

ИНФОРМАТИКА



Понятие информации, общая характеристика процессов сбора, передачи, обработки и накопления информации

Лекция 1

Учебные вопросы лекции

Понятие информации, общая характеристика процессов сбора, передачи, обработки и накопления информации.

- Вопрос 1. Предмет информатики, основные задачи информатики. Место информатики в ряду других фундаментальных наук. История становления информатики
- Вопрос 2. Теоретические основы информатики. Сигналы. Данные. Информация. Информационный процесс
- Вопрос 3. Файлы и файловая структура
- Вопрос 4. Количество и качество информации. Единицы представления, измерения и хранения информации. Кодирование информации. Системы счисления

1.

Предмет информатики, основные задачи информатики.

Место информатики в ряду других фундаментальных наук.

История становления информатики.

Предмет информатики, основные задачи информатики

- ▣ ***Информатика — это техническая наука, систематизирующая приемы создания, хранения, воспроизведения, обработки и передачи данных средствами вычислительной техники, а также принципы функционирования этих средств и методы управления ими.***
- ▣ Из этого определения видно, что информатика очень близка к технологии, поэтому ее предмет нередко называют *информационной технологией*.
- ▣ «Информатика» (от франц.) = «информация» + «автоматика», что выражает ее суть как науки об автоматической обработке информации.

Предмет информатики, основные задачи информатики

- Предмет информатики составляют следующие понятия:
- • аппаратное обеспечение средств вычислительной техники;
- • программное обеспечение средств вычислительной техники;
- • средства взаимодействия аппаратного и программного обеспечения;
- • средства взаимодействия человека с аппаратными и программными средствами.
- Как видно из этого списка, в информатике особое внимание уделяется вопросам *взаимодействия*. Для этого даже есть специальное понятие — *интерфейс*. Методы и средства взаимодействия человека с аппаратными и программными средствами называют *пользовательским интерфейсом*. Соответственно, существуют *аппаратные интерфейсы, программные интерфейсы и аппаратно-программные интерфейсы*.

Предмет информатики, основные задачи информатики

- ▣ *Основной задачей информатики является систематизация приемов и методов работы с аппаратными и программными средствами вычислительной техники. Цель систематизации состоит в выделении, внедрении и развитии передовых, наиболее эффективных технологий, в автоматизации этапов работы с данными, а также в методическом обеспечении новых технологических исследований.*
- ▣ Информатика — практическая наука. Ее достижения должны проходить подтверждение практикой и приниматься в тех случаях, когда они соответствуют критерию повышения эффективности.

Предмет информатики, основные задачи информатики

В составе основной задачи информатики сегодня можно выделить следующие направления для практических приложений:

- архитектура вычислительных систем;
- интерфейсы вычислительных систем;
- программирование;
- преобразование данных;
- защита информации;
- автоматизация;
- стандартизация (обеспечение совместимости между аппаратными и программными средствами).

Тема 2. Теоретические основы информатики.

**Сигналы. Данные. Информация.
Информационный процесс.**

Теоретические основы информатики

- Теоретические основы информатики включают в себя :
- ❖ исследования процессов создания, накопления и обработки информации;
- ❖ исследования методов преобразования информации в данные и знания;
- ❖ создание и исследование информационных моделей, моделей данных и знаний, методов работы со знаниями, методов машинного обучения и обнаружения новых знаний; исследования принципов создания и функционирования аппаратных и программных средств автоматизации указанных процессов.

Сигналы. Данные

- Все процессы в природе сопровождаются сигналами. Зарегистрированные сигналы образуют данные. Данные преобразуются, транспортируются и потребляются с помощью методов. При взаимодействии данных и адекватных им методов образуется информация.
- Информация — это динамический объект, образующийся в ходе информационного процесса. Он отражает диалектическую связь между объективными данными и субъективными методами. Свойства информации зависят как от свойств данных, так и от свойств методов.

Сигналы. Данные

- Данные различаются *типами*, что связано с различиями в физической природе сигналов, при регистрации которых образовались данные. В качестве средства хранения и транспортировки данных используются *носители данных*. Для удобства операций с данными их структурируют.
- Наиболее широко используются следующие структуры: *линейная, табличная и иерархическая* — они различаются методом адресации к данным. При сохранении данных образуются данные нового типа — *адресные данные*.

Сигналы. Данные

- Вопросами систематизации приемов и методов создания, хранения, воспроизведения, обработки и передачи данных средствами вычислительной техники занимается техническая наука — *информатика*.
- С целью унификации приемов и методов работы с данными в вычислительной технике применяется универсальная система кодирования данных, называемая *двоичным кодом*. Элементарной единицей представления данных в двоичном коде является *двоичный разряд (бит)*. Другой, более крупной единицей представления данных является *байт*.

