# ELEMENTARY HOMOMORPHISMS AND A SOLUTION OF THE DOL SEQUENCE EQUIVALENCE PROBLEM

by

A. Ehrenfeucht\*

G. Rozenberg\*\*

#CU-CS-115-77

September, 1977

<sup>\*</sup>A. Ehrenfeucht, Department of Computer Science, University of Colorado at Boulder, Boulder, Colorado 80309 USA

<sup>\*\*</sup>G. Rozenberg, Department of Mathematics, University of Antwerp, U.I.A., Wilrijk, Belgium

## ABSTRACT

This paper continues the research (started in [3]) on elementary DOL systems. In particular we provide an alternative (and simpler than the one presented in [1]) proof that the DOL sequence equivalence problem is decidable.

#### I. INTRODUCTION

The notion of a homomorphism on a free monoid is a basic notion in the formal language theory, and it is certainly a central notion in the theory of L systems (see, e.g., [4]). It must be clear to anyone following the development of the theory of L systems that our knowledge about the basic properties of homomorphisms is rather poor. This may be due to the fact that we have only few basic techniques to deal with them (among these techniques, e.g., the growth function arguments as in [5] and subword complexity arguments as in [2]).

In [3] we have proposed a technique which we believe is a basic technique to deal with homomorphisms (several examples of its usefulness were demonstrated in [3]). This paper elaborates on this technique as well as extends it. As an example of its application we provide a proof of decidability of the DOL sequence equivalence problem. The decidability of this problem was demonstrated in [1]. Our new proof is entirely different from the one in [1]. We believe that this proof is simpler, more elementary and sheds an additional light on the nature of this problem.

We assume the reader to be familiar with rudiments of formal language theory and in particular with the rudiments of the theory of L systems.

Since the problems considered in this paper are trivial otherwise, we consider here only DOL systems generating infinite languages.

#### II. PRELIMINARIES

Mostly we will use the standard formal language theoretic notation and terminology. Perhaps the following requires an explanation.

- 1) For a finite set Z, #Z denotes its cardinality. N and N<sup>+</sup> denote the set of nonnegative and positive integers, respectively. Given finite alphabets  $\Sigma$  and  $\Delta$ ,  $\text{HOM}(\Sigma,\Delta)$  denotes the set of all homomorphisms from  $\Sigma^*$  into  $\Delta^*$ . If H  $\subseteq$  HOM( $\Sigma,\Sigma$ ) then SemH denoted the semigroup of homomorphisms generated by H. The composition of homomorphisms h<sub>1</sub>,...,h<sub>k</sub> is written as h<sub>k</sub>...h<sub>1</sub>. For a homomorphism h in  $\text{HOM}(\Sigma,\Delta)$  and K  $\subseteq$   $\Sigma^*$ ,
- $\underline{im}_h K = \{\alpha \in \Delta^* : (\exists \beta)_K [h(\beta) = \alpha]\}.$  Also  $\underline{maxrh} = max\{|\alpha| : \alpha \in \underline{im}_h \Sigma\}$
- 2) For a word  $\alpha$ ,  $|\alpha|$  denotes its length and  $\alpha \setminus \beta$  denotes the word resulting from  $\alpha$  by cutting off its suffix  $\beta$ . For a sequence of words  $\tau = \omega_0, \omega_1, \ldots \text{ and a nonnegative integer i we use } \underline{i}(\tau) \text{ to denote the i'th element of } \tau \text{ (hence } \underline{i}(\tau) = \omega_i \text{)}.$
- 3) For an automaton we use the notation  $A = (\Sigma, Q, \delta, q_{in}, F)$  and A is a finite automaton if its set of states Q is finite. In specifying automata we always assume that all the transitions not specified by its definition lead to a special "dead state" D (all transitions from D lead to itself and  $D \notin F$ ). The language of A is denoted by T(A). SUCC(Q) denotes the set of all states from which one can get to a final state. For a Q in Q and Q in Q in Q in Q at a started at Q is the sequence of states encountered when following Q in A starting with Q. Also PRED(Q) denotes the set of all states in Q from which Q can be reached in one transition step.
- 4) For a DOL system G =  $(\Sigma,h,\omega)$ , E(G) denotes its sequence  $\omega_0,\omega_1,\ldots$

The following terminology concerning homomorphisms will be used in the sequel.

Definition 1. Let  $h, g \in HOM(\Sigma, \Delta)$ .

- 1) Let K be a language over  $\Sigma$ . We say that h,g are K-<u>equal</u>, denoted as h = K g, if for every word  $\alpha$  in K  $h(\alpha) = g(\alpha)$ . In such a case we say that K is an <u>identifying set</u> for h,g.
- 2) The <u>maximal identifying set for</u> h,g, denoted as MID(h,g), is defined by MID(h,g) = {  $\alpha \in \Sigma^*$  : h( $\alpha$ ) = g( $\alpha$ ) }.
- 3) Let  $\tau = \alpha_0, \alpha_1, \ldots$  be a sequence of words over  $\Sigma$ . We say that h,g are  $\tau$ -equal, denoted h =  $_{\tau}$  g, if for every  $i \ge 0$  h( $\alpha_i$ ) = g( $\alpha_i$ ).

<u>Definition 2</u>. Let h,g  $\in$  HOM( $\Sigma$ , $\Delta$ ) and let  $\alpha$  in  $\Sigma^*$  be such that h( $\alpha$ ) is a prefix of g( $\alpha$ ) or g( $\alpha$ ) is a prefix of h( $\alpha$ ) (in particular it can be that h( $\alpha$ ) = g( $\alpha$ )). Then the <u>delay of</u> h,g <u>on</u>  $\alpha$ , denoted <u>del</u>h,g( $\alpha$ ), is the word over  $\Delta$  such that

either  $h(\alpha) del_{h,q}(\alpha) = g(\alpha) \underline{or} g(\alpha) \underline{del}_{h,q}(\alpha) = h(\alpha)$ .

The following result follows directly from the above definition.

<u>Lemma 1</u>. Let h,g be homomorphisms on  $\Sigma^*$  and  $\alpha, \bar{\alpha}, \beta \in \Sigma^*$ .

If  $\underline{\text{del}}_{h,g}(\alpha) = \underline{\text{del}}_{h,g}(\bar{\alpha})$  then  $\underline{\text{del}}_{h,g}(\alpha\beta) = \underline{\text{del}}_{h,g}(\bar{\alpha}\beta)$ .

#### III. ELEMENTARY LANGUAGES

In this section we generalize the idea of an elementary homomorphism from [3] by introducing a notion of an elementary language. Then we prove a basic combinatorial property of elementary languages.

Let us start by recalling from [3] the notion of an elementary homomorphism.

