

REPETITIVE PROPERTIES  
OF DOL LANGUAGES

by

A. Ehrenfeucht\*

and

G. Rozenberg\*\*

CU-CS-233-82

September, 1982

\*Department of Computer Science University of Colorado at  
Boulder, Boulder, Colorado 80309

\*\*Institute of Applied Mathematics and Computer Science,  
University of Leiden, Leiden, The Netherlands

All correspondence to second author.

This research was supported by NSF grant MCS 79-03838.

ANY OPINIONS, FINDINGS, AND CONCLUSIONS  
OR RECOMMENDATIONS EXPRESSED IN THIS  
PUBLICATION ARE THOSE OF THE AUTHOR AND  
DO NOT NECESSARILY REFLECT THE VIEWS OF THE  
NATIONAL SCIENCE FOUNDATION.

THIS MATERIAL IS BASED UPON WORK SUPPORTED  
IN PART BY THE NATIONAL SCIENCE FOUNDATION  
UNDER GRANT NO. MCS 79-03838

## REPETITIVE PROPERTIES OF DOL LANGUAGES

by

A. Ehrenfeucht  
Department of Computer Science  
University of Colorado  
Boulder, Colorado, 80309

and

G. Rozenberg  
Institute of Applied Math. and Comp. Sci.  
University of Leiden  
Leiden, The Netherlands

### INTRODUCTION

The investigation of the combinatorial structure of languages forms an important part of formal language theory. One of the most basic combinatorial structures of languages is the repetition of subwords (in words of a language). Roughly speaking, the investigation of repetitions of subwords can be divided into two (certainly not disjoint!) directions.

- (1) The investigation of languages where repetitions of subwords (in the words of the language) are forbidden. This area was initiated by Thue ([T]) in 1906 and since then this area was a subject of active investigation in numerous areas of mathematics and in formal language theory (see, e.g., [BEM], [C], [MH], and [S1]).
- (2) The investigation of languages where repetitions of subwords (must) occur. The most classical example here is the class of context-free languages where the celebrated "pumping lemma" forces arbitrary long repetitions to be present in an infinite context-free language.

Recently one notices a revival of interest in area (1) ("Thue problems") among formal language theorists (see, e.g., [B], [H], [K], [S2]). In particular it was discovered that the theory of nonrepetitive sequences of Thue [T] is strongly related to the theory of DOL systems (see, e.g., [RS]). As a matter of fact it was pointed out in [B] that most (if not all) examples of the so called square-free

sequences constructed in the literature are either DOL sequences or their codings. Thus by now quite a lot is known about DOL languages (sequences) not containing repetitions of subwords (see also [ER3]).

On the other hand very little is known on DOL languages containing repetitive structures. The pumping-like properties do not hold for DOL languages and "detecting" repetitiveness in a DOL language becomes a challenging problem.

This paper is devoted to the study of repetitiveness in DOL languages. Let us first make the notion of repetitiveness (of subwords in a language) more precise. We say that a language  $K \subseteq \Sigma^*$  is *repetitive* if for each  $n \geq 1$  there exists a word  $w \in \Sigma^+$  such that  $w^n$  is a subword of  $K$ . We say that  $K$  is *strongly repetitive* if there exists a word  $w \in \Sigma^+$  such that  $w^n$  is a subword of a word of  $K$  for each  $n \geq 1$ . It is easily seen that there exist repetitive languages that are not strongly repetitive, while on the other hand each strongly repetitive language is obviously repetitive. By the pumping lemma infinite context-free languages are strongly repetitive.

We demonstrate that

- (1) a DOL language is repetitive if and only if it is strongly repetitive and
- (2) it is decidable whether or not an arbitrary DOL system generates a repetitive language.

This paper is not a full paper. It states all the main and technical results, however the proofs are mostly omitted. The full version of this paper will be published elsewhere.

## 1. PRELIMINARIES

We assume the reader to be familiar with the basic theory of DOL systems (see, e.g., [RS]). We will use the standard notation and terminology concerning

DOL systems (as used in [RS]).

Perhaps recalling the following notational matters will make the reading of this paper easier.

$\mathbf{N}$  denotes the set of nonnegative integers and  $\mathbf{N}^+$  denotes the set of positive integers. For a set  $A$ ,  $\#A$  denotes its cardinality.  $\Lambda$  denotes the empty word. For a nonempty word  $w$ ,  $first(w)$  denotes its first letter and  $last(w)$  denotes its last letter; for  $n \in \mathbf{N}$ ,  $pref_n(w)$  denotes the prefix of  $w$  of length  $n$  and  $sub_n(w)$  denotes the set of subwords (segments) of  $w$  of length  $n$ . Then  $sub(w)$  denotes the set of all subwords of  $w$  and for a language  $K$ ,  $sub(K) = \bigcup_{x \in K} sub(x)$ . For a DOL system  $G = (\Sigma, h, \omega)$ ,  $E(G)$  denotes its sequence,  $L(G)$  its language and  $maxr(G) = \max\{|z| : h(a) = z \text{ for some } a \in \Sigma\}$ . A letter  $a \in \Sigma$  is called *alive* if  $h^n(a) \neq \Lambda$  for all  $n \in \mathbf{N}^+$ . We will use  $T(G)$  to denote the (infinite) derivation tree corresponding to  $E(G)$ . For a node  $x$  in  $T(G)$ ,  $lb(x)$  denotes its label,  $anc(x)$  its direct ancestor and  $anc^2(x)$  the direct ancestor of  $anc(x)$ . Let  $E(G) = \omega_0, \omega_1, \dots$ . For a node  $x$  on the level  $r \geq 0$  of  $T(G)$  (counted top-down) and (an occurrence of) a subword  $z$  of  $\omega_s$  where  $s \geq r$  we use  $contr_z(x)$  to denote the contribution of  $x$  to  $z$ ; similarly if  $u$  is (