Revised copy April 1982

The ω-sequence Equivalence Problem for DOL Systems is Decidable

Karel Culik II
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

Tero Harju
Department of Mathematics
University of Turku
Turku, Finland

Research Report CS-81-02
January 1981

Faculty of Mathematics

University of Waterloo
Waterloo, Ontario, Canada
N2L 3G1

Abstract

The following problem is shown to be decidable. Given are homomorphisms h_1 and h_2 from Σ^* to Σ^* and strings σ_1 and σ_2 over Σ such that $h_i^n(\sigma_i)$ is a proper prefix of $h_i^{n+1}(\sigma_i)$ for $i = 1, 2$ and all $n \geq 0$, i.e. for $i = 1, 2$, h_i generates from σ_i an infinite string α_i with prefixes $h_i^n(\sigma_i)$ for all $n \geq 0$. Test whether $\alpha_1 = \alpha_2$. From this result easily follows the decidability of limit language equivalence (ω-equivalence) for DOL systems.

1. Introduction

Since the old work of Thue [15], infinite words (ω-words) have been investigated. Apart from being of interest in its own right, the theory of infinite words has often been able to shed light on some problems concerning ordinary finite words and languages of them. As regards infinite words associated to finite automata, the reader is referred to [8], and as regards those associated to context-free grammars, the reader is referred to [11].

Iterated morphisms (in other words: DOL systems) provide a very suitable framework for studying certain problems dealing with infinite words, see [13] and [14]. This problem area is closely connected with problems concerning morphisms in general, see [5]. For instance, it has been shown in [6] that our main result, the decidability of ω-sequence equivalence problem for DOL systems, implies the decidability of the ordinary DOL sequence equivalence problem, which was for a long time the best known open problem in the area of L systems [4]. This indicates that our main result is a hard one, especially when the attempts to reduce it to DOL equivalence have not succeeded. However, we are using some auxiliary results and refinements of techniques from [4].

We consider both words and infinite words also referred to as <u>ω-words</u>, over a finite alphabet Σ . An <u>ω-language</u> is a set of ω-words. If $L \subseteq \Sigma^*$, then $\lim(L)$ is the ω-language consisting of the ω-words with arbitrary long prefixes belonging to L .

The limit language equivalence problem (or ω-equivalence problem) for a family of languages is the decision problem of whether $\lim(L_1) = \lim(L_2)$ for any two effectively given languages L_1 and L_2

from the family. We show that this problem is decidable for DOL languages, given by DOL systems. It was conjectured to be decidable in [6], where it was reduced to the following problem.

Given are endomorphisms $h_1, h_2 : \Sigma^* \to \Sigma^*$ and words $\sigma_1, \sigma_2 \in \Sigma$, such that $h_i^n(\sigma_i)$ is a proper prefix of $h_i^{n+1}(\sigma_i)$ for $i = 1, 2$ and for all $n \geq 0$; i.e. morphisms h_1 and h_2 generate two infinite words α_1 with prefixes $h_1^n(\sigma_1)$ for all $n \geq 0$ and α_2 with prefixes $h_2^n(\sigma_2)$ for all $n \geq 0$. Decide whether $\alpha_1 = \alpha_2$? We will use this reduction and also a very useful lemma from [6] concerning "combinations" of morphisms (our Theorem 2.6).

Our approach generalizes and extends the techniques used in [4] to prove the decidability of the DOL-sequence equivalence problem. Similar notions of normal systems, simple systems, common subalphabets and combinations of morphisms as in [4] are used, however the situation at a number of places is more difficult and new techniques need to be deviced. The basic strategy remains the same, we show that for every pair of ω-equivalent DOL systems we can construct a finite number of pairs of DOL systems each of them ω-equivalent with "bounded balance".

The main goal of section 2 is to show that without loss of generality we can restrict ourselves to normal 1-systems. In the next section 1-simple systems are introduced and it is shown, using linear algebraic arguments, that ω-equivalent 1-simple systems have combinations with bounded balance. The last section contains the most crucial arguments showing essentially that the general case can be reduced to the case of 1-simple normal systems.

2. Preliminaries

For notations and definitions in language theory not explained here we refer to [12]. We shall also assume familiarity with the results in [4].

The entity $|x|$ denotes (i) the absolute value of a complex number x ; (ii) the length of a word x ; (iii) the vector $(|x_1|, \ldots, |x_k|)$ if x is a real-valued vector $(x_1, \ldots, x_k)$.

Let x and y be two words over a finite alphabet. If x is a prefix (a postfix, resp.) of y then we denote $x <_{pr} y$ ($x <_{po} y$, resp.). A word x is periodic if it is of the form $x = y^n y_1$, where $n \geqq 2$ and $y_1 <_{pr} y$. The words x and y are comparable if either $x <_{pr} y$ or $y <_{pr} x$. The empty word is denoted by e and the free monoid generated by a set Σ is denoted by Σ^* .

If $h_1, \ldots, h_k$ are endomorphisms on Σ^* then $<h_1, \ldots, h_k>$ denotes the monoid generated by $h_1, \ldots, h_k$ under the operation of composition of morphisms.

An infinite word is called an ω-word and a set of ω-words is said to be an ω-language. To each language L (of finite words) we associate an ω-language $\mathrm{Lim}(L)$, the limit language of L , which consists of the ω-words α having arbitrarily long prefixes belonging to L . Clearly if L is finite then $\mathrm{Lim}(L) = \phi$.

A language L is semi-convergent if $\mathrm{Lim}(L) \neq \phi$, convergent if each word in L is a prefix of some ω-word in $\mathrm{Lim}(L)$. Furthermore L is said to be uniformly convergent if $\#\mathrm{Lim}(L) = 1$, i.e. L has an unique limit word.

The <u>limit language equivalence problem</u> (or ω-<u>equivalence problem</u>) for a family of languages means the decision problem $\mathrm{Lim}(L_1) =? \mathrm{Lim}(L_2)$ for any (effectively given) L_1 and L_2 from the family.

We shall prove that the limit language equivalence problem is decidable for DOL languages.

For the proof of this result we shall first reduce the problem to a simplified form in this chapter. A <u>DOL system</u> is a construct $G = (\Sigma, h, \sigma)$, where Σ is a finite alphabet, h is an endomorphism on Σ^* and $\sigma \in \Sigma^*$. Denote

$$L(G) = \{h^n(\sigma) : n \geqq 0\} ,$$

the language generated by G .

<u>Lemma 2.1</u>

Given a DOL system $G = (\Sigma, h, \sigma)$, it is decidable if $\mathrm{Lim}(G)$ (= $\mathrm{Lim}(L(G))$) is empty or not. Furthermore $\mathrm{Lim}(G)$ is always finite.

<u>Proof</u>

See [6]. □

The system G (as defined above) is <u>prefix-preserving</u> if $\sigma <_{pr} h(\sigma)$. The following was shown in [6].

<u>Theorem 2.2</u>

The limit language equivalence problem is decidable for DOL systems iff it is decidable for prefix-preserving DOL systems.

□

In fact it was shown in [6] that

$$\mathrm{Lim}(L(G)) = \bigcup_{i=1}^{n} \mathrm{Lim}(L(G_i)) ,$$

where G_i , $i = 1,2,\ldots,n$, are subsystems of G and G_i is prefix-preserving for each i . From [6] we take also

Theorem 2.3

A prefix-preserving DOL system is uniformly convergent or its limit language is empty.

