

УДК 004.315

О. В. Самощенко, канд. техн. наук, доц.
В. В. Лапко, канд. техн. наук, доц.
Г. Е. Маргієв, аспірантДВНЗ «Донецький національний технічний університет», м. Покровськ, Україна
aleksandr.samoshchenko@gmail.com, margievge@gmail.com

Алгоритм додавання та віднімання чисел з рухомою комою при представленні порядків в системі з додатним нулем

Отримані логічні вирази для визначення знаку результату та способу обчислення модуля результату. Встановлено відповідність ознак порівняння модулів, порядків та мантис початкових даних з рухомою комою при представленні порядків в коді з додатним нулем. Показано, що сукупність ознак порівняння модулів, порядків та мантис чисел з рухомою комою при представленні порядків в системі з додатним нулем визначається відніманням модулів початкових даних як чисел з фіксованою комою. Обґрунтовані алгоритми коректування порядків операндів в коді з додатним нулем при вирівнюванні порядків початкових чисел. Запропоновані алгоритми перетворення порядку результату та способи фіксації переповнення поля порядків за нормалізації мантиси результату. Розроблені алгоритми та математичні моделі операцій додавання та віднімання чисел з рухомою комою обґрунтовані методами програмного моделювання операцій в середовищі ACTIVE-VHDL

Ключові слова: модуль, порядок, мантиса, код з додатним нулем, денормалізація, нормалізація, переповнення, полярність переповнення.

DOI: 10.31474/1996-1588-2019-1-28-79-88

Вступ

Операції додавання та віднімання чисел з рухомою комою становлять значну частину машинних операцій в обчислювальних системах. Аналіз модулів чисел з рухомою комою суттєво спрощується за умови представлення порядків зміщеним кодом з додатним нулем [1-5], двійкове зображення якого пропорційне значенню порядку. Через це сукупність порядку та мантиси утворює двійковий позиційний код з фіксованою комою, значення якого пропорційне модулю числа з рухомою комою. Це дозволяє операції порівняння модулів, порядків та мантис в процедурах додавання та віднімання чисел з рухомою комою зводити до простої арифметичної операції віднімання беззнакових чисел з фіксованою комою. Через обмеженість розрядної сітки операційних суматорів порядку та мантис при обчисленні результату операцій додавання та віднімання чисел з рухомою комою додатково необхідно фіксувати наявність та полярність переповнення. Крім того, при перетвореннях порядків та мантис операції необхідно виконувати таким чином, щоб результат на основних виходах суматорів формувався безпосередньо в коректному форматі. В даній роботі пропонується математичний апарат та алгоритм синтезу способів реалізації процедур додавання та віднімання чисел з рухомою комою при представленні порядків в коді з додатним нулем,

що забезпечують мінімізацію апаратних та часових витрат. Адекватність розробленого математичного опису операційного пристрою підтверджена моделюванням операцій додавання та віднімання чисел з додатним нулем засобами мови VHDL [6].

Загальна методика виконання операцій додавання та віднімання чисел з рухомою комою

У загальному випадку операції додавання та віднімання чисел з рухомою комою зводяться до розв'язання задачі:

$$C_{ПК} = A_{ПК} D B_{ПК} \quad (1)$$

де $A_{ПК}$, $B_{ПК}$, $C_{ПК}$ – прямі коди змінних з рухомою комою задачі: $A_{ПК} = [NA]A$, $B_{ПК} = [NB]B$, $C_{ПК} = [NC]C$; NA , NB , NC – відповідно знакові розряди змінних $A_{ПК}$, $B_{ПК}$, $C_{ПК}$, нуль та одиниця в котрих визначають додатне або від'ємне значення змінних задачі; $A = 2^{PA} \cdot MA$, $B = 2^{PB} \cdot MB$, $C = 2^{PC} \cdot MC$ – модулі змінних $A_{ПК}$, $B_{ПК}$, $C_{ПК}$ в системі з рухомою комою; $(|P_A|, |P_B|, |P_C|) \in [0; 2^m - 1]$ – порядки модулів A , B та C ; m – формат (або довжина) поля порядків, $(MA, MB, MC) \in [2^{-1}; 1 - 2^{-n}]$ – нормалізовані коди мантиси модулів A , B та C ; n – формат (або довжина) поля мантиси модулів A , B та C ; $D \in [0; 1]$ – код операції, нуль та одиниця котрого визначають відповідно

операції додавання та віднімання початкових даних $A_{ПК}$ та $B_{ПК}$.

Знак та модуль результату операції (1), згідно табл.1, визначаються сукупністю співвідношень:

$$C = \begin{cases} D_M = NA \oplus NB \oplus D; \\ A + B, \text{ при } D_M = 0; \\ |A - B|, \text{ при } D_M = 1; \end{cases}$$

$$\omega = \begin{cases} 1, \text{ при } A \geq B; \\ 0, \text{ при } A < B; \end{cases} \quad (2)$$

$$NC = \begin{cases} NA, \text{ при } \omega = 1 \\ NB \oplus D, \text{ при } \omega = 0. \end{cases}$$

де D_M – код операції процедури обчислення модуля результату; ω – ознака порівняння модулів початкових даних.

Таблиця 1 - Алгоритм формування модулю та знаку результату

NA	0	0	0	0	1	1	1	1
D	0	0	1	1	0	0	1	1
NB	0	1	0	1	0	1	0	1
D_M	0	1	1	0	1	0	0	1
NC	A=B	0	*	*	0	*	1	*
	A>B	0	0	0	0	1	1	1
	A<B	0	1	1	0	0	1	0

Ознака порівняння модулів ω визначається відношенням мантис та модулів початкових даних.

