

ЛЕКЦИЯ 3-4

Представление информации в ЭВМ, методы кодирования

Представление символьной информации.


Представление числовой информации.

Представление графической информации.

Звук в памяти компьютера

Представление символьной информации

Для кодирования одного символа в ЭВМ используется 8 бит (8 разрядный), что позволяет закодировать 256 различных символов. Стандартный набор из 256 символов называется ASCII. Набор ASCII был разработан в США Американским Национальным Институтом Стандартов (ANSI), но может быть использован и в других странах, поскольку вторая половина из 256 стандартных символов, т.е. 128 символов, могут быть с помощью специальных программ заменены на другие, в частности на символы национального алфавита, в нашем случае - буквы кириллицы. Объем информации, необходимый для запоминания одного символа ASCII - 1 байт.



Трудности кодирования связаны с ограничением – набором из 256 кодов. Очевидно, что если кодировать символы не 8 разрядным числом, то и диапазон станет больше. Такая система, основанная на 16 разрядном кодировании получила название универсальной - UNICODE (UNiversalCODE). $2^{16} = 65\ 536$ различных символов. Во второй половине 90 – годов технические средства достигли необходимого уровня и сегодня осуществлен переход и программных средств и документов на универсальную систему кодирования. Единая стандартизация форматов наводит порядок в международном кодировании алфавитов различных языков.

Представление числовой информации. Для представления чисел в памяти компьютера используются два формата: формат с фиксированной точкой (*Fixed*) и формат с плавающей точкой (*Float*). В формате с фиксированной точкой представляются только целые числа, в формате с плавающей точкой - вещественные числа (дробные и целые).

Целые числа. Множество целых чисел, представимых в памяти ЭВМ, ограничено. Диапазон значений зависит от размера ячеек памяти, используемых для их хранения. В k -разрядной ячейке может храниться 2^k различных значений целых чисел.

Например, в компьютере используется 16-разрядная ячейка (2 байта). Диапазон хранимых чисел - всего в 16-разрядной ячейке может храниться $2^{16} = 65536$ различных значений. Если используются только положительные числа - диапазон значений от 0 до 65 535 (от 0 до $2^k - 1$); Если используются как положительные, так и отрицательные числа - диапазон значений от -32768 до 32767. Чтобы получить внутреннее представление целого положительного числа N , хранящегося в k -разрядном машинном слове, необходимо: перевести число N в двоичную систему счисления: полученный результат дополнить слева незначащими нулями до k -разрядов.

Для записи внутреннего представления целого отрицательного числа $(-N)$ необходимо: получить внутреннее представление положительного числа N ; получить обратный код этого числа заменой 0 на 1, 1 на 0; к полученному числу прибавить 1. Данная форма представления целого отрицательного числа называется дополнительным кодом. Использование дополнительного кода позволяет заменить вычитание, на операцию сложения уменьшаемого числа с дополнительным кодом вычитаемого.

Например, получим внутреннее представление целого отрицательного числа «-1607».

- внутреннее представление положительного числа (прямой код): 0000 0110 0100 0111
- обратный код: 1111 1001 1011 1000
- результат прибавления 1: 1111 1001 1011 1001 - это внутреннее двоичное представление числа «-1607». Шестнадцатеричная форма которого - F9B9.

Двоичные разряды в ячейке памяти нумеруются от 0 до k справа налево. Старший, k -й разряд во внутреннем представлении любого положительного числа равен 0, отрицательного - 1. Поэтому этот разряд называется знаковым разрядом.

Вещественные числа. Формат с плавающей точкой использует представление вещественного числа R в виде произведения мантиссы m на основание системы счисления n в некоторой целой степени p , которую называют порядком: $R = m \cdot n^p$.

В ЭВМ используют нормализованное представление числа в форме с плавающей точкой. Мантисса должна удовлетворять условию: $0,1 \leq m < 1$.

В памяти компьютера мантисса представляется как целое число, содержащее только значащие цифры (0 и запятая не хранятся). Следовательно, внутреннее представление вещественного числа сводится к представлению пары целых чисел: мантиссы и порядка.

В разных типах ЭВМ применяются различные варианты представления чисел в форме с плавающей запятой.

