

Дискретная математика

Домашнее задание №6

«Сложение чисел с плавающей запятой»

Вариант №59

Выполнил: Нодири Хисравхон (гр. Р3131)

СЛОЖЕНИЕ ЧИСЕЛ С ПЛАВАЮЩЕЙ ЗАПЯТОЙ

Формат Ф1 (с 12 битной мантиссой)

$$A = 62.84 = 0.3ED7..._{16} \approx 0.\underbrace{3ED}_{\text{мантисса}}_{16} * 16^2$$

$$X_A = 2 + 64 = \underbrace{66}_{\text{характеристика}}$$

0	1	0	0	0	1	0	0	0	1	1	1	1	0	1	1	0	1
0	1					7	8										19

$$B = 58.92 = 0.3AE B..._{16} \approx 0.\underbrace{3AF}_{\text{мантисса}}_{16} * 16^2$$

$$X_B = 2 + 64 = \underbrace{66}_{\text{характеристика}}$$

0	1	0	0	0	1	0	0	0	1	1	1	0	1	1	1	1	1
0	1					7	8										19

Сравнение порядков:

$$\begin{array}{r} 1\ 0\ 0\ 0\ 0\ 1\ 0 = X_A \\ - 1\ 0\ 0\ 0\ 0\ 1\ 0 = X_B \\ \hline 0\ 0\ 0\ 0\ 0\ 0\ 0 = (X_A - X_B)_{\text{доп}} \\ X_A - X_B = 0, X_C = X_A = 66 \end{array}$$

Оба операнда положительные:

$$\begin{array}{r} 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \\ 0.0\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 0\ 1 = M_A \\ + 0.0\ 0\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ 1 = M_B \\ \hline 0.0\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0 = M_C \end{array}$$

Результат нормализован.

0	1	0	0	0	0	1	0	0	1	1	1	1	0	0	1	1	0	0
0	1					7	8											19

$$C^* = S_M * M_M * 16^{P_M} = 121.75$$

$$C_T = A + B = 121.7600000000000000051159076975$$

$$\Delta C = C_T - C^* = 0.0100000000000000051159076975$$

$$\delta C = \left| \frac{\Delta C}{C_T} \right| * 100\% \approx 0.008213\%$$

Формат Ф2 (с 11 битной мантиссой)

$$A = 62.84 = 0.1111101101011..._2$$

$$A \approx 0.\underbrace{111110110110}_{\text{мантисса}}_2 * 2^6$$

$$X_A = 6 + 128 = \underbrace{134}_{\text{характеристика}}$$

0	1	0	0	0	0	1	1	0	1	1	0	1	1	0	1	1	0
0	1						8	9									19

$$B = 58.92 = 0.1110101110101..._2$$

$$A \approx 0.\underbrace{111010111011}_{\text{мантисса}}_2 * 2^6$$

$$X_B = 6 + 128 = \underbrace{134}_{\text{характеристика}}$$

0	1	0	0	0	0	1	1	0	1	1	0	1	1	0	1	1	0
0	1						8	9									19

Сравнение порядков:

$$\begin{array}{r} 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0 = X_A \\ - 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0 = X_B \\ \hline 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 = (X_A - X_B)_{\text{доп}} \\ X_A - X_B = 0, X_C = X_A = 134 \end{array}$$

Оба операнда положительные:

$$\begin{array}{r} 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \\ 0.1\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 0 = M_A \\ + 0.1\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 1 = M_B \\ \hline 1.1\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1 = M_C \end{array}$$

Результат денормализован влево.

$$M_C = 1.111001110001$$

$$M_C \rightarrow 1 = 0.1111100111000$$

Характеристику результата нужно увеличить на 1:

$$X'_C = X_C + 1 = 135$$

0	1	0	0	0	0	1	1	1	1	0	0	1	1	0	0	0	0
0	1						8	9									19

$$C^* = S_M * M_M * 2^{P_M} = 121.75$$

$$C_T = A + B = 121.7600000000000000051159076975$$

$$\Delta C = C_T - C^* = 0.0100000000000000051159076975$$

$$\delta C = \left| \frac{\Delta C}{C_T} \right| * 100\% \approx 0.008213\%$$

А положительно, В отрицательно

[illegible]