За рівних мантис чисел ($'MA/'MB = 1$) ознака порівняння модулів операндів, згідно (1) та (2) визначається системою нерівностей:

$$\omega = \begin{cases} 1 \text{ при } 2^{P_A - P_B} = 1: \frac{A}{B} = 1 \\ 1 \text{ при } 2^{P_A - P_B} > 1: \frac{A}{B} > 1 \\ 0 \text{ при } 2^{P_A - P_B} < 1: \frac{A}{B} < 1 \end{cases} \quad (3)$$

при $'MA = 'MB$.

Згідно (3), отримаємо:

$$\omega = \begin{cases} 1 \text{ при } P_A - P_B \geq 0 \\ 0 \text{ при } P_A - P_B < 0 \end{cases}, \quad (4)$$

при $'MA = 'MB$.

Якщо $'MA \neq 'MB$ ознака порівняння модулів операндів ω , згідно (1) та (2), описується системою нерівностей:

$$\omega = \begin{cases} 1 \text{ при } 2^{P_A - P_B} \frac{'MA}{'MB} > 1: \frac{A}{B} > 1 \\ 0 \text{ при } 2^{P_A - P_B} \frac{'MA}{'MB} < 1: \frac{A}{B} < 1 \end{cases}, \quad (5)$$

при $'MA \neq 'MB$.

У випадку, коли $'MA > 'MB$, згідно (1), відношення нормалізованих мантис змінюється в межах:

$$1 > \frac{'MA}{'MB} > 2^{-1}$$

Через це, згідно (5) отримаємо:

$$\omega = \begin{cases} 1 \text{ при } 2^{P_A - P_B} > \left(\frac{'MA}{'MB} > 1\right): \frac{A}{B} > 1 \\ 0 \text{ при } 2^{P_A - P_B} < \left(\frac{'MA}{'MB} < 1\right): \frac{A}{B} < 1 \end{cases},$$

при $'MA > 'MB$.

Звідки випливає рівносілля система нерівностей:

$$\omega = \begin{cases} 1 \text{ при } 2^{P_A - P_B} \frac{'MB}{'MA} > 1: \frac{A}{B} > 1 \\ 0 \text{ при } 2^{P_A - P_B} \frac{'MB}{'MA} < 1: \frac{A}{B} < 1 \end{cases},$$

при $'MA > 'MB$.

Через це:

$$\omega = \begin{cases} 1 \text{ при } P_A - P_B > 0: \frac{A}{B} > 1 \\ 0 \text{ при } P_A - P_B < 0: \frac{A}{B} < 1 \end{cases}, \quad (6)$$

при $'MA > 'MB$.

При $'MA > 'MB$ відношення мантис визначається інтервалом:

$$1 < \frac{'MB}{'MA} < 2 \quad (7)$$

Якщо $'MA < 'MB$, згідно (2) та (5), отримаємо

$$\omega = \begin{cases} 1 \text{ при } 2^{P_A - P_B} > \left(\frac{'MB}{'MA} < 2\right): A > B \\ 0 \text{ при } 2^{P_A - P_B} < \left(\frac{'MB}{'MA} < 2\right): A < B \end{cases},$$

при $'MA < 'MB$.

Звідси випливає рівносілля система співвідношень:

$$\omega = \begin{cases} 1 \text{ при } 2^{P_A - P_B} \geq 2: \frac{A}{B} > 1 \\ 0 \text{ при } 2^{P_A - P_B} < 2: \frac{A}{B} < 1 \end{cases},$$

при $'MA < 'MB$.

Через це при $'MA < 'MB$, згідно (7), остаточно отримаємо:

$$\overline{P_B^{DH}(m)} \cdot 2^{m-1} + \dots + \overline{P_B^{DH}(2)} \cdot 2^1 + \overline{P_B^{DH}(1)} \cdot 2^0 = \overline{P_B^{DH}(m)} \cdot \overline{P_B^{DH}(2)} \cdot \overline{P_B^{DH}(1)} = \overline{P_B^{DH}(m, 1)} - \text{порозрядна інверсія коду } P_B^{DH}(m, 1).$$

Для реалізації співвідношення (15) на входах двійкового суматора, згідно (16), необхідно скомпонувати двійкові коди [5]:

$$\begin{aligned} a_p(m, 1) &= P_A^{DH}(m, 1); \\ b_p(m, 1) &= \overline{P_B^{DH}(m, 1)}; \quad e_p = 1, \end{aligned} \quad (17)$$

де $a_p(m, 1)$, $b_p(m, 1)$ – основні входи операційного суматора (17) довжиною m біт; e_p – вхідний перенос суматора порядків (17) з ваговим коефіцієнтом 2^0 .

Повна сума на виводах суматора порядків (17) утворює поліном

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) \cdot 2^{m-1} + S_p(m-1, 1) &= \\ &= P_A^{DH}(m, 1) + S_p(m, 1) + 1 = \\ &= \frac{V}{2} + P_A(m, 1) + KE(m, 1) - \\ &\quad - \left(\frac{V}{2} + P_B(m, 1) \right) + 1 = \\ &= P_A(m, 1) - P_B(m, 1) + V - 1 + 1 = \\ &= (P_A(m, 1) - P_B(m, 1)) + V, \end{aligned} \quad (18)$$

де E_p – вихідний перенос суматора порядків; $S_p(m, 1)$ – код суми на основних виводах суматора порядків (17).

