

Сложение в формате с плавающей точкой

Михаил Шихов
m.m.shihov@gmail.com

Лекция по дисциплине «информатика»
(12 апреля 2016 г.)

Содержание

1 Порядок

- Примеры представления
- Правила сложения
- Примеры сложения

2 Характеристика

- Примеры представления
- Правила сложения
- Примеры сложения

Формат с плавающей точкой

$$X = m_X \cdot 2^{p_X},$$

где m_X — нормализованная мантисса числа X , p_X — порядок числа X , подобранный так, чтобы m_X была нормализованной.

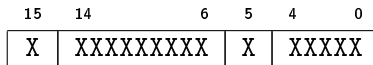
Правила нормализации $X \neq 0$

Мантисса m_X получается из двоичного представления X переносом точки в такую позицию, чтобы целая часть была равна нулю, а в старшем разряде дробной части была единица:

$$(0. \underbrace{1xxx \cdots xxx}_\text{разряды мантиссы})_2$$

порядок p_X определяет на сколько разрядов нужно передвинуть запятую в мантиссе, чтобы получить исходное число.

Формат для примеров



- Разряды нормализованной мантииссы в прямом коде хранятся в разрядах [15 : 6].
- Порядок в прямом коде хранится в разрядах [5 : 0].

Целое, без потерь

$$-9 = (-1001)_2 = (-.1001) \cdot 2^4$$

15	14	6	5	4	0
1	100100000	0	00100		

$$(-.100100000)_2 \cdot 2^4 = -9$$

Целое, с потерями

$$999 = (1111100111)_2 = (.1111100111) \cdot 2^{10}$$

15	14	6	5	4	0
0	111110011	0	01010		

$$(.1111100111)_2 \cdot 2^{10} = \boxed{998}$$

Дробное, без потерь

$$0.171875 = (0.001011)_2 = (.1011) \cdot 2^{-2}$$

15	14	6	5	4	0
0	101100000	1	00010		

$$(.101100000)_2 \cdot 2^{-2} = 0,171875$$

Дробное, с потерями

$$-0.2 = (-0.[0011])_2 = (-.[1100]) \cdot 2^{-2}$$

15	14	6	5	4	0
1	110011001	1	00010		

$$(-.110011001)_2 \cdot 2^{-2} = \boxed{-0.19970703125}$$

Ноль

15	14	6	5	4	0
0	0000000000	0	00000		

Правила сложения

$$X + Y = m_X \cdot 2^{p_X} + m_Y \cdot 2^{p_Y}$$

- 1 Порядки чисел выравниваются до большего, мантисса числа с меньшим порядком сдвигается вправо на модуль разности порядков.

$$\begin{cases} (p_X - p_Y) \geq 0, & m'_Y \leftarrow (m_Y \gg |p_X - p_Y|), m'_X \leftarrow m_X, \\ (p_X - p_Y) < 0, & m'_X \leftarrow (m_X \gg |p_X - p_Y|), m'_Y \leftarrow m_Y. \end{cases}$$

- 2 Получившиеся мантиссы складываются $m_R = m'_X + m'_Y$. При этом порядок результата: $p_R = \max(p_X, p_Y)$.
- 3 Выполняется нормализация результата, если он получился не нормализованным.

13+29

15	14	6	5	4	0	+	15	14	6	5	4	0
0	110100000	0	00100				0	111010000	0	00101		

m	p	прим.
0,110100000	0,00100	$X = 13$, ПК
0,111010000	0,00101	$Y = 29$, ПК
	$ \begin{array}{r} 00,00100 \\ + 11,11011 \\ \hline 11,11111 \end{array} $	$p_X - p_Y < 0$, в МДК, денормализуется m_X
0,011010000	0,00101	X' , денормализованное
$ \begin{array}{r} 00,011010000 \\ + 00,111010000 \\ \hline 01,010100000 \end{array} $		$m_R = m'_X + m'_Y$, в МДК, ПРС!
$ \begin{array}{r} \underline{01},010100000 \\ 0,101010000 \end{array} $	$ \begin{array}{r} 0,00101 \\ 0,00110 \end{array} $	Нормализовать! $m_R \leftarrow m_R \gg 1$; $p_R \leftarrow p_R + 1$ Рез-т!

