Low maximal pattern complexity of infinite permutations

Theoretical Computer Science 412 (2011) pp.2911-2921

S. V. Avgustinovich[*], A. Frid[*], T. Kamae[†], P. Salimov[‡]

August 13, 2013

Abstract

An infinite permutation α is a linear ordering of $\mathbb{N}$. We study properties of infinite permutations analogous to those of infinite words and showing some resemblance and some difference between permutations and words. In this paper, we define maximal pattern complexity $p^*_\alpha(n)$ for infinite permutations and show that this complexity function is ultimately constant if and only if the permutation is ultimately periodic; otherwise its maximal pattern complexity is at least n, and the value $p^*_\alpha(n) \equiv n$ is reached on a large family of permutations constructed with the use of Sturmian words. We also conjecture that there are no other infinite permutations of maximal pattern complexity equal to n.

1 Infinite permutations

Let S be a finite or countable ordered set: we shall consider S equal either to $\mathbb{N}$, or to some finite subset of $\mathbb{N}$, where $\mathbb{N} = \{0, 1, 2, \cdots\}$. Let $\mathcal{A}_S$ be the set of all sequences of pairwise distinct reals defined on S. Define an equivalence relation $\sim$ on $\mathcal{A}_S$ as follows: let $a, b \in \mathcal{A}_S$, where $a = \{a_s\}_{s\in S}$ and $b = \{b_s\}_{s\in S}$; then $a \sim b$ if and only if for all $s, r \in S$ the inequalities $a_s < a_r$ and $b_s < b_r$ hold or do not hold simultaneously. An equivalence class from $\mathcal{A}_S/\sim$ is called an *(S-)permutation*. If an S-permutation α is realized by a sequence of reals a, we denote it by $\alpha = \overline{a}$. In particular, a $\{1, \dots, n\}$-permutation always has a representative with all values in $\{1, \dots, n\}$, i. e., can be identified with a usual permutation from S_n.

In equivalent terms, a permutation can be considered as a linear ordering of S which may differ from the "natural" one. That is, for $i, j \in S$, the natural order between them corresponds to $i < j$ or $i > j$, while the ordering

[*]Sobolev Institute of Mathematics SB RAS Koptyug av., 4, 630090 Novosibirsk, Russia
[†]Satakedai 5-9-6, 565-0855, Japan
[‡]Novosibirsk State University, 2 Pirogova st., 630090 Novosibirsk, Russia

we intend to define corresponds to $\alpha_i < \alpha_j$ or $\alpha_i > \alpha_j$. We shall also use the symbols $\gamma_{ij} \in \{<, >\}$ meaning the relations between α_i and α_j, so that we by definition have $\alpha_i \gamma_{ij} \alpha_j$ for all $i \neq j$.

We are interested in properties of infinite permutations analogous to those of infinite words, for example, periodicity and complexity. A permutation $\alpha = \{\alpha_s\}_{s \in S}$ is called *t-periodic* if for all i and j such that $i, j, i+t, j+t \in S$ we have $\gamma_{ij} = \gamma_{i+t,j+t}$. An $\mathbb{N}$-permutation is called *ultimately t-periodic* if these equalities hold provided that $i, j > n_0$ for some n_0. This definition is analogous to that for words: an infinite word $w = w_1 w_2 \cdots$ on an alphabet Σ is t-periodic if $w_i = w_{i+t}$ for all i and is ultimately t-periodic if $w_i = w_{i+t}$ for all $i \geq n_0$ for some n_0.

In the previous paper by Fon-Der-Flaass and Frid [4], all periodic $\mathbb{N}$-permutations have been characterized; in particular, it has been shown that there exists an infinite number of distinct t-periodic permutations for each $t \geq 2$. For example, for each n the permutation with a representative sequence

$$-1,\ 2n-2,\ 1,\ 2n,\ 3,\ 2n+2, \dots$$

is 2-periodic, and all such permutations are distinct. So, the situation with periodicity differs from that for words, since the number of distinct t-periodic words on a finite alphabet of cardinality q is clearly finite (and is equal to q^t).