При $P_A(m, 1) = P_B(m, 1)$, згідно (18), отримаємо

$$\begin{aligned} E_p + S_p(m, 1) &= \\ V + (P_A(m, 1) - P_B(m, 1)) &= \\ V + 0 = 2^m + 0, \end{aligned} \quad (19)$$

Отже, за рівності порядків початкових даних, згідно (19), на штатних виводах суматора (17) формується комбінація

$$E_p Z_{SP} = 11, \quad (20)$$

$$\text{де } Z_{SP} = \begin{cases} 1 & \text{при } S_p(m, 1) = 0, \\ 0 & \text{при } S_p(m, 1) \neq 0. \end{cases}$$

Таким чином, ознака рівності порядків початкових даних, згідно (20), описується логічним виразом:

$$Z_p = E_p \cdot Z_{SP}, \quad (21)$$

$$\text{де } Z_p = \begin{cases} 1 & \text{при } P_A(m, 1) = P_B(m, 1), \\ 0 & \text{при } P_A(m, 1) \neq P_B(m, 1). \end{cases}$$

У загальному випадку область визначення m -розрядних двійкових кодів $P_A^{DH}(m, 1) = V/2 + P_A(m, 1)$ та $P_B^{DH}(m, 1) = V/2 + P_B(m, 1)$ описується нерівністю

$$0 \leq (P_A^{DH}(m, 1), P_B^{DH}(m, 1)) \leq V - 1.$$

Отже, порядки початкових даних обмежені інтервалом числової осі:

$$-\frac{V}{2} \leq (P_A(m, 1), P_B(m, 1)) \leq \frac{V}{2} - 1. \quad (22)$$

Звідси випливає, що за $P_A(m, 1) \neq P_B(m, 1)$ різниця порядків (14) обмежена інтервалом:

$$-(V - 1) \leq (P_A(m, 1), P_B(m, 1)) \leq V - 1. \quad (23)$$

В області додатних значень різниці (14) при $0 < (P_A(m, 1) - P_B(m, 1)) \in [1; V - 1]$ повна сума на виводах суматора порядків (17), згідно (18), змінюється в межах:

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) &= \\ &= \begin{cases} V + P_A(m, 1) - P_B(m, 1) = V + 1 = 2^m + 1 & \text{при } P_A(m, 1) - P_B(m, 1) = 1; \\ V + P_A(m, 1) - P_B(m, 1) = V + V - 1 = & \\ 2^m + 2^m - 1 = 2^m + 2^{m-1} + 2^{m-1} - 1 & \\ \text{при } P_A(m, 1) - P_B(m, 1) = V - 1. \end{cases} \end{aligned}$$

За можливих додатних значущих значеннях різниці порядків (14) на виводах суматора (17) формується стан

$$E_p = 1. \quad (24)$$

В області від'ємних значень різниці (14) при $0 > (P_A(m, 1) - P_B(m, 1)) \in [-(V - 1); -1]$.

Повна сума на виводах суматора (17), згідно (18), змінюється в діапазоні

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) &= \\ &= \begin{cases} V + P_A(m, 1) - P_B(m, 1) = V - V + 1 = 1 & \text{при } P_A(m, 1) - P_B(m, 1) = -(V - 1); \\ V + P_A(m, 1) - P_B(m, 1) = V - 1 = 2^m - 1 = & \\ = 2^{m-1} + 2^{m-1} - 1 & \\ \text{при } P_A(m, 1) - P_B(m, 1) = -1. \end{cases} \end{aligned}$$

Таким чином, за від'ємних значень різниці порядків на виводах суматорів (17) формується стан

$$E_p = 0. \quad (25)$$

В загальному випадку при $MA(n, 1) \geq MB(n, 1)$, згідно (21) та (24)–(25), нерівності (4) та (6) рівносильні співвідношенням:

$$\omega = E_P;$$

$$E_P = \begin{cases} 1 & \text{при } P_A(m, 1) > P_B(m, 1) \\ 0 & \text{при } P_A(m, 1) < P_B(m, 1) \end{cases} \quad (26)$$

при $Z_P = 0$

При 'MA(n,1)<'MB(n,1), згідно (8), ознака порівняння модулів початкових даних ω визначається полярністю різниці:

$$P_A(m, 1) - P_B(m, 1) - 1. \quad (27)$$

Для реалізації виразу за допомогою двійкового суматора представимо (27) у вигляді

$$\begin{aligned} (P_A(m, 1) - P_B(m, 1) - 1)_{mV} &= \\ &= (V + P_A(m, 1) - P_B(m, 1) - 1)_{mV} = \\ &= \left(\frac{V}{2} + P_A(m, 1) - \left(\frac{V}{2} + P_B(m, 1) \right) + V - 1 \right)_{mV} = \\ &= (P_A^{DH}(m, 1) + (V - 1) - P_B^{DH}(m, 1))_{mV} = \\ &= (P_A^{DH}(m, 1) + KE(m, 1) - P_B^{DH}(m, 1))_{mV} = \\ &= (P_A^{DH}(m, 1) + \overline{P_B^{DH}(m, 1)})_{mV}. \end{aligned} \quad (28)$$

Для формування коду (28) на входах двійкового суматора необхідно скомпонувати набір двійкових кодів [5]:

$$\begin{aligned} a_p(m, 1) &= P_A^{DH}(m, 1); \\ b_p(m, 1) &= \overline{P_B^{DH}(m, 1)}; e_p = 0, \end{aligned} \quad (29)$$

