A High-Speed Realization of Chinese Remainder Theorem

Shuangching Chen and Shugang Wei
Department of Computer Science
Gunma University
Tenjin-cho 1-5-1, Kiryu-shi, Gunma 376-8515
Japan

Abstract: This paper proposes a novel technique for the Chinese remainder theorem (CRT) with the moduli \((2^n - 1, 2^n, 2^n + 1)\). Hardware implementation of the proposed CRT algorithm utilizes two kinds of parallel adders. One is referred to as modulo signed-digit \(m\) adder (MSDA) which performs a fast propagation-free addition and allows for the annihilation of carry or borrow chains using redundant binary number representation. Another implements parallel prefix adder which is the evolution of carry-lookahead adder (CLA). Compared to 16-digit Piestrak’s high-speed converter, the computation time is shortened by 34%.

Key Words: Residue number system, Signed-digit number, Carry-lookahead adder, Parallel prefix adder, Chinese remainder theorem.

1 Introduction

The residue number system (RNS) is an integer number system whose most important property is that additions, subtractions and multiplications are inherently carry-free [1, 2]. The residue number architecture with the three moduli \((2^n - 1, 2^n, 2^n + 1)\) has been widely used, since the residue addition can be performed by a binary adder [3, 4].

The residue number arithmetic can not be used for some applications, because the RNS does not have weights in the residue digits. Residue-to-binary number conversions are the crucial steps for any successful RNS application. For general moduli sets, residue-to-binary number conversions are based on the Chinese Remainder Theorem or Mixed Radix Conversion.

The three moduli \((2^n - 1, 2^n, 2^n + 1)\) can be performed efficiently with limited amount or even without ROM [5, 6, 7]. The complexity of the conversion has been greatly reduced by using compact forms with the multiplicative inverse and the properties of modular arithmetic. However, since the residue arithmetic mod \((2^n + 1)\) requires \((n + 1)\) bits to represent \((2^n + 1)\) states, it is not easy to perform mod \((2^n + 1)\) arithmetic with an end-around-carry adder. To overcome the drawbacks, several residue-to-binary arithmetic converters for the moduli set \((2^n - 1, 2^n, 2^n - 1)\) have been proposed [8, 9].

In this paper, we present a novel CRT algorithm using the redundant representation to improve the computation time. The primary advantage is that it is easy to represent \((2^n + 1)\) states for mod \((2^n + 1)\) with a signed-digit (SD) number representation. Wei et al. [10] have shown that it is efficient to use the SD number representation to design the three moduli \((2^n - 1, 2^n, 2^n + 1)\) adders without carry propagation. The multiplicative inverses of three moduli \((2^n - 1, 2^n, 2^n + 1)\) are also used to simplify our conversion algorithm.

2 Preliminaries

2.1 Residue Number System

In this paper, we consider a residue number system which has a set of relatively prime moduli, \(\{2^n, 2^n - 1, 2^n + 1\}\). A residue digit with respect to a modulus \(m_i\) is represented by the number set:

\[ l_{m_i} = \{0, \ldots, (m_i - 1)\} \tag{1} \]

Let \(M = \prod_{i=1}^{3} m_i = 2^n(2^n - 1)(2^n + 1)\). Szabo et al. [1] proved that if \(0 \leq A < M\), then the integer \(A\) has an one-to-one correspondence to its RNS representation. The integer \(A\) is uniquely represented by a 3-tuple \((a_1, a_2, a_3)\), where

\[ a_i = |A|_{m_i} = A - [A/m_i] \times m_i, \tag{2} \]

for \(i = 1, 2, 3\). In the above equation, \([A/m_i]\) is the integer part, and each residue digit is defined as the remainder of least magnitude when \(A\) is divided by \(m_i\).
2.2 the Extended Dynamic Range for SD Number Representation

