A new proof for the decidability of D0L ultimate periodicity

Vesa Halava\textsuperscript{1}, Tero Harju\textsuperscript{1} and Tomi Kärki\textsuperscript{1,2}

\textsuperscript{1}Department of Mathematics
University of Turku, 20014 Turku, Finland
\textsuperscript{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 \cite{6,7}. The challenging and fruitful study of these systems in the 70s and 80s created many new results and notions \cite{9}. In this paper we consider the important problem of recognizing ultimately periodic D0L sequences.

Let \( \mathcal{A} \) be a finite alphabet and denote the empty word by \( \varepsilon \). A DOL system is a pair \((h,u)\), where \( h: \mathcal{A}^* \rightarrow \mathcal{A}^* \) is a morphism and \( u \) is a finite word over \( \mathcal{A} \). The language of the DOL 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 DOL 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 \( x = uv^\omega = uvvy\cdots \), 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 DOL 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 \( a \notin \mathcal{A} \) and \( h': (\mathcal{A} \cup \{a\})^* \rightarrow (\mathcal{A} \cup \{a\})^* \) where \( h'(a) = ay \) and \( h'(b) = h(b) \) for every \( b \in \mathcal{A} \). The limit \( h^\omega(u) \) is ultimately periodic if and only if \( h^\omega(a) \) is.

The decidability of the ultimate periodicity question for DOL sequences was proven by T. Harju and M. Linna \cite{4} and, independently, by J.-J. Pansiot \cite{8}; see also a more recent proof of J. Honkala \cite{5}. In
the binary case the problem was effectively solved by Séébold [10]. Here we show how the proof of [11] 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 \( h : \mathcal{A}^* \to \mathcal{A}^* \), we call a letter \( b \in \mathcal{A} \) 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 \( \mathcal{A}_f \), the set on infinite letters by \( \mathcal{A}_i \) and the set of recurrent letters by \( \mathcal{A}_r \). Also, denote by \( \mathcal{A}_1 \) the subset of \( \mathcal{A} \) which consists of the infinite letters occurring infinitely many times in \( h^\omega(a) \), i.e., \( \mathcal{A}_1 = \mathcal{A}_f \cap \mathcal{A}_r \).

Let us shortly describe how the sets \( \mathcal{A}_f \), \( \mathcal{A}_i \) and \( \mathcal{A}_r \) can be constructed. Note that if \( b \) is a mortal letter, i.e., \( h^n(b) = \varepsilon \) for some \( n \geq 1 \), then \( h^{i\mathcal{A}}(b) = \varepsilon \). Denote \( \hat{h} = h^{i\mathcal{A}} \) and denote the set of the mortal letters by \( \mathcal{M} \). 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 \( u = h^n(u) \) for some \( p \geq 1 \). Clearly, \( \{h^n(b) \mid n \geq 0\} \) 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 \( b \in \mathcal{M} \). Moreover, let \( \mathcal{B} = \mathcal{A} \setminus \mathcal{M} \) and let \( g : \mathcal{B}^* \to \mathcal{B}^* \) be a morphism defined by \( g(b) = \mu h(b) \), where

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

Now \( g \) is non-erasing, and \( b \in \mathcal{A}_f \) 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 \( U_1 = \{b \in \mathcal{B} \mid g^i(b) \in \mathcal{B} \text{ for every } i \geq 0\} \). If \( U_i = \{b \in \mathcal{B} \mid g(b) \in U_{i-1} \} \), then \( U_{i-1} \subseteq U_i \) and

\[
\mathcal{A}_f \setminus \mathcal{M} = \bigcup_{i=1}^\infty U_i = U_{|\mathcal{A}|}.
\]

Hence, we can effectively calculate \( \mathcal{A}_f \) and \( \mathcal{A}_1 = \mathcal{A} \setminus \mathcal{A}_f \). In order to find the recursive letters, we construct a graph \( G \) where the set of vertices is \( \mathcal{A} \) 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 \( \mathcal{A} = \{a_1, \ldots, a_d\} \). 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 \( x = h^\omega(a) \), where \( h \) is prolongable on the word \( u \). This is based on the fact that there exist integers \( r \) and \( q \) such that

\[
|h^r(b)| \equiv |h^{r+q}(b)| \pmod{p}
\]

for every letter \( b \in \mathcal{A} \). 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 \( M^n \pmod{p} \), 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 \( j \)-th column of \( M^n \), we conclude that the sequence \( (|h^n(a_j)|)_{n \geq 0} \pmod{p} \) is ultimately periodic for every \( a_j \in \mathcal{A} \) and \([11]\) 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 \( \{(a, i) \mid a \in \mathcal{A}, 0 \leq i < p\} \) and there is an edge from \((c, i)\) to \((d, j)\) if, for some \( b \) in \( x \), the
Namely, if we say that a vertex \( h \) is the \( k \)th letter of \( h(b) \), then \( h \) belongs to the same \( \mathcal{A} \)-set of \( h^{\omega}(c) \) at position congruent to \( j \) (mod \( p \)) in \( x \); see Figure 1.