15	14	6	5	4	0
0	101010000	0	00110		

-17+14

15	14	6	5	4	0	15	14	6	5	4	0
1	100010000	0	00101	+		0	111000000	0	00100		

m	p	прим.
1,100010000 0,111000000	0,00101 0,00100	$X = -17$, ПК $Y = 14$, ПК
	+ 00,00101 11,11100 00,00001	$p_X - p_Y \geq 0$, МДК, денормализуется m_Y
0,011100000	0,00101	Y' , денормализованное
+ 11,011110000 00,011100000 11,111010000		$m_R = m'_X + m'_Y$, МДК
11,111010000	0,00101	Получить модуль мантиссы для представления в ПК!
00,000110000	0,00101	Нормализовать модуль! $m_R \leftarrow m_R \ll 1$; $p_R \leftarrow p_R - 1$
00,001100000	0,00100	Нормализовать модуль! $m_R \leftarrow m_R \ll 1$; $p_R \leftarrow p_R - 1$
00,011000000	0,00011	Нормализовать модуль! $m_R \leftarrow m_R \ll 1$; $p_R \leftarrow p_R - 1$
1,110000000	0,00010	Рез-т!

15	14	6	5	4	0
1	110000000	0	00010		

-2+-2

15	14	6	5	4	0	+	15	14	6	5	4	0
1	100000000	0	00010				1	100000000	0	00010		

m	p	прим.
1,100000000	0,00010	$X = -2$, ПК
1,100000000	0,00010	$Y = -2$, ПК
	+ 00,00010 11,11110 00,00000	$p_X - p_Y = 0$, порядки одинаковы
+ 11,100000000 11,100000000 11,000000000		$m_R = m'_X + m'_Y$, МДК
11,000000000	0,00010	Получить модуль мантиссы для представления в ПК!
01,000000000	0,00010	Нормализовать модуль! $m_R \leftarrow m_R \gg 1$; $p_R \leftarrow p_R + 1$
1,100000000	0,00011	Рез-т!

15	14	6	5	4	0
1	100000000	0	00011		

ПРС (переполнение разрядной сетки)

$$\begin{array}{|c|c|c|c|} \hline 15 & 14 & 6 & 5 & 4 & 0 \\ \hline 0 & 111100000 & 0 & 11111 \\ \hline \end{array} + \begin{array}{|c|c|c|c|} \hline 15 & 14 & 6 & 5 & 4 & 0 \\ \hline 0 & 110000000 & 0 & 11101 \\ \hline \end{array}$$

m	p	прим.
0,111100000	0,11111	X , ПК
0,110000000	0,11101	Y , ПК
	$ \begin{array}{r} + 00,11111 \\ 11,00011 \\ \hline 00,00010 \end{array} $	$p_X - p_Y \geq 0$, МДК, денормализуется m_Y
0,001100000	0,11111	Y' , денормализованное
$ \begin{array}{r} + 00,111100000 \\ 00,001100000 \\ \hline 01,001000000 \end{array} $		$m_R = m'_X + m'_Y$, МДК, ПРС мантиссы!
<u>01</u> ,001000000	0,11111	Нормализовать! $m_R \leftarrow m_R \gg 1$; $p_R \leftarrow p_R + 1$
<u>00</u> ,100100000	?,?????	ПРС порядка — настоящий ПРС в формате с ПЗ!

Генерация ошибки вычислений!

ПМР (потеря младших разрядов)

15	14	6	5	4	0	+	15	14	6	5	4	0
0	100001000	1	11110				1	111100000	1	11111		

m	p	прим.
0,100001000	1,11110	X , ПК
1,111100000	1,11111	Y , ПК
	$ \begin{array}{r} + 11,00010 \\ 00,11111 \\ \hline 00,00001 \end{array} $	$p_X - p_Y \geq 0$, МДК, денормализуется m_Y
1,011110000	1,11110	Y' , денормализованное
$ \begin{array}{r} + 00,100001000 \\ 11,100010000 \\ \hline 00,000011000 \end{array} $		$m_R = m'_X + m'_Y$, МДК
<u>00,000011000</u>	1,11110	Нормализовать! $m_R \leftarrow m_R \ll 1$; $p_R \leftarrow p_R - 1$
<u>00,000110000</u>	1,11111	Нормализовать! $m_R \leftarrow m_R \ll 1$; $p_R \leftarrow p_R - 1$
<u>00,001100000</u>	?,?????	p_R за пределом представления отрицательных чисел в ПК!

15	14	6	5	4	0
0	0000000000	0	00000		

$$X + Y = X, Y \neq 0?$$

15	14	6	5	4	0	+	15	14	6	5	4	0
0	111000000	0	11000				0	110000000	0	01110		

m	p	прим.
0,111000000	0,11000	$X = 7 \cdot 2^{21}$, ПК
0,110000000	0,01110	$Y = 3 \cdot 2^{12}$, ПК
	$ \begin{array}{r} 00,11000 \\ + 11,10010 \\ \hline 00,01010 \end{array} $	$ p_X - p_Y \geq 9$, МДК, m_Y денормализуется в 0
0,000000000	0,11000	$Y' = 0?$

15	14	6	5	4	0
0	111000000	0	11000		

Характеристика

$$X = m_X \cdot 2^{p_X}.$$

Диапазон представления порядка p_X в n -разрядной сетке будет¹:

$$p_X \in [-2^{n-1}, +(2^{n-1} - 1)]$$

Характеристика получается из порядка прибавлением фиксированной поправки Δ , такой, что левая граница представления обращается в ноль. Таким образом,

характеристика c_X — всегда положительное число.

$$c_X = p_X + \Delta, \tag{1}$$

где^a $\Delta = +2^{n-1}$, а $c_X \in [0, 2^n - 1]$.