<u>Definition 3.</u> Let  $h \in HOM(\Sigma,\Sigma)$ . We say that h is <u>simplifiable</u> if there exists an alphabet  $\Delta$  with  $\#\Delta < \#\Sigma$  and homomorphisms  $f \in HOM(\Sigma,\Delta)$ ,  $g \in HOM(\Delta,\Sigma)$  such that h = gf; otherwise h is called <u>elementary</u>.

This notion generalizes to DOL systems as follows.

<u>Definition 4</u>. Let  $G_1 = (\Sigma, h, \omega)$  and  $G_2 = (\overline{\Sigma}, \overline{h}, \overline{\omega})$  be two DOL systems. We say that  $G_2$  is a simplification of  $G_1$  if

- (i)  $\#\bar{\Sigma} < \#\Sigma$ , and
- (ii) there exist homomorphisms  $f \in HOM(\Sigma, \overline{\Sigma})$ ,  $g \in HOM(\overline{\Sigma}, \Sigma)$  such that  $f(\omega) = \overline{\omega}$ , h = gf and  $\overline{h} = fg$ .

If a DOL system does not have a simplification then we call it elementary.

The following easy result from [3] underlies a lot of our considerations about decidability of various problems considered in this paper.

#### Lemma 2.

- (i) It is decidable whether or not an arbitrary homomorphism is elementary.
- (ii) It is decidable whether or not an arbitrary DOL system is elementary.

The following result was proved in [3].

Theorem 1. If h is a noninjective homomorphism then h is simplifiable.

The idea of a simplifiable (elementary) homomorphism can be generalized to sets as follows.

<u>Definition 5</u>. Let  $\Sigma$  be a finite alphabet and let  $U \subseteq \Sigma^*$  be a finite language. We say that U is <u>simplifiable</u> if there exists  $Z \subseteq \Sigma^*$  such that #Z < #U and  $U \subseteq Z^*$ . Otherwise U is called <u>elementary</u>.

Note that for a homomorphism  $h \in HOM(\Sigma,\Sigma)$ ,  $\underline{im}_h\Sigma$  is simplifiable or elementary if h is simplifiable or elementary respectively. This is a reason that elementary sets are of interest to us, and we move now to investigate them.

As a direct consequence of Theorem 1 we get the following result.

Lemma 3. Let  $U = \{u_1, \dots, u_k\}$  be simple. If  $i_1, \dots, i_n, j_1, \dots, j_m$  are elements of  $\{1, \dots, k\}$  such that

$$u_{i_1} \dots u_{i_n} = u_{j_1} \dots u_{j_m}$$
 then  $m = n$  and  $i_t = j_t$  for  $1 \le t \le n$ .

The next result turns out to be an important technical result for this paper.

Theorem 2. Let  $U = \{u_1, \ldots, u_k\}$  be elementary. If  $i_1, \ldots, i_n$  is the sequence of indices from  $\{1, \ldots, k\}$  such that there exist a word  $\gamma$  and a sequence of indices  $j_1, \ldots, j_m$  from  $\{1, \ldots, k\}$  where  $i_1 \neq j_1$  and

$$u_{i_1} \dots u_{i_n}^{\gamma} = u_{j_1} \dots u_{j_m}^{\gamma}$$
, then

$$|u_{i_1}...u_{i_n}| \leq |u_1...u_k| - k.$$

#### Proof.

Let us construct the finite sequence of sequences of words  $\tau_1, \tau_2, \dots$  (each of which contains (k + 2) words) inductively as follows.

(i) Let 
$$\tau_1 = u_1, 1, \dots, u_k, 1, u_{i_1}, 1, \dots u_{i_{n_1}}, 1^{\gamma}, u_{j_1}, 1, \dots u_{j_{m_1}}, 1$$
where  $u_1, 1 = u_1, \dots, u_k, 1 = u_k, u_{i_1}, 1 = u_{i_1}, \dots, u_{i_{n_1}}, 1 = u_{i_n}, \dots, u_{i_{n_1}}, \dots, u_{i_{n_$ 

(ii) Now, for  $\ell \ge 1$ , given a sequence

$$\begin{array}{l} {}^{\tau_{\ell}} = u_{1,\ell}, \ldots, u_{k,\ell}, u_{i_{1},\ell}, \ldots u_{i_{n},\ell}^{\gamma,u} j_{1}, \ell \ldots u_{j_{m},\ell}, \\ \\ \text{where } (1,\ell), \ldots, (k,\ell), (i_{1},\ell), \ldots, (i_{n_{\ell}},\ell), (j_{1},\ell), \ldots, (j_{m_{\ell}},\ell) \in \{1,\ldots,k\}, \\ \\ \text{satisfying conditions} \end{array}$$

- 1)  $(\underline{k+1})(\tau_0) = (\underline{k+2})(\tau_0)$ ,
- 2)  $(i_1, \ell) \neq (j_1, \ell)$ , and
- 3) {  $\underline{1}(\tau_{\ell}),\underline{2}(\tau_{\ell}),\ldots,\underline{k}(\tau_{\ell})$  } is an elementary set we define

From the conditions 2 and 3 above it follows that  $w_{i_1,\ell} \neq w_{j_1,\ell}$  and so from the condition 1 above it follows that either  $w_{i_1,\ell}$  is a proper prefix of  $w_{j_1,\ell}$  or  $w_{j_1,\ell}$  is a (proper) prefix of  $w_{i_1,\ell}$ .

(ii.1) Assume that  $\xi_{\ell}$  is a nonempty word such that  $u_{i_1,\ell}\xi_{\ell} = u_{j_1,\ell}$ . Then

for 
$$1 \le t \le k$$
,  $\underline{t}(\tau_{\ell+1}) = \begin{cases} \underline{t}(\tau_{\ell}) & \text{for } t \ne j_1, \\ \zeta_{\ell} & \text{for } t = j_1, \end{cases}$ 

$$(\underline{k+1})(\tau_{\ell+1}) = u_{i_2}, \ell \cdots u_{i_{n_{\ell}}}, \ell^{\gamma}, \text{ and}$$

$$(\underline{k+2})(\tau_{\ell+1}) = \tau_{i_2}, \ell \cdots u_{i_{n_{\ell}}}, \ell^{\gamma}, \text{ and}$$

$$(\underline{k+2})(\tau_{\ell+1}) = \zeta_{\ell} u_{j_2,\ell} \dots u_{j_{m_{\ell}},\ell}$$

(ii.2) Assume that 
$$z_{\ell}$$
 is a nonempty word such that  $u_{j_1,\ell}z_{\ell}=u_{j_1,\ell}$ . Then

$$\text{for 1 } \leq \ t \leq \ k \,, \ \underline{t}(\tau_{\ell+1}) \ = \begin{cases} \underline{t}(\tau_{\ell}) \ \text{for } t \neq i_1 \,, \\ \\ \zeta_{\ell} \ \text{for } t = i_1 \,, \end{cases}$$

$$(\underline{k+1})(\tau_{\ell+1}) = \zeta_{\ell} u_{i_{2},\ell} \dots u_{i_{n_{\varrho}},\ell} \gamma$$
, and

$$(\underline{k+2})(\tau_{\ell+1}) = u_{\underline{j}_2,\ell} \dots u_{\underline{j}_{m_0},\ell}$$

