On the Delta Sequence of the Thue-Morse Sequence

T.W. Cusick[a], H. Fredricksen[b], P. Stănică[b*]

[a]Department of Mathematics, The State University of New York
Buffalo, NY 14260-2900, USA; cusick@buffalo.edu
[b]Department of Applied Mathematics, Naval Postgraduate School
Monterey, CA 93943, USA; {half,pstanica}@nps.edu

February 27, 2007

Abstract

In this note, we investigate the delta sequence associated to the classical Thue-Morse sequence and prove a conjecture about the delta sequence. Further, we generalize the Thue-Morse sequence and show some results about this new sequence and its associated delta sequence.

1 Motivation and Some Definitions

The *Thue-Morse (TM) sequence* $T = (t_n)_{n\geq 0}$ is defined as the limit of iterates $\varphi^n(0)$, where the map φ is defined by $\varphi(0) = 01$, $\varphi(1) = 10$. We denote the 2^n-length initial segment of the TM sequence by T_{2^n}. Furthermore, the TM sequence can also be generated by:

$$
\begin{aligned}
&T_1 = t_0 = 0,\\
&T_{2^n} = T_{2^{n-1}}\overline{T_{2^{n-1}}},\ n \geq 1.
\end{aligned}
$$

or

$$
\begin{aligned}
&T_1 = t_0 = 0,\\
&T_{2^n} = T_{2^{n-1}} r(\overline{T_{2^{n-1}}}),\ \text{for } n \text{ odd}.\\
&T_{2^n} = T_{2^{n-1}} r(T_{2^{n-1}}),\ \text{for } n \text{ even},
\end{aligned}
$$

*Research supported by the Naval Postgraduate School RIP funding.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **27 FEB 2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE
On the Delta Sequence of the Thue-Morse Sequence

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School,Department of Applied Mathematics,Monterey,CA,93943

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Australasian Journal of Combinatorics 39 (2007), 293-300.

14. ABSTRACT
In this note, we investigate the delta sequence associated to the classical Thue-Morse sequence and prove a conjecture about the delta sequence. Further, we generalize the Thue-Morse sequence and show some results about this new sequence and its associated delta sequence.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**9**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

where $r(\cdot)$ is the map that reverses the bits of the argument, and $\overline{B}$ is the complement of B. Moreover, the TM sequence can also be generated by using the bit expansion of the position, that is,

$$\text{if } i = \sum_j b_j 2^j, \text{ then } t_i = \sum_j b_j \pmod 2, \tag{1}$$

that is, $T = (t_n)_{n\geq 0}$ counts the number of 1's (mod 2) in the base-2 representation of n. The first few terms of the Thue-Morse sequence are

$$T = 0110100110010110100 10\cdots$$

x_1	x_2	x_3	x_4	f
0	0	0	0	0
0	0	0	1	1
0	0	1	0	1
0	0	1	1	1
0	1	0	0	0
0	1	0	1	0
0	1	1	0	1
0	1	1	1	1
1	0	0	0	0
1	0	0	1	1
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	1
1	1	1	0	1
1	1	1	1	1

Table 1: Truth table of a Boolean function

Let $\mathbb{F}_2^n$ be the vector space of dimension n over the two element field $\mathbb{F}_2$. Let us denote the addition operator over $\mathbb{F}_2$ by $\oplus$, and the direct product by "$\cdot$". The vectors consisting of all 1, respectively, all 0 (of some length) are denoted by $\mathbf{1}$, respectively, $\mathbf{0}$. By abuse of notation, when there is no danger of confusion, we sometimes use $\mathbf{1}, \mathbf{0}$ to denote a binary string consisting of all 1, respectively, all 0. A Boolean function on n variables may be viewed as a mapping from $\mathbb{F}_2^n$ into $\mathbb{F}_2$. We order $\mathbb{F}_2^n$ lexicographically, and denote $\mathbf{v}_0 = (0, \ldots, 0, 0)$, $\mathbf{v}_1 = (0, \ldots, 0, 1)$, $\mathbf{v}_{2^n-1} = (1, \ldots, 1, 1)$. We interpret a Boolean function $f(x_1, \ldots, x_n)$ as the output column of its *truth table*, i.e., a binary string of length 2^n, $f = [f(\mathbf{v}_0),\ f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_{2^n-1})]$. In Table 1 we present the truth table of a 4-variable Boolean function.