A set $T = \{0, m_1, \dots, m_{k-1}\}$ of cardinality k, where $0 = m_0 < m_1 < \cdots < m_{k-1}$, is called a ($k$-)window. It is natural to define *T-factors* of an infinite permutation α as projections of α to $T+n$, $n \in \mathbb{N}$, considered as permutations on T. Such a projection is denoted by $\alpha_{T+n} = \alpha_n \alpha_{n+m_1} \cdots \alpha_{n+m_{k-1}}$. We call the number of distinct T-factors of α the *T-complexity* of α and denote it by $p_\alpha(T)$.

In particular, if $T = \{0, 1, 2, \dots, k-1\}$, then T-factors of α are called just *factors* of α and are analogous to factors (or subwords) of infinite words. They are denoted by $\alpha_{[i..i+k)} = \alpha_i \alpha_{i+1} \cdots \alpha_{i+k-1}$, and their number is called the *factor complexity* $f_\alpha(n)$ of α. This function is analogous to the subword complexity $f_w(n)$ of infinite words which is equal to the number of different words $w_{[i..i+n)}$ of length n occurring in an infinite word w (see [3] for a survey). However, not all the properties of these two functions are similar [4]. Consider in particular the following classical lemma.

Theorem 1 *An infinite word w is ultimately periodic if and only if $f_w(n) = C$ for some constant C and all sufficiently large n. If w is not ultimately periodic, then $f_w(n)$ is strictly growing and fits $f_w(n) \geq n+1$.*

Only the first statement of Theorem 1 has an analogue for permutations; as for the second one, the situation with permutations is completely different.

Theorem 2 *[4] Let α be an $\mathbb{N}$-permutation; then $f_\alpha(n) \leq C$ if and only if α is ultimately periodic. At the same time, for each unbounded growing* *function $g(n)$, there exists a $\mathbb{N}$-permutation α with $f_\alpha(n) \leq g(n)$ for all $n \geq N_0$ which is not ultimately periodic.*

The supporting example of a permutation with low complexity can be defined by the inequalities $\alpha_{2n} < \alpha_{2n+2} < \alpha_{2n+1} < \alpha_{2n+3}$ for all $n \geq 0$, and $\alpha_{2n_k} < \alpha_{2k+1} < \alpha_{2n_k+2}$ for some sequence $\{n_k\}_{k=0}^{\infty}$ which grows sufficiently fast.

In this paper we study the properties of another complexity function, namely, *maximal pattern complexity*

$$p^*_\alpha(n) = \max_{\#T=n} p_\alpha(T).$$

The analogous function $p^*_w(n)$ for infinite words was defined in 2002 by Kamae and Zamboni [6] where the following statement was proved:

Theorem 3 *[6] An infinite word w is not ultimately periodic if and only if $p^*_w(n) \geq 2n$ for some n.*

Infinite words of maximal pattern complexity $2n$ include rotation words [6] and also some words built by other techniques [7]. The classification of all words of maximal pattern complexity $2n$ is an open problem [5].

In this paper, we prove analogous results for infinite permutations and state a conjecture that in the case of permutations, lowest maximal pattern complexity is achieved only in the "Sturmian" case.

2 Lowest complexity

First of all, we prove a lower bound for the maximal pattern complexity of a non-periodic infinite permutation.

Theorem 4 *An infinite permutation α is not ultimately periodic if and only if $p^*_\alpha(n) \geq n$ for any n.*

Proof. Clearly, if a permutation is ultimately periodic, its maximal pattern complexity is ultimately constant, and thus the "if" part of the proof is obvious. Now suppose that $p^*_\alpha(l) < l$ for some l; we shall prove that α is ultimately periodic.

