Формы представления числовой информации

Число 0,028 можно записать так: $28 \cdot 10^{-3}$, или 0,03 (с округлением), или $2{,}8 \cdot 10^{-2}$ и т.д. Разнообразие форм записи одного числа может послужить причиной затруднений для работы компьютера. Во избежание этого нужно или создать специальные алгоритмы распознавания числа, или указывать каждый раз форму его записи. Второй путь проще.

Существует две формы записи чисел: естественная и нормальная. При естественной форме число записывается в естественном натуральном виде, например, 12560 – целое число, 0,00345 – правильная дробь, 4,475 – неправильная дробь. При нормальной форме запись одного числа может принимать разный вид в зависимости от ограничений, накладываемых на ее форму. Например, число 12560 может быть записано так: $12560 = 1{,}256 \cdot 10^4 = 0{,}1256 \cdot 10^5 = 125600 \cdot 10^{-1}$ и т.д.

Автоматное представление числа – представление числа в разрядной сетке компьютера.

Представление целых чисел в компьютере

Естественная форма представления числа в компьютере характеризуется тем, что положение его разрядов в автоматном представлении остается всегда постоянным независимо от величины самого числа. Эта форма записи чисел также называется *представлением чисел с фиксированной точкой*.

Так как числа бывают положительные и отрицательные, то формат (разрядная сетка) автоматного представления числа разбивается на знаковую часть и поле числа. В знаковую часть записывается 0 для положительного числа и 1 для отрицательного.

Если компьютер работает только с целыми числами, то в разрядной сетке один разряд отводится под знак числа, а последующие разряды образуют поле числа.

Диапазон представления чисел в этом случае находится в пределах от $-(2^n-1)$ до (2^n-1), где n — количество разрядов беззнаковой части.

Если в результате операции появится число, по абсолютному значению превышающее максимально допустимое число, то возникает *переполнение разрядной сетки* компьютера, что нарушает нормальное функционирование программы.

Обычно целые положительные числа представляются в компьютере в прямом коде, а отрицательные — в дополнительном коде.

Прямой код двоичного числа

Прямой код двоичного числа $A = a_1 a_2 \ldots a_к$ — такое автоматное представление числа $[A_п] = a_з a_1 a_2 \ldots a_к$, старший разряд которого является знаковым. Знак числа $a_з = 0$ для положительных чисел, и $a_з = 1$ для отрицательных чисел.

ример 1

Пусть размер разрядной сетки равен 8

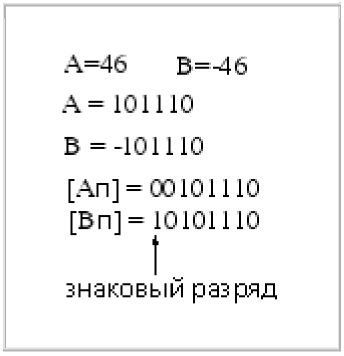

Рис. 1. Прямой код

В прямом коде в разрядную сетку компьютера можно записать следующее максимальное по абсолютному значению число: $A_п^{max} = 0111\ldots1 = 2^{n-1} - 1$,

где n — количество разрядов разрядной сетки компьютера.

Диапазон изменения машинного представления целых чисел для прямого кода лежит в пределах $-(2^{n-1} - 1) < [A_п] < (2^{n-1} - 1)$.

Обратный код двоичного числа

Обратный код отрицательного двоичного числа $A = -a_1 a_2 \ldots a_к$ — такое автоматное представление этого числа $[A_о] = 1\bar{a}_1\bar{a}_2\ldots\bar{a}_к$, для которого $\bar{a}_i = 0$, если $a_i = 1$ и $\bar{a}_i = 1$, если $a_i = 0$. То есть обратный код числа является инверсным изображением самого числа, в котором все разряды исходного числа принимают инверсное (обратное) значение, т.е. все нули заменяются на единицы и наоборот.

Пример 2

$A = -0101110$

$[A_п] = 10101110$

$[A_о] = 11010001$

Дополнительный код двоичного числа

Дополнительный код отрицательного двоичного числа $A = -a_1 a_2 \ldots a_к$ – такое автоматное представление этого числа $[A_д] = 1\bar{a}_1\bar{a}_2\ldots\bar{a}_к$, для которого $\bar{a}_i = 0$, если $a_i = 1$ и $\bar{a}_i = 1$, если $a_i = 0$, за исключением последнего значащего разряда, для которого $\bar{a}_i = 1$, если $a_i = 1$.

Пример 3

$A = -0101110$

$[A_п] = 10101110$

$[A_д] = 11010010$

Дополнительный код является математическим дополнением основанию q ПСС:

$|A| + [A_д] = q^{n+1}$,

где $|A|$ — абсолютное значение числа A, n — количество разрядов в представлении числа.

