Permutation complexity of the fixed points of some uniform binary morphisms

Alexander Valyuzhenich

Novosibirsk State University, Novosibirsk, Russia

graphkiper@mail.ru

An infinite permutation δ is a linear order on the set $\mathbb{N}$. We study the properties of infinite permutations generated by fixed points of some uniform binary morphisms, and find the formula for their complexity.

1 Introduction

The notion of an infinite permutation was introduced in [1], where were investigated the periodic properties and low complexity of permutations. Similarly to the definition of subword complexity of infinite words, we can introduce the notion of the factor complexity of a permutation as the number of distinct subpermutations of a given length. The notion of a permutation generated by an infinite non-periodic word was introduced in [3]. In [4] Makarov calculated the factor complexity of permutations generated by a well-known family of Sturmian words. In [5] Widmer calculated the factor complexity of the permutation generated by the Thue-Morse word. In this paper we find a formula for the factor complexity of permutations generated by the fixed points of binary uniform morphisms from a some class. Since the Thue-Morse word belongs to this class, we obtain an alternative way to compute the factor complexity of the Thue-Morse permutation.

In Section 2 we introduce basic definitions to be used below. In Section 3, we introduce the class of morphisms Q for which in Section 9 we state the main theorem of this paper. In Sections 4–8 we state some auxiliary assertions needed to prove the main theorem. In Section 10 we give an alternative proof of the formula for the factor complexity of the permutation generated by the Thue-Morse word.

2 Basic definitions

Let Σ be a finite alphabet. Everywhere below we will use only the two-letter alphabet $\Sigma = \{0, 1\}$.

A right infinite word over the alphabet Σ is a word of the form $\omega = \omega_1 \omega_2 \omega_3 \ldots$, where each $\omega_i \in \Sigma$. A (finite) word u is called a subword of a (finite or infinite) word v if $v = s_1 u s_2$ for some words s_1 and s_2 which may be empty. The set of all finite subwords of the word ω is denoted by $F(\omega)$. For the word ω we define the binary real number $R_\omega(i) = 0.\omega_i \omega_{i+1} \ldots = \sum_{k \geq 0} \omega_{i+k} 2^{-(k+1)}$. The mapping $h : \Sigma^* \longrightarrow \Sigma^*$ is called a morphism if $h(xy) = h(x)h(y)$ for any words $x, y \in \Sigma^*$. We say that ω is a *fixed point* of a morphism φ if $\varphi(\omega) = \omega$. Clearly, every morphism is uniquely determined by the images of symbols, which we call *blocks*. A morphism is called *uniform* if its blocks are of the same length.

We say that a morphism $\varphi : \Sigma^* \longrightarrow \Sigma^*$ is *marked* if its blocks are of the form $\varphi(a_i) = b_i x c_i$, where x is an arbitrary word, b_i and c_i are symbols of the alphabet Σ, and all b_i (as well as all c_i) are distinct. In what follows, we will consider only uniform marked morphisms with blocks of length l.

An *interpretation* of a word $u \in \Sigma^*$ under the morphism φ is a triple $s = \langle v, i, j \rangle$, where $v = v_1 \ldots v_k$ is a word over the alphabet Σ, i and j are nonnegative integers such that $0 \leq i < |\varphi(v_1)|$ and $0 \leq j < |\varphi(v_k)|$, and the word obtained from $\varphi(v)$ by erasing i symbols to the left i and j symbols to the right is u. Moreover, if v is a subword of ω, then s is called an interpretation on ω. The word v is called an *ancestor* of the word u. In what follows we shall consider only interpretations of subwords of the word ω. We say that (u_1, u_2) is a *synchronization point* of $u \in F(\omega)$ if $u = u_1 u_2$ and $\forall v_1, v_2 \in \Sigma^*, \forall s \in F(\omega) \exists s_1, s_2 \in F(\omega)$ such that $[v_1 u v_2 = \varphi(s) \Rightarrow (s = s_1 s_2, v_1 u_1 = \varphi(s_1), u_2 v_2 = \varphi(s_2))]$. A fixed point $\omega = \varphi(\omega)$ of the morphism φ is called *circular* ([2]), if any subword v of word ω of length at least L_ω contains at least one point of synchronization.

For uniform morphisms, this means the uniqueness of the partition of the word v into blocks.

