Method of Numbers' Dichotomic Decomposition for Generation of Equal Probability Binary Sets

Rabah AlShboul

Computer Science Department
Al Albayt University
Mafraq, Jordan

Vitaliy A. Romankevich
Department of Applied Mathematics
Igor Sikorsky Kyiv Polytechnic Institute
Kyiv, Ukraine

Annotation
In this paper a method for analyzing the components of a sequential dichotomic decomposition of integers is proposed. Some conditions that allow to generate a sequence of equiprobable binary vectors of given dimensions are formulated. The variety of vectors is limited to a certain and predetermined numerical value. That value is chosen for the dichotomic decomposition and is not necessarily related to the power of 2. It is proved that it's always possible. The method is intended for using in controlled specialized means for structural generation of pseudo-random binary vectors.

Key words:
dichotomic decomposition, equiprobable binary vectors, pseudo-random generators.

1. Introduction

Already known methods and corresponding specialized structural means for the formation of pseudo-random (PR) sequences are concerned, as a rule, with certain definite values of such a parameter as the number of different output binary sets being formed by the generators. So in classical generators based on shift registers with linear feedback [1-4], the number of different output binary sets is determined by the power of 2. In many other known implementations such problem is being solved by generating a sufficiently large PR numbers with the subsequent division of the obtained values by an amount n - a given parameter. That limits from above the power of set of different elements (numbers) in generated sequence. Peculiarity of such approach appears in complexity of its structural implementation as well as in the multistage (multi-step) procedures for dividing of multi-digit integers. That means a relatively slow performance. At the same time the problem of obtaining sequences of equiprobable PR binary sets for arbitrary values of the number of different output vectors n has considerable both theoretical and practical interest.

Considering abovementioned conditions it is possible to formulate the purpose of this article as solving the problem of theoretical substantiation of the possibility and analysis of the features of organizing the structural formation of pseudo-random sequences of equiprobable binary sets by using sequential dichotomic decomposition procedures. It is assumed that there are no restrictions on the number of different binary sets in generated sequence which means that this number can be arbitrary. For example, it can be determined only by specific number of processor elements in a fault-safe multiprocessor system when simulating the effect of possible failures in such systems and is in no way not delimited by power of 2.

2. Decomposition Procedure

Proposed method of obtaining a given number of PR vectors is based on a sequential dichotomic division of an integer positive number n (we consider the numbers between powers of 2) on a series of add-ons on the basis of relation

$$n=2q+r,$$

where q is the quotient of the number n by 2, r is the remainder of division. Obviously r = 0 if n - even and r = 1 for odd n. Or else:

$$\begin{cases} q = n/2, \text{ if } r = 0, \\ q = (n-1)/2, \text{ if } r = 1. \end{cases}$$

In other words, for each possible value n> 1, two positive integers n0 and ne can be put in correspondence, such as:

Manuscript received February 5, 2019
Manuscript revised February 20, 2019

$$\begin{cases} n = n_0 + n_e, \\ n_e = \dfrac{n-1}{2} + 1 = \dfrac{n+1}{2}; \text{if } \ r = 1, \\ n_0 = \dfrac{n-1}{2}; \text{if } \ r = 1, \\ n_0 = n_e = n/2; \text{if } \ r = 0. \end{cases}$$

If the process of division by 2 is continued for the obtained values of n0 and n_e, then in the end result the values q=1 will be achieved. Obviously, this requires $m=\log_2 n$ cycles. Note that in fact $m=[\log_2 n]$, where $[\log_2 n]$ means an integer positive number, such as:

$$m > \log_2 n > m-1$$

Examples of the results of such sequential decomposing for numbers 17, 20 and 31 are given in the table 1.

Table 1:

Step i	max (n0, ne)		
0	31	20	17
1	16	10	9
2	8	5	5
3	4	3	3
4	2	2	2
5	1	1	1

In other words if after every (any) division step i=1,2,...m we get a pair of numbers n_{0i} and n_{ei} for the desired number n_{i-1} (divisible), then:

$$\begin{cases} n_{0i} = n_{ei} = \dfrac{n_{i-1}}{2}; \text{if } \ n_{i-1} - \text{even}, \\ n_{0i} = \dfrac{n_{i-1}+1}{2},\ n_{ei} = \dfrac{n_{i-1}-1}{2}; \text{if } \ n_{i-1} - \text{odd}. \end{cases}$$

3. Features of Decomposition Tree

To simplify on the one hand and generalize on the other, the decomposition procedure for a number a1 can be represented as a tree structure (Fig. 1). Let's denote by j=1,2,...7 indexes depicting this structure's members.