Сигналы. Данные

- Основной единицей хранения данных является *файл*. *Файл представляет собой последовательность байтов, имеющую собственное имя. Совокупность файлов образует файловую структуру, которая, как правило, относится к иерархическому типу.*
- *Полный адрес файла в файловой структуре является уникальным и включает в себя собственное имя файла и путь доступа к нему.*

Носители данных

- Данные — диалектическая составная часть информации. Они представляют собой зарегистрированные сигналы. При этом физический метод регистрации может быть любым: механическое перемещение физических тел, изменение их формы или параметров качества поверхности, изменение электрических, магнитных, оптических характеристик, химического состава и (или) характера химических связей, изменение состояния электронной системы и многое другое. В соответствии с методом регистрации данные могут храниться и транспортироваться на носителях различных видов.
- Самым распространенным носителем данных, хотя и не самым экономичным, по-видимому, является бумага. На бумаге данные регистрируются путем изменения оптических характеристик ее поверхности.

Носители данных

- Изменение оптических свойств (изменение коэффициента отражения поверхности в определенном диапазоне длин волн) используется также в устройствах, осуществляющих запись лазерным лучом на пластмассовых носителях с отражающим покрытием (*CD и DVD-ROM*). В качестве носителей, использующих изменение магнитных свойств, можно назвать магнитные ленты и диски.
- Регистрация данных путем изменения химического состава поверхностных веществ носителя широко используется в фотографии. На биохимическом уровне происходит накопление и передача данных в живой природе.
- Носители данных интересуют нас не сами по себе, а постольку, поскольку свойства информации весьма тесно связаны со свойствами ее носителей.

Носители данных

- Любой носитель можно характеризовать параметром *разрешающей способности* (количеством данных, записанных в принятой для носителя единице измерения) и *динамическим диапазоном* (логарифмическим отношением интенсивности амплитуд максимального и минимального регистрируемого сигналов). От этих свойств носителя нередко зависят такие свойства информации, как полнота, доступность и достоверность.
- Например, для обычного потребителя доступность информации в книге заметно выше, чем той же информации на компакт-диске, поскольку не все потребители обладают необходимым оборудованием.

Носители данных

- И наконец, известно, что визуальный эффект от просмотра слайда в проекторе намного больше, чем от просмотра аналогичной иллюстрации, напечатанной на бумаге, поскольку диапазон яркостных сигналов в проходящем свете на два-три порядка больше, чем в отраженном.
- Задача преобразования данных с целью смены носителя относится к одной из важнейших задач информатики. В структуре стоимости вычислительных систем устройства для ввода и вывода данных, работающие с носителями информации, составляют до половины стоимости аппаратных средств.

Операции с данными

- В ходе информационного процесса данные преобразуются из одного вида в другой с помощью методов. Обработка данных включает в себя множество различных операций. По мере развития научно-технического прогресса и общего усложнения связей в человеческом обществе трудозатраты на обработку данных неуклонно возрастают. Прежде всего это связано с постоянным усложнением условий управления производством и обществом. Второй фактор, также вызывающий общее увеличение объемов обрабатываемых данных, тоже связан с научно-техническим прогрессом, а именно с быстрыми темпами появления и внедрения новых носителей данных, средств их хранения и доставки.

Операции с данными

- В структуре возможных операций с данными можно выделить следующие основные:
- • *сбор данных* — накопление информации с целью обеспечения достаточной полноты для принятия решений;
- • *формализация данных* — приведение данных, поступающих из разных источников, к одинаковой форме, чтобы сделать их сопоставимыми между собой, то есть повысить их уровень доступности;
- • *фильтрация данных* — отсеивание "лишних" данных, в которых нет необходимости для принятия решений; при этом должен уменьшаться уровень "шума", а достоверность и адекватность данных должны возрасти;
- • *сортировка данных* — упорядочение данных по заданному признаку с целью удобства использования повышает доступность информации;

Операции с данными

- • *архивация данных* — организация хранения данных в удобной и легкодоступной форме; служит для снижения экономических затрат по хранению данных и повышает общую надежность информационного процесса в целом;
- • *защита данных* — комплекс мер, направленных на предотвращение утраты, воспроизведения и модификации данных;
- • *транспортировка данных* — прием и передача (доставка и поставка) данных между удаленными участниками информационного процесса; при этом источник данных в информатике принято называть *сервером*, а потребителя — *клиентом*;
- • *преобразование данных* — перевод данных из одной формы в другую или из одной структуры в другую.

Операции с данными

- Преобразование данных часто связано с изменением типа носителя: например книги можно хранить в обычной бумажной форме, но можно использовать для этого и электронную форму, и микрофотопленку.
- Необходимость в многократном преобразовании данных возникает также при их транспортировке, особенно если она осуществляется средствами, не предназначенными для транспортировки данного вида данных. В качестве примера можно упомянуть, что для транспортировки цифровых потоков данных по каналам телефонных сетей (которые изначально были ориентированы только на передачу аналоговых сигналов в узком диапазоне частот) необходимо преобразование цифровых данных в некое подобие звуковых сигналов, чем и занимаются специальные устройства — *телефонные модемы*.

