ON COMBINATIONAL LOGIC FOR SIGN DETECTION IN RESIDUE NUMBER SYSTEMS

D.K. Banerji
Department of Computer Science
University of Ottawa
Ottawa, Ontario
Canada K1N 6N5

SUMMARY

This paper is concerned with the algebraic sign detection of a number in a residue number system. The proposed solution is applicable only to nonredundant systems. The method utilizes a systematic decomposition of the sign function S that is based on some special properties of S. Starting with the canonical sum-of-products expression for S, we transform the expression to a form whose realization is simpler than the canonical form realization and, if possible, also simpler than the minimal sum-of-products realization. In some cases, the proposed method yields savings as high as 85% compared to the minimal sum-of-products realization for S.

INTRODUCTION

Residue number systems have been of interest to mathematicians for a very long time. However, the use of the system to carry out machine computation has attracted attention only within the last two decades. The most notable feature of the residue number system is that in the operations of addition, subtraction, and multiplication, any digit of the result is determined solely by the corresponding digits of the operands. This results in the elimination of carries from one residue position to another. However, one of the drawbacks of the system is the fact that the algebraic sign of any number in an arbitrary residue code is a function of all the residue digits. This makes the sign detection process quite cumbersome and slow.

It is the purpose of this paper to investigate the sign detection problem, which deserves special attention because it is also closely related to the problems of relative magnitude comparison and overflow detection.

RESIDUE CODES

Residue codes and their properties have been widely discussed in literature[1,4-10], and for this reason they will not be dealt with in detail here. Only the essentials will be briefly reviewed.

Definition 1: Let $M=\{m_1,m_2,\ldots,m_n\}$ be an ordered set of positive integers, where $m_i \geq 2$ for $i=1,2,\ldots,n$. The m_i are called "moduli" or "radices," and the corresponding ordered set $(x_1,x_2,\ldots,x_n)$ of least positive residues of a natural number X, with respect to the moduli, forms the residue representation or code for that number, where the least positive residue of X with respect to m_i is denoted by $|X|_{m_i}=x_i$. For example, if $M=\{2,3,5\}$ and X=14, then $|14|_2=0$, $|14|_3=2$, and $|14|_5=4$. Thus 14 is represented by (0,2,4) in this system.

In order to avoid redundancy (unless redundancy is desirable), the moduli of a residue number system must be pair-wise relatively prime i.e., the greatest common divisor of each pair of moduli must be unity. If this is so, the number of integers that can be coded uniquely in a system with moduli $\{m_1,m_2,\ldots,m_n\}$ equals the product $m_1m_2\ldots m_n$. This is a direct consequence of the Chinese Remainder Theorem[6]. In the case $M=\{2,3,5\}$, therefore, a total of 30 integers can be coded uniquely. These can correspond to the natural numbers 0 through 29.

A convenient way of representing negative integers is as follows. The residue number range is divided into two parts. One part is assigned to positive integers, and the other to negative integers. The negative integers are then represented in radix complement form, defined in terms of additive inverse. Thus -X is represented by X', where X' is the additive inverse of X, and is defined as follows. If $X=(x_1,x_2,\ldots,x_n)$, then $X'=(x_1',x_2',\ldots,x_n')$ where $x_i'=m_i-x_i$ for $i=1,2,\ldots,n$. Thus for $M=\{2,3,5\}$, -14 is represented by (0,1,1).

SIGN DETERMINATION

The problem of sign determination is a major problem encountered in using residue arithmetic for computation. Attaching a sign bit to a residue number is of little help because the magnitude of a residue number is not readily available, and therefore, after adding a positive and a negative number, for instance, the sign of the result is not immediately known. We have seen earlier how the whole range of representation in a residue system can be divided into two parts to represent positive and negative integers. One obvious way, therefore, is to convert a given residue number to its natural number form which will fall either in the positive or the negative region of the representation. However, this is not an attractive solution for the problem because it is slow, and therefore would tend to offset the advantage of speed in a residue arithmetic unit.

