

А. С. Янко, В. М. Звездін

Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна

РОЗРОБКА МЕТОДІВ ТА АЛГОРИТМІВ СТИСНЕННЯ ІНФОРМАЦІЙНИХ ДАНИХ ТАБЛИЧНИХ СТРУКТУР

Анотація. Розглянуто властивості непозиційної системи числення в системі залишкових класів (СЗК): незалежність залишків, рівноправність залишків і малорозрядність залишків. Використання основних властивостей СЗК дає можливість створення ефективних методів контролю, діагностики та виправлення помилок даних при введенні мінімальної інформаційної та часової надмірності, і побудови відмовостійких і швидкодіючих комп'ютерних систем обробки даних (КСОД) реального часу. Основна властивість СЗК, а саме малорозрядність залишків дозволяє застосовувати табличні методи реалізації арифметичних операцій. При застосуванні методів спеціального кодування інформації в СЗК, здійснюється скорочення таблиць ПЗП, стиснення даних та зменшення кількості необхідного обладнання. На основі даних досліджень було розроблено методи та алгоритми стиснення інформаційних даних табличних структур у СЗК.

Ключові слова: комп'ютерна система обробки даних, непозиційна кодова структура, система залишкових класів, система числення, табличний принцип.

Вступ

Зростаючі вимоги до сучасних комп'ютерних систем обробки даних (КСОД) зумовлюють необхідність постійного пошуку, розробки та впровадження методів швидкої обробки даних. Необхідно відзначити що, незважаючи на інтенсивний розвиток сучасних інформаційних технологій, що застосовуються при створенні апаратно-програмних засобів КСОД, в даній області залишається багато не вирішених науково-технічних завдань і проблем. В системах залишкових класів (СЗК) кожне число, представляється у вигляді декількох малорозрядних позиційних чисел, які є залишками від ділення вихідного числа на взаємно прості основи. У звичайній позиційній двійковій системі виконання операцій (наприклад, складання двох чисел) відбувалося послідовно за розрядами, починаючи з молодшого. При цьому утворюється перенесення в наступний старший розряд, що і визначає порозрядну послідовність обробки. В СЗК з'явилася можливість розпаралелити цей процес: всі операції над залишками в кожній основі виконуються окремо і незалежно (паралельно), отже, в зв'язку з їх малою розрядністю, нескладно і швидко. Мала розрядність залишків забезпечує можливість реалізації табличної арифметики, при якій результат операції не вираховується кожен раз, а, одного разу розрахований, поміщається в запам'ятовуючий пристрій (ЗП) і при необхідності зчитується з нього, тобто, операція в СЗК при табличній арифметиці виконується за один період синхронізуючої частоти (машинний такт).

Аналіз останніх досліджень і публікацій. Численні публікації останніх років свідчать про те, що перспективи створення КСОД реального часу на основі використання системи залишкових класів (СЗК) відкриває широкі можливості використання комп'ютерних систем з високим ступенем стиснення даних за рахунок основних властивостей системи числення і розпаралелювання процесу обробки цілочисельних даних. Результати досліджень в області створення швидкодіючих КСОД відомих авторів (Валаха М., Свободи А., Сабо Н., Акушського І. Я., Юдицького Д. І., Николайчука Я. М., Долгова О. І., Торгашо-

ва В. А., Амербаєва В. М., Shimbo A., Paulier P., Thornton M. A., Dreschler R., Miller D. M. та ін.) показали, що використання СЗК як системи числення КСОД, призначеної для реалізації цілочисельних арифметичних операцій додавання, віднімання і множення чисел, в додатному числовому діапазоні, істотно підвищує швидкодію вирішення задач певного класу, а методи, алгоритми та засоби реалізації арифметичних операцій на основі табличного принципу дозволяють не лише час підвищити швидкодію виконання операцій, а зменшити об'єм даних (зжати інформаційні потоки даних). За останні роки були розроблені наступні КСОД в СЗК: бортовий комп'ютер Star (США); спеціалізовані процесори ДПФ (США, Південна Корея); ряд військових спеціалізованих бортових комп'ютерів (США, Японія); спеціалізовані процесори ЦГЗ (США); комп'ютери Sprint для робототехніки (США, Японія); в китайській компанії "Trp Display Technology (Wuhan, China) Co., Ltd" при розробці та впровадженні бездротової сенсорної мережі системи контролю стану промислового устаткування при виготовленні моніторів; на підприємстві ТОВ "Релком-Поділля" при розробці системи відеоспостереження на основі бездротових мультимедійних сенсорних мереж; в корпорації "Cypress Semiconductors" при розробці апаратно-програмного забезпечення для модулів CY8CKIT-050 PsoC 5 і CyFi (CYRF7936), які можуть бути використані в бездротових сенсорних мережах [1].