□

By Theorem 2.3. $\#\mathrm{Lim}(L(G)) = 1$ if G is prefix-preserving and $L(G)$ is infinite.

A DOL system $G = (\Sigma, h, \sigma)$ is a 1-system if it is prefix-preserving and furthermore

(i) $\sigma \in \Sigma$, (denote $\Sigma_c = \Sigma - \{\sigma\}$) ,

(ii) $h(\sigma) \in \sigma\Sigma_c^*$ and $h(\Sigma_c) \subseteq \Sigma_c^*$,

(iii) if $a \in \Sigma_c$ then a occurs infinitely many times in the unique limit word of G .

The subset Σ_c of Σ is called a core (core alphabet) of G .

We note here that if $G = (\Sigma, h, \sigma)$ is prefix-preserving and if $h(\sigma) = \sigma x$ then

$$h^{n+1}(\sigma) = h^n(\sigma)h^n(x) ,$$

for all $n \geqq 0$.

Lemma 2.4

Two prefix-preserving DOL systems $G_i = (\Sigma, h, \sigma_i)$, $i = 1, 2$ define the same limit word iff $\sigma_1 <_{pr} \sigma_2$ or $\sigma_2 <_{pr} \sigma_1$.

Proof

Immediate since G_1 and G_2 have the same morphism h to be iterated.

□

The next lemma reveals that we may restrict ourselves to 1-systems.

Lemma 2.5

The limit language equivalence problem is decidable for DOL systems iff it is decidable for 1-systems.

Proof

Let $G_i = (\Sigma, h_i, \sigma_i)$, $i = 1, 2$ be two prefix-preserving DOL systems. By Lemma 2.1 we may suppose that they both define an ω-word : α_i , correspondingly. Let Δ_i be the set of symbols which occur infinitely often in α_i . The sets Δ_1 and Δ_2 can be constructed effectively. If $\Delta_1 \neq \Delta_2$ then $\alpha_1 \neq \alpha_2$. Suppose thus that $\Delta_1 = \Delta_2$, and let $h_i(\sigma_i) = \sigma_i x_i$ for $i = 1, 2$. Now, there exists an integer n_o such that

$$h_i^n(x_i) \in \Delta_i^*$$

for $n \geq n_o$, $i = 1, 2$. Assume that $h_i^{n_o+1}(\sigma_i) <_{pr} \alpha_j$ for $i,j = 1, 2$. Otherwise we conclude that $\alpha_1 \neq \alpha_2$. Let $h_j^{n_o+1}(\sigma_j)$ be of maximum length of the two words $h_1^{n_o+1}(\sigma_1)$ and $h_2^{n_o+1}(\sigma_2)$.

By Lemma 2.4 the systems $(\Sigma, h_i, h_j^{n_o+1}(\sigma_j))$ define the limit words α_i , correspondingly for $i = 1, 2$. Thus

(1) $$\alpha_i = h_j^{n_o}(\sigma_j)h_j^{n_o}(x_j)\overline{\alpha}_i$$

for some ω-word $\overline{\alpha}_i$, $i = 1,2$. Let σ be a new symbol and set $\overline{\Sigma} = \{\sigma\} \cup \Delta_1$. Define two DOL systems

$$\overline{G}_i = (\overline{\Sigma}, \overline{h}_i, \sigma) , \quad i = 1, 2 ,$$

where $\overline{h}_i(\sigma) = \sigma h_j^{n_o}(x_j)$ and $\overline{h}_i(a) = h_i(a)$ for $a \in \Delta_1$. Here j is fixed as above. By (1) the system $\overline{G}_i$ defines the limit word

$$\alpha_i' = \sigma h_j^{n_o}(x_j)\overline{\alpha}_i$$

for $i = 1, 2$. Thus $\alpha_1 = \alpha_2$ if and only if $\alpha_1' = \alpha_2'$ and $h_1^{n_o+1}(\sigma_1)$ and $h_2^{n_o+1}(\sigma_2)$ are comparable. The claim follows since $\overline{G}_1$ and $\overline{G}_2$ are both 1-systems.

□

The following important result was proved in [6] (Thm 6 in [6]).

<u>Theorem 2.6</u>

Let $G_i = (\Sigma, h_i, \sigma)$ be two 1-systems for $i = 1, 2$ and $h \in \langle h_1, h_2 \rangle$. Then

$$\mathrm{Lim}(L(G_1)) = \mathrm{Lim}(L(G_2)) = \{\alpha\}$$

iff

$$\mathrm{Lim}(L(G_1')) = \mathrm{Lim}(L(G_2')) = \{\alpha\} ,$$

where $G_i' = (\Sigma, h_i h, \sigma)$ for $i = 1,2$. □

This result yields immediately to the following:

Lemma 2.7

Let G_i be as above and $g_i \in \langle h_1, h_2 \rangle$ for $i = 1, 2$. The systems G_1 and G_2 define a common limit word α iff the 1-systems $G_i' = (\Sigma, h_i g_i, \sigma)$ define the limit word α for both $i = 1, 2$.

□

For the next reduction we need some notations and facts from [4]. Let x be a word in Σ^* and define

$$\min(x) = \{a : a \text{ occurs in } x , a \in \Sigma\} .$$

Let $G = (\Sigma, h, \sigma)$ be a 1-system and $m : P(\Sigma) \to P(\Sigma)$ a function, where $P(\Sigma)$ is the set of subsets of Σ, such that

$$m(\phi) = \phi ,$$

$$m(\{a\}) = \min(h(a)) \quad \text{for } a \in \Sigma ,$$

$$m(A \cup B) = m(A) \cup m(B) .$$

The 1-system G is said to be normal if

$$a \in m^j(b) , \ j > 0 \text{ implies } a \in m(b)$$

holds for every $a,b \in \Sigma_c$. The following result immediately follows from [4, Lemma 2].

Lemma 2.8

(i) For each 1-system $G = (\Sigma, h, \sigma)$ one can find effectively an integer k such that the 1-system $G^k = (\Sigma, h^k, \sigma)$ is normal.

(ii) For each pair of normal 1-systems $G_i = (\Sigma, h_i, \sigma)$, $i = 1,2$ one can find effectively an integer k such that the 1-systems $G_i^k = (\Sigma, h_i(h_1h_2)^k, \sigma)$, $i = 1,2$ are normal.

□

The morphism $(h_1h_2)^k$ in (ii) was called a <u>normal combination</u> of (G_1, G_2) in [4].

Combining Lemmas 2.5, 2.7 and 2.8(i) we obtain

<u>Lemma 2.9</u>

The limit language equivalence problem is decidable for DOL systems iff it is decidable for normal 1-systems.

□

Let $G_i = (\Sigma, h_i, \sigma)$ be two DOL systems for $i = 1, 2$. Define a mapping $\beta : \Sigma^* \to \mathbb{Z}$ by setting

$$\beta(x) = |h_1(x)| - |h_2(x)| \quad .$$

The integer $\beta(x)$ is called the <u>balance</u> of the word x (with respect to h_1 and h_2). The systems G_1 and G_2 are of <u>bounded balance</u> if there exists an integer k such that

$$|\beta(x)| \leq k$$

whenever x is a prefix of a word in $L(G_1)$. The following result has been shown in [7].