Операции с данными

- Приведенный здесь список типовых операций с данными далеко не полон. Миллионы людей во всем мире занимаются созданием, обработкой, преобразованием и транспортировкой данных, и на каждом рабочем месте выполняются свои специфические операции, необходимые для управления социальными, экономическими, промышленными, научными и культурными процессами. Полный список возможных операций составить невозможно, да и не нужно. Сейчас нам важен другой вывод: *работа с информацией может иметь огромную трудоемкость и ее надо автоматизировать.*

Информация. Информационный процесс

Информация. Информационный процесс

- Термин "информация" происходит от латинского *informatio*, что в переводе означает изложение, разъяснение. В обыденной жизни под этим словом понимают сведения, передаваемые людьми устным, письменным или другим образом.
- Информация – это неубывающий ресурс жизнеобеспечения. Деятельность отдельных людей, групп, коллективов и организаций сейчас все в большей степени начинает зависеть от их информированности и способности эффективно использовать имеющуюся информацию.

Информация. Информационный процесс

- ▣ **Информационный процесс** - совокупность последовательных действий (операций), производимых над информацией (в виде данных, сведений, фактов, идей, гипотез, теорий и пр.), для получения какого-либо результата (достижения цели).
- ▣ Информация проявляется именно в информационных процессах. Информационные процессы всегда протекают в каких-либо системах (социальных, социотехнических, биологических и пр.).
- ▣ **Наиболее обобщенными информационными процессами являются сбор, преобразование, использование информации.**

Информация. Информационный процесс

- **К основным информационным процессам, изучаемым в курсе информатики, относятся: поиск, отбор, хранение, передача, кодирование, обработка и защита информации.**
- Информационные процессы, осуществляемые по определенным информационным технологиям, составляет основу информационной деятельности человека.
- Компьютер является универсальным устройством для автоматизированного выполнения информационных процессов.

Вопросы для самоконтроля

- 1. Как вы можете объяснить бытовой термин «переизбыток информации»? Что имеется в виду: излишняя полнота данных; излишняя сложность методов; неадекватность поступающих данных и методов, имеющих в наличии?
- 2. Как вы понимаете термин «средство массовой информации»? Что это? Средство массовой поставки данных? Средство, обеспечивающее массовое распространение методов? Средство, обеспечивающее процесс информирования путем поставки данных гражданам, обладающим адекватными методами их потребления?
- 3. Как вы полагаете, являются ли данные товаром? Могут ли методы быть товаром?
- 4. На примере коммерческих структур, обеспечивающих коммуникационные услуги, покажите, как взаимодействуют между собой маркетинг данных и маркетинг методов? Можете ли вы привести примеры лизинга данных и методов?
- 5. Как вы понимаете диалектическое единство данных и методов? Можете ли вы привести примеры аналогичного единства двух понятий из других научных дисциплин: естественных, социальных, технических?
- 6. Как вы понимаете динамический характер информации? Что происходит с ней по окончании информационного процесса?
- 7. Можем ли мы утверждать, что данные, полученные в результате информационного процесса, адекватны исходным? Почему? От каких свойств исходных данных и методов зависит адекватность результирующих данных?
- 8. Что такое *вектор данных*? Является ли список номеров телефонов населенного пункта вектором данных? Является ли вектором данных текстовый документ, закодированный двоичным кодом, если он не содержит элементов оформления?
- 9. Является ли цифровой код цветного фотоснимка вектором данных? Если нет, то чего ему не хватает?
- 10. Как вы понимаете следующие термины: *аппаратно-программный интерфейс, программный интерфейс, аппаратный интерфейс*? Как бы вы назвали специальность людей, разрабатывающих аппаратные интерфейсы? Кад называется специальность людей, разрабатывающих программные интерфейсы?
- 11. На основе личных наблюдений сделайте вывод о том, какими средствами может пользоваться преподаватель для обеспечения интерфейса с аудиторией. Можете ли вы рассмотреть отдельно методические и технические средства, имеющиеся в его распоряжении? Может ли преподаватель рассматривать *вашу тетрадь и авторучку* как свое средство обеспечения интерфейса? Если да, то в какой мере?

3.

Файлы и файловая структура

Единицы представления данных

- Существует множество систем представления данных. С одной из них, принятой в информатике и вычислительной технике, двоичным кодом, мы уже знакомы.
- Наименьшей единицей такого представления является бит (*двоичный разряд*). Совокупность двоичных разрядов, выражающих числовые или иные данные, образует некий битовый рисунок. Практика показывает, что с битовым представлением удобнее работать, если этот рисунок имеет регулярную форму. В настоящее время в качестве таких форм используются группы из восьми битов, которые называются *байтами*.

