DECIDABILITY OF SEQUENCE EQUIVALENCE FOR QUASI-SIMPLE DOL-SYSTEMS AND ALL SYSTEMS OVER TWO LETTERS*

K. Culik II and I. Fris†
Department of Computer Science
University of Waterloo

CS-76-14

March 1976

* This research was supported by the National Research Council of Canada, Grant No. A7403.

† On leave from the University of New England, Armidale, N.S.W., Australia.

THE DECIDABILITY OF SEQUENCE EQUIVALENCE FOR QUASI-SIMPLE DOL-SYSTEMS AND ALL SYSTEMS OVER TWO LETTERS*

by

K. Culik II and I. Fris†
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

Abstract. The notion of a quasi-simple pair of DOL-systems is introduced. It is shown that for a quasi-simple pair of DOL-systems the sequence equivalence is decidable. As an application it is shown that the sequence equivalence problem is decidable for (all) DOL systems over two-letter alphabet.

---

* This research was supported by the National Research Council of Canada, Grant No. A7403.

† On leave from the University of New England, Armidale, N.S.W., Australia.

Introduction

This paper is a sequel of [1]. We assume that the reader is familiar with [1] and we will use its terminology and notation.

In the first section we will sharpen the linear algebraic part (first half of section 3) of [1]. The proofs are somewhat simplified, worked out in greater detail and the unnecessary assumptions are dropped. This will allow us to relax the sufficient conditions on a pair of equivalent DOL-systems to have bounded balance, namely, we will require that at least one member of the pair, rather than both, is so called primitive. Further, we note that the proof of Theorem 3.2 from [1] goes through unchanged using these weaker assumptions. Therefore, we can apply the Theorem 2.1 from [1] and we have an algorithm to decide whether two DOL-systems are sequence equivalent if at least one of them is primitive.

The case when neither system of a pair is primitive can in some instances be reduced to the case when one is. This is done by suitable composition of the homomorphisms. The pair of DOL-systems for which this is possible is called quasi-simple. Each pair of simple DOL systems from [1] is quasi-simple and so we have extended the family of pairs of DOL-systems for which we can decide sequence equivalence and therefore also language equivalence using the result from [2].

Finally, as an application of our main result we show that the sequence equival nce problem is decidable for (all) DOL-systems over two-letter alphabet. This is done by considering separately a few singular cases which are not quasi-simple.

1. Quasi-simple pairs of DOL-systems

Let $G = \langle\Sigma,h,\sigma\rangle$ be a DOL-system.

Definition 1 We call G *irreducible* if for each $a,b \in \Sigma$, $a \neq b$ there is $i \geq 1$ such that $h^i(a) = ubv$ for some $u,v \in \Sigma^*$. More precisely denoting $m(a) = \bigcup_{i=0}^{\infty} \min(h^i(a))$ the condition for irreducibility is

$$b \in m(a) \text{ for all } a,b \in \Sigma, \text{ i.e. } m(a) = \Sigma \text{ for all } a \in \Sigma.$$

Definition 2 A real $n\times n$ matrix M with non-negative elements $m_{ij} \geq 0$ is called *irreducible* if there is no permutation of its rows and the same permutation of its columns which puts M into the form

$$M = \begin{pmatrix} N & 0 \\ P & Q \end{pmatrix},$$

where N, Q are square matrices, and 0 denotes a zero-matrix, possibly rectangular. Equivalently, $M = (m_{ij})$ is irreducible if its indexes cannot be split into two disjoint non-empty sets Σ_1,Σ_2 such that $a \in \Sigma_1$, $b \in \Sigma_2 \Rightarrow m_{ab} = 0$.

Lemma 1 A DOL-system G is irreducible iff its growth matrix is irreducible.