The novelty of our work consists of the Boolean functions approach on the TM sequence, which enables us to resolve several questions on the TM sequence. We do not claim that some of our results cannot be obtained by working with the sequence directly, however our approach is elegant and brings into play the powerful tool of Boolean functions.

2 Delta sequence of the TM-sequence

We define

$$\begin{aligned} S = \{A = 0,0,1,1;\ \bar{A} = 1,1,0,0;\ B = 0,1,0,1; \bar{B} = 1,0,1,0; \\ C = 0,1,1,0;\ \bar{C} = 1,0,0,1;\ D = 0,0,0,0; \bar{D} = 1,1,1,1\}. \end{aligned} \tag{2}$$

Our Theorem 1 will give an alternate definition for the TM sequence, and it can be deduced from the generation algorithm (1) and the following lemma.

Lemma 2.1. (Folklore Lemma [8, Lemma 3.7.2]) *Any affine function $f = [t_1, \ldots, t_{2^n}]$ on n variables, $n \geq 2$, is a linear string of length 2^n made up of 4-bit blocks $I_1, \ldots, I_{2^{n-2}}$ given as follows:*

1. *The first block I_1 is one of $A, B, C, D, \bar{A}, \bar{B}, \bar{C}$ or $\bar{D}$.*
2. *The second block I_2 is I_1 or $\bar{I}_1$.*
3. *The next two blocks I_3, I_4 are I_1, I_2 or $\bar{I}_1$, $\bar{I}_2$.*

....

$n-1$. *The 2^{n-3} blocks $I_{2^{n-3}+1}, \ldots, I_{2^{n-2}}$ are $I_1, \ldots, I_{2^{n-3}}$ or $\bar{I}_1, ..., \bar{I}_{2^{n-3}}$.*

Theorem 1. *The initial segment of length 2^n, $n \geq 2$, of the TM sequence is the truth table of the Boolean function*

$$f(x_1, x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \cdots \oplus x_n,$$

defined on $\mathbb{F}_2^n$ (ordered lexicographically).

Proof. By the Folklore Lemma it is easy to see that $x_1 \oplus \cdots \oplus x_n = C\bar{C} \cdots$, which is exactly the initial segment of length 2^n of the TM sequence. □

In [4] the following *delta-j sequence* (we will call it *delta sequence*, if j is understood from the context) is associated to the TM sequence: For $j \geq 1$, we define

$$\delta_i^{(j)} = t_i \oplus t_{i+j}.$$

Various results were proved in [4] by working with the delta-sequence, in particular it was proved that T has the *nonoverlap* property (also known as the BBb property), that is, the subsequence BBb, where B is a block of bits of any > 0 length, and b is the first bit of B, does not appear in the TM sequence. The nonoverlap property was originally proved by Thue in his seminal papers from 1906 and 1912 [9, 10]. The result has been rediscovered in [6] and other places (see [2, 3] for surveys of results on the TM sequence).

