TWO ANALOGS OF THUE-MORSE SEQUENCE

VLADIMIR SHEVELEV

Abstract. We introduce and study two analogs of one of the best known sequence in Mathematics : Thue-Morse sequence. The first analog is concerned with the parity of number of runs of 1's in the representation of nonnegative integers in binary (or in base 2). The second one is connected with the parity of number of 1's in the representation of nonnegative integers in so-called negabinary (or in base -2). We give for them some recurrent and structure formulas and consider several interesting difficult problems.

1. Introduction

Let $T = \{t_n\}|_{n\geq 0}$ be Thue-Morse sequence (or it is called also Prouhet-Thue -Morse sequence, Allouche and Shallit [2]). It is defined as the parity of number of $1's$ in binary representation of n. T is the sequence A010060 [8]. There are well known the following formulas for T ([2]):
i) (A recurrent formula). $t_0 = 0$, for $n > 0$, $t_{2n} = t_n$ and $t_{2n+1} = 1 - t_n$.
ii) (Structure formula). Let A_k denote the first 2^k terms; then $A_0 = 0$ and for A_{k+1}, $k >= 0$, we have a concatenation $A_{k+1} = A_kB_k$, where B_k is obtained from A_k by interchanging $0's$ and $1's$;
iii) (A relation which is equivalent to ii)). For $0 <= k < 2^m$, $t_{2^m+k} = 1-t_k$;

Some much more general formulas one can find in author's article [6]. In this paper we introduce and study two analogs of T :
1) Let $R = \{r_n\}|_{n\geq 0}$ be the parity of number of runs of $1's$ in the binary representation of n. R is our sequence A268411 in [8].
2) Let $G = \{g_n\}|_{n\geq 0}$ be the parity of number of $1's$ in the negabinary representation (or in base -2 of n. (cf. [9] and sequence A039724 [8]; see also [4, p.101], [5, p.189])). G is our sequence A269027 in [8].

We say several words concerning the appearance of the sequences G and R. For the first time, the numerical base -2 was introduced by *V.Grünwald* in 1885 (see [10] and references there). The author was surprised that during 130 years nobody considered a natural analog of Thue-Morse sequence in the base -2 , and he decided to study this sequence. Unexpectedly, it turned out that it has very interesting properties (although less canonical than Thue-Morse sequence). Moreover, it definitely has astonishing (yet unproved)

1991 *Mathematics Subject Classification.* 11B83.

joint properties with Thue-Morse sequence (see below our problems 3,4).

Concerning sequence R, it first appears in some-what exotic way. Let $u(n)$ be characteristic $(0,1)$ function of a sequence $S = \{1 = s_1 < s_2 < s_3 < ...\}$. In the main result of our recent paper [7] there appears the following series $f(x) = \sum_{i\geq 2}(u(i) - u(i-1))x^i$. It is easy to see that if to ignore zero coefficients (when $u(i) = u(i-1)$), then other coefficients form alternative $(-1,1)$ sequence. The author decided to introduce a numerical system based on base 2 with such an order of digits. So he considered the "balanced binary" representation of n which is obtained from the binary representation of n by replacing every 2^j by $2^{j+1} - 2^j$. The system was named "balanced" since the digital sum of every n in this system equals 0. For example $7 = 4+2+1 = (8-4)+(4-2)+(2-1) = 8-1 = (1,0,0,-1)_b$. The natural question: "how many pairs $1,-1$ are contained in the balanced binary representation of n?" is easily answered: this number equals the number of runs of $1's$ in the binary representation of n. This sequence modulo 2 is R.

2. Main results

In this paper we prove the following.

Theorem 1. *The following recursion holds.* $r_0 = 0, r_{2n} = r_n$; *for even* n, $r_{2n+1} = 1 - r_n$; *for odd* n, $r_{2n+1} = r_n$.

Theorem 2. *Let* R_k *denote the first* 2^k *terms of* R; *then* $R_1 = \{0,1\}$ *and for* $k >= 1$, *we have a concatenation* $R_{k+1} = R_kS_k$, *where* S_k *is obtained from* R_k *by complementing the first* 2^{k-1} $0's$ *and* $1's$ *and leaving the rest unchanged.*

Theorem 3. *The following recursion holds.* $g_0 = 0, g_{4n} = g_n$, $g_{4n+1} = 1-g_n$, $g_{4n+2} = 1 - g_{n+1}$, $g_{4n+3} = g_{n+1}$.