- (iii)  $\tau_{\ell+1}$  satisfies the conditions 1, 2 and 3 from above. We prove it as follows (since cases (ii.1) and (ii.2) are symmetric we assume that (ii.1) holds).
- (iii.1) Since  $(\underline{k+1})(\tau_{\underline{\ell}}) = (\underline{k+2})(\tau_{\underline{\ell}})$  and both  $(\underline{k+1})(\tau_{\underline{\ell+1}}), (\underline{k+2})(\tau_{\underline{\ell+1}})$  result by erasing the prefix  $w_{\underline{1},\underline{\ell}}$  in  $(\underline{k+1})(\tau_{\underline{\ell}})$  and  $(\underline{k+2})(\tau_{\underline{\ell}})$  respectively, we get  $(\underline{k+1})(\tau_{\underline{\ell+1}}) = (\underline{k+2})(\tau_{\underline{\ell+1}})$ .
- (iii.2) By the construction (and Lemma 3) we have  $(j_1, \ell+1) = (j_1, \ell)$  and  $(i_1, \ell+1) = (i_2, \ell)$ . There are two possibilities to consider.
- (iii.2.1)  $(i_2, \ell) \neq (j_1, \ell)$ . Then  $(i_1, \ell+1) \neq (j_1, \ell+1)$ .
- (iii.2.2)  $(i_2,\ell) = (j_1,\ell)$ . Then  $u_{i_1,\ell+1} = u_{i_2,\ell} = u_{j_1,\ell} = u_{i_1,\ell} \neq 0$   $\zeta_{\ell} = w_{j_1,\ell+1} \text{ and so } (i_1,\ell+1) \neq (j_1,\ell+1).$
- (iii.3) {  $\underline{1}(\tau_{\ell+1}), \ldots, \underline{k}(\tau_{\ell+1})$  } must be elementary since otherwise {  $\underline{1}(\tau_{\ell}), \ldots, \underline{k}(\tau_{\ell})$  } would not be elementary.
- (iii.4) Since obviously  $\tau_{\text{1}}$  satisfies condition 1, 2, and 3, (iii) holds.
- (iv) We terminate the construction on  $\tau_{\ell+1}$  such that  $((\underline{k+1})\tau_{\ell+1}) \setminus \gamma = \Lambda$ . Obviously the construction always terminate.
- (v) It follows directly from the construction that, for every  $\ell \geq 1$ ,

$$|((\underline{k+1})(\tau_{\ell})) \setminus \gamma| - |u_{1,\ell} \dots u_{k,\ell}| = |((\underline{k+1})(\tau_{1})) \setminus \gamma| - |u_{1,1} \dots u_{k,1}| = |u_{i_{1}} \dots u_{i_{n}}| - |u_{1} \dots u_{k}|.$$

(vi) Now we prove the theorem by contradiction as follows.

Assume that  $|u_{i_1}...u_{i_n}| > |u_1...u_k| - k$ , thus  $|u_{i_1}...u_{i_n}| - |u_1...u_k| > -k$ .

Hence if  $\tau_{\ell}$  is such that  $((\underline{k+1})(\tau_{\ell})) \setminus \gamma = \Lambda$ , then (v) implies that  $|u_{1,\ell}...u_{k,\ell}| < k$ ; a contradiction.

#### IV. HOMOMORPHIC INVERSES OF A LANGUAGE

In this section we briefly investigate the inverse image of a language by a (not necessarily finite) set of homomorphisms. Formally it is defined as follows.

<u>Definition 6</u>. Let K be a language,  $K \subseteq \Sigma^*$ , and let H be a set of homomorphisms from  $\Sigma^*$  into  $\Sigma^*$ . The H-<u>inverse</u> of K, denoted as I(K,H), is defined by  $I(K,H) = \{ \alpha \in \Sigma^* : (\forall h)_H [h(\alpha) \in K] \}$ .

Theorem 3. Let  $K \subseteq \Sigma^*$  and let H be a set of homomorphisms from  $\Sigma^*$  into  $\Sigma^*$ . If K is regular then I(K,H) is regular.

#### Proof.

Let  $A = (\Sigma, Q, \delta, q_{in}, F)$  be a finite automaton such that T(A) = K.

For each h in H let  $A_h = (\Sigma, Q, \delta_h, q_{in}, F)$  be a finite automaton such that  $(\forall q)_0(\forall a)_{\wedge}[\delta_h(a) = \delta(h(a))].$ 

Clearly:

- 1.  $(\forall \alpha)_{\Sigma^*}[h(\alpha) \in T(A) \text{ if and only if } \alpha \in T(A_h)]$  and
- 2. {  $A_h : h \in H$  } is a finite set.

But regular languages are closed under finite intersection, and so  $I(K,H) = \bigcap_{h \in H} T(A_h) \text{ is a regular set.}$ 

Corollary 1. It is decidable whether or not  $L(G) \subseteq T(A)$  for an arbitrary DOL system G and an arbitrary finite automaton A.

#### Proof.

Let  $G = (\Sigma, h, \omega)$  and let  $H = Sem\{h\}$ .

Note that one can effectively construct an automaton  $A_H$  such that  $T(A_H) = I(L(G),H)$ . This follows from the proof of Theorem 3 and from the observation that, for each  $n \geq 2$ ,  $A_h$  is obtained from  $A_{h}^{n-1}$  in the same way as  $A_h$  is obtained from  $A_h$ .

Hence to decide whether L(G)  $\subseteq$  T(A) it suffices to check whether  $\omega \in I(L(G),H)$  which can be effectively done because the membership question for finite automata is decidable.

## V. THE (h,g)-AUTOMATON

The maximal identifying sets for a given set of homomorphisms turn out to play a major role in investigating DOL systems. In this section we present a construction which given a pair of homomorphisms h,g provides a (not necessarily finite) automaton, called the (h,g)-automaton, which accepts precisely MID(h,g). Then we prove that if h,g are elementary then the (h,g)-automaton is finite (and so MID(h,g) is regular).