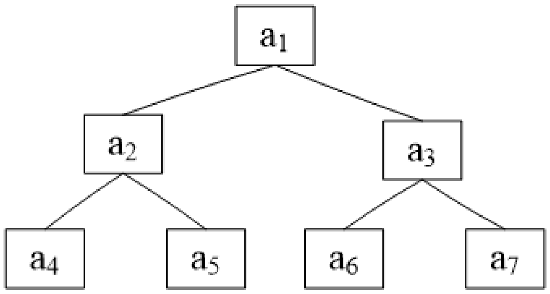

Fig. 1

Decomposition conditions can be written as:

$$[(a_1) \bmod 2 = 0] \rightarrow \left(a_2 = a_3 = \frac{a_1}{2}\right),$$
$$[(a_1) \bmod 2 = 1] \rightarrow \left[\left(a_2 = \frac{a_1+1}{2}, a_3 = \frac{a_1-1}{2}\right) \vee \left(a_2 = \frac{a_1-1}{2}, a_3 = \frac{a_1+1}{2}\right)\right].$$

It is enough to verify that the difference $\Delta = |a_2 - a_3|$ belongs to set {0,1} for j=2,3..., which means $\Delta = |a_2 - a_3|$.

For the definiteness:

$$\max(a_2, a_3) = \frac{a_1+1}{2} \ if \ a_1 - \text{odd},$$
$$\min(a_2, a_3) = \frac{a_1-1}{2} \ if \ \ a_1 - \text{odd}.$$

In this case

$$\max(a_4, a_5, a_6, a_7) = \frac{\max(a_2, a_3) + r_1}{2},$$
$$\min(a_4, a_5, a_6, a_7) = \frac{\min(a_2, a_3) + r_2}{2},$$

where r1(2)=0(1) is the sign of parity of corresponding max and min values. Therefore:

$$\forall j = 4,5,6,7 \{\Delta i_{\max} = \left|\max(a_4,\ldots a_7) - \min(a_4,\ldots a_7)\right| = \left| \frac{\frac{a_1+1+r_1}{2}}{2} - \frac{\frac{a_1-1+r_2}{2}}{2} \right| =$$
$$= \left|\frac{a1+1+2r1-a1+1+2r2}{4}\right| = \left|\frac{r1+r2+1}{2}\right| = 1\},$$

because (r_1=1)→(r_2=0) or (r_2=1) → (r_1=0).

When the structure is expanded by adding branches, there is preserved condition that the difference of any two decomposition elements belonging to one level does not exceed a value of 1.

So in general case we can formulate following suggestion.

Statement 1. For any step of decomposition i=1,2,...m, the value

$$\Delta q_{i\max} = \left| q_{i\max} - q_{i\min} \right| \leq 1,$$

which means that for any pair of quotients of the same level $\Delta qi \in \{0,1\}$.

Proof. At the first stage of the decomposition n=2q1+r1, q1 – is a quotient, r1 – remainder, $r_1 \in \{0,1\}$.

Consequently, $q_1 = \frac{(n-r_1)}{2}$, which means $q_1 \in \left\{\frac{n}{2}, \frac{n-1}{2}\right\}$.

For the second stage q1=2q2+r2, wherein $q_2 = \frac{q_1 - r_2}{2} = \frac{n-r_1}{4} - \frac{r_2}{2}$.

For the third stage (q2=2q3+r3):

$$q_3 = \frac{n-r_1}{8} - \frac{r_2}{4} - \frac{r_3}{2}.$$

Similarly for an arbitrary stage i we can write:

$$q_i = \frac{n-r_1}{2^i} - \frac{r_2}{2^{i-1}} - \frac{r_3}{2^{i-2}} - \ldots - \frac{r_i - 1}{4} - \frac{r_i}{2}.$$

In this case three options are possible (two limiting and one intermediate).

a) All r_i=0, which means n and intermediate dividends are even numbers or n=2^i·k, k - positive integer. Then it follows that:

$$\forall i = 1, 2, \ldots m \left(\Delta q_i = \left| k - k \right| = 0 \right).$$

b) All r_i=1, wherein n and obtained dividends are odd, that is:

$$q_i = \frac{n-1}{2^i} - \frac{1}{2^{i-1}} - \frac{1}{2^{i-2}} - \ldots - \frac{1}{4} - \frac{1}{2}.$$