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

\[
\begin{align*}
i &\equiv |h'(x_0 \cdots x_{l-1})| + m - 1 \pmod{p}, \\
j &\equiv |h'^{+q}(x_0 \cdots x_{l-1})| + |h'(y_1 \cdots y_{m-1})| + m' - 1 \pmod{p}.
\end{align*}
\]

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

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

We say that a vertex \((c, i) \in V\) is an initial vertex if there exists a letter \( b = x_l \) such that \( 0 \leq l < |h'(a)| \), \( c \) is the \( m \)th letter of \( h'(b) \) 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 \( h: \mathcal{A}^* \to \mathcal{A}^* \) 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 \( h: \mathcal{A}^* \to \mathcal{A}^* \) be a morphism and \( u, v \in \mathcal{A}^* \). If there is a positive integer \( n \) such that \( h^n(u) = h^n(v) \), then \( h^{|\mathcal{A}|}(u) = h^{|\mathcal{A}|}(v) \).

This theorem can be proved by induction on the size of the alphabet and the induction step is based on elementary morphisms. A morphism \( h: \mathcal{A}^* \to \mathcal{B}^* \) is called elementary if there do not exist an alphabet \( \mathcal{B} \) smaller than \( \mathcal{A} \) and two morphisms \( f: \mathcal{A}^* \to \mathcal{B}^* \) and \( g: \mathcal{B}^* \to \mathcal{A}^* \) 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, \( (fg)^{|x|} f(u) = (fg)^{|x|} f(v) \). This proves the claim, since \( (fg)^{|x|+1} f(u) = (fg)^{|x|+1} f(v) \) and \( |x| \geq |y| + 1 \).

Using Theorem 1 and Theorem 2 and following the guidelines in Theorem 3, we give a new proof for the decidability of the DOL 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 DOL sequences.

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

If \( \mathcal{A} = \emptyset \), 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 \( \mathcal{A} \neq \emptyset \). 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 \( b \in \mathcal{A} \). We may write

\[
h^\omega(a) = u_0bu_1bu_2 \cdots ,
\]

where \( u_i \in (\mathcal{A} \setminus \{b\})^* \). 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 \( c \in \mathcal{A} \) 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 \( i \leq |\mathcal{A}| \). Hence, we may assume that for each infinite letter \( c \) the letter \( b \) occurs in \( h^i(c) \) for some \( i \leq |\mathcal{A}| \).

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 \(|\mathcal{A}| \) steps, there must be arbitrarily long words from \( \mathcal{A} \) in \( U \). This is possible only if for some \( c \in \mathcal{A} \) and integer \( s \leq |\mathcal{A}| \) we have \( h^s(c) = v_1c_2v_2 \), where for \( i = 1 \) or \( i = 2 \) we have \( v_i \in \mathcal{A}^* \) and \( h^s(v_i) \neq \varepsilon \) for every \( n \geq 0 \). This is a property that we can effectively check. Note that if \( h^s(v_i) = \varepsilon \) for some \( n \geq 0 \), then \( h^{|\mathcal{A}|}(v_i) = \varepsilon \). On the other hand, if there exists \( c \in \mathcal{A} \) 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 \( \mathcal{A} \), 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 \( h^n(bu_i) \in w_n.\mathcal{A}^* \), where \( w_n \) is a conjugate of \( v \). Moreover, by the primitivity of \( v \) and \( w_n \), we have

\[
h^n(bu_i) \in w_n^t \quad \text{for all } u_i \in U'.
\]

Namely, assume that \( h^n(bu_i) = w^t_nw' \), 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 \( h^n(bu_i)b \in w^t_nw'w_n.\mathcal{A}^* \) is a prefix of \( w^\omega_n \), which implies that the word \( w_n \) occurring after \( w' \) occurs inside \( w^2_n \). 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 $m = |\sigma|$. 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 $m = |\sigma|$ 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. \qed

References

[1] K. Culik, II (1978): The decidability of $\upsilon$-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.