Повна сума на штатних виводах суматора (29) утворює поліном

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) &= \\ &= P_A^{DH}(m, 1) + \overline{P_B^{DH}(m, 1)} = \\ &= \frac{V}{2} + P_A(m, 1) + V - 1 - \frac{V}{2} - P_B(m, 1) = \\ &= V + P_A(m, 1) - P_B(m, 1) - 1. \end{aligned} \quad (30)$$

Область визначення різниці (27), згідно (23), обмежена інтервалом:

$$-V \leq P_A(m, 1) - P_B(m, 1) - 1 \leq V - 2. \quad (31)$$

В області додатних значень різниці (27), згідно (31), при $P_A(m, 1) - P_B(m, 1) - 1 \in [0; V-2]$ повна сума на виводах суматора (29), згідно (30), змінюється в межах:

$$E_p \cdot 2^m + S_p(m, 1) =$$

$$= \begin{cases} V + P_A(m, 1) - P_B(m, 1) - 1 = V + 0 = 2^m \\ \text{при } P_A(m, 1) - P_B(m, 1) - 1 = 0; \\ V + P_A(m, 1) - P_B(m, 1) - 1 = V + V - 2 = \\ = 2^m + 2^m - 2 \\ \text{при } P_A(m, 1) - P_B(m, 1) - 1 = V - 2. \end{cases}$$

При додатних значеннях різниці (27) на виводах суматора (29) формується стан

$$E_P = 1. \quad (32)$$

В області від'ємних значень різниці (27) при

$$0 > P_A(m, 1) - P_B(m, 1) - 1 \in [-V; -1],$$

згідно (31), повна сума на виводах суматора (29), згідно (30), змінюється в діапазоні:

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) &= \\ &= \begin{cases} V + P_A(m, 1) - P_B(m, 1) - 1 = V - V = 0 \\ \text{при } P_A(m, 1) - P_B(m, 1) - 1 = -V; \\ V + P_A(m, 1) - P_B(m, 1) - 1 = V - 1 = 2^m - 1 = \\ = 2^{m-1} + 2^{m-1} - 1 \\ \text{при } P_A(m, 1) - P_B(m, 1) - 1 = -1. \end{cases} \end{aligned}$$

За від'ємних значень різниці (27) на виході суматора (29) формується стан

$$E_P = 0. \quad (33)$$

У випадку збігу порядків початкових даних на основних виводах суматора (29), згідно (30), формується двійковий код

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) &= \\ V + P_A(m, 1) - P_B(m, 1) - 1 &= V - 1 = 2^m - 1 = \\ &= KE(m, 1) < 2^m \end{aligned}$$

Отже, за рівності порядків початкових даних на виводах суматора (29), згідно (30), формується комбінація

$$E_P + K_{SP} = 01, \quad (34)$$

$$\text{де } K_{SP} = \begin{cases} 1 & \text{при } S_p(m, 1) = KE(m, 1); \\ 0 & \text{при } S_p(m, 1) \neq KE(m, 1). \end{cases}$$

Таким чином, ознака рівності порядків Z_P , згідно (34), визначається логічним виразом:

$$Z_P = \overline{E_P} \cdot Z_{SP}, \quad (35)$$

$$\text{де } Z_P = \begin{cases} 1 & \text{при } P_A(m, 1) = P_B(m, 1); \\ 0 & \text{при } P_A(m, 1) \neq P_B(m, 1). \end{cases}$$

Згідно (32)–(33) та (35), при 'MA(n,1)<'MB(n,1) система нерівностей (8) рівносильна співвідношенням:

$$\omega = E_p;$$

$$E_p = \begin{cases} 1 & \text{при } P_A(m, 1) > P_B(m, 1) \\ 0 & \text{при } P_A(m, 1) < P_B(m, 1) \end{cases} \quad (36)$$

$$\text{при } Z_p = 0$$

У загальному випадку для формування ознак порівняння модулів та порядків початкових даних при $'MA(n, 1) \geq 'MB(n, 1)$ та $'MA(n, 1) < 'MB(n, 1)$, згідно (13), (17) та (29), в суматорі порядків необхідно виконати сукупність операцій:

$$\begin{aligned} a_p(m, 1) &= P_A^{DH}(m, 1); \\ b_p(m, 1) &= P_B^{DH}(m, 1); \\ e_p &= \begin{cases} 1 & \text{при } 'MA(n, 1) \geq 'MB(n, 1) : E_M = 1; \\ 0 & \text{при } 'MA(n, 1) < 'MB(n, 1) : E_M = 0. \end{cases} \end{aligned}$$

Звідки випливає, що для реалізації однокрокового алгоритму формування ознак порівняння модулів, порядків та мантис початкових даних достатньо використовувати каскадну схему додавання у складі суматора порядків та суматора мантис:

$$\begin{aligned} E_m^{AB} \cdot 2^0 + S_m^{AB}(n, 1) &= \\ = 'MA(n, 1) + 'MB(n, 1) + 1 \cdot 2^{-n}; \\ e_p^{AB} &= E_m^{AB}; \\ E_p^{AB} \cdot 2^m + S_p^{AB}(n, 1) &= \\ = P_A^{DH}(m, 1) + P_B^{DH}(m, 1) + e_p^{AB}. \end{aligned} \quad (37)$$