числовое (буквенное)значение	Двоичный код
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
A	1010
B	1011
C	1100
D	1101
E	1110
F	1111

Единицы представления данных

Десятичное число	Двоичное число	байт
1	1	0000 0001
2	10	0000 0010
...
255	11111111	1111 1111

- Понятие о байте как группе взаимосвязанных битов появилось вместе с первыми образцами электронной вычислительной техники.
- Долгое время оно было *машинно-зависимым*, то есть для разных вычислительных машин длина байта была разной.

Единицы представления данных

- Во многих случаях целесообразно использовать не восьмиразрядное кодирование, а 16-разрядное, 24-разрядное, 32-разрядное и более. Группа из 16 взаимосвязанных бит (двух взаимосвязанных байтов) в информатике называется словом. Соответственно, группы из четырех взаимосвязанных байтов (32 разряда) называются удвоенным словом, а группы из восьми байтов (64 разряда) — учетверенным словом. Пока, на сегодняшний день, такой системы обозначения достаточно.

Единицы измерения данных

- Существует много различных систем и единиц измерения данных. Каждая научная дисциплина и каждая область человеческой деятельности может использовать свои, наиболее удобные или традиционно устоявшиеся единицы. В информатике для измерения данных используют тот факт, что разные типы данных имеют универсальное двоичное представление и потому вводят свои единицы данных, основанные на нем.
- Наименьшей единицей измерения является байт. Поскольку одним байтом, как правило, кодируется один символ текстовой информации, то для текстовых документов размер в байтах соответствует лексическому объему в символах (пока исключение представляет универсальная кодировка *UNICODE*).

Кодирование текстовой информации

- Текстовую информацию кодируют двоичным кодом через обозначение каждого символа алфавита определенным целым числом. С помощью восьми двоичных разрядов возможно закодировать 256 различных символов. Данного количества символов достаточно для выражения всех символов английского и русского алфавитов.
- В первые годы развития компьютерной техники трудности кодирования текстовой информации были вызваны отсутствием необходимых стандартов кодирования. В настоящее время, напротив, существующие трудности связаны с множеством одновременно действующих и зачастую противоречивых стандартов.
- Для английского языка, который является неофициальным международным средством общения, эти трудности были решены. Институт стандартизации США выработал и ввел в обращение *систему кодирования ASCII* (*American Standard Code for Information Interchange* – стандартный код информационного обмена США).
- Для кодировки русского алфавита были разработаны несколько вариантов кодировок:
 - 1) Windows-1251 – введена компанией *Microsoft*; с учетом широкого распространения операционных систем (ОС) и других программных продуктов этой компании в Российской Федерации она нашла широкое распространение;
 - 2) КОИ-8 (Код Обмена Информацией, восьмизначный) – другая популярная кодировка русского алфавита, распространенная в компьютерных сетях на территории Российской Федерации и в российском секторе Интернет;
 - 3) ISO (International Standard Organization – Международный институт стандартизации) – международный стандарт кодирования символов русского языка. На практике эта кодировка используется редко.
- Ограниченный набор кодов (256) создает трудности для разработчиков единой системы кодирования текстовой информации. Вследствие этого было предложено кодировать символы не 8-разрядными двоичными числами, а числами с большим разрядом, что вызвало расширение диапазона возможных значений кодов. Система 16-разрядного кодирования символов называется *универсальной* – UNICODE. Шестнадцать разрядов позволяет обеспечить уникальные коды для 65 536 символов, что вполне достаточно для размещения в одной таблице символов большинства языков.
- Несмотря на простоту предложенного подхода, практический переход на данную систему кодировки очень долго не мог осуществиться из-за недостатков ресурсов средств вычислительной техники, так как в системе кодирования UNICODE все текстовые документы становятся автоматически вдвое больше. В конце 1990-х гг. технические средства достигли необходимого уровня, начался постепенный перевод документов и программных средств на систему кодирования UNICODE.

Единицы измерения данных

- Более крупная единица измерения — килобайт (Кбайт). Условно можно считать, что 1 Кбайт примерно равен 1000 байт. Условность связана с тем, что для вычислительной техники, работающей с двоичными числами, более удобно представление чисел в виде степени двойки и потому на самом деле 1 Кбайт равен 2^{10} байт (1024 байт).
- Однако всюду, где это не принципиально, с инженерной погрешностью (до 3 %) "забывают" о "лишних" байтах.
- В килобайтах измеряют сравнительно небольшие объемы данных. Условно можно считать, что одна страница неформатированного машинописного текста составляет около 2 Кбайт.

Единицы измерения данных

- Более крупные единицы измерения данных образуются добавлением префиксов *мега-*, *гига-*, *тера-*; в более крупных единицах пока нет практической надобности.
- 1 Мбайт - 1024 Кбайт = 10^{20} байт
- 1 Гбайт = 1024 Мбайт = 10^{30} байт
- 1 Тбайт = 1024 Гбайт = 10^{40} байт
- При переходе к более крупным единицам "инженерная" погрешность, связанная с округлением, накапливается и становится недопустимой, поэтому на старших единицах измерения округление производится реже.