In [2] it was proved that the nonperiodic fixed points of uniform binary morphisms with $w_1 = 0$ are circular, except for the case when $\varphi(1) = 1^n$.

An *occurrence* of word $u \in \Sigma^*$ in the word ω is a pair $(u, m+1)$ such that $u = \omega_{m+1} \omega_{m+2} \ldots \omega_{m+n}$. It is easy to see that a word can have many different occurrences.

Since the fixed point ω is circular, the interpretations of all occurrences of the word u are the same and equal to $\langle v, i, j \rangle$.

An occurrence $(v, p+1)$ of a word v of length k is called the *ancestor of the occurrence* $(u, m+1)$ of the word u if $m = pl + i$ and $0 \leq i < l$.

It is easy to see that for $|u| \geq L_\omega$ word u has exactly one ancestor, and any occurrence $(u, m+1)$ of the word u also has exactly one ancestor.

Let $|u| \geq L_\omega$. A sequence $u_0, u_1, \ldots, u_m$ of subwords of word ω is called a *chain of ancestors* of the word u, if u_{i+1} is the ancestor of u_i for any $0 \leq i \leq m-1$ and $u_0 = u$.

The chain of ancestors of word u will be denoted as $u \to u_1 \to \ldots \to u_m$.

We say that u is a *descendant* of v if v belongs to a chain of ancestors of u.

Now we introduce the main object of this paper.

Let ω be a right infinite nonperiodic word over the alphabet Σ.

We define the *infinite permutation* generated by the word ω as the ordered triple $\delta = \langle \mathbb{N}, <_\delta, < \rangle$, where $<_\delta$ and $<$ are linear orders on $\mathbb{N}$.

The order $<_\delta$ is defined as follows: $i <_\delta j$ if and only if $R_\omega(i) < R_\omega(j)$, and $<$ is the natural order on $\mathbb{N}$.

Since ω is a non-periodic word, all $R_\omega(i)$ are distinct, and the definition above is correct.

We define a function $\gamma : R^2 \to \{<, >\}$, which for two different real numbers reveals their relation.

We say that a permutation $\pi = \pi_1 \ldots \pi_n \in S_n$ is a *subpermutation* of length n of an infinite permutation δ if $\gamma(\pi_s, \pi_t) = \gamma(R(i+s), R(i+t))$ for $1 \leq s < t \leq n$ and for a fixed positive integer i.

We define the set $Perm(n) = \{\pi(i, n+i-1) | i \geq 1\}$, where $\pi(i, n+i-1) = \pi_i \ldots \pi_{n+i-1}$ is subpermutation induced by the sequence $R(i), \ldots, R(n+i-1)$ (in the sense that $\gamma(\pi_{i+s_1}, \pi_{i+s_2}) = \gamma(R(i+s_1), R(i+s_2))$ for $0 \leq s_1 < s_2 \leq n-1$).

Now we define the *permutation complexity* of the word ω (or equivalently, the factor complexity of the permutation δ_ω) as $\lambda(n) = |Perm(n)|$. We say that an occurrence $(u, m+1)$ of the word u *generates* a permutation π if π is induced by a sequence of numbers $R(m+1) \ldots R(m+n)$. A subword u of the word ω *generates* the permutation π if there is an occurrence $(u, m+1)$ of this word which generates π.

3   Morphisms considered

We say that a uniform marked binary morphism φ with blocks of length l belongs to the class Q if one of the following conditions is fulfilled: either $\varphi(0) = X = 01^n0x1, \varphi(1) = Y = 10^m1y0$, where $n, m \in \mathbb{N}$, both the word 1^n and the word 0^m is included in both blocks morphism exactly once, and the word X (Y) does not end by $1^{n-1}(0^{m-1})$; or $\varphi(0) = 01^n, \varphi(1) = 10^n$, where $n = l - 1$.

It is easy to see that the fixed point $\omega = \lim\limits_{n \to \infty} \varphi^n(0)$ of any morphism φ which belongs to the class Q is circular because $\varphi(1) \neq 1^n$ for any n.

Everywhere below the word $\varphi(0)$ will be called the block of the first type, and $\varphi(1)$ is called the block of the second type.

Example. The morphism $\varphi(0) = 011101, \varphi(1) = 100010$ belongs to Q, whereas the morphism $\varphi(0) = 01011, \varphi(1) = 10000$ does not belong to Q.