<u>Theorem 2.10</u>

If G_1 and G_2 are DOL sequence equivalent, then G_1 and G_2 are of bounded balance. □

The boundedness of the balance concerns also the subwords of $L(G_1)$, [3].

Lemma 2.11

If G_1 and G_2 are of bounded balance then there exists an integer k such that $|\beta(x)| \leqq k$ for each subword x of $L(G_1)$.

□

In the next chapter we shall need some facts and notations about the growth matrices of DOL systems, [12].

Let $G = (\Sigma, h, \sigma)$ be a DOL system and A its growth matrix. Hence

$$[x]A\eta = |h(x)| ,$$

where $[x]$ is the Parikh-vector of the word $x \in \Sigma^*$ and $\eta = (1, \ldots, 1)^T$ is a column vector (of suitable order) consisting entirely of ones.

If G is a 1-system then there is a permutation matrix B such that

$$BAB^T = \begin{pmatrix} I & C \\ 0 & D \end{pmatrix} , \qquad (2)$$

where I is a 1 x 1 identity matrix, C is a nonzero vector and D is a square matrix. Without contradiction we may identify A and BAB^T .

For the following matrix theoretical results we refer to [9]. Let M be a nonnegative matrix and let $\rho(M)$ be the real and positive characteristic value of M such that

$$\rho(M) \geqq |\nu|$$

for any other characteristic value ν .

Furthermore denote by $a_{i,j}^{(n)}$ the (i, j) element of M^n. The matrix M is irreducible if for each (i, j) there is an integer n such that $a_{i,j}^{(n)} > 0$. M is said to be primitive if there is an integer n such that M^n is positive, i.e. if $a_{i,j}^{(n)} > 0$ for each (i, j).

Theorem 2.12

If M is primitive then $\rho(M)$ is a simple characteristic value and greater than the absolute value of any other characteristic value. Furthermore M has a positive characteristic vector ν (called the maximal (characteristic) vector) corresponding to $\rho(M)$ and any non-negative characteristic vector of M is a scalar multiple of ν.

□

Lemma 2.13

Let $G = (\Sigma, h, \sigma)$ be a 1-system and (2) its growth matrix. We have

$$\rho(A) = \rho(D) \geqq 1$$

and if D is primitive then $\rho(D) > 1$.

Proof

Since G is a 1-system we have

$$[\sigma]A^n < [\sigma]A^{n+1}.$$

for each $n \geqq 0$. The claim follows from this.

□

3. 1-Simple Systems

A 1-system $G = (\Sigma, h, \sigma)$ is called 1-simple if the growth matrix of h restricted to Σ_c is primitive. Thus the growth matrix of a 1-simple system may be transformed by suitable permutations of rows and columns into the form

$$(1) \qquad A = \begin{pmatrix} I & C \\ 0 & P \end{pmatrix} ,$$

where I is a 1x1 identity matrix, P is a primitive (square) matrix, and C is a non-zero vector.

For the primitive matrix P we have the following property from [1]

$$(2) \qquad \lim_{n\to\infty} \frac{xP^n}{\rho(P)^n} = d_x \cdot \nu ,$$

where ν is the normalized positive characteristic vector of P corresponding to $\rho(P)$ and d_x is a scalar constant depending only on x. By (2) we obtain

$$\lim_{n\to\infty} \frac{xP^n\eta}{\rho(P)^n} = d_x$$

and thus

$$(3) \qquad \lim_{n\to\infty} \frac{xP^n}{xP^n\eta} = \nu .$$

Here the limit ν is independent of the start vector x.

We shall show that the corresponding limit exists for A, too.

Lemma 3.1

$$\lim_{n\to\infty} \frac{yA^n}{yA^n\eta} = (0, \nu) = \nu_o .$$

Proof

Suppose first that $y = (1, 0)$ and assume that $yA = (1, x)$, where the vector x is non-zero because of (1). By the form of the matrix A we have

$$yA^n = \left(1, \sum_{i=0}^{n-1} xP^i\right).$$

Let $\varepsilon > 0$ be a real number. There are integers r and n_o such that for $n \geqq n_o$

(4) $$\frac{\sum_{i=0}^{r} xP^i}{\sum_{i=0}^{n-1} xP^i\eta} < \varepsilon\cdot\eta^T$$

and

(5) $$\left|\frac{xP^i}{xP^i\eta} - \nu\right| < \varepsilon\cdot\eta^T$$

whenever $i \geqq r$. Using (4) and (5) it is not difficult to show

$$\left|\frac{\sum_{i=0}^{n-1} xP^i}{\sum_{i=0}^{n-1} xP^i\eta} - \nu\right| \leqq K\cdot\varepsilon\cdot\eta^T$$

for some $K > 0$.

This proves the claim for $y = (1, 0)$. If $y = (0, x)$ for some non-zero x then the claim holds true by (3). The general case follows from these two cases when a vector is represented as a sum of vectors of the above form. Since this general statement is never used in what follows we omit the details here.

□

We shall generalize Lemma 3.1 to the limit word α generated by the morphism h. Let α_n be the prefix of α of length n. In particular $\alpha_1 = \sigma$.

Lemma 3.2

$$\lim_{n\to\infty} \frac{[\alpha_n]}{n} = \nu_0 .$$

Proof

Let $\varepsilon > 0$ be a real number. By Lemma 3.1 there is an integer n_0 such that

$$\left| \frac{[h^n(\alpha_1)]}{|h^n(\alpha_1)|} - \nu_0 \right| < \varepsilon\eta^T \tag{6}$$

for $n \geqq n_0$. Consider a prefix β of α of the form

$$\beta = h^n(\alpha_1)h^{n_0}(x) , \tag{7}$$

where $n \geqq n_0$ and β is a prefix of $h^{n+1}(\alpha_1)$. We have

$$\left|\frac{[\beta]}{|\beta|} - \nu_o\right| = \left|\frac{[h^n(\alpha_1)] + [h^{n_o}(x)]}{|h^n(\alpha_1)| + |h^{n_o}(x)|} - \nu_o\right|$$

$$\leq \frac{\left|\frac{[h^n(\alpha_1)]}{|h^n(\alpha_1)|} - \nu_o\right| \cdot |h^n(\alpha_1)| + \left|\frac{[h^{n_o}(x)]}{|h^{n_o}(x)|} - \nu_o\right| \cdot |h^{n_o}(x)|}{|h^n(\alpha_1)| + |h^{n_o}(x)|}$$

$$< \quad \varepsilon \cdot \eta^T \quad ,$$

by (6) and the fact that

$$\frac{[h^{n_o}(x)]}{|h^{n_o}(x)|} = \left(0, \frac{[x]P^{n_o}}{[x]P^{n_o}\eta}\right) .$$

For any prefix α_n , $n \geqq n_o$, there are prefixes β_1 and β_2 of the form (7) such that β_1 is a prefix of α_n , α_n is a prefix of β_2 and

$$0 \leqq |\beta_2| - |\beta_1| < \max\{|h^{n_o}(a)| : a \in \Sigma\} .$$

Since

$$\frac{\max\{|h^{n_o}(a)| : a \in \Sigma\}}{|h^n(\alpha_1)|} < \varepsilon ,$$

for $n \geqq n_1$, the claim follows.