4.
Количество и качество информации. Единицы
представления, измерения и хранения
информации.
Кодирование информации.
Системы счисления.

Количество и качество информации

- Возможность и эффективность использования обуславливаются такими основными её потребительскими *показателями качества*, как репрезентативность, содержательность, достаточность, доступность, актуальность, своевременность, точность, достоверность, устойчивость.
- *Репрезентативность* информации связана с правильностью её отбора и формирования в целях адекватного отражения свойств объекта. Важнейшее значение здесь имеют:
 1. Правильность концепции, на базе которой сформулировано исходное понятие.
 2. Обоснованность отбора существенных признаков и связей отображаемого явления.

Количество и качество информации

- Нарушение репрезентативности информации отражает семантическую ёмкость, равную отношению количества семантической информации в сообщении к объёму обрабатываемых данных.
- С увеличением содержательности информации растёт семантическая пропускная способность информационной системы, так как для получения одних и тех же сведений требуется преобразовать меньший объём данных.
- Наряду с коэффициентом содержательности, отражающим семантический аспект, можно использовать и коэффициент информативности, характеризующийся отношением количества информации к объёму данных.

Количество и качество информации

- *Достаточность (полнота)* информации означает, что она содержит минимальный, но достаточный для принятия правильного решения состав (набор показателей). Понятие полноты информации связано с её смысловым содержанием (семантикой) и прагматикой. Как неполная, т.е. недостаточная для принятия правильного решения, так и избыточная информация снижает эффективность принимаемых пользователем решений.
- *Доступность* информации восприятию пользователя обеспечивается выполнением соответствующих процедур её получения и преобразования. Например, в информационной системе информация преобразовывается к доступной и удобной для восприятия пользователя форме. Это достигается, в частности, и путём согласования её семантической формы с тезаурусом (терминологическим словарным запасом)

Количество и качество информации

- *Актуальность* информации определяется степенью сохранения ценности информации для управления в момент её использования и зависит от динамики изменения её характеристик и от интервала времени, прошедшего с момента возникновения данной информации.
- *Своевременность* информации означает её поступление не позже заранее назначенного момента времени, согласованного с временем решения поставленной задачи.

Количество и качество информации

- Точность информации определяется степенью близости получаемой информации к реальному состоянию объекта, процесса, явления и т.п. Для информации, отображаемой цифровым кодом, известны четыре классификационных понятия точности:
 1. Формальная точность, измеряемая значением единицы младшего разряда числа.
 2. Реальная точность, определяемая значением единицы последнего разряда числа, верность которого гарантируется.
 3. Максимальная точность, которую можно получить в конкретных условиях функционирования системы.
 4. Необходимая точность, определяемая функциональным назначением показателя.

Количество и качество информации

- ▣ *Достоверность* информации определяется её свойством отражать реально существующие с необходимой точностью. Измеряется достоверность информации доверительной вероятностью необходимой точности, т.е. вероятностью того, что отображаемое информацией значение параметра отличается от истинного значения этого параметра в пределах необходимой точности.
- ▣ Устойчивость информации отражает её способность реагировать на изменения исходных данных без нарушения необходимой точности. Устойчивость информации, как и репрезентативность, обусловлена выбранной методикой её отбора и формирования.

Количество и качество информации

- В заключение следует отметить, что такие параметры качества информации, как репрезентативность, содержательность, достаточность, доступность, устойчивость, целиком определяются на методическом уровне разработки информационных систем.
- Параметры актуальности, своевременности, точности и достоверности обуславливаются в большей степени также на методическом уровне, однако на их величину существенно влияет и характер функционирования системы, в первую очередь её надёжность. При этом параметры актуальности и точности жёстко связаны соответственно с параметрами своевременности и достоверности.

Кодирование информации



Кодирование информации

- Для автоматизации работы с данными, относящимися к различным типам, очень важно унифицировать их форму представления – для этого обычно используется прием *кодирования*, то есть выражение данных одного типа через данные другого типа. Естественные человеческие *языки* – это не что иное, как системы кодирования понятий для выражения мыслей посредством речи.
- К языкам близко примыкают *азбуки* (системы кодирования компонентов языка с помощью графических символов). История знает интересные, хотя и безуспешные попытки создания «универсальных» языков и азбук.
- Та же проблема универсального средства кодирования достаточно успешно реализуется в отдельных отраслях техники, науки и культуры. В качестве примеров можно привести систему записи математических выражений, телеграфную азбуку, морскую флажковую азбуку, систему Брайля для слепых и многое другое.

