On permutation complexity of fixed points of uniform binary morphisms

Alexandr Valyuzhenich

Novosibirsk State University,
Novosibirsk, Russia

September 12, 2011

Basic definitions

Let $\omega = \omega_1\omega_2\omega_3\ldots$ be an infinite word where $\omega_i \in \Sigma = \{0, 1\}$. Then ω corresponds to the binary real number $R_\omega(i) = 0.\omega_i\omega_{i+1}\ldots = \sum_{k\geq 0} \omega_{i+k}2^{-(k+1)}$.

A mapping $h : \Sigma^* \longrightarrow \Sigma^*$ is called a morphism if $h(xy) = h(x)h(y)$ for any words $x, y \in \Sigma^*$.

A word ω is a *fixed point* of a morphism φ if $\varphi(\omega) = \omega$.

Example. The word $\omega = \lim\limits_{n\to\infty} \varphi^n(0)$ is the Thue-Morse word when $\varphi(0) = 01, \varphi(1) = 10$

$$\omega_{TM} = 0110100110010110...$$

Ancestors and descendants

$$u' = u'_0 u'_1 \cdots u'_n$$ ancestor

$$\downarrow \varphi$$

$$\varphi(u'_0)\,\varphi(u'_1)\;\cdots\;\varphi(u'_n)$$

$$u = s\varphi(u'_1)\;\cdots\;\varphi(u'_{n-1})p$$ descendant

When the fixed point is *circular*, each sufficiently long word has a unique ancestor.

Basic definitions

An *occurrence* of a word $u \in \Sigma^*$ in the word ω is a pair (u, m) such that $u = \omega_{m+1}\omega_{m+2}\ldots\omega_{m+n}$.

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

Basic definitions

The *infinite permutation* generated by the word ω is the ordered triple $\alpha = \langle \mathbb{N}, <_\alpha, < \rangle$, where $<_\alpha$ and $<$ are linear orders on $\mathbb{N}$. The order $<_\alpha$ is defined as follows: $i <_\alpha j$ if and only if $R_\omega(i) < R_\omega(j)$, and $<$ is the natural order on $\mathbb{N}$.

We define a function $\gamma : \mathbb{R}^2 \setminus \{(a,a) | a \in \mathbb{R}\} \to \{<, >\}$, which for two different real numbers reveals their relation: $\gamma(a,b) = <$ if and only if $a < b$.

Basic definitions

A permutation $\pi = \pi_1 \dots \pi_n$ is a *subpermutation* of length n of an infinite permutation α if $\gamma(\pi_s, \pi_t) = \gamma(R_\omega(i+s), R_\omega(i+t))$ for $1 \leq s < t \leq n$ and for a fixed positive integer i.

$Perm(n)$ is the set of all subpermutations of α_ω of length n.

The *permutation complexity* of a word is $\lambda(n) = |Perm(n)|$.

Basic definitions

We say that an occurrence (u, m) of the word u *generates* a permutation π if π is induced by a sequence of numbers $R_\omega(m+1)\dots R_\omega(m+n)$.

A subword u of the word ω *generates* the permutation π if there is an occurrence (u, m) of this word which generates π. The permutation that is generated by the occurrence of (u, m) is denoted by $\pi(u, m)$.

Example. The subword $u = 010$ of the Thue-Morse word generate permutations 132 and 231, because $\pi(u, 3) = 231$ and $\pi(u, 10) = 132$.

Considered morphisms

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

Example. Each morphism $\varphi(0) = 01^{2n}01^n, \varphi(1) = 10^{2n}10^n$ for $n \geq 2$ belongs to Q.

Example. Morphism $\varphi(0) = 01011, \varphi(1) = 10000$ does not belong to Q.

The properties of Q

Property 1

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

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

Property 2

Let ω be a fixed point of the morphism $\varphi \in Q$. Let $\omega_i = \omega_j$, where $i \equiv i'(mod\ l)$,$j \equiv j'(mod\ l)$ and $0 \leq i', j' \leq l-1$. 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.

Conjugacy of permutations

Let $z = z_1 z_2 \dots z_k$ be a permutation of length k, where $z_i \in \{1, 2, \dots, k\}$.
An *element* of the permutation z is the number z_i, where $1 \leq i \leq k$.

Definition

Permutations $x = x_1 x_2 \dots x_k$ and $y = y_1 y_2 \dots y_k$ are *conjugate* 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 conjugacy by $x \sim y$.

Example. There are exactly two pairs of conjugate permutations among the permutations of length 3: $132 \sim 231$ and $213 \sim 312$.

Bad words

Let u be an arbitrary subword of the word ω, N_u is the set of all pairs of conjugate permutations, and M_u be the set of all remaining permutations generated by u. The number of permutations generated by u is denoted by $f(u)$.

Definition

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 conjugate permutations.

Example. The subword $u = 010$ of the Thue-Morse word is bad, because its occurrences $(u, 3)$ and $(u, 10)$ generate permutations $\pi(u, 3) = 231$ and $\pi(u, 10) = 132$.

The properties of bad words

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 $u \longrightarrow u_1 \longrightarrow u_2 \longrightarrow \ldots \longrightarrow u_m = a$, 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

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

Narrow words

Definition

A word u with $|u| = n \geq L_\omega$ will be called *narrow* if its chain of ancestors is $u = u_0 \to \ldots \to u_{p-1} \to u_p \to \ldots \to u_m = a$, where $a \in A$, u_p is a bad word and $|u_{p-1}| < (|u_p| - 1)l + 1$ for some $p \in \{1, \ldots, m\}$.

