View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Theoretical Computer Science 432 (2012) 94–97

Contents lists available at SciVerse ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Note

Marked D0L systems and the 2*n*-conjecture

Juha Honkala

Department of Mathematics, University of Turku, FI-20014 Turku, Finland

ARTICLE INFO

Article history:
Received 23 May 2011
Accepted 16 January 2012
Communicated by J. Karhumäki

Keywords:
D0L system
D0L sequence equivalence problem
Marked morphism
Decidability

ABSTRACT

We show that to test the equivalence of two D0L sequences over an n-letter alphabet generated by marked morphisms it suffices to compare the first $2n + 1$ initial terms of the sequences. Under an additional condition it is enough to consider the $2n$ initial terms.
© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Let X be a finite alphabet, let $g : X^* \to X^*$ and $h : X^* \to X^*$ be morphisms and let $w \in X^*$ be a word. The problem of deciding whether or not

$$g^i(w) = h^i(w)$$

for all $i \geq 0$ is known as the D0L sequence equivalence problem. This problem is decidable which was first proved by Culik II and Fris [2]. For other solutions see [4,3,9–12,5]. However, the D0L sequence equivalence problem remains one of the most intriguing problems concerning free monoid morphisms. This is illustrated by the fact that in the general case all known algorithms to decide the equivalence of two given D0L sequences by comparing initial terms of the sequences require a great amount of work. On the other hand it is possible that it suffices to consider the $2n$ initial terms of the sequences where n is the cardinality of the underlying alphabet (see [13]). At least, no counterexamples are known. The claim that this simple algorithm always gives the correct answer is called the $2n$-conjecture. The validity of the $2n$-conjecture has been shown by Karhumäki if $n = 2$ (see [6]). All other cases remain open.

In this paper we study the sequence equivalence problem for marked D0L systems. A D0L system is marked if the underlying morphism $g : X^* \to X^*$ is marked meaning that the first letters of the images of the letters of X are different. We prove that to test the equivalence of two D0L sequences over an n-letter alphabet generated by marked morphisms it is enough to consider the $2n + 1$ initial terms. We will also give an additional condition which makes it possible to replace the bound $2n + 1$ by $2n$.

We assume that the reader is familiar with the basics concerning D0L systems (see [7,8]).

2. Definitions and results

We use standard language-theoretic notation and terminology. The *cardinality* of a finite set X is denoted by $\text{card}(X)$. Two words u and v are called *comparable* if u is a prefix of v or vice versa. The first letter of a nonempty word w is denoted by $\text{first}(w)$.

E-mail address: juha.honkala@utu.fi.

0304-3975/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2012.01.037

Let X be a finite alphabet. A set $L \subseteq X^+$ is called *marked* if $\text{first}(w_1) \neq \text{first}(w_2)$ whenever $w_1, w_2 \in L$ and $w_1 \neq w_2$. If $L \subseteq X^+$ is marked, then $\text{card}(L)$ is less than or equal to $\text{card}(X)$.

A nonerasing morphism $g : X^* \to X^*$ is called *marked* if $\text{first}(g(a)) \neq \text{first}(g(b))$ whenever $a, b \in X$ and $a \neq b$. Marked morphisms were introduced by Karhumäki and used, e.g., in [6].

A *D0L system* is a triple $G = (X, g, w)$, where X is a finite alphabet, $g : X^* \to X^*$ is a morphism and $w \in X^*$ is a word. The *sequence* $S(G)$ generated by G consists of the words

$$w, g(w), g^2(w), g^3(w), \ldots .$$

The set $L(G) = \{g^n(w) \mid n \geq 0\}$ is called the *language* of G. A D0L system $G = (X, g, w)$ is called *reduced* if there is no proper subset Y of X such that $L(G) \subseteq Y^*$. In the sequel it is tacitly assumed that all D0L systems under consideration are reduced.

Two D0L systems G and H are called *sequence equivalent* if $S(G) = S(H)$. In other words, the D0L systems $G = (X, g, u)$ and $H = (X, h, v)$ are sequence equivalent if

$$g^i(u) = h^i(v) \quad \text{for all } i \geq 0.$$

We now state our main result.

