

**Федеральное государственное автономное образовательное учреждение высшего  
образования «Национальный исследовательский университет ИТМО»**

Факультет программной инженерии и компьютерной техники

**Домашнее задание №6**

По дискретной математике

Вариант №85

Выполнил:

Ступин Тимур Русланович

Группа № Р3108

Проверил:

Поляков Владимир Иванович

Санкт-Петербург 2023

## Содержание

Исходный данные .....	3
1.    Формат $\Phi 1$ .....	3
2.    Формат $\Phi 2$ .....	5

## Исходный данные

№	A	B
85	304,5	256,2

### 1. Формат Ф1

$$A = (304,5)_{10} = (130,8)_{16} = (0,1308)_{16} \times 16^3$$

$$X_A = P_A + 64 = 3 + 64 = (67)_{10} = (1000011)_2$$

0	1	0	0	0	0	1	1	0	0	0	1	0	0	1	1	0	0	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

1 округление  
к ближайшему

$$B = (256,2)_{10} = (100,(3))_{16} = (0,100(3))_{16} \times 16^3$$

$$X_B = P_B + 64 = 3 + 64 = (67)_{10} = (1000011)_2$$

0	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

0 округление  
к ближайшему

$$1) \quad X_A = \underline{1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1}$$

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

$$(X_A - X_B)_{\text{пр.}} = \underline{0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0}$$

$$(X_A - X_B) = 0; \quad X_C = X_B = 3.$$

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

$$3) \quad M_A = \begin{array}{r} .0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \\ M_B = +.0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \\ \hline M_C = .0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \end{array}$$

$$M_B = \begin{array}{r} .0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \\ \hline M_C = .0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \end{array}$$

$$M_C = \begin{array}{r} .0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \end{array}$$

Результат сложения нормализован

C

0	1	0	0	0	0	1	1	0	0	1	0	0	0	1	1	0	0	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

$$C^* = M_C \cdot 16^{P_C} = (0,231)_{16} \cdot 16^3 = (231,0)_{16} = 561,0.$$

$$\Delta C = C_T - C^* = 560,7 - 561,0 = -0,3$$

где  $\Delta C$  – абсолютная погрешность;

$C_T$  – точное значение;

$C^*$  – приближенное значение.

$$\delta C = \left| \frac{\Delta C}{C_T} \right| \cdot 100\% = \left| \frac{-0,3}{560,7} \right| \cdot 100\% = 0,054\%,$$

где  $\delta C$  – относительная погрешность.

Погрешность полученного результата обусловлена неточностью представления операндов.

б)  $A < 0, B > 0$ .

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

$$\begin{array}{r} 3) \quad M_B = \quad . \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{1} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \overset{\curvearrowright}{0} \\ M_A = \quad . 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \\ \hline M_C = \quad . 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \end{array}$$

Результата сложения мантисс отрицателен и представлен в дополнительном коде:

$$C_{\text{доп}} = .1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1$$

Переведем в прямой код:  $C_{\text{пр}} = .0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1$

Результат сложения денормализован вправо.

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

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

$$C \quad \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$C^* = M_C \cdot 16^{P_C} = (-0,31)_{16} \cdot 16^2 = (-31,0)_{16} = -49,0.$$

$$\Delta C = C_T - C^* = -48,3 - (-49,0) = 0,7.$$

$$\delta C = \left| \frac{\Delta C}{C_T} \right| \cdot 100\% = \left| \frac{0,7}{-48,3} \right| \cdot 100\% = 1,45\%.$$

Погрешность полученного результата обусловлена неточностью представления операндов.

в)  $A > 0, B < 0$ .

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

$$\begin{array}{r} 3) \quad M_A = \quad . 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \\ M_B = \quad . 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \\ \hline M_C = \quad . 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \end{array}$$

Результат сложения денормализован вправо.

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

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

C	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

$$C^* = M_C \cdot 16^{P_C} = (0,31)_{16} \cdot 16^2 = (31,0)_{16} = 49,0.$$

$$\Delta C = C_T - C^* = 48,3 - 49,0 = -0,7.$$

$$\delta C = \left| \frac{\Delta C}{C_T} \right| \cdot 100\% = \left| \frac{-0,7}{-48,3} \right| \cdot 100\% = 1,45\%.$$

Погрешность полученного результата обусловлена неточностью представления операндов.