Кодирование информации

□ **C O M P U T E R**

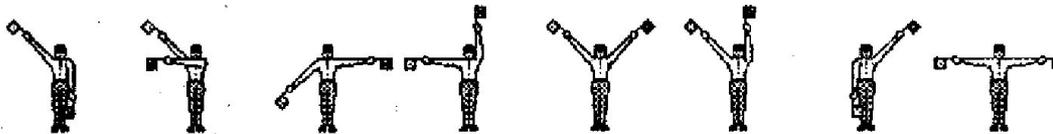
□ 43 4F 4D 50 55 54 45 52 Код ASCII

□ -. - . - - - - - . - . - - - - - - . . - - - . - . Код Морзе

□



Код Брайля



Код морской сигнальный

□ Рис. 1.2. Примеры различных систем кодирования

Кодирование информации

- Своя система существует и в вычислительной технике – она называется *двоичным кодированием* и основана на представлении данных последовательностью всего двух знаков: 0 и 1. Эти знаки называются *двоичными цифрами*, по-английски – *binary digit*, или, сокращенно, *bit* (*бит*).
- Одним битом могут быть выражены два понятия: 0 или 1 (*да* или *нет*, *черное* или *белое*, *истина* или *ложь* и т. п.). Если количество битов увеличить до двух, то уже можно выразить четыре различных понятия: 00 **01** 10 **11**.
- Тремя битами можно закодировать восемь различных значений: 000 **001** 010 **011** 100 **101** 110 **111**.
- Увеличивая на единицу количество разрядов в системе двоичного кодирования, мы увеличиваем в два раза количество значений, которое может быть выражено в данной системе.

Кодирование целых и действительных чисел

- Для кодирования целых чисел от 0 до 255 достаточно иметь 8 разрядов двоичного кода (8 бит).
- 0000 0000 = 0
- 0000 0001 = 1
-
- 1111 1110 = **254**
- 1111 1111 = **255**.
- Шестнадцать бит позволяют закодировать целые числа от 0 до **65 535**,
- а 24 бита – уже более **16,5** миллионов разных значений.

Кодирование целых и действительных чисел

- Для кодирования действительных чисел используют 80-разрядное кодирование. При этом число предварительно преобразуется в *нормализованную форму*:
- $3,1415926 = 0,31415926 * 10^1$
- $300\ 000 = 0,3 * 10^6$
- $123\ 456\ 789 = 0,123456789 * 10^9$.
- Первая часть числа называется *мантиссой*, а вторая – *характеристикой (порядком)*. Большую часть из 80 бит отводят для хранения мантиссы (вместе со знаком) и некоторое фиксированное количество разрядов отводят для хранения характеристики (тоже со знаком).

Кодирование текстовых данных

- Если каждому символу алфавита сопоставить определенное целое число (например порядковый номер), то с помощью двоичного кода можно кодировать и текстовую информацию. Восемью двоичных разрядов достаточно для кодирования 256 различных символов. Этого хватит, чтобы выразить различными комбинациями восьми битов все символы английского и русского алфавитов как строчные, так и прописные, а также знаки препинания, символы основных арифметических действий и некоторые общепринятые специальные символы, например символ «\$».

Кодирование текстовых данных

- Технически это выглядит очень просто, однако всегда существовали достаточно веские организационные сложности. В первые годы развития вычислительной техники они были связаны с отсутствием необходимых стандартов, а в настоящее время вызваны, наоборот, избытком одновременно действующих и противоречивых стандартов. Для того чтобы весь мир одинаково кодировал текстовые данные, нужны единые таблицы кодирования, а это пока невозможно из-за противоречий между символами национальных алфавитов, а также противоречий корпоративного характера. Для английского языка, захватившего де-факто нишу международного средства общения, противоречия уже сняты.

Кодирование текстовых данных

- Институт стандартизации США (*ANSI – American National Standard Institute*) ввел в действие систему кодирования *ASCII (American Standard Code for Information Interchange – стандартный код информационного обмена США)*.
- 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 32 кода базовой таблицы, начиная с нулевого, отданы производителям аппаратных средств (в первую очередь производителям компьютеров и 19 печатающих устройств). В этой области размещаются так называемые *управляющие коды*, которым не соответствуют никакие символы языков, и, соответственно, эти коды не выводятся ни на экран, ни на устройства печати, но ими можно управлять тем, как производится вывод прочих данных.

Кодирование текстовых данных

- Начиная с кода 32 по код 127 размещены коды символов английского алфавита, знаков препинания, цифр, арифметических действий и некоторых вспомогательных символов.
- Аналогичные системы кодирования текстовых данных были разработаны и в других странах. Так, например, в СССР в этой области действовала система кодирования КОИ-7 (*код обмена информацией, семизначный*). Однако поддержка производителей оборудования и программ вывела американский код *ASCII* на уровень международного стандарта, и национальным системам кодирования пришлось «отступить» во вторую, расширенную часть системы кодирования, определяющую значения кодов со 128 по 255.