Consider a fixed point $\omega = \varphi(\omega)$ of a morphism $\varphi \in Q$. Then ω is divided into blocks, which are the images of its symbols. Such a partition is called *correct*.

Lemma 1 *Let ω be a fixed point of the morphism φ, where $\varphi \in Q$. Then the following statements are true:*

1. *Let $\omega_i = \omega_j = 0$ and $i \equiv 1 \bmod l$, $j \not\equiv 1 \bmod l$. Then $R_\omega(i) > R_\omega(j)$.*

2. *Let $\omega_i = \omega_j = 1$ and $i \equiv 1 \bmod l$, $j \not\equiv 1 \bmod l$. Then $R_\omega(i) < R_\omega(j)$.*

Lemma 2 *Let ω be a fixed point of the morphism φ, where $\varphi \in Q$. Let $\omega_i = \omega_j$, where $i \equiv i'(\bmod\ l), j \equiv j'(\bmod\ l)$ and $0 \leq i', j' \leq l - 1$, where i' and j' are fixed. If $i' \neq j'$, or if ω_i and ω_j lie in blocks of different types in the correct partition ω into blocks, then the relation $\gamma(R_\omega(i), R_\omega(j))$ is uniquely defined by i', j' and the types of respective blocks.*

Lemma 3 *Let $\omega_i = \omega_j$ and $R_\omega(i) < R_\omega(j)$. Then inequality $R((i-1)l+r) < R((j-1)l+r)$ holds for all $1 \leq r \leq l$.*

4   Equivalence of permutations

In this section we introduce the concept of equivalent permutations. Let $z = z_1z_2 \ldots z_k$ be a permutation belonging to S_k.

An *element* of the permutation z is the number z_i, where $1 \leq i \leq k$.

We will say that two permutations $x = x_1x_2 \ldots x_k$ and $y = y_1y_2 \ldots y_k$ are *equivalent* if they differ only in relations of extreme elements, i.e $\gamma(x_1, x_k) \neq \gamma(y_1, y_k)$, but $\gamma(x_i, x_j) = \gamma(y_i, y_j)$ for all other i, j. We will denote this equivalence by $x \sim y$.

Lemma 4 *Let x be a finite permutation and $x = x_1 \ldots, x_k$. Then the permutation y such that $x \sim y$ exists if and only if $|x_1 - x_k| = 1$.*

5   Bad, narrow and wide words

For an arbitrary subword v of the word ω we define the sets M_v and N_v, where N_v is the set of all pairs of equivalent permutations, and M_v is the remaining set of permutations generated by v.

A word u will be called *bad* if the set N_u is not empty, i.e, if u generates at least one pair of equivalent permutations.

Let u be a subword of word ω of length $|u| = n$. The number of permutations generated by u is denoted by $f(u)$.

Lemma 5 *Let u be a word of length $|u| = n \geq L_\omega$, and u' be the ancestor of u. Then the following statements are true:*

1. $f(u) \leq f(u')$,
2. *If $N_{u'} = \emptyset$, then $N_u = \emptyset$ and $f(u) = f(u')$.*

Lemma 6 *Let u be a bad word and $|u| = n \geq L_\omega$, and u' be the ancestor of u. Then u' is a bad word and $f(u) = f(u'), m_u = m_{u'}, n_u = n_{u'}$.*

It is worth noting that from the proof of Lemma 6 it follows that if u is a bad word of length $|u| = n \geq L_\omega$, then $|u| \equiv 1 \pmod{l}$.

The set of all words of length less than L_ω, having descendants of length at least L_ω is denoted by A.

The set of bad words of length n, whose chain of ancestors is $u \to u_1 \to u_2 \to \ldots \to u_m = a$, where $m \in \mathbb{N}$ (m is not fixed) and $a \in A$, is denoted by $F_a^{bad}(n)$. The cardinality of the set $F_a^{bad}(n)$ is denoted by $C_a^{bad}(n)$.

Lemma 7 *Let $u \in F_a^{bad}(n)$, where $n \geq L_\omega$. Then $f(u) = m_a + 2n_a$.*

A word u will be called *narrow* if its chain of ancestors is $u \to \ldots \to u_{k-1} \to u_k \to \ldots \to u_m = a$, where $a \in A$, u_k is the first bad word in the chain of ancestors, and for the interpretation $\langle u_k, i, j \rangle$ of the word u_{k-1} we have $i + 1 > l - j$.

