

ФОРМЫ
ПРЕДСТАВЛЕНИЯ
ЧИСЕЛ:
ПРЕДСТАВЛЕНИЕ
СИМВОЛЬНОЙ
ИНФОРМАЦИИ И
ДЕСЯТИЧНЫХ ЧИСЕЛ

Воробьев
Константин
ВТИПО-102(с)

ПРЕДСТАВЛЕНИЯ ЧИСЕЛ В ЭВМ

- Любая информация (чисел, команды, алфавитно-цифровые записи и.) представляется в ЭВМ в виде двоичных кодов (двоичных слов) фиксированной или переменной длины. Отдельные элементы двоичного кода, имеющие значение 0 или 1, называют разрядами или битами. В ЭВМ слова часто разбивают на части, называемые слогами или байтами. В современных ЭВМ широко используется байт, содержащий 8 бит (разрядов).
- Двоичный разряд представляется в ЭВМ некоторым техническим устройством, например триггером, двух различных состояниям которого приписывают значения 0 и 1. Набор соответствующего количества таких устройств служит для представления много разрядного двоичного числа (слова)
- В ЭВМ применяют две формы представления чисел; с фиксированной точкой и с плавающей точкой. Эти формы называют также соответственно естественной и полупологарифмической.
- При представлении чисел с фиксированной точкой положение точки фиксируется в определенном месте относительно разрядов числа. Обычно подразумевается, что точка находится или перед старшим цифровым разрядом, или после младшего. В первом случае могут быть представлены только числа, которые по модулю меньше 1, во втором – только целые числа.

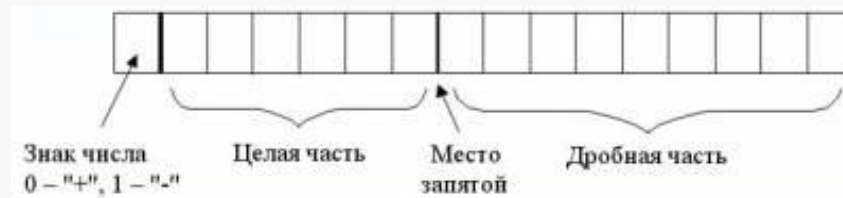
ПРЕДСТАВЛЕНИЯ ДЕСЯТИЧНЫХ ЧИСЕЛ В ЭВМ

- Машинным изображением числа называют его представление в разрядной сетке ЭВМ. В вычислительных машинах применяются две формы представления чисел:
- Естественная форма или форма с фиксированной запятой (точкой);
- Нормальная форма или форма с плавающей запятой (точкой);
- Всякое десятичное число, прежде чем оно попадает в память компьютера, преобразуется по схеме:

$$X_{10} \rightarrow X_2 = M_1 \times [10_2]^r$$

После этого осуществляется ещё одна важная процедура:

- мантисса с её знаком заменяется кодом мантиссы с её знаком;
- порядок числа с его знаком заменяется кодом порядка с его знаком.
- В форме с фиксированной запятой в разрядной сетке выделяется строго определенное число разрядов для целой и для дробной частей числа. Левый (старший) разряд хранит признак знака (0 – "+", 1 – "-") и для записи числа не используется



ПРЕДСТАВЛЕНИЯ ДЕСЯТИЧНЫХ ЧИСЕЛ В ЭВМ

■ Алгоритм представления числа с плавающей запятой:

- перевести число из r -ичной системы счисления в двоичную;
- представить двоичное число в нормализованной экспоненциальной форме;
- рассчитать смещённый порядок числа;
- разместить знак, порядок и мантиссу в соответствующие разряды сетки.

При представлении информации в виде десятичных многоразрядных чисел каждая десятичная цифра заменяется двоично-десятичным кодом. Для ускорения обмена информацией, экономии памяти и удобства операций над десятичными числами предусматриваются специальные форматы их представления: зонный (распакованный) и упакованный. Зонный формат используется в операциях ввода-операций. Для этого в ЭВМ имеются специальные команды упаковки и распаковки десятичных чисел.

- Прямой код
- Представление числа в привычной форме "знак"- "величина", при которой старший разряд ячейки отводится под знак, а остальные - под запись числа в двоичной системе, называется *прямым кодом* двоичного числа. Например, прямой код двоичных чисел 1001 и -1001 для 8-разрядной ячейки равен 00001001 и 10001001 соответственно.
- Положительные числа в ЭВМ всегда представляются с помощью прямого кода. Прямой код числа полностью совпадает с записью самого числа в ячейке машины. Вообще, положительные числа в прямом, обратном и дополнительном кодах изображаются одинаково — двоичными кодами с цифрой 0 в знаковом разряде.

- Например:



КОДИРОВАНИЕ ДЕСЯТИЧНЫХ ЧИСЕЛ И АЛФАВИТНО-ЦИФРОВОЙ ИНФОРМАЦИИ.

