Sequence entropy and the maximal pattern complexity of infinite words

Ergodic Theory and Dynamical Systems 22-4 (2002), pp.1191-1199

Teturo Kamae* and Luca Zamboni†

Abstract: For an infinite word $\alpha = \alpha_0\alpha_1\alpha_2\cdots$, over a finite alphabet A we define the maximal pattern complexity by

$$p^*_\alpha(k) = \sup_\tau \sharp\{\alpha_{n+\tau(0)}\alpha_{n+\tau(1)}\cdots\alpha_{n+\tau(k-1)};\ n = 0,1,2,\cdots\}$$

where the "sup" is taken over all subsequences $0 = \tau(0) < \tau(1) < \cdots < \tau(k-1)$ of integers of length k. We prove that α is eventually periodic if and only if $p^*_\alpha(k) \le 2k-1$ for some k. Infinite words α with $p^*_\alpha(k) = 2k$ for any k are called pattern Sturmian words and are studied.

1 Introduction

An increasing sequence of integers

$$\tau : 0 = \tau(0) < \tau(1) < \cdots < \tau(k-1)$$

with $k = 1,2,\cdots$ is called a *window* of size k. For $k = 1,2,\cdots$, we denote by $\langle k\rangle$ the window of size k such that

$$\langle k\rangle(i) = i \quad (i = 0,1,\cdots,k-1).$$

*Department of Mathematics, Osaka City University, Osaka, 558-8585 Japan (kamae@sci.osaka-cu.ac.jp)

†Department of Mathematics, P.O.Box 311430, University of North Texas, Denton, TX 76203-1430, USA (luca@unt.edu)

Let $\alpha = \alpha_0\alpha_1\alpha_2\cdots$ be an infinite word over a finite alphabet A with $\sharp A \geq 2$, where $\sharp A$ denotes the number of elements in A. Let τ be a window of size k. We denote by $\alpha[n+\tau]$ the word $\alpha_{n+\tau(0)}\alpha_{n+\tau(1)}\cdots\alpha_{n+\tau(k-1)}$ over A of length k. A finite word $\eta_0\eta_1\cdots\eta_{k-1}$ is called a *τ-factor* of α if $\eta_0\eta_1\cdots\eta_{k-1} = \alpha[n+\tau]$ for some $n = 0,1,2,\cdots$. The set of τ-factors of α is denoted by $F_\alpha(\tau)$. We also denote $F_\alpha(k) := F_\alpha(\langle k\rangle)$.

We define the *complexity* $p_\alpha(k)$ in the usual sense and the *maximal pattern complexity* $p^*_\alpha(k)$ as a function on $k \in \{1,2,3,\cdots\}$ by

$$\begin{aligned} p_\alpha(k) &= \sharp F_\alpha(k) \\ p^*_\alpha(k) &= \sup_\tau \sharp F_\alpha(\tau) \end{aligned}$$

where the "sup" is taken over all windows τ of size k.

For windows τ and τ' of size k and $k+1$, respectively, such that $\tau(i) = \tau'(i)$ for $i = 0,1,\cdots,k-1$, we call τ' an *immediate extension* of τ.

An infinite word α is called *recurrent* if for any $L \geq 1$, there exists $M \geq 1$ such that

$$\alpha_i = \alpha_{i+M} \quad (i = 0,1,\cdots,L-1). \tag{1}$$

It is easy to see that if α is recurent, then there exist infinitely many M's as above.

It is known that α is eventually periodic if and only if $p_\alpha(k) \leq k$ for some k. In this paper, we prove the following theorem.

Theorem 1. *An infinite word α over a finite set A is eventually periodic if and only if $p^*_\alpha(k) \leq 2k-1$ for some $k = 1,2,\cdots$.*

Proof. The "only if" part holds since if α is eventually periodic, then $p^*_\alpha(k)$ is bounded in $k = 1,2,\cdots$.

