数理解析研究所講究録
1166 巻 2000 年 27-34

A Word Length Controlled DT0L System with a Periodic Control Function

Takashi Asayama (朝山　隆志)
Taishin Yasunobu Nishida (西田　泰伸)*
Faculty of Engineering, Toyama Prefectural University,
Kosugi-machi, 939-0398 Toyama, Japan

Abstract

We have introduced a new controlled DT0L system, called a word length controlled DT0L system, or a wlcDT0L system for short. A wlcDT0L system is a DT0L system with a control function which maps from the set of nonnegative integers to the set of tables. A wlcDT0L system derives exactly one word from a given word by iterating the table which is the value of the control function of the length of the given word. Thus a wlcDT0L system generates a sequence of words which starts from the axiom. In this paper we prove that a wlcDT0L system with a periodic control function generates a finite combination of D0L sequences.

1 Introduction

In this paper we will discuss farther property of word length controlled (or wlc for short) DT0L systems which have been first introduced in [5]. We prove that a wlcDT0L system with a periodic control function generates a finite combination of D0L sequences.

We briefly explain the motivation of introducing the wlcDT0L systems. M. Andraşiu, *et al.* have suggested the idea that a slender languages, which has at most k words of the same length for some nonnegative integer k, can be used as a key of a cryptosystem [1]. Then many researchers have investigated slenderness condition of known language families eagerly. Slender

*Corresponding author. Email: nishida@pu-toyama.ac.jp

context-free languages are characterized by L. Ilie [3, 4]. Slender context-free languages have the form $\bigcup_{\text{finite}}\{uv^iwx^iy \mid i \geq 0\}$ and are easily inferable [11]. G. Păun and A. Salomaa have proved that all D0L languages are slender [8]. There are some other subfamilies of 0L languages which consist of slender languages only [6, 7, 2]. But all known slender languages in these subfamilies have periodic structures, which are fatal weakness as a key of cryptosystems. So we seek new language families which contain complex, preferably like random sequences, languages. The wlcDT0L systems are good candidates for keys of cryptosystems.

2 Preliminaries

Let Σ be a finite alphabet. The element of Σ is called a letter. The set of all finite words over Σ including the empty word λ is denoted by Σ^*. For a word $w \in \Sigma^*$, the length of w is denoted by $|w|$. Let a be a letter in Σ. We denote by $|w|_a$ the number of occurrences of a in w.

Let $\Sigma = \{a_1, \ldots, a_n\}$ be a finite alphabet and let $w \in \Sigma^*$. The *Parikh vector* π of w is an n-dimensional vector given by

$$\pi = (|w|_{a_1}, \ldots, |w|_{a_n}).$$

Let S be an arbitrary set. The cardinality of S is denoted by $\mathrm{card}(S)$.

We denote by $\mathbb{N}$ the set of nonnegative integers and $\mathbb{N}_+$ the set of positive integers.

Let Σ and Γ be finite alphabets. A mapping h from Σ^* to Γ^* is said to be a morphism if h satisfies

$$h(uv) = h(u)h(v)$$

for every $u, v \in \Sigma^*$. A morphism from Σ^* to Σ^* is called a morphism over Σ. A morphism h is said to be λ-free if for every $a \in \Sigma$, $h(a) \neq \lambda$. Let h be a morphism over Σ. For every $n \in \mathbb{N}$ and $w \in \Sigma^*$, h^n is defined by

$$\begin{aligned} h^0(w) &= w \quad \text{and} \\ h^n(w) &= h(h^{n-1}(w)) \quad \text{for } n > 0. \end{aligned}$$

A triplet $G = \langle \Sigma, h, w \rangle$ is said to be a D0L system if Σ is a finite alphabet, h is a morphism over Σ, and $w \in \Sigma^*$. A D0L system G generates a sequence of words (w_i) where $w_i = h^i(w)$. A D0L system is called a PD0L system if h is λ-free.

A triplet $G = \langle \Sigma, \Pi, w \rangle$ is said to be a DT0L system if Σ is a finite alphabet, Π is a finite set of morphisms over Σ, and $w \in \Sigma^*$. A DT0L system G generates a set of words W_k in k steps for $k \in \mathbb{N}$ as follows:

$$W_k = \begin{cases} \{w\} & \text{if } k = 0 \\ \{u \mid u = h_1 \cdots h_k(w) \text{ where } h_1, \ldots, h_k \in \Pi\} & \text{otherwise} \end{cases}.$$

So there may be at most c^k words which are generated in k steps where $c = \mathrm{card}(\Pi)$. A DT0L system is called a PDT0L system if every morphism in Π is λ-free.

We assume the reader is familiar with the rudiments of formal language theory and theory of L systems, see, for example, [9, 10].

3 Definitions of word length controlled DT0L systems

A word length controlled DT0L system first appears in [5]. Here we give the definition.

Definition 1 *A* word length controlled DT0L system, *or a* wlcDT0L system *for short, is a 4-tuple $\langle \Sigma, \Pi, w, f \rangle$ where Σ is a finite alphabet, Π is a set of morphisms over Σ called the set of tables, $w \in \Sigma^*$ is the axiom, and f is a partial recursive function from $\mathbb{N}$ to Π called the* control function.

A derivation by a wlcDT0L system is defined as follows.

Definition 2 *Let $G = \langle \Sigma, \Pi, w, f \rangle$ be a wlcDT0L system. Let x and y be words over Σ. Then G* directly derives *y from x if $y = f(|x|)(x)$. If $f(|x|)$ is not defined, then G derives nothing from x.*

By Definition 2, a wlcDT0L system $G = \langle \Sigma, \Pi, w, f \rangle$ generates a sequence of words $w = w_0, w_1, \ldots, w_i, \ldots$ which is given by $w_{i+1} = f(|w_i|)(w_i)$ for $i \in \mathbb{N}$. The sequence (w_i) is called the sequence generated by G.

A wlcDT0L system $G = \langle \Sigma, \Pi, w, f \rangle$ is said to be a wlcPDT0L system if every morphism $h \in \Pi$ is λ-free.

Now we give an example of a wlcPDT0L system.

Example 1 *Let $G = \langle \{A, a, b\}, \{h_1, h_2\}, A, f \rangle$ be a wlcPDT0L system where*

$$h_1(A) = aA, \; h_1(a) = a, \; h_1(b) = b,$$

$$h_2(A) = bA,\ h_2(a) = b,\ h_2(b) = a$$

and

$$f(n) = \begin{cases} h_1 & \textit{if } n \textit{ is a prime number} \\ h_2 & \textit{otherwise} \end{cases}.$$

The first few words in the sequence generated by G is as follows:

$$w_0 = A,\ w_1 = h_2(A) = bA,\ w_2 = h_1(w_1) = baA,\ w_3 = h_1(w_2) = baaA,$$

$$w_4 = h_2(w_3) = abbbA,\ w_5 = h_1(w_4) = abbbaA,\ w_6 = h_2(w_5) = baaabbA,$$

$$w_7 = h_1(w_6) = baaabbaA, \ldots.$$

Since f is a total recursive function, the sequence (w_i) is infinite. We cannot characterize (w_i) because we do not have an entire characterization of prime numbers.

4 Periodic control function

In this section we consider a wlcDT0L system with a periodic control function. Our goal is to establish Theorem 1, which insists the sequence generated by a wlcDT0L system with a periodic control function is made of finite number of D0L sequences. First we define this concept clearly.

Definition 3 *Let $G = \langle \Sigma, \Pi, w, f \rangle$ be a wlcDT0L system and let (w_i) be the sequence generated by G. The sequence (w_i) is said to be a* finite combination *of D0L sequences if there are k D0L systems $G_j = \langle \Sigma, h_j, u^{(j)} \rangle$ $(j = 0, 1, \ldots, k-1)$ and a nonnegative integer n_0 such that for every $n \geq n_0$, there exist $0 \leq p$ and $0 \leq j \leq k-1$ satisfying*

$$n = n_0 + pk + j \quad \text{and} \quad w_n = u_p^{(j)}$$

where $u_p^{(j)}$ is the p-th word in the sequence generated by G_j.

Example 2 *Let $G = \langle \{a, b, c\}, \{h_1, h_2\}, a, f \rangle$ be a wlcDT0L system in which*

$$h_1(a) = ab, \quad h_1(b) = bc, \quad h_1(c) = c,$$

$$h_2(a) = a, \quad h_2(b) = h_2(c) = \lambda \text{ and}$$

$$f(n) = \begin{cases} h_1 & \textit{if } n \leq 10 \\ h_2 & \textit{if } n > 10 \end{cases}.$$

Then G is a finite combination of D0L sequences generated by D0L systems $G_i = \langle \{a, b, c\}, h, w_i \rangle$ $(i = 1, \ldots, 5)$ where

$$h(a) = a, \quad h(b) = b, \quad h(c) = c$$

and

$$w_1 = a, \quad w_2 = ab, \quad w_3 = abbc, \quad w_4 = abbcbcc, \quad w_5 = abbcbccbccc,$$

because the sequence generated by G begins

$$a,\ ab,\ abbc,\ abbcbcc,\ abbcbccbccc,\ a, \ldots.$$

The next example shows another wlcDT0L system of finite combination of D0L sequences.

Example 3 *Let $G = \langle \{a, b\}, \{h_1, h_2\}, a, f \rangle$ be a wlcDT0L system where*

$$h_1(a) = ba,\ h_1(b) = ab, \quad h_2(a) = a,\ h_2(b) = ab$$

and

$$f(x) = \begin{cases} h_1 & \text{if } x = 2m + 1 \text{ (odd number)} \\ h_2 & \text{if } x = 2m \text{ (even number)} \end{cases}.$$

The first few words generated by G is

a	ba
aba	$baabba$
$abaaababa$	$baabbababaabbaabba$
$\ldots.$	

Then there are two D0L systems $G_1 = \langle \{a, b\}, g_1, a \rangle$ and $G_2 = \langle \{a, b\}, g_2, ba \rangle$ such that

$$g_1(a) = aba,\ g_1(b) = aab$$

and

$$g_2(a) = ba,\ g_2(b) = baab.$$

Now it is obvious that the sequence generated by G is a finite combination of the D0L sequences generated by G_1 and G_2.

The control functions of the wlcDT0L systems in the above examples are periodic. We can generalize these examples as follows.

Theorem 1 *Let $G = \langle \Sigma, \Pi, w, f \rangle$ be a wlcDT0L system. If f is ultimately periodic, that is, there exist positive integers n_0 and p such that for every integer $n \geq n_0$, $f(n) = f(n+p)$ holds, then the sequence (w_n) generated by G is a finite combination of D0L sequences.*

Proof. Let $\Sigma = \{a_1, a_2, \ldots, a_l\}$ and $\Pi = \{h_1, h_2, \ldots, h_k\}$. Let M_i $(i = 1, 2, \ldots, k)$ be the growth matrix corresponding to h_i, that is, the pq element a_{pq} of M_i is given by $a_{pq} = |h_i(a_p)|_{a_q}$. Let (w_n) be the sequence generated by G and let π_n be the Parikh vector of w_n. Let $\overline{\pi_n}$ and $\overline{M_i}$ be the image of π_n and M_i to the residue class ring of modulo p, that is, the i-th element $\overline{x_i}$ of $\overline{\pi_n}$ satisfies $\overline{x_i} \equiv x_i \bmod p$ where x_i is the i-th element of π_n for every $i = 1, \ldots, l$ and $\overline{a_{pq}}$ of $\overline{M_i}$ satisfies $\overline{a_{pq}} \equiv a_{pq} \bmod p$ where a_{pq} is the pq element of M_i.

Now it is obvious that for every $x, y \geq n_0$ $x \equiv y \bmod p$ if and only if $f(x) = f(y)$. For every $n \geq n_0$ we have

$$\pi_{n+1} = \pi_n M$$

and

$$\overline{\pi_{n+1}} = \overline{\pi_n} \overline{M_{f(|w_n|)}}.$$

Since $\overline{\pi_n}$ vary over a finite set, there exist integers $n \geq n_0$ and $k_0 \leq p^l$ such that $\overline{\pi_n} = \overline{\pi_{n+k_0}}$. Therefore we have that $|w_n| \equiv |w_{n+k_0}| \bmod p$ because $|w_n| = \pi_n \eta$ where η is the column vector $\eta = (1, 1, \ldots, 1)^T$. Now we have the equation

$$f(|w_n|) = f(|w_{n+k_0}|).$$

Then we have

$$\begin{aligned} \overline{\pi_{n+k_0+1}} &= \overline{\pi_{n+k_0}} \overline{M_{f(|w_{n+k_0}|)}} \\ &= \overline{\pi_n} \overline{M_{f(|w_n|)}} \\ &= \overline{\pi_{n+1}}. \end{aligned}$$

This means that the sequence $(|w_{n'}| \bmod p)$ has period k_0 for $n' \geq n$. Let k_1 be the least common multiple of k_0 and p. Then for every $0 \leq j < k_1$ the same morphism is iterated to w_{n+j+ik_1} for every $i \geq 0$. This completes the proof. □

We note that the reverse of Theorem 1 is not true. For example the wlcDT0L system $G = \langle \{a, b\}, \{h_1, h_2\}, a, f \rangle$ where f is given by

$$f(n) = \begin{cases} h_1 & \text{if } n \text{ is not a prime number} \\ h_2 & \text{if } n \text{ is a prime number} \end{cases}$$

and h_1 and h_2 are given by

$$h_1(a) = h_2(a) = ab,\ h_1(b) = h_2(b) = b$$

is an instance of counter-examples because the sequence generated by G is a finite combination of D0L sequences but the control function is not periodic.

There is another related question of Theorem 1, that is, for every D0L sequences $(u_i^{(0)}), \ldots, (u_i^{(k-1)})$ whether or not there exists a wlcDT0L system G such that the sequence generated by G is a finite combination of the given D0L sequences. The answer is no. The D0L sequences $(a^2, a^2, \ldots)$ and $(a^3, a^3, \ldots)$ which are generated by D0L systems $\langle\{a\}, h, a^2\rangle$ and $\langle\{a\}, h, a^3\rangle$ with $h(a) = a$ serve an example. Since no morphisms map a^2 to a^3 nor a^3 to a^2, there are no wlcDT0L systems which generate the sequence $(\ldots, a^2, a^3, a^2, a^3, \ldots)$. This example shows that the finite language $\{a^2, a^3\}$ cannot be generated by any wlcDT0L system.

References

[1] M. Andraşiu, G. Păun, J. Dassow, and A. Salomaa (1990) Language-theoretic problems arising from Richelieu cryptosystems, *Theoret. Comput. Sci.* **116**, 339–357.

[2] J. Honkala, On slender 0L languages over the binary alphabet, *Acta Informatica,* to appear.

[3] L. Ilie (1994) On a conjecture about slender context-free languages, *Theoret. Comput. Sci.* **132**, 427–434.

[4] L. Ilie, On length of words in context-free languages, *Theoret. Comput. Sci.,* to appear.

[5] T. Y. Nishida (1999) Word length controlled DT0L systems and slender languages, in: J. Karhumäki, H. Maurer, G. Păun, and G. Rozenberg, ed., *Jewels are Forever, Contributions on Theoretical Computer Science in Honor of Arto Salomaa*, Springer, Berlin, 213–221.

[6] T. Y. Nishida and A. Salomaa (1996) Slender 0L languages, *Theoret. Comput. Sci.* **158**, 161–176.

[7] T. Y. Nishida and A. Salomaa (2000) Note on slender 0L languages, *Theoret. Comput. Sci.* **233**, 279–286.

[8] G. Păun and A. Salomaa (1992) Decision problems concerning the thinness of D0L languages, *Bull. EATCS* **46**, 171–181.

[9] G. Rozenberg and A. Salomaa (1980) *The Mathematical Theory of L Systems*, Academic Press, New York.

[10] A. Salomaa (1973) *Formal Languages*, Academic Press, New York.

[11] Y. Takada and T. Y. Nishida (1996) A note on grammatical inference of slender context-free languages, in: L. Miclet and C. de la Higuera, ed., *Grammatical Inference: Learning Syntax from Sentences*, Lecture Notes in Artifical Intelligence, Vol. 1147, Springer, Berlin, 117–125.