У каскадному суматорі (37) через наявність залежності $e_p^{AB} = E_m^{AB}$ одночасно формуються коректні ознаки порівняння мантис, порядків та модулів початкових даних як при $'MA(n, 1) \geq 'MB(n, 1)$, так і при $'MA(n, 1) < 'MB(n, 1)$:

$$\begin{aligned} \omega &= E_p^{AB}; \quad Z_p = E_p^{AB} Z_{SP}^{AB} \vee E_p^{AB} K_{SP}^{AB}; \\ E_m^{AB} &= \begin{cases} 1 & \text{при } 'MA(n, 1) \geq 'MB(n, 1) \\ 0 & \text{при } 'MA(n, 1) < 'MB(n, 1) \end{cases}; \\ E_p^{AB} &= \begin{cases} 1 & \text{при } P_A(m, 1) > P_B(m, 1) \\ 0 & \text{при } P_A(m, 1) < P_B(m, 1) \end{cases} \end{aligned} \quad (38)$$

$$\begin{aligned} \text{де } Z_{SP}^{AB} &= \begin{cases} 1 & \text{при } S_p^{AB}(m, 1) = 0, \\ 0 & \text{при } S_p^{AB}(m, 1) \neq 0; \end{cases} \\ K_{SP}^{AB} &= \begin{cases} 1 & \text{при } S_p^{AB}(m, 1) = KE(m, 1), \\ 0 & \text{при } S_p^{AB}(m, 1) \neq KE(m, 1). \end{cases} \end{aligned}$$

Вирівнювання порядків початкових даних

Для визначення мантиси результату у випадку нерівності порядків операндів, згідно правилам додавання та віднімання чисел з рухомою комою [1-2], порядки початкових даних спочатку треба вирівняти збільшенням меншого коду та денормалізацією відповідної мантиси числа. Якщо наприклад, згідно (38), $P_A(m, 1) > P_B(m, 1)$, процедура вирівнювання порядків зводиться до циклічного виконання сукупності операцій:

$$P_B(m, 1) + 1; \quad (39)$$

$$'MB(n, 1) = R1('MB(n, 1)), \quad (40)$$

де $R1('MB(n, 1))$ – опис процедури зсуву мантиси $'MB(n, 1)$ на один розряд праворуч.

Оскільки порядки початкових даних представлені в коді з додатним нулем, двійковий код (39) також представимо в коді з додатним нулем:

$$\begin{aligned} (P_B(m, 1) + 1)^{DH} &= \frac{V}{2} + P_B(m, 1) + 1 = \\ &= P_B^{DH}(m, 1) + 1. \end{aligned} \quad (41)$$

Для реалізації операції (41) на входах суматора порядків треба скомпонувати набір двійкових кодів:

$$\begin{aligned} a_p(m, 1) &= 0; \quad b_p(m, 1) = P_B^{DH}(m, 1); \\ e_p &= 1, \end{aligned} \quad (42)$$

Тоді на основних виходах суматора (42),

$$S_p(m, 1) = P_B^{DH}(m, 1) = (P_B(m, 1) + 1)^{DH}, \quad (43)$$

де $P_B(m, 1) + 1 \in [-V/2; V/2 - 1]$, оскільки інкремент меншого порядку не може перевищувати області визначення нормованого значення (22) більшого порядку.

Цикл (40)–(41) завершується при рівності перетвореного порядку $P_B^{DH}(m, 1)$ і більшого порядку $P_A^{DH}(m, 1)$. В суматорі порядків порівняння кодів $P_B^{DH}(m, 1)$ та $P_A^{DH}(m, 1)$ виконується згідно (17). Процедура вирівнювання порядків (40)–(41) завершується при фіксації в суматорі (17) стану $Z_p=1$.

Якщо, згідно (38), $P_A(m, 1) < P_B(m, 1)$, процедура вирівнювання порядків здійснюється аналогічним чином.

Формування та нормалізація мантиси результату

При однакових порядках та після вирівнювання порядків початкових даних поліном повної суми мантиси результату на виходах суматора мантис, згідно (2) та (38), визначається сукупністю співвідношень:

$$\begin{aligned} E_m^C \cdot 2^0 + S_m^C(n) + S_m^C(n-1, 1) &= \\ \begin{cases} 'MA(n, 1) + 'MB(n, 1) & \text{при } D_M=0; \\ ('MA(n, 1) - 'MB(n, 1))_{m1} & \text{при } D_M E_M^{AB}=1; \\ ('MB(n, 1) - 'MA(n, 1))_{m1} & \text{при } D_M \overline{E_M^{AB}}=1, \end{cases} \end{aligned} \quad (44)$$

де E_m^C , $S_m^C(n, 1)$ – відповідно ознаки порушення нормалізації мантиси результату ліворуч та праворуч.

Звідси, згідно (10), отримаємо:

$$E_m^C \cdot 2^0 + S_m^C(n) + S_m^C(n-1, 1) = \quad (45)$$

$$= \begin{cases} 'MA(n, 1) + 'MB(n, 1) \text{ при } D_M=0; \\ ('MA(n, 1) + 'MB(n, 1) + 1 \cdot 2^{-n})_{m1} \\ \text{при } D_M E_M^{AB}=1; \\ ('MA(n, 1) + 'MB(n, 1) + 1 \cdot 2^{-n})_{m1} \\ \text{при } D_M \overline{E_M^{AB}}=1. \end{cases}$$