We consider a sign function S which has a value 0 for positive integers and 1 for negative integers, i.e, S=0 for $0 \leq X \leq [\frac{M}{2}]-1$, and S=1 for $[\frac{M}{2}] \leq X \leq M-1$, where $M=\prod_{i=1}^{n} m_i$.

In considering the sign detection problem, it could be expected that to get one bit of information (sign) all the residue information is not required. Szabo[8] has proved that such a scheme is impossible and in the general case no reduction of information from any residue digit is possible without loss of sign information. A method for sign detection, which falls within the limitations imposed by Szabo's coding theorem, has been proposed by Banerji & Brzozowski[11]; however, this method is still too slow compared to the speed of the three arithmetic operations mentioned earlier. The quantitative effect of the frequency of sign detection on the overall arithmetic speed in a residue machine has been discussed by Banerji[12,13]. The need exists to investigate methods to speed up the process of sign detection if residue arithmetic is to offer a viable alternative to fixed-base arithmetic. Towards this end, we develop in this paper a systematic approach to the design of combinational logic for sign detection. A few years ago such an approach would have appeared impractical because of the amount of logic required. (Recall that the sign function S is a function of all the residue digits of a number). However, in an era when entire processors are available on a single LSI chip, the combinational logic approach for sign detection is certainly worthy of serious attention.

Our method is based on a systematic decomposition of the sign function S that takes into account the cyclic structure of S. To the best of the author's knowledge, this is the first such approach to the sign detection problem. For a given switching function there exist many possible decompositions. However, for a function with a large number of variables (like S) it is impractical, if not impossible, to try all possible decompositions in order to find out which one is best. The decomposition that we have derived is by no means unique. But, as stated earlier, it does take advantage of the cyclic structure of S.

Informally, the starting point of our method is the canonical sum-of-products expression for S. We then rearrange the expression to arrive at a realization for S which is simpler than the canonical form realization and, if possible, also simpler than the minimal sum-of-products realization. We now describe the method in more details.

Formal Description of the Method

Consider a residue system $M=\{m_1,m_2,\ldots m_n\}$ and a number X with residue code $(x_1,x_2,\ldots x_n)$ i.e.,

$\langle X\rangle_M = (x_1,x_2,\ldots,x_n)$ where $x_i = |X|_{m_i}$, $i=1,2,\ldots n$.

Consider a typical modulus m_i. The binary representation $(\delta_{i,p_i-1},\ldots,\delta_{i,1},\delta_{i,0})$ of x_i is assumed to be stored in a register of p_i bits. In our approach we first completely decode x_i by producing binary signals ψ_{i,j_i} for each j_i, $0\le j_i\le m_i-1$, such that $\psi_{i,j_i}=1$ iff $x_i=j_i$.

Clearly, the sign function can be expressed as a sum of products of the form

$$\theta = \psi_{1,j_1}\psi_{2,j_2}\cdots\psi_{n,j_n}$$

where $(j_1,j_2,\ldots,j_n)$ represents a positive number. Thus $\bar{S}$ consists of $[\frac{M}{2}]$ * such products θ.

The canonical form of $\bar{S}$ requires a large amount of logic; our approach is to factor the expression for $\bar{S}$ so as to reduce this logic.

Let (M_1,M_2) be any partition on M such that $M_1 = \{m_1,m_2,\ldots,m_k\}$, $M_2 = \{m_{k+1},\ldots,m_n\}$. Let $M_1=\prod_{i=1}^{k} m_i$, and $M_2=\prod_{i=k+1}^{n} m_i$.

<u>Lemma 1</u>: For the set of all positive numbers $X, X+M_1, \ldots, X+tM_1$, where $0\le X\le M_1-1$, the corresponding products $\theta_o,\theta_1,\ldots,\theta_t$, respectively agree in k positions. The converse is also true.

<u>Proof</u>: The verification of this is routine. ■

If $\theta_o,\theta_1,\ldots,\theta_t$ have a common factor, then we can write $\Gamma = \theta_o+\theta_1+\ldots+\theta_t = \alpha(\beta_o+\beta_1+\ldots+\beta_t)$ (1)

We shall call Γ an <u>elementary term</u>, α its <u>head</u> and $(\beta_o+\beta_1+\ldots+\beta_t)$ its <u>tail</u>.