□

<u>Lemma 3.3</u>

If $G_i = (\Sigma, h_i, \sigma)$, $i = 1, 2$, are 1-simple systems which are limit language equivalent then the maximal characteristic vectors of the two systems are equal.

Proof

This follows immediately from Lemmas 3.1 and 3.2 and from the equality in (3).

□

We shall now turn to investigate the balance of two limit language equivalent 1-simple systems. The previous result is not sufficient for showing any bounds for the balance since it does not say anything about the maximal characteristic values of the two growth matrices. Thus a lemma of the following kind is needed. So, in the following lemma we add an assumption which implies the equality of the maximal characteristic values. It will suit our purposes later.

Lemma 3.4

If $G_1 = (\Sigma, h, \sigma)$ and $G_2 = (\Sigma, g, \sigma)$ are limit language equivalent and 1-simple and there are morphisms h_1 and h_2 such that $h = h_1 h_2$ and $g = h_2 h_1$, then the maximal characteristic values and vectors of the two systems are equal.

Proof

The claim follows by Lemma 3.3 and by the fact that the characteristic values of AB and BA are the same for any matrices A and B .

□

Lemma 3.5

Let $G_i = (\Sigma, h_i, \sigma)$ be two limit language equivalent 1-simple systems with a common maximal characteristic value and let α be their common limit word. If α_n denotes the prefix of α of length n then

$$\lim_{n\to\infty} \frac{|\beta(\alpha_n)|}{n} = 0 ,$$

where β is the balance function of the pair (G_1, G_2) .

Proof

Let $\varepsilon > 0$ be a real number. By Lemma 3.3 there is an integer n_0 such that

$$\left|\frac{[\alpha_n]}{n} - \nu_0\right| < \varepsilon\eta^T \tag{8}$$

for $n \geqq n_0$ and for the maximal characteristic vector ν_0 common to G_1 and G_2 . Furthermore by the definition of β

$$\frac{|\beta(\alpha_n)|}{n} = \left|\frac{[\alpha_n](A_1 - A_2)\eta}{n}\right| ,$$

where A_i is the growth matrix corresponding to G_i $(i = 1, 2)$. Thus

$$\frac{|\beta(\alpha_n)|}{n} = \left|\frac{([\alpha_n] - n\cdot\nu_0 + n\cdot\nu_0)(A_1 - A_2)\eta}{n}\right|$$

$$\leqq \left|\frac{([\alpha_n] - n\cdot\nu_0)(A_1 - A_2)\eta}{n}\right| + \left|\frac{n\cdot\nu_0(A_1 - A_2)\eta}{n}\right| .$$

Here $\nu_0(A_1 - A_2) = 0$ since ν_0 is a characteristic vector of A_1 and A_2 which corresponds to the same characteristic value. Hence by (8) we have

$$\frac{|\beta(\alpha_n)|}{n} < \varepsilon\eta^T(A_1 + A_2)\eta ,$$

where $\eta^T(A_1 + A_2)\eta$ is a constant scalar independent of n . This proves the claim.

□

Let G_1 and G_2 be as in Lemma 3.5 and let

$$c_i = \max\{|\beta(x)| : x <_{pr} h_1^i(\sigma)\} .$$

By the previous lemma we have

(9) $$\lim_{n\to\infty} \frac{c_n}{\lambda^{n-k}} = 0$$

for each k . By Lemma 3.1

$$\lim_{n\to\infty} \frac{|h_1^n(a)|}{\lambda^n} > 0$$

for all $a \in \Sigma_c$.

(Here λ denotes the maximal characteristic value of the growth matrices for G_1 and G_2 .) Thus for each $k \geqq 0$ there exists an integer n_k such that

(10) $$|h_1^{n-k}(a)| > c_n$$

for all $a \in \Sigma_c$ and $n \geqq n_k$.

Let us now fix (for the rest of this section) two limit language equivalent 1-systems

$$G_i = (\Sigma, h_i, \sigma) , \quad i = 1, 2$$

In the next two lemmas we show that the balance of the systems

(Σ, h_1h_2, σ) and (Σ, h_2h_1, σ) is bounded whenever they are 1-simple.

Lemma 3.6

If $G_{12} = (\Sigma, h_1h_2, \sigma)$ and $G_{21} = (\Sigma, h_2h_1, \sigma)$ are 1-simple, then either they are both exponentially growing or $\#\Sigma_c = 1$ and

$$h_1h_2(a) = h_2h_1(a) = a ,$$

where $\Sigma_c = \{a\}$. In the latter case the systems G_{12} and G_{21} are DOL equivalent and their balance is zero.

Proof

If $\#\Sigma_c > 1$ then the systems are exponentially growing since they are 1-simple. Suppose that $\Sigma_c = \{a\}$. If $|h_1h_2(a)| > 1$ and $|h_2h_1(a)| > 1$ then the systems grow exponentially.

Otherwise $h_1h_2(a) = a$ or $h_2h_1(a) = a$, which implies that $h_1h_2(a) = h_2h_1(a) = a$ since $h_i(a) \in \Sigma_c^*$ for $i = 1, 2$. Furthermore if

$$h_1(\sigma) = \sigma a^r , \quad h_2(\sigma) = \sigma a^s$$

then

$$h_1h_2(\sigma) = \sigma a^{r+s} = h_2h_1(\sigma) \quad .$$

Thus the systems are DOL equivalent. Moreover

$$h_1h_2(\sigma a^n) = \sigma a^{r+s+n} = h_2h_1(\sigma a^n)$$

for each $n \geqq 0$ and hence the balance of the systems G_{12} and G_{21} is zero. In fact $G_{12} = G_{21}$. □

The following lemma is analogous to Theorem 2 in [4], however, its proof is considerably harder. It uses the "shifting lemma" from [2] which says that the existence of a local strict maximum of balance would imply the existence of a substring of the form v^2 at the point of the maximum. However, unlike in [4], this does not immediately yield a contradiction.

Lemma 3.7

If $G_{12} = (\Sigma, h_1h_2, \sigma)$ and $G_{21} = (\Sigma, h_2h_1, \sigma)$ are 1-simple (and corresponding to the limit language equivalent 1-systems G_1 , G_2) and they are exponentially growing then their balance is bounded.

Proof

Assume the contrary, i.e. that the balance is unbounded (Assumption 1). Denote $g_1 = h_1h_2$, $g_2 = h_2h_1$ and let

$$w_{n_0,0} = g_1^{n_0}(\sigma) = \sigma a_1a_2\ldots a_r ,$$

where n_0 is large enough, $a_i \in \Sigma_c$, $1 \leqq i \leqq r$, such that if ab is a subword of any $g_1^n(\sigma)$, $n \geqq 1$, then ab is a subword of $w_{n_0,0}$, $a,b \in \Sigma_c$. Furthermore we write

$$w_{j,i} = g_1^j(\sigma)g_1^j(a_1a_2\ldots a_i)$$

for $i = 1,2,\ldots,r$ and $j \geqq n_0$. Thus $w_{j,i}$ is always a prefix of the limit word α for each j and i .

By Assumption 1 for each $n > 0$ there exists integers i , j ($1 \leqq i \leqq r$, $j \geqq n$) and a prefix u of α such that $u <_{pr} w_{j,i}$ and

(i) $|\beta(w)| < |\beta(u)|$ for $w <_{pr} u$,

(ii) $|\beta(w)| \leqq |\beta(u)|$ for $w <_{pr} w_{j,i}$.