Рассмотрим внутреннее представление вещественного числа в 4-х байтовой ячейке памяти. В ячейках должна содержаться следующая информация о числе: знак числа, порядок и значащие цифры мантииссы. В старшем бите 1-го байта хранится знак числа: 0- плюс, 1- минус. Оставшиеся 7 бит первого байта содержат машинный порядок. В следующих трех байтах хранятся значащие цифры мантииссы (всего 24 разряда). Порядок может быть и отрицательным и положительным. Разумно эти 128 значений разделить поровну между положительными и отрицательными значениями порядка от – 64 до 63.

Для записи внутреннего представления вещественного числа необходимо:

- перевести модуль данного числа в двоичную систему счисления с 24 значащими цифрами;
- нормализовать двоичное число;
 - найти машинный порядок в двоичной системе счисления по формуле: $Mr_2 = p_2 + 100\ 0000_2$ (Mr - машинный порядок, p - математический порядок в двоичной системе счисления);
 - учитывая знак числа, вписать его представление в 4-х байтовом машинном слове.
- Пример. Записать внутреннее числа 250,1875. $250,1875 = 11111010, 0011000000000000_2$ (перевели в двоичную систему счисления)
- $0, 111110100011000000000000_2 * 10_2^{1000}$ (мантиисса, основание системы счисления $2_{10} = 10_2$ и порядок $8_{10} = 1000_2$ записаны в двоичной системе)
- Вычислим машинный порядок в двоичной системе счисления: $Mr_2 = 1000 + 100\ 0000 = 100\ 1000$;
- Запишем в 4-х байтовой ячейке памяти с учетом знака числа: Шестнадцатеричная форма: 48FA3000.

Представление графической информации. Существует два принципиально разных подхода к тому, каким образом можно представить изображение в виде нулей и единиц (оцифровать изображение):

Растровое представление. При использовании растровой графики с помощью определенного числа бит кодируется цвет каждого элемента изображения - пикселя. При использовании растрового способа в ЭВМ под каждый пиксель отводится определенное число бит, называемое битовой глубиной. Каждому цвету соответствует определенный двоичный код (т.е. код из нулей и единиц). Например, если битовая глубина равна 1, т.е. под каждый пиксель отводится 1 бит, то 0 соответствует черному цвету, 1 - белому, а изображение может быть только черно-белым. Все многообразие красок на экране получается путем смешивания трех базовых цветов: красного, синего и зеленого. Цветные дисплеи, использующие этот принцип называются RGB (Red-Green-Blue)-мониторами.

Если битовая глубина равна 2, т.е. под каждый пиксель отводится 2 бита, 00- соответствует черному цвету, 01- красному, 10 - синему, 11- черному, т.е. в рисунке может использоваться четыре цвета. Далее, при битовой глубине 3 можно использовать 8 цветов, при 4 - 16 и т.д. Поэтому, графические программы позволяют создавать изображения из 2, 4, 8, 16, 32, 64, ..., 256, и т.д. цветов.

Пример. Сколько бит видеопамати занимает информация об одном пикселе на черно-белом экране (без полутонов)?

Для ответа воспользуемся формулой $K=2^N$, связывающей число цветов, воспроизводимых на экране дисплея (K) и число бит, отводимых в видеопамати под каждый пиксель (N). Для нашего случая $K=2$. Следовательно $2^N=2$, отсюда $N=1$ бит на пиксель. Величину N называют битовой глубиной.

Видеопамать – оперативная память, хранящая видеоинформацию о графическом изображении. Основным недостатком растровой графики является большой объем памяти, требуемый для хранения изображения. Это объясняется тем, что нужно запомнить цвет каждого пикселя, общее число которых может быть очень большим. Например, одна фотография среднего размера в памяти компьютера занимает несколько Мегабайт, т.е. столько же, сколько несколько сотен (а то и тысяч) страниц текста.

Векторное представление. При использовании векторной графики в памяти ЭВМ сохраняется математическое описание каждого графического примитива - геометрического объекта (например, отрезка, окружности, прямоугольника и т.п.), из которых формируется изображение. Основным недостатком векторной графики является невозможность работы с высококачественными художественными изображениями, фотографиями и фильмами.