Для обчислення абсолютних значень мантиси результату на входах суматора мантис (44) необхідно скомпонувати набори двійкових кодів:

$$\begin{aligned} a_m^c(n, 1) &= 'MA(n, 1) \oplus D_M \cdot E_m^{AB}; \\ b_m^c(n, 1) &= 'MB(n, 1) \oplus D_M \cdot \overline{E_m^{AB}}; \\ e_m^c(n, 1) &= D_M. \end{aligned} \quad (46)$$

При нормалізованому поліномі на виході суматора (44) коди мантис та порядку модуля результату, згідно (2) та (38), формуються у вигляді:

$$\begin{aligned} 'MC(n, 1) &= 'S_M^c(n, 1); \\ P_C^{DH}(m, 1) &= \end{aligned} \quad (47)$$

$$= \begin{cases} P_A^{DH}(m, 1) \text{ при } E_P^{AB} \cdot \overline{Z_P^{AB}} = 1; \\ P_B^{DH}(m, 1) \text{ при } \overline{E_P^{AB}} \cdot \overline{Z_P^{AB}} = 1; \\ P_A^{DH}(m, 1) + P_B^{DH}(m, 1) \text{ при } Z_P^{AB} = 1. \end{cases} \quad (48)$$

При порушенні нормалізації ліворуч у вихідному коді суматора мантис (44) нормалізація результату досягається зсувом поліному та інкрементуванням порядку результату:

$$\begin{aligned} 'MC(n, 1) &= R1(E_P^c S_M^c(n, 1)) = 'E_P^c S_M^c(n, 2) \\ P_C^{DH}(m, 1) &= (P_C^{DH}(m, 1))^{DH} \end{aligned} \quad (49)$$

Для реалізації операції за допомогою двійкового суматора представимо вираз (49) у вигляді:

$$\begin{aligned} P_C^{DH}(m, 1) &= (P_C^{DH}(m, 1) + 1)^{DH} = \\ &= \frac{V}{2} + P_C^{DH}(m, 1) + 1 = P_C^{DH}(m, 1) + 1, \end{aligned} \quad (50)$$

де $P_C^{DH}(m, 1) = V/2 + P_C^{DH}(m, 1)$ – зміщений код поля порядку модулю результату.

Область визначення функції $P_C^{DH}(P_C(m, 1))$, згідно (22), обмежена інтервалом $[-V/2; V/2-1]$. В загальному випадку змінна $P_C(m, 1) + 1$ змінюється в діапазоні $[-(V/2-1); V/2]$. Тому при $P_C(m, 1) + 1 = V/2$ необхідно фіксувати додатне переповнення m -розрядного поля порядків модулю результату.

При корекції порядку за допомогою двійкового суматора вихідний код суматора, згідно (50), утворює поліном:

$$\begin{aligned} E_P \cdot 2^m + S_P(m, 1) &= P_C^{DH}(m, 1) + 1 = \\ &= \frac{V}{2} + P_C(m, 1) + 1. \end{aligned} \quad (51)$$

При коректних значеннях суми $P_C(m, 1) + 1 \in [-V/2; V/2-1]$ повна сума (51) змінюється в межах:

$$E_P \cdot 2^m + S_P(m, 1) = \quad (52)$$

$$= \begin{cases} \frac{V}{2} + P_C(m, 1) + 1 = \frac{V}{2} - \frac{V}{2} = 0 \\ \text{при } P_C(m, 1) + 1 = -\frac{V}{2}; \\ \frac{V}{2} + P_C(m, 1) + 1 = \\ = \frac{V}{2} + \frac{V}{2} - 1 = 2^{m-1} + 2^{m-1} - 1 \\ \text{при } P_C(m, 1) = \frac{V}{2} - 1. \end{cases}$$

За відсутності додатного переповнення, згідно (52), вихідний перенос суматора порядків

$$E_P = 0. \quad (53)$$

В області додатного переповнення суматора (51) вихідний поліном суматора утворює двійковий код

$$\begin{aligned} E_P \cdot 2^m + S_P(m, 1) &= \frac{V}{2} + P_C(m, 1) + 1 = \\ &= \frac{V}{2} + \frac{V}{2} = 2^m. \end{aligned} \quad (54)$$

При переповненні m -розрядного поля порядків модуля результату, згідно (54), вихідний перенос суматора

$$E_P = 1. \quad (55)$$

Таким чином, переповнення розрядної сітки суматора порядку, згідно (53) та (55), описується логічним виразом

$$ДПП = E_P. \quad (56)$$

де ДПП – ознака додатного переповнення суматора порядків (51).

При порушенні нормалізації праворуч у вихідному коді суматора мантис (44) (при $E_M^c S_M^c(n) = 00$) нормалізація результату досягається виконанням сукупності операцій:

$$\begin{aligned} 'MC(n, 1) &= 'S_M^c(n, 1); \\ 'MC(n, 1) &= L1('MC(n, 1)) \end{aligned} \quad (57)$$

$$P_C^{DH}(m, 1) = (P_C(m, 1) - 1)^{DH} \quad (58)$$

де $L1('MC(n, 1))$ – зсув ліворуч абсолютного значення поля мантис модуля результату.

Операції (57)–(58) циклічно повторюються до утворення в полі мантис нормалізованого коду $'MC(n, 1) = '1MC(n-1, 1)$.