Thus we select those prefixes u of α whose balance is maximal with respect to the words $w_{j,i}$. It was shown in [2] (or Thm 2 in [4]) that if u satisfies the conditions (i) and (ii) for sufficiently large j ($j \geqq n_1$) then there exists a nonempty word v such that $\alpha = u'v^2\alpha'$, where $u = u'v$, and

(11) $$|v| < |g_1^j(a)| \text{ , for all } a \in \Sigma \text{ .}$$

Note that $u <_{pr} w_{j,i}$ for some i and thus $u <_{pr} g_1^{j+n_0}(\sigma)$. The "shift" word v is of length at most $|\beta(u)|$ by [2]. The inequality (11) follows now from (10).

Suppose that $\beta(u) > 0$, the case $\beta(u) < 0$ is quite analogous. From the above we deduce that $\beta(v) > 0$ since the case (i) implies that $|\beta(u')| < |\beta(u)|$. Thus

$$\beta(u) < \beta(uv)$$

and hence uv cannot be a prefix of $w_{j,i}$ by (ii). However, by (11)

(12) $$uv <_{pr} \mathrm{suc}(w_{j,i}) \text{ ,}$$

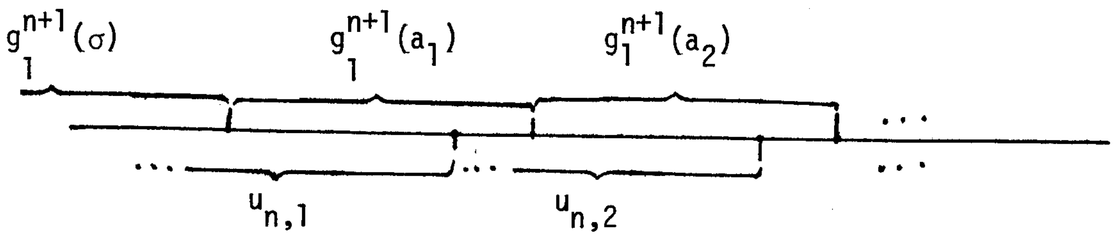
where $\mathrm{suc}(w_{j,i}) = w_{j,i+1}$ if $i < r$ and $\mathrm{suc}(w_{j,i}) = w_{j+1,1}$ if $i = r$. By (12) it follows that $\mathrm{suc}(w_{j,i})$ has a prefix u_1 satisfying the conditions (i) and (ii) such that $w_{j,i} <_{pr} u_1$. Proceeding inductively we find a sequence $\{u_{t,m}\}$, $t \geqq j$, $1 \leqq m \leqq r$, starting from u such that for each t and m

(13) $$\beta(u_{t,m}) < \beta(suc(u_{t,m}))$$

and

(14) $$w_{t,m} <_{pr} u_{t,m} <_{pr} suc(w_{t,m}) \quad ,$$

where $suc(u_{t,m})$ is defined analogously to $suc(w_{t,m})$.



Consider the subsequence where $t = j+1$ and $m = 1,2,\ldots,r$ in (13) and (14). We may write for $m = 1,2,\ldots,r$

$$g_1^{j+1}(a_m) = v_m z_m \quad ,$$

where $\beta(v_m) > 0$ and $\beta(z_m) \leqq 0$, and $u_{j,m} = u_{j,m-1} z_{m-1} v_m$.

By the above considerations we have that $\beta(z_{m-1} v_m) > 0$ and thus

$$|\beta(v_m)| > |\beta(z_{m-1})|$$

for each m . We may conclude that if $x = b_1 b_2 \ldots b_s$ is a subword of $w_{n_o,o}$ then

$$\beta(g_1^{j+1}(x)) \geqq s + \beta(v)$$

where $v = z_m$ for some m (namely for $b_s = a_m$) . This implies that

$$\beta(g_1^{j+1}(x)) > 0$$

whenever $|x| > \max\{|\beta(z_m)| : m = 1,2,\ldots,r\}$. Since j is fixed the right hand side of this inequality is also fixed. Therefore there exists an integer k such that

$$\beta(g_1^{k}(a)) > 0$$

for each $a \in \Sigma_c$. Thus for $n \geqq 0$

$$\beta(g_1^{n+k}(a)) \geqq |g_1^{n}(a)|$$

for each $a \in \Sigma_c$.

On the other hand there is a constant α such that

$$|g_1^{n+k}(a)| < \alpha^k |g_1^{n}(a)|$$

and hence

$$\beta(g_1^{n+k}(a)) > |g_1^{n+k}(a)|/\alpha^k$$

for $n \geqq 0$. This contradicts Lemma 3.5 (see also the limit in (9)) and proves the claim.

$\square$

4. The General Case

Given a 1-system $G = (\Sigma, h, \sigma)$ a set $\pi \subsetneq \Sigma_c$ is called a subalphabet of G if $h(\pi) \subseteq \pi^*$. Denote $\Omega = \Sigma - \pi$ and let x^Ω be a word in Ω^* obtained by deleting the symbols from π in x. Furthermore set $h^\Omega(x) = h(x)^\Omega$ and $G^\Omega = (\Omega, h^\Omega, \sigma)$. A set π is called a common subalphabet of the 1-systems G_1 and G_2 if it is a subalphabet of both of them. Note, that in distinction with [4] we are not requiring that $\pi \neq \phi$.

Lemma 4.1

If G_1 and G_2 are limit language equivalent then so are G_1^Ω and G_2^Ω for any common subalphabet π. Moreover, if G_1 is normal, then so is G_1^Ω.

Proof

Immediate by the definitions.

□

Let us fix for Lemmas 4.2-4.4 two normal 1-systems

$$G_i = (\Sigma, h_i, \sigma), \quad i = 1,2$$

which are limit language equivalent.

Lemma 4.2

There is a morphism $h \in \langle h_1, h_2 \rangle$ and a common subalphabet π of

$$G_{1,i} = (\Sigma, h_i h, \sigma), \quad i = 1,2$$

such that $G_{1,1}$ and $G_{1,2}$ are normal and $G^{\Omega}_{1,1}$ and $G^{\Omega}_{1,2}$ are propagating for $\Omega = \Sigma - \pi$. Furthermore, π and h can be found effectively.

Proof

If h_1 and h_2 are already propagating then $\pi = \phi$ and h is the identity morphism. Otherwise we define a sequence of pairs of morphisms $(g_{1,i}, g_{2,i})$ as follows. Set $g_{j,o} = h_j$ for $j = 1,2$. Suppose now that

$$\pi_t = \{a : g_{j,t}(a) = e\}$$

is nonempty for $t \geqq 0$ and for $j = 1$ or 2 . Then we define

$$g_{j,t+1} = g_{j,t}(g_{i,t} \; g_{j,t})^k$$

and

$$g_{i,t+1} = g_{i,t}(g_{i,t} \; g_{j,t})^k ,$$

where $i \neq j$ and the two morphisms are normal (by the choice of k , see Lemma 2.8). Now,

$$g_{1,t+1}(\pi_t) = g_{2,t+1}(\pi_t) = \{e\}$$

and thus π_t is a common subalphabet of the systems $(\Sigma, g_{i,t+1}, \sigma)$, $i = 1,2$. Furthermore $\pi_{t-1} \subseteq \pi_t$. Hence there is an integer $p \leqq \#\Sigma$ such that $\pi_p = \pi_{p+1}$. Moreover $\Sigma_c - \pi_p \neq \phi$ since otherwise the morphisms $g_{1,p}$ and $g_{2,p}$ would not define an infinite limit word contradicting Theorem 2.6. This proves the lemma when we choose $\pi = \pi_p$.