Proof Let $M = (m_{a,b})_{a,b\in\Sigma}$ be the growth matrix of G. Then $m_{ab} = 0$ iff $b \notin \min(h(a))$ by the definition of M. Let $\Sigma_1 \subset \Sigma$, then $m_{a,b} = 0$ for all $a \in \Sigma_1$, $b \notin \Sigma_1$ iff for all $a \in \Sigma_1$ $\min(h(a)) \subset \Sigma_1$, i.e. iff

(1) $\qquad h(\Sigma_1) \subset \Sigma_1.$

Thus M is irreducible if there is $\Sigma_1 \subset \Sigma$ such that (1), and such that Σ_1, $\Sigma-\Sigma_1$ are both non-empty. Conversely if $a \in \Sigma_1$, $b \notin \Sigma_1$, then (1) implies $b \in m(a)$.

If a matrix M is irreducible some of its powers might not be.

<u>Definition 3</u> An $n \times n$ matrix M is called *primitive* if all the matrices M^i, $i = 1,\dots,n$ are irreducible.

If a matrix is primitive then all its powers are irreducible. Note that a matrix is primitive iff some of its powers are positive, i.e. using the terminology of [1] system G is simple iff its growth matrix is primitive. Also note that in fact it is not necessary to examine all powers $M, M^2, \dots, M^n$ but only M^k for k divides n.

Let $G_i = \langle \Sigma, h_i, \sigma \rangle$ be two DOL-systems. We define *balance* β for $w \in \Sigma^*$ as in [1]:

(2) $$\beta(w) = \big| |h_1(w)| - |h_2(w)| \big| .$$

We will use the same symbol β to denote the balance defined on vectors, if $[w]$ is the Parikh vector of w we put $\beta([w]) = \beta(w)$. Let M_1, M_2 be the growth matrices of G_1, G_2. Obviously

(3) $$\beta(x) = |xM_1 - xM_2| \quad \text{for a vector } x,$$

where $|x| = \Sigma |x_i|$ (or simply Σx_i - as all $x_i \geq 0$).

This shows that $\beta(x)$ is well defined.

<u>Lemma 2</u> Suppose that G_1, G_2 are two DOL-systems which are Parikh equivalent, and suppose that the growth matrix of G_1 is primitive. Then for every $\varepsilon > 0$ there is an $n_0 > 0$ such that for every $w \in \Sigma^*$

(4) $$\beta(h_1^n(w)) \leq \varepsilon |h_1^n(w)| \quad \text{for all } n \geq n_0 .$$

Proof Translated into the language of linear algebra (4) becomes

(5) $\beta(xM_1^n) \le \varepsilon|xM_1^n|$ for all vectors $x = (x_1,\dots,x_n)$, such that $x_i \ge 0$.

We will prove (5) for all non-negative vectors, not only those with x_i integers.

From the primitivity of M_1 it follows that M_1 has a positive characteristic value r such that $r > |r_i|$ for all other characteristic values r_i. Moreover, there is a single characteristic vector u belonging to r, and $u > 0$ (i.e. $u = (u_1,\dots,u_n)$, $u_i > 0$ for all $1 \le i \le n$). The same applies to the transposed matrix M_1^T. Denote the single characteristic vector of M_1^T belonging to r by v, thus $v > 0$. Finally we will use (x,y) for the scalar product $\Sigma x_i y_i$. As $u,v > 0$ $(u,v) > 0$. Normalize them such that $(u,v) = 1$. If x is any vector and we write

(6) $x = (x,v)u + w_0$,

then $(w_0,v) = 0$; moreover, if $x \ge 0$ and $x \ne 0$ then $(x,v) \ne 0$. Multiplying (6) by M_1 we get $xM_1 = (x,v)uM_1 + w_0M_1 = (x,v)ru + w_1$, where $w_1 = w_0M_1$ and $(w_1,v) = (w_0M_1,v) = (w_0,vM_1^T) = r(w_0,v) = 0$. By induction

(7) $xM_1^n = ar^nu + w_n$, and $w_n \in W$

where the space $W = \{w:(w,v) = 0\}$ is invariant with respect to M_1, and $a = (x,v) \ne 0$.