The number of such elementary terms in $\bar{S}$ is a function of the partition on M. The effect of this partition on the complexity of realization of $\bar{S}$ will be discussed later.

Some Useful Results

In this section we give some results which will prove to be useful in later analysis.

<u>Lemma 2</u>: For all integers I,J,K such that $J,K\ge 1$,

$$\left[\frac{[\frac{I}{J}]}{K}\right] = \left[\frac{I}{JK}\right].$$

* "[]" denotes the ceiling operator i.e., [I] is the <u>least integer</u> $\ge$I

<u>Proof</u>: For any real Y, $[Y]\ge Y>[Y]-1$, and if L is some integer such that $L\ge Y$, then $L\ge[Y]$. Hence we get

$$\left[\frac{I}{JK}\right] \ge \frac{I}{JK} > \left[\frac{I}{JK}\right]-1$$

or $$K\left[\frac{I}{JK}\right]\ge \frac{I}{J} > K\left(\left[\frac{I}{JK}\right]-1\right)$$

or $$K\left[\frac{I}{JK}\right]\ge\left[\frac{I}{J}\right]> K\left(\left[\frac{I}{JK}\right]-1\right)$$

or $$\left[\frac{I}{JK}\right]\ge \frac{[\frac{I}{J}]}{K} >\left[\frac{I}{JK}\right]-1.$$

From this it is clear that $[\frac{I}{JK}]$ denotes the least integer $\ge \frac{[\frac{I}{J}]}{K}$ i.e., $[\frac{I}{JK}] = \left[\frac{[\frac{I}{J}]}{K}\right]$. ■

<u>Lemma 3</u>: In an elementary term the number n_t of terms in the tail (called the <u>size</u> of the tail) satisfies:

$$n_t\le \left[\frac{M_2}{2}\right].$$

<u>Proof</u>: The number of terms in the tail is t+1, where $X_t = X+tM_1\le [\frac{M}{2}]-1$ (by assumption). Thus

$$t \le ([\tfrac{M}{2}]-1-X)\frac{1}{M_1} \le ([\tfrac{M}{2}]-1)\frac{1}{M_1}, \text{ since } X\ge 0.$$

So, $$t\le \frac{1}{M_1}[\tfrac{M}{2}] - \frac{1}{M_1}$$

$$\le \left[\frac{1}{M_1}[\tfrac{M}{2}]\right] - \frac{1}{M_1} = \left[\frac{M}{2M_1}\right] - \frac{1}{M_1},$$ by Lemma 2. Since t must be an integer, the highest value that t can have is $[\frac{M}{2M_1}]-1 = [\frac{M_2}{2}]-1$. Thus $n_t = t+1\le [\frac{M_2}{2}]$, which proves the result. ■

This bound can be met as the next result shows. Unless otherwise specified, we assume that $\bar{S}$ consists of M_1 elementary terms formed by considering all $X_t = X+tM_1$, $t\ge 0$, $0\le X\le M_1-1$ and $X_t\le[\frac{M}{2}]-1$.

<u>Lemma 4</u>:

i) If M_2 is even, then each tail is of size $\frac{M_2}{2}$.

ii) If M_2 is odd, then there exist $[\frac{M_1}{2}]$ tails of size $[\frac{M_2}{2}]$ and $M_1-[\frac{M_1}{2}]$ tails of size $[\frac{M_2}{2}]-1$.