Theorem 1. *Let X be an alphabet having n letters. Let $g : X^* \to X^*$ and $h : X^* \to X^*$ be marked morphisms and let $w \in X^*$ be a word. Then*

$$g^i(w) = h^i(w) \quad \text{for all } i \geq 0$$

if and only if

$$g^i(w) = h^i(w) \quad \text{for } i = 0, 1, \ldots, 2n.$$

Theorem 1 will be proved in the following section.

It is an open problem whether the bound $2n$ in Theorem 1 can be replaced by the bound $2n - 1$ required for the $2n$-conjecture. The following example shows that for marked D0L systems over a binary alphabet it is necessary to test four initial terms.

Example 1. Let $X = \{a, b\}$ and define the morphisms $g : X^* \to X^*$ and $h : X^* \to X^*$ by $g(a) = bba$, $g(b) = abbaa$ and $h(a) = bbaabba$, $h(b) = a$. Then g and h are marked morphisms and

$$g^i(ba) = h^i(ba)$$

for $i = 0, 1, 2$, but

$$g^3(ba) \neq h^3(ba).$$

Note that this example is essentially the same as Example I.1.5 in [7] due to Nielsen; only the order of letters is reversed to obtain marked morphisms.

Under an additional condition the bound required for the $2n$-conjecture can be obtained.

Let $g : X^* \to X^*$ and $h : X^* \to X^*$ be morphisms. Then the pair (g, h) is called *1-incomparable* if there is a letter $a \in X$ such that $g(a)$ and $h(a)$ are not comparable.

Theorem 2. *Let X be an alphabet having n letters. Let $g : X^* \to X^*$ and $h : X^* \to X^*$ be marked morphisms and let $w \in X^*$ be a word. Assume that the pair (g, h) is 1-incomparable. Then*

$$g^i(w) = h^i(w) \quad \text{for all } i \geq 0$$

if and only if

$$g^i(w) = h^i(w) \quad \text{for } i = 0, 1, \ldots, 2n - 1.$$

Theorem 2 will be proved in the following section.

3. Proofs

In this section we assume that X is an alphabet having n letters and we assume that $g : X^* \to X^*$ and $h : X^* \to X^*$ are marked morphisms.

Define

$$E(g, h) = \{w \in X^* \mid g(w) = h(w)\}$$

and

$$D = \{w \in X^* \mid g^i(w) \in E(g, h) \text{ for } i = 0, 1, \ldots, n - 1\}.$$

The sets $E(g, h)$ and D are submonoids of X^*. The next two lemmas show that as a submonoid of X^* the set D is generated by a marked set.

Lemma 3. *Let $w, u_1, u_2 \in X^*$. If $wu_1 \in D$ and $wu_2 \in D$ but $w \notin D$, then first$(u_1) =$ first(u_2).*

Proof. Suppose $wu_1 \in D$ and $wu_2 \in D$ but $w \notin D$. Because $w \notin D$, there is an integer $i \in \{0, 1, \ldots, n-1\}$ such that $g^i(w) \notin E(g, h)$. Because $wu_1 \in D$, the words $gg^i(w)$ and $hg^i(w)$ are comparable. Assume that $hg^i(w)$ is a prefix of $gg^i(w)$, say $gg^i(w) = hg^i(w)v$, where $v \in X^*$. Because $g^i(w) \notin E(g, h)$, v is nonempty. Now