If $p^*_\alpha(1) = 1$, then α is clearly periodic. Assume that $p^*_\alpha(1) \geq 2$ and $p^*_\alpha(k) \leq 2k-1$ for some $k = 2,3,\cdots$. Then, there exists $k = 1,2,\cdots$ such that $p^*_\alpha(k+1) \leq p^*_\alpha(k)+1$. Let τ be a window of size k such that $\sharp F_\alpha(\tau) = p^*_\alpha(k)$. Then we have

$$\sharp F_\alpha(\tau') \leq p^*_\alpha(k+1) \leq p^*_\alpha(k)+1 = \sharp F_\alpha(\tau)+1$$

for any immediate extension τ' of τ. Thus, our theorem follows from the following Theorem 2. $\square$

Theorem 2. *For an arbitrary infinite word α over a finite set A, if there exists $k = 1, 2, 3, \cdots$ and a window τ of size k such that*

$$\sharp F_\alpha(\tau') \leq \sharp F_\alpha(\tau) + 1$$

holds for any immediate extension τ' of τ, then α is eventually periodic.

The proof differs according to whether α is recurrent (Section 3) or not (Section 4).

An infinite word α with $p_\alpha(k) = k + 1$ for any k is known as a *Sturmian word* and has been studied extensively (Valerie Berthé [1] for a nice recent survey). We prove that for a Sturmian word α, $p^*_\alpha(k) = 2k$ holds for any k. Moreover, any infinite word of the labels given by a partition into 2 nonempty intervals of the circle along an orbit of an irrational rotation has this property (Example 2). An infinite word α with $p^*_\alpha(k) = 2k$ for any k is called a *pattern Sturmian word*. Another family of pattern Sturmian words is given in Example 5.

It follows from Kushnirenko [2] that an infinite word α which induces a dynamical system with a partially continuous spectrum satisfies that $\limsup_{k\to\infty}(1/k)\log p^*_\alpha(k) > 0$ (Corollary 1). In fact, Thue-Morse sequence α satisfies that $p^*_\alpha(k) = 2^k$ (Example 1), while $p_\alpha(k)$ is known to be of linear order in k.

2 Sequence entropy

Let A be a finite set with $\sharp A \geq 2$. Let $\mathbb{N} := \{0, 1, 2, \cdots\}$ and $A^{\mathbb{N}}$ be the product space. Let X_n ($n \in \mathbb{N}$) be the projection $A^{\mathbb{N}} \to A$ defined by $X_n(\alpha) = \alpha(n)$ for any $\alpha \in A^{\mathbb{N}}$. Let T be the shift on the space $A^{\mathbb{N}}$ and μ be a T-invariant probability Borel measure on $A^{\mathbb{N}}$. Let $\tau : 0 = \tau(0) < \tau(1) < \tau(2) < \cdots$ be an infinite sequence of integers. We define a *sequence entropy* $h(\mu,\tau)$ of μ with respect to τ by

$$h(\mu,\tau) := \limsup_{k\to\infty} \frac{1}{k} H(X_{\tau(0)}, X_{\tau(1)}, \cdots, X_{\tau(k-1)})$$

where $X_0,\ X_1,\ X_2,\ \cdots$ are considered as random variables on the probability space $(A^{\mathbb{N}},\mu)$ and $H(\cdots)$ is the Shannon's entropy of random variables.

Theorem 3 (Kushnirenko[2]). *The sequence entropy $h(\mu,\tau)$ is 0 for any τ if and only if the dynamical system $(A^{\mathbb{N}},\mu,T)$ has a discrete spectrum, that is, the eigen-functions with respect to the isometry on $L^2(A^{\mathbb{N}},\mu)$ induced by the shift T span the whole space $L^2(A^{\mathbb{N}},\mu)$.*

Corollary 1. *For $\alpha \in A^{\mathbb{N}}$, assume that*

$$\mu_\alpha := w\text{-}\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i\alpha}$$

exists, where δ_x is the unit measure at $x \in A^{\mathbb{N}}$ and the "w-lim" implies the weak limit on the space of measures. Assume further that the dynamical system $(A^{\mathbb{N}},\mu_\alpha,T)$ has a partially continuous spectrum. That is, it has not a discrete spectrum. Then, we have

$$\limsup_{k\to\infty} (1/k)\log p^*_\alpha(k) > 0.$$

Proof. Assume that $(A^{\mathbb{N}},\mu_\alpha,T)$ has a partially continuous spectrum. We consider X_n's as random variables on the probability space $(A^{\mathbb{N}},\mu_\alpha)$. Let τ be a window of size k. Since the random variable $X_{\tau(0)}X_{\tau(1)}\cdots X_{\tau(k-1)}$ on the space of words over A of length k has a distribution which is supported by $F_\alpha(\tau)$, we have