Definition 7. Let h,g  $\in$  HOM( $\Sigma$ , $\Delta$ ). Then the (h,g)-automaton, denoted as  $A_{h,g}$ , is a (not necessarily finite) deterministic automaton which equals the initial connected component of the automaton ( $\Sigma$ ,Q, $\delta$ ,q<sub>in</sub>,F) where

$$Q = Q_1 \cup Q_{2h} \cup Q_{2g}$$
 with  $Q_1 = \{ [\Lambda], D_{g} \}$ ,

$$Q_{2h} = \{ [h, \alpha] : \alpha \in \Delta^+ \text{ and } ((\exists \bar{\alpha})_{\Delta^*} [\alpha \bar{\alpha} \in (\underline{im}_h \Sigma)^+]) \},$$

$$Q_{2g} = \{ [g,\alpha] : \alpha \in \Delta^{+} \text{ and } ((\exists \bar{\alpha})_{\Delta^{*}} [\alpha \bar{\alpha} \in (\underline{im}_{h}\Sigma)^{+}]) \},$$

$$q_{in} = [\Lambda],$$

$$F = \{ [\Lambda] \},$$

and  $\delta$  is defined as follows:

for every a  $\in \Sigma$ ,

(i) 
$$\delta(D,a) = D$$
,

(ii) 
$$\delta([\Lambda],a) =$$

$$\begin{cases} \frac{\text{if }}{\text{if }} \text{ } h(a) = \text{g(a)} \ \underline{\text{then}} \ [\Lambda], \\ \frac{\text{if }}{\text{if }} \left( (\exists \alpha)_{\Delta^+} (\exists \bar{\alpha})_{\Delta^*} [ \ h(a)_{\alpha} = \text{g(a)} \ \text{and} \ \alpha \bar{\alpha} \in \left( \underline{\text{im}}_h \Sigma \right)^+ ] \right) \ \underline{\text{then}} \ [h, \alpha], \\ \text{if } \left( (\exists \alpha)_{\Delta^+} (\exists \bar{\alpha})_{\Delta^*} [ \ \text{g(a)}_{\alpha} = \text{h(a)} \ \text{and} \ \alpha \bar{\alpha} \in \left( \underline{\text{im}}_g \Sigma \right)^+ ] \right) \ \underline{\text{then}} \ [g, \alpha], \\ \text{else D.} \end{cases}$$

(iii) 
$$\delta([h,\alpha],a) = \frac{\inf \alpha g(a) = h(a) \underline{then} [\Lambda],}{\inf ((\exists \beta)_{\Delta} + (\exists \overline{\beta})_{\Delta} * [\alpha g(a) \beta = h(a) and \beta \overline{\beta} \in (\underline{im}_{g} \Sigma)^{+}]) \underline{then} [g,\beta],}{\inf ((\exists \beta)_{\Delta} + (\exists \overline{\beta})_{\Delta} * [\alpha g(a) = h(a) \beta and \beta \overline{\beta} \in (\underline{im}_{h} \Sigma)^{+}]) \underline{then} [h,\beta],}{\underline{else} D.}$$
(iv)  $\delta([g,\alpha],a) = \frac{\inf \alpha h(a) = g(a) \underline{then} [\Lambda],}{\inf ((\exists \beta)_{\Delta} + (\exists \overline{\beta})_{\Delta} * [\alpha h(a) \beta = g(a) and \beta \overline{\beta} \in (\underline{im}_{h} \Sigma)^{+}]) \underline{then} [h,\beta],}{\underline{if} ((\exists \beta)_{\Delta} + (\exists \overline{\beta})_{\Delta} * [\alpha h(a) \beta = g(a) \beta and \beta \overline{\beta} \in (\underline{im}_{g} \Sigma)^{+}]) \underline{then} [g,\beta],}{\underline{else} D.}$ 

We will use the following terminology concerning  $A_{h,q}$ .

- 1) If  $g \in Q_2$  then  $tailq = \alpha$  where either  $q = [h, \alpha]$  or  $q = [g, \alpha]$ .
- 2) A state  $q \in Q$  is called <u>deterministic</u> if there is at most one a in  $\Sigma$  such that  $\delta(q,a) \neq D$ . We use DET(Q) to denote the set of all deterministic states in Q.

The following two technical results explain the usefulness of the automaton  $\mathbf{A}_{\text{h.g.}}$ 

Lemma 4. 
$$T(A_{h,g}) = MID(h,g)$$
.

Proof.

It follows from an easy observation that, for every  $\alpha \in \Sigma^+$ ,

- 1)  $\underline{\text{del}}_{h,g}(\alpha) \neq \Lambda \text{ and } h(\alpha)\underline{\text{del}}_{h,g}(\alpha) = g(\alpha) \text{ if and only if}$
- $\delta( [\Lambda], \alpha) = [h, \alpha],$
- 2)  $\underline{\text{del}}_{h,g}(\alpha) \neq \Lambda \text{ and } g(\alpha)\underline{\text{del}}_{h,g}(\alpha) = h(\alpha) \text{ if and only if}$
- $\delta( [\Lambda], \alpha) = [g, \alpha],$
- 3)  $\underline{\text{del}}_{h,g}(\alpha) = \Lambda \text{ if and only if } \delta( [\Lambda],x ) = [\Lambda].$

This observation implies that, for every  $\alpha \in \Sigma^+$ ,

 $\alpha \in T(\bar{A}_{h,g})$  if and only if  $\alpha \in MID(h,g)$ .

 $\underline{\text{Lemma 5}}$ . If h and q are elementary homomorphisms then SUCC(Q) is a finite set.

Proof.

(i)  $(\exists e)_{\mathbb{N}^+}(\forall q)_{\mathbb{Q}_2} [\underline{if} | \underline{tailq}| > e \underline{then} q \in DET(\mathbb{Q}) ].$ 

Proof of (i).

We prove it by contradiction.

(i.1) Let  $e_h = (2 \max_h (h) + |h(a_1, a_2...a_n)| - n)$ , where  $\Sigma = \{a_1, ..., a_n\}$ , and assume  $q = [h, \alpha] \notin DET(Q)$  but  $|\alpha| > e_h$ .

Thus

$$(\exists a_i, a_j)_{\Sigma} [a_i \neq a_j, \delta(q, a_i) \neq D \text{ and } \delta(q, a_j) \neq D] \dots$$
 (\*).

Since  $|\alpha| > e_h > \underline{\text{maxrh}}$  only the third and the fourth case of the conditional definition of  $\delta([h,\alpha],a)$  apply to  $\delta(q,a_i)$  and  $\delta(q,a_j)$ . Hence from (\*) it follows that

$$(\exists \bar{\beta}_{i})_{\Delta^{*}} (\exists t_{1})_{N^{+}} (\exists i_{1}, \dots, i_{t_{1}})_{\{1, \dots, n\}} [\alpha g(a_{i}) \bar{\beta}_{i} = h(a_{i}) h(a_{i_{1}}) \dots h(a_{i_{1}})] \dots$$
 and 
$$(**)$$

$$(\exists \bar{\beta}_{j})_{\Delta^{*}}(\exists t_{2})_{N^{+}}(\exists j_{1},...,j_{t_{2}})_{\{1,...,n\}}[\alpha g(a_{j})\bar{\beta}_{i} = h(a_{j})h(a_{j_{1}})...h(a_{j_{t_{2}}})]...$$

(\*\*\*)

Let p be the maximal positive integer such that

$$(\exists \gamma)_{\Delta *}[h(a_i)h(a_{i_1})...h(a_{i_p})\gamma = \alpha].$$

(Note that since  $|\alpha| > 2$ maxrh such a p exists.)