A residue number $X$ can be represented by an $n$-digit radix-two SD number representation as follows:

$$X = x_{n-1}2^{n-1} + x_{n-2}2^{n-2} + \cdots + x_0,$$

where $x_i \in \{-1, 0, 1\}$, and $X$ can be denoted as $(x_{n-1}, x_{n-2}, \cdots, x_0)_{SD}$. To simplify the manipulation of the modular operation in an SD number representation, we apply the definition that each residue digit has the following redundant residue number set:

$$L_{m_i} = \{-2^n - 1, \cdots, 0, \cdots, (m_i - 1), \cdots, 2^n - 1\}.$$

Thus, $X$ must be in $L_{m_i}$ when it is expressed in an $n$-digit SD number representation. Obviously,

$$-X = -(x_{n-1}, x_{n-2}, \cdots, x_0)_{SD}$$

$$= (-x_{n-1}, -x_{n-2}, \cdots, -x_0)_{SD}$$

is also in $L_{m_i}$.

Definition 1 Let $Y$ be an SD number representation and $m$ be a modulus. Then $y = \langle Y \rangle_m$ is defined as an integer in $L_m$. When $|Y|_m \neq 0$, $y$ has one of two possible values given by equations

$$y = \langle Y \rangle_m = |Y|_m,$$

and

$$y = \langle Y \rangle_m = |Y|_m - $sign$ (|Y|_m) \times m,$$

where

$$\text{sign}(s) = \begin{cases} -1 & s < 0 \\ 1 & s \geq 0 \end{cases}.$$

When $|Y|_m = 0$ and $m = 2^n - 1$, there are three possible values for $y$, that is, $-m$, 0 and $m$. However, it is difficult to know if $Y$ is in $l_m$, because of the redundancy of the SD number representation.

The numbers as the intermediate results calculated in $L_m$ are used for fast residue arithmetic. If necessary for a final result, they can be converted into $l_m$.

2.3 Modulo $m$ Signed-Digit Adder

A novel residue arithmetic hardware algorithm using a radix-two SD number representation has been proposed to implement the modulo $m$ multiplication for the symmetric RNS [10, 11]. It is the key to increase the computation of such modular addition.

Figure 1 illustrates a circuit diagram of an $n$-digit modulo $m$ Signed-Digit adder (MSDA) with $n$ SD full adders (SDFAs), where $m = 2^n + \mu$ and $\mu \in \{-1, 0, 1\}$. One S DFA consists of ADD1 and ADD2. ADD1 generates the intermediate sum and the intermediate carry, and ADD2 sums the low intermediate carry and the intermediate sum. Let $c_i$ and $w_i$ be the carry and the intermediate sum of the $i$th digit position, respectively. Their values are determined by Table 1 with respect to the values of $x_i$, $y_i$, $x_{i-1}$, $y_{i-1}$.

Thus, the modulo $m$ addition can be performed in parallel without the carry propagation. We use $\otimes$ to mean an 1-by-1 multiplier.

Figure 1: Modulo $m$ Signed-Digit adder (MSDA)

3 Chinese Remainder Theorem

A number $X$ can be generally represented as $X = (x_n, x_{n-1}, \cdots, x_1)$ in an RNS, where $0 \leq x_i < m_i$. To convert $(x_n, x_{n-1}, \cdots, x_1)$ into the binary number representation $X_B$, the following CRT is generally used.

$$X_B = \sum_{i=1}^{n} \left( \frac{1}{m_i} \right) \left( x_i \right)$$

(7)
where
\[ \hat{m}_i = \frac{M}{m_i}, \]

\[ M = \prod_{i=1}^{n} m_i \]  

\( \{m_1, m_2, \cdots, m_n\} \) relatively prime integers

\[ \frac{1}{m_i} \mod m_i \] multiplicative inverse of \( \hat{m}_i \)

### 3.1 CRT for Moduli \( (2^n - 1, 2^n, 2^n + 1) \) Using Redundant Binary Number Architecture

The proposed conversion is shown in Fig.2, and the procedure is performed in the following two stages. In the first stage, the main conversion is based on the SD number architecture and we have no carry propagation during addition. In the second stage, for one-to-one correspondence, we convert numbers in \( L_m \) to those in \( L_m \) and the output is in the binary number representation.