$$\begin{aligned} h(\mu,\tau) &= \limsup_{k\to\infty} \frac{1}{k} H(X_{\tau(0)}, X_{\tau(1)}, \cdots, X_{\tau(k-1)}) \\ &\le \limsup_{k\to\infty} \frac{1}{k} \log \sharp F_\alpha(\tau) \\ &\le \limsup_{k\to\infty} \frac{1}{k} \log p^*_\alpha(k). \qquad (2) \end{aligned}$$

By Theorem 3, there exists an infinite sequence $\tau : 0 = \tau(0) < \tau(1) < \tau(2) < \cdots$ such that $h(\mu, \tau) > 0$. Therefore, by (2)

$$\limsup_{k\to\infty}(1/k)\log p^*_\alpha(k) > 0.$$

□

Remark 1. The converse of Corollay 1 is not true. In fact, in [4], it is proved that for any irrational number θ, there exists a closed set S in $\mathbb{R}/\mathbb{Z}$ such that the symbolic representation α of the rotation by angle θ with respect to the patition $\{S, S^c\}$ has the maximal pattern complexity $p^*_\alpha(k) = 2^k$ $(k = 1, 2, \cdots)$ for almost all starting point of the rotation.

Example 1. Let $\alpha = 0110100110010110\cdots$ be the Thue-Morse sequence over $A := \{0, 1\}$. Then, we have $p^*_\alpha(k) = 2^k$ for $k = 1, 2, \cdots$.

Proof. Note that $\alpha_n \equiv \xi(n) \pmod 2$, where $\xi(n)$ is the number of digits 1 in the 2-adic representation of n. Let $\tau(i) = 2^{2i} - 1$ $(i = 0, 1, 2, \cdots)$. For any $k = 1, 2, \cdots$, let U_k be the set of $u := \sum_{i=0}^{k} u_i 2^{2i} + 1$ with $(u_0, u_1, \cdots, u_k) \in \{0, 1\}^{k+1}$ satisfying that $\sum_{i=0}^{k} u_i \equiv 1 \pmod 2$. Then for any $u \in U_k$ and $i < k$, we have

$$\xi(u + \tau(i)) \equiv u_i \pmod 2.$$

Hence we have

$$\alpha_{u+\tau(0)}\alpha_{u+\tau(1)}\cdots\alpha_{u+\tau(k-1)} = u_0 u_1 \cdots u_{k-1}.$$

Since $\{u_0 u_1 \cdots u_{k-1};\ u \in U_k\} = \{0, 1\}^k$, we have

$$p^*_\alpha(k) = \sharp\{u_0 u_1 \cdots u_{k-1};\ u \in U_k\} = 2^k.$$

□

3 Recurrent case

We prove Theorem 2 in the case that α is recurrent. Let α be a recurrent infinite word over a finite set A. Assume that there exist $k = 1, 2, \cdots$ and a window τ of size k such that

$$\sharp F_\alpha(\tau') \le \sharp F_\alpha(\tau) + 1$$

holds for any immediate extension τ' of τ.

If $\sharp F_\alpha(1) = 1$, then α is clearly periodic. Therefore, assume that $\sharp F_\alpha(1) \geq 2$.

Suppose that α is not eventually periodic. Take L such that

$$\{\alpha[n+\tau];\ n = 0, 1, \cdots, L-1\} = F_\alpha(\tau). \tag{3}$$

Since α is recurrent, there exists M with $M > \tau(k-1)$ such that (1) holds for this L. Define a window τ' of size $k+1$ which is an immediate extension of τ by

$$\tau'(i) = \begin{cases} \tau(i) & (i = 0, 1, \cdots, k-1) \\ M & (i = k). \end{cases}$$

Then, for any n with $0 \leq n \leq L-1$ we have

$$\alpha_{n+\tau'(k)} = \alpha_{n+M} = \alpha_n = \alpha_{n+\tau'(0)}. \tag{4}$$

For each $b \in F_\alpha(1)$, let

$$G_b = \{\xi_1\xi_2\cdots\xi_{k-1};\ b\xi_1\xi_2\cdots\xi_{k-1} \in F_\alpha(\tau)\}$$