Then (\*\*) implies that

$$|h(a_i)h(a_{i_1})...h(a_{i_p})| \ge |\alpha| - \underline{maxr}(h) > |h(a_1...a_n)| - n.$$
 (\*\*\*\*)

But (\*\*\*) implies that

$$(\exists \overline{\gamma})_{\Delta^{+}}[h(a_{i})h(a_{i})...h(a_{i}_{p})\overline{\gamma} = h(a_{j})h(a_{j})...h(a_{j}_{t_{2}})].$$

Since  $i \neq j$  and  $\underline{im}_{\Sigma}h$  is elementary (because h is elementary), this together with (\*\*\*\*) contradicts Theorem 2.

- (i.2) Similarly if  $e_g = (2maxr(g) + |h(a_1...a_n)| n)$  and we assume that  $q = [g,\alpha] \notin DET(Q)$  but  $|\alpha| > e_q$ , then we get a contradiction.
- (i.3) Thus (i) holds if we set  $e = \max\{e_h, e_q\}$ .
- (ii)  $(\exists r)_{N+}(\forall q)_{Q_2}[\underline{if} | \underline{tailq}| > r \underline{then} PRED(q) \subseteq DET(Q)]$ Proof of (ii).

This follows from (i) and from the obvious fact that

- $(\exists s)_{N^{+}}(\forall q, \bar{q})_{Q_{2}}[\underline{if} \delta(q, a) = \bar{q} \text{ for some a in } \Sigma$   $\underline{then} \parallel |\underline{tailq}| |\underline{tail\bar{q}}| \parallel < s ].$
- (iii) Let  $r_0$  be the smallest positive integer satisfying (ii) and larger than e from (i). Then a state q from  $Q_2$  is called <u>long</u> if  $|tailq| > r_0$  and it is called <u>short</u> otherwise; LONG(Q) and SHORT(Q) denote the corresponding subsets of  $Q_2$ .
- (iv) Let  $\mathbf{q} \in (\mathsf{SHORT}(\mathsf{Q}))$  on (DET(Q)). Let  $\tau = \mathsf{q}, \mathsf{q}_1, \mathsf{q}_2, \ldots, \mathsf{q}_{in}$  and  $\bar{\tau} = \mathsf{q}, \bar{\mathsf{q}}_1, \bar{\mathsf{q}}_2, \ldots, \mathsf{q}_{in}$  be two traces starting at q where  $\mathsf{q} = \bar{\mathsf{q}}_1$  is long. Let  $\mathsf{i}_0$  be the smallest index such that  $\mathsf{q}_i$  is short. Then for every  $\mathsf{i} \leq \mathsf{i}_0$   $\mathsf{q}_i = \bar{\mathsf{q}}_i$ .

Proof of (iv).

This follows directly from (ii).

(v) According to (iv) if  $q \in (SHORT(Q)) \cap (DET(Q)) \cap SUCC(Q)$  and one considers a trace starting with q, leading through long states and ending with a short state, then this trace is unique. Let TRANS(q) denote the set of states appearing in this unique trace. (Thus in the notation of (iv) we have  $TRANS(q) = \{q,q_1,\ldots,q_{j_0}\}$ .) Note that TRANS(q) is a finite set.

# Proof of (vi).

This follows directly from (iv).

(vii) Since both components of the right hand side of the containment relation in (vi) are finite, the lemma follows.

Theorem 4. If h,g are elementary homomorphisms then MID(h,g) is regular.

#### Proof.

Let  $A_{h,g} = (\Sigma,Q,\delta,q_{in},F)$  and let  $\psi$  be a function from Q into Q defined by

$$\psi(q) = \begin{cases} [\Lambda] & \underline{if} \ q = [\Lambda], \\ q & \underline{if} \ q \in SUCC(Q), \\ D & \text{otherwise.} \end{cases}$$

Let  $B_{h,g} = (\Sigma, \psi(Q), \delta_{\psi}, [\Lambda], \{[\Lambda]\})$  where  $\delta_{\psi}$  is defined as follows:  $(\forall q, \bar{q})_Q(\forall a)_{\Sigma}[\delta_{\psi}(q, a) = \bar{q} \text{ if and only if } \delta(\psi(q), a) = \bar{q}].$  Obviously  $T(B_{h,g}) = T(A_{h,g})$  and so the theorem follows from Lemma 4 and Lemma 5.

Given two homomorphisms h,g we can effectively construct a sequence of finite automata approximating  $A_{h,g}$ , hence also approximating MID(h,g). This is done as follows.

Choose a positive integer k.

Let  $A_{h,g}^{(1)}$  be the automaton constructed in the same was as  $A_{h,g}$  except that we take as states [A],D and all short states; all other states of  $A_{h,g}$  (and transitions leading to and from them) are discarded.

Let for  $\ell \geq 1$ ,

 $A_{h,g}^{(\ell+1)}$  be the automaton constructed in the same was as  $A_{h,g}^{(1)}$  but we enlarge the set of states by including all the states q from Q for which  $|\underline{\text{tailq}}| \leq r_0 + \ell$  · k where  $r_0$  is the constant (separating short states from long) from the proof of Lemma 5. Among these states we discard the ones from which we cannot reach (staying within this new set of states only!!) either [A] or D.

In this way we get the sequence

$$A_{h,g}^{(1)}, A_{h,g}^{(2)}, A_{h,g}^{(3)}, \dots$$

This sequence is denoted as  $A_{k,h,g}$  or simply as  $A_{h,g}$  whenever the choice of k is not important. It is called the <u>approximating sequence of  $A_{h,g}$ </u>, and the corresponding sequence of languages

$$T(A_{h,g}^{(1)}), T(A_{h,g}^{(2)}), T(A_{h,g}^{(3)}), \dots$$

is called the approximating sequence of MID(h,g).

Obviously the following holds.

Lemma 6. The construction of  $A_{h,g}$  is effective and moreover  $\bigcup_{i \geq 1} T(A_{h,g}^{(i)}) = MID(h,g).$ 

#### VI. THE DOL SEQUENCE EQUIVALENCE PROBLEM

In this section using the techniques developed so far we prove that the DOL sequence equivalence problem is decidable. Since our proof is completely different (and simpler) than the proof in [1] we believe that it sheds new light on the nature of this problem.

We start by considering elementary DOL systems.

Lemma 7. Given two elementary DOL systems  ${\bf G_1}$  and  ${\bf G_2}$  it is decidable whether or not  $E({\bf G_1})=E({\bf G_2})$ .

Proof.

Let 
$$G_1 = (\Sigma, h, \omega)$$
 and  $G_2 = (\Sigma, g, \omega)$  and let  $H = \underline{Sem}\{h\}$ ,  $G = \underline{Sem}\{g\}$ .  
Let  $A_{h,g} = A_{h,g}^{(1)}, A_{h,g}^{(2)}, \dots$ 

The following algorithm decides whether or not  $E(G_1) = E(G_2)$ .