**RNS with binary number representation**

\[ x \rightarrow (x_3, x_2, x_1)_{\text{RNS}} \]

**Converter based on signed-digit architecture in \( L_m \)**

\[ x_{\text{SD}} \rightarrow k_3, k_2, k_1 \]

**Signed-digit to binary in \( L_m \)**

\[ x_B \rightarrow b_3, b_2, b_1 \]

**Binary**

Figure 2: Residue-to-binary number conversions

Let \( m_1 = 2^n, m_2 = 2^n - 1, m_3 = 2^n + 1 \). Then \( \hat{m}_1 = 2^{2n} - 1, \hat{m}_2 = 2^{2n} + 2^n, \hat{m}_3 = 2^{2n} - 2^n \). We use the modulo operation \( \langle \cdot \rangle_{m} \) to deal with \( \cdot \mod m \), and then the multiplicative inverse is shown as follows:

\[ \langle \frac{1}{m_1} \rangle_{m_1} = -1, \]  

\[ \langle \frac{1}{m_2} \rangle_{m_2} = 2^{n-1}, \]  

\[ \langle \frac{1}{m_3} \rangle_{m_3} = 2^{n-1} + 1. \]

The proofs of (8), (9), (10) are based on the fact that \( \langle \hat{m}_i \rangle_{m_i} \mod m_i = 1 \). Let \( x_1, x_2, x_3 \) be residue numbers for modulo \( 2^n, 2^n - 1, 2^n + 1 \). Because a binary representation is one in the redundant number representation, we don’t need any computation for the conversion from the binary number representation to the redundant one. Thus the calculation of CRT can be written as

\[ X_{SD} = \langle A + B + C \rangle M \]  \( \text{(11)} \)

where \( M = \prod_{i=1}^{3} m_i \) and

\[ A = (2^{2n} - 1)\langle x_1 \rangle_{2^n}, \]  \( \text{(12)} \)

\[ B = (2^{2n} + 2^n)\langle x_2 \rangle_{2^n - 1}, \]  \( \text{(13)} \)

\[ C = (2^{2n} - 2^n)\langle x_3 \rangle_{2^n + 1} + x_3 \]  \( \text{(14)} \)

where \( X_{SD} \) means that \( X \) is in the SD number representation. Then the conversion result is \( X_{SD} = k_3 2^{2n} + k_2 2^n + k_1 \). Let \( TA = \langle x_1 \rangle_{2^n} = x_1 \), \( TB = \langle x_2 \rangle_{2^n - 1} \), and \( TC = \langle x_3 \rangle_{2^n + 1} + x_3 \). \( X \) can be written as

\[ X_{SD} = \left\langle (-2^{2n} + 1)TA + (2^n + 2^n)TB + (2^{2n} - 2^n)TC \right\rangle_{2^{2n+1}(2^n - 1)} \]

\[ = \left\langle -2^n(2^n)(TA) + 2^n(2^n + 1)TB + 2^n(2^n - 1)TC \right\rangle_{2^{2n+1}(2^n - 1)} + TA \]

\[ = 2^n \left\langle (-2^nTA) + (2^n + 1)TB \right\rangle_{2^{2n-1}} + TA \]

\[ = 2^n \left\langle (E + F)_{2^{2n-1}} + GA \right\rangle_{2^{2n-1}} + TA \]  \( \text{(15)} \)

where \( E = -2^nTA, F = (2^n + 1)TB \) and \( GA = (2^n - 1)TC \). Let \( GB = (E + F)_{2^{2n-1}} \). It is clear that \( k_1 = TA = x_1 \) since \( 0 \leq x_1 < 2^n \). The calculation of \( X_{SD} \) can be considered as how to compute the magnitude of \( 2^{2n} \) and \( 2^n, k_3 \) and \( k_2 \).