Then by (3) and (4), we have

$$\bigcup_{b\in F_\alpha(1)} bG_b b \subset F_\alpha(\tau').$$

Supose that

$$F_\alpha(\tau') = \bigcup_{b\in F_\alpha(1)} bG_b b.$$

Then, the first and last letter of each word in $F_\alpha(\tau')$ are equal. Therefore, for any $n \in \mathbb{N}$, we have $\alpha_n = \alpha_{n+M}$ since $\tau'(0) = 0$ and $\tau'(k) = M$. Thus, α is periodic, contradicting with our supposition.

Suppose that

$$F_\alpha(\tau') = \bigcup_{b\in F_\alpha(1)} bG_b b \cup \{\xi\}$$

with some word ξ of length $k+1$ beginning by letter c and ending by letter d. If $c = d$, we have the same contradiction as above. Assume that $c \neq d$. Then, the possible combinations of $\alpha_n \alpha_{n+M}$ for each $n \in \mathbb{N}$ are limited to bb for each $b \in F_\alpha(1)$ or cd. For $n \in \mathbb{N}$ with $0 \le n < M$, consider the infinite word $\alpha_n \alpha_{n+M} \alpha_{n+2M} \cdots$. Then, we have only the following possibilities for it:

$$\begin{array}{ll} bbbb\cdots & (b \in F_\alpha(1)) \\ \underbrace{c\cdots c}_{m} ddd\cdots & (m \in \mathbb{N}). \end{array}$$

Therefore, for each $n \in \mathbb{N}$ with $0 \le n < M$, there exists $m(n)$ such that $\alpha_{i+M} = \alpha_i$ for any i with $i \equiv n \pmod{M}$ and $i \ge m(n)$. Let $m_0 = \max_{0 \le n < M} m(n)$. Then, $\alpha_{i+M} = \alpha_i$ holds for any $i \ge m_0$, which implies that α is eventually periodic, contradicting with our supposition.

Thus, $F_\alpha(\tau')$ must contain at least 2 more elements which are not contained in $\bigcup_{b \in F_\alpha(1)} bG_b b$. Since $\bigcup_{b \in F_\alpha(1)} bG_b = F_\alpha(\tau)$, we have $\sharp F_\alpha(\tau') \ge \sharp F_\alpha(\tau) + 2$, contradicting with our assumption which completes the proof of Theorem 2 in the case that α is recurrent.

4 Nonrecurrent case

Now assume that α is not recurrent. We may further assume that the shift of α, $T^n\alpha = \alpha_n \alpha_{n+1} \alpha_{n+2} \cdots$ is not recurrent for any $n \in \mathbb{N}$.

This is because if $T^n\alpha$ is recurrent for some $n \in \mathbb{N}$ and α is not eventually periodic, then it is already prove that for any window τ of size k, there exists τ' which is an immediate extension of τ such that

$$\sharp F_{T^n\alpha}(\tau') \ge \sharp F_{T^n\alpha}(\tau) + 2,$$

which implies that

$$\sharp F_\alpha(\tau') \ge \sharp F_\alpha(\tau) + 2$$

since in the following commutative diagram, the projections π, φ are surjective while the inclusions i and j are injective.

$$\begin{array}{ccc} F_\alpha(\tau') & \xrightarrow{\pi} & F_\alpha(\tau) \\ i\uparrow & & \uparrow j \\ F_{T^n\alpha}(\tau') & \xrightarrow{\varphi} & F_{T^n\alpha}(\tau) \end{array}$$

Assume that
(i) $T^n\alpha$ *is not recurrent for any* $n \in \mathbb{N}$, and that
(ii) *there exists* $k = 1, 2, 3, \cdots$ *and a window* τ *of size* k *such that*

$$\sharp F_\alpha(\tau') \leq \sharp F_\alpha(\tau) + 1$$

holds for any immediate extension τ' *of* τ.
We claim that α is eventually periodic from these assumptions. Suppose to the contrary that α is not eventually periodic; we shall derive a contradiction.

Take any K with $K > 2\tau(k-1)$.