INPUT:  $G_1, G_2$ .

#### PROCEDURE:

- 1. Set n: = 0.
- 2. Compute  $h^n(\omega), g^n(\omega)$  and construct a finite automaton  $B_n$  such that  $T(B_n) = I(T(A_{h,q}^{(n)}), H) \cap I(T(A_{h,q}^{(n)}), G)$ .
- 3. If  $h^{n}(\omega) \neq g^{n}(\omega)$  output  $E(G_{1}) \neq E(G_{2})$ .
- 4. If  $\omega \in T(B_n)$  output  $E(G_1) = E(G_2)$ .
- 5. Else set n: = n + 1, goto 2.

The effectiveness and correctness of this algorithm follows from Theorem 3, Corollary 1 (and their proofs) and its termination follows from Lemma 6.

Theorem 5. It is decidable whether or not h = T g where h,g are arbitrary elementary homomorphisms and T is a DOL sequence.

#### Proof.

Let  $H = \underline{Sem}\{h\}$ ,  $G = \underline{Sem}\{g\}$  and let g be a homomorphism and  $\omega$  a word such that  $\tau = \omega, f(\omega), f^2(\omega), \ldots$ 

Let 
$$F = \underline{Sem}\{f\}$$
 and let  $A_{h,g} = A_{h,g}^{(1)}, A_{h,g}^{(2)}, \dots$ 

The following algorithm decides whether or not h = g.

INPUT:  $h,g,\tau$ .

#### PROCEDURE:

- 1. Set n: = 0.
- 2. Compute h( $f^n(\omega)$ ),  $g(f^n(\omega))$  and construct a finite automaton  $B_n$  such that  $T(B_n) = I(T(A_{h,q}^{(n)}), F)$ .
- 3. If h( $f^n(\omega)$ )  $\neq g(f^n(\omega))$  output h  $\neq_{\hat{T}} g$ .
- 4. If  $\omega \in T(B_n)$  output  $h =_{\tau} g$ .
- 5. Else set n: = n + 1, goto 2.

The effectiveness and correctness of this algorithm follows from Theorem 3, Corollary 1 (and their proofs) and its termination follows from Lemma 6.

Since it is clear that given a subclass  $\mathcal D$  of the class of DOL systems (of the class of homomorphisms), the sequence equivalence problem for homomorphic images of sequences from  $\mathcal D$  is decidable if and only if it is decidable whether or not two arbitrary homomorphisms from  $\mathcal D$  are equal on an arbitrary DOL sequence  $^{1}$ ), Theorem 5 implies the following result.

Theorem 6. It is decidable whether or not  $h(E(G_1)) = g(E(G_2))$  where h,g are arbitrary homomorphisms and  $G_1, G_2$  are arbitrary elementary DOL systems.

The second part of this result is proved as follows. Let us generate systematically all sequences  $i_1, \ldots, i_k$  from  $\{1,2\}^+$ . For each of them let us find whether or not there exists  $p_1, p_2$ , f satisfying conditions of the lemma. If we succeed we are done; if not we move to the next sequence. The first part of this proof guarantees that we will eventually succeed.

Now we are able to prove the main result of this section.

Theorem 7. The DOL sequence equivalence problem is decidable.

Proof.

Let 
$$G_1 = (\Sigma, h_1, \omega)$$
 and  $G_2 = (\Sigma, h_2, \omega)$  be DOL systems with  $E(G_1) = \omega_0^{(1)}, \omega_1^{(1)}, \ldots$  and  $E(G_2) = \omega_0^{(2)}, \omega_1^{(2)}, \ldots$ 

Since by Lemma 7 the DOL equivalence problem is decidable for elementary DOL systems, let us assume that at least one of  $G_1, G_2$  is simplifiable.

Let  $i_1, \ldots, i_k, p_1, p_2, fp_1, fp_2$  satisfy the statement of Lemma 8 for  $h_1, h_2$ .

Let 
$$g_1 = h_1 h_1 \dots h_i$$
 and  $g_2 = h_2 h_1 \dots h_i$  and let for  $1 \le i \le 2$ ,  $0 \le j \le k+1$ ,  $G_{i,j} = (\Sigma, g_i, \omega_j^{(i)})$ .

- (i)  $E(G_1) = E(G_2)$  if and only if, for every  $0 \le j \le k + 1$ ,  $E(G_{1,j}) + E(G_{2,j})$ .

  Proof of (i).
- 1) Obviously  $E(G_1) = E(G_2)$  implies that  $E(G_{1,j}) = E(G_{2,j})$  for every  $0 \le j \le k+1$ .
- 2) Assume that, for every  $0 \le j \le k+1$ ,  $E(G_{1,j}) = E(G_{2,j})$ . Then if we assume that  $E(G_1) \ne E(G_2)$  we get a contradiction as follows.

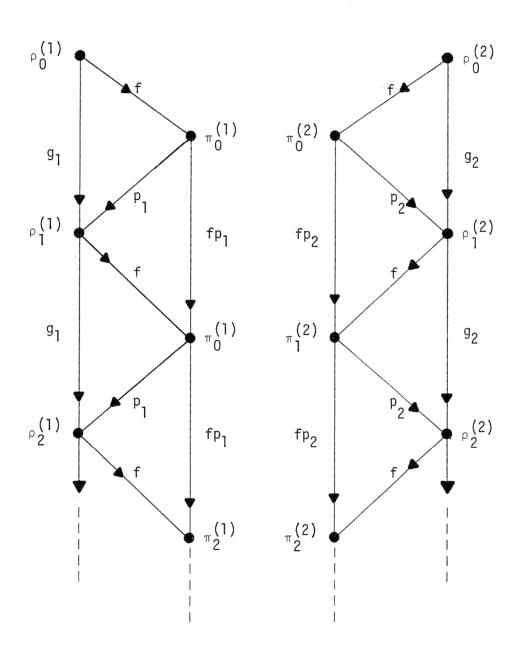
Let m be the minimal integer such that  $\omega_{m}^{(1)} \neq \omega_{m}^{(2)}$ .

If  $m \le k + 1$  then  $E(G_{1,m}) \ne E(G_{2,m})$ ; a contradiction.

If m > k + 1 then for some  $\ell \ge 1$  we have  $0 \le m - \ell \cdot (k+1) \le k+1$  and  $\omega_{m}^{(1)} = \omega_{m}^{(2)} - \ell(k+1)$ .

But then the  $\ell$ 'th element in  $E(G_{1,m} - \ell \cdot (k+1))$  and the  $\ell$ 'th element in  $E(G_{2,m} - \ell \cdot (k+1))$  are different; a contradiction.

(ii) Hence we can decide whether  $E(G_1)=E(G_2)$  if we can decide whether  $E(G_{1,j})=E(G_{2,j})$  for every  $0 \le j \le k+1$ . Let us then fix a j and let  $H_1=G_{1,j}=(\Sigma,g_1,\rho)$  and  $H_2=G_{2,j}=(\Sigma,g_2,\rho)$ . We get the following situation.



Let  $\bar{H}_1=(\Theta,fp_1,\pi_0^{(1)})$  and  $\bar{H}_2=(\Theta,fp_2,\pi_0^{(2)})$  where  $\Theta$  is the alphabet through which  $g_1$  and  $g_2$  are simplified. Note that  $\bar{H}_1,\bar{H}_2$  are simple DOL systems for which by Lemma we can decide whether or not  $E(\bar{H}_1)=E(\bar{H}_2)$ . But if  $E(\bar{H}_1)\neq E(\bar{H}_2)$  then  $E(H_1)\neq E(H_2)$  (note that  $E(\bar{H}_1)$  and  $E(\bar{H}_2)$  are obtained from  $E(H_1)$  and  $E(H_2)$  respectively by the same homomorphism f). Thus  $E(H_1)=E(H_2)$  if and only if  $E(\bar{H}_1)=E(\bar{H}_2)$  and  $p_1,p_2$  are equal on  $E(\bar{H}_1)$ . Since  $p_1,p_2$  are elementary, Theorem 5 implies that it is decidable whether  $p_1,p_2$  are equal on  $E(\bar{H}_1)$ .

Thus the theorem holds.

#### VII. ON AN "EXPLICIT" SOLUTION OF THE DOL SEQUENCE EQUIVALENCE PROBLEM

In a sense a positive solution to the DOL sequence equivalence problem seems to be natural if (given two arbitrary DOL systems  $G_1, G_2$ ) it provides a way to compute a constant  $C_{G_1, G_2}$  such that  $G_1$  and  $G_2$  are equivalent if and only if the first  $C_{G_1, G_2}$  elements in  $E(G_1)$  and  $E(G_2)$  are identical.

In this section we analyze our solution and provide an algorithm to compute such a constant explicitely. Clearly we realize that this is not the best constant (we conjecture that such a constant should not exceed  $r \cdot n$ ! where r is the maximal length of the image of an alphabet involved under one of the homomorphisms and n is the size of the alphabet). Still we believe that the analysis we give to find such a constant provides some additional insight into the problem involved.

Theorem 8. There exists an algorithm which given arbitrary two DOL systems  $G_1$  and  $G_2$  provides a constant  $C_{G_1,G_2}$  such that  $E(G_1)=E(G_2)$  if and only if the first  $C_{G_1,G_2}$  elements of  $E(G_1)$  and  $E(G_2)$  are equal.

#### Proof.

The proof of this result goes through a careful analysis of automata in the approximating sequence  $A_{h,g}$  and an analysis of the proof of Lemma 8. Accordingly we have the following two Lemmas.

Lemma 9. Let  $G_1=(\Sigma,h_1,\omega)$  and  $G_2=(\Sigma,h_2,\omega)$  be two elementary DOL systems with  $E(G_1)=\omega_0^{(1)},\omega_1^{(1)},\ldots$  and  $E(G_2)=\omega_0^{(2)},\omega_1^{(2)},\ldots$ . Let  $n=\#\Sigma$ ,  $m=\max\{\max n,\max n\}$ ,  $s=|\omega|$  and r be such that  $\max\{|\tan n|: q\in Q_2\}\le r$  where  $Q_2$ ,  $\tan n$  are defined as in Section V. Let  $\Psi$  be a function from  $N^+$  into  $N^+$  defined by  $\Psi(x)=s\cdot m^{X^n} \cdot x$  + x. Then  $E(G_1)=E(G_2)$  if and only if  $\omega_u^{(1)}=\omega_u^{(2)}$ , where

$$u = (\Psi^2 \cdot n^r_{(2n^r)})^n \cdot \Psi^2 \cdot n^r_{(2 \cdot n^r)}.$$

Proof.

It suffices to show that if  $\omega_u^{(1)}=\omega_u^{(2)}$  then for every  $v\geq u$ ,  $\omega_v^{(1)}=\omega_v^{(2)}$ .

Let  $B_1, B_2, \ldots$  be a sequence approximating  $A_{h_1, h_2}$  such that (recall that  $A_{h_1, h_2} = A_{h_1, h_2}^{(1)}, A_{h_1, h_2}^{(2)}, \ldots$ ):

 $B_1 = A_{h_1,h_2}^{(1)}$ , i.e., it consists of all the short states of  $A_{h_1,h_2}$ , and for  $i \ge 1$ ,

 $B_{i+1}$  is the first automaton encountered in  $A_{h_1,h_2}$  which is different from  $B_i$ , i.e., it consists of more states.

Let  $\omega_{k_1}^{(1)}, \omega_{k_2}^{(1)}, \ldots$  be the subsequence of  $E(G_1)$  obtained as follows: for  $i \geq 1$ ,  $\omega_{k_1}^{(1)}$  is the first element of  $E(G_1)$  that does not belong to  $T(B_i)$ . Clearly there exists an  $\ell$  such that  $T(B_{\ell}) = T(A_{h_1,h_2})$ . Let us estimate  $\ell$  first.

(i)  $\ell \leq 2 \cdot n^r$ .

Proof of (i).

Note that in  $A_{h_1,h_2}$  each short state which has a transition to a long state and can lead back to a short state gives rise to a unique trace (path) in  $A_{h_1,h_2}$ . Hence the number of these paths is bounded by the number of short states which is clearly bounded by  $2 \cdot n^r$ . But each  $B_{i+1}$  results from  $B_i$  by adding at least one such path to  $B_i$ . Hence the bound.

Let  $d(B_i)$  denote the number of states in  $B_i$ .

(ii)  $d(B_1) \le 2 \cdot n^r$ .

Proof of (ii).

Obvious.

(iii) 
$$k_i \leq d(B_i)^n \cdot d(B_i)$$
.

## Proof of (iii).

Let us consider the sequence of automata  $\tau_i = B_i$ ,  $(B_i)_{h_1}$ ,  $(B_i)_{h_1}^2$ ,... (see the notation from the proof of Theorem 3). Then  $k_i$  is the smallest integer j such that  $\omega \notin (B_j)_j$ . However the number of all different automata in  $h_1$   $n \cdot d(B_i)$  and hence the bound.  $n \cdot d(B_i)$ 

$$(iv) \quad d(B_{i+1}) \leq s \cdot m \qquad \qquad + d(B_{i})$$

# Proof of (iv).

Going from B<sub>i</sub> to B<sub>i+1</sub> we get  $\omega_{k_i}$  to be in  $T(B_{i+1})$ . In other words  $\omega_{k_i}$  gives rise to a trace starting with  $q_{in}$  then coming to a short state from which a unique path leads through the sequence of long states (which were not in B<sub>i</sub>!!) and come back to a short state (which was in B<sub>i</sub>). Thus we added no more than  $|\omega_{k_i}|$  new states to B<sub>i</sub> to obtain B<sub>i+1</sub>. Hence  $d(B_{i+1}) \leq |\omega_{k_i}| + d(B_i).$  But  $|\omega_{k_i}| \leq s \cdot m^i$  and so  $d(B_{i+1}) \leq s \cdot m^i + d(B_i)$ .