For the first time this statement was formulated by R. Israel in our sequence A269027 [8].

Theorem 4. *Let* G_k *denote the first* 2^k *terms of* G; *then* $G_0 = 0$ *and for even* $k >= 0$, *we have a concatenation* $G_{k+1} = G_kF_k$, *where* F_k *is obtained from* G_k *by complementing its* $0's$ *and* $1's$; *for odd* $k >= 1$, *we have a concatenation* $G_{k+1} = G_kH_k$, *where* H_k *is obtained from* G_k *by complementing its last* $(2/3)(2^{k-1} - 1)$ $0's$ *and* $1's$.

3. Proof of Theorem 1

Proof. 1) Trivially, $r_{2n} = r_n$.
2) Let $n = 2k,\ \ 2n+1 = 4k+1$ which ends on 00...01, where the number of zeros ≥ 1, then the last 1 forms a new run of 1's. So, $r_{2n+1} = 1 - r_{4k} = 1 - r_n$.
3) Let n be odd such that $n-1 = 2^m l$, where l is odd, $m \geq 1$ and $2n+1 = 2^{m+1}l + 3$.
3a) Let $m = 1$. Then $4l$ ends on two zeros and the adding of 3 does not form a new run. So, $r_{2n+1} = r_{4l+3} = r_{4l} = r_{2l} = r_{2l+1} = r_n$.
3b) Let $m \geq 2$. Then $2^{m+1}l$ ends on ≥ 3 zeros and the adding of 3 forms a new run. So, $r_{2n+1} = 1 - r_{2^{m+1}l} = 1 - r_{2^m l} = 1 - (1 - r_n) = r_n$. ☐

4. Proof of Theorem 2

Proof. It is easy to see that Theorem 2 is equivalent to the formula

$$r_{n+2^k} = \begin{cases} 1 - r_n, & 0 \leq n \leq 2^{k-1} - 1 \\ r_n, & 2^{k-1} \leq n \leq 2^k - 1. \end{cases} \tag{1}$$

In case when $n \in [2^{k-1}, 2^k - 1]$ in the binary expansion of n the maximal weight of 1 is 2^{k-1}. After addition of 2^k this new 1 continues the previous run of 1's in which there is 1 of the weight 2^{k-1}. So, in this case the number of runs of 1's does not change and $r_{n+2^k} = r_n$. In opposite case when $n \in [0, 2^{k-1} - 1]$ after addition of 2^k this new 1 forms a new run and the number of runs is increased on one, so $r_{n+2^k} = 1 - r_n$. ☐

5. Proof of Theorem 3

In binary expansion of n, we call *even 1's* the 1's with the weight $2^{2k},\ \ k \geq 0$, and other 1's we call *odd 1's*. In conversion from base 2 to base -2 an important role plays the parity of 1's in binary, since only every odd 1 with weight $2^{2k+1},\ \ k >= 0$, we should change by two 1's with weights $2^{2k+2},\ \ 2^{2k+1}$, which corresponds to the equality

$$2^{2k+1} = (-2)^{2k+2} + (-2)^{2k+1}.$$

For example $7 = 2^2 + 2 + 1 => 2^2 + 2^2 - 2 + 1 = 2^3 - 2 + 1 = 2^4 - 2^3 - 2 + 1 = 11011_{-2}$.

Proof. 1)Since multiplication n by 4 does not change the parity of 1's, then, evidently, $g_{4n} = g_n$.
2) Again evidently $g_{4n+1} = 1 - g_{4n} = 1 - g_n$.
3) Note also that $g_{2n+1} = 1 - g_{2n}$. Hence, $g_{4n+3} = g_{2(2n+1)+1} = 1 - g_{4n+2}$.
4) It is left to prove that $g_{4n+3} = g_{n+1}$ (then $g_{4n+2} = 1 - g_{n+1}$).

4a) Let n be even $= 2m$. We should prove that $g_{8m+3} = g_{2m+1}$. Note that $8m+3$ ends in the binary on 100...011, where the number of zeros≥ 1. Since 011 in binary converts to 111_{-2}, then $g_{8m+3} = 1 - g_{8m} = 1 - g_{2m} = 1 - (1 - g_{2m+1}) = g_{2m+1}$.
4b) Let n be odd and ends on even number of 1's. We need lemma.

Lemma 1. *For $m \geq 2$,*

$$2^m - 1 = \begin{cases} 100...011_{-2}, & m \textit{ is even} \\ 1100...011_{-2}, & m \textit{ is odd,} \end{cases} \tag{2}$$

where in the 0's run we have $m-2$ zeros.

