателем)

Проблемы **синтаксического** уровня:

Создание теоретических основ, построение связей, основные показатели функционирования которых были бы близки к предельно возможным, а также совершенствование существующих систем с целью повышения эффективности их использования.

Проблемы **семантического** уровня:

Связаны с формализацией смысла передаваемой информации, то есть оценкой её качества.

Проблемы **прагматического уровня**:

Последствия от получения и использования информации.

## Формы представления информации, её виды и свойства

**ДискретныеАналоговые**

Виды информации по представлению: **графическая, звуковая, текстовая, символьная, экономическая.**

Виды информации по восприятию: **аудиальная, визуальная, органолептическая.**

Свойства информации: **достоверность, полнота, точность, ценность, понятность, актуальность, своевременность.**

# Формы представления и преобразования информации

Отображение внешней информации во внутренней называется **кодированием**.

**Код –** способ отображения и множество понятий.

В ЭВМ используются 2 формы представления двоичных чисел: естественная (фиксированная точка, обычно целые числа) и нормальная (плавающая точка).

Общая форма записи естественных чисел: 0|000 0000, где самый левый бит – знак, а остальные – модуль.

**Переполнение разрядной сетки**  **–** явление, при котором из-за слишком большого модуля числа теряются его старшие биты. Переполнение происходит, когда количество значимых разрядов модуля превышает n-1, где n – число бит, выделенных на хранение числа.

Диапазон модулей чисел, которые могут быть представлены в n-разрядной сетке – от нуля до 2n-1 – 1.

Представление чисел с плавающей точкой состоит из порядка и мантиссы. Порядок указывает действительное положение запятой в числе. Точность представленного значения зависит от количества значимых битов в мантиссе. Признаком нормализованного числа служит наличие единицы в старшем разряде модула мантиссы.

## Машинные коды двоичного числа

ЭВМ в целях упрощения выполнения арифметических операций применяет специальные коды для преставления чисел. При помощи этих кодов упрощается определение знака результата операции. Операция вычитания сводится к арифметическому сложению, облегчается выработка признаков переполнения разрядной сетки. Для представления чисел со знаком, в ЭВМ используют прямой, обратный и дополнительный код. Прямой код – это обычный код числа, обратный код – побитово инвертированный прямой код, дополнительный код – обратный код + 1.

# Представление символьной информации

**Таблица кодировки** – это стандарт, который ставит в соответствие каждому символу свой порядковый номер. Таблица кодировки устанавливает связь между внешним символьным алфавитом компьютера и внутренним двоичным представлением символов.

В кодировке ASCII один символ занимает 1 байт, в Unicode – минимум 2 байта. В настоящее время самой актуальной является таблица Unicode, самой распространённой кодировкой является UTF-8 (Unicode Transformation Format). Согласно Unicode, в первом байте кодируется символ, а во втором – его признак. Для русских букв существует 5 стандартов.

# ИЗМЕРЕНИЕ ИНФОРМАЦИИ

Существует два подхода к измерению информации. На синтаксическом уровне для оценки количества информации используют **вероятностные** методы, которые рассматривают количество информации исключительно как меру уменьшения неопределённости. Общий принцип вероятностного подхода: «Если сообщение приводит к уменьшению неопределённости наших знаний, то можно утверждать, что такое сообщение содержит информацию».

Формула Хартли – H = log2N, где H – количество информации, N – вероятность событий

Формула Шеннона – H = , где p – вероятность событий

Количество информации – числовая характеристика сигнала, отражающая ту степень неопределённости, которая исчезает после получения сообщения в виде данного сигнала. Эту меру неопределённости в теории информации называют энтропией.

Не всякий текст, записанный двоичными символами, допускает измерение объёма информации в вероятностном смысле. Если некоторое сообщение допускает измеримость количества информации в обоих смыслах, то они необязательно совпадают, при этом вероятностное количество информации не может быть больше объёмного.

## Семантическая мера информации

Информация рассматривается по её содержанию, отражающему состояние отдельного объекта или системы в целом. При этом не учитывается её полезность для получателя информации. Данный подход предполагает, что для использования и понимания полученной информации получатель должен обладать тезаурусом.

Тезаурус – совокупность определений, которыми распологает пользователь или система.

В зависимости от соотношений между смысловым содержанием сообщения и тезаурусом пользователя изменяется количество информации. При этом характер такой зависимости не поддаётся строгому математическому описанию, а сводится к рассмотрению трёх основных условий при которых тезаурус пользователя:

1. Стремится к нулю.
2. Стремится к бесконечности
3. Согласован со смысловым содержанием – сообщение понятно и несёт новые сведения.

Относительной мерой семантической информации может служить коэффициент содержательности, который определяется как отношение количества семантической информации к её объёму.

Прагматическая мера информации – на данном уровне рассматривается информация с точки зрения её полезности для достижения потребителем поставленной практической цели. Данный подход при определении полезности информации основан на расчёте приращения вероятности достижения цели до и после получения информации. Определяющий ценность информации.

1. Количество информации равно нулю, когда P1 = P0
2. Количество информации положительно, когда P1 > P0 – полученная информация уменьшает исходную неопределённость и увеличивает вероятность достижения цели.
3. Сумма значения информации является отрицательной, когда P1 < P0 увеличивает исходную неопределённость и уменьшает вероятность достижения цели.