Кодирование текстовых данных

- Отсутствие единого стандарта в этой области привело к множественности одновременно действующих кодировок. Только в России можно указать три действующих стандарта кодировки и еще два устаревших.
- Так, например, кодировка символов русского языка, известная как кодировка *Windows-1251*, была введена «извне» – компанией Microsoft, но, учитывая широкое распространение операционных систем и других продуктов этой компании в России, она глубоко закрепилась и нашла широкое распространение. Эта кодировка используется на большинстве локальных компьютеров, работающих на платформе Windows. Де-факто она стала стандартной в российском секторе World Wide Web.

Кодирование текстовых данных

- Другая распространенная кодировка носит название КОИ-8 (*код обмена информацией, восьмизначный*) – ее происхождение относится ко временам действия Совета Экономической Взаимопомощи государств Восточной Европы. На базе этой кодировки ныне действуют кодировки КОИ8-Р (русская) и КОИ8-У (украинская). Сегодня кодировка КОИ8-Р имеет широкое распространение в компьютерных сетях на территории России и в некоторых службах российского сектора Интернета. В частности, в России она де-факто является стандартной в сообщениях электронной почты и телеконференций.
- Международный стандарт, в котором предусмотрена кодировка символов русского алфавита, носит название кодировки *ISO (International Standard Organization – Международный институт стандартизации)*. На практике данная кодировка используется редко.

Кодирование текстовых данных

- В связи с изобилием систем кодирования текстовых данных, действующих в России, возникает задача межсистемного преобразования данных – это одна из распространенных задач информатики.
- Если проанализировать организационные трудности, связанные с созданием единой системы кодирования текстовых данных, то можно прийти к выводу, что они вызваны ограниченным набором кодов (256). В то же время очевидно, что если, например, кодировать символы не восьмиразрядными двоичными числами, а числами с большим количеством разрядов, то и диапазон возможных значений кодов станет намного больше. Такая система, основанная на 16-разрядном кодировании символов, получила название *универсальной* – *UNICODE*. Шестнадцать разрядов позволяют обеспечить уникальные коды для 65 536 различных символов — этого поля достаточно для размещения в одной таблице символов большинства языков планеты.

Кодирование графических данных

- Если рассмотреть с помощью увеличительного стекла черно-белое графическое изображение, напечатанное в газете или книге, то можно увидеть, что оно состоит из мельчайших точек, образующих характерный узор, называемый *растром*. Поскольку линейные координаты и индивидуальные свойства каждой точки (яркость) можно выразить с помощью целых чисел, то можно сказать, что растровое кодирование позволяет использовать двоичный код для представления графических данных. Общепринятым на сегодняшний день считается представление черно-белых иллюстраций в виде комбинации точек с 256 градациями серого цвета, и, таким образом, для кодирования яркости любой точки обычно достаточно восьмиразрядного двоичного числа.

Кодирование графических данных

- Для кодирования цветных графических изображений применяется *принцип декомпозиции* произвольного цвета на основные составляющие. В качестве таких составляющих используют три основных цвета: красный (*Red, R*), зеленый (*Green, G*) и синий (*Blue, B*). На практике считается (хотя теоретически это не совсем так), что любой цвет, видимый человеческим глазом, можно получить путем механического смешения этих трех основных цветов. Такая система кодирования называется системой *RGB* по первым буквам названий основных цветов.
- Если для кодирования яркости каждой из основных составляющих использовать по 256 значений (восемь двоичных разрядов), как это принято для полутоновых черно-белых изображений, то на кодирование цвета одной точки надо затратить 24 разряда.

Кодирование графических данных

- При этом система кодирования обеспечивает однозначное определение 16,5 млн. различных цветов, что на самом деле близко к чувствительности человеческого глаза. Режим представления цветной графики с использованием 24 двоичных разрядов называется *полноцветным (True Color)*.
- Если уменьшить количество двоичных разрядов, используемых для кодирования цвета каждой точки, то можно сократить объем данных, но при этом диапазон кодируемых цветов заметно сокращается. Кодирование цветной графики 16-разрядными двоичными числами называется режимом *High Color*.

Системы счисления



Системы счисления

- Система счисления – это совокупность правил и приемов записи чисел с помощью набора цифровых знаков (алфавита). Различают два типа систем счисления: **позиционные**, когда значение каждой цифры числа определяется ее местом (позицией) в записи числа; и **непозиционные**, когда значение цифры в числе не зависит от ее места в записи числа.
- Примером непозиционной системы счисления являются римские цифры: IX, IV, XV, LX и т.д., а примером позиционной системы счисления можно назвать арабские цифры, используемые нами повседневно: 12, 67, 329 и т.д. Позиционные системы счисления характеризуются *основанием* – количеством знаков или символов, используемых в разрядах для изображения числа в данной системе счисления.