Proof. Let m be even. Then we have

$$2^m - 1 = 2^{m-1} + 2^{m-2} + ... + 2 + 1 =$$
$$(2^m - 2^{m-1}) + 2^{m-2} + (2^{m-2} - 2^{m-3}) + ... + (16 - 8) + 4 + (4 - 2) + 1 =$$
$$2^m + (-2^{m-1} + 2^{m-1}) + (-2^{m-3} + 2^{m-3}) + ... + (-8+8) - 2 + 1 = 100...011_{-2}$$

such that zeros correspond to exponents $m-1, m-2, ..., 3, 2$, i.e., we have $m-2$ zeros. Now let m be odd. Then $m-1$ is even and, using previous result, we have

$$2^m - 1 = 2^{m-1} + 2^{m-2} + ... + 2 + 1 =$$
$$2^{m-1} + 2^{m-1} + (-2^{m-2} + 2^{m-2}) + (-2^{m-4} + 2^{m-4}) + ... + (-8+8) - 2 + 1 =$$
$$2^{m+1} - 2^m - 2 + 1 = 1100...011_{-2}$$

with also $m-2$ zeros. □

Corollary 1.

$$g_{2^m-1} = \begin{cases} 0, & m = 0 \\ 1, & m = 1 \\ 1, & m \geq 2 \textit{ is even} \\ 0, & m \geq 3 \textit{ is odd.} \end{cases} \tag{3}$$

Let n in the binary ends on even $m \geq 2$ 1's. Then $4(n+1)$ ends on 100...0 with $m+2$ zeros and thus the end of $4(n+1)$ equals $100...0_{-2}$ with $m+2$ zeros. On the other hand, $4n+3$ ends on $m+2 \geq 4$ 1's: 011...1, so, by (2), the end of $4n+3$ equals $10...011_{-2}$ with m zeros. Since all the previous binary digits for $4n+3$ and $4(n+1)$ are the same (indeed, $4n+4-(4n+3) = 1 = 10...0_2 - 01...1_2$), then, continuing conversion from base 2 to base -2, we obtain the equality $g_{4n+3} = g_{n+1}$.

4c) Finally, let n be odd and ends on odd number m of 1's. Consider in more detail the end of n. If n ends on $001..1_2$, then the proof does not differ from the previous case, since $4(n+1) = ...010...0_2$ with $m+2$ zeros, thus the end of $4(n+1)$ equals $110...0_{-2}$; on the other hand, $4n+3 = ...001...111_2$ and, by (2) the end of $4n+3$ equals $110...011$ with m zeros and we again conclude that $g_{4n+3} = g_{n+1}$.

Now let n ends on $01...101...1$, where the first run contains k 1's while the second run contains odd m 1's. Then $4n+3$ ends on $01...101...111$, while $n+1$ ends on $01...110...0$, where the run of 1's contains $k+1$ 1's which is followed by the run of m 0's. Let us pass in two last ends to base -2. We need a lemma which is proved in the same way as Lemma 1.

Lemma 2. *For odd* $m \geq 3$,

$$2^{m+k+1} - 2^m - 1 = \begin{cases} 10...010...011_{-2}, & k \text{ is even} \\ 110...010...011_{-2}, & k \text{ is odd,} \end{cases} \tag{4}$$

where the first (from the left to the right) 0's run has k *zeros, while the second 0's run has* $m-2$ *zeros;*

$$2(2^{k+1} - 1) = \begin{cases} 10...010_{-2}, & k \text{ is even} \\ 110...010_{-2}, & k \text{ is odd,} \end{cases} \tag{5}$$

where the first 0's run has k *zeros.*

So, by (4), for the end of $4n+3$ we have

$$01...101...111_2 = 2^{(m+2)+k+1} - 2^{m+2} - 1 =$$
$$\begin{cases} 10...010...011_{-2}, & k \text{ is even} \\ 110...010...011_{-2}...010...011_{-2}, & k \text{ is odd,} \end{cases}$$

where the 0's runs have k and m zeros respectively.