$$vgg^i(u_j) = hg^i(u_j)$$

for $j = 1, 2$. Hence

$$\text{first}(hg^i(u_1)) = \text{first}(hg^i(u_2)).$$

Because h and g are marked, it follows that first$(u_1) =$ first(u_2). □

Lemma 4. *Let*

$$P = (D \setminus \varepsilon) \setminus (D \setminus \varepsilon)^2.$$

Then P is a marked set and

$$D = P^*. \tag{1}$$

Proof. Eq. (1) follows because D is a submonoid of X^*. To prove that P is marked, suppose on the contrary that there are nonempty words $w_1, w_2 \in P$ such that first$(w_1) =$ first(w_2) and $w_1 \neq w_2$. Suppose first that one of w_1 and w_2 is a prefix of the other, say w_1 is a prefix of w_2. Then there is a nonempty word w_3 such that $w_2 = w_1w_3$. Because $w_1, w_2 \in D$, the definition of D implies that $w_3 \in D$. But then w_2 is a product of two nonempty words of D which contradicts the assumption that $w_2 \in P$. Hence the words w_1 and w_2 have to be incomparable. Then we can write $w_1 = wu_1$ and $w_2 = wu_2$ where $w, u_1, u_2 \in X^*$ are nonempty words and first$(u_1) \neq$ first(u_2). If $w \in D$, then $u_1 \in D$ and w_1 would be a product of two nonempty words of D which is not possible. Hence $w \notin D$. Now Lemma 3 implies that first$(u_1) =$ first(u_2) which is a contradiction. □

The following lemma shows that if $w \in D$, then the length of $gg^i(w)$ equals the length of $hg^i(w)$ for all $i \geq 0$. To prove this fact we use results concerning rational sequences. For the definition and basic properties of rational sequences we refer to [1,14].

Let $f : X^* \to X^*$ be a morphism and let $X = \{x_1, x_2, \ldots, x_n\}$. Then the *matrix M associated with f* is defined by

$$M = (m_{ij})_{1 \leq i,j \leq n},$$

where m_{ij} is the number of occurrences of x_i in $f(x_j)$.

Lemma 5. *If $w \in D$, then $|gg^i(w)| = |hg^i(w)|$ for all $i \geq 0$.*

Proof. Fix a word $w \in D$ and define the sequence $(a_i)_{i \geq 0}$ by

$$a_i = |gg^i(w)| - |hg^i(w)|$$

for $i \geq 0$. Let M and N be the matrices associated with g and h, respectively, and let β be the Parikh vector of w. Let $\alpha = \eta(M - N)$, where η is the row vector whose all entries equal 1. Then

$$a_i = \alpha M^i \beta^T$$

for $i \geq 0$. Hence $(a_i)_{i \geq 0}$ is a rational sequence of rank at most n. Because $w \in D$, we have $a_i = 0$ for $i = 0, 1, \ldots, n-1$. Now the claim follows by Corollary II.3.6 in [1]. □

The next two lemmas show that if w is a word such that $g^i(w) \in D$ for $i = 0, 1, \ldots$, card(P), then $g^i(w) \in D$ for all $i \geq 0$. To prove this we write the words $g^i(w)$, $i = 0, 1, \ldots$, card(P), as products of words of P and show that if R is the set of words of P which appear in these products, then $g(R) \subseteq R^*$.

Lemma 6. *Let t be a positive integer and let $w_1, \ldots, w_t \in P$. Suppose $g(w_1w_2 \cdots w_t) \in D$. Then $g(w_\alpha) \in D$ for $\alpha = 1, \ldots, t$.*

Proof. Because $w_1, \ldots, w_t \in D$, Lemma 5 implies that

$$|gg^i(w_\alpha)| = |hg^i(w_\alpha)| \tag{2}$$

for all $i \geq 0$ and $\alpha = 1, \ldots, t$. Because $g(w_1w_2 \cdots w_t) \in D$ we have

$$gg^ig(w_1w_2 \cdots w_t) = hg^ig(w_1w_2 \cdots w_t)$$

for $i = 0, 1, \ldots, n-1$. This together with (2) implies that

$$gg^ig(w_\alpha) = hg^ig(w_\alpha)$$

for $i = 0, 1, \ldots, n-1$ and $\alpha = 1, \ldots, t$. Hence $g(w_\alpha) \in D$ for $\alpha = 1, \ldots, t$. □

Lemma 7. *Let $w \in X^*$ be a nonempty word. Suppose $g^i(w) \in D$ for $i = 0, 1, \ldots$, card(P). Then $g^i(w) \in D$ for all $i \geq 0$.*

Proof. For $i = 0, 1, \ldots, \mathrm{card}(P)$, let P_i be the smallest subset of P such that

$$\{g^j(w) \mid j = 0, 1, \ldots, i\} \subseteq P_i^*.$$

Then

$$P_0 \subseteq P_1 \subseteq \cdots \subseteq P_{\mathrm{card}(P)}.$$

Because P_0 contains at least one word and $P_{\mathrm{card}(P)}$ contains at most $\mathrm{card}(P)$ words, there is an integer $\beta \in \{0, 1, \ldots, \mathrm{card}(P) - 1\}$ such that

$$P_\beta = P_{\beta+1}.$$

By Lemma 6 we have

$$g(P_\beta) \subseteq P_{\beta+1}^*.$$

Hence

$$g(P_\beta) \subseteq P_\beta^*.$$

It follows inductively that

$$g^i(g^\beta(w)) \in P_\beta^*$$

for all $i \geq 0$. This implies the claim. □

Now we are in a position to prove Theorems 1 and 2.

