

3. Кодирование данных в ЭВМ

Автор: Александр
25.08.2014 16:43

В ЭВМ применяется двоичная система счисления, т.е. все числа в компьютере представляются с помощью нулей и единиц, поэтому компьютер может обрабатывать только информацию, представленную в цифровой форме.

Для преобразования числовой, текстовой, графической, звуковой информации в цифровую необходимо применить кодирование.

Кодирование – это преобразование данных одного типа через данные другого типа. В ЭВМ применяется система двоичного кодирования, основанная на представлении данных последовательностью двух знаков: 1 и 0, которые называются двоичными цифрами (binary digit – сокращенно bit).

Целые числа кодируются двоичным кодом довольно просто (путем деления числа на два). Для кодирования нечисловой информации используется следующий алгоритм: все возможные значения кодируемой информации нумеруются и эти номера кодируются с помощью двоичного кода.

Кодирование чисел

Есть два основных формата представления чисел в памяти компьютера. Один из них используется для кодирования целых чисел, второй (так называемое представление числа в формате с плавающей точкой) используется для задания некоторого подмножества действительных чисел.

Кодирование целых чисел производится через их представление в двоичной системе счисления: именно в этом виде они и помещаются в ячейке. Один бит отводится при

3. Кодирование данных в ЭВМ

Автор: Александр
25.08.2014 16:43

этом для представления знака числа (нулем кодируется знак "плюс", единицей – "минус").

Для кодирования действительных чисел существует специальный формат чисел с плавающей запятой. Число при этом представляется в виде: $M \cdot q^p$, где M – мантисса, p – порядок числа, q – основание системы счисления. Если при этом мантисса M удовлетворяет условию $|M| < q$, то число N называют нормализованным.

Кодирование координат

Закодировать можно не только числа, но и другую информацию, например, о том, где находится некоторый объект. Величины, определяющие положение объекта в пространстве, называются **координатами**. В любой **системе координат** есть начало отсчёта, единица измерения, масштаб, направление отсчёта, или оси координат. Примеры систем координат – декартовы координаты, полярная система координат, шахматы, географические координаты.

Кодирование текста

Для представления текстовой информации используется таблица нумерации символов или таблица кодировки символов, в которой каждому символу соответствует целое число (порядковый номер). Восемь двоичных разрядов могут закодировать 256 различных символов.

Существующий стандарт ASCII (сокращение от American Standard Code for Information Interchange – американский стандартный код для обмена информацией; 8 – разрядная система кодирования) содержит две таблицы кодирования – базовую и расширенную. Первая таблица содержит 128 основных символов, в ней размещены коды символов английского алфавита, а во второй таблице кодирования содержатся 128 расширенных символов.

Так как в этот стандарт не входят символы национальных алфавитов других стран, то в

3. Кодирование данных в ЭВМ

Автор: Александр
25.08.2014 16:43

каждой стране 128 кодов расширенных символов заменяются символами национального алфавита. В настоящее время существует множество таблиц кодировки символов, в которых 128 кодов расширенных символов заменены символами национального алфавита.

Так, например, кодировка символов русского языка Widows – 1251 используется для компьютеров, работающих под ОС Windows. Другая кодировка для русского языка – это КОИ – 8, которая также широко используется в компьютерных сетях и российском секторе Интернет.

В настоящее время существует универсальная система UNICODE, основанная на 16 – разрядном кодировании символов. Эта 16 – разрядная система обеспечивает универсальные коды для 65536 различных символов, т.е. в этой таблице могут разместиться символы языков большинства стран мира.

Кодирование графической информации

В видеопамяти находится двоичная информация об изображении, выводимом на экран. Почти все создаваемые, обрабатываемые или просматриваемые с помощью компьютера изображения можно разделить на две большие группы – **растровую** и **векторную** графику.

Растровые изображения представляют собой однослойную сетку точек, называемых пикселями (pixel, от англ. picture element). Код пикселя содержит информации о его цвете.

Для описания черно-белых изображений используются оттенки серого цвета, то есть при кодировании учитывается только яркость. Она описывается одним числом, поэтому для кодирования одного пикселя требуется от 1 до 8 бит: чёрный цвет – 0, белый цвет – $N = 2^k - 1$, где k – число разрядов, которые отводятся для кодирования цвета. Например, при длине ячейки в 8 бит это $256 - 1 = 255$. Человеческий глаз в состоянии различить от 100 до 200 оттенков серого цвета, поэтому восьми разрядов для этого вполне хватает.