Звук в памяти компьютера. Приемы и методы работы со звуковой информацией пришли в ВТ наиболее поздно (90-е годы). Поэтому методы кодирования звуковой информации двоичным кодом не стандартизированы. Существуют корпоративные стандарты:

1) Метод FM (Frequency Modulation). Физическая природа звука – колебания в определенном диапазоне частот, передаваемые звуковой волной через воздух (или другую упругую среду). Сложный звук можно разложить на последовательность простейших гармонических сигналов разных частот, каждый из которых представляет собой правильную синусоиду, а следовательно может быть описан числовым кодом.

Аудио адаптер (звуковая карта) - специальное устройство, предназначенное для преобразования электрических колебаний звуковой частоты в числовой двоичный код при вводе звука и для обратного преобразования (из числового кода в электрические колебания) при воспроизведении звука. В процессе записи звука аудиоадаптер с определенным периодом измеряет амплитуду электрического тока и заносит в регистр двоичный код полученной величины. Затем полученный код из регистра переписывается в оперативную память компьютера. Качество компьютерного звука определяется характеристиками аудио адаптера: частотой дискретизации и разрядностью.

Частота дискретизации - количество измерений входного сигнала за 1 сек. Частота измеряется в Гц. Одно измерение за одну секунду соответствует частоте 1 Гц. 1000 измерений за 1 сек. - 1 кГц. Характерные частоты дискретизации аудио адаптеров: 11 кГц, 22 кГц, 44 кГц и др.

Разрядность регистра - число бит в регистре аудио адаптера. Разрядность определяет точность измерения входного сигнала. Чем больше разрядность, тем меньше погрешность каждого отдельного преобразования величины электрического сигнала и обратно. Если разрядность равна 8 (16), то при измерении входного сигнала может быть получено $2^8=256$ ($2^{16}=65\,536$) различных значений. Очевидно, что 16-разрядный адаптер точнее кодирует и воспроизводит звук, чем 8-разрядный.

2) Метод таблично-волнового (Wave-Table) синтеза соответствует современному уровню развития ВТ. В таблицах хранятся образцы звуков для множества различных музыкальных инструментов (и не только для них). В технике такие образцы называют сэмплами. Числовые коды выражают тип инструмента, номер модели, высоту тона, продолжительность и т.д.

Звуковым файлом называют файл, хранящий звуковую информацию в двоичной форме, как правило, информация в звуковых файлах подвергается сжатию. Можно оценить объем моноаудиофайла с длительностью звучания 1 с. при среднем качестве звука (16 бит, 22 кГц). Для этого 16 бит необходимо умножить на 22 000, что дает результат 43 Кбайта.

Литература

- Вычислительные системы, сети и телекоммуникации. Пятибратов и др. - ФИС, 2000.
- Сергеев Н.П., Вашкевич Н.П. Основы вычислительной техники. –М.:ВШ, 1988.
- Гук М. Процессоры фирмы Intel от 8086 до PENTIUM II. -Санкт-Петербург: Питер-Пресс, 1998.
- Гук М. Аппаратные средства IBM PC. –Санкт –Петербург: Питер, 2000.
- Пухальский Г.И., Новосельцева Т.Я. Цифровые устройства: Учебное пособие для вузов. _СПб.: Политехника, 1996. -866с.
- Самофалов К.Г., Корнейчук И.В., Тарасенко В.П. Цифровые электронные вычислительные машины. –Киев: ВИЦА ШКОЛА, 1983.
- Гук М., Юров В. PENTIUM 4 Athlon и Duron. –Санкт –Петербург: Питер, 2001.
- Томпсон Р.Б., Томпсон Б.Ф. Железо ПК: Энциклопедия. –СПб.: Питер, 2003.
- В.Л.Бройдо, О.П.Ильина. Архитектура ЭВМ и систем. Питер, 2006, 703с.
- Таненбаум Э. Архитектура компьютера. СПб, Питер, 2005, 685с.
- Мюллер М. Аппаратные средства ПК. М. Радио, 2004.
- Утепбергенов, И.Т. Архитектура компьютерных систем: Учебное пособие. / Алматы: Экономика, 2010.- 265 с.