It is interesting to note that, independently, in [1], a different approach was taken, which arrives to the same delta sequence. Kimberling proposed a problem in American Mathematical Monthly on the sequence $\mathbf{c} = (c_k)_k$ defined by

$$c_0 = 1;\ c_{k+1} = \begin{cases} c_k + 1 \ if\ (c_k+1)/2 \notin \mathbf{c} \\ c_k + 2 \ otherwise. \end{cases}$$

Later, Plouffe and Zimmermann [7] proposed the following problem (which was found by a method that goes back to Euler):

$$\sum_{k\geq 0} c_k x^k = \frac{1}{1-x}\prod_{j\geq 1}(1+x^{e_j}),$$

where $\mathbf{c}$ is the sequence of Kimberling and $\mathbf{e} = (e_j)_j$ is defined by

$$e_1 = 1;\ e_{j+1} = \begin{cases} 2e_j + 1 \ if\ j\ is\ even \\ 2e_j - 1 \ if\ j\ is\ odd. \end{cases}$$

The conjecture was proven in [1] by a method that uses the ever-present TM sequence. Furthermore, if one defines the characteristic function of c,

$$\chi(k) = \begin{cases} 1 \ if\ k \in \mathbf{c} \\ 0 \ otherwise, \end{cases}$$

then one can show [1, Lemma 3] that

$$\chi(k) = t_k \oplus t_{k-1},$$

that is, $\chi(k)$ is the same as the delta-1 sequence $\delta^{(1)}_{k-1}$.

Fredricksen, in [4], proved that $\delta^{(1)}_k$ is 1 if and only if $k+1 = (1+2\ell)2^{2j}$, for some integers ℓ, j, and Proposition 1 and Lemma 1 of [1] state similar results about $\chi(k)$. Fredricksen proved that $\delta^{(2)}_k$ is the dilated by 2 sequence of $\delta^{(1)}_k$, and observed that $\delta^{(4)}_k$ is the dilated by 4 sequence of $\delta^{(1)}_k$. For example, $\delta^{(1)}_k = 101110\ldots$ and $\delta^{(2)}_k = 110011111100\ldots$, that is, $\delta^{(2)}_k$ contains twice every bit of $\delta^{(1)}_k$.

Consequently, he proposed a conjecture, which we prove in our main result of this section.

Theorem 2. *The delta sequence $\delta^{(2j)}$ is the dilated by two sequence of the delta sequence $\delta^{(j)}$.*

Proof. To prove the claim it is sufficient (and necessary) to show that

$$\delta_{2i}^{(2j)} = \delta_{2i+1}^{(2j)} = \delta_i^{(j)}. \tag{3}$$

Let f denote the linear function in Theorem 1 for some fixed n. It is sufficient to show

$$f(\mathbf{v}_{2i}) \oplus f(\mathbf{v}_{2i+2j}) = f(\mathbf{v}_{2i+1}) \oplus f(\mathbf{v}_{2i+2j+1}) = f(\mathbf{v}_i) \oplus f(\mathbf{v}_{i+j}), \tag{4}$$

for this function, since then (3) follows for all i and j by letting n tend to infinity. Observe that $v_{2\ell+1} = v_{2\ell} \oplus v_1$ (there is no carry). Since f is linear, we obtain

$$\begin{aligned}
&f(\mathbf{v}_{2i+1}) \oplus f(\mathbf{v}_{2i+2j+1}) \\
&= f(\mathbf{v}_{2i} \oplus \mathbf{v}_1) \oplus f(\mathbf{v}_{2i+2j} \oplus \mathbf{v}_1) \\
&= f(\mathbf{v}_{2i}) \oplus f(\mathbf{v}_1) \oplus f(\mathbf{v}_{2i+2j}) \oplus f(\mathbf{v}_1) \\
&= f(\mathbf{v}_{2i}) \oplus f(\mathbf{v}_{2i+2j}).
\end{aligned}$$

We are left with checking that $f(\mathbf{v}_{2i}) \oplus f(\mathbf{v}_{2i+2j}) = f(\mathbf{v}_i) \oplus f(\mathbf{v}_{i+j})$. We prove the latest claim, by showing that

$$f(\mathbf{v}_{2\ell}) = f(\mathbf{v}_\ell), \tag{5}$$

for any ℓ, in particular, for $\ell = i$, and $\ell = i + j$. Equation (5) follows from the observation that $\mathbf{v}_{2\ell}$ is obtained from $\mathbf{v}_\ell$ by moving the leftmost 0 bit to the rightmost location of the string. That is, the Hamming weight of $\mathbf{v}_{2\ell}$ is the same as the Hamming weight of $\mathbf{v}_\ell$, which implies, again using Theorem 1 that $f(\mathbf{v}_{2\ell}) = f(\mathbf{v}_\ell)$. The theorem is proved. □

See the remark after Theorem 4 for an alternative approach to infer the truth of Theorem 1.