(v) Now we complete the proof of the lemma as follows.

From (iv) we have that  $d(B_{i+1}) \leq \Psi(d(B_i))$ , hence from (i) it follows that  $d(B_{\ell}) \leq \Psi^2 \cdot n^r (2 \cdot n^r).$ 

Then from (iii) it follows that

$$k_{\ell} \leq (\Psi^{2} \cdot n^{r}(2 \cdot n^{r}))^{n} \cdot \Psi^{2} \cdot n^{r}(2 \cdot n^{r}).$$

But clearly

$$\omega_{k_{\varrho}}^{(1)} \in T(B_{\ell})$$
 if and only if

$$\omega_{k_{\ell}}^{(1)} \in T(A_{h_1,h_2})$$
 if and only if

for every  $v \ge k_{\ell}$ ,  $\omega_{v}^{(1)} \in T(A_{h_1,h_2})$ 

if and only if, for every  $v \ge k_{\ell}$ ,  $\omega_{V}^{(1)} = \omega_{V}^{(2)}$ .

Hence the lemma holds.

Lemma 10. Let  $h_1,h_2 \in HOM(\Sigma,\Sigma)$  where  $\#\Sigma = n$ . Let  $\Phi$  be a function from  $N^+$  into  $N^+$  defined by  $\Phi(x) = 2 \cdot \Phi(x-1) + 1$  for  $x \ge 1$ . Then for some  $k \le \Phi^n(1)$  there exists a sequence  $i_1,\ldots,i_k$  of elements from  $\{1,2\}$  and homomorphisms  $f,p_1,p_2$  satisfying the statement of Lemma 8.

#### Proof.

Since the result is trivial when both  $h_1$  and  $h_2$  are elementary, let us assume that at least one of  $h_1,h_2$  is simplifiable (assume, e.g., that  $h_1$  is).

Then  $h_1$  = pf for some  $f \in HOM(\Sigma, \Theta_1)$ ,  $p \in HOM(\Theta_1, \Sigma)$  where  $\#\Theta_1 < \#\Sigma$  and  $h_1h_1 = p_1f$ ,  $h_2h_1 = p_2f$  for some  $p_1, p_2 \in HOM(\Theta_1, \Sigma)$ .

Thus let us assume inductively that we are given a sequence  $i_1,\ldots,i_k$  in  $\{1,2\}^+$  such that

$$h_{i_1}...h_{i_k} = pf$$
 for some  $f \in HOM(\Sigma, \Theta_{i_1}...i_k)$ ,  $p \in HOM(\Theta_{i_1}...i_k, \Sigma)$  where

$$h_1 h_1 \dots h_{i_k} = p_1 f, h_2 h_1 \dots h_{i_k} = p_2 f \text{ for } p_1, p_2 \in HOM(\Theta_{i_1 \dots i_k}, \Sigma).$$

If  $p_1, p_2, fp_1, fp_2$  satisfy the statement of the lemma we are done. Otherwise there can be two reasons why the statement of the lemma is not (yet) true.

1) At least one of  $p_1, p_2$  is simplifiable.

Thus there exists an alphabet  $\Delta$  with  $\#\Delta < \#\Theta_{\mbox{$1$}}\dots \mbox{$i$}_k$  such that at least one of  $h_1h_1\dots h_{\mbox{$i$}_k}$ ,  $h_2h_1\dots h_{\mbox{$i$}_k}$  can be simplified through  $\Delta$ . For example  $h_1h_1\dots h_{\mbox{$i$}_k} = \bar{p}_1\bar{f} \text{ where } \bar{f} \in \text{HOM}(\Sigma,\Delta) \text{ and } \bar{p} \in \text{HOM}(\Delta,\Sigma).$ 

But then also

 $h_1h_1\dots h_i h_1h_1\dots h_i$  and  $h_2h_1\dots h_i h_1h_1\dots h_i$  can be simplified through  $\Delta$ . Thus in this case we have got a new sequence  $(i_1,\dots,i_k,1,i_1,\dots,i_k)$  of length 2k+1 and a new smaller alphabet  $\Delta=\Theta_1\dots i_k 1i_1\dots i_k$  satisfying our inductive assumption.

2) At least one of fp<sub>1</sub>,fp<sub>2</sub> is simplifiable through an alphabet  $\triangle$  with  $\#\triangle < \#\Theta_{i_1...i_k}$ . (Assume fp<sub>1</sub> is.)

Then also

 $p_1 f p_1 f$  and  $p_2 f p_1 f$  are simplifiable through  $\Delta$ .

Hence we have got a new sequence  $(i_1, ..., i_k, 1, i_1, ..., i_k)$  of length 2k + 1 and a new smaller alphabet  $\Delta = \Theta_1, ..., i_k$  satisfying our inductive assumption.

Since we can iterate our procedure at most  $n=\#\Sigma$  times (and from the proof of the lemma we know that for a "minimal alphabet" the conditions of the lemma are satisfied)  $\Phi^n(1)$  yields us a desired upper bound, M and the Lemma holds.

Now the two previous lemma together with the proof of Theorem 6 yield  ${^{C}_{G_1,G_2}} = 2(\Phi^n(1) + 1) + (\Psi^2 \cdot n^r(2 \cdot n^r))^n \cdot \Psi^2 \cdot n^r(2 \cdot n^r)$ 

#### REFERENCES

- [1]. Culik II, K. and Fris, I., 1977. The decidability of the equivalence problem for DOL systems, <u>Information and Control</u>, to appear.
- [2]. Ehrenfeucht, A., Lee, K.P. and Rozenberg, G., 1974. Subword complexities of various classes of deterministic developmental languages without interactions, Theoretical Computer Science, 1:59-76.
- [3]. Ehrenfeucht, A. and Rozenberg, G., 1977. Simplifications of homomorphisms, submitted for publication.
- [4]. Rozenberg, G. and Salomaa, A., 1974. The mathematical theory of L systems, In: J. Tou (ed.), <u>Advances in Information Systems Science</u>, vol. 6, Plenum Press, New York, 161-206.
- [5]. Salomaa, A. and Soittola, M., 1977, <u>Automata-Theoretic Aspects of</u>
  Formal Power Series, Springer Verlag, to appear.

# FOOTNOTES

1) The foregoing implication was known to K. Culik II (personal communication).

### ACKNOWLEDGMENTS

The authors greatfully acknowledge the financial support by the NSF Grant Number MCS75-09972 and by IBM Belgium. They are also very much indebted to L. Fosdick for making the work on this and other papers possible.