## 2. Формат Ф2.

$$A = (304,5)_{10} = (130,8)_{16} = (100110000.1)_2 = (0.1001100001) \times 2^9$$

$$X_A = P_A + 128 = 9 + 128 = (137)_{10} = (10001001)_2$$

0	1	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0	1	0	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

$$B = (256,2)_{10} = (100,(3))_{16} = (0.1000000000001) \times 2^9$$

$$X_B = P_B + 128 = 9 + 128 = (137)_{10} = (10001001)_2$$

0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

$$\begin{array}{r} 1) \quad X_A = \underline{1 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1} \\ \quad \quad X_B = \underline{1 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1} \\ (X_A - X_B)_{\text{пр.}} = 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \\ (X_A - X_B) = 0; \quad X_C = X_B = 9. \end{array}$$

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

$$\begin{array}{r} 3) \quad M_A = \overset{\curvearrowright}{.} 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \\ \quad \quad M_B = \overset{+}{.} 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \\ \quad \quad M_C = 1.0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1 \end{array}$$

Результат сложения денормализован влево.

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

Т.к. выполнен сдвиг мантиисы вправо, то характеристику результата нужно увеличить на 1 ( $X_C = X_C + 1 = 10$ ).

C	0	1	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

$$C^* = M_C \cdot 2^{P_C} = (0,10001100001)_2 \cdot 2^{10} = (1000110000,1)_2 = 560,5.$$

$$\Delta C = C_T - C^* = 560,7 - 560,5 = 0,2.$$

$$\delta C = \left| \frac{\Delta C}{C_T} \right| \cdot 100\% = \left| \frac{0,2}{560,7} \right| \cdot 100\% = 0,036\%.$$

Погрешность полученного результата обусловлена неточностью представления операндов и потерей значащих разрядов при нормализации мантиссы.

б)  $A < 0, B > 0$ .

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

$$\begin{array}{r} 3) \ M_B = \quad .1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1 \\ \underline{M_A = \quad .1\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0} \\ M_C = \quad .1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1 \end{array}$$

Результата сложения мантисс отрицателен и представлен в дополнительном коде:

$$C_{\text{доп}} = .1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1$$

Переведем в прямой код:  $C_{\text{пр}} = .0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1$

Результат сложения денормализован вправо.

$$4) \overset{\leftarrow 3}{M}_C = .1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0$$

Т.к. выполнен сдвиг мантиссы влево, то характеристику результата нужно уменьшить на 3 ( $X_C = X_C - 3 = 6$ ).

$$C \quad \boxed{1 \mid 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0 \mid 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0}$$

$$C^* = M_C \cdot 2^{P_C} = -(0,110000011)_2 \cdot 2^6 = -(110000,011)_2 = -48,375.$$

$$\Delta C = C_T - C^* = -48,3 - (-48,375) = 0,075.$$

$$\delta C = \left| \frac{\Delta C}{C_T} \right| \cdot 100\% = \left| \frac{0,075}{-48,3} \right| \cdot 100\% = 0,16\%.$$

Погрешность полученного результата обусловлена неточностью представления операндов

в)  $A > 0, B < 0$ .

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

$$\begin{array}{r} 3) \ M_A = \quad .1\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0 \\ \underline{M_B = \quad .1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1} \\ M_C = \quad .0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1 \end{array}$$

Результат сложения денормализован вправо.

$$4) \overset{\leftarrow 3}{M}_C = .1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0$$

Т.к. выполнен сдвиг мантиссы влево, то характеристику результата нужно уменьшить на 3 ( $X_C = X_C - 3 = 6$ ).

$C$

1	1	0	0	0	0	1	1	0	1	0	0	0	0	0	1	1	0	0	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

$$C^* = M_C \cdot 2^{P_C} = (0,110000011)_2 \cdot 2^6 = (110000,011)_2 = 48,375.$$

$$\Delta C = C_T - C^* = 48,3 - 48,375 = -0,075.$$

$$\delta C = \left| \frac{\Delta C}{C_T} \right| \cdot 100\% = \left| \frac{-0,075}{48,3} \right| * 100\% = 0,16\%.$$

Погрешность полученного результата обусловлена неточностью представления операндов

В формате  $\Phi 2$  результаты получились точнее из-за того, что операнды представлены точнее и при нормализации результата сдвиг производился на один двоичный разряд, а не на четыре.