Для коректування поля порядків модуля результату за допомогою суматора порядків представимо співвідношення (58) у вигляді:

$$\begin{aligned} (P_C(m, 1) - 1)^{DH} &= \frac{V}{2} + P_C(m, 1) - 1 = \\ &= P_C^{DH}(m, 1) - 1 = \\ &= (P_C^{DH}(m, 1) - 1 + V)_{mv} = \\ &= (P_C^{DH}(m, 1) - KE(m, 1))_{mv}. \end{aligned} \quad (59)$$

Значення $P_C^{DH}(P_C(m, 1))$, згідно (22), визначається інтервалом $[-V/2; V/2-1]$. Двійковий код $P_C(m, 1) - 1$ змінюється в діапазоні $[-(V/2+1); V/2-2]$, котрий більше області визначення коду з додатним нулем в полі порядків модуля результату. Тому при $P_C(m, 1) - 1 = -(V/2 - 1)$ в суматорі порядків (59) необхідно

фіксувати від'ємне переповнення m -розрядної сітки поля порядків.

Повна сума на виході суматора порядку, згідно (59), утворює поліном:

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) &= \\ &= P_C^{DH}(m, 1) + KE(m, 1) = \\ &= \frac{V}{2} + P_C(m, 1) + V - 1 = \\ &= V + \frac{V}{2} + P_C(m, 1) - 1. \end{aligned} \quad (60)$$

За відсутності від'ємного переповнення повна сума (60), згідно (22), змінюється в діапазоні:

$$E_p \cdot 2^m + S_p(m, 1) = \begin{cases} V + \frac{V}{2} + P_C(m, 1) - 1 = V + \frac{V}{2} - \frac{V}{2} = V = 2^m, \\ \text{при } P_C(m, 1) - 1 = -\frac{V}{2}; \\ V + \frac{V}{2} + P_C(m, 1) - 1 = V + \frac{V}{2} + \frac{V}{2} - 1 = \\ = V + V - 1 = 2^m + 2^m - 1 = \\ = 2^m + 2^{m-1} + 2^{m-1} - 1 > 2^m, \\ \text{при } P_C(m, 1) = \frac{V}{2} - 1 \end{cases}$$

Отже, за відсутності від'ємного переповнення суматора порядків (60) вихідний перенос

$$E_p = 1. \quad (61)$$

В області від'ємного переповнення при $P_C(m, 1) - 1 = -(V/2 + 1)$ поліном повної суми суматора (60) утворює двійковий код

$$\begin{aligned} E_p \cdot 2^m + S_p(m, 1) &= P_C^{DH}(m, 1) + KE(m, 1) = \\ &= \frac{V}{2} + P_C(m, 1) + V - 1 = \\ &= V + \frac{V}{2} + P_C(m, 1) - 1 = \\ &= V + \frac{V}{2} - \frac{V}{2} - 1 = V - 1 = 2^m - 1 = \\ &= 2^{m-1} + 2^{m-1} - 1 < 2^m. \end{aligned}$$

Отже, за від'ємного переповнення m -розрядної сітки поля порядків вихідний перенос суматора порядків (60)

$$E_p = 0. \quad (62)$$

Таким чином, факт від'ємного переповнення суматора порядків (60), згідно (61)–(62), описується логічним виразом

$$\text{ВПП} = \bar{E}_p. \quad (63)$$

де ВПП – ознака від'ємного переповнення суматора порядків.

Коли виникає ВПП суматора порядків (60) процедура нормалізації мантиси результатів (57)–(58) завершується і в полі порядків та мантис формується нечисло з рухомою комою.

Моделювання розроблених алгоритмів додавання та віднімання чисел з рухомою комою при представленні порядків в системі з додатним нулем виконано з використанням програмної моделі операційного пристрою (рис.1).

Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
+ ▢ A	5E																			3C
+ ▢ B	3C																			5A
▢ D																				
+ ▢ C	E8																			4E

Рисунок 1 – Результати моделювання операцій додавання та віднімання чисел з рухомою комою при представленні порядків в системі з додатним нулем

Висновки

Основні результати роботи зводяться до наступного:

- отримані залежності для обчислення знаку та способу обчислення модуля результату;
- встановлена відповідність ознак порівняння модулів чисел, порядків, мантис початкових даних з рухомою комою при представленні порядків в коді з додатним нулем;
- показано, що сукупність ознак порівняння модулів порядків та мантис чисел з рухомою комою при представленні порядків в коді з додатним нулем визначається обчисленням

модулів початкових даних як чисел з фіксованою комою;

- обґрунтований алгоритм коректування порядків операндів в коді з додатним нулем при вирівнюванні порядків початкових чисел;
- запропонований алгоритм перетворення порядку результату в коді з додатним нулем та спосіб фіксації переповнення поля порядків при нормалізації мантиси результату;
- розроблені машинні алгоритми операцій додавання та віднімання чисел з рухомою комою обґрунтовані методами програмного моделювання в середовищі ACTIVE-VHDL;

– запропоновані алгоритми додавання та перспективними для створення віднімання чисел з рухомою комою є конкурентоспроможних арифметичних пристроїв.

Список літератури

1. Карцев М.А. Арифметика цифровых машин. – М.: Наука, 1969. – 576с.
2. Каган, Б. М. Электронные вычислительные машины и системы : Учеб.пособие для вузов / Б.М. Каган. - 3-е изд. перераб. и доп. - Л. : Энергоатомиздат, 1991. – 590 с.
3. Степанов А.Н. Архитектура вычислительных систем и компьютерных сетей. – СПб.: Питер, 2007. – 509 с.
4. Крейгон Х. Архитектура компьютера и ее реализация. Москва: Мир, 2004. 416 с.
5. Святный В.А. Математическое описание компьютерных операций суммирования и вычитания целых чисел при смещенных кодах операндов / В.А.Святный, В.В.Лапко, А.В.Самощенко. // Наукові праці Донецького національного університету. Серія «Інформатика, кібернетика та обчислювальна техніка». – 2016. – Вип.1(22). – с.75-83.
6. Сергиенко, А.М. VHDL для проектирования вычислительных устройств /А.М.Сергиенко. – К.: «Корнейчук», 2003. – 208с.