3 Generalized Thue-Morse sequences

Let $\epsilon := \epsilon_1 \epsilon_2 \cdots$ be a sequence of $\epsilon_i \in \{0, 1\}$ bits (possibly infinite). Define a function r_{ϵ_i} on arbitrary bit-blocks B, in the following way:

$$r_{\epsilon_i}(B) = \begin{cases} B \ if \ \epsilon_i = 0 \\ \overline{B} \ if \ \epsilon_i = 1. \end{cases} \tag{6}$$

We introduce the *generalized Thue-Morse* sequence $T^\epsilon = (t_n^\epsilon)_{n\geq 0}$ (we call it the ϵ-*TM sequence*) by the following algorithm ($T^\epsilon_{2^i}$ is the binary string made up of the first 2^i bits of T^ϵ):

$$\begin{aligned} T_1^\epsilon &= t_0 \in \{0,1\} \\ T^\epsilon_{2^i} &= T^\epsilon_{2^{i-1}}\, r_{\epsilon_i}(T^\epsilon_{2^{i-1}}) \end{aligned} \tag{7}$$

The classical Thue-Morse sequence is T^ϵ, where $\epsilon = 11\cdots$.

Theorem 3. *Given an initial segment T_{2^n} of length 2^n of a generalized Thue-Morse sequence, there exists an affine Boolean function f (if $t_0 = 0$, then f is linear) on n variables, such that T_{2^n} is the truth table of f.*

Proof. First, assume $t_0 = 0$. Then the choices $\epsilon_1\epsilon_2 = 01, 10, 11, 00$ give, respectively, the initial segments A, B, C, D of length 4. Now by the Folklore Lemma the resulting generalized Thue-Morse sequences all have their initial segments T_{2^n} given by the corresponding initial segments of some linear function. If $t_0 = 1$, then the same argument leads to an affine function f with $f(\mathbf{0}) = 1$. $\square$

We call such a sequence T_{2^n} as in Theorem 3, the TM-sequence *associated* to f, and the Boolean function f – sometimes, labeled f_T – is the *companion* of T_{2^n}.

Define the *delta*-(ϵ, j) *sequence* associated to T^ϵ, in the same way as before, that is,

$$\delta_i^{(\epsilon,j)} = t_i^\epsilon \oplus t_{i+j}^\epsilon.$$

Let $\epsilon := \epsilon_1\epsilon_2\cdots$ be an infinite bit string.

Theorem 4. *The delta sequence $\delta^{(\epsilon,2j)}$ satisfies*

$$\delta_{2i}^{(\epsilon,2j)} = \delta_{2i+1}^{(\epsilon,2j)}, \tag{8}$$

for any i, j. In general, if $f(x_1, \ldots, x_n) = x_{i_1}\oplus\cdots\oplus x_{i_k}$, then the delta sequence associated to f satisfies

$$\delta_i^{(\epsilon,j)} = 1 \text{ if and only if } wt(\pi_{i_1,\ldots,i_k}(v_i))\oplus wt(\pi_{i_1,\ldots,i_k}(v_{i+j})) = 1,\ \forall i, j$$

where $\pi_{i_1,\ldots,i_k}(\mathbf{v})$ is the length-k projection on the coordinates $i_1, \ldots, i_k$ of the vector $\mathbf{v}$.