□

The next lemma appears already in [4] for DOL equivalence.

Lemma 4.3

If G_1 and G_2 are propagating then they have a nonempty common subalphabet or the morphisms h_1h_2 and h_2h_1 are 1-simple.

Proof

The proof given in [4] (Lemma 5 in [4]) proves the present claim when we note that $\Delta_1 = \Delta_2 = \Sigma_c$, where Δ_i is the subset of Σ_c of symbols which occur infinitely often in $L(G_i)$, $i = 1, 2$. We remind here that G_1 and G_2 are 1-systems and thus $\Delta_i = \Sigma_c$ for $i = 1, 2$.

□

The previous lemmas have the following corollary.

Lemma 4.4

There is a morphism $h \in \langle h_1, h_2 \rangle$ and a common subalphabet π of normal

$$\hat{G}_i = (\Sigma, h_i h, \sigma) , \quad i = 1,2$$

such that $(h_1hh_2h)^{\Omega}$ and $(h_2hh_1h)^{\Omega}$ are 1-simple for $\Omega = \Sigma - \pi$. Furthermore π and h can be found effectively.

Proof (i) By Lemma 4.2 we find a combination g_1 of h_1 and h_2 and a common subalphabet $\pi_{1,1}$ such that the 1-systems

$$G_{1,i} = (\Sigma, h_i g_1, \sigma) , \quad i = 1,2$$

are normal and $G_{1,1}^{\Omega_{1,1}}$ and $G_{1,2}^{\Omega_{1,1}}$ are propagating for $\Omega_{1,1} = \Sigma - \pi_{1,1}$.

(ii) By Lemma 4.3 the morphisms $h_1g_1h_2g_1$ and $h_2g_1h_1g_1$ are 1-simple or there is a common subalphabet $\pi_{1,2} \neq \phi$ of $G_{1,1}^{\Omega_{1,1}}$ and $G_{1,2}^{\Omega_{1,1}}$. If the first case holds the claim follows. Otherwise we observe that the set

$$\pi_1 = \pi_{1,1} \cup \pi_{1,2}$$

is a common subalphabet of $G_{1,1}$ and $G_{1,2}$. We may assume that $\pi_{1,2}$ is a maximal common subalphabet. If $G_{1,i}^{\Omega_1}$ are propagating for $\Omega_1 = \Sigma - \pi_1$, $i = 1,2$ then the claim follows by Lemma 4.3. Suppose that $G_{1,i}^{\Omega_1}$ is nonpropagating for $i = 1$ or 2 . Then we start the procedure from (i) on again but now for $G_{1,1}$ and $G_{1,2}$.

Thus we obtain a sequence of pairs of normal 1-systems

$$G_{j,i} = (\Sigma, h_i g_j, \sigma) , \quad j = 1,2,\ldots$$

and a sequence

$$\pi_j = \pi_{j,1} \cup \pi_{j,2} , \quad j = 1,2,\ldots$$

of common subalphabets of $G_{j,1}$ and $G_{j,2}$ such that $\pi_j \subseteq \pi_{j+1}$. Thus this procedure will terminate for some $j \leqq \#\Sigma$ when $\pi_j = \pi_{j+1}$. Denote $\Omega_j = \Sigma - \pi_j$ for $j = 1,2,\ldots$.

<u>Claim 4.5</u>

$\Omega_j - \{\sigma\} \neq \phi$ for each $j \geqq 1$.

Proof of the claim.

Assume the contrary and let j be the least integer such that $\Omega_j = \{\sigma\}$ and thus $\pi_j = \Sigma_c$. By the beginning of the proof we know that $j > 1$. Denote $\Omega'_{j-1} = \Omega_{j-1} - \{\sigma\}$. By the assumption $\Omega_j - \{\sigma\} = \phi$ it follows that

$$(h_i g_j)^{\Omega'_{j-1}}(\Omega'_{j-1}) = \{e\} , \quad i = 1,2$$

and thus

$$h_i g_j(a) \in \pi^*_{j-1} , \quad i = 1,2$$

for each $a \in \Omega'_{j-1}$. Because

$$h_i g_j(\pi_{j-1}) \subseteq \pi^*_{j-1}$$

it follows that there are no symbols $b \in \Sigma_c$ such that

$$a \in \min(h_i g_j(b)) , \quad i = 1,2$$

if $a \in \Omega'_{j-1}$. Hence the symbols in Ω'_{j-1} would occur only finitely many times in the limit word of $h_i g_j$. This contradicts the fact that $h_i g_j$ is a 1-system and the fact that $\Omega'_{j-1} \neq \phi$. Thus the claim has been proved. By Claim 4.5 the systems obtained in (i) are all propagating and thus the systems $G_{j,i}^{\Omega_j}$ are nontrivial for $j = 1,2,\ldots,\#\Sigma$. Hence $G_{j,1}^{\Omega_j}$ and $G_{j,2}^{\Omega_j}$ are propagating for some $j \leq \#\Sigma$ and the lemma follows by Lemma 4.3.

□

The previous lemmas further reduce the general problem:

Lemma 4.6

The limit language equivalence problem is decidable for DOL systems iff it is decidable for pairs of normal 1-systems

$$G_i = (\Sigma, h_i, \sigma) , \quad i = 1,2$$

with the following property (P) : G_1 and G_2 have a common subalphabet π such that $(h_1h_2)^{\Omega}$ and $(h_2h_1)^{\Omega}$ are 1-simple for $\Omega = \Sigma - \pi$.

Proof

The Lemmas 4.2-4.4 are clearly effective in the construction of the common subalphabets and of the morphisms. The lemma follows now by Theorem 2.6 and Lemma 2.9.

□

In the next lemma we prove that the property (P) implies bounded balance of the "commutator systems".

Lemma 4.7

Let $G_i = (\Sigma, h_i, \sigma)$, $i = 1,2$, be two normal limit language equivalent 1-systems with the property (P) . Then the 1-systems $G_{12} = (\Sigma, h_1h_2, \sigma)$ and $G_{21} = (\Sigma, h_2h_1, \sigma)$ have bounded balance.

