Информатика

Лекция 2

Представление вещественных чисел

- Вещественные числа обычно представляются в виде чисел с плавающей запятой. Числа с плавающей запятой — один из возможных способов представления действительных чисел, который является компромиссом между точностью и диапазоном принимаемых значений.
- Число с плавающей запятой состоит из набора отдельных двоичных разрядов, условно разделенных на так называемые **знак** (англ. *sign*), **порядок** (англ. *exponent*) и **мантиссу** (англ. *mantis*).
- В наиболее распространённом формате (стандарт IEEE 754) число с плавающей запятой представляется в виде набора битов, часть из которых кодирует собой мантиссу числа, другая часть — показатель степени, и ещё один бит используется для указания знака числа (0 — если число положительное, 1— если число отрицательное).
- При этом порядок записывается как целое число в коде со сдвигом, а мантисса — в нормализованном виде, своей дробной частью в двоичной системе счисления.

Представление вещественных чисел

0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Знак и порядок								Знак и мантисса																							

- Знак — один бит, указывающий знак всего числа с плавающей точкой. Порядок и мантисса — целые числа, которые вместе со знаком дают представление числа с плавающей запятой в следующем виде:

$$(-1)^S * M * B^P$$

где S — знак, B— основание, P— порядок, а M— мантисса.

Представление вещественных чисел

- Более простым вариантом представления вещественных чисел является вариант с фиксированной точкой, когда целая и вещественная части хранятся отдельно.
- Например, на целую часть отводится всегда X бит и на дробную отводится всегда Y бит.
- Такой способ в архитектурах процессоров не присутствует. Отдаётся предпочтение числам с плавающей запятой, как компромиссу между диапазоном допустимых значений и точностью.

Нормальная форма

- **Нормальной формой** (англ. *normal form*) числа с плавающей запятой называется такая форма, в которой мантисса (без учёта знака) в десятичной системе находится на полуинтервале [0,1).
- Такая форма записи имеет недостаток: некоторые числа записываются неоднозначно, поэтому распространена также другая форма записи — **нормализованная** (англ. *normalized*), в которой мантисса десятичного числа принимает значения от 1 (включительно) до 10 (не включительно), а мантисса двоичного числа принимает значения от 1 (включительно) до 2 (не включительно).

Представление вещественных чисел

- То есть в мантиссе слева от запятой до применения порядка находится ровно один знак.
- В такой форме любое число записывается единственным образом.
- Так как старший двоичный разряд (целая часть) мантиссы вещественного числа в нормализованном виде всегда равен «1», то его можно не записывать, сэкономив таким образом один бит, что и используется в стандарте IEEE 754.
- В позиционных системах счисления с основанием большим, чем 2 (в троичной, четверичной и др.), этого замечательного свойства нет (ведь целая часть там может быть не только единицей).

Вещественные числа

Одинарная точность

1 бит	8 бит	23 бита
Зн	Порядок	Мантисса

Двойная точность

1 бит	11 бит	52 бита
Зн	Порядок	Мантисса

Расширенная точность

1 бит	15 бит	64 бита
Зн	Порядок	Мантисса

Представление вещественных чисел

Тип	Диапазон значений	Размер(байт)
float	3.4E–38 … 3.4E+38	4
double	1.7E–308 … 1.7E+308	8
long double	3.4E–4932 … 3.4E+4932	10

Число 0

- В нормализованной форме числа с плавающей точкой невозможно представить ноль.
- Поэтому для его представления зарезервированы специальные значения мантиссы и порядка — число считается нулём, если все его биты, кроме знакового, равны нулю.
- При этом в зависимости от значения бита знака ноль может быть как положительным, так и отрицательным.

Арифметика нуля со знаком

- $\frac{-0}{|x|} = -0$ (если $x \neq 0$)
- $(-0) \cdot (-0) = +0$
- $|x| \cdot (-0) = -0$
- $x + (\pm 0) = x$
- $(-0) + (-0) = -0$
- $(+0) + (+0) = +0$
- $\frac{-0}{-\infty} = +0$
- $\frac{|x|}{-0} = -\infty$ (если $x \neq 0$)

Неопределенность (*NaN*)

- **NaN** — это аббревиатура от фразы "*not a number*". NaN является результатом арифметических операций, если во время их выполнения произошла ошибка (примеры см. ниже).
- В IEEE 754 NaN представлен как число, в котором все двоичные разряды порядка — единицы, а мантисса не нулевая.
- Любая операция с NaN возвращает NaN. При желании в мантиссу можно записывать информацию, которую программа сможет интерпретировать. Стандартом это не оговорено и мантисса чаще всего игнорируется.

Получение NaN

- $\infty + (-\infty) = NaN$
- $0 \times \infty = NaN$
- $\frac{\pm 0}{\pm 0} = \frac{\pm \infty}{\pm \infty} = NaN$
- $\sqrt{x} = NaN$, где $x < 0$

Бесконечности

- Как и нули со знаком, бесконечности позволяют получить хотя бы близкий к правильному результат вычисления в случае переполнения.
- Согласно стандарту IEEE 754 число с плавающей запятой считается равным бесконечности, если все двоичные разряды его порядка — единицы, а мантисса равна нулю.
- Знак бесконечности определяется знаковым битом числа.

Умножение и деление

- Для того, чтобы умножить два вещественных числа в нормализованной форме необходимо перемножить их мантиссы, сложить порядки, округлить и нормализовать полученное число.
- Соответственно, чтобы произвести деление нужно разделить мантиссу делимого на мантиссу делителя и вычесть из порядка делимого порядок делителя. Затем точно так же округлить мантиссу результата и привести его к нормализованной форме.

Сложение и вычитание

- Идея метода сложения и вычитания чисел с плавающей точкой заключается в приведении их к одному порядку.
- Сначала выбирается оптимальный порядок, затем мантиссы обоих чисел представляются в соответствии с новым порядком, затем над ними производится сложение/вычитание, мантисса результата округляется и, если нужно, результат приводится к нормализированной форме.