Proof. Without loss of generality we may assume that $t_0 = 1$. Suppose that $f_T(x_1, \ldots, x_n) = x_{i_1} \oplus \cdots \oplus x_{i_k}$ is the companion of the initial segment T_{2^n} of $\delta^{(\epsilon,j)}$. To prove (8) it suffices to show

$$f_T(v_{2i}) \oplus f_T(v_{2i+2j}) = f_T(v_{2i+1}) \oplus f_T(v_{2i+2j+1}), \tag{9}$$

Since f_T is linear, this follows by the argument used in the proof of Theorem 2. To prove (9), we use the facts that

$$\delta_i^{(\epsilon,j)} = 1 \text{ if and only if } f_T(v_i) \oplus f_T(v_{i+j}) = 1$$

(from Theorem 3) and

$$f_T(v_i) = wt(\pi_{i_1,\ldots,i_k}(v_i)) \pmod 2$$

(from the form of the linear function f_T). $\square$

An alternative approach, suggested by one reader of the paper is to observe that $t_{2n} \equiv t_n \pmod 2$, and $t_{2n+1} \equiv 1 + t_n \pmod 2$, and so,

$$\begin{aligned} \delta_{2i}^{2j} &\equiv t_{2i} + t_{2i+2j} \equiv t_i + t_{i+j} \equiv \delta_i^j \pmod 2, \\ \delta_{2i+1}^{2j} &\equiv t_{2i+1} + t_{2i+2j+1} \equiv (1 + t_i) + (1 + t_{i+j}) \\ &\equiv t_i + t_{i+j} \equiv \delta_i^j \pmod 2. \end{aligned}$$

A further analysis of the binary expansion of n, say $n = \sum_{k \geq 0} e_k(n) 2^n$, implies

$$t_n^\epsilon = t_0 + \sum_{k \geq 0} e_k(n) \epsilon_{k+1} \pmod 2,$$

(using induction on N that the relation is true for all $n \in [0, 2^N)$, using some properties of the Thue-Morse sequence). Now, for $k \geq 1$, we have $e_k(2n) = e_k(2n+1) (= e_{k-1}(n))$, $e_0(2n) = 0$ and $e_0(2n+1) = 1$. We can now get $\delta_{2i}^{(\epsilon,2j)} = \delta_{2i+1}^{(\epsilon,2j)}$.

Next, we shall investigate the nonoverlap property of the generalized Thue-Morse sequence, and prove our main result of this section.

Theorem 5. *The ϵ-TM sequence satisfies the nonoverlap property if and only if $\epsilon = \mathbf{1}$.*

Proof. If the sequence ϵ is not identically 1, it contains a 0, hence either the block 00 or one of the blocks 011 or 010. It thus suffices to show that, if the sequence ϵ contains one of the blocks 00, 011, 010, then the corresponding ϵ-TM sequence contains an overlap. For easy writing, we denote by $B_i := T^{\epsilon}_{2^i}$, $i \geq 0$ (obviously, $B_0 = \{t_0\}$). We distinguish the following cases.
Case (i) If $\epsilon_i = 0$ and $\epsilon_{i+1} = 0$, for some $i \geq 1$, then

$$\begin{aligned} B_{i+1} &= B_i r_{\epsilon_{i+1}}(B_i) = B_i B_i \\ &= B_{i-1} r_{\epsilon_i}(B_{i-1}) B_{i-1} r_{\epsilon_i}(B_{i-1}) \\ &= B_{i-1} B_{i-1} B_{i-1} B_{i-1}, \end{aligned}$$

which contains the cube $B_{i-1}B_{i-1}B_{i-1}$, hence an overlap.
Case (ii) If $\epsilon_i = 0$, $\epsilon_{i+1} = 1$, and $\epsilon_{i+2} = 1$, for some $i \geq 1$, then

$$\begin{aligned} B_{i+2} &= B_{i+1}\overline{B_{i+1}} = B_i \overline{B_i}\,\overline{B_i} B_i \\ &= B_{i-1}B_{i-1}\overline{B_{i-1}}\,\overline{B_{i-1}}\,\overline{B_{i-1}}\,\overline{B_{i-1}}B_{i-1}B_{i-1}, \end{aligned}$$