Пример 4

$A = -0101110$

$[A_д] = 11010010$

$|A| = 00101110$

$|A| + [A_д] = 00101110 + 11010010 = 100000000 = 10^8$

Дополнительный код двоичного числа можно получить из обратного кода, прибавив к младшему разряду единицу.

Пример 5

$[A_д] = [A_о] + 1 = 11010001 + 1 = 11010010$

В компьютере операцию вычитания заменяют операцией сложения в соответствии со следующей формулой:

$A - B = A + (-B);$

$[A_п] + [B_д] = [A_п] + q^{n+1} - |B|;$

$A - B = [A_п] - |B| = [A_п] + [B_д] - q^{n+1}$

Таким образом, для выполнения вычитания двух целых чисел выполняется сложение прямого кода уменьшаемого с дополнительным кодом вычитаемого, при этом единица переноса из старшего разряда игнорируется.

Пример 6

$35 - 11$ заменяется на $35 + 89$ (дополнительный код до 100 - 11) = 124
Ответ : 24 (1 в старшем разряде игнорируется).

Представление чисел в форме с плавающей точкой

В нормальной форме число представляется следующим образом:

$A = m_a \cdot q^{p_a},$

где m_a — мантисса числа A; p_a — порядок числа A.

Такое представление чисел неоднозначно; для определенности обычно вводят некоторые ограничения. Наиболее распространено и удобно для представления в компьютере ограничение вида

$$\frac{1}{q} < m_a < 1, \qquad (1)$$

где q — основание системы счисления.

Нормализованная форма представления чисел — форма представления чисел, для которых справедливо условие (1).

Так как положение разрядов для изображения мантиссы в автоматном представлении числа не постоянно, такую форму представления чисел называют формой *представления числа с плавающей точкой* (ПТ). Формат автоматного представления числа с ПТ должен содержать знаковые части и поля мантиссы и порядка. Выделяются специальные разряды для изображения знака числа (мантиссы) и знака порядка. Кодирование знаков такое же, как для формы с фиксированной точкой. Этот формат является машинно-зависимым.

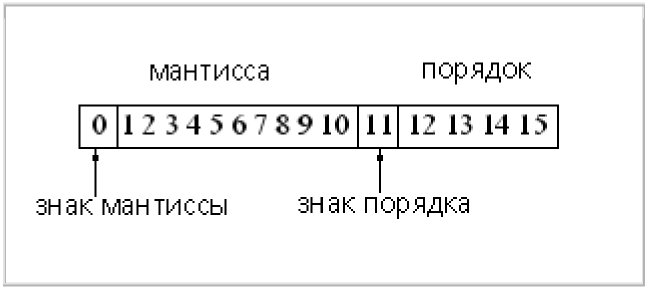

Рис. 2. Пример формата с плавающей точкой

Пример 7

Пусть в данную разрядную сетку надо записать числа A_1 = -10110,1111 и A_2 = 0,0001100010111.

Эти числа надо записать в нормальной форме:

$A_1 = -0.101101111 \cdot 10^{101}$

$A_2 = 0.1100010111 \cdot 10^{-11}$

Мантисса числа A_1(со знаком) — $[m_1]$ = 1 101101111

Порядок числа A_1(со знаком) — [p1]=0 0101

Мантисса числа A_2(со знаком) — [m2] = 0 110001011

Порядок числа A_2(со знаком) — [p2]=1 0011

$[A_1]$=1 101101111 0 0101

$[A_2]$=0 110001011 1 0011

Погрешности представления числовой информации

Представление числовой информации в ЭВМ, как правило, влечет за собой появление погрешностей (ошибок), величина которых зависит от формы машинного представления числа и от длины разрядной сетки ЭВМ.

Абсолютная погрешность представления числа — разность между истинным значением входной величины A и ее значением, полученным из машинного представления A_M:

$\triangle[A] = A - A_M$.

Относительная погрешность представления числа определяется следующим образом:

$$\delta[A] = \left|\frac{\triangle[A]}{A_M}\right|$$

Входные величины независимо от количества значащих цифр могут содержать грубые ошибки, возникающие из-за опечаток, ошибочных отсчетов показаний каких-либо приборов, некорректной постановки задачи или отсутствия более полной и точной информации. Например, часто принимают $\pi = 3.14$. Однако эта величина может быть получена с более высокой точностью. Если принять, что точное значение $\pi = 3.14139265$, то абсолютная погрешность равна $\triangle[\pi] = 0.00159265$.

Часто некоторая величина в одной ПСС имеет конечное значение, а в другой ПСС становится бесконечной величиной. Например, дробь $\frac{1}{10}$ имеет конечное десятичное представление, но, будучи переведена в двоичную ПСС, становится бесконечной дробью.

Следовательно, при переводе чисел из одной ПСС в другую неизбежно возникают погрешности, оценить которые нетрудно, если известны их истинные значения.

Максимальная погрешность перевода десятичной информации в двоичную не будет превышать единицы младшего разряда разрядной сетки ЭВМ. Минимальная погрешность будет равна 0.