**Conversion algorithm A**

Input: \( x_1, x_2, x_3 \) (binary numbers)

Output: \( k_1, k_2, k_3 \) (SD numbers)

(1) Procedure for \( k_1 \)

\[ k_1 = x_1; \]

(2) Procedure for \( k_2 \) and \( k_3 \)

(2A) \( TA = x_1; \)
\[ TB = (2^{n-1}x_3)2^{n-1}; \]
\[ TC_1 = (2^{n-1}x_3)2^{n+1}; \]
(2B) \[ E = 2^n(-TA); \]
\[ F = (2^n+1)TB \]
(2C) \[ TC = (TC_1 + x_3)2^{n+1}; \]
\[ GB = (E + F)2^{n+1}; \]
(2D) \[ GA = (2^n-1)TC; \]
(2E) \[ k_32^n + k_2 = (GB + GA)2^{n-1}; \]

In (2A), \( TB \) and \( TC_1 \) are evaluated by performing the end-around-carry shift of \((n-1)\) digit positions to left. \(-TA\) is directly derived by changing the signs of nonzero digits. The word length of \( E, F \) and \( GA \) are twice of \( n \). In (2C), one \( 2n \)-digit and one \( n \)-digit SD additions are performed in parallel. In (2E), \( k_32^n + k_2 \) is evaluated by one \( 2n \)-digit MSDA. Therefore, the above algorithm needs two \( 2n \)-digit MSDAs, one \( n \)-digit MSDA. Note that the conversion is based on SD number system and the legitimate range of \( X_{SD} \) is \([-\(M-1\), \((M-1)\)] by the definition of the SD number for RNS.

### 3.2 Converting \( X_{SD} \) into binary number representation \( X_B \)

Then we convert \( X_{SD} \) into \( l_m \) for one-to-one correspondence. Because \( k_1 = x_1 \geq 0 \), to convert \( X_{SD} \) into \( l_m \) can be cantered as to convert \( k_32^n + k_22^n \) into \( l_m \). We use the property of \( \text{mod} \ 2^n \) arithmetic, and this leads to that we can carry out the conversion through \( 2^n - 1 \) parallel prefix adder.

**Definition 2** Let \( H \) and \( J \) be two integers in the binary number representation. If \( H + J = 2^n - 1 \), then \( J \) is 1’s complement of \( H \).

**Property 1** Let \( R = (r_{n-1}, \ldots, r_0) \) and \( r_i \in \{0, -1\} \). Let \( T = (t_{n-1}, \ldots, t_0) \). If \( R + (2^n - 1) = T \), then \( t_i \in \{0, 1\} \).

From Property 1, we know \( T \geq 0 \). It is enough to use the binary number representation to represent \( T \). We use \( Q = (q_{n-1}, \ldots, q_0) \) to represent \( R \) by the following steps: (1) If \( r_i = -1 \) then \( q_i = 1 \). (2) If \( r_i = 0 \) then \( q_i = 0 \). Since \( J = (2^n - 1) + (-H) \), then we can calculate \( T \) by using 1’s complement representation. For example: when \( R = (0, -1, 0, -1) \), then \( Q = (0, 1, 0, 1) \). By Definition 2, \( T \) is 1’s complement of \( Q \), and then \( T = (1, 0, 1, 0) \). We know \( R + (2^n - 1) = T \).

Let \( D = k_32^n + k_2 \) and \( D^+ \) and \( D^- \) be \( 2n \)-digit SD numbers for the positive digits and the negative digits, respectively. For example: if \( D = (-1, -1, -1, 0, 0, 0, 0, 0) \), then \( D^+ = (0, 0, 0, 0, 0, 0, 0, 0, 0) \) and \( D^- = (-1, -1, -1, 0, 0, 0) \). Thus \( k_32^n + k_2 \) can be written as

\[
D = |D^+ + D^-|_{2^{2n-1}} = |D^+ + D^- + (2^{2n-1})|_{2^{2n-1}} = |D^+ + DA|_{2^{2n-1}}
\]

where \( DA = D^+ + (2^{2n-1}) \geq 0 \) and \( D^+ \geq 0 \). Therefore, \( D \geq 0 \). We can use 1’s complement representation to calculate \( DA \). Since \( D^+ \geq 0 \) and \( DA \geq 0 \), it is enough to use the binary number representation to represent \( D^+ \) and \( DA \). Then, we use the binary number representation, \( DP \) and \( DN \), to express \( D^+ \) and \( DA \).