<u>Proof</u>: For a given X such that the corresponding elementary term Γ_X has a tail of maximum size $[\frac{M_2}{2}]$, we must have $X+tM_1\le[\frac{M}{2}]-1$, where $n_t=t+1=[\frac{M_2}{2}]$. Hence, $t = [\frac{M_2}{2}]-1$ and

$$X\le[\tfrac{M}{2}] - 1 - ([\tfrac{M_2}{2}]-1)M_1 \quad (2)$$

i) If M_2 is even, then from (2) we get $X\le M_1-1$. Therefore, all the M_1 elementary terms have tails of size $[\frac{M_2}{2}] = \frac{M_2}{2}$.

ii) If M_2 is odd, the right hand side of (2) equals $[\frac{M_1}{2}]-1$. Hence, for all X such that $0\le X\le[\frac{M_1}{2}]-1$, the corresponding Γ_X has a tail of size $[\frac{M_2}{2}]$ i.e., there exist $[\frac{M_1}{2}]$ such tails. Now, consider $X=M_1-1$.

Then $$J = X+([\tfrac{M_2}{2}]-2)M_1$$
$$= M_1-1+([\tfrac{M_2}{2}]-2)M_1$$
$$= ([\tfrac{M_2}{2}]-1)M_1-1<\frac{M_2}{2}\cdot M_1-1$$

So $$J < \frac{M_1M_2}{2} -1\le \left[\frac{M_1M_2}{2}\right] -1$$

Hence $J \le [\frac{M}{2}] - 1$.

It is easily shown that for $X = [\frac{M_1}{2}]$, $X + ([\frac{M_2}{2}] - 1)M_1 = [\frac{M}{2}]$. Therefore, for all X, $[\frac{M_1}{2}] \le X \le M_1 - 1$, Γ_X has a tail of size $[\frac{M_2}{2}] - 1$ i.e., there exist $M_1 - [\frac{M_1}{2}]$ such tails. ∎

We now want to establish the maximum number of <u>distinct</u> tails of size n_t in $\bar{S}$.

<u>*Lemma 5:*</u> The maximum number of distinct tails of a given size n_t is M_2.

<u>Proof:</u> Let Γ_X be the elementary term formed by considering the positive numbers $X, X+M_1, \ldots, X+tM_1$. Then $n_t = t+1$. Let the residue code for X be (j_1, j_2), where $j_1 = |X|_{M_1}$ and $j_2 = |X|_{M_2}$. Then for all the numbers, the corresponding residues with respect to M_2 are:

$j_2, |j_2+M_1|_{M_2}, \ldots, |j_2+tM_1|_{M_2}$ i.e., the residues are <u>uniquely</u> determined by j_2. Hence the tail of Γ_X is also uniquely determined by j_2. Since $0 \le j_2 \le M_2 - 1$, there can be <u>at most</u> M_2 distinct tails of size n_t. This proves the lemma. ∎

This bound can be met. For example, if $M_1=5$, $M_2=4$, the number of distinct tails of size 2 is $4 = M_2$.

Let T_X and T_Y denote the tails of elementary terms Γ_X and Γ_Y respectively. Then we show the following for $X \ne Y$, $0 \le X$, $Y \le M_1 - 1$.

<u>Theorem 1:</u>

i) If $T_X = T_Y$ then $|X|_{M_2} = |Y|_{M_2}$.

ii) If $|X|_{M_2} = |Y|_{M_2}$, then T_X and T_Y differ in at most one element.

<u>Proof:</u> i) The proof is by contradiction. Let the elements of T_X and T_Y be $|X|_{M_2}, |X+M_1|_{M_2}, \ldots, |X+KM_1|_{M_2}$ and $|Y|_{M_2}, |Y+M_1|_{M_2}, \ldots, |Y+KM_1|_{M_2}$ respectively, where either $K = [\frac{M_2}{2}] - 1$ or $K = [\frac{M_2}{2}] - 2$. Suppose $|Y|_{M_2} \ne |X|_{M_2}$. Since $T_X = T_Y$, there must exist some $t > 0$, such that $|Y|_{M_2} = |X+tM_1|_{M_2}$. Therefore, $|Y+(K-t)M_1|_{M_2} = |X+KM_1|_{M_2}$ and $|Y+(K-t+1)M_1|_{M_2} = |X+(K+1)M_1|_{M_2}$. But we must have $|Y+(K-t+1)M_1|_{M_2} = |X|_{M_2}$, otherwise $|X|_{M_2}$ will not appear in T_Y. Therefore,

$$|X|_{M_2} = |X+(K+1)M_1|_{M_2}$$

or $$|(K+1)M_1|_{M_2} = 0$$

or $$(K+1)M_1 = LM_2 \text{ for some } L>0.$$

Since M_1 and M_2 are relatively prime, it is clear that $(K+1)$ must be a multiple of M_2. This is impossible since $K+1 < M_2$. Therefore $|Y|_{M_2} = |X|_{M_2}$.