Since α is not eventually periodic, there exists at least one element $\xi \in F_\alpha(K)$ having more than one extension in $F_\alpha(K+1)$.

Suppose that there exists $\xi = \xi_0\xi_1 \cdots \xi_{K-1} \in F_\alpha(K)$ such that ξ has more than 2 extensions in $F_\alpha(K+1)$. Define a window τ' which is an immediate extension of τ with $\tau'(k) = K$. Then, $\xi_{\tau(0)}\xi_{\tau(1)} \cdots \xi_{\tau(k-1)} \in F_\alpha(\tau)$ has more than 2 extensions in $F_\alpha(\tau')$ which implies that $\sharp F_\alpha(\tau') \geq \sharp F_\alpha(\tau) + 2$, contradicting with our assumption (ii).

Suppose that there exists $\xi, \eta \in F_\alpha(K)$ with $\xi \neq \eta$ such that both of ξ and η have 2 extensions in $F_\alpha(K+1)$. Then, there exists $n \in \mathbb{N}$ with $0 \leq n < K - \tau(k-1)$ such that $\xi[n+\tau] \neq \eta[n+\tau]$. Define a window τ' which is an immediate extension of τ by $\tau'(k) = K - n$. Since both of $\xi[n+\tau]$ and $\eta[n+\tau]$ have 2 extensions in $F_\alpha(\tau')$, we have $\sharp F_\alpha(\tau') \geq \sharp F_\alpha(\tau) + 2$, contradicting with our assumption (ii).

Thus, we have

Lemma 1. *For any* $K > 2\tau(k-1)$, *there exists a unique element in* $F_\alpha(K)$ *which has exactly 2 extensions in* $F_\alpha(K+1)$ *while any other element in* $F_\alpha(K)$ *has a unique extension in* $F_\alpha(K+1)$.

Fix $K > 2\tau(k-1)$. There exists a $\xi \in F_\alpha(K)$ which occurs at least twice in α. Hence, there exist $m > n \geq 0$ such that both $T^m\alpha$ and $T^n\alpha$ begins in ξ. Since $T^n\alpha$ is not recurrent, there exists a shortest prefix u' of $T^n\alpha$ which occurs only once in $T^n\alpha$. Since ξ is a prefix of $T^n\alpha$ which occurs twice in $T^n\alpha$, it follows that the length of u' is larger than K, the length of ξ. Let $L+1$ be the length of u'. Then, $L \geq K > 2\tau(k-1)$ and by the minimality of the length of u', the prefix u of u' of length L occurs more than once in $T^n\alpha$. But since u' occurs only once, it follows that there exists letters $a \neq b$ such that ua and ub are in $F_\alpha(L+1)$. Thus by Lemma 1, $u \in F_\alpha(L)$ has exactly two extenions in $F_\alpha(L+1)$ (namely, ua and ub one of which is u'), and all other elements in $F_\alpha(L)$ have a unique extension in $F_\alpha(L+1)$. Since u' does not occur in $T^{n+1}\alpha$, all elements in $F_{T^{n+1}\alpha}(L)$ have a unique extension in $F_{T^{n+1}\alpha}(L+1)$, which implies that $T^{n+1}\alpha$ is eventually periodic. Thus, we have a contradiction that α is eventually periodic, which completes the proof.

5 Examples

Example 2. Let θ be an irrational number and $[a, b)$ be an interval with $0 < b - a < 1$. Define an infinite word $\alpha = \alpha_0\alpha_1\alpha_2\cdots$ over $\{0, 1\}$ by

$$\alpha_n = \begin{cases} 1 & n\theta \in [a,b) \pmod 1 \\ 0 & n\theta \notin [a,b) \pmod 1. \end{cases}$$

Then, we have $p^*_\alpha(k) = 2k$ for any $k = 1, 2, 3, \cdots$. Note that α is recurrent.