### Conversion algorithm B

**Input:** \( k_1, k_2, k_3 \) (SD numbers)

**Output:** \( b_1, b_2, b_3 \) (binary numbers)

Let \( X_B = b_32^{2n} + b_22^n + b_1 \),
\[ D = (d_{2n-1}, \ldots, d_0), \]
\[ DP = (dp_{2n-1}, \ldots, dp_0) \]
\[ DN = (dn_{2n-1}, \ldots, dn_0). \]

**Procedure for \( b_1 \)**

\[ b_1 = k_1; \]

**Procedure for converting \( D = k_32^n + k_2 \) into \( l_m \)**

(1) for \( i = 0 \) to \((2n - 1)\);

- if \( d_i = 0 \), then \( dp_i = 0 \) and \( dn_i = 1 \);
- else if \( d_i = 1 \), then \( dp_i = 1 \) and \( dn_i = 1 \);
- else \( dp_i = 0 \) and \( dn_i = 0 \);

(2) \[ b_32^n + b_2 = |DP + DN|_{2^{2n-1}}; \]

Consider how to calculate \( |DP + DN|_{2^{2n-1}} \).

A good idea to perform modulo \( 2^n - 1 \) adder has been proposed [13] by using Kogge-Stone tree structure which uses the associative operator ‘\( ⊓ ⊔ \)’ defined in Brent et al. [12] to implement the carry computation and the method is referred to as parallel prefix adder. Brent et al. [12] defined \( G_1 \) and \( P_1 \) as “block carry generate” and “block carry propagate”, respectively. They have shown that it suffices to compute all the \( G_1, G_2, P_1, P_2 \) for computing carry propagation. To calculate \( |DP + DN|_{2^{2n-1}} \), the above algorithm needs one modulo \((2^n - 1)\) parallel prefix adder.
4 Hardware Realization and Performance Evaluation

4.1 Hardware Realization

In hardware implementation, an SD digit $x$ is encoded as a 2-bit binary code defined as $x = [x^s, x^n]$, where $x^s$ is the sign and $x^n$ is the absolute value. This demands more hardware resources, but it also reallocates space, which effects the overall speed. We use a hardware description language, VHDL, to design the residue arithmetic circuits for the implementation of the proposed converters. Then, we performed a simulation under the conditions of 1-$\mu$m CMOS gate array technology.

The proposed converter based on conversion algorithms A and B, whose main blocks consist of three MSDAs and one modulo $(2^{2n} - 1)$ prefix adder, is shown in Fig. 3. Block $E$ is used to get $2^n(-x_1)$. Block $F$ is used to get $(2^n + 1)TB$ where $TB = (2^{n-1}x_2)_{2n-1}$. As mentioned before, the $TC1$ and $TB$ are designed as the shift-left operation. $GA$ combines $2^nTC$ and $(-TC)$. The division block is used to convert $L_n$ to $b_n$, which divides itself to two positive numbers including '0'. The modulo $(2^{2n} - 1)$ prefix adder[13] is used to sum the two positive numbers, such that $b_12^n + b_2$ is $[0, 2^{2n} - 2]$. Now we evaluate our proposed converters. Our aim is to enable high-speed conversion, so drawbacks in terms of area are not problems. The main process in Fig. 2 is on MSDA and modulo $2^{2n} - 1$ parallel prefix adder. The area of MSDA is the proportion of $n$ and the delay is independent of $n$ [11, 10]. The area of modulo $2^{2n} - 1$ parallel prefix adder [13] is $6n\log2n + 8n$ additions and the delay is $O(2\log2n + 3)$. Therefore, our conversion (Fig. 3) is faster than that by Piestrak [6] when $n$ is large.

