8.3. ЗАДАНИЕ 6
СЛОЖЕНИЕ ЧИСЕЛ С ПЛАВАЮЩЕЙ ЗАПЯТОЙ

1. Заданные числа ***A*** и ***B*** представить в форме с плавающей запятой в разрядных сетках форматов *Ф1* и *Ф2* с укороченной мантиссой (12 двоичных разрядов).
   Примечание: общее число разрядов в формате – 20.
2. Выполнить операцию сложения заданных чисел со следующими комбинациями знаков операндов: “++”, “+–”, “–+” в разрядных сетках форматов *Ф1* и *Ф2*.
3. Результаты представить в форматах операндов, перевести в десятичную систему счисления и проверить их правильность.
4. Определить абсолютную и относительную погрешности результатов и обосновать их причину.
5. Сравнить погрешности результатов аналогичных операций для форматов Ф*1* и Ф*2* и объяснить причины их сходства или различия.
   Варианты задания приведены в табл. 6 Приложения 1.

Пример операции сложения.

***A* = 0,3; *B* = 15,8.**

1. Формат *Ф1* (число разрядов мантиссы *m* =12).

$A = (0{,}3)_{10} = \underbrace{(0{,}4(C))_{16}}_{MA} \cdot 16^0$

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

1 округление к ближайшему

$B = (15{,}8)_{10} = (F{,}(C))_{16} = \underbrace{(0{,}F(C))_{16}}_{MB} \cdot 16^1$

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

1 округление к ближайшему

$$
\begin{array}{lr}
1) \quad X_A = & -\ 1\ 0\ 0\ 0\ 0\ 0\ 0 \\
X_B = & 1\ 0\ 0\ 0\ 0\ 0\ 1 \\
\hline
(X_A\text{-}X_B)_{доп.} = & 1\ 1\ 1\ 1\ 1\ 1\ 1
\end{array}
$$

$(X_A\text{-}X_B) = -1;\ \ X_C = X_B = 1.$

а) Оба операнда положительные (A>0, B>0):

$$
\begin{array}{lr}
2{,}3)\ \ \overset{4}{\rightarrow} M_A = & .\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0 \\
M_B = & +\ .\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 1 \\
\hline
M_C = & 1.\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1
\end{array}
$$

Результат сложения денормализован влево.

4) $\overset{4}{\overrightarrow{M_C}} = .\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1$

Т.к. выполнен сдвиг мантиссы влево, то характеристику результата нужно увеличить на 1 ($X_C = X_C + 1 = 2$).

C

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

$C^* = M_C \cdot 16^{P_C} = (0{,}101)_{16} \cdot 16^2 = (10{,}1)_{16} = 16{,}0625.$

Несмотря не то, что оба операнда за счет округления были представлены с избытком, результат получился представленным с недостатком. Этот факт можно объяснить потерей значащих младших разрядов сначала у первого операнда при выравнивании порядков, а затем и у результата при его нормализации.

В принципе, погрешность полученного результата можно объяснить следующими факторами:

- неточным представлением операндов;
- потерей значащих разрядов мантиссы одного из операндов при уравнивании порядков;
- потерей значащих разрядов мантиссы результата при его нормализации сдвигом мантиссы вправо.

$\Delta C = C_{\text{Т}} - C^* = 16{,}1 - 16{,}0625 = 0{,}0375,$

где ΔC –абсолютная погрешность;

$C_{\text{Т}}$ –точное значение;

C^* - приближенное значение.

$$\delta C = \left|\frac{\Delta C}{C_T}\right| \cdot 100\% = \left|\frac{0{,}0375}{16{,}1}\right| . 100\% = 0{,}22\%,$$

где δC – относительная погрешность.

б) $A<0$, $B>0$.

Сложение мантисс будем проводить их прямым вычитанием. В качестве уменьшаемого используем мантиссу положительного операнда (B);

$$\begin{array}{r l}
2,3)\ M_B = & .\ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1\ \ 0\ \ 1 \\
- \\
\underline{\overset{4}{\overrightarrow{M_A}} =} & \underline{.\ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1\ \ 0\ \ 0} \\
M_C = & .\ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1
\end{array}$$

Результат сложения нормализован.

C

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

$C^* = M_C \cdot 16^{P_C} = (0{,}F81)_{16} \cdot 16^1 = (F{,}81)_{16} = 15{,}5039.$

$\Delta C = C_{\text{Т}} - C^* = 15{,}5 - 15{,}5039 = -\ 0{,}0039,$

$$\delta C = \left|\frac{\Delta C}{C_T}\right| \cdot 100\% = \left|\frac{0{,}0039}{15{,}5}\right| . \ 100\% = 0{,}025\%.$$

2. Формат *Ф2*.

$A = (0{,}3)_{10} = (0{,}4(C))_{16} = (0{,}100110011001100)_2 \cdot 2^{-1}$

0	0	1	1	1	1	1	1	1	0	0	1	1	0	0	1	1	0	0	1
19	18							11	10										0

$B = (15{,}8)_{10} = (F,(C))_{16} = (0{,}111111001100)_2 \cdot 2^4$

0	1	0	0	0	0	1	0	0	1	1	1	1	1	0	0	1	1	0	0
19	18							11	10										0

1) X_A = _0 1 1 1 1 1 1 1

$\underline{X_B = 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0}$

$(X_A\text{-}X_B)_{\text{доп.}} = 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1$

$(X_A\text{-}X_B)_. = -5$; $X_C = X_B = 4$.

а) Оба операнда положительные (A>0, B>0):

2,3) $\overrightarrow{\text{M}_A}^{5}$ = . 0 0 0 0 0 1 0 0 1 1 0 0

\+

M_B = . 1 1 1 1 1 1 0 0 1 1 0 0

M_C = 1 . 0 0 0 0 0 0 0 1 1 0 0 0

Результат сложения денормализован влево.

4) $\overrightarrow{\text{M}_C}^{1}$ = . 1 0 0 0 0 0 0 0 1 1 0 0

Т.к. выполнен сдвиг мантиссы влево, то характеристику результата нужно увеличить на 1 ($X_C = X_C + 1 = 5$).

C

0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0
19	18							11	10										0

$C^* = \text{M}_C \cdot 2^{\text{P}c} = (0{,}10000000111)_2 \cdot 2^5 = (10000{,}000111)_2 = 16{,}109375.$

$\Delta C = C_\text{T} - C^* = 16{,}1 - 16{,}\ 109375 = 0{,}009375,$

$$\delta C = \left|\frac{\Delta C}{C_T}\right| \cdot 100\% = \left|\frac{0{,}009375}{16{,}1}\right| . \ 100\% = 0{,}058\%.$$

б) A<0, B>0.

2,3) M_B = . 1 1 1 1 1 1 0 0 1 1 0 0

$\overrightarrow{\text{M}_A}^{5}$ = . $\overline{0}$ 0 0 0 0 1 0 0 1 1 0 0

M_C = . 1 1 1 1 1 0 0 0 0 0 0 0

C	0	1	0	0	0	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0
	19	18							11	10										0

$C^* = \text{М}_C \cdot 2^{\text{Р}c} = (0{,}11111_2 \cdot 2^4 = (1111{,}1)_2 = 15{,}5.$

$\Delta C = C_\text{Т} - C^* = 15{,}5 - 15{,}5 = 0.$

В формате *Ф2* результаты получились точнее из-за того, что операнды представлены точнее и при нормализации результата сдвиг производился на один двоичный разряд, а не на четыре.