ii) If $|X|_{M_2} = |Y|_{M_2}$, then $|X+tM_1|_{M_2} = ||X|_{M_2} + tM_1|_{M_2} = ||Y|_{M_2} + tM_1|_{M_2} = |Y+tM_1|_{M_2}$. Since, in general the size of a tail is either $[\frac{M_2}{2}]$ or $[\frac{M_2}{2}] - 1$, it is clear that T_X and T_Y differ in at most one element. ∎

<u>Example 1:</u> Let $M_1 = 10$ and $M_2 = 3$. Then $T_1 = \psi_{2,1} + \psi_{2,2} = T_4$ and $|1|_3 = |4|_3 = 1$. Also $|7|_3 = 1$ and $T_7 = \psi_{2,1}$. T_1 and T_7 differ by one element.

<u>Corollary 1:</u> If $M_1 < M_2$, then all tails are distinct.

<u>Proof:</u> For $M_1 < M_2$, there exist no X and Y, $0 \le X$, $Y \le M_1 - 1$, such that $|X|_{M_2} = |Y|_{M_2}$. Hence, from Theorem 1, $T_X \ne T_Y$ ∎

For example, if $M_1=2$, $M_2=3$, both the tails are distinct.

<u>Corollary 2:</u> If $M_1 > M_2$ and M_2 is even, then each tail appears either $[\frac{M_1}{M_2}]$ times or $[\frac{M_1}{M_2}] - 1$ times in $\bar{S}$.

<u>Proof:</u> Consider some X, $0 \le X \le M_1 - 1$, with residue code (j_1, j_2). Choose $Y = j_2 + tM_2 \le M_1 - 1$, $t > 0$. Then $|Y|_{M_2} = j_2 = |X|_{M_2}$. Since all tails have the same size (M_2 is even), by Theorem 1 we get $T_X = T_Y$. We have

$$Y \le M_1 - 1$$

or $$j_2 + tM_2 \le M_1 - 1$$

or $$t \le \frac{M_1 - 1 - j_2}{M_2} \le \frac{M_1 - 1}{M_2} \text{ since } j_2 \ge 0.$$

So $$t \le \frac{M_1}{M_2} - \frac{1}{M_2}$$

or $$t \le [\frac{M_1}{M_2}] - \frac{1}{M_2}.$$

Since t must be an integer, the highest value t can have is $[\frac{M_1}{M_2}] - 1$. Therefore, T_X appears <u>at most</u> $[\frac{M_1}{M_2}]$ times.

To show that T_X appears <u>at least</u> $[\frac{M_1}{M_2}] - 1$ times, consider $j_2 = M_2 - 1$ (highest value of j_2). Then

$$J = M_2 - 1 + ([\frac{M_1}{M_2}] - 2)M_2 = ([\frac{M_1}{M_2}] - 1)M_2 - 1 < \frac{M_1}{M_2} \cdot M_2 - 1 \text{ (since } [Y] - 1 < Y).$$

Hence $J < M_1 - 1$ and T_X appears <u>at least</u> $[\frac{M_1}{M_2}] - 1$ times. ∎

<u>Example 2:</u> Let $M_1=9$, $M_2=4$. The tail $(\psi_{2,0} + \psi_{2,1})$ appears $3 = [\frac{9}{4}]$ times and all others appear $2 = [\frac{9}{4}] - 1$ times.