- Современные ЭВМ обрабатывают не только числовую, но и текстовую, другими словами алфавитно-цифровую информацию, содержащую цифры, буквы, знаки препинания, математические и другие символы. Именно такой характер имеют экономическая, планово-производственная, учетная информация, а также тексты программ на алгоритмических языках и другая информация. Характер этой информации такой, что ее представления требуются слова переменной длины.
- Возможность ввода, обработки и вывода алфавитно-цифровой информации важна и для решения чисто математических задач, так как позволяет оформить результаты вычислений в удобной форме в виде таблиц с нужными заголовками и пояснениями.
- Совокупность всех символов, используемых в вычислительной системе, представляет собой ее алфавит. Символу соответствует машинная единица информации слог. Так называют группу двоичных разрядов, служащую для представления символа в машине (двоичный код символа). Применяются различные варианты кодирования символов, использующие коды разной длины.
- При выборе способа кодирования необходимо учитывать объем алфавита символов и требования, связанные с облегчением автоматической обработки данных.
- Деловая информация в среднем содержит почти вдвое больше цифр, чем букв. Поэтому наряду с общей системой кодирования алфавитно-цифровых символов (десятичные цифры, буквы и другие знаки) в машинах сохраняют также отдельную систему кодирования для данных, состоящих только из десятичных цифр.
- Необходимо, чтобы память машины эффективно использовала при размещении в ней как алфавитно-цифровой, так и десятичной информации.
- Наибольшее распространение получило представление алфавитно-цифровой информации с помощью 8-разрядных слогов (рис.1.3), называемых байтами. С помощью байта можно кодировать 256 различных символов.
- Для представления алфавитно-цифровых символов в памяти машины и на носителях информации в ЭВМ и в некоторых других ЭВМ используется двоичный код для обработки информации (ДКОИ), а в микропроцессорах К1810, персональных компьютерах и в микроЭВМ –расширенный за счет букв русского алфавита код ASCII.

ПРЕДСТАВЛЕНИЕ СИМВОЛЬНОЙ ИНФОРМАЦИИ В ЭВМ

- В отличие от обычной словесной формы, принятой в письменном виде, символьная информация хранится и обрабатывается в памяти ЭВМ в форме цифрового кода. Например, можно обозначить каждую букву числами, соответствующими ее порядковому номеру в алфавите: А - 01, Б - 02, В - 03, Г - 04, ... , Э - 30, Ю - 31, Я - 32. Точно так же можно договориться обозначать точку числом 33, запятую - 34 и т.д. Так как в устройствах автоматической обработки информации используются двоичные коды, то обозначения букв надо перевести в двоичную систему. Тогда буквы будут обозначаться следующим образом: А - 000001, Б - 000010, В - 000011, Г - 000100, ... , Э - 011110, Ю - 011111, Я - 100000. При таком кодировании любое слово можно представить в виде последовательности кодовых групп, составленных из 0 и 1. Например, слово ЭВМ выглядит так: 011110000011001110.
- При преобразовании символов (знаков) в цифровой код между множествами символов и кодов должно иметь место взаимнооднозначное соответствие, т.е. разным символам должны быть назначены разные цифровые коды, и наоборот. Это условие является единственным необходимым требованием при построении схемы преобразования символов в числа. Однако существует ряд практических соглашений, принимаемых при построении схемы преобразования исходя из соображений наглядности, эффективности, стандартизации. Например, какое бы число ни назначили коду для знака 0 (не следует путать с числом 0), знаку 1 удобно назначить число, на единицу большее, чем код 0, и т.д. до знака 9. Аналогичная ситуация возникает и при кодировке букв алфавита: код для Б на единицу больше кода для А, а код для В на единицу больше кода для Б и т. д. Таким образом, из соображений наглядности и легкости запоминания целесообразно множества символов, упорядоченных по какому-либо признаку (например, лексико-графическому), кодировать также с помощью упорядоченной последовательности чисел.

ПРЕДСТАВЛЕНИЕ СИМВОЛЬНОЙ ИНФОРМАЦИИ В ЭВМ

- При назначении кодов знакам надо также учитывать соглашения, касающиеся стандартизации кодировки. Можно назначить знаковые коды по своему выбору, но тогда возникнут трудности, связанные с необходимостью обмена информацией с другими организациями, использующими кодировку, отличную от нашей. В настоящее время существует несколько широко распространенных схем кодирования. Например, код BCD (Binary-Coded Decimal) — двоично-десятичный код используется для представления чисел, при котором каждая десятичная цифра записывается своим четырехбитовым двоичным эквивалентом. Этот код может оказаться полезным, когда нужно преобразовать строку числовых знаков, например, строку из числовых знаков «2537» в число 2537, над которым затем будут производиться арифметические действия. Расширением этого кода является EBCDIC (Extended Binary-Coded Decimal Interchange Code) -расширенный двоично-десятичный код обмена информацией, который преобразует как числовые, так и буквенные строки.
- В ЭВМ типа PDP (или CM) применяется код ASCII (AmericanJ Standard Code for Information Interchange) - американский стандартный код обмена информацией. Этот код генерируется некоторыми внешними устройствами (принтером, АЦПУ) и используется для обмена данными между ними и оперативной памятью ЭВМ. Например когда нажимаем на терминале клавишу G, то в результате этого действия код ASCII для символа G (1000111) передается в ЭВМ. А если надо этот символ распечатать на АЦПУ, то его код ASCII должен быть послан на печатающее устройство.
- Отечественной версией кода ASCII является код КОИ-7 семибитовый код обмена информацией, который совпадает с ним, за исключением букв русского алфавита.