SIGN DETECTION IN NON-REDUNDANT RESIDUE NUMBER SYSTEM WITH REDUCED INFORMATION

SAROJ KAUSHIK

Centre for Computer Science and Engineering
Indian Institute of Technology, Delhi-110016, India.

ABSTRACT

A necessary and sufficient condition for sign detection in Non-Redundant Residue Number System by reducing the information of a residue digit has been obtained. The function to reduce the information of a residue digit x_p corresponding to a modulus m_p has been assumed to be periodic with the period length $\hat{m}_p$, where $\hat{m}_p = M/m_p$ and $M = \prod_{i=1}^{n} m_i$. A sequential method for determining the sign of a number is shown to demonstrate the applicability of the results thus proved.

INTRODUCTION

Sign detection is one of those functions which has nullified the advantages of residue arithmetic over conventional positional arithmetic. Numerous attempts have been made to reduce the time taken by this process. It was in sequal to these efforts that the possibility of reducing the information provided by the residue digits was considered. Such a reduction can be done in two ways - one by considering fewer digits than those in the residue representation of the number and the other, by reducing the information of a digit. It has been established by Szabo and Tanaka[1] that it is impossible to adopt the first strategy. However, for the second case, it has been proved in his coding theorem that all the information from a residue digit must be used in any sign determination process, provided the modulus m_p is smaller than M, where, $M = \prod_{i=1}^{n} m_i$. In the corollary to this theorem, it is proved that the sign detection is impossible if the p^{th} residue digit is coded into less than $\hat{m}_p$ states, where, $\hat{m}_p = M/m_p$. This yields a positive result. It shows that it is possible to reduce the information from a residue digit but only within a certain limit.

In this paper, we have used a periodic function with period length $\hat{m}_p$ to reduce the information of the residue digit x_p and have obtained necessary and sufficient conditions under which sign function is determined. Application of this theorem is shown by a sequential method for determining the sign of a number.

RESIDUE CODE

Consider an ordered set of n positive integers $(m_1, m_2, \ldots, m_n)$ such that $m_i \geqslant 2$ for any i, $1 \leqslant i \leqslant n$ are relatively prime to each other. These m_i's are called moduli and the corresponding ordered n-tuple $(x_1, x_2, \ldots, x_n)$ of least non-negative residues of a number X with respect to the moduli is called the residue representation of X. Such a representation of numbers forms the Residue Number System (RNS). Since all moduli are relatively prime, each $X \in [0,M)$, where $M = \prod_{i=1}^{n} m_i$, is uniquely represented in the RNS. We denote the residue representation of a number X with respect to the moduli $m_1, m_2, \ldots, m_n$ as

$$X \leftrightarrow (x_1, x_2, \ldots, x_n),$$

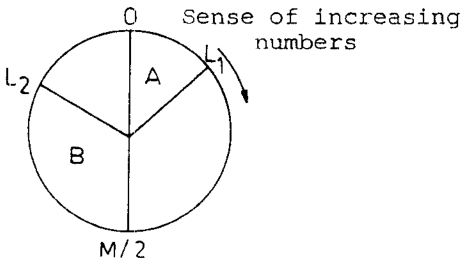
where $x_i = |X|_{m_i}$, $i = 1, 2, \ldots, n$.

REDUCTION OF INFORMATION

We shall first prove three lemmas on the basis of which the necessary and sufficient conditions shall be derived. Let L_1 and L_2 be the end points of two sectors A and B in the residue ring of M numbers. Sector A is traversed when proceeding from L_2 in the sense of increasing numbers and B is the remaining sector (Fig.1). Assume that L_2 belongs to A and L_1 belongs to B. Let m_p be any modulus such that $\beta\hat{m}_p < m_p$, where $\beta \geqslant 2$.

CH1892-9/83/0000/0024$01.00 © 1983 IEEE

Fig.1: Partitioning of the residue ring M



Since we intend to reduce the information of p^{th} residue digit to $\hat{m}_p$ states, construct a function $g(x_p)$ such that it may take on $\hat{m}_p$ values and is periodic with period length $\hat{m}_p$. Now consider a function f as

$$f(x_1, x_2, \dots, x_{p-1}, g(x_p), x_{p+1}, \dots, x_n) \cdot$$

It is a function of all x_i, $i \neq p$ and of $g(x_p)$ which maps all residue representations of the elements in A into one set of points and residue representation of the elements in B into a disjoint set of points. Then the function f will be the desired sign function. Now, with this definition of g and f we prove the following three lemmas and a theorem.

Lemma 1: There exists two points X and Y, $0 \leqslant X, Y < M$ such that $f(X) = f(Y)$ and $X \epsilon A$, $Y \epsilon B$ if $|L_1 - L_2|_M < m_p$ or $|L_2 - L_1|_M < m_p$ holds.

Proof: Let $d = |L_1 - L_2|_M < m_p$.

There will be two cases:

Case 1: $|L_2|_{m_p} \geqslant \hat{m}_p$.

Choose $X = L_2$ and $Y = L_2 - \hat{m}_p$.

Clearly $X \epsilon A$ and $Y \epsilon B$.

Now

$$|Y|_{m_p} = |L_2 - \hat{m}_p|_{m_p} = |X - \hat{m}_p|_{m_p} ,$$
$$= ||X|_{m_p} - \hat{m}_p|_{m_p} = |X|_{m_p} - \hat{m}_p,$$

since

$$|L_2|_{m_p} = |X|_{m_p} \geqslant \hat{m}_p \cdot$$

Then

$$g(|Y|_{m_p}) = g(|X|_{m_p} - \hat{m}_p) = g(|X|_{m_p}) \cdot$$

Also

$$|Y|_{\hat{m}_p} = |X - \hat{m}_p|_{\hat{m}_p} = |X|_{\hat{m}_p} \cdot$$

Therefore,

$$f(|X|_{\hat{m}_p}, g(|X|_{m_p})) = f(|Y|_{\hat{m}_p}, g(|Y|_{m_p})),$$

and hence $f(X) = f(Y)$.

Case 2: $|L_2|_{m_p} < \hat{m}_p$.

For $d \leqslant \hat{m}_p$, choose $X = L_2$ and $Y = X + \hat{m}_p$, then X and Y lie in different sectors.

Now,

$$|Y|_{m_p} = |X + \hat{m}_p|_{m_p}$$
$$= ||X|_{m_p} + \hat{m}_p|_{m_p} = |X|_{m_p} + \hat{m}_p ,$$

since $\beta\hat{m}_p < m_p$, $\beta \geqslant 2$.

Therefore,

$$g(|Y|_{m_p}) = g(|X|_{m_p} + \hat{m}_p) = g(|X|_{m_p}).$$

Also $|Y|_{\hat{m}_p} = |X + \hat{m}_p|_{\hat{m}_p} = |X|_{\hat{m}_p}$.

Hence $f(X) = f(Y)$.

If $d > \hat{m}_p$ and $|L_2|_{m_p} \neq 0$.

Then choose $Y = L_2 - 1$ and $X = Y + \hat{m}_p$.

Again $X \epsilon A$ and $Y \epsilon B$.

Now,

$$|Y|_{m_p} = |L_2 - 1|_{m_p} = ||L_2|_{m_p} - 1|_{m_p} = |L_2|_{m_p} - 1,$$

since $|L_2|_{m_p} < \hat{m}_p$.

Also,

$$|X|_{m_p} = |Y + \hat{m}_p|_{m_p} = ||Y|_{m_p} + \hat{m}_p|_{m_p} ,$$
$$= |Y|_{m_p} + \hat{m}_p ,$$

since $|Y|_{m_p} + \hat{m}_p = |L_2|_{m_p} - 1 + \hat{m}_p < m_p$.

Hence,

$$g(|X|_{m_p}) = g(|Y|_{m_p} + \hat{m}_p) = g(|Y|_{m_p}).$$

Further,

$$|X|_{\hat{m}_p} = |Y + \hat{m}_p|_{\hat{m}_p} = |Y|_{\hat{m}_p} ,$$

and consequently $f(X) = f(Y)$.

If $|L_2|_{m_p} = 0$ and $d \leqslant \beta\hat{m}_p$, then choose $X = L_2$ and $Y = X + \beta\hat{m}_p$. Clearly $X \epsilon A$ and $Y \epsilon B$.

Also

$$g(|Y|_{m_p}) = g(|X + \beta\hat{m}_p|_{m_p}),$$
$$= g(|X|_{m_p} + \beta\hat{m}_p) = g(|X|_{m_p}).$$

And $|Y|_{\hat{m}_p} = |X|_{\hat{m}_p}$, therefore $f(X) = f(Y)$.

If $|L_2|_{m_p} = 0$ and $d > \beta\hat{m}_p$.

Choose $Y = L_1$ and $X = L_1 - \hat{m}_p$.

Now $|X|_{m_p} = |Y - \hat{m}_p|_{m_p} = |Y|_{m_p} - \hat{m}_p$, since $|Y|_{m_p} > \hat{m}_p$.

Therefore, $g(|X|_{m_p}) = g(|Y|_{m_p})$.

Also $|X|_{\hat{m}_p} = |Y|_{\hat{m}_p}$ and hence the sign function is equal.

So, we conclude that if either of two conditions holds, there exists atleast two numbers X and Y, $0 \leqslant X, Y < M$ such that

$X \in A$, $Y \in B$ and $f(X) = f(Y)$. In otherwords, information of p^{th} digit cannot be reduced to m_p states if either of the following holds.

$$|L_1 - L_2|_M < m_p ,$$
$$|L_2 - L_1|_M < m_p .$$

Lemma 2: There exists atleast two numbers X and Y such that $X \in A$, $Y \in B$ and $f(X)= f(Y)$ if $L_2 \neq \alpha m_p$, $\alpha \geqslant 0$.

Proof: Assume $L_2 = \alpha m_p + t$, $\alpha \geqslant 0, 1 \leqslant t < m_p$. There are two cases:

Case 1: $\hat{m}_p \leqslant t < m_p$,

Choose $X = L_2 \leftrightarrow (x_1, x_2, \ldots, t, \ldots, x_n)$,

where $x_i = |\alpha m_p + t|_{m_i}$, $i = 1, 2, \ldots, n$,

and $Y = X - \hat{m}_p \leftrightarrow (x_1, x_2, \ldots, t - m_p, \ldots, x_n)$.

Clearly $X \in A$ and $Y \in B$.

Now,

$$g(|Y|_{m_p}) = g(|X - \hat{m}_p|_{m_p}) = g(t - \hat{m}_p),$$
$$= g(t) = g(|X|_{m_p}).$$

Also $|Y|_{\hat{m}_p} = |X - \hat{m}_p|_{\hat{m}_p} = |X|_{\hat{m}_p}$,

therefore,

$$f(X) = f(Y).$$

Case 2: $0 \leqslant t < \hat{m}_p$.

Choose $Y = \alpha m_p + r$, $r < t$, and $X = Y + \hat{m}_p$.

Here $X \in A$ and $Y \in B$.

$$g(|X|_{m_p}) = g(|Y + \hat{m}_p|_{m_p}) = g(||Y|_{m_p} + \hat{m}_p|_{m_p}),$$
$$= g(|r + \hat{m}_p|_{m_p}) = g(r + \hat{m}_p),$$

since $r + \hat{m}_p < 2\hat{m}_p < m_p$.

Then $g(|X|_{m_p}) = g(r) = g(|Y|_{m_p})$.

Further, $|X|_{\hat{m}_p} = |Y|_{\hat{m}_p}$ and consequently,

$f(X) = f(Y)$.

Lemma 3: There exists two points X and Y such that $X \in A$, $Y \in B$ and $f(X) = f(Y)$ if

$$|L_1 - L_2|_M \neq q\, m_p, \ q \geqslant 1.$$

Proof: Assume that $|L_1 - L_2|_M = q\, m_p + h$, $0 < h < m_p$. Two cases arise:

Case 1: L_2 is a multiple of m_p i.e., $L_2 = s\, m_p$ for some s, $s \geqslant 0$. If $\hat{m}_p \leqslant h < m_p$, then choose $Y = s\, m_p + q\, m_p + h$ and $X = Y - \hat{m}_p$.

Here $X \in A$ and $Y \in B$.

Now,

$$|Y|_{\hat{m}_p} = |X|_{\hat{m}_p} .$$

Therefore, $f(X) = f(Y)$.

If $1 \leqslant h < \hat{m}_p$, then choose $X = sm_p + qm_p + h - 1$ and $Y = X + \hat{m}_p$. Clearly $X \in A$ and $Y \in B$, since length of each interval is greater than m_p. In this case also $f(X) = f(Y)$.

Case 2: If $L_2 \neq$ multiple of m_p.

Proof of this case is exactly same as that of lemma 2.

Theorem:

There does not exist any pair of points $0 \leqslant X, Y < M$ such that $X \in A$, $Y \in B$ and $f(X) = f(Y)$ if and only if the following conditions hold true.

1. L_2 = multiple of m_p
2. $|L_1 - L_2|_M = tm_p$, $t \geqslant 1$.

Proof:

In order to prove the sufficiency of the above conditions, we assume that the conditions (1) and (2) are true and then we show that there does not exist any pair of numbers $0 \leqslant X, Y < M$ such that $X \in A, Y \in B$ and $f(X) = f(Y)$.

Let $X = L_2 + u$, $u < tm_p$.

Then $X \in A$. and $X \leftrightarrow (x_1, x_2, \ldots, |u|_{m_p}, \ldots, x_n)$, where

$$x_i = |L_2 + u|_{m_i}, i = 1, 2, \ldots, n.$$

Choose $Y = X + t_1 \hat{m}_p$ for all those t_1 such that $Y \in B$. Then, $Y \leftrightarrow (x_1, \ldots, |u + t_1 \hat{m}_p|_{m_p}, \ldots, x_n)$.

Now,

$$g(|Y|_{m_p}) = g(|u + t_1 \hat{m}_p|_{m_p}) = g(u + t_1 \hat{m}_p - t_2 m_p),$$

for some $0 \leqslant t_2 < \hat{m}_p$.

Also,

$$g(|X|_{m_p}) = g(|u|_{m_p}) = g(u - t_3\, m_p),$$

for some $0 \leqslant t_3 < \hat{m}_p$. If $g(|X|_{m_p}) = g(|Y|_{m_p})$, then $g(u - t_3 m_p) = g(u + t_1 \hat{m}_p - t_2 m_p)$.

$\Rightarrow u - t_3 m_p = u + t_1 \hat{m}_p - t_2 m_p + t_4 \hat{m}_p$, for for some integer t_4,

or,

$$(t_2 - t_3) m_p = (t_1 + t_4) \hat{m}_p .$$

Now $t_2 \neq t_3$ due to the conditions (1) and (2) and $(t_2 - t_3) < \hat{m}_p$. Hence $(t_2 - t_3) m_p = (t_1 + t_4) \hat{m}_p$ does not hold since m_p and $\hat{m}_p$ are relatively prime.

Therefore, $g(|X|_{m_p}) \neq g(|Y|_{m_p})$ and consequently $f(X) \neq f(Y)$ for $X \in A$ and $Y \in B$.

Necessity of condition (1): Assume that (1) does not hold i.e., $L_2 \neq$ multiple of m_p.

Then by lemma 2, there exists two numbers $0 \leqslant X, Y < M$ such that $X \in A$, $Y \in B$ and $f(X) = f(Y)$. This contradicts the fact that there does not exist any numbers X and Y such that X and Y belong to different sectors with equal sign function.

Necessity of condition (2): Assume on contrary that $|L_1 - L_2|_M \neq$ multiple of m_p. Then two cases arise:

Case 1: $|L_1 - L_2|_M < m_p$. Then lemma 1 contradicts the assumption.

Case 2: $|L_1 - L_2|_M > m_p$. Then lemma 3 proves the existence of atleast one pair of numbers $0 \leq X, Y < M$ such that $X \in A, Y \in B$ and $f(X) = f(Y)$, hence again a contradiction to the assumption. Hence the theorem.

Remark 1: For $\beta = 1$, the number X has a unique representation if

$$m_p - \hat{m}_p \leq |X|_{m_p} \leq \hat{m}_p \, .$$

Remark 2: If we consider condition (2) as if $|L_1-L_2|_M$ or $|L_2-L_1|_M$ is not a multiple of m_p then the sign function can be defined by introducing some check, say if sign function for two numbers lying in the different sectors is equal, then $X \in A$ only if

$$|X|_{m_p} < |L_1|_{m_p} \, .$$

In the next section, we present a sequential method to detect the sign of a number which is used to demonstrate the applicability of the result provedaabove. It is based on the sequential method proposed by Szabo and Tanaka [1].

SEQUENTIAL SIGN DETERMINATION

Let $m_1, m_2, \ldots, m_n$ be n mutually prime moduli and $X \longleftrightarrow (x_1, x_2, \ldots, x_n)$ be a number whose sign is to be determined. Assume that X is non-negative in the range $[0, M/2)$ and negative in the range $[M/2, M)$, so that this assumption satisfies the condition $|L_1-L_2|_M > m_p$, where m_p is that modulus whose information is to be reduced to have $\hat{m}_p$ output states.

First define a number q to be the largest number which satisfies the relationship $\prod_{i=1}^{q} m_i < \sqrt{M}$. So by coding theorem, the set of moduli $(m_1, m_2, \ldots, m_q)$ can be regarded as a composite modulus and hence up to unit q, every unit must contain a decoding net. Next form a composite modulus of size $m_p = \prod_{i=1}^{s} m_i$, for some $s, q+1 \leq s \leq n-1$, the unit s will have $\hat{m}_p = \prod_{i=s+1}^{n} m_i$ output states according to the theorem proved in this paper.

Denote the class of numbers which have the first j residue digits $x_1, x_2, \ldots, x_j$ to be the same as that of number X by $C^X_{x_1 x_2 \cdots x_j}$. This class contains $\prod_{i=j+1}^{n} m_i$ members and includes non-negative as well as negative numbers. These members may be generated in numerical order by successively adding $\prod_{i=1}^{j} m_i$ to the smallest non-negative number of the class until M is exceeded. Let y_j^1 be the smallest non-negative member of this class. The next larger members are, say $y_j^2, y_j^3, \ldots, y_j^{\bar{m}}$, where $\bar{m} = \prod_{i=j+1}^{n} m_i$. The sign of any member of this class may be determined in the following way.

$$y_j^t \leq C^X_{x_1 x_2 \cdots x_j} \text{ is non-negative iff}$$

$$t \leq \left\lceil \prod_{i=j+1}^{n} m_i/2m \right\rceil^{*}$$

or if $t = \left\lceil \prod_{i=j+1}^{n} m_i/2 \right\rceil + 1$, then if $(x_1, x_2, \ldots, x_j)$ is non-negative.

As said earlier, there will be decoding nets up to unit q. In the (q+1) unit the residue digits 1 through (q+1) of X are known and thus X can be identified as a member of the class $C^X_{x_1 x_2 \cdots x_{q+1}}$. The lowest member of this class may be obtained by table look-up having $\prod_{i=1}^{q+1} m_i$ entries, say it is y^1_{q+1}. Let p be the serial number of the location of $(x_1, x_2, \ldots, x_{q+1})$ in the above table. The (q+1) unit then transmits $|y^1_{q+1}|_{m_{q+2}}, |y^1_{q+1}|_{m_{q+3}}, \ldots, |y^1_{q+1}|_{m_n}$ to the (q+2) unit. In this unit, next larger members of the class are obtained by adding $|\prod_{i=1}^{q+1} m_i|_{m_j}$ to $|y^1_{q+1}|_{m_j}$, $j = q+2, q+3, \ldots, n$.

It is sufficient to find out only first m_{q+2} members of this class, since we want to choose the smallest member such that the residue digit corresponding to modulus m_{q+2} is equal to x_{q+2}. Let this member be denoted by $y_{q+1}^{t_{q+1}}$ for some t_{q+1}.

Write $y^1_{q+2} = y_{q+1}^{t_{q+1}}$.

Likewise proceeding we can find

$$C^X_{x_1 x_2 \cdots x_{q+3}}, C^X_{x_1 x_2 \cdots x_{q+4}}, \ldots, C^X_{x_1 x_2 \cdots x_n} \, .$$

Let $y_{n-1}^{t_{n-1}}$ be the smallest member of the class $C^X_{x_1 x_2 \cdots x_{n-1}}$ for some t_{n-1} which has n^{th} residue digit equal to x_n. Finally perform the following procedure.

Step 1: Set $j \leftarrow n-1$.

Step 2: If $t_j \leq \lceil m_{j+1}/2 \rceil$ then X is non-negative else if m_{j+1} is even then X is negative and stop. If m_{j+1} is not an even modulus and

* $\lceil I \rceil$ denotes ceiling of I, i.e., smallest integer$\leq$

$t_j = \lceil m_{j+1}/2 \rceil + 1$ then go to next step else X is negative. Stop.

Step 3: Set $j \longleftarrow j-1$. If $j > q+1$ then go to step 2 else if $p < \lceil \prod_{i=1}^{q+1} m_i/2 \rceil$ then X is non-negative else negative. Stop.

The following example has been considered which makes use of the sequential method suggested above.

Example: Determine the sign of a number $X \longleftrightarrow (10,4,1,1)$ in the Non-redundant Residue Number System of moduli $m_1 = 11$, $m_2 = 7$, $m_3 = 5$ and $m_4 = 3$.

Solution: Find $M = \prod_{i=1}^{4} m_i = 1155$ and $\sqrt{M} = 33.985$. Here $q = 1$, since $m_1 m_2 > \sqrt{M}$. Therefore, unit 1 has a decoding net i.e., it must transmit the total amount of information received. In unit 2, the information from unit 1 and the second residue digit are received. This information consists of $m_1\ m_2$ states. However, since $m_1 m_2 > \sqrt{M}$, theorem says that it is sufficient to transmit the last two residue digits of the smallest number of the class $C^X_{x_1 x_2}$. This can be found by the table 1, which has $m_1 m_2$ entries. This table would be of the form

TABLE 1

S. No.	First 2 residue digits of X		Last 2 residue digits of X	
	x_1	x_2	x_3	x_4
0 1 2 ⋮	0 1 2 ⋮	0 1 2 ⋮	0 1 2 ⋮	0 1 2 ⋮
32	10	4	2	2
⋮ 76	⋮ 10	⋮ 6	⋮ 1	⋮ 1

Hence the smallest number of the class $C^X_{x_1 x_2}$ is any y_2^1 i.e., $y_2^1 \rightarrow (2,2)$.

Here p, the serial number of location of y_2^1 in Table 1 is 32.

Next numbers are generated by adding successively $|m_1 m_2|_{m_j}$ to $|y_2^1|_{m_j}$, $j = 3,4$. We get,

$y_2^2 \rightarrow (4,1)$, $y_2^3 \rightarrow (1,0)$, $y_2^4 \rightarrow (3,2)$ and $y_2^5 \rightarrow (0,1)$.

Here $y_2^3 \rightarrow (1,0)$ is such that $|y_2^3|_{m_3} = 1 = x_3$, and $\alpha_2 = 3$. Write y_3^1 as y_2^3 with the last residue digit only.

$y_3^1 \rightarrow (0)$, $y_3^2 \rightarrow (1)$ and $y_3^3 \rightarrow (2)$.

Here y_3^2 is such that $|y_3^2|_{m_4} = 1 = x_4$, and $\alpha_3 = 2$.

Next we proceedaaccording to the procedure.

Set $j = 4-1 = 3$.

As $\alpha_3 = 2 = \lceil m_4 /2 \rceil + 1$, step 3 will be performed.

Set $j = 3-1 = 2$. Now $2 > q+1 = 2$, check $p = 32 < \lceil 77/2 \rceil$.

Since it is so, the number X is positive.

REFERENCES

[1] SZABO, N.S. and TANAKA, R.I.(1967), Residue Arithmetic and its applications to computer technology, New York, McGraw-Hill.

[2] SZABO, N.S. (1962), Sign Detection in Non-Redundant Residue Number System, IRE Trans. Electron., Computers, Vol. EC-11, No. 4, pp. 494-500.

[3] KAUSHIK, S. (1980), On the Arithmetic Operations and Error Correction in Residue Code, Ph.D. Thesis, Indian Institute of Technology, New Delhi.