Proof. We consider $I_1 := [a, b)$ as a subset in the torus $\mathbb{R}/\mathbb{Z}$ and I_0 be its complement in the torus. Thus, $\mathcal{I} := \{I_0, I_1\}$ is a partition of the torus. Take an arbtrary window $\tau:\ 0 = \tau(0) < \tau(1) < \cdots < \tau(k-1)$ of size $k \geq 1$ and define the points on the torus :

$$\begin{array}{l} a - \tau(0),\ a - \tau(1)\theta,\ a - \tau(2)\theta,\ \cdots,\ a - \tau(k-1)\theta \\ b - \tau(0),\ b - \tau(1)\theta,\ b - \tau(2)\theta,\ \cdots,\ b - \tau(k-1)\theta \end{array}$$

considered in modulo 1. Let K be the number of different points in this list. Then, K is the number of nonempty elements in the partition

$$(\mathcal{I} - \tau(0)\theta) \vee (\mathcal{I} - \tau(1)\theta) \vee \cdots \vee (\mathcal{I} - \tau(k-1)\theta),$$

where "$\vee$" implies the common refinement of the partitions, if each element in the partition is connected, which we assume since otherwise, the number of non-empty elements in the partition is less than K. If $n\theta$ is in a element of this partition, say

$$n\theta \in (I_{c_0} - \tau(0)\theta) \cap (I_{c_1} - \tau(1)\theta) \cap \cdots \cap (I_{c_{k-1}} - \tau(k-1)\theta) \pmod 1,$$

then we have $\alpha_{n+\tau(i)} = c_i$ $(i = 0, 1, \cdots, k-1)$. Since the set of $\{n\theta;\ n = 0, 1, 2, \cdots\}$ is dense in the torus, this implies that

$$K = \sharp\{\alpha_{n+\tau(0)}\alpha_{n+\tau(1)} \cdots \alpha_{n+\tau(k-1)};\ n \in \mathbb{N}\}.$$

Since $K \leq 2k$ for any τ and $K = 2k$ for some τ, we have $p^*_\alpha(k) = 2k$ for any $k = 1, 2, 3, \cdots$. It is clear that α is recurrent. □

Example 3. Let α be a Sturmian word over $\{0, 1\}$. Then, it is known [1] that it is represented as Example 2 with $b - a = \theta \pmod 1$. Thus, α is recurrent and $p^*_\alpha(k) = 2k$ for any $k = 1, 2, 3, \cdots$.

Example 4. (Yu-Mei Xue [3]) Let θ and η be rationally independent irrational numbers. Consider the rotation $R : (\mathbb{R}/\mathbb{Z})^2 \to (\mathbb{R}/\mathbb{Z})^2$ by $(x, y) \mapsto (x + \theta, y + \eta)$. Let S be a closed convex subset of $(0, 1)^2$ with diameter less than $1/2$. Let α be the infinite word over $\{0, 1\}$ defined by

$$\alpha_n = \begin{cases} 1 & (n\theta, n\eta) \in S \\ 0 & (n\theta, n\eta) \notin S. \end{cases}$$

Then, $p^*_\alpha(k) \leq k^2 - k + 2$ $(k = 1, 2, 3, \cdots)$. Moreover, if S is a closed ball with diameter less than $1/2$ and larger than 0, then $p^*_\alpha(k) = k^2 - k + 2$ $(k = 1, 2, 3, \cdots)$.

Example 5. Let $0 < c_1 < c_2 < c_3 < \cdots$ be a sequence of integers such that $c_{k+1} > 2c_k$ $(k = 1, 2, 3, \cdots)$. Define an infinite word α over $\{0, 1\}$ by

$$\alpha_n = \begin{cases} 1 & n = c_k \text{ for some } k = 1, 2, \cdots \\ 0 & \text{otherwise.} \end{cases}$$

Then, we have $p^*_\alpha(k) = 2k$ $(k = 1, 2, 3, \cdots)$. Note that α is not recurrent.

Proof. For any $k = 1, 2, 3, \cdots$ and any window $\tau : \ 0 = \tau(0) < \tau(1) < \tau(2) < \cdots < \tau(k-1)$ of size k, it is clear that any word over $\{0, 1\}$ with length k which contains 1 at most once belongs to $F_\alpha(\tau)$. There are $k+1$ such words.