3. Кодирование данных в ЭВМ

Автор: Александр
25.08.2014 16:43

Цветные изображения воспринимаются нами как сумма трёх основных цветов – красного, зелёного и синего. Например, сиреневый = красный + синий; жёлтый = красный + зелёный; оранжевый = красный + зелёный, но в другой пропорции. Поэтому достаточно закодировать цвет тремя числами – яркостью его красной, зелёной и синей составляющих. Этот способ кодирования называется **RGB** (Red – Green – Blue). Его используют в устройствах, способных излучать свет (мониторы). При рисовании на бумаге действуют другие правила, так как краски сами по себе не испускают свет, а только поглощают некоторые цвета спектра. Если смешать красную и зелёную краски, то получится коричневый, а не жёлтый цвет. Поэтому при печати цветных изображений используют метод **CMY** (Cyan – Magenta – Yellow) – голубой, сиреневый, жёлтый цвета. При таком кодировании красный = сиреневый + жёлтый; зелёный = голубой + жёлтый.

В противоположность растровой графике **векторное** изображение многослойно. Каждый элемент такого изображения – линия, прямоугольник, окружность или фрагмент текста – располагается в своем собственном слое, пиксели которого устанавливаются независимо от других слоев. Каждый элемент векторного изображения является объектом, который описывается с помощью специального языка (математических уравнения линий, дуг, окружностей и т.д.) Сложные объекты (ломаные линии, различные геометрические фигуры) представляются в виде совокупности элементарных графических объектов.

Объекты векторного изображения, в отличие от растровой графики, могут изменять свои размеры без потери качества (при увеличении растрового изображения увеличивается зернистость).

Кодирование звука

Как всякий звук, музыка является не чем иным, как звуковыми колебаниями, зарегистрировав которые достаточно точно, можно этот звук безошибочно воспроизвести. Нужно только непрерывный сигнал, которым является звук, преобразовать в последовательность нулей и единиц. С помощью микрофона звук можно превратить в электрические колебания и измерить их амплитуду через равные промежутки времени (несколько десятков тысяч раз в секунду). Каждое измерение записывается в двоичном коде. Этот процесс называется **дискретизацией**. Устройство для выполнения дискретизации называется аналогово-цифровым преобразователем (АЦП). Воспроизведение такого звука ведётся при помощи цифро-аналогового преобразователя (ЦАП). Полученный ступенчатый сигнал сглаживается и преобразуется

3. Кодирование данных в ЭВМ

Автор: Александр
25.08.2014 16:43

в звук при помощи усилителя и динамика. На качество воспроизведения влияют **частота дискретизации и разрешение** (размер ячейки, отведённой под запись значения амплитуды). Например, при записи музыки на компакт-диски используются 16-разрядные значения и частота дискретизации 44 032 Гц.

Описанный способ кодирования звуковой информации достаточно универсален, он позволяет представить любой звук и преобразовывать его самыми разными способами. Но бывают случаи, когда выгодней действовать по-иному.

Издавна используется достаточно компактный способ представления музыки – нотная запись. В ней с помощью специальных символов указывается высота и длительность, общий темп исполнения и как сыграть. Фактически, такую запись можно считать алгоритмом для музыканта, записанным на особом формальном языке. В 1983 г. ведущие производители компьютеров и музыкальных синтезаторов разработали стандарт, определивший такую систему кодов. Он получил название MIDI (Musical Instrument Digital Interface). При таком кодировании запись компактна, легко меняется инструмент исполнителя, тональность звучания, одна и та же запись воспроизводится как на синтезаторе, так и на компьютере.

Конечно, такая система кодирования позволяет записать далеко не всякий звук, она годится только для инструментальной музыки. Но есть у нее и преимущества: чрезвычайно компактная запись, естественность для музыканта (практически любой MIDI-редактор позволяет работать с музыкой в виде обычных нот), легкость замены инструментов, изменения темпа и тональности мелодии.

Есть и другие форматы записи музыки. Среди них – формат MP3, позволяющий с очень большим качеством и степенью сжатия кодировать музыку, при этом вместо 18 – 20 музыкальных композиций на стандартном компакт-диске (CDROM) помещается около 200. Одна песня занимает примерно 3,5 Mb, что позволяет пользователям сети Интернет легко обмениваться музыкальными композициями.