Системы счисления

Системы счисления	Значения							
Десятеричная	0	1	2	3	4	5	6	7
Двоичная	0	1	10	11	100	101	110	111
Восьмеричная	0	1	2	3	4	5	6	7
Шестнадцатеричная	0	1	2	3	4	5	6	7
Десятеричная	8	9	10	11	12	13	14	15
Двоичная	1000	1001	1010	1011	1100	1101	1110	1111
Восьмеричная	10	11	12	13	14	15	16	17
Шестнадцатеричная	8	9	A	B	C	D	E	F

Контрольные вопросы

1. Дайте определение информатики, как науки.
2. Какие понятия составляют предмет информатики?
3. *Что является Основной задачей информатики?*
4. Дайте определение информации.
5. Каковы единицы хранения информации?
6. **Как кодируются данные в вычислительной технике?**
7. **Что такое двоичный код?**
8. **Что такое бит?**
9. **Сколько различных символов (значений) можно закодировать тремя битами?**
10. **Сколько различных символов (значений) можно закодировать одним байтом?**



Контрольные вопросы

11. Сколько двоичных разрядов требуется для кодирования 65 536 целых чисел?
12. Что представляет собой нормализованная форма действительных чисел?
13. Сколько разрядов используется для кодирования в системе ASCII?
14. Какие две таблицы закреплены в системе ASCII?
15. Какие коды размещены в базовой таблице системы ASCII?
16. Какие коды размещены в расширенной таблице системы ASCII?
17. Какие системы кодирования текстовой информации вы знаете?
18. Как кодируется черно-белое изображение?
19. Как кодируются цветные графические изображения?
20. Сколько двоичных разрядов требуется для режимов True Color и High Color?

Кодирование звуковой информации

- Приемы и методы работы со звуковой информацией пришли в вычислительную технику наиболее поздно. К тому же, в отличие от числовых, текстовых и графических данных, у звукозаписей не было столь же длительной и проверенной истории кодирования. В итоге методы кодирования звуковой информации двоичным кодом далеки от стандартизации. Множество отдельных компаний разработали свои корпоративные стандарты, но если говорить обобщенно, то можно выделить два основных направления.
- Метод FM (*Frequency Modulation*) основан на том, что теоретически любой сложный звук можно разложить на последовательность простейших гармонических сигналов разных частот, каждый из которых представляет собой правильную синусоиду, а следовательно, может быть описан числовыми параметрами, то есть кодом.

Кодирование звуковой информации

- В природе звуковые сигналы имеют непрерывный спектр, то есть являются аналоговыми. Их разложение в гармонические ряды и представление в виде дискретных цифровых сигналов выполняют специальные устройства — аналогово-цифровые преобразователи {АЦП). Обратное преобразование для воспроизведения звука, закодированного числовым кодом, выполняют цифро-аналоговые преобразователи {ЦАП). При таких преобразованиях неизбежны потери информации, связанные с методом кодирования, поэтому качество звукозаписи обычно получается не вполне удовлетворительным и соответствует качеству звучания простейших электромузыкальных инструментов с окрасом, характерным для электронной музыки. В то же время, данный метод кодирования обеспечивает весьма компактный код, и потому он нашел применение еще в те годы, когда ресурсы средств вычислительной техники были явно недостаточны.

Кодирование звуковой информации

- Метод таблично-волнового (Wave-Table) синтеза лучше соответствует современному уровню развития техники. Если говорить упрощенно, то можно сказать, что где-то в заранее подготовленных таблицах хранятся образцы звуков для множества различных музыкальных инструментов (хотя не только для них). В технике такие образцы называют сэмплами. Числовые коды выражают тип инструмента, номер его модели, высоту тона, продолжительность и интенсивность звука, динамику его изменения, некоторые параметры среды, в которой происходит звучание, а также прочие параметры, характеризующие особенности звука. Поскольку в качестве образцов используются "реальные" звуки, то качество звука, полученного в результате синтеза, получается очень высоким и приближается к качеству звучания реальных музыкальных инструментов.

Литература по курсу «Информатика»

- 1. Информатика: Базовый курс: учебник для студентов / под ред. С. В. Симоновича. СПб.: Питер, 2002.
- 2. *Левин А. Ш.* Самоучитель работы на компьютере/А. Ш. Левин. 8-е изд. СПб.: Питер, 2004.
- 3. *Леонтьев В. П.* Новейшая энциклопедия персонального компьютера 2003 / В. П. Леонтьев. М.: ОЛМА-Пресс, 2003.
- 4. *Могилев А. В.* Информатика: учеб. пособие для студентов/ А. В. Могилев, Н. И. Пак, Е. К. Хеннер; под ред. Е. К. Хеннера. М.: Академия, 2001.
- 5. *Мураховский В. И.* Железо персонального компьютера: Практическое руководство / В. И. Мураховский, Г. А. Евсеев. М.: ДЕСС КОМ, 2001.
- 6. *Олифер В. Г.* Компьютерные сети. Принципы, технологии, протоколы: учебник для студентов / В. Г. Олифер, Н. А. Олифер. СПб.: Питер, 2001.

□