^aОпять же только в случае использования дополнительного кода

¹Если использовать дополнительный код

Свойства n -разрядной характеристики

- Характеристика — положительное число.
- Разность характеристик равна разности порядков.
- Если в процессе нормализации (или денормализации) порядок увеличивается (или уменьшается), то то же самое происходит и с характеристикой.
- Если для работы с характеристиками использовать ДК или МДК, о ПРС при нормализации легко судить по знаковому разряду: он не должен быть 1.
- Если используется поправка $\Delta = 2^{n-1}$, то характеристика получается из дополнительного кода порядка инверсией знакового разряда.

Формат для примеров

15	14	6	5	0
X	XXXXXXXXXX	XXXXXX		

- Разряды нормализованной мантииссы в прямом коде хранятся в разрядах [15 : 6].
- Характеристика хранится в разрядах [5 : 0].
- $\Delta = 2^5 = 32 = (100000)_2$

Целое, без потерь

$\Delta = 2^5 = 32$

$$-9 = (-1001)_2 = (-.1001) \cdot 2^4$$

15	14	6	5	0
1	100100000	100100		

$$(-.100100000)_2 \cdot 2^{36-\Delta} = -9$$

Целое, с потерями

$\Delta = 2^5 = 32$

$$999 = (1111100111)_2 = (.1111100111) \cdot 2^{10}$$

15	14	6	5	0
0	111110011	101010		

$$(.111110011)_2 \cdot 2^{42-\Delta} = \boxed{998}$$

Дробное, без потерь

$\Delta = 2^5 = 32$

$$0.171875 = (0.001011)_2 = (.1011) \cdot 2^{-2}$$

15	14	6	5	0
0	101100000	011110		

$$(.101100000)_2 \cdot 2^{30-\Delta} = 0,171875$$

Дробное, с потерями

$\Delta = 2^5 = 32$

$$-0.2 = (-0.[0011])_2 = (-.[1100]) \cdot 2^{-2}$$

15	14	6	5	0
1	110011001	011110		

$$(-.110011001)_2 \cdot 2^{30-\Delta} = \boxed{-0.19970703125}$$

Ноль

15	14	6	5	0
0	0000000000	000000		

Правила

Исходя из формулы (1), почти не меняются

$$X + Y = m_X \cdot 2^{c_X - \Delta} + m_Y \cdot 2^{c_Y - \Delta}$$

- 1 Характеристики чисел выравниваются до большей, мантисса числа с меньшей характеристикой сдвигается вправо на модуль разности характеристик.

$$\begin{cases} (c_X - c_Y) \geq 0, & m'_Y \leftarrow (m_Y \gg |c_X - c_Y|), m'_X \leftarrow m_X, \\ (c_X - c_Y) < 0, & m'_X \leftarrow (m_X \gg |c_X - c_Y|), m'_Y \leftarrow m_Y. \end{cases}$$

- 2 Получившиеся мантиссы складываются $m_R = m'_X + m'_Y$. При этом характеристика результата: $c_R = \max(c_X, c_Y)$.
- 3 Выполняется нормализация результата, если он получился не нормализованным.

13+57

$$\begin{array}{|c|c|c|} \hline 15 & 14 & 6 \quad 5 \quad 0 \\ \hline 0 & 110100000 & 100100 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 15 & 14 & 6 \quad 5 \quad 0 \\ \hline 0 & 111001000 & 100110 \\ \hline \end{array}$$

m	c	прим.
0,110100000	100100	$X = 13$
0,111001000	100110	$Y = 57$
	$ \begin{array}{r} + 0,100100 \\ 1,011010 \\ \hline 1,111110 \end{array} $	$c_X - c_Y < 0$, ДК, денормализуется m_X
0,001101000	0,100110	X' , денормализованное
$ \begin{array}{r} + 00,001101000 \\ 00,111001000 \\ \hline 01,000110000 \end{array} $		$m_R = m'_X + m'_Y$, в МДК, ПРС!
<u>01</u> ,000110000	0,100111	Нормализовать! $m_R \leftarrow m_R \gg 1$; $c_R \leftarrow c_R + 1$
0,100011000	100111	Рез-т!