which contains the cube $\overline{B_{i-1}}\,\overline{B_{i-1}}\,\overline{B_{i-1}}$, hence an overlap.
Case (iii) If $\epsilon_i = 0$, $\epsilon_{i+1} = 1$, and $\epsilon_{i+2} = 0$, for some $i \geq 1$, the next bit is either $\epsilon_{i+3} = 0$ and then (i) shows that there is an overlap in B_{i+3}, or $\epsilon_{i+3} = 1$, in which case

$$\begin{aligned} B_{i+3} &= B_{i+2}\overline{B_{i+2}} = B_{i+1}B_{i+1}\overline{B_{i+1}}\,\overline{B_{i+1}} \\ &= B_i\overline{B_i}B_i\overline{B_i}\,\overline{B_i}B_i\overline{B_i}B_i \\ &= B_{i-1}B_{i-1}\overline{B_{i-1}}\,\overline{B_{i-1}}\,B_{i-1}B_{i-1}\overline{B_{i-1}}\,\overline{B_{i-1}} \\ &\quad \overline{B_{i-1}}\,\overline{B_{i-1}}B_{i-1}B_{i-1}\overline{B_{i-1}}\,\overline{B_{i-1}}B_{i-1}B_{i-1} \end{aligned}$$

which contains the cube $\overline{B_{i-1}}\,\overline{B_{i-1}}\,\overline{B_{i-1}}$, hence an overlap.
The theorem is proved. □

It would be an interesting problem to investigate what patterns are avoided in the ϵ-TM sequence.

Acknowledgments. The authors would like to acknowledge the very careful and constructive comments of the referee, in particular the short approach for the proof of our last theorem. H. F. acknowledges support from the National Security Agency under contract RMA54. Research of P. S. was supported in part by a RIP grant from Naval Postgraduate School.

References

[1] J.-P. Allouche, A. Arnold, J. Berstel, S. Brlek, W. Jockusch, S. Plouffe, B.E. Sagan, A relative of the Thue-Morse sequence, in *Formal power series and algebraic combinatorics* (Montreal, PQ, 1992), Discrete Math. 139, 455–461, 1995.

[2] J.-P. Allouche, J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, In C. Ding, T. Helleseth,and H. Niederreiter, editors, *Sequences and Their Applications, Proceedings of SETA'* 98, Springer-Verlag, 1–16, 1999.

[3] J.-P. Allouche, J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, 387-391, 2003.

[4] H. Fredricksen, Gray codes and the Thue-Morse-Hedlund sequence, *J. Combin. Math. Combin. Comput.* 11, 3–11, 1992.

[5] G.A. Hedlund, Remarks on the work of Axel Thue on sequences, *Nordisk Mat. Tidskr.* 15, 148–150, 1967.

[6] M. Morse, A one-to-one representation of geodesics on a surface of negative curvature, *Amer. J. Math.* 43, 35–51, 1921.

[7] S. Plouffe, P. Zimmermann, Quelques conjectures, preprint, 1992.

[8] P. Stănică, *Chromos, Boolean Functions and Avalanche Characteristics*, Ph.D. Thesis, State University of New York at Buffalo, 1998.

[9] A. Thue, Über unendliche Zeichenreichen, *Norske Vid Selsk. Skr. I. Mat. Nat. Kl. Christiana* 7 (1906), 1–22. Reprinted in "Selected mathematical papers of Axel Thue", T. Nagell ed. Universitetsforlaget, Oslo, 139–158, 1977.

[10] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Norske Vid Selsk. Skr. I. Mat. Nat. Kl. Christiana* 7 (1912), 1–67. Reprinted in "Selected mathematical papers of Axel Thue", T. Nagell ed. Universitetsforlaget, Oslo, 413–478, 1977.