The area comparison and delay time comparison are shown in Table 2. In Table 2, because our converters are based on the SD number representation, we use two bits to express one-digit SD number. This leads to that we need more hardware than the binary converters. Compared to 16-digit Piestrak’s high-speed converters [6], our method is fast and the computation time is shortened by 34%. The proposed converter based on parallel adders (MSDA, parallel prefix adder) is very fast when $n \geq 16$. The Andraos-Ahmad algorithm [5], which introduces compact forms of multiplicative inverses to simplify CRT, is an efficient technique. The algorithm uses four adders, two of which operate in parallel, to convert the moduli $(2^n + 1, 2^n, 2^n - 1)$ residue number into their binary equivalent. Recently, Piestrak [6] has suggested a simplification of the Andraos-Ahmad technique: the value of $-r_1$ modulo $2^{2n-1}$ can be easily obtained by the manipulation of $r_1$. Piestrak proposed two methods. The first method, referred to as the cost-effective (CE) version, uses two $2n$-bit carry save adders (CSAs) and one $2n$ carry propagation adder (CPA) with an end-around-carry to calculate $A + B + C - r_1$ of Andraos et al.[5]. The second method, which is referred to as the high-speed (HS) version, uses two $2n$-bit CSAs and two parallel $2n$-bit CPAs followed by a multiplexer. The CE converter needs $(4n + 1)$ full adder (FA) and the delay time is $(2t_{FA} + 2t_{CPA(2n)})$. The HS converter needs $(6n + 1)FA$ and $2n$-bit multiplexer (MUX), and the delay time is $(2t_{FA} + t_{CPA(2n)} + t_{MUX})$.

As mentioned before, mod $2^{2n} + 1$ requires $(n+1)$ bits to represent $2^n + 1$ states. Hiasat et al.[8] proposed using mod $2^{2n-1} - 1$ to instead of mod $2^{2n} + 1$. The main advantage is that arithmetic mod $2^{2n-1} - 1$ is more efficient than that mod $2^n + 1$, especially in the end-around-carry is positive number. Wang et al. [9] improved the method of Hiasat et al. [8], and saved one stage of modulo subtraction for computing $Y$ in Hiasat et al. [8]. The converter [9] needs $(7n - 3)FA$ and $(3n-7)$ half adder (HA), and the delay time is $(3n + 2)t_{FA}$.
Table 2: Performance of the CRT number converters

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>area (gate)</td>
<td>delay (ns)</td>
<td>area (gate)</td>
<td>delay (ns)</td>
</tr>
<tr>
<td>4</td>
<td>240</td>
<td>20.06</td>
<td>345</td>
<td>15.25</td>
</tr>
<tr>
<td>8</td>
<td>480</td>
<td>32.38</td>
<td>689</td>
<td>22.56</td>
</tr>
<tr>
<td>16</td>
<td>960</td>
<td>57.02</td>
<td>1377</td>
<td>37.19</td>
</tr>
</tbody>
</table>

The moduli \((2^n - 1, 2^n, 2^n - 1)\) have more beneficial than the three-moduli \((2^n - 1, 2^n, 2^n + 1)\) in RNS, because the end-around-carry adder of modulo \(2^n - 1\) operation is very simply. However, in the RNS-to-Binary conversion, the CRT based on the moduli \((2^n - 1, 2^n, 2^n - 1)\) is more complex than that based on moduli \((2^n - 1, 2^n, 2^n + 1)\). In fact, the Table 2 shows the computation time of [9] is longer than others.

5 Conclusion

The CRT algorithm based on MSDA and parallel prefix adder has been proposed. The proposed converter has been demonstrated throughout the paper using examples and analysis. The simulations show that the proposed schemes are high-speed architectures. Therefore, the conversion scheme will have less computation time than binary converters.

References:


