Journal of Universal Computer Science, vol. 7, no. 10 (2001), 886-892
submitted: 3/3/01, accepted: 23/10/01, appeared: 28/10/01

Three Variants of the DT0L Sequence Equivalence Problem

Juha Honkala
Department of Mathematics
University of Turku
FIN-20014 Turku, Finland
juha.honkala@utu.fi
and
Turku Centre for Computer Science (TUCS)
Lemminkäisenkatu 14
FIN-20520 Turku, Finland

Abstract: We discuss three variants of the DT0L sequence equivalence problem. One of the variants generalizes the ω-sequence equivalence problem of D0L systems for DT0L systems.

Key Words: Lindenmayer system, equivalence problem, decidability

Category: F.4.3

1 Introduction

The ω-sequence equivalence problem for D0L systems was shown to be decidable by Culik II and Harju [2]. This is one of the deepest results concerning morphisms of free monoids. Another deep result is the decidability of the sequence equivalence problem for DT0L systems (see Culik II and Karhumäki [3] and Honkala [6]). In this note we discuss three generalizations of these problems. The first one is the DT0L ω-sequence equivalence problem. The decidability status of this problem remains open. We also introduce two closely related variants. The first is shown to be undecidable while the second gives a nontrivial decidable generalization of the DT0L sequence equivalence problem.

It is assumed that the reader is familiar with the basics concerning L systems (see Rozenberg and Salomaa [8,9]). In the proofs we will use results concerning rational series (see Berstel and Reutenauer [1] and Salomaa and Soittola [10]). For infinite words generated by D0L and DT0L systems see also Culik II and Salomaa [4]. In particular, [4] gives a condition guaranteeing that a given DT0L system defines a unique infinite word.

2 The DT0L ω-sequence equivalence problem

Let X be a finite alphabet and X^* be the *free monoid generated* by X. The *length* of a word $w \in X^*$ is denoted by $|w|$. By definition the length of the *empty* *word* ε is zero. If $u, v \in X^*$ we denote $u \leq v$ if u is a prefix of v. Two words $u, v \in X^*$ are called *comparable* if $u \leq v$ or $v \leq u$. If $u = x_1 x_2 \ldots x_t$ where $x_i \in X$ for $1 \leq i \leq t$, we denote $u^T = x_t \ldots x_2 x_1$.

Suppose $(v_n)_{n \geq 0}$ is a sequence of words such that $v_n \leq v_{n+1}$ for all $n \geq 0$. Then we say that

$$\lim v_n$$

exists. If there is an integer n_0 such that $v_n = v_{n_0}$ for all $n \geq n_0$ then $\lim v_n = v_{n_0}$. Otherwise $\lim v_n$ equals the unique infinite word having v_n as a prefix for all $n \geq 0$.

A *DT0L system* is an $(n+2)$-tuple $G = (X, g_1, \ldots, g_n, w)$ where X is a finite alphabet, $w \in X^*$ is a word and $g_i : X^* \longrightarrow X^*$ is a morphism for $1 \leq i \leq n$. If $n = 1$, G is called a *D0L system*. For $n \geq 1$, let $X_n = \{1, 2, \ldots, n\}$ be an alphabet with n letters. If $G = (X, g_1, \ldots, g_n, w)$ is a DT0L system and $u = i_1 i_2 \ldots i_t \in X_n^*$ where $i_j \in X_n$, $1 \leq j \leq t$, denote

$$s_G(u) = g_{i_1} g_{i_2} \ldots g_{i_t}(w).$$

By definition, $s_G(\varepsilon) = w$. Two DT0L systems $G_1 = (X, g_1, \ldots, g_n, w_1)$ and $G_2 = (X, h_1, \ldots, h_n, w_2)$ are called *sequence equivalent* if $s_G(u) = s_H(u)$ for all words $u \in X_n^*$.

We say that a DT0L system $G = (X, g_1, \ldots, g_n, w)$ is *prolongable* if

$$w \leq g_i(w)$$

for all $1 \leq i \leq n$. Let $G = (X, g_1, \ldots, g_n, w)$ be a prolongable DT0L system and $(i_k)_{k \geq 1}$ be a sequence such that $i_k \in \{1, \ldots, n\}$ for all $k \geq 1$. Then the word

$$g_{i_1} g_{i_2} \cdots g_{i_k}(w)$$

is clearly a prefix of

$$g_{i_1} g_{i_2} \cdots g_{i_{k+1}}(w)$$

for any $k \geq 1$. Hence

$$\lim g_{i_1} g_{i_2} \ldots g_{i_k}(w)$$

exists. We denote

$$\omega_G(\alpha) = \lim g_{i_1} g_{i_2} \ldots g_{i_k}(w)$$

where $\alpha = (i_k)_{k \geq 1}$. Depending on α, $\omega_G(\alpha)$ is a finite or an infinite word. Two prolongable DT0L systems $G = (X, g_1, \ldots, g_n, w_1)$ and $H = (X, h_1, \ldots, h_n, w_2)$ are called *ω-sequence equivalent* if

$$\omega_G(\alpha) = \omega_H(\alpha)$$

for all sequences $\alpha = (i_k)_{k \geq 1}$ such that $i_k \in \{1, \ldots, n\}$ for $k \geq 1$.

Each pair of sequence equivalent DT0L systems gives in a canonical way a pair of ω-sequence equivalent DT0L systems.

Proposition 1. *Let $G = (X, g_1, \ldots, g_n, w)$ and $H = (X, h_1, \ldots, h_n, w)$ be DT0L systems. Choose two new letters $a, b \notin X$ and extend g_i and h_i, $1 \leq i \leq n$, by*

$$g_i(a) = h_i(a) = awb, \quad g_i(b) = h_i(b) = b, \quad 1 \leq i \leq n.$$

Then the DT0L systems $\overline{G} = (X \cup \{a, b\}, g_1, \ldots, g_n, a)$ and $\overline{H} = (X \cup \{a, b\}, h_1, \ldots, h_n, a)$ are ω-sequence equivalent if and only if G and H are sequence equivalent.

Proof. The systems $\overline{G}$ and $\overline{H}$ are prolongable. Let $\alpha = (i_k)_{k \geq 1}$ be a sequence such that $i_k \in \{1, \ldots, n\}$ for all $k \geq 1$. Then we have

$$g_{i_1} g_{i_2} \ldots g_{i_k}(a) = awbg_{i_1}(w)bg_{i_1}g_{i_2}(w)b \ldots bg_{i_1}g_{i_2} \ldots g_{i_{k-1}}(w)b$$

and

$$h_{i_1} h_{i_2} \ldots h_{i_k}(a) = awbh_{i_1}(w)bh_{i_1}h_{i_2}(w)b \ldots bh_{i_1}h_{i_2} \ldots h_{i_{k-1}}(w)b$$

for any $k \geq 1$. Hence

$$\lim g_{i_1} g_{i_2} \ldots g_{i_k}(a) = awbg_{i_1}(w)bg_{i_1}g_{i_2}(w)b \ldots bg_{i_1}g_{i_2} \ldots g_{i_k}(w)b \ldots$$

and, similarly,

$$\lim h_{i_1} h_{i_2} \ldots h_{i_k}(a) = awbh_{i_1}(w)bh_{i_1}h_{i_2}(w)b \ldots bh_{i_1}h_{i_2} \ldots h_{i_k}(w)b \ldots.$$

This implies the claim. □

It is now natural to pose

The DT0L ω-sequence equivalence problem. Is it decidable whether or not two given prolongable DT0L systems $G = (X, g_1, \ldots, g_n, w_1)$ and $H = (X, h_1, \ldots, h_n, w_2)$ are ω-sequence equivalent?

Culik II and Harju [2] have shown that the ω-sequence equivalence problem is decidable for D0L systems. For DT0L systems with more than one tables the problem remains open. By Proposition 1, the DT0L ω-sequence equivalence problem also generalizes the DT0L sequence equivalence problem.

The DT0L ω-sequence equivalence problem might turn out to be undecidable. This would be interesting because the problem is a common generalization of two very important decidable problems. If the DT0L ω-sequence equivalence problem is decidable there appears to be two possibilities. Firstly, it might be that the DT0L ω-sequence equivalence problem can be reduced to D0L ω-sequence equivalence and DT0L sequence equivalence problems. This is certainly the case for some instances of the problem. On the other hand, if no such reduction exists in the general case the problem is very difficult. This follows because the known solutions of the D0L ω-sequence equivalence problem and DT0L sequence equivalence problem are entirely different and a decision method for the DT0L ω-sequence equivalence problem would solve both of them.

3 The weak DT0L ω-sequence equivalence problem

In this section we discuss an undecidable problem closely related to the DT0L ω-sequence equivalence problem.

Suppose $G = (X, g_1, \ldots, g_n, w_1)$ and $H = (X, h_1, \ldots, h_n, w_2)$ are DT0L systems which are not necessarily prolongable. We say that G and H are *weakly ω-sequence equivalent* if for any $k \geq 0$ and $1 \leq i_1, \ldots, i_k \leq n$, the words

$$g_{i_1} g_{i_2} \ldots g_{i_k}(w_1)$$

and

$$h_{i_1} h_{i_2} \ldots h_{i_k}(w_2)$$

are comparable.

If the above systems G and H are prolongable and the words $\omega_G(\alpha)$ and $\omega_H(\alpha)$ are infinite for all sequences α, then G and H are ω-sequence equivalent if and only if they are weakly ω-sequence equivalent. For prolongable systems in general, ω-sequence equivalence implies weak ω-sequence equivalence but not vice versa. We will prove that weak ω-sequence equivalence is undecidable for DT0L systems. The DT0L systems considered in the proof are not prolongable. Therefore the proof does not apply to the DT0L ω-sequence equivalence problem.

Theorem 2. *It is undecidable whether or not two DT0L systems G and H with three tables are weakly ω-sequence equivalent.*

Proof. We will show that if we could decide weak ω-sequence equivalence for DT0L systems with three tables we could decide whether or not a $\mathbf{Z}$-rational series $r \in \mathbf{Z}^{\mathrm{rat}} \ll X_2^* \gg$ has a positive coefficient. Since such an algorithm does not exist for $\mathbf{Z}$-rational series (see Salomaa and Soittola [10]), there is no algorithm for the weak DT0L ω-sequence equivalence problem.

Suppose $r \in \mathbf{Z}^{\mathrm{rat}} \ll X_2^* \gg$. Recall that every $\mathbf{Z}$-rational series can be expressed as the difference of two DT0L series (see Salomaa and Soittola [10]). Hence there exist a DT0L system $F = (A, f_1, f_2, w)$ and two morphisms $\alpha_1 : A^* \longrightarrow b^*$, $\alpha_2 : A^* \longrightarrow b^*$ where $b \notin A$ is a new letter such that

$$(r, i_1 \ldots i_t) = |\alpha_1 f_{i_t} \ldots f_{i_1}(w)| - |\alpha_2 f_{i_t} \ldots f_{i_1}(w)|$$

for any $t \geq 0$, $1 \leq i_1, \ldots, i_t \leq 2$. Choose a new letter $\$ \notin A \cup b$ and extend f_i and α_i by

$$f_i(b) = \alpha_i(b) = b, \quad f_i(\$) = \alpha_i(\$) = \$$$

for $i = 1, 2$. Define two new DT0L systems F_1 and F_2 by

$$F_1 = (A \cup b \cup \$, f_1, f_2, \alpha_1, w)$$

and

$$F_2 = (A \cup b \cup \$, f_1, f_2, \alpha_2, w\$).$$

Now let $u \in X_3^*$ be a word and denote $v_1 = s_{F_1}(u)$ and $v_2 = s_{F_2}(u)$. Then there exist integers $q \geq 0$, $1 \leq j_1, \ldots, j_q \leq 2$ such that either

$$v_1 = f_{j_q} \ldots f_{j_1}(w) \text{ and } v_2 = f_{j_q} \ldots f_{j_1}(w)\$$$

or

$$v_1 = \alpha_1 f_{j_q} \ldots f_{j_1}(w) = b^{k_1} \text{ and } v_2 = \alpha_2 f_{j_q} \ldots f_{j_1}(w)\$ = b^{k_2}\$$$

where

$$k_i = |\alpha_i f_{j_q} \ldots f_{j_1}(w)| \text{ for } i = 1, 2.$$

Therefore F_1 and F_2 are weakly ω-sequence equivalent if and only if no coefficient of r is positive. This concludes the proof. □

Culik II and Salomaa [4] have shown that there is no algorithm for deciding whether or not in a given DT0L language some word is a prefix of another one.

4 The near equivalence of DT0L systems

In this section we discuss a decidable variant of the DT0L sequence equivalence problem which is closely connected to the DT0L ω-sequence equivalence problem. Suppose $G = (X, g_1, \ldots, g_n, w_1)$ and $H = (X, h_1, \ldots, h_n, w_2)$ are arbitrary DT0L systems. We say that G and H are *nearly equivalent* if there exists a positive integer K such that the following two conditions hold for any $u \in X_n^*$:
(i) The words $s_G(u)$ and $s_H(u)$ are comparable.
(ii) $||s_G(u)| - |s_H(u)|| \leq K$.
The near equivalence of D0L systems is discussed in Honkala [7] where it is seen to be a very natural notion in the study of infinite words.

Theorem 3. *The near equivalence is decidable for DT0L systems.*

Proof. Suppose $G = (X, g_1, \ldots, g_n, w_1)$ and $H = (X, h_1, \ldots, h_n, w_2)$ are DT0L systems. Define the power series $r \in \mathbf{Z} \ll X_n^* \gg$ by

$$(r, u) = |s_G(u)| - |s_H(u)|, \quad u \in X_n^*.$$

Because G and H are DT0L systems, r is $\mathbf{Z}$-rational. Hence it is decidable whether or not r has infinitely many different coefficients (see Berstel and Reutenauer [1]). If it does, G and H are not nearly equivalent. We proceed with the assumption that r takes only finitely many different values. Then there effectively exist a positive integer t, integers $a_1, \ldots, a_t$ and pairwise disjoint regular languages $L_1, \ldots, L_t \subseteq X_n^*$ such that $X_n^* = L_1 \cup \ldots \cup L_t$ and

$$r = \sum_{i=1}^{t} a_i \mathrm{char}(L_i)$$

(see Salomaa and Soittola [10]).

Now G and H are nearly equivalent if and only if for any $u \in X_n^*$ one of $s_G(u)$ and $s_H(u)$ is a prefix of the other. To decide whether this is true, define first the words $e_G(u), e_H(u) \in X^*$ for $u \in X_n^*$ as follows. If $u \in L_i$, $1 \leq i \leq t$, and $a_i \geq 0$, the word $e_G(u)$ is empty and $e_H(u)$ is the suffix of $s_G(u)$ of length a_i. If $u \in L_i$, $1 \leq i \leq t$, and $a_i < 0$, the word $e_G(u)$ is the suffix of $s_H(u)$ of length $-a_i$ and the word $e_H(u)$ is empty. It is easy to see that G and H are nearly equivalent if and only if

$$s_G(u)e_G(u) = s_H(u)e_H(u) \tag{1}$$

for all $u \in X_n^*$.

Now, for any word $v \in X^*$ the set

$$\{u \in X_n^* \mid s_G(u) \in X^*v\}$$

is regular (see Ginsburg and Rozenberg [5]). Hence, for any i, $1 \leq i \leq t$, there exist an integer $k_i \geq 1$ and pairwise disjoint regular languages L_{ij}, $1 \leq j \leq k_i$, such that

$$L_i = \bigcup_{j=1}^{k_i} L_{ij}$$

and $e_G(u)$ (respectively $e_H(u)$) is the same word for any $u \in L_{ij}$. Consequently there is a finite deterministic automaton with the initial state q_0 and input alphabet X_n such that if

$$q_0 u^T = q_0 v^T$$

then

$$e_G(u) = e_G(v), \quad e_H(u) = e_H(v)$$

for $u, v \in X_n^*$. (Here $q_0 u^T$ is the state which the automaton reaches by reading u^T.) It is easy to see that there is a DT0L system F with n tables such that

$$s_F(u) = q_0 u^T$$

for $u \in X_n^*$. Putting all this together we see that $(s_G(u)e_G(u))_{u \in X_n^*}$ and $(s_H(u)e_H(u))_{u \in X_n^*}$ are HDT0L sequences. Hence we can decide whether or not (1) holds for all $u \in X_n^*$. $\square$

References

[1] J. Berstel and C. Reutenauer, Rational Series and Their Languages (Springer, Berlin, 1988).

[2] K. Culik II and T. Harju, The ω-sequence equivalence problem for D0L systems is decidable, J. Assoc. Comput. Mach. 31 (1984) 282-298.

[3] K. Culik II and J. Karhumäki, A new proof for the D0L sequence equivalence problem and its implications, in G. Rozenberg and A. Salomaa (Eds.), The Book of L, Springer, Berlin, 1986, pp. 63-74.

[4] K. Culik II and A. Salomaa, On infinite words obtained by iterating morphisms, Theoret. Comput. Sci. 19 (1982) 29-38.

[5] S. Ginsburg and G. Rozenberg, T0L schemes and control sets, Inform. and Control 27 (1975) 109-125.

[6] J. Honkala, A short solution for the HDT0L sequence equivalence problem, Theoret. Comput. Sci. 244 (2000) 267-270.

[7] J. Honkala, On infinite words generated by polynomial D0L systems, Discrete Appl. Math., to appear.

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

[9] G. Rozenberg and A. Salomaa (Eds.), Handbook of Formal Languages, Vol. 1-3 (Springer, Berlin, 1997).

[10] A. Salomaa and M. Soittola, Automata-Theoretic Aspects of Formal Power Series (Springer, Berlin, 1978).