We now summarize the previous results and then use them for estimating the combinational logic required in a given residue system. For moduli M_1 and M_2, we get the following: *

1. There are M_1 elementary terms.
2. There are at most M_2 distinct tails of a given size.
3. Each tail contains either $[\frac{M_2}{2}]$ or $[\frac{M_2}{2}] - 1$ elements.
4. For $M_1 < M_2$ all tails are distinct.
5. For $M_1 > M_2$, M_2 even, each tail appears at least $[\frac{M_1}{M_2}] - 1$ times. In this case at least $[\frac{M_1}{M_2}] - 1$ elementary terms can be combined to form a "head sum" of $[\frac{M_1}{M_2}] - 1$ elements.

   There can be at most M_2 such head sums.

In the following example, the previous results are used to estimate the complexity of combinational logic for sign detection in a given residue system.

<u>Example 3:</u> Let $m_1=19$, $m_2=20$ and $m_3=21$. Let $M_1=19$ and $M_2=20 \times 21=420$. We assume a maximum fan-in of 5. An

---

* M_1 and M_2 can be either individual moduli or composite moduli (i.e., products of other moduli).

analysis shows that the logic required in this case is $\simeq$ 1530 AND/OR gates, and the time required for sign detection, $T_S=9\Delta$, where Δ denotes single gate delay.

The logic requirement can be reduced by more than 50 per cent by choosing M_1 and M_2 properly. Let $M_1=19X21$ and $M_2=20$. In this case, the total requirement = 645 AND/OR gates and $T_S=7\Delta$.

In Table I, we list the logic complexity of sign detection for several residue systems. The figures represent orders of magnitude. From the Table we note that except for very small values of M, our method provides tremendous savings as compared to minimal cover realization for S.

CONCLUSION

We have developed a systematic approach to the combinational logic realization for S. This appears to be the first such approach to the sign detection problem. The properties of the sign function have been studied and used in estimating the combinational logic requirements. The method is shown to provide great savings in comparison to the minimal cover realization for S.

Table I. Combinational Logic Complexity for Sign Detection

MODULI	Range M	Number of Gates: Using Minimal Cover	Number of Gates: Using our Method
2,3,5	30	13	30
2,3,5,7	210	74	78
8,9,11	792	300	170
19,20,21	7980	2750	645
2,3,5,7,11,13	30,030	10,000	1500

REFERENCES

[1] H. Aiken and W. Semon, "Advanced Digital Computer Logic," Computing Labs., Harvard University, Cambridge, Mass., Rept. WADS TR-59-472, July 1959

[2] P.W. Cheney, "A Digital Correlator Based on the Residue Number System," IRE Trans. Electronic Computers, vol. EC-10, p.p. 63-70, March 1961.

[3] W.E. Deskins, Abstract Algebra. New York: Collier-Macmillan, 1964.

[4] I. Flores, The Logic of Computer Arithmetic. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

[5] H.L. Garner, "The Residue Number System, "IRE Trans. Electronic Computers, vol. EC-8, p.p. 140-147, June 1959.

[6] G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers. New York: Oxford, 1954.

[7] A. Svoboda, "The Numerical System of Residual Classes (SRC)," Computer Prog. Czechoslovakia.

[8] N. Szabo, "Sign Detection in Nonredundant Residue Systems," IRE Trans. Electronic Computer, vol. EC-11, p.p. 494-500. August 1962.

[9] "Modular Arithmetic Computing Techniques," Wright-Patterson Air Force Base, Ohio, Rept. ASD-TDR-63-280, May 1963.

[10] N.S. Szabo and R.I. Tanaka, Residue Arithmetic and its Applications to Computer Technology, New York: McGraw-Hill, 1967.

[11] D.K. Banerji and J.A. Brzozowski, "Sign Detection in Residue Number Systems", IEEE Transactions on Computers, vol. C-18, No. 4, April 1969, p.p. 313-320.

[12] D.K. Banerji, "Residue Arithmetic in Computer Design," Ph.D. thesis, Dept. of Applied Analysis and Computer Science, University of Waterloo, Waterloo, Ontario, Canada, March 1971.

[13] D.K. Banerji, " On the Use of Residue Arithmetic for Computation," IEEE Transactions on Computers, Vol. C-23, No. 12, December 1974, p.p. 1315-1317.