Therefore:

$$\forall i = 1, 2, \ldots m \left(\Delta q_i = \left| q_i - q_i \right| = 0 \right).$$

c) In this (intermediate variants between a) and b)) case $q_{i\max} > q_i > q_{i\min}$,

however, since $q_{i\max} = \frac{n}{2^i}$, and , then

$$\Delta q_i < \Delta q_{i\min} = \left| q_{i\max} - q_{i\min} \right| = \frac{n}{2^i} - \frac{n}{2^i} + \frac{1}{2^i} + \frac{1}{2^{i-1}} + \ldots + \frac{1}{2} + 1.$$

Thus, the value of Δq_i ranges $2 > \Delta q_i \geq 0$ and as Δq_i - a positive integer then $\Delta qi \in \{0,1\}$. Therefore statement 1 is proved, i.e.

$$\forall i = 1, 2, \ldots m \left(\Delta q_i \leq 1 \right).$$

Return to the tree shown in pic. 1. Let us prove next.

Statement 2. At each step i of decomposition a sum of components belonging to subset 2i is constant and equal to n, that is:

$$\forall i = 1, 2, \ldots m \left(\sum_{j=1}^{2^i} a_{ij} = n \right).$$

Indeed, for the first division step: $q_1 = \frac{n-r_1}{2}$, which means:

$$(r_1 = 0) \rightarrow \left(q_1 = \frac{n}{2}, a_1 = a_2 = \frac{n}{2} \right),$$

$$(r_1 = 1) \rightarrow \left(q_1 = \frac{n-1}{2} \right).$$

From last condition it follows that $a_1 = \frac{n-1}{2}$ and $a_2 = \frac{n+1}{2}$ or in the general case any term of decomposition can be represented as the relation $\frac{n \pm r_1}{2}$.

Therefore, a_1+a_2=n.

For the second decomposition step result of division of a_1 belongs to set {a_3, a_4}, and result of dividing a_2 - to the set {a_5, a_6}.

Then:

$$a_3 = \frac{a_1 + r_2}{2}, a_4 = \frac{a_1 - r_2}{2}, a_5 = \frac{a_2 + r_3}{2}, a_6 = \frac{a_2 - r_3}{2}.$$

Sum of these terms of decomposition for i = 2 is:

$$\sum_{k=3}^{6} a_k = \frac{a_1 + r_2 + a_1 - r_2 + a_2 + r_3 + a_2 - r_3}{2} = \frac{2(a_1 + a_3)}{2} = n$$

For an arbitrary step i, we find sum Si as: $S_i = \sum_{k=1}^{2^i} a_{ik}$. Under conditions of the division $\forall k = 1, 2, ...2^i \left(a_{ik} + a_{i(k+1)} = a_{(i-1)k} \right)$, and besides $a_{ik} = \frac{a_{(i-1)k} + r_k}{2},\ a_{i(k+1)} = \frac{a_{(i-1)k} - r_k}{2}.$ It was shown above that a1+a2=a11+a12=n, from which, using the well-known induction rules, we can deduce that:

$$\sum_{k=1}^{2^1} a_{1k} = \sum_{k=1}^{2^2} a_{2k} = \sum_{k=1}^{2^3} a_{3k} = ... = \sum_{k=1}^{2^i} a_{ik} = n$$

, so the validity of statement 2 is proved.

4. Probabilities of paths in the graph

Let's assume that the positive integer n is decomposed into the components at the levels which are corresponding to levels of the tree and each j-th number of a_{ij} level i, i = 1,2, ..m^{-1}, j = 1,2, ... 2^i is the sum:

$$a_{ij} = a_{(i+1)k} + a_{(i+1)(k+1)} ,$$

where k=2j+1, and besides:

$$a_{(i+1)k} = \frac{a_{ij} - r_{ij}}{2},\ a_{(i+1)(k+1)} = \frac{a_{ij} + r_{ij}}{2} ,$$

where r_{ij} = res(a_{ij})mod2 is an attribute of parity of the number a_{ij}.

When decomposition tree is built it can be assumed that at each level i = 1,2, ... m^{-1} it is carried out a random (pseudo-random) selection of one of the level's decomposition members i + 1, as it is shown at Fig. 2. Such selection will be needed of an external probabilistic source. Wherein probability of choosing the number a_{ij} after the level i-1 is p(a_{ij}) so probabilities of choosing each of the corresponding components of number a_{ij} (after step i) are p_1=p($a_{(i+1)k}$) and p2=p($a_{(i+1)(k+1)}$).