Let Λ_k be the set of words over $\{0, 1\}$ of length k containing 1 at least twice belonging to $F_\alpha(\tau)$. For $\eta := \eta_0\eta_1 \cdots \eta_{k-1} \in \Lambda_k$, let $\phi(\eta)$ be the maximum j such that $\eta_j = 1$. Let $\eta \in \Lambda_k$ satisfy that $\phi(\eta) = j$. Then, $\eta_j = 1$ and there exists i with $0 \le i < j$ such that $\eta_i = 1$. Therefore, there exists $n \in \mathbb{N}$ such that $n + \tau(i) = c_a, \ n + \tau(j) = c_b$ for some $a, \ b = 1, 2, 3, \cdots$. Let d be the minimum d such that $\tau(j) \le c_d$.

We prove that $b = d$. Suppose that this is not the case. Since $\tau(j) \le c_b$ and $d \ne b$, $b \ge d + 1$ holds. Since $c_a - \tau(i) = c_b - \tau(j)$ and $\tau(j) > \tau(i)$, we have $c_a < c_b$. On the other hand, since

$$c_a = c_b - \tau(j) + \tau(i) \ge c_b - c_d \ge c_b - c_{b-1} > c_{b-1}.$$

we have $c_a \ge c_b$, which contradicts with $c_a < c_b$.

Thus, we have $b = d$. This implies that $n = c_d - \tau(j)$ and n is determined by j. Thus, $\eta = \alpha_{n+\tau(0)}\alpha_{n+\tau(1)} \cdots \alpha_{n+\tau(k-1)}$ holds and η is determined by $\phi(\eta)$. This implies that Λ_k has at most $k-1$ elements, since the image of ϕ is contained $\{1, 2, \cdots, k-1\}$. Thus, the number of words with length k which appears in α through τ is at most $k + 1 + k - 1 = 2k$.

For any $k = 1, 2, 3, \cdots$, define a window $\tau : \ 0 = \tau(0) < \tau(1) < \tau(2) < \cdots < \tau(k-1)$ of size k inductively by

$$\tau(0) = 0, \quad \tau(1) = c_1$$
$$\tau(i+1) = c_{2i+1} - c_{2i} + \tau(i) \quad (i = 1, 2, \cdots, k-2)$$

Then, Λ_k defined above for this τ consists of $k-1$ elements :

$$\begin{array}{rcl} 1\,1\,0\,0\,0 & \cdots & 0\,0\,0 \\ 1\,1\,0\,0 & \cdots & 0\,0\,0 \\ 1\,1\,0 & \cdots & 0\,0\,0 \\ 1\,1 & \cdots & 0\,0\,0 \\ & & \cdots \\ * & & 1\,1\,0 \\ & & 1\,1 \end{array}$$

Thus, we have $\sharp F_\alpha(k) = k+1+k-1 = 2k$. This completes the proof that $p^*_\alpha(k) = 2k$ $(k = 1, 2, 3, \cdots)$. □

There are unsolved problems:

Problem 1: What is the general structure of recurrent pattern Sturmian words?

Problem 2: Is it true that an exponential growth in maximal pattern complexity of an infinite word over 2 letters forces the full maximal pattern complexity 2^k?

Acknowledgment: The author would like to thank Prof. Wen Zhi-Ying (Beijing), Prof. Wu Jun (Wuhan), Prof. Sébastien Ferenczi (Marseille) and Prof. Michel Mendès-France (Bordeaux) for useful discussions on the matter.

The authors were informed by Prof. A. Restivo that the term "pattern complexity" is used by A. Restivo and S. Salemi [5] in a different sense. Since it was after sending the final version, the change of terminology was impossible. The authors apologize for the potential confusion and thank Prof. A. Restivo for communicating to them.

References

[1] Valerie Berthé, Sequences of low complexity: automatic and Sturmian sequences, London Math. Soc. Lecture Note Ser. 279, Cambridge Univ. Press, 2000 pp.1-34

[2] A.G. Kushnirenko, On metric invariants of entropy type, Russian Math. Surveys 22-5 (1967), pp.53-61

[3] Yu-Mei Xue, Transformations with discrete spectrum and sequence entropy, Master Thesis at Osaka City University, 2000 (in Japanese)

[4] Teturo Kamae & Luca Zamboni, Maximal pattern complexity for discrete systems, *Ergodic Theory and Dynamical Systems* 22-4 (2002), pp.1201-1214.

[5] A. Restivo and S. Salemi, Binary pattern in infinite binary words, in Formal and Natural Computing, Lecture Notes in Computer Science vol. 2300, Springer, 2002