For the corresponding end of $4(n+1)$ having $m+2$ zeros at the end, by (5), we have (since $m+1$ is even):

$$01...110...0_2 = 2^{m+1} \cdot 1...110_2 =$$
$$2^{m+1} \cdot \begin{cases} 10...010_{-2}, & k \text{ is even} \\ 110...010_{-2}, & k \text{ is odd,} \end{cases} =$$
$$\begin{cases} 10...010...0_{-2}, & k \text{ is even} \\ 110...010...0_{-2}, & k \text{ is odd,} \end{cases}$$

where the 0's runs have k and $m+2$ zeros respectively.
Since all the previous binary digits for $4n+3$ and $4(n+1)$ are the same ($4n+4-(4n+3) = 1 = 1...110...0_2 - 1...101...1_2$), then, continuing conversion from base 2 to base -2, we obtain the equality $g_{4n+3} = g_{n+1}$. $\square$

6. Proof of Theorem 4

Proof. It is easy to see that Theorem 4 is equivalent to the formula

$$(6)\qquad g_{2^k+m}=\begin{cases}1-g_m, & k\ \ is\ \ even\geq 2 \text{ and } 2^{k-1}\leq m<2^k\\ g_m, & k\ \ is\ \ odd\geq 1 \text{ and } 0\leq m<2^k-\frac{2}{3}(2^{k-1}-1)\\ 1-g_m, & k\ \ is\ \ odd\geq 3 \text{ and } 2^k-\frac{2}{3}(2^{k-1}-1)\leq m<2^k.\end{cases}$$

1) Let k be even≥ 2 and $2^{k-1}\leq m<2^k$. We use induction over k. For $k=2,\ 2\leq m<4$, (6) is true: $g_{4+2}=1-g_2,\ \ g_{4+3}=1-g_3=0$; also it is easy verify (6) for $k=4$. Suppose that (6) is true for $k-2$. We write in binary $1\vee m$ instead of 2^k+m.
1a) Let $m=4x$. By the induction supposition $g_{1\vee x}=1-g_x$. But also, by the condition, $g_{1\vee m}=g_{1\vee x}=1-g_x$ and $g_m=g_x$. So $g_{1\vee m}=1-g_m$.
1b) Let $m=4x+1$. By the induction supposition $g_{1\vee x}=1-g_x$. But also, by the condition (since $g_{4n+1}=1-g_n$) we have $g_{1\vee m}=1-g_{1\vee x}=g_x$ and $g_m=1-g_x$. So $g_{1\vee m}=1-g_m$.
1c) Let $m=4x-1$. By the induction supposition $g_{1\vee x}=1-g_x$. But also, by the condition (since $g_{4n-1}=g_{4(n-1)+3}=g_n$) $g_{1\vee m}=g_{1\vee x}=1-g_x$ and $g_m=g_x$. So $g_{1\vee m}=1-g_m$.
1d) Let $m=4x-2$. By the induction supposition $g_{1\vee x}=1-g_x$. But also, by the condition (since $g_{4n-2}=g_{4(n-1)+2}=1-g_n$) $g_{1\vee m}=1-g_{1\vee x}=g_x$ and $g_m=1-g_x$. So $g_{1\vee m}=1-g_m$.
The proof in the following two points is the same, except for the bases of induction. Therefore in the points 2),3) we give the bases of induction only.
2) Let k be odd≥ 1 and $0\leq m<2^k-\frac{2}{3}(2^{k-1}-1)$. For $k=1$, we have $m=0,1$ and (6) is true: $g_{2+0}=g_0=0$ and $g_{2+1}=g_1=1$; for $k=3$, we have $m=0,1,2,3,4,5$ and (6) is true: $g_{8+0}=g_0=0$, $g_{8+1}=g_1=1$, $g_{8+2}=g_2=0$, $g_{8+3}=g_3=1$, $g_{8+4}=g_4=1$, $g_{8+5}=g_5=0$.
3) Let k be odd≥ 3 and $\frac{2}{3}(2^{k-1}-1)\leq m<2^k$. For $k=3$, we have $m=6,7$ and (6) is true: $g_{8+6}=1-g_6=0$ and $g_{8+7}=1-g_7=1$. for $k=5$, we have $m=22,23,...,31$ and (6) is true: $g_{32+22}=1-g_{22}=1$, $g_{32+23}=1-g_{23}=0$ ... $g_{32+31}=1-g_{31}=1$. □

7. Several difficult author's problems

1) [6]. For which positive numbers a,b,c, for every nonnegative n there exists $x\in\{a,b,c\}$ such that $t_{n+x}=t_n$?
This problem was solved in [6] only partially. For example, for every $a\geq 1,\ \ k\geq 0$, the triple $\{a,a+2^k,a+2^{k+1}\}$ is suitable. However, there are other infinitely many solutions.