Lemma 8 *Let u be a narrow word with $|u| = n \geq L_\omega$, u' be an ancestor of u, and u' be a bad word. Then $n_u = 0$ and $f(u) = m_{u'} + n_{u'}$.*

The set of narrow words of length n whose chain of ancestors is $u \to u_1 \to u_2 \to \ldots \to u_m = a$, where $m \in \mathbb{N}$ (m is not fixed), is denoted by $F_a^{nar}(n)$. The cardinality of the set $F_a^{nar}(n)$ is denoted by $C_a^{nar}(n)$.

Lemma 9 *Let $u \in F_a^{nar}(n)$, where $|u| = n \geq L_\omega$. Then $f(u) = m_a + n_a$.*

A word u will be called *wide* if its chain of ancestors is $u \to \ldots \to u_{k-1} \to u_k \to \ldots \to u_m = a$, where $a \in A$, u_k is the first bad word in the chain of ancestors, and for the interpretation $\langle u_k, i, j \rangle$ of the word u_{k-1} we have $i + 1 < l - j$.

Lemma 10 *Let u be a wide word with $|u| = n \geq L_\omega$, u' be an ancestor of u, and u' be a bad word. Then $n_u = 0$ and $f(u) = m_{u'} + 2n_{u'}$.*

The set of wide words of length n whose chain of ancestors is $u \to u_1 \to u_2 \to \ldots \to u_m = a$, where $m \in \mathbb{N}$ (m is not fixed), is denoted by $F_a^{wide}(n)$. The cardinality of the set $F_a^{wide}(n)$ is denoted by $C_a^{wide}(n)$.

Lemma 11 *Let $u \in F_a^{wide}(n)$, where $|u| = n \geq L_\omega$. Then $f(u) = m_a + 2n_a$.*

6   Algorithm for finding $f(u)$

Suppose that u is a subword of ω and $|u| = n$. In this section, we calculate $\sum_{|u|=n} f(u)$. The set of all subwords of length n of the word ω, whose chain of ancestors is $u \to u_1 \to u_2 \to \ldots \to u_m = z$, where $m \in \mathbb{N}$ (m is not fixed), is denoted by $F_z(n)$. The cardinality of the set $F_z(n)$ is denoted by $C_z(n)$.

We consider the set A introduced in the previous section. Let $A = A_1 \cup A_2$ be a partition of set A, where A_1 is the set of bad words belonging to the set A, and A_2 is the set of remaining words of A. Thus, for a word u there are two opportunities:

1) $a \in A_2$. In this case, Lemma 5 implies that $f(u) = f(a) = m_a$,

2) $a \in A_1$. In this case, there are two cases: if u is a bad or a wide word, due to Lemmas 7 and 11 we obtain $f(u) = f(a) = m_a + 2n_a$. If u is a narrow word, then by Lemma 4 we obtain $f(u) = m_a + n_a$.

Theorem 1 $\sum_{|u|=n} f(u) = \sum_{a_1 \in A_1} [C_{a_1}^{nar}(n)(m_{a_1} + n_{a_1}) + (C_{a_1}^{bad}(n) + C_{a_1}^{wide}(n))(m_{a_1} + 2n_{a_1})] + \sum_{a_2 \in A_2} C_{a_2}(n) m_{a_2}$.

Now for the calculation of $\sum_{|u|=n} f(u)$ it remains to compute $C_a^{nar}(n)$, $C_a^{wide}(n)$, $C_a^{bad}(n)$, $C_a(n)$. Let $n = xl + r$, where $0 \leq r \leq l - 1$. It is easy to see that for $xl + r \geq L_\omega$ the following recurrence relations hold:

1. $C_a^{bad}(xl + 1) = l C_a^{bad}(x + 1)$. We note that the remark to Lemma 6 implies that $C_a^{bad}(xl + r) = 0$ for $r \neq 1$.

2. $C_a^{nar}(xl + r) = (r - 1)C_a^{nar}(x + 2) + (r - 1)C_a^{bad}(x + 2) + (l - r + 1)C_a^{nar}(x + 1)$ for $r \geq 1$.
   $C_a^{nar}(xl) = (l - 1)C_a^{nar}(x + 1) + (l - 1)C_a^{bad}(x + 1) + C_a^{nar}(x)$.