Since $p^*_\alpha(1) = 1$ (there is exactly one permutation of length one), and the function p^* is non-decreasing, we see that $p^*_\alpha(l) < l$ implies that $p^*_\alpha(n+1) = p^*_\alpha(n)$ for some $n \leq l$. Consider an n-window $T = (0, m_1, \ldots, m_{n-1})$ such that $p_\alpha(T) = p^*_\alpha(n)$; the equality $p^*_\alpha(n+1) = p_\alpha(T)$ means that for each $T' = (0, m_1, \ldots, m_{n-1}, m_n)$ with $m_n > m_{n-1}$ we have $p_\alpha(T) = p^*_\alpha(T')$, that is, each T-permutation which occurs in α can be extended to a T'-permutation which occurs in α by a unique way. Clearly, there exist two equal factors of length $2m_{n-1}$ in α: say,

$$\alpha_{[k..k+2m_{n-1})} = \alpha_{[k+t..k+t+2m_{n-1})}.$$

We shall prove that α is ultimately t-periodic, namely, that $\gamma_{ij} = \gamma_{i+t,j+t}$ for all i, j with $k \leq i < j$. The proof will use the induction on the pair i, j starting by the pairs i, j with $k \leq i < j < k + 2m_{n-1}$, for which our statement holds since $\alpha_{[k..k+2m_{n-1})} = \alpha_{[k+t..k+t+2m_{n-1})}$.

Now for the induction step: for some $M \geq 2m_{n-1}$, suppose that $\gamma_{ij} = \gamma_{i+t,j+t}$ for all $k \leq i < j < k + M$, that is, $\alpha_{[k..k+M)} = \alpha_{[k+t..k+t+M)}$. We are going to prove that $\gamma_{i,k+M} = \gamma_{i+t,k+t+M}$ for all $i \in \{k, \dots, k + M - 1\}$, and thus $\alpha_{[k..k+M+1)} = \alpha_{[k+t..k+t+M+1)}$.

Indeed, condider the case $i \in \{k, \dots, k + M - m_{n-1} - 1\}$ first. Then α_{T+i} is a T-factor of $\alpha_{[k..k+M)}$ and α_{T+i+t} is a T-factor of $\alpha_{[k+t..k+t+M)}$ standing at the same position. So, these T-factors of α are equal, and due to the choice of T, so are their extensions $\alpha_{T'+i}$ and $\alpha_{T'+i+t}$, where $T' = (0, m_1, \dots, m_{n-1}, M - i)$. In particular, the first and last elements of $\alpha_{T'+i}$ and $\alpha_{T'+i+t}$ are in the same relationship: $\gamma_{i,k+M} = \gamma_{i+t,k+t+M}$, which is what we needed.

Now if $i \in \{k + M - m_{n-1}, \dots, k + M - 1\}$, we consider $\alpha_{T+i-m_{n-1}}$ which is a T-factor of $\alpha_{[k..k+M)}$ with the last element α_i, and $\alpha_{T+i+t-m_{n-1}}$ which is a T-factor of $\alpha_{[k+t..k+t+M)}$ with the last element α_{i+t}. They are equal, and so are their extensions $\alpha_{T'+i-m_{n-1}}$ and $\alpha_{T'+i+t-m_{n-1}}$, where $T' = (0, m_1, \dots, m_{n-1}, M - i + m_{n-1})$. In particular, the next to last and the last elements of these T-permutations are in the same relationship: $\gamma_{i,k+M} = \gamma_{i+t,k+t+M}$ for all $i \in \{k, \dots, k + M - 1\}$; together with the induction hypothesis it means that $\alpha_{[k..k+M+1)} = \alpha_{[k+t..k+t+M+1)}$. Repeating the induction step we get that $\gamma_{ij} = \gamma_{i+t,j+t}$ for all $k \leq i < j$, that is, the permutation α is ultimately t-periodic. □

3 Sturmian permutations

A one-side infinite word $w = w_0 w_1 w_2 \cdots$ on the alphabet $\{0, 1\}$ is called *Sturmian* if its subword complexity $f_w(n)$ is equal to $n+1$ for all n. Sturmian words have a number of equivalent definitions [1]; we shall need two of them. First, Sturmian words are exactly aperiodic *balanced* words which means that for each length n, the number of 1s in factors of w of length n takes only two successive values. Second, Sturmian words are exactly irrational *mechanical* words which means that there exists some irrational $\sigma \in (0, 1)$ and some $\rho \in [0, 1)$ such that for all i we have

$$w_i = \lfloor \sigma(i + 1) + \rho \rfloor - \lfloor \sigma i + \rho \rfloor \text{ or}$$
$$w_i = \lceil \sigma(i + 1) + \rho \rceil - \lceil \sigma i + \rho \rceil.$$

These definitions coincide if $\sigma i + \rho$ is never integer; if it is for some (unique) i, the sequences built by these two formulas differ in at most two successive positions. So, we distinguish *lower* and *upper* Sturmian words according to the choice of $\lfloor \cdot \rfloor$ or $\lceil \cdot \rceil$ in the definition.