Example. The subword $u = 1100$ of the Thue-Morse word is narrow, because its chain of ancestors is $u \to u' = 010$, u' is a bad word and $|u| < 2(|u'| - 1) + 1 = 5$.

The properties of narrow words

The set of narrow words of length n whose chain of ancestors is $u \longrightarrow u_1 \longrightarrow u_2 \longrightarrow \ldots \longrightarrow u_m = a$, 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

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

Wide words

Definition

A word u with $|u| = n \geq L_\omega$ will be called *wide* if its chain of ancestors is $u = u_0 \to \ldots \to u_{p-1} \to u_p \to \ldots \to u_m = a$, where $a \in A$, u_p is a bad word and $|u_{p-1}| > (|u_p| - 1)l + 1$ for some $p \in \{1, \ldots, m\}$.

Example. The subword $u = 011001$ of the Thue-Morse word is wide, because its chain of ancestors is $u \to u' = 010$, u' is a bad word and $|u| > 2(|u'| - 1) + 1 = 5$.

The properties of wide words

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

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

Special words

Definition

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

Lemma

Let $u = u_1 \dots u_n$ and $v = v_1 \dots v_n$ 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 equal permutations.

So, two words can generate equal permutations only if they are $v0$ and $v1$ for some special word v. The number of common permutations generated by some occurrences of words $v0$ and $v1$ is denoted by $g(v)$.

Special words

Consider a special word v of length $n-1$. Let a be the first letter of v and $b = \{0,1\} \setminus a$.

Definitions

- Let k_v be the number of permutations of M_{va} which also belong to H_{vb}.
- Let t_v be the number of permutations of M_{va} each of which is conjugate to some permutation of H_{vb}.
- Let r_v be the number of permutations of N_{va} which also belong to H_{vb}.

Examples

Example. For the subword $u = 010$ of the Thue-Morse word $k_{010} = t_{010} = r_{010} = 0$, because the words 0101 and 0100 generate different nonconjugate permutations 1324 and 3421.

Example. For the subword $u = 01$ of the Thue-Morse word $k_{01} = t_{01} = 0$ and $r_{01} = 1$, because 010 generate conjugate permutations 132 and 231, and 011 generate permutation 132.

Algorithm for finding $g(v)$

The set of all special words of length less than L_ω, with the special descendants of the length greater than L_ω is denoted by Z.

Definition

The set of all special subwords of length n of the word ω, whose chain of ancestors is $v \to v_1 \to v_2 \to \ldots \to v_m = z$ and $z \in Z$, is denoted by $B_z(n)$. The cardinality of the set $B_z(n)$ is denoted by $S_z(n)$.

Lemma

Let $v \in B_z(n-1)$ with $n \geq L_\omega + 1$. Then the following statements are true:

- If $n \neq l^p|z| + 1$ for any positive integer p, then $g(v) = k_z + t_z + r_z$.
- If $n = l^p|z| + 1$ for some positive integer k, then $g(v) = k_z + r_z$.

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

Theorem

$$\sum_{v \in B(n-1)} g(v) = \sum_{z \in Z} [S_z(n-1)(k_z + t_z + r_z)(1 - \delta(n,z)) + (k_z + r_z)\delta(n,z)].$$

The main theorem

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 $A_2 = A \setminus A_1$.

Theorem

Let ω be a fixed point of the morphism φ, where $\varphi \in Q$. Then the permutation complexity of ω is calculated as follows:
$\lambda(n) = \sum_{a_1 \in A_1} [C_{a_1}^{nar}(n)(m_{a_1} + n_{a_1}) + (C_{a_1}^{bad}(n) + C_{a_1}^{wide}(n))(m_{a_1} + 2n_{a_1})] + \sum_{a_2 \in A_2} C_{a_2}(n) m_{a_2} - \sum_{z \in Z} [S_z(n-1)(k_z + t_z + r_z)(1 - \delta(n,z)) + (k_z + r_z)\delta(n,z)]$ for $n \geq L_\omega$.

Recurrence relations

- $C_a^{bad}(xl+1) = lC_a^{bad}(x+1)$ and $C_a^{bad}(xl+r) = 0$ for $r \neq 1$.
- $C_a^{nar}(xl+r) =$ $(r-1)C_a^{nar}(x+2)+(r-1)C_a^{bad}(x+2)+(l-r+1)C_a^{nar}(x+1)$ for $r \geq 1$.
- $C_a^{nar}(xl) = (l-1)C_a^{nar}(x+1)+(l-1)C_a^{bad}(x+1)+C_a^{nar}(x)$.
- $C_a^{wide}(xl+r) =$ $(r-1)C_a^{wide}(x+2)+(l-r+1)C_a^{wide}(x+1)+(l-r+1)C_a^{bad}(x+1)$ for $r \geq 2$.
- $C_a^{wide}(xl+1) = lC_a^{wide}(x+1)$.
- $C_a^{wide}(xl) = (l-1)C_a^{wide}(x+1)+C_a^{wide}(x)+C_a^{bad}(x)$.
- $C_a(xl+r) = (r-1)C_a(x+2)+(l-r+1)C_a(x+1)$ for $r \geq 1$.
- $C_a(xl) = (l-1)C_a(x+1)+C_a(x)$.

It is easy to see that $A_1 = \{010, 101\}$ and $A_2 = \{00, 01, 10, 11, 001, 011, 100, 101\}$.

We note that $Z = \{01, 10, 010, 101\}$.

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