Результат денормализован вправо.

$$M_C = 0.0000000111110$$

$$M_C \leftarrow 4 = 0.001111100000$$

Характеристику результата нужно уменьшить на 1: $X'_C = X_C - 1 = 65$

[illegible]

$$\begin{aligned} C^* &= S_M * M_M * 16^{P_M} = 3.875 \\ C_T &= A + B = 3.920000000000001705302565824 \\ \Delta C &= C_T - C^* = 0.045000000000001705302565824 \\ \delta C &= \left| \frac{\Delta C}{C_T} \right| * 100\% \approx 1.147959\% \end{aligned}$$

А отрицательно, В положительно

$$\begin{array}{r} 0.001111101101 = M_A \\ - 0.0001110101111 = M_B \\ \hline 0.000000111110 = M_C \end{array}$$

Результат денормалізований вправо.

$$\begin{aligned} M_C &= 0.000000111110 \\ M_C \leftarrow 4 &= 0.001111100000 \end{aligned}$$

Характеристику результата нужно уменьшить на 1: $X'_C = X_C - 1 = 65$

1	1	0	0	0	0	0	1	0	0	1	1	1	1	1	0	0	0	0	
0	1						7	8											19

$$\begin{aligned} C^* &= S_M * M_M * 16^{P_M} = -3.875 \\ C_T &= A + B = -3.920000000000001705302565824 \\ \Delta C &= C_T - C^* = -0.045000000000001705302565824 \\ \delta C &= \left| \frac{\Delta C}{C_T} \right| * 100\% \approx 1.147959\% \end{aligned}$$

[illegible]

Результат денормализован вправо.

$$\begin{aligned} M_C &= 0.000011111011 \\ M_C \leftarrow 4 &= 0.111110110000 \end{aligned}$$

Характеристику результата нужно уменьшить на 4:
 $X'_C = X_C - 4 = 130$

0	1	0	0	0	0	1	0	1	1	1	1	0	1	1	0	0	0	0		
0	1						8	9												19

$$\begin{aligned} C^* &= S_M * M_M * 2^{P_M} = 3.921875 \\ C_T &= A + B = 3.920000000000001705302565824 \\ \Delta C &= C_T - C^* = -0.001874999999998294697434176 \\ \delta C &= \left| \frac{\Delta C}{C_T} \right| * 100\% \approx 0.047832\% \end{aligned}$$

А отрицательно, В положительно

[illegible]

Результат денормализован вправо.

$$\begin{aligned} M_C &= 0.000011111011 \\ M_C \leftarrow 4 &= 0.111110110000 \end{aligned}$$

Характеристику результата нужно уменьшить на 4:
 $X'_C = X_C - 4 = 130$

1	1	0	0	0	0	0	1	0	1	1	1	1	0	1	1	0	0	0	0	
0	1							8	9											19

$$\begin{aligned} C^* &= S_M * M_M * 2^{P_M} = -3.921875 \\ C_T &= A + B = -3.920000000000001705302565824 \\ \Delta C &= C_T - C^* = 0.001874999999998294697434176 \\ \delta C &= \left| \frac{\Delta C}{C_T} \right| * 100\% \approx 0.047832\% \end{aligned}$$

Причины возникновения погрешности:

1. Неточное представление операндов.
2. Потеря значащих разрядов мантиссы одного из операндов при уравнивании порядков.
3. Потеря значащих разрядов мантиссы результата при его нормализации сдвигом мантиссы вправо.

В формате Ф2 результаты точнее, потому что операнды представлены точнее, и сдвиг при нормализации результата производится на любое число бит, не обязательно кратное 4, за счет чего теряется меньше значащих разрядов.