Метою статті є розробка методів та алгоритмів стиснення інформаційних даних на основі табличного принципу з використанням властивостей непозиційних кодових структур у СЗК з урахуванням зменшення кількості необхідного обладнання.

Виклад основного матеріалу

Дослідження впливу основних властивостей системи залишкових класів на архітектуру та принципи функціонування спеціалізованих процесорів. Із вищерозглянутих прикладів випливають основні властивості системи залишкових класів:

- 1) незалежність залишків;
- 2) рівноправність залишків;
- 3) малорозрядність залишків.

Розглянемо як ці властивості впливають на структуру та принцип функціонування спеціалізованого обчислюваного пристрою [2].

1) *Незалежність залишків* дає можливість побудови ЕОМ у вигляді набору (по числу залишків СЗК) інформаційно-незалежних трактів, що працюють паралельно у часі при такій побудові ЕОМ обчислювана система в СЗК має модульність конструкції, що дозволяє здійснювати ремонт і технічне обслуговування не перериваючи розв'язування задач, і для здійснення профілактичних заходів ЕОМ не потрібен висококваліфікований обслуговуючий персонал. Окрім цього помилки, що виникли у тракті m_i , не “розмножуються” в інші тракти ЕОМ, при цьому байдуже чи мала місце на цій підставі однократна чи багаторазова, чи навіть пачка помилок довжиною більш $m_i - 1$ двійкових розрядів. Таким чином, помилка, що виникла в довільному тракті m_i ЕОМ у СЗК або збережеться в цьому тракті до кінця обчислень або у процесі подальших обчислень самоусунеться (наприклад, якщо після виникнення збою в залишку a_i проміжний результат стане числом, що має нульову цифру в залишку по m_i). У цьому випадку за допомогою СЗК можна побудувати систему виправлення помилок при введенні мінімальної надлишковості, що використовує динаміку обчислювального процесу, увівши поняття альтернативної сукупності.

Основна ідея визначення помилкового залишку $a_i = a_i + \Delta a_i$ полягає в тому, що для одержаної в результаті операції послідовності неправильних операндів A_i ($i = 1, 2, 3, \dots, p$) у динаміці обчислювального процесу, не перериваючи розв'язання задачі, послідовно визначаються умовні альтернативні сукупності $W(A) = W_{i-1}(A) \wedge W_i(A)$. За визначений час умовна альтернативна сукупність стягується до помилкового залишку (або двох залишків m_i і m_n). Після цього відомими методами проводиться корекція спотвореного залишку a_i . Відмінною рисою даного методу корекції помилок є можливість виправляти помилки без зупинки обчислень, що можливо для ЕОМ, які функціонують в реальному часі. Детальне дослідження розглянутої особливості СЗК дозволяє зробити висновки про те, що пристрої, які функціонують у СЗК, відносяться до таких об'єктів, які легко контролювати і легко діагностувати. Відзначена особливість ЕОМ, що функціонує в СЗК, сприяє розробці ефективних методів контролю і діагностики.

2) *Рівноправність залишків*. Будь-який залишок a_i числа A_k у СЗК несе інформацію про все вихідне число, що дає можливість чисто програмними методами замінити спотворений тракт по модулю m_i на справний (контрольний) тракт по модулю m_i ($m_i < m_i$), не перериваючи розв'язання задачі. Окрім того, ЕОМ в СЗК з двома контрольними основами зберігає свою працездатність при відмові будь-яких двох обчислювальних трактів. При виникненні третьої чи навіть четвертої відмови, ЕОМ все ще може виконувати програму при деякому зменшенні точності чи швидкості обчислень, тобто ЕОМ в СЗК є винятково “живучою”, наближаючись в цьому плані до живих організмів. Відзначимо, що дана особливість обумовлює одну із самих чудових властивостей СЗК: та сама ЕОМ може мати різну надійність при розв'язанні за-

дач в залежності від вимог, які висуваються до точності, обсягу пам'яті і швидкодії машини при їх розв'язанні, тобто в процесі розв'язання різних задач на ЕОМ у СЗК можливе здійснення “обмінних” операцій між точністю, швидкістю і надійністю.