Now let us define a *Sturmian permutation* $\alpha(w, x, y) = \alpha = \overline{a}$ associated with a Sturmian word w and positive numbers x and y by its representative sequence a, where a_0 is a real number and for all $i \geq 0$ we have

$$a_{i+1} = \begin{cases} a_i + x, & \text{if } w_i = 0, \\ a_i - y, & \text{if } w_i = 1. \end{cases}$$

Clearly, such a permutation is well-defined if and only if we never have $kx \neq ly$ if k is the number of 0s and l is the number of 1s in some factor of w; and in particular if x and y are rationally independent.

Note that a factor of w of length n corresponds to a factor of α of length $n+1$, and the correspondence is one-to-one. So, we have $f_\alpha(n) = n$ for all n. In fact, we are going to prove that the maximal pattern complexity of α is also equal to n, and thus the lower bound in Theorem 4 is precise.

Theorem 5 *For each Sturmian permutation α we have $p^*_\alpha(n) \equiv n$.*

Proof. Let us start with the situation when $x = \sigma$ and $y = 1 - \sigma$. This case has been proved by M. Makarov in [9], but we give a proof here for the sake of completeness.

If we take $a_0 = \rho$, then by the definition of the Sturmian word, $a_i = \{\sigma i + \rho\}$ holds in the case that w is a lower Sturmian word, and $a_i = 1 - \{1 - \sigma i - \rho\}$ holds in the case that w is an upper Sturmian word. In what follows, we consider lower Sturmian words without loss of generality.

Consider a k-window $T = \{0, m_1, \ldots, m_{k-1}\}$ and the set of T-factors $\alpha_{T+n} = \overline{\{\sigma n + \rho\}, \{\sigma(n + m_1) + \rho\}, \ldots, \{\sigma(n + m_{k-1}) + \rho\}}$ for all n. Since the set of $\{\sigma n + \rho\}$, $n \in \mathbb{N}$, is dense in $[0, 1]$, the set of T-factors is equal to the set of all permutations $\overline{t, \{t + \sigma m_1\}, \ldots, \{t + \sigma m_{k-1}\}}$ with $t \in [0, 1]$

Let us arrange the points $\{t + \sigma m_i\}$ $(i = 0, \ldots, k-1)$ on the unit circle, that is the interval $[0, 1]$ with the points 0 and 1 identified (recall that $m_0 = 0$ by definition). Then, the arrangement partitions the unit circle into k arcs. Since the arrangements for different t's are different only by rotations, the permutation defined by the points is determined by the arc containing the point $0 = 1$. Since the number of arcs is k, there are exactly k different permutations defined by the points $\{t + \sigma m_i\}$ $(i = 0, \ldots, k-1)$ with different t's. Thus, $p_\alpha(T) = k$. Since the window T was chosen to be arbitrary, we have $p^*_\alpha(k) = k$.

Now consider the general case of arbitrary x and y. Let us keep the notation γ_{ij} for the relation between $\alpha(w, \sigma, 1 - \sigma)_i$ and $\alpha(w, \sigma, 1 - \sigma)_j$, and denote the relation between $\alpha(w, x, y)_i$ and $\alpha(w, x, y)_j$ by δ_{ij}.