Let us denote r_1 the modulus of largest characteristic number of M_1 on W. Because $u \notin W$ we have $r > r_1$. Let r_0 be such that $r > r_0 > r_1$, then

(8) $$|w_n| = |w_0M_1^n| \leq r_0^n|w_0|.$$

From this it follows that $\frac{1}{a}\lim \frac{xM_1^n}{r^n} = u$ for all $x \geq 0$, $x \neq 0$.

If we start in particular with $s = [\sigma]$, the Parikh vector of the initial string we see that

$$u \in \{s, sM_1, sM_1^2, \ldots\} = \{s, sM_2, sM_2^2, \ldots\}$$

where the equality of sets follows from the assumption that G_1, G_2 are Parikh equivalent. But this gives $uM_1 = uM_2$, i.e. $\beta(u) = 0$.

Applying this to (7) and using (8) we get

$$\beta(xM_1^n) = \beta(w_n) \leq \beta r_0^n|w_0|$$

where $\beta = \max_i \beta(\ell_i)$ for the unit vectors $\ell_1, \ell_2, \ldots, \ell_n$. On the other hand,

$$|xM_1^n| = |ar^n u + w_n| \geq |a|r^n|u| - r_0^n|w_0| \geq \frac{|(x,v)|}{2} r^n|u|$$

if n is large enough. Thus

$$\frac{\beta(xM_1^n)}{|xM_1^n|} \leq \left(\frac{r_0}{r}\right)^n \frac{2\beta}{|u|} \frac{|w_0|}{|(x,v)|}.$$

Here $\frac{|w_0|}{|(x,v)|}$ is invariant with respect to the transformation $x \to \alpha x$ and $|(x,v)| \neq 0$, so it is bounded by a constant. This proves (5).

<u>Lemma 3</u> Under the same assumptions as in Lemma 1, for every $\varepsilon > 0$ there is ℓ such that for every G_1-prefix w

(9) $$|w| > \ell \Rightarrow \beta(w) \leq \varepsilon|w|.$$

Proof Using Lemma 2, given $\frac{\varepsilon}{2}$ we find n_0. Let w be any G_1-prefix, i.e. $h_1^n(\sigma) = wx$ for suitable n,x. If

(10) $n \geq n_0$

then denote $u = h_1^{n-n_0}(\sigma)$, and let $u = u_1 a u_2$, $a \in \Sigma$ be such that $h_1^{n_0}(u_1)$ is a prefix of w but w is a proper prefix of $h_1^{n_0}(u_1 a)$, i.e. $w = h_1^{n_0}(u_1)x_1$, $h_1^{n_0}(u_1 a) = wx_2$, $x_1, x_2 \in \Sigma^*$. Now

$$\beta(w) \leq \beta(h_1^{n_0}(u_1))\beta(x_1) \leq \frac{\varepsilon}{2}|h_1^{n_0}(u_1)| + \beta(h_1^{n_0}(a)) \leq \frac{\varepsilon}{2}|w| + \beta q^{n_0}$$

where $q = \max_i \sum_j m_{ij}$ is the length of the longest $h_1(\sigma)$, $\sigma \in \Sigma$. To prcve (9) it is sufficient to take $\frac{q^{n_0}\beta}{|w|} \leq \frac{\varepsilon}{2}$, i.e. to take

$$\ell > q^{n_0}\max(\frac{2}{\varepsilon}\beta, 1)$$

the second case in max-function guarantees (10).

Lemma 4 Let again the assumptions be as in Lemma 1 then the balance is bounded for all G_1-prefixes.

Proof Theorem 3.2 of [1] proves this under the stronger assumption, namely that both G_1, G_2 are simple. However, only the following assumptions were actually used in the proof

a) $h_1^n(x)$ is exponentially growing for each $x \in \Sigma$

b) (9) is true for sufficiently long strings.

Both assumptions are true, so the proof goes through.

The following theorem has also been proved in [1].

Theorem If the pair of DOL-systems G_1, G_2 has bounded balance, then the semialgorithms for equivalence of G_1 and G_2 terminate.

From this it follows that the equivalence problem is decidable for such pairs G_1, G_2 in which at least one of the systems is primitive. This result can be somewhat strengthened as follows.

Definition 4 Let G_1, G_2 be a pair of DOL-systems, $G_i = \langle \Sigma, h_i, \sigma \rangle$. The pair is called *quasi-simple* if there exists

(11) $h = h_{i_n} h_{i_{n-1}} \dots h_{i_2} h_{i_1}$, where $i_j \in \{1,2\}$ for $j = 1,\dots,n$

such that $G = \langle \Sigma, h, \sigma \rangle$ is simple.

Theorem The equivalence problem is decidable on the class of quasi-simple pairs of DOL-systems.

Proof We use the standard method to show that equivalence for the original pair can be established by establishing

(i) the equivalence for n pairs

$$G_1^{(i)} = \langle \Sigma, h, \sigma^{(i)} \rangle \ , \quad G_2^{(i)} = \langle \Sigma, \bar{h}, \sigma^{(i)} \rangle \ ,$$

where $\sigma^{(i)} = h_1^i(\sigma)$, $i = 0,1,\dots,n-1$, and if h was given by (11) then $\bar{h} = h_{3-i_n} h_{i_{n-1}} \dots h_1$;

and (ii) $h_1^j(\sigma) = h_2^j(\sigma)$ for $j = 1,2,\dots,n-1$.

Obviously, if G_1 is equivalent to G_2 then all $G_1^{(i)}$ are equivalent to the corresponding $G_2^{(i)}$. To show the converse assume that G_1 is not equivalent to G_2, so that there is j such that $(h_1^j(\sigma) \neq h_1^j(\sigma))$ and assume that our j is the smallest possible. In order to simplify the notation assume $i_n = 1$. Every ℓ can be written as

$$\ell = n\, x_\ell + r_\ell \quad \text{with } 0 \le r < n$$

and we have $h^{x_\ell}(\sigma^{(r_\ell)}) = h_1^\ell(\sigma)$ for all $\ell \le j$, similarly

$$\bar{h}^{x_\ell}(\sigma^{(r_\ell)}) = h_2^\ell(\sigma) \text{ for } \ell \le j.$$

This shows that $G_1^{(r_j)}$ is not equivalent to $G_2^{(r_j)}$.

At the same time the systems $G_1^{(i)}$ are irreducible so the problem for them is decidable.

Note that it is decidable whether a given pair of DOL-systems is quasi-simple. This is so because in the first place this depends only on the growths matrices of both systems and secondly not on actual elements of the matrices but only on whether the element is or is not zero. Thus we may start with given growth matrices M_1, M_2. On each step take all the matrices from previous steps multiplied (on the right) by M_1 or M_2 considering two matrices different only if one has zero element where the second has a non-zero one (thus we consider in fact the matrices over a boolean semi-ring $\{0,1\}$). As there is only 2^{n^2} of such matrices we must either reach a matrix with all elements positive or reach a step when all matrices taken on this step were already considered. In the former case the pair is quasi-simple, in the latter it is not.

2. The case of two letter alphabet

Theorem Equivalence problem is decidable for DOL-systems with $\#\Sigma = 2$.

Proof It is sufficient to consider the case when neither of the two systems is primitive. For 2 symbols the system G is reducible if - after possible permutation - its growth matrix is of the form

$$M_1 = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix},$$

we may assume $c \neq 0$, otherwise $h(w)$ is a power of one letter only, and this case need not be considered (see [2]).

Let us consider separately the following three cases:

(i) $a = 0$. Let us write $\Sigma = \{\sigma_1,\sigma_2\}$, $h_1(\sigma_1) = \varepsilon$, $h_1(\sigma_2) = v$, and let (p,q) be the Parikh vector of the axiom w. The sequence generated by G_1 is obviously

$L = (w, v^q, v^{qc}, v^{qc^2}, \ldots)$.

Let $G_2 = \langle\Sigma, g_2, w\rangle$ be a system generating the same sequence, i.e. in particular $h_2(v^q) = v^{cq}$, but this implies $h_2(v) = v^c$, and from this it follows that $h_2(v^{cq^i}) = v^{cq^{i+1}}$. Thus to test whether G_1 eq G_2 it is enough to test whether $h_1(w) = h_2(w)$ and $h_1(v) = h_2(v)$.