3. $C_a^{wide}(xl + r) = (r - 1)C_a^{wide}(x + 2) + (l - r + 1)C_a^{wide}(x + 1) + (l - r + 1)C_a^{bad}(x + 1)$ for $r \geq 2$.
   $C_a^{wide}(xl + 1) = l C_a^{wide}(x + 1)$.
   $C_a^{wide}(xl) = (l - 1)C_a^{wide}(x + 1) + C_a^{wide}(x) + C_a^{bad}(x)$.

4. $C_a(xl + r) = (r - 1)C_a(x + 2) + (l - r + 1)C_a(x + 1)$ for $r \geq 1$.
   $C_a(xl) = (l - 1)C_a(x + 1) + C_a(x)$.

7 Special words

Recall that the subword v of word ω is called special if $v0$ and $v1$ are also subwords of ω.

Note that the unique interpretation of any special word v of length at least L_w is equal to $\langle v', i, 0 \rangle$.

Indeed, if $j > 0$, then v is uniquely complemented to the right to a full block, and thus only one of the words $v0$ and $v1$ is a subword of ω.

Lemma 12 *Let $(v0, m_1)$ and $(v1, m_2)$ be some occurrences of words $v0$ and $v1$, where v is a special word with the ancestor v'. Let $(v'0, m'_1)$ and $(v'1, m'_2)$ be the ancestors of occurrences of $(v0, m_1)$ and $(v1, m_2)$, where $(v'0, m'_1)$ and $(v'1, m'_2)$ generate the same permutation. Then $(v0, m_1)$ and $(v1, m_2)$ also generate the same permutation.*

Lemma 13 *Let $(v0, m_1)$ and $(v1, m_2)$ be some occurrences of words $v0$ and $v1$, where v is a special word with the ancestor v'. Let $(v'0, m'_1)$ and $(v'1, m'_2)$ be the ancestors of occurrences of $(v0, m_1)$ and $(v1, m_2)$, where $(v'0, m'_1)$ and $(v'1, m'_2)$ generate different non-equivalent permutations. Then $(v0, m_1)$ and $(v1, m_2)$ also generate different non-equivalent permutations.*

Lemma 14 *Let $(v0, m_1)$ and $(v1, m_2)$ be some occurrences of words $v0$ and $v1$, where v is a special word with the ancestor v'. Let $(v'0, m'_1)$ and $(v'1, m'_2)$ be the ancestors of occurrences of $(v0, m_1)$ and $(v1, m_2)$. Then the following statements are true:*

1) If $(v'0, m'_1)$ and $(v'1, m'_2)$ generate equivalent permutations and $|v| = l|v'|$, then $(v0, m_1)$ and $(v1, m_2)$ also generate equivalent permutations.

2)If $(v0, m_1)$ and $(v1, m_2)$ generate equivalent permutations, then $(v'0, m'_1)$ and $(v'1, m'_2)$ also generate equivalent permutations and $|v| = l|v'|$.

3)If $(v'0, m'_1)$ and $(v'1, m'_2)$ generate equivalent permutations and $|v| < l|v'|$, then $(v0, m_1)$ and $(v1, m_2)$ generate the same permutation.

We will consider special word v of length $n - 1$.

Let v, without loss of generality, start with 0. Then $v1$ cannot generate equivalent permutations.

The number of permutations of the set M_{v0} which also belong to M_{v1}, is denoted by k_v.

The number of permutations of the set M_{v0}, such that each of their is equivalent to some permutation of the set M_{v1}, is denoted by t_v.

The number of pairs of permutations of a set N_{v0} such that a permutation of the pair is equivalent, and the other is equal to some permutation of the set M_{v1}, is denoted by r_v.

Lemma 15 *Let v be a special word with the ancestor v', and $|v| < l|v'|$. Then $k_v = k_{v'} + t_{v'} + r_{v'}$, $t_v = 0$ and $r_v = 0$.*

Lemma 16 *Let v be a special word with the ancestor v', and $|v| = l|v'|$. Then $k_v = k_{v'}$, $t_v = t_{v'}$ and $r_v = r_{v'}$.*

Lemma 17 *Let v be a special word with the ancestor v', and $t_{v'} = r_{v'} = 0$. Then $k_v = k_{v'}$ and $t_v = r_v = 0$.*

8 Algorithm for finding $g(v)$

Suppose v is a special word and $|v| = n - 1$. The set of all the special words of length n is denoted by $B(n)$. The number of common permutations generated by some occurrences of words $v0$ and $v1$ is denoted by $g(v)$.

In this section, we calculate $\sum_{|v|=n-1} g(v)$.

The set of all special subwords of length n of the word ω, whose chain of ancestors is $v \to v_1 \to v_2 \to \ldots \to v_m = z$, is denoted by $B_z(n)$. The cardinality of the set $B_z(n)$ is denoted by $S_z(n)$. It is clear that the chain of ancestors of word v consists of special words. The set of all special words of length less than L_ω, with the descendants of the length greater than L_ω is denoted by B.

Lemma 18 *Let $xl + r \geq L_\omega$, $0 \leq r \leq l - 1$, and b be a special word. Then the following recurrence relations hold:*

1)$S_b(xl + r) = S_b(x + 1)$ if $r > 0$.
2)$S_b(xl) = S_b(x)$ if $r = 0$.
3)$S_b(l^k|b|) = 1$ if $k \geq 1$.

Now we calculate $g(v)$.

Lemma 19 *Let $v \in B_b(n - 1)$. Then the following statements are true:*
1)If $n \neq l^k|b| + 1$ for any positive integer k, then $g(v) = k_b + t_b + r_b$.
2)If $n = l^k|b| + 1$ for some positive integer k, then $g(v) = k_b + r_b$.

Let us introduce the function $\delta(n, b)$: if $n = l^s|b| + 1$ for some positive integer s, then $\delta(n, b) = 0$, otherwise $\delta(n, b) = 1$.

Theorem 2 $\sum_{v \in B(n-1)} g(v) = \sum_{b \in B} [S_b(n-1)(k_b + t_b + r_b)\delta(n, b) + (k_b + r_b)(1 - \delta(n, b))]$.

9 The main theorem

In this section we state the main theorem of this article. For this purpose we use the following Lemma (Lemma 1 from [3]).

Lemma 20 *Let $u = u_1 \ldots u_n$ and $v = v_1 \ldots v_n$ be two subwords of word ω and $u_i \neq v_i$ for some $1 \leq i \leq n-1$. Then u and v do not generate the same permutations.*

We can now prove the main theorem of this paper:

Theorem 3 *Let ω be a fixed point of the morphism φ, where $\varphi \in Q$. Then the permutation complexity of ω is calculated as follows: $\lambda(n) = \sum_{a_1 \in A_1}[C_{a_1}^{nar}(n)(m_{a_1} + n_{a_1}) + (C_{a_1}^{bad}(n) + C_{a_1}^{wide}(n))(m_{a_1} + 2n_{a_1})] + \sum_{a_2 \in A_2} C_{a_2}(n)m_{a_2} - \sum_{b \in B}[S_b(n-1)(k_b + t_b + r_b)\delta(n,b) + (k_b + r_b)(1 - \delta(n,b))].$*

10 Permutation complexity of the Thue-Morse Word

In [5] Widmer calculated the factor complexity of the permutation generated by the Thue-Morse word. In this section, we present an alternative proof of the formula for permutation complexity of the Thue-Morse word. Let $n = 2^k + b$, where $0 < b \leq 2^k$. Note that the length L_ω of the synchronization of the Thue-Morse word is 4. It is also easy to understand that $A_1 = \{010, 101\}$ and $A_2 = \{00, 01, 10, 11, 001, 011, 100, 101\}$ (for example, 010 generates two equivalent permutations 132 and 231 due to occurrences of words $\omega_{11}\omega_{12}\omega_{13}$ and $\omega_4\omega_5\omega_6$, respectively). Thus, we obtain:

$\sum_{|u|=n} f(u) = \sum_{a_1 \in A_1}[C_{a_1}^{nar}(n)(m_{a_1} + n_{a_1}) + (C_{a_1}^{bad}(n) + C_{a_1}^{wide}(n))(m_{a_1} + 2n_{a_1})] + \sum_{a_2 \in A_2} C_{a_2}(n)m_{a_2} = \sum_{a_1 \in A_1}[C_{a_1}^{nar}(n) + 2(C_{a_1}^{bad}(n) + C_{a_1}^{wide}(n))] + \sum_{a_2 \in A_2} C_{a_2}(n) = C_{010}^{bad}(n) + C_{010}^{wide}(n) + C_{101}^{bad}(n) + C_{101}^{wide}(n) + C(n).$

It is clear that for $C_a^{bad}(n)$ and $C_a^{wide}(n)$ the following recurrence relations hold:

$C_a^{bad}(2n+1) = 2C_a^{bad}(n+1)$, $C_a^{wide}(2n+1) = 2C_a^{wide}(n+1)$, $C_a^{wide}(2n) = C_a^{wide}(n+1) + C_a^{wide}(n) + C_a^{bad}(n)$. Hence it is easy to see that $C_{010}^{bad}(2^k + 1) = C_{101}^{bad}(2^k + 1) = 2^{k-1}$ for $k > 0$; for other n we have $C_{010}^{bad}(n) = C_{101}^{bad}(n) = 0$.

Let us prove by induction on n that for $2^k + 2 \leq n < 3 * 2^{k-1} (k > 2)$ the relation $C_{010}^{wide}(n) = C_{101}^{wide}(n) = 2^{k-1} - b + 1$ holds. The base $n = 10$ follows from the relations $C_{010}^{wide}(10) = C_{010}^{wide}(6) + C_{010}^{wide}(5) + C_{010}^{bad}(5) = 1 + 0 + 2 = 3 = 2^{3-1} - 2 + 1$ and $C_{010}^{wide}(11) = 2C_{010}^{bad}(6) = 2 = 2 = 2^{3-1} - 3 + 1$.

Let us prove the induction step. If $n = 2^k + 2b'$, then by the induction hypothesis we have $C_{010}^{wide}(2^{k-1} + b') = C_{101}^{wide}(2^{k-1} + b') = 2^{k-2} - b' + 1$ and $C_{010}^{wide}(2^{k-1} + b' + 1) = C_{101}^{wide}(2^{k-1} + b' + 1) = 2^{k-2} - b'$. Hence we obtain $C_{010}^{wide}(n) = C_{101}^{wide}(n) = C_{010}^{wide}(2^{k-1} + b') + C_{010}^{wide}(2^{k-1} + b' + 1) + C_{010}^{bad}(2^{k-1} + b') = 2^k - 2b' + 1$. If $n = 2^k + 2b' + 1$, then by the induction hypothesis we have $C_{010}^{wide}(2^{k-1} + b' + 1) = C_{101}^{wide}(2^{k-1} + b' + 1) = 2^{k-2} - b'$. Hence we obtain $C_{010}^{wide}(n) = C_{101}^{wide}(n) = 2C_{010}^{wide}(2^{k-1} + b' + 1) = 2(2^{k-2} - b') = 2^{k-1} - 2b'$. The induction step is proved.

Similarly, we prove by induction on n that for $3 * 2^{k-1} + 1 \leq n \leq 2^{k+1} + 1$ the relation $C_{010}^{wide}(n) = C_{101}^{wide}(n) = 0$ holds. The base $n = 7$ and $n = 8$ follow from the relations $C_{010}^{wide}(8) = C_{101}^s(8) = C_{010}^{wide}(5) + C_{010}^{wide}(4) + C_{010}^{bad}(4) = 0 + 0 + 0 = 0$, $C_{010}^{wide}(7) = C_{101}^{wide}(7) = 2C_{010}^{wide}(4) = 0$. Let us prove the induction step. If $n = 2^k + 2b'$, then by the induction hypothesis we have $C_{010}^{wide}(2^{k-1} + b') = C_{101}^{wide}(2^{k-1} + b') = 0$ and $C_{010}^{wide}(2^{k-1} + b' + 1) = C_{101}^s(2^{k-1} + b' + 1) = 0$. Hence we obtain $C_{010}^{wide}(n) = C_{101}^{wide}(n) = C_{010}^{wide}(2^{k-1} + b') + C_{010}^{wide}(2^{k-1} + b' + 1) + C_{010}^{bad}(2^{k-1} + b' + 1) = 0 + 0 + 0 = 0$. If $n = 2^k + 2b' + 1$, then by the induction hypothesis we have $C_{010}^{wide}(2^{k-1} + b' + 1) = C_{101}^{wide}(2^{k-1} + b' + 1) = 0$. Hence we obtain $C_{010}^{wide}(n) = C_{101}^{wide}(n) = 2C_{010}^{wide}(2^{k-1} + b' + 1) = 2 \cdot 0 = 0$. The induction step is proved.

Thus we have proved the following:

1) If $2^k + 2 \leq n < 3 * 2^{k-1}$, then $\sum_{|u|=n} f(u) = 2(2^{k-1} - b + 1) + 4(2^k + b - 1) - 2^k = 2^{k+2} + 2b - 2$.

2) If $3 * 2^{k-1} + 1 \leq n < 2^{k+1} + 1$, then $\sum_{|u|=n} f(u) = 2(n-1) + 2^{k+1} = 2^{k+2} + 2b - 2$.

Therefore the equality $\sum_{|u|=n} f(u) = C(n) = 2(n-1) + 2^{k+1} = 2^{k+2} + 2b - 2$ holds for all $n \geq 6$.

It is clear that for $S_b(n)$ the following recurrence relations hold:

$S_b(2n+1) = S_b(n+1)$ and $S_b(2n) = S_b(n)$.

Taking into account that $S_{01}(3) = S_{10}(3) = 1$ and $S_{01}(4) = S_{10}(4) = 1$, it is easy to prove by induction that for $n \geq 3$ the equality $S_{01}(n) = S_{10}(n) = 1$ holds.

It is also easy to see that $B = \{01, 10, 010, 101\}$.

In addition word 010 generates two equivalent permutations 132 and 231, while word 011 generates only a single permutation 132. Hence $k_{01} = t_{01} = 0$ and $r_{01} = 1$. Similarly $k_{10} = t_{10} = 0$ and $r_{10} = 1$.

Words 0101 and 0100 generate different inequivalent permutations 1324 and 3421. Hence $k_{010} = t_{010} = 0$ and $r_{010} = 0$. Similarly $k_{101} = t_{101} = 0$ and $r_{101} = 0$. Therefore for $n \neq 2^k + 1$ by Theorem 2 the following relation holds:

$\sum_{v \in B(n-1)} g(v) = \sum_{b \in B}[S_b(n-1)(k_b + t_b + r_b)\delta(n,b) + (k_b + r_b)(1 - \delta(n,b))] = S_{01}(n-1)(k_{01} + t_{01} + r_{01}) + S_{10}(n-1)(k_{10} + t_{10} + r_{10}) = 1 + 1 = 2$.

For $n = 2^k + 1$ by Theorem 2 the following relation holds:

$\sum_{v \in B(n-1)} g(v) = \sum_{b \in B}[S_b(n-1)(k_b + t_b + r_b)\delta(n,b) + (k_b + r_b)(1 - \delta(n,b))] = k_{01} + r_{01} + k_{10} + r_{10} = 2$.

Thus the formula for the permutation complexity of the Thue-Morse word is

$\lambda(n) = \sum_{|u|=n} f(u) - \sum_{b \in B(n-1)} g(b) = 2^{k+2} + 2b - 2 - 2 = 2(2^{k+1} + b - 2)$ for $n = 2^k + b$, where $0 < b \leq 2^k$.

11 Acknowledgements

References

[1] D. G. Fon-Der-Flaass & A. E. Frid (2007): *On periodicity and low complexity of infinite permutations. European J. Combin.* 28(8), pp. 2106–2114, doi:10.1016/j.ejc.2007.04.017.

[2] A.E. Frid (2000): *On Combinatorial Properties of fixed points of morphisms.* Ph.D. thesis, Ph.D. thesis, Novosibirsk.

[3] M. A. Makarov (2006): *On permutations generated by infinite binary words. Sib. Èlektron. Mat. Izv.* 3, pp. 304–311 (electronic).

[4] M. A. Makarov (2009): *On permutations generated by Sturmian words. Sibirsk. Mat. Zh.* 50(4), pp. 850–857, doi:10.1007/s11202-009-0076-6.

[5] S Widmer (2011): *Permutation complexity of the Thue-Morse word. Advances in Applied Mathematics* 47(2), pp. 309–329, doi:10.1016/j.aam.2010.08.002.