References

1. Kartsev, M.A. (1969) Arithmetic of digital computers [Arifmetika tsifrovyykh mashin], Nauka, Moscow, 576 p.
2. Kagan B.M. (1991) Digital computers and systems [Electronnyye vychislitel'nyye mashiny i sistemy], Energoatomizdat, Moscow, 680 p.
3. Stepanov, A.N. (2007) Architecture of computer systems and computer networks. [Arkhitektura vy`chislitel'ny`x sistem i komp'yuterny`x setej]. SPB, Piter, 509p.
4. Krejgon, H. (2004). Computer architecture and implementation. [Arkhitektura komp'yutera i ee realizaciya]. Moskva, Mir, 416
5. Sviyatny, V.A. , Lapko, V.V., Samoschenko, A.V.(2016) Mathematical description of computer operations of summation and subtraction of integers with offset operand codes [Matematicheskoye opisaniye komp'yuternykh operatsiy summirovaniya i vychitaniya tselykh chisel pri smeshchennykh kodakh operandov] Naukovi p atsi DonNTU: Info matyka, Kybe netyka ta obchysliuvalna tekhnika, №1 (22), Donetsk National Technical University, Krasnoarmiysk, P. 75-83.
6. Sergienko, A.M. (2003) VHDL for designing computing devices. [VHDL dlya proektirovaniya vy`chislitel'ny`x ustrojstv]. Kiev, Korneychuk, 208p.

Надійшла до редакції 20.05.2019

А.В. САМОЩЕНКО, В.В. ЛАПКО, Г.Э.МАРГИЕВ

Донецкий национальный технический университет, г. Покровск

АЛГОРИТМ СЛОЖЕНИЯ И ВЫЧИТАНИЯ ЧИСЕЛ С ПЛАВАЮЩЕЙ ЗАПЯТОЙ ПРИ ПРЕДСТАВЛЕНИИ ПОРЯДКОВ В СИСТЕМЕ С ПОЛОЖИТЕЛЬНЫМ НУЛЕМ

Сформулированы логические выражения для определения знака результата и способа вычисления модуля результата. Установлено соответствие признаков сравнения модулей, порядков и мантисс исходных данных с плавающей запятой при представлении порядков в коде с положительным нулем. Показано, что совокупность признаков сравнения модулей, порядков и мантисс чисел с плавающей запятой при представлении порядков в системе с положительным нулем определяется вычитанием модулей исходных данных как чисел с фиксированной запятой. Обоснованы алгоритмы корректирования порядков операндов в коде с положительным нулем при выравнивании порядков исходных чисел. Предложены алгоритмы преобразования порядка результата и способы фиксирования переполнения порядков при нормализации мантиссы результата. Разработанные алгоритмы и математические модели операций сложения и вычитания чисел с плавающей запятой при представлении порядков в системе с положительным нулем подтверждены методами программного моделирования в среде ACTIVE-VHDL
Ключевые слова: модуль, порядок, мантисса, код с положительным нулем, денормализация, нормализация, переполнение, знак переполнения.

O. V. SAMOSHCHENKO, V.V. LAPKO, H. E. MARHIEV

Donetsk National Technical University, Pokrovsk

ALGORITHM OF ADDITION AND SUBTRACTION FOR FLOATING-POINT NUMBERS WITH REPRESENTED IN POSITIVE-ZERO CODE EXPONENTS

Operations of floating-point numbers addition and subtraction constitute a significant part of machine operations in computer systems. An analysis of floating point numbers modules can be significantly simplified with representation of their exponents with shifted positive-zero codes, binary images of which are proportional to the

exponents. Thus the aggregate of exponent and mantissa makes a fixed-point binary positional code, value of which is proportional to the module of the floating-point number. This allows operations of comparing modules, exponents and mantissas in procedures of floating-point numbers addition and subtraction reduce to simple arithmetic operation of fixed-point unsigned numbers subtraction.

Due to the limited bit capacity of adders for exponents and mantissas during calculating of floating-point numbers addition and subtraction results it is additionally necessary to fix a rise and polarity of an overflow sign. Besides that, when converting exponents and mantissas, the operations have to be performed in such a way, that the result at the adder main outputs is formed directly in the correct format.

In the given article, logical expressions for calculation of result's sign and the result's module calculating method were obtained. Correspondence between flags of modules, exponents and mantissas of initial floating-point data comparison, when representation of exponents with positive-zero code is determined. It is shown that a set of flags of modules, exponents and mantissas of initial floating-point data comparison, when representation of exponents with positive-zero code is determined by subtracting the initial data modules as fixed-point numbers.

Correction algorithms for positive-zero coded exponents of operands during exponents matching are proved. The result's exponent transformation algorithms and exponents overflow sign fixing methods during result mantissa normalization are proposed. The developed algorithms and mathematical models of floating-point numbers addition and subtraction operations are proved by software simulation in ACTIVE-VHDL simulation tool.

The proposed algorithm of addition and subtraction for floating-point numbers is promising for the creation of competitive arithmetic devices.

Keywords: module, exponent, mantissa, positive-zero code, denormalization, normalization, overflow, overflow sign.