2) Conjecture[6]. Let $u(n) = (-1)^{t_n}|_{n\geq 0}$ and a be a positive integer. Let $\{l_0 < l_1 < l_2 < ...\}$, $\{m_0 < m_1 < m_2 < ...\}$ be integer sequences for which $u(l_i + a) = -u(l_i)$, $u(m_i + a) = u(m_i)$. Let $\beta_a(n) = u(l_n)$, $\gamma_a(n) = u(m_n)$. Then the sequences β_a, γ_a are periodic, of the smallest period $2^{v(a)+1}$, where $v(a)$ is such that $2^v(a)||a$. They satisfy $\beta_a = -\gamma_a$.
This conjecture was proved by Allouche [1].
3) (in A268866 [8]). Let $v(n)$ be the maximal number k such that $g_r = t_{r+n}$, $r = 0, 1..., k-1$(if $k = 0$, there is no equality already for $r = 0$.) Let $\{a(n)\}$ be the sequence of records in the sequence $\{v(n)\}$. Conjecture: 1) Let $l(n)$ be the position in $\{v(n)\}$ corresponding to $a(n)$. Then $l(n) = (2/3)(4^n - 1)$, if n is even, and $l(n) = (2/3)(4^{n-1} - 1) + 3 \cdot 4^{n-1}$, if n is odd; 2) $a(n) = 2l(n) + 2$, if n is even, and $a(n) = (7l(n) + 12)/11$, if n is odd.
4) (A dual problem: in A269341 [8]). Let $w(n)$ be the maximal number k such that $g_r = 1 - t_{r+n}$, $r = 0, 1..., k-1$. Let $\{b(n)\}$ be the sequence of records in the sequence $\{w(n)\}$. Denote by $m(n)$ the position in $\{w(n)\}$ corresponding to $b(n)$. Then $m(0) = 0, m(1) = 1$. Conjecture: 1) for even $n \geq 2$, $m(n) = (2/3)(4^{n-1}-1)$; for odd $n \geq 3$, $m(n) = (2/3)(4^{n-2}-1)+3\cdot 4^{n-2}$; 2) for even $n \geq 2$, $b(n) = 2m(n)+2$; for odd $n \geq 3$, $b(n) = (7m(n)+12)/11$.
5) Together with the Thue-Morse constant $P = 0.01101001100101..._2$ which is given by the concatenated digits of the Thue-Morse sequence A010060 and interpreted as a binary number, consider the constant $S = 0.01111011100..._2$ and the constant $F = 0.0101101001..._2$ which are given by the concatenated digits of the sequences $R = A268411$ and $G = A269027$ respectively and interpreted as binary numbers. Dekking [3] proved that P is a transcendent number. We ask, are the constants S and F transcendent?

The author hopes that this paper will help to solve at least the problems 3 and 4. The paper is connected with the following sequences in [8]: A000695, A010060, A039724, A069010, A020985, A022155, A203463, A268382, A268383, A268411, A268412, A268415, A268865, A268866, A268272, A268273, A268476, A268477, A268483, A269003, A269027, A269340, A269341, A269458, A269528, A269529.

8. Acknowledgement

The author is very grateful to Peter J. C. Moses for the numerical calculations.

References

[1] J.-P. Allouche, Thue, Combinatorics on words, and conjectures inspired by the Thue-Morse sequence, J. de Theorie des Nombres de Bordeaux, 27, no. 2 (2015), 375-388.

[2] J.-P. Allouche and J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, http://www.lri.fr/∼allouche/bibliorecente.html

[3] F. M. Dekking, Transcendence du nombre de Thue-Morse. Comptes Rendus de l'Academie des Sciences de Paris 285 (1977) 157-160.

[4] M. Gardner, Knotted Doughnuts and Other Mathematical Entertainments. Freeman, NY, 1986.

[5] D. E. Knuth, The Art of Computer Programming. Addison-Wesley, 1998, Vol. 2 (3rd ed.), 204-205.

[6] V. Shevelev, Equations of the form $t(x + a) = t(x)$ and $t(x + a) = 1 - t(x)$ for Thue-Morse sequence, Preprint 2009 and 2012, available electronically at http://arxiv.org/abs/0907.0880.

[7] V. Shevelev, *Exponentially S-numbers,* arXiv:1510.05914 [math.NT], 2015.

[8] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences* http://oeis.org.

[9] Eric Weisstein, Negabinary (MathWorld).

[10] Wikipedia, Negative base.

Department of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. e-mail:shevelev@bgu.ac.il