(ii) $b = 0$. Let us write $w = \sigma_1^{i_0}\sigma_2^{j_1}\sigma_1^{i_1}\ldots\sigma_2^{j_k}\sigma_1^{i_k}$, $k \geq 0$, where $i_0, i_k \geq 0$, $i_s, j_s > 0$, for the rest. In this case,

$L = (\ldots, \sigma_1^{i_0 a^\ell}\sigma_2^{j_1 c^\ell}\ldots\sigma_1^{i_k a^\ell}, \ldots | \ell = 0,1,\ldots)$

in particular, the number of "runs" of each symbol is constant, and both a and c are determined uniquely. Thus such a system is equivalent only with itself if both $a,c > 1$. The case $a = 1$ or $c = 1$ can give an additional equivalence of $\begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix}$ with $\begin{pmatrix} 0 & 0 \\ 1 & b-1 \end{pmatrix}$ which can easily be tested.

(iii) So we may now assume $a \neq 0$, $b \neq 0$, $c \neq 0$. There are two possibilities for G_2 which is also reducible.

a) $M_2 = \begin{pmatrix} d & e \\ 0 & f \end{pmatrix}$. As before, we may assume that all elements d,e,f are $\neq 0$.

In this case $M_1 M_2$ is irreducible.

b) $M_2 = \begin{pmatrix} d & 0 \\ e & f \end{pmatrix}$. From the Parikh equivalence we get $f = c$.

We will show that

(1) $a > 1 \Rightarrow a = d$.

From this, by symmetry we get $d > 1 \Rightarrow a = d$, and as the cases $a = 0$ or $d = 0$ were already covered under (i), we may ignore them. On the other hand, $a = d$ implies $M_1 = M_2$. Obviously, if the homomorphisms are same on one letter, they are same on the other one.

In order to prove (1) let us write $\Sigma = \{\sigma_1, \sigma_2\}$,

$h_1(\sigma_1) = \sigma_1^a$

$h_2(\sigma_2) = \sigma_1^g \psi \sigma_1^h$; $g,h \geq 0$ and ψ both starts and ends with σ_2.

[Note that we assume $\sigma_2 \in h_2(\sigma_2)$ as $c \neq 0$.]

We will investigate substrings of $w_k = h_1^k(w)$ consisting of σ_1's only - let us call such a substring *maximal* if it cannot be lengthened and still contains σ_1's only. In more details if $w_k = xyz$, $y \in \{\sigma_2\}^+$ and either

x is empty - i.e. y is a prefix or x ends with σ_2 and at the same time z is empty - i.e. y is a suffix or z starts with σ_2.

Now if w_k contains a maximal string of the length ℓ, then $w_{k+1}, w_{k+2}, w_{k+3}, \ldots$ will contain a maximal string of the length $a\ell+i$, $a^2\ell+i(a+1)$, $a^3\ell+i(a^2+a+1), \ldots$ where $i = g$ if our string was a prefix, $i = h$ for suffix, or $i = g+h$ if the string was neither suffix nor prefix.

In order to see that such a sequence of string lengths determines a uniquely it is sufficient to consider the longest maximal string in w_k, and take k large enough so that the length is larger than $h_2(\sigma_2)$. This is possible if $a > 1$.

Finally, there is a case of M irreducible, but M_1^2 reducible, in this case $M_1 = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$ and this case can be reduced to the case (ii) by considering M_1^2 instead of M_1. See [2] for more details.

References

[1] K. Culik II, On the decidability of the sequence equivalence problem for DOL-systems, Theoretical Computer Science to appear, also Res. Rep. CS-75-24, Dept. of Computer Science, University of Waterloo, September 75.

[2] M. Nielsen, On the decidability of some equivalence problems for DOL-systems, Inf. and Control 25, 166-193 (1974).

[3] F.R. Gantmacher, The Theory of Matrices, vol.2, Chelsea, New York.