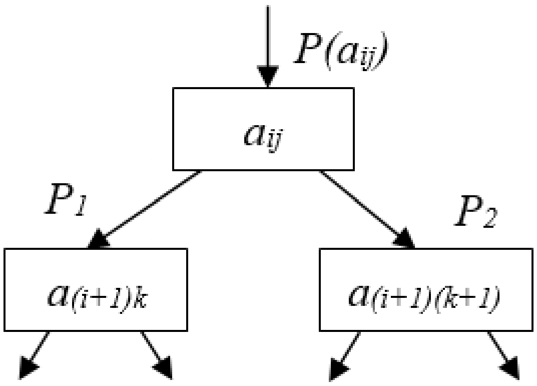

Fig. 2

Consider set of paths from a vertex $a_{(i+1)k}$ of the tree graph to the end "1" vertices as it shown on Fig.3. Obviously for intermediate terms $a_{(i+1)k}$ there are $a_{(i+1)k}$ paths l(q), q=1,2,… $a_{(i+1)k}$ in the tree graph. Let us prove next statement.

Statement 3. To fulfill the requirements of equal probability 1/n of all n tree graph paths starting from (n) to the end (1) vertices next necessary and sufficient conditions should be satisfied:

$$\forall i = 1, 2, ..., m-1,\ j = 1, 2, ...2^i \left\{ P\left[a_{(i+1)k} \right] = \frac{a_{(i+1)k}}{a_{ij}} \right\},$$

where k=2j+1.

Proof. Analyzing the structure on pic. 3, we can see that in order to fulfill requirement of equal probability of paths, the following ratio must be true:

$$\forall q, i, j \left\{ P\left[l(q) \right] \cdot P_1 \cdot P(a_{ij}) = \frac{1}{n} \right\},$$

According to formulation of statement 3 for any pair of paths q and s we can write:

$$\forall q, s = 1, 2...a_{(i+1)k},\ q \neq s \left\{ P\left[l(q) \right] \cdot P_1 \cdot P(a_{ij}) = P\left[l(s) \right] \cdot P_1 \cdot P(a_{ij}) \right\} ,$$

which means that

$$\forall q, s, q \neq s \left\{ P\left[l(s) \right] = P\left[l(q) \right] \right\}.$$

Continuing analysis of the structure we can note that the number of different paths starting from vertex $a_{(i+1)k}$ is exactly equal to the value of component $a_{(i+1)k}$. However, since:

$$\sum_{d=1}^{a_{(i+1)k}} P\left[l(d)\right]=1,$$

consequently therefore

$$\forall q=1,2...a_{(i+1)k}\left\{a_{(i+1)k}\cdot P\left[l(q)\right]=1\right\}.$$

In other words following condition is true:

$$\forall q=1,2...a_{(i+1)k}\left\{P\left[l(q)\right]=\frac{1}{a_{(i+1)k}}\right\}\cdot$$

Doing in the same way as we did for the value of a(i+1)k, it can also be stated that for vertex a(i+1)(k+1) corresponding probability of selection (transition) will be equal to

$$\forall s\left\{P\left[l(s)\right]=\frac{1}{a_{(i+1)(k+1)}}\right\}.$$

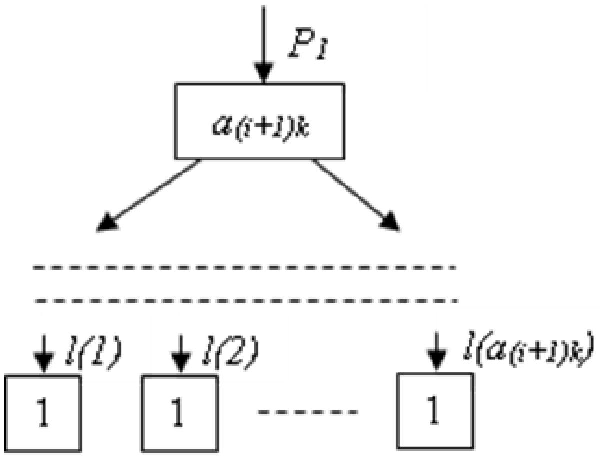

Fig. 3

Obviously, the requirement of equal probability 1 /n for all n possible paths leads to necessity of fulfilling the relation

$$\forall d=1,2,...a_{(i+1)k,}\ f=1,2,...a_{(i+1)(k+1)}\left\{P_1\cdot P\left[l(d)\right]=P_2\cdot P\left[l(f)\right]\right\}\cdot$$

But from this it follows that

$$\frac{P_1}{a_{(i+1)k}}=\frac{P_2}{a_{(i+1)(k+1)}}.$$

Taking into account that P1 + P2 = 1 we can write:

$$\frac{P_1}{a_{(i+1)k}}=\frac{1-P_1}{a_{(i+1)(k+1)}},$$

or

P1a(i+1)k=(1-P1(·a(i+1)(k+1) , which means that a(i+1)k = P1 [$a_{(i+1)k}$ + $a_{(i+1)(k+1)}$].