Дальнейшее развитие данного подхода базируется на статической теории информации и теории решений, при этом кроме вероятностных характеристик достижения цели после получения информации вводятся функции потерь и оценка полезности информации производится в результате минимизации функции потерь. Максимальной ценностью обладает то количество информации, которое уменьшает потери до 0 при достижении поставленных целей.

# Структуры данных

Представление данных определяется исходя из средств и возможностей, допускаемых компьютером и его ПО.

Важную роль играют свойства самих данных, операций, которые должны выполняться над ними.

Выделяют две группы данных:

1. Простые (неструктурированные) типы данных
   1. Целые числа
   2. Числа с плавающей точкой
   3. Символьный тип данных
   4. Логический тип (булевы переменные, true-false)
2. Структурированные типы данных
   1. Массив (регулярный тип) – однородный набор величин одного и того же типа, называемых компонентами массива, объединённых одним общим именем и адресуемых вычисляемым индексом. Компонентами массива могут быть не только простейшие данные, но и структурные. Массивы статичны – то есть имеют заранее определяемый размер, который в последствии не изменяется. Данные, хранящиеся в массиве, находятся в оперативной памяти – это ускоряет доступ в ходе решения задач, но налагает ограничения на объём возможной информации, организованной в виде массива.
   2. Запись – неоднородная упорядоченная статическая структура прямого доступа. Запись – набор именованных компонент, а именно – полей (чаще всего – разного типа), объединённых одним общим именем и адресуемых с помощью как *имени записи*, так и *имён полей.* Записи и массивы обладают одним общим свойством – произвольным доступом к компонентам.
   3. Множество – совокупность каких-либо объектов, являющихся его элементами. операции над множествами – пересечение, объединение, вычитание. проверка принадлежности. Различия между множеством и массивом - размер множества заранее не ограничен, не существует иного способа доступа к элементам множество кроме как проверка принадлежности.
   4. Очередь – упорядоченный набор следующих друг за другом компонент, доступ к которым происходит по следующим правилам:
      1. Новые компоненты могут добавляться лишь в хвост очереди
      2. Значение компонент могут читаться (извлекаться) лишь в порядке следования компонент – от головы к хвосту очереди.
      3. Размер очереди заранее не оговаривается, не ограничен, для запоминания компонент очереди испольуется внешнее запоминающее устройство большой ёмкости. Иногда используют другое название очереди – файл.
   5. Стек – структура данных, в которой элемент, который первым в неё помещался, выходит последним, и наоборот.
   6. Иерархическая организация данных.

Важным отличием компьютерных чисел от математических является ограниченность диапазона. Для каждой компьютерной системы свойство своё самое большое по модулю допустимое целое число. Прибавив единицу к самому большому допустимому положительному числу мы получим модуль самого малого допустимого отрицательного числа. Это связано с особенностями позиционирования в ячейках памяти.

Структурированные типы данных классифицируют по следующим признакам:

* Однородность/Неоднородность
* Упорядоченность/Неупорядоченность
* Прямой/Последовательный доступ
* Статическая/Динамическая структура

Эти признаки противостоят друг другу лишь внутри пары, вне её они могут сочетаться.

# Архитектура эвм

Архитектура представляет собой совокупность общих принципов организации аппаратно-программных средств и их характериситик, определяющая функциональные возможности ЭВМ при решении соответствюущих классов задач. Все ЭВМ основаны на структуре Джона фон Нейман. Фон Неймановоской называется организация ЭВМ, при которой вычислительная машина состоит из двух основных частей: линейно-адресуемая память, слова которой хранят команды и элементы данных, и процессора, выполняющего эти команды. В основе моделей вычислений фон Неймана лежит принцип последовательной передачи управления (счётчик команд) и концепция переменной (идентификатор)

С Х Е М А



## Принципы фон Неймана

1. Принцип двоичного кодирования – вся информация, поступающая в ЭВМ, кодируется с помощью двоичных сигналов и разделяется на единицы, называемые словами.
2. Принцип однородности памяти – программы и данные хранятся в однрой и той же памяти, над командами можно выполнять те же действия, что и над данными.
3. Принцип параллельной организации вычислений – операции над числом проводятся по всем разрядам одновременно.
4. Принцип адресуемости памяти – основная память состоит из пронумерованных ячеек, процессору в произвольный момент времени доступна любая ячейка. Отсюда следует возможность давать имена областям памяти так, чтобы к хранящимся в них значениям можно было обращаться и изменять их в процессе выполнения программы.
5. Принцип программного управления – программа состоит из набора команд, которые выполняются процессором автоматически, друг за другом в определённой последовательности.
6. Принцип жёсткости архитектуры – неизменяемость в процессе работы топологии, архитектуры, списка команд.

## Следствия из принципов

1. Программа уже не постоянная часть ЭВМ
2. Компьютер состоит из арифметико-логического устройства, выполняющего арифметические и логические операции (да ладно?), устройства управления, предназначенного для управления организацией программы, и запоминающих устройств.
3. Внешнее устройство ввода-вывода данных.