3) *Малорозрядність залишків* дозволяє застосовувати табличні методи реалізації арифметичних операцій. У цьому випадку більшість арифметичних операцій здійснюється за один такт, що різко підвищує швидкість використання раціональних операцій. Одночасно табличні методи використання арифметичних операцій дозволяють створити на базі матричних схем високонадійні обчислювальні пристрої [3].

Отже, розглянуті властивості СЗК, при використанні її в операційному пристрої (ОПІ), дозволяють значно підвищити ефективність функціонування ЕОМ. Розробка методів та алгоритмів стиснення інформаційних даних табличних структур. Під табличною реалізацією табличних операцій $c_i = f(a_i, b_i)$ розуміється організація такої таблиці, в якій кожній комбінації вхідних величин a_i та b_i відповідає одне і тільки одне значення вихідної величини c_i . Нехай $[0, X)$ – діапазон зміни величини a_i ; $[0, Y)$ – діапазон зміни величини b_i ; $[0, Z)$ – діапазон зміни величини c_i . У цьому разі основні характеристики таблиць будуть представлені таким чином:

- надлишковість таблиці:

$$I = X^2 - Z; \quad (1)$$

- коефіцієнт використання таблиці:

$$\gamma = (Z / X^2) \cdot 100\%; \quad (2)$$

- коефіцієнт надлишковості таблиці:

$$W = (X^2 / Z) \cdot 100\%. \quad (3)$$

В подальшому будуть розглянуті методи, які дозволяють покращити основні характеристики таблиць арифметичних операцій в СЗК. Пошук шляхів спрощення структури ЕОМ привів до необхідності побудови алгоритмів реалізації модульних операцій, які дозволяють підвищити ефективність застосування табличної арифметики [4].

Розглянемо алгоритм реалізації операції модульного множення. Складемо таблицю із числових значень $a_i \cdot b_i \pmod{m_i}$. Ця таблиця симетрична відносно діагоналей, вертикалі і горизонталі, що проходять між числами $(m_i - 1)/2$ і $(m_i + 1)/2$. Симетричність таблиці відносно лівої діагоналі визначається комутативністю операції множення, симетричність відносно правої діагоналі визначається тим, що:

$$(m_i - a_i)(m_i - b_i) \equiv a_i b_i \pmod{m_i}. \quad (4)$$

Симетричність відносно вертикалі і горизонталі визначається із умови кратності суми симетричних чисел:

$$a_i b_i + a_i (m_i - b_i) \equiv 0 \pmod{m_i}; \quad (5)$$

$$a_i b_i + b_i (m_i - a_i) \equiv 0 \pmod{m_i}. \quad (6)$$

Щоб відновити таблицю модульного множення $a_i \cdot b_i \pmod{m_i}$, достатньо мати числову інформацію тільки про восьму її частину. Звідси випливає можливість скорочення таблиці (кількості схем співпадання ПЗП) модульного множення. Відмітимо, що зменшення таблиці у вісім разів призводить до необхідності проведення попереднього аналізу величин вхідних

операндів a_i і b_i , що збільшує час реалізації арифметичної операції. Для найбільш ефективної реалізації операції $a_i \cdot b_i \pmod{m_i}$ застосовуються методи спеціального кодування, які дозволяють в чотири рази зменшити таблицю множення. Розв'язання поставленої задачі можливе в результаті застосування спеціальних кодів. Розглянемо варіант виконання операції модульного множення з допомогою коду табличного множення [5].

Нехай дано вхідні операнди a_i і b_i . Значення a_i (b_i), що лежать в діапазоні $[0; (m_i - 1)/2]$, можуть бути закодовані довільним способом, а значення a_i (b_i), що лежать в діапазоні $[(m_i + 1)/2; m_i - 1]$, кодуються як $m_i - a_i$ ($m_i - b_i$). Для розрізнення діапазонів вводиться індекс:

$$\gamma_a(\gamma_b) = 0, \text{ якщо } 0 < a_i(b_i) < \frac{m_i-1}{2}; \quad (7)$$

$$\gamma_a(\gamma_b) = 1, \text{ якщо } \frac{m_i+1}{2} < a_i(b_i) < m_i - 1. \quad (8)$$

Алгоритм визначення результату операції модульного множення з допомогою коду табличного множення (КТМ) наступний: якщо задано два операнди в КТМ $a_i = (\gamma_a, a'_i)$, $b_i = (\gamma_b, b'_i)$, то для того, щоб отримати добуток цих чисел по модулю m_i , достатньо знайти добуток $a'_i \cdot b'_i \pmod{m_i}$ і інвертувати його узагальнений індекс γ_b , якщо γ_a відрізняється від γ_b , тобто:

$$a_i \cdot b_i \pmod{m_i} = (\gamma_i \cdot a'_i \cdot b'_i \pmod{m_i}), \quad (9)$$