Proof

Let π be the common subalphabet of G_1 and G_2 such that G_{12}^{Ω} and G_{21}^{Ω} are 1-simple for $\Omega = \Sigma - \pi$. The 1-simple systems G_{12}^{Ω} and G_{21}^{Ω} are limit language equivalent and thus by Lemmas 3.6 and 3.7 their balance is bounded: (β^{Ω} denotes the balance with respect to morphisms restricted to Ω)

$$|\beta^{\Omega}(x)| \leq B$$

for all subwords x of the limit word α . Moreover by Lemma 3.6 there is a constant k such that for each $a \in \Omega$

(1) $$|(h_1h_2)^k(a)^\Omega| > B^2 .$$

Note that we write $h(a)^\Omega$ instead of more precise $(h(a))^\Omega$. Define $g_1 = (h_1h_2)^k$ and $g_2 = (h_2h_1)(h_1h_2)^{k-1}$. Clearly, the balance of g_1^Ω and g_2^Ω is also bounded by B . Denote $g_1(\sigma)^\Omega = \sigma x$ and

$$g_1^n(x)^\Omega = a_n x_n ,$$

where $a_n \in \Omega$ for $n \geqq 1$. For each $n \geqq 1$ there are two words u_n and v_n such that

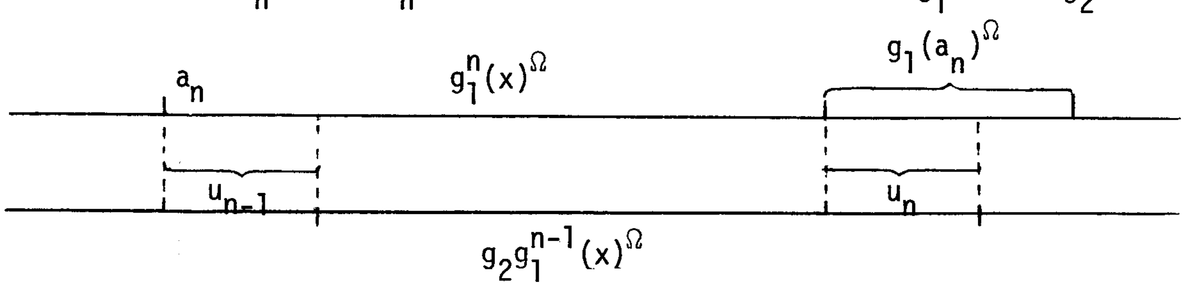
(a) $g_1^n(\sigma)^\Omega u_n = g_2 g_1^{n-1}(\sigma)^\Omega v_n$;

(b) $|u_n v_n| \leqq B$;

(c) $u_n = e$ or $v_n = e$;

(d) $u_n <_{pr} g_1(a_n)^\Omega$, $v_n <_{pr} g_2(a_n)^\Omega$.

Here the words u_n and v_n are the "balance words" for g_1 and g_2 .



Claim 4.8

There are integers n_o and p such that

$$u_n = u_{n+ip} , \quad v_n = v_{n+ip}$$

whenever $n \geqq n_0$ and $i \geqq 0$.

Proof of the claim.

Note first that there are integers n_1 and q such that $a_n = a_{n+iq}$ whenever $n \geqq n_1$ and $i \geqq 0$. Thus, for $n \geqq n_1$ and $k \geqq j$,

(2) $$g_1^{n+jq}(a_n)^{\Omega} <_{pr} g_1^{n+kq}(a_n)^{\Omega} .$$

Suppose that there is an integer $m \geqq n_1$ such that

(3) $$u_{m+iq} \neq u_{m+(i+1)q} \quad \text{or} \quad v_{m+iq} \neq v_{m+(i+1)q}$$

for infinitely many integers i . If no such m exists then the claim follows immediately. By (2) we obtain for each $j \geqq 0$ and $k > j$ that

$$v_{m+kq} g_1^{m+jq}(a_m)^{\Omega} <_{pr} u_{m+kq} g_2 g_1^{m+(j+1)q}(a_m)^{\Omega} .$$

Here $|u_{m+kq}\ v_{m+kq}| \leqq B$ and thus the assumption (2) implies that the words $g_1^{m+jq}(a_m)^{\Omega}$ are periodic for $j \geqq 0$. Moreover by (2) the periods of these words must be equal for each $j \geqq 0$. Thus we may write

$$g_1^{m+jq}(a_m)^{\Omega} = z^{t_j} z_j ,$$

where z is the minimal period, $|z| \leqq B$ and $z_j <_{pr} z$. The morphism g_1^{Ω} is 1-simple and thus

$$\min(g_1^{m}(a_m)^{\Omega}) = \Omega - \{\sigma\} .$$

Let $a \in \Omega - \{\sigma\}$. The words $g_1^{jq}(a)$ are subwords in $g_1^{m+jq}(a)$ for $j \geqq 0$ and hence they are periodic, too. Thus

$$g_1^{jq}(a)^\Omega = z_{j,a} z^{t_{j,a}} \hat{z}_{j,a}$$

for $a \in \Omega - \{\sigma\}$, $j \geqq 0$, where $z_{j,a} <_{po} z$ and $\hat{z}_{j,a} <_{pr} z$. The subwords $z_{j,a}$ and $\hat{z}_{j,a}$ form a periodic sequence for $j = 0,1,\ldots$ since their lengths are bounded by B . Let n_2 and p_1 be integers such that

(4) $$z_{j,a} = z_{j+p_1,a} , \quad \hat{z}_{j,a} = \hat{z}_{j+p_1,a}$$

whenever $j \geqq n_2$ and $a \in \Omega - \{\sigma\}$. By (1) the balance of z must be zero (with respect to g_1^Ω and g_2^Ω), i.e. $\beta^\Omega(z) = 0$, since otherwise we would have for sufficiently large n that

$$|\beta^\Omega(g_1^n(a_m))| > B .$$

The periodicity of the sequences in (4) implies now that

$$\beta^\Omega(g_1^{jq}(a)^\Omega) = \beta^\Omega(g_1^{(j+ip_1)q}(a)^\Omega)$$

for $j \geqq n_2$, $i \geqq 0$ and $a \in \Omega - \{\sigma\}$. Hence,

$$\beta^\Omega(g_1^{jq}(y)^\Omega) = \beta^\Omega(g_1^{jq+ip_1q}(y)^\Omega)$$

for $y \in (\Omega - \{\sigma\})^*$ and $j \geqq n_2$, $i \geqq 0$. Set $y = g_1^s(x)$, $s \geqq 0$. Thus, we have

$$\beta^\Omega(g_1^n(x)^\Omega) = \beta^\Omega(g_1^{n+ip}(x)^\Omega)$$

for $n \geqq n_2q$ and $p = p_1q$. This proves the claim since $g_1(\sigma) = \sigma x$.

In order to complete our proof of Lemma 4.7 our next task is to transform the 1-systems (Σ, g_i, σ) , $i = 1,2$ into two sequence equivalent DOL systems. Let $a^{(i)}$, $i = 0,1,\ldots,p-1$, be new symbols for each $a \in \Omega - \{\sigma\}$ and denote

$$\hat{\Sigma} = \Sigma \cup \{a^{(i)} : a \in \Omega - \{\sigma\},\ 0 \leqq i \leqq p-1\} \cup \{\hat{\sigma}\}$$

We write $<j,p>$ for the integer k , where $0 \leqq k \leqq p-1$ and $j \equiv k \pmod p$. Define two morphisms φ_1 and φ_2 as follows.

(i) $$\varphi_j(\hat{\sigma}) = \sigma y a^{(1)} , \quad j = 1,2 ,$$

if $g_1^{n_o}(\sigma) = \sigma y a y_1$, where $(ay_1)^{\Omega} = a$;

(ii) $$\varphi_j(a) = g_j(a) , \quad j = 1,2 ,$$

if $a \in \Sigma$

(iii) $$\varphi_1(a^{(i)}) = yb^{(<i+1,p>)} ,$$

if $g_1(a) = yby_1$, where $(by_1)^{\Omega} = b$;

(iv) $$\varphi_2(a^{(i)}) = yb^{(<i+1,p>)} ,$$

if $u_{n_o+i} \neq e$ and $g_2(a) = yby_1$, where $(by_1)^{\Omega} = bu_{n_o+i}$;

(v) $$\varphi_2(a^{(i)}) = yb^{(<i+1,p>)} ,$$

if $u_{n_o+i} = e$ and $yb <_{pr} g_2(a)g_2(a_{n_o+i})$ such that $b \in \Omega$ and

$$(yb)^{\Omega} = g_2(a)^{\Omega} v_{n_o+i} .$$

(Note that possibly $u_{n_o+i} = v_{n_o+i} = e$.) By (i)-(iii) we have: If $g_1^{n_o+i}(\sigma) = yay_1$ and $(ay_1)^{\Omega} = a$ (i.e. if $a \in \Omega$ and $y_1 \in \pi^*$), then

$\varphi_1^i(\hat{\sigma}) = ya^{(<i+1,p>)}$. In particular

$$g_1^n(\sigma) <_{pr} \varphi_1^i(\hat{\sigma}) \tag{5}$$

for $n < i+n_0$. Furthermore, by (a) there are words y_n and z_n such that

$$g_1^n(\sigma)y_n = g_2g_1^{n-1}(\sigma)z_n$$

for $n \geqq 1$, where $y_n^\Omega = u_n$ and $z_n^\Omega = v_n$. Thus by (iv), (v) and (d)

$$\varphi_1^i(\hat{\sigma}) = \varphi_2\varphi_1^{i-1}(\hat{\sigma})$$

for $i \geqq 1$. Hence the systems $(\hat{\Sigma}, \varphi_i, \hat{\sigma})$, $i = 1,2$ are sequence equivalent DOL systems and their balance is bounded by Theorem 2.10. The prefix property (5) implies that the morphisms g_1 and g_2 are of bounded balance on $L(G_{12})$, i.e. the systems (Σ, g_i, σ) , $i = 1,2$ have bounded balance. Finally, we reason that the systems (Σ, h_1h_2, σ) and (Σ, h_2h_1, σ) have bounded balance since

$$(h_1h_2)^n = g_1^{m_1}(h_1h_2)^{m_2} \quad ,$$

where $n = km_1 + m_2$, $m_2 < k$.

□

Let h and g be two endomorphisms on Σ^* . The <u>compatibility language</u> of h and g is defined by

$$\mathrm{Com}(h,g) = \{x : h(x) <_{pr} g(x) \quad \text{or} \quad g(x) <_{pr} h(x)\} .$$

Note the resemblance of Com(h,g) and Eq(h, g) as defined in [12].

In fact the equality language $Eq(h, g)$ is included in $Com(h, g)$. Denote by $Com_k(h, g)$ the subset of $Com(h, g)$ which has balance at most k , i.e.

$$Com_k(h, g) = \{x : x \in Com(h, g), x_1 <_{pr} x \rightarrow |\beta(x_1)| \leqq k\} .$$

The following lemma is an analogy of the result for bounded equality languages. Since the proof is obvious and similar to that given for $Eq_k(h, g)$, see [2] or [12], we just state the result.

<u>Lemma 4.9</u>

For each $k \geqq 0$ and endomorphisms h and g , the set $Com_k(h, g)$ is regular and can be effectively constructed.

□

Finally we collect the above results into the main theorem.

<u>Theorem 4.10</u>

The limit language equivalence problem is decidable for DOL systems.

<u>Proof</u>

By Lemma 4.6 we may restrict ourselves to normal 1-systems which have the property (P). Let $G_i = (\Sigma, h_i, \sigma)$ be two of such systems for $i = 1,2$.

We use two semidecision procedures, one for non-equivalence and one for equivalence.

(A) The first semidecision procedure computes $h_1^n(\sigma)$ and $h_2^n(\sigma)$ for $n = 0,1,\ldots$ and checks for each n if these words are comparable. Thus if G_1 and G_2 are not limit language equivalent then the procedure finds an integer n such that $h_1^n(\sigma)$ and $h_2^n(\sigma)$ are incomparable.

(B) The procedure for equivalence constructs the regular languages $\mathrm{Com}_k(h_1h_2, h_2h_1) = C_k$ for $k = 0,1,\ldots$ inductively. For each k one checks if

$$L(G_1) \subseteq C_k \quad (6)$$

The above inclusion can be checked effectively since $L(G_1)$ is a DOL language and C_k is a regular set by Lemma 4.9. If G_1 and G_2 are limit language equivalent, then by Lemma 4.7 the systems $G_{12} = (\Sigma, h_1h_2, \sigma)$ and $G_{21} = (\Sigma, h_2h_1, \sigma)$ have bounded balance and thus (6) holds for some $k \geqq 0$. On the other hand if (6) holds for some k then G_1 and G_2 are limit language equivalent since $L(G_1)$ is an infinite prefix-preserving language.

□

Acknowledgement

The authors are grateful to J. Karhumaki for comments to the preliminary version of this paper.

References

1. Bellman, R. "Introduction to Matrix Analysis". McGraw-Hill, New York, 1960.

2. Culik, II, K. On the decidability of the sequence equivalence problem for DOL-systems, Theoret. Comput. Sci. 3 (1976), 75-84.

3. Culik, II, K. The ultimate equivalence problem for DOL systems, Acta Inform. 10 (1978), 79-84.

4. Culik, II, K. and Fris, I. The decidability of the equivalence problem for DOL systems, Inform. and Control 35 (1977), 20-39.

5. Culik, II, K. "Homomorphisms: decidability, equality and test sets", in R. Book (Ed.), Formal Language Theory, Perspectives and Open Problems, 167-194, Academic Press, 1980.

6. Culik, II, K. and Salomaa, A. "On infinite words obtained by iterating morphisms", submitted to Theoret. Comput. Sci., also University of Waterloo, Department of Computer Science Research Report CS-80-29. (Rev. Feb.81)

7. Ehrenfeucht, A. and Rozenberg, G. "Every two equivalent DOL systems have a regular true envelope", Theoret. Comput. Sci. 10 (1980), 45-52.

8. Eilenberg, S. Automata, Languages and Machines, Vol. A. Academic Press, New York, 1974.

9. Gantmacher, F.R. The Theory of Matrices, Vol. II, Chelsea, New York, 1959.

10. Linna, M. "The decidability of the DOL prefix problem", Internat. J. of Comput. Math. 6 (1977), 127-142.

11. Nivat, M. "Infinite words, infinite trees, infinite computations", in J.W. Barleer and J. van Leeuwen (Ed), "Foundations of Computer Science", III.2, 3-52, Mathematisch Centrum, Amsterdam, 1979.

12. Rozenberg, G. and Salomaa, A. The Mathematical Theory of L Systems, Academic Press, New York, 1980.

13. Salomaa, A. "Morphisms of free monoids and language theory", Proc. Conf. on Formal Languages in Santa Barbara, December 1979, Academic Press, in preparation.

14. Salomaa, A. Jewels in formal language theory, Computer Science Press, in preparation.

15. Thue, A. "Über unendliche Zeichenreichen", Videnskapsselsk Skrifter I, Kristiania (1906), 1-22.