Proof of Theorem 1. Suppose $g^i(w) = h^i(w)$ for $i = 0, 1, \ldots, 2n$. Then

$$gg^i(w) = g^{i+1}(w) = h^{i+1}(w) = hh^i(w) = hg^i(w)$$

for $i = 0, 1, \ldots, 2n - 1$. Hence $g^i(w) \in D$ for $i = 0, 1, \ldots, n$. Because $\mathrm{card}(P) \leq n$, Lemma 7 implies that $g^i(w) \in D$ for all $i \geq 0$. Hence $gg^i(w) = hg^i(w)$ for all $i \geq 0$. It follows inductively that $g^i(w) = h^i(w)$ for all $i \geq 0$. □

Proof of Theorem 2. If the pair (g, h) is 1-incomparable, there is a letter $a \in X$ such that $g(a)$ and $h(a)$ are incomparable. Hence no word of $E(g, h)$ begins with a. It follows that no word of D or P begins with a. Because P is marked, we have $\mathrm{card}(P) \leq n - 1$.

The rest of the proof of Theorem 2 is analogous with the proof of Theorem 1.

If $g^i(w) = h^i(w)$ for $i = 0, 1, \ldots, 2n - 1$ we get $gg^i(w) = hg^i(w)$ for $i = 0, 1, \ldots, 2n - 2$. Hence $g^i(w) \in D$ for $i = 0, 1, \ldots, n - 1$. Now Lemma 7 implies that $g^i(w) \in D$ for all $i \geq 0$. This again implies that $g^i(w) = h^i(w)$ for all $i \geq 0$. □

References

[1] J. Berstel, C. Reutenauer, Rational Series and Their Languages, Springer, Berlin, 1988.
[2] K. Culik II, I. Fris, The decidability of the equivalence problem for D0L-systems, Inform. Control 35 (1977) 20–39.
[3] K. Culik II, J. Karhumäki, A new proof for the D0L sequence equivalence problem and its implications, in: G. Rozenberg, A. Salomaa (Eds.), The Book of L, Springer, Berlin, 1986, pp. 63–74.
[4] A. Ehrenfeucht, G. Rozenberg, Elementary homomorphisms and a solution of the D0L sequence equivalence problem, Theoret. Comput. Sci. 7 (1978) 169–183.
[5] J. Honkala, A short solution for the HDT0L sequence equivalence problem, Theoret. Comput. Sci. 244 (2000) 267–270.
[6] J. Karhumäki, On the equivalence problem for binary D0L systems, Inform. Control 50 (1981) 276–284.
[7] G. Rozenberg, A. Salomaa, The Mathematical Theory of L Systems, Academic Press, New York, 1980.
[8] G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, Vols. 1–3, Springer, Berlin, 1997.
[9] K. Ruohonen, Equivalence problems for regular sets of word morphisms, in: G. Rozenberg, A. Salomaa (Eds.), The Book of L, Springer, Berlin, 1986, pp. 393–401.
[10] K. Ruohonen, Test sets for iterated morphisms, Report 49, Tampere University of Technology, Tampere, 1986.
[11] K. Ruohonen, Explicit test sets for iterated morphisms in free monoids and metabelian groups, Theoret. Comput. Sci. 330 (2005) 171–191.
[12] K. Ruohonen, D0L sequence equivalence is in P for fixed alphabets, Theor. Inform. Appl. 42 (2008) 361–374.
[13] A. Salomaa, D0L equivalence: The problem of iterated morphisms, EATCS Bulletin 4 (1978) 5–12.
[14] A. Salomaa, M. Soittola, Automata-Theoretic Aspects of Formal Power Series, Springer, Berlin, 1978.