При використанні даного алгоритму ПЗП, який реалізує операцію модульного множення, таблиці конструктивно зменшуються в чотири рази. При виконанні операції табличними методами в де яких випадках можливе додаткове зменшення обладнання за

рахунок того, що будується не єдина таблиця для модульних операцій, а k значно менших таблиць, які дозволяють дати відповіді по кожному з k розрядів результату, де k – розрядність регістра, що необхідна для зберігання цифри по основі, яка розглядається [6].

Висновки

Основна властивість СЗК, а саме малорозрядність залишків дозволяє застосовувати табличні методи реалізації арифметичних операцій. У цьому випадку більшість арифметичних операцій здійснюється за один такт, що різко підвищує швидкість використання раціональних операцій. Одночасно табличні методи використання арифметичних операцій дозволяють створити на базі матричних схем високонадійні обчислювальні пристрої КСОД. При застосуванні методів спеціального кодування інформації в СЗК, метою яких являється скорочення таблиць ПЗП, які реалізують табличні операції. Кількість обладнання при табличній побудові ОпП у СЗК скорочується мінімум в два рази в порівнянні з аналогами реалізованими на елементній базі з використанням двійкової позиційної системи числення. Також при виконанні модульної операції можливе додаткове скорочення обладнання за рахунок організації не єдиної таблиці (ПЗП), яка реалізує результат операції в машинному коді, а значно менших таблиць, які дають відповіді по кожному із k розрядів, де k – розрядність регістра, що необхідна для зберігання цифри за основою, яка розглядається.

СПИСОК ЛІТЕРАТУРИ

1. Виноградов И.М. Основы теории чисел. – Наука, 1981. – 176 с.
2. Акушкин И.Я., Юдицкий Д.И. Машинная арифметика в остаточных классах. – М.: Сов. радио, 1968. – 444 с.
3. Торгашев В.А. Система остаточных классов и надежность ЭВМ. – М.: Сов. Радио, 1973. – 118 с.
4. Барсов В.И., Краснобаев В.А., Фурман И.А., Малиновский М.Л., Шевченко В.В. Система обработки информации и управления АСУ ТП на основе применения кодов в модулярной арифметике. Х.: МОН, УИПА, 2009. 159 с.
5. В. А. Краснобаев, С. В. Сомов, А. С. Янко. Основные свойства непозиционной системы счисления // Системи управління, навігації та зв'язку. – 2013. – Вип. 1 (25). – С. 110–113.
6. Krasnobayev V., Yanko A. and Koshman S. A. Method for arithmetic comparison of data represented in a residue number system. Cybernetics and Systems Analysis, vol. 52, Issue 1, 2016, pp. 145–150.

Received (Надійшла) 18.12.2020

Accepted for publication (Прийнята до друку) 03.02.2021

Development of methods and algorithms for compressing information data of tabular structures

A. Yanko, V. Zvezdina

The properties of the non-positional number system in the system of residual classes (SRC) are considered: independence of residues, equality of residues and low-bit residuals. Use of the main properties of SRC gives the chance to create effective methods of control, diagnostics and correction of errors of data at introduction of the minimum information and time redundancy, and construction of fault-tolerant and high-speed computer systems of data processing (CSDP) of real time. The research of influence of the basic properties of SRC on architecture and principles of functioning of modern CSDP is considered. The main property of SRC, namely low-bit residuals allows you to use tabular methods of arithmetic operations. In this case, most arithmetic operations are performed in one cycle, which dramatically increases the speed of rational operations. At the same time, tabular methods of using arithmetic operations allow to create on the basis of matrix schemes highly reliable computing devices of computer data processing systems. When applying the methods of special coding of information in SRC, the purpose of which is to reduce the tables of RAM that implement tabular operations. The amount of equipment in the tabular construction of the operating device in the SRC is reduced by at least twice in comparison with analogues implemented on an element basis using a binary positional number system. Also, when performing a modular operation, it is possible to further reduce the equipment by organizing not a single table (RAM), which implements the result of the operation in machine code, but much smaller tables that give answers for each of the k bits, where k – is the bit size required for storing the figure on the basis under consideration. Based on the research data, methods and algorithms for compressing information data of tabular structures in SRC were developed.

Keywords: computer data processing system, non-positional code structure, number system, system of residual classes, tabular principle.