In accordance with adopted method of dichotomic decomposition of numbers into terms, the following condition is fulfilled:

$$\forall i,j,k\left[a_{(i+1)k}+a_{(i+1)(k+1)}=a_{ij}\right],$$

from which it follows that:

$$\forall i,j,k\left\{P_1=P\left[a_{(i+1)k}\right]=\frac{a_{(i+1)k}}{a_{ij}}\right\}\cdot$$

Thus, validity of statement 3 is proved.

By analogy and according to statement 3, we can verify that:

$$\forall i,j,k\left\{P_2=1-P_1=P\left[a_{(i+1)(k+1)}\right]=\frac{a_{(i+1)(k+1)}}{a_{ij}}\right\}\cdot$$

Let's select a certain path in tree-like graph which begins from initial vertex a_0=n and ends on some finite vertex a_m=1 (pic. 4).

Taking into account the fact that m=]log$_2$n[, a_m=1 and a_0 = n, so in accordance with statement 3, we can write that:

$$\forall i=1,2,...m\left(P_i=\frac{a_i}{a_{i-1}}\right)\cdot$$

Probability P of passage through each particular path can be found from relation

$$P=\prod_{i=1}^{m}P_i=\frac{a_1}{a_0}\cdot\frac{a_2}{a_1}\cdot\frac{a_3}{a_2}\cdot\frac{a_4}{a_3}\cdot...\cdot\frac{a_{m-1}}{a_{m-2}}\cdot\frac{a_m}{a_{m-1}}=\frac{a_m}{a_0}=\frac{1}{n}\cdot$$

If on each of the m=]log$_2$n[levels of the dichotomic graph we form the value of some bit in accordance with the probabilities that were calculated for particular level, then as a result we obtain a binary set of dimension m, that belongs to set n, and, as it was shown above, all the components of such set will have an equal probability of 1 / n.

5. Conclusion

Method for analyzing the components of a sequential dichotomic decomposition of integers was considered. The method is supposed to be used in procedures of structural synthesis of sequence generators of pseudo-random equiprobable binary sets which belong to certain set of arbitrary number of components. Also this method can be applied to construction of high-speed hardware binary signal generators with variable (programmable) probability of one (zero) values [5-7]. A certain value is represented by a number of parity bits of decomposition components - for specifying and minimizing Boolean parity functions for synthesis of their realization circuit. The same parity features are proposed to be used to control the structural means for multiplexing both binary signals with equal probability and signals with given probabilities at the outputs of external generators. In this case the multiplexing elements set output bit states of equiprobable binary sets.

Literature

[1] Gill A. Linear sequential circuits / Arthur Gill. 1966, New York: McGraw-Hill.

[2] Ярмолик В.Н. Проектирование самотестируемых СБИС / Ярмолик В.Н. и др. - Т.1.- Минск – БУИР - 2001.- 159 с.

[3] Agrawal P. Agrawal V.D. Probabilistic analysis of random test generation method for irredundant combinational logic networks. // IEEE Trans. on Comput.- 1975.- Vol. C-24.- № 7.-P. 681-695.

[4] Chaowen Yu, Sudhakar M. Reddy, Irith Pomeranz Circuit Independent Weighted Pseudo-Random BIST Pattern Generator. // 14th Asian test Symposium (ATS'05). December 2005, pp. 132-137.

[5] Rabah AlShboul, Vitaliy A. Romankevich. Structural Means Generating Pseudorandom Sequences Of Fixed Weight Binary Patterns // IJCSNS International Journal of Computer Science and Network Security. – 2017. – Vol. 17, No.10. – P. 62 – 66.

[6] A. Fúster-Sabater, P. Caballero-Gil. Linear solutions for cryptographic nonlinear sequence generators // Physics Letters A Vol. 369, Is. 5-6, 1 Oct. 2007, pp. 432-437

[7] B. Barak and S. Halevi. An architecture for robust pseudo-random generation and Applications to /dev/random // In ACM, editor, Proc. Computing and Communication Security (CCS).- 2005.- pp. 179-188.