$$\begin{array}{|c|c|c|} \hline 15 & 14 & 6 \quad 5 \quad 0 \\ \hline 0 & 100011000 & 100111 \\ \hline \end{array}$$

-17+14

15	14	6	5	0	+	15	14	6	5	0
1	100010000	100101				0	111000000	100100		

<i>m</i>	<i>c</i>	прим.
1,100010000 0,111000000	100101 100100	$X = -17$ $Y = 14$
	+ 0,100101 1,011100 0,000001	$c_X - c_Y \geq 0$, ДК, денормализуется m_Y
0,011100000	0,100101	Y' , денормализованное
+ 11,011110000 00,011100000 11,111010000		$m_R = m'_X + m'_Y$, МДК
11,111010000	0,100101	Получить модуль мантиссы для представления в ПК!
00,000110000	0,100101	Нормализовать модуль! $m_R \leftarrow m_R \ll 1$; $c_R \leftarrow c_R - 1$
00,001100000	0,100100	Нормализовать модуль! $m_R \leftarrow m_R \ll 1$; $c_R \leftarrow c_R - 1$
00,011000000	0,100011	Нормализовать модуль! $m_R \leftarrow m_R \ll 1$; $c_R \leftarrow c_R - 1$
1,110000000	100010	Рез-т!

15	14	6	5	0
1	110000000	100010		

ПРС (переполнение разрядной сетки)

$$\begin{array}{|c|c|c|} \hline 15 & 14 & 6 \quad 5 \quad 0 \\ \hline 0 & 111100000 & 111111 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 15 & 14 & 6 \quad 5 \quad 0 \\ \hline 0 & 110000000 & 111101 \\ \hline \end{array}$$

m	c	прим.
0,111100000	111111	X
0,110000000	111101	Y
	$ \begin{array}{r} 0,111111 \\ + 1,000011 \\ \hline 0,000010 \end{array} $	$c_X - c_Y \geq 0$, ДК, денормализуется m_Y
0,001100000	0,111111	Y' , денормализованное
$ \begin{array}{r} 00,111100000 \\ + 00,001100000 \\ \hline 01,001000000 \end{array} $		$m_R = m'_X + m'_Y$, МДК, ПРС мантиссы!
<u>01</u> ,001000000	0,111111	Нормализовать! $m_R \leftarrow m_R \gg 1$; $c_R \leftarrow c_R + 1$
<u>00</u> ,100100000	<u>1</u> ,000000	$c_R < 0$, выход за правую границу представления — ПРС!

Генерация ошибки вычислений!

ПМР (потеря младших разрядов)

$$\begin{array}{|c|c|c|} \hline 15 & 14 & 6 & 5 & 0 \\ \hline 0 & 100001000 & 000010 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 15 & 14 & 6 & 5 & 0 \\ \hline 1 & 111100000 & 000001 \\ \hline \end{array}$$

m	c	прим.
0,100001000	000010	X
1,111100000	000001	Y
	$ \begin{array}{r} + 0,000010 \\ 1,111111 \\ \hline 0,000001 \end{array} $	$c_X - c_Y \geq 0$, МДК, денормализуется m_Y
1,011110000	0,000010	Y' , денормализованное
$ \begin{array}{r} + 00,100001000 \\ 11,100010000 \\ \hline 00,000011000 \end{array} $		$m_R = m'_X + m'_Y$, МДК
<u>00</u> ,000011000	0,000010	Нормализовать! $m_R \leftarrow m_R \ll 1$; $c_R \leftarrow c_R - 1$
<u>00</u> ,000110000	0,000001	Нормализовать! $m_R \leftarrow m_R \ll 1$; $c_R \leftarrow c_R - 1$
<u>00</u> ,001100000	0,000000	Нормализовать! $m_R \leftarrow m_R \ll 1$; $c_R \leftarrow c_R - 1$
<u>00</u> ,011000000	<u>1</u> ,111111	$c_R < 0$, выход за левую границу представления — ПМР!

$$\begin{array}{|c|c|c|} \hline 15 & 14 & 6 & 5 & 0 \\ \hline 0 & 000000000 & 000000 \\ \hline \end{array}$$

1)

Придумать правила определения ПРС/ПМР при работе

- с порядками в прямом коде;
- с характеристиками.

2)

Разработать собственный 10-разрядный формат² и сложить в нем числа:

- ❶ 9 и -11 ;
- ❷ 10 и 7;
- ❸ 0.625 и 0.75;
- ❹ -0.625 и 0.375.

²Преподавателю: обязательно проследить, чтобы были использованы и порядки и характеристики

3)

Придумать пример из последовательности трех чисел, сумма которых зависисит от порядка суммирования.

Советы самоучке

Стандарт на формат с плавающей точкой IEEE 754.

Библиография I