Recall that the *weight* of a binary word is the number of 1's in it. Note that by the definition of α, we have $\delta_{i,i+n} = \delta_{j,j+n}$ if $w_{[i..i+n)}$ and $w_{[j..j+n)}$ have the same weight. Note also that the weight of a factor of w of length n is either equal to $\lfloor n\sigma \rfloor$ or to $\lceil n\sigma \rceil$. So, in $\alpha(w, \sigma, 1-\sigma)$, words $w_{[i..i+n)}$ and $w_{[j..j+n)}$ of the same length n but of different weight always correspond to $\gamma_{i,i+n} \neq \gamma_{j,j+n}$, since $(n - \lfloor n\sigma \rfloor)\sigma - \lfloor n\sigma \rfloor(1-\sigma) = n\sigma - \lfloor n\sigma \rfloor > 0$ and $(n - \lceil n\sigma \rceil)\sigma - \lceil n\sigma \rceil(1-\sigma) = n\sigma - \lceil n\sigma \rceil < 0$.

Now let us fix an arbitrary k-window $T = \{0 = m_0, m_1, \ldots, m_{k-1}\}$ and two positions i and j such that $\alpha(w, \sigma, 1-\sigma)_{T+i} = \alpha(w, \sigma, 1-\sigma)_{T+j}$. Let us prove that $\alpha(w, x, y)_{T+i} = \alpha(w, x, y)_{T+j}$. Indeed, for all $p, r \in \{0, \ldots, k-1\}$ with $p < r$, we have $\gamma_{i+m_p, i+m_r} = \gamma_{j+m_p, j+m_r}$. Due to the arguments above this means that the weight of $w_{[i+m_p..i+m_r)}$ is equal to the weight of $w_{[j+m_p..j+m_r)}$, and thus $\delta_{i+m_p, i+m_r} = \delta_{j+m_p, j+m_r}$. Since a T-permutation is determined by the relations between pairs of its elements, these equalities for all p and r mean that $\alpha(w, x, y)_{T+i} = \alpha(w, x, y)_{T+j}$. So, we have $p_{\alpha(w,x,y)}(T) \leq p_{\alpha(w,\sigma,1-\sigma)}(T)$ and thus $p^*_{\alpha(w,x,y)}(k) \leq p^*_{\alpha(w,\sigma,1-\sigma)}(k) = k$; at the same time, $p^*_{\alpha(w,x,y)}(k) \geq k$ since this permutation is not ultimately periodic. So, $p^*_{\alpha(w,x,y)}(k) = k$, and the theorem is proved. □.

4 Concluding remark

At the moment we conjecture that the described Sturmian permutations are the only permutations of maximal pattern complexity $p^*_\alpha(n) = n$. We hope to prove it is a subsequent work.

References

[1] J. Berstel, P. Séébold, Sturmian words, in: M. Lothaire, Algebraic Combinatorics on Words, Cambridge University Press, 2002. P. 40–97.

[2] J. A. Davis, R. C. Entringer, R. L. Graham, and G. J. Simmons, On permutations containing no long arithmetic progressions, Acta Arithmetica 34 (1977), 81-90.

[3] S. Ferenczi, Complexity of sequences and dynamical systems, Discrete Math. 206 (1999), 145–154.

[4] D. G. Fon-Der-Flaass, A. E. Frid, On periodicity and low complexity of infinite permutations, European Journal of Combinatorics 28 (2007), 2106–2114.

[5] T. Kamae, H. Rao, Bo Tan, Yu-M. Xue, Language Structure of Pattern Sturmian Word, Discrete Mathematics, Discrete Mathematics 306 (2006), 1651–1668.

[6] T. Kamae, L. Zamboni, Sequence entropy and the maximal pattern complexity of infinite words, Ergodic Theory and Dynamical Systems 22 (2002), 1191–1199.

[7] T. Kamae, L. Zamboni, Maximal pattern complexity for discrete systems, Ergodic Theory and Dynamical Systems 22 (2002), 1201–1214.

[8] M. Makarov, On permutations generated by infinite binary words, Siberian Electronic Mathematical Reports 3 (2006) 304–311 [in Russian, English abstract].

[9] M. Makarov, On Sturmian permutations, accepted to Siberian Mathematical Journal [in Russian].

[10] G. Rote, Sequences with subword complexity $2n$, Journal of Number Theory 46 (1993), 196–213.