A new proof for the decidability of D0L ultimate periodicity

Vesa Halava[1], Tero Harju[1]and Tomi Kärki[1,2]

[1]Department of Mathematics
University of Turku, 20014 Turku, Finland
[2]Department of Teacher Education
University of Turku, PO Box 175, 26101 Rauma, Finland

We give a new proof for the decidability of the D0L ultimate periodicity problem based on the decidability of p-periodicity of morphic words adapted to the approach of Harju and Linna.

1 Introduction

L systems were originally introduced by A. Lindenmayer to model the development of simple filamentous organisms [6, 7]. The challenging and fruitful study of these systems in the 70s and 80s created many new results and notions [9]. In this paper we consider the important problem of recognizing ultimately periodic D0L sequences.

Let $\mathscr{A}$ be a finite alphabet and denote the empty word by ε. A *D0L system* is a pair (h,u), where $h\colon \mathscr{A}^* \to \mathscr{A}^*$ is a morphism and u is a finite word over $\mathscr{A}$. The *language* of the D0L system is $L(h,u) = \{h^i(u) \mid i \geq 0\}$ and the *limit set* $\lim L(h,u)$ consists of all infinite words w such that for all n there is a prefix of w longer than n belonging to $L(h,u)$. Clearly, if the limit set is non-empty, then one can effectively find integers p and q such that $h^p(u)$ is a proper prefix of $h^{p+q}(u)$ and

$$\lim L(h,u) = \bigcup_{i=0}^{q-1} \lim L(h^q, h^{p+i}(u)),$$

where $|\lim L(h^q, h^{p+i}(u))| = 1$. Hence, we may restrict to D0L systems (h,u) where h is prolongable on u, i.e., $h(u) = uy$ and $h^n(y) \neq \varepsilon$ for all integers $n \geq 0$. In this case, $h^n(u)$ is a prefix of $h^{n+1}(u)$ and the limit is the following fixed point of h:

$$h^\omega(u) = \lim_{n\to\infty} h^n(u) = uyh(y)h^2(y)\cdots .$$

An infinite word x is ultimately periodic if it is of the form $x = uv^\omega = uvvv\cdots$, where u and v are finite words. The length $|u|$ is a *preperiod* and the length $|v|$ is a *period* of x. An infinite word x is *ultimately p-periodic* if $|v| = p$. The smallest period of x is called *the period* of x.

Now we are ready to formulate the *D0L ultimate periodicity problem*: *Given a morphism h prolongable on u, decide whether $h^\omega(u)$ is ultimately periodic.* Note that in this problem we may assume that u is a letter. Indeed, if $h(u) = uy$, then instead of (h,u) we may consider (h',a) where $a \notin \mathscr{A}$ and $h'\colon (\mathscr{A} \cup \{a\})^* \to (\mathscr{A} \cup \{a\})^*$ where $h'(a) = ay$ and $h'(b) = h(b)$ for every $b \in \mathscr{A}$. The limit $h^\omega(u)$ is ultimately periodic if and only if $h'^\omega(a)$ is.

The decidability of the ultimate periodicity question for D0L sequences was proven by T. Harju and M. Linna [4] and, independently, by J.-J. Pansiot [8]; see also a more recent proof of J. Honkala [5]. In

P. Ambrož, Š. Holub and Z. Masáková (Eds.):
8th International Conference WORDS 2011
EPTCS 63, 2011, pp. 147–151, doi:10.4204/EPTCS.63.20

the binary case the problem was effectively solved by Séébold [10]. Here we show how the proof of [4] can be simplified using a recent result concerning the decidability of the p-periodicity problem.

Before giving the proof, we introduce the following notation. Given a morphism $h\colon \mathscr{A}^* \to \mathscr{A}^*$, we call a letter $b \in \mathscr{A}$ *finite* if $\{h^n(b) \mid n \geq 0\}$ is a finite set. Otherwise, b is an *infinite letter*. Moreover, we say that a letter b is *recurrent* in $h^\omega(a)$ if it occurs infinitely often in $h^\omega(a)$. For a given morphism h prolongable on a and for an infinite word $h^\omega(a)$, denote the set of finite letters by $\mathscr{A}_F$, the set on infinite letters by $\mathscr{A}_I$ and the set of recurrent letters by $\mathscr{A}_R$. Also, denote by $\mathscr{A}_1$ the subset of $\mathscr{A}$ which consists of the infinite letters occurring infinitely many times in $h^\omega(a)$, i.e., $\mathscr{A}_1 = \mathscr{A}_I \cap \mathscr{A}_R$.

Let us shortly describe how the sets $\mathscr{A}_F$, $\mathscr{A}_I$ and $\mathscr{A}_R$ can be constructed. Note that if b is a mortal letter, i.e., $h^n(b) = \varepsilon$ for some $n \geq 1$, then $h^{|\mathscr{A}|}(b) = \varepsilon$. Denote $\hat{h} = h^{|\mathscr{A}|}$ and denote the set of the mortal letters by $\mathscr{M}$. Note also that b is a finite letter if and only if there exists a word $u \in \{h^n(b) \mid n \geq 0\}$ such that $u = h^p(u)$ for some $p \geq 1$. Clearly, $\{\hat{h}^n(b) \mid n \geq 0\}$ is finite if and only if $\{h^n(b) \mid n \geq 0\}$ is finite. Hence, by replacing h with $\hat{h}$ we may assume that $h(b) = \varepsilon$ if $b \in \mathscr{M}$. Moreover, let $\mathscr{B} = \mathscr{A} \setminus \mathscr{M}$ and let $g\colon \mathscr{B}^* \to \mathscr{B}^*$ be a morphism defined by $g(b) = \mu h(b)$, where

$$\mu(b) = \begin{cases} \varepsilon, & \text{if } b \in \mathscr{M}, \\ b, & \text{otherwise.} \end{cases}$$

Now g is non-erasing, and $b \in \mathscr{A}_F$ if and only if $\{g^n(b) \mid n \geq 0\}$ is finite. Namely, for any $n \geq 0$, we know by the definition of g that the word $h^n(b)$ can be obtained by inserting a finite number of mortal letters to $g^n(b)$. The set $\{g^n(b) \mid n \geq 0\}$ is finite if and only if for some n all letters in $g^n(b)$ belong to $U_1 = \{b \in \mathscr{B} \mid g^i(b) \in \mathscr{B} \text{ for every } i \geq 0\}$. If $U_i = \{b \in \mathscr{B} \mid g(b) \in U_{i-1}^*\}$, then $U_{i-1} \subseteq U_i$ and

$$\mathscr{A}_F \setminus \mathscr{M} = \bigcup_{i=1}^{\infty} U_i = U_{|\mathscr{A}|}.$$

Hence, we can effectively calculate $\mathscr{A}_F$ and $\mathscr{A}_I = \mathscr{A} \setminus \mathscr{A}_F$. In order to find the recursive letters, we construct a graph G where the set of vertices is $\mathscr{A}$ and there is an edge from b to c if c occurs in the image $h(b)$. Let $h(a) = ax$. If there are infinitely many paths from a letter in x to the letter b, then b occurs infinitely many times in $h^\omega(a)$.

2 Decidability of the p-periodicity problem

Let $p \geq 1$, and let $x = (x_n)_{n \geq 0}$ be an infinite word over $\mathscr{A} = \{a_1, \ldots, a_d\}$. For $0 \leq k \leq p-1$, we say that the letters occurring infinitely many times in positions x_n, where $n \equiv k \pmod{p}$, form the *k-set of x modulo p*. It was shown in [3] that these k-sets can be effectively constructed for $x = h^\omega(u)$, where h is prolongable on the word u. This is based on the fact that there exist integers r and q such that

$$|h^r(b)| \equiv |h^{r+q}(b)| \pmod{p} \tag{1}$$

for every letter $b \in \mathscr{A}$. The incidence matrix of h is the matrix $M = (m_{i,j})_{1 \leq i,j \leq d}$ where $m_{i,j}$ denotes the number of occurrences of a_i in $h(a_j)$. The sequence of matrices $M^n \bmod p$, where the entries are the residues modulo p, must be ultimately periodic. Since $|h^n(a_j)| \pmod{p}$ is the sum of the elements in the jth column of M^n, we conclude that the sequence $(|h^n(a_j)|)_{n \geq 0} \pmod{p}$ is ultimately periodic for every $a_j \in \mathscr{A}$ and (1) follows.

In order to find the k-sets of x modulo p we construct a directed graph $G_h = (V, E)$ where the set of vertices V is $\{(a,i) \mid a \in \mathscr{A},\ 0 \leq i < p\}$ and there is an edge from (c,i) to (d,j) if, for some b in x, the

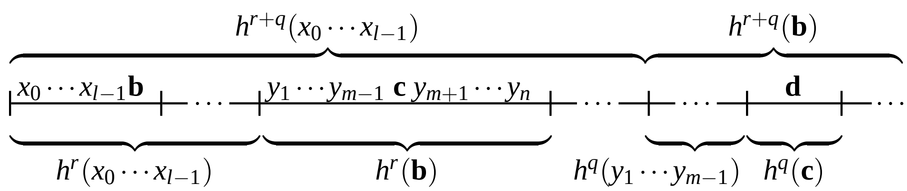

Figure 1: Images $h^r(b)$ and $h^{r+q}(b)$.

letter c occurs in the image $h^r(b)$ at position congruent to i (mod p) in x, and the letter d occurs in the image $h^q(c)$ at position congruent to j (mod p) in x; see Figure 1.

It is possible to construct such a graph by calculating the images $h^r(b)$ and $h^{r+q}(b)$ for every $b \in \mathscr{A}$. Namely, if $b = x_l$ and c is the mth letter of $h^r(b) = y_1 \cdots y_n$ and d is the m'th letter of $h^q(c)$, then we have

$$i \equiv |h^r(x_0 \cdots x_{l-1})| + m - 1 \pmod{p}, \tag{2}$$
$$j \equiv |h^{r+q}(x_0 \cdots x_{l-1})| + |h^q(y_1 \cdots y_{m-1})| + m' - 1 \pmod{p}. \tag{3}$$

By (1), we have $|h^{r+q}(x_0 \cdots x_{l-1})| \equiv |h^r(x_0 \cdots x_{l-1})|$ (mod p), which together with (2) and (3) implies

$$j \equiv |h^q(y_1 \cdots y_{m-1})| + i + m' - m \pmod{p}.$$

We say that a vertex $(c,i) \in V$ is an *initial vertex* if there exists a letter $b = x_l$ such that $0 \le l < |h^r(a)|$, c is the mth letter of $h^r(b)$ and i satisfies (2). A vertex (c,k) is called *recurrent* if there exist infinitely many paths starting from some initial vertex and ending in (c,k). By construction, this means that c belongs to the k set of x modulo p.

Given a coding g and a morphism $h\colon \mathscr{A}^* \to \mathscr{A}^*$ prolongable on a, it is easy to see that the morphic word $g(h^\omega(a))$ is ultimately p-periodic if and only if $g(b) = g(c)$ for all pairs of letters (b,c) such that b and c belong to the same k-set of $h^\omega(a)$ modulo p. Since the k-sets of $h^\omega(a)$ can be effectively constructed, we have the following result proved in [3].

Theorem 1. *Given a positive integer p, it is decidable whether a morphic word $g(h^\omega(a))$ is ultimately p-periodic.*

3 Decidability of the D0L ultimate periodicity problem

Before the decidability proof, we give the following result proved in [1, 2]; see also [5].

Theorem 2. *Let $h\colon \mathscr{A}^* \to \mathscr{A}^*$ be a morphism and $u, v \in \mathscr{A}^*$. If there is a positive integer n such that $h^n(u) = h^n(v)$, then $h^{|\mathscr{A}|}(u) = h^{|\mathscr{A}|}(v)$.*

This theorem can be proved by induction on the size of the alphabet and the induction step is based on elementary morphisms. A morphism $h\colon \mathscr{A}^* \to \mathscr{A}^*$ is called *elementary* if there do not exist an alphabet $\mathscr{B}$ smaller than $\mathscr{A}$ and two morphisms $f\colon \mathscr{A}^* \to \mathscr{B}^*$ and $g\colon \mathscr{B}^* \to \mathscr{A}^*$ such that $h = gf$.

Since elementary morphisms are injective, the claim is clear if h is elementary. Now assume that $h = gf$ as above. Then $h^n(u) = h^n(v)$ implies that $(fg)^n f(u) = (fg)^n f(v)$ and, by induction, $(fg)^{|\mathscr{B}|} f(u) = (fg)^{|\mathscr{B}|} f(v)$. This proves the claim, since $(gf)^{|\mathscr{B}|+1}(u) = (gf)^{|\mathscr{B}|+1}(v)$ and $|\mathscr{A}| \geq |\mathscr{B}| + 1$.

Using Theorem 1 and Theorem 2 and following the guidelines in [4] we give a new proof for the decidability of the D0L ultimate periodicity problem. The difference between the original proof of Harju and Linna and this proof is that we employ a new method obtained from p-periodicity as stated in Theorem 1.

Theorem 3. *The ultimate periodicity problem is decidable for D0L sequences.*

Proof. As explained above, it suffices to show that we can decide whether $h^\omega(a)$ is ultimately periodic for a given morphism $h\colon \mathscr{A}^* \to \mathscr{A}^*$ prolongable on a. Without loss of generality, we assume that every letter of $\mathscr{A}$ really occurs in $h^\omega(a)$. Otherwise, we could consider a restriction of h. Recall also that $\mathscr{A}_1$ is the subset of $\mathscr{A}$ which consists of the infinite letters occurring infinitely many times in $h^\omega(a)$.

If $\mathscr{A}_1 = \emptyset$, then the sequence is ultimately periodic. Namely, if $h(a) = ay$ and y contains infinite letters, then every image $h^n(y)$ contains infinite letters and there must be at least one infinite letter occurring infinitely many times in $h^\omega(a) = ayh(y)h^2(y)\cdots$, which means that $\mathscr{A}_1 \neq \emptyset$. Therefore, there is only one infinite letter and it is the letter a occurring once in the beginning of the word. Hence, $h(a) = ay$ where y consists of finite letters. Then there must be integers n and p such that $h^{n+p}(y) = h^n(y)$. Thus $|h^n(y)h^{n+1}(y)\cdots h^{n+p-1}(y)|$ is a period of $h^\omega(a)$.

Assume now that $b \in \mathscr{A}_1$. We may write

$$h^\omega(a) = u_0 b u_1 b u_2 \cdots,$$

where $u_i \in (\mathscr{A} \setminus \{b\})^*$. If the set $U = \{u_i \mid i \geq 0\}$ is infinite then $h^\omega(a)$ cannot be ultimately periodic. Note that if there exists a $c \in \mathscr{A}_I$ such that the letter b does not occur in any $h^i(c)$, then U is infinite. This property is clearly decidable since if a letter occurs in $h^i(c)$ for some i, then it occurs in the image for $i \leq |\mathscr{A}|$. Hence, we may assume that for each infinite letter c the letter b occurs in $h^i(c)$ for some $i \leq |\mathscr{A}|$.

Next we show that we may decide if U is infinite or not. First assume that U is infinite. Then there are arbitrarily long words in U. Since each infinite letter from $h^\omega(a)$ produces an occurrence of b in at most $|\mathscr{A}|$ steps, there must be arbitrarily long words from $\mathscr{A}_F$ in U. This is possible only if for some $c \in \mathscr{A}_I$ and integer $s \leq |\mathscr{A}|$ we have $h^s(c) = v_1 c v_2$, where for $i = 1$ or $i = 2$ we have $v_i \in \mathscr{A}_F^+$ and $h^n(v_i) \neq \varepsilon$ for every $n \geq 0$. This is a property that we can effectively check. Note that if $h^n(v_i) = \varepsilon$ for some $n \geq 0$, then $h^{|\mathscr{A}|}(v_i) = \varepsilon$. On the other hand, if there exists $c \in \mathscr{A}_I$ satisfying the above conditions, the set U is clearly infinite. Hence, the finiteness of U can be verified and the finite set U can be effectively constructed.

Now assume that $h^\omega(a)$ is ultimately periodic, i.e., $h^\omega(a) = uv^\omega$, where v is primitive. Consider a subset U' of U containing the elements u_i occurring infinitely many times in $h^\omega(a)$. Since b is in $\mathscr{A}_I$, there exists an integer N such that $|h^n(b)| \geq |v|$ for every $n \geq N$. Hence, let $n \geq N$. Since bu_i with $u_i \in U'$ occurs in the periodic part of the sequence, we conclude that $h^n(bu_i) \in w_n\mathscr{A}^*$, where w_n is a conjugate of v. Moreover, by the primitivity of v and w_n, we have

$$h^n(bu_i) \in w_n^* \quad \text{for all } u_i \in U'. \tag{4}$$

Namely, assume that $h^n(bu_i) = w_n^t w'$, where t is some positive integer and w' is a proper prefix of w_n, i.e., w' is non-empty and $w' \neq w_n$. Then $h^n(bu_i b) \in w_n^t w' w_n \mathscr{A}^*$ is a prefix of w_n^ω, which implies that the word w_n occurring after w' occurs inside w_n^2. Since w_n is primitive, this is impossible.

Take now any two words u_i and $u_j \in U'$. By (4), we conclude that there exists m such that $h^\ell(bu_ibu_j) = h^\ell(bu_jbu_i)$ for all $\ell \geq m$. Moreover, by Theorem 2, we know that we may choose $m = |\mathscr{A}|$. Note that if the above does not hold for some u_i and u_j in U', then $h^\omega(a)$ cannot be ultimately periodic. Hence, let $m = |\mathscr{A}|$ and

$$h^m(bu_ibu_j) = h^m(bu_jbu_i),$$

for every $u_i, u_j \in U'$. Then the words $h^m(bu_i)$ and $h^m(bu_j)$ commute and by transitivity we can find a primitive word z such that

$$h^\ell(bu_i) \in z^* \quad \text{for all } u_i \in U',\ \ell \geq m.$$

This implies that $h^\omega(a)$ is ultimately $|z|$-periodic. Since we can test the ultimate $|z|$-periodicity of $h^\omega(a)$ by Theorem 1, the ultimate periodicity problem of $h^\omega(a)$ is decidable. □

References

[1] K. Culik, II (1978): *The decidability of υ-local catenativity and of other properties of* D0L *systems*. *Information Processing Lett.* 7(1), pp. 33–35, doi:`10.1016/0020-0190(78)90035-2`.

[2] A. Ehrenfeucht & G. Rozenberg (1978): *Simplifications of homomorphisms*. *Inform. and Control* 38(3), pp. 298–309, doi:`10.1016/S0019-9958(78)90095-5`.

[3] V. Halava, T. Harju, T. Kärki & M. Rigo (2010): *On the periodicity of morphic words*. In Y. Gao, H. Lu, S. Seki & S. Yu, editors: *Developments in Language Theory, 14th International Conference, DLT 2010, London, ON, Canada, August 17-20, 2010. Proceedings*, *Lecture Notes in Comput. Sci.* 6224, Springer, Berlin, pp. 209–217, doi:`10.1007/978-3-642-14455-4_20`.

[4] T. Harju & M. Linna (1986): *On the periodicity of morphisms on free monoids*. *RAIRO Inform. Théor. Appl.* 20(1), pp. 47–54.

[5] J. Honkala (2008): *Cancellation and periodicity properties of iterated morphisms*. *Theoret. Comput. Sci.* 391(1-2), pp. 61–64, doi:`10.1016/j.tcs.2007.10.030`.

[6] A. Lindenmayer (1968): *Mathematical models for cellular interactions in development I. Filaments with one-sided inputs*. *J. Theoret. Biol.* 18, pp. 280–299, doi:`10.1016/0022-5193(68)90079-9`.

[7] A. Lindenmayer (1968): *Mathematical models for cellular interactions in development II. Simple and branching filaments with two-sided inputs*. *J. Theoret. Biol.* 18, pp. 300–315, doi:`10.1016/0022-5193(68)90080-5`.

[8] J.-J. Pansiot (1986): *Decidability of periodicity for infinite words*. *RAIRO Inform. Théor. Appl.* 20(1), pp. 43–46.

[9] G. Rozenberg & A. Salomaa (1986): *The Book of* L. Springer, Berlin.

[10] P. Séébold (1988): *An effective solution to the* D0L*-periodicity problem in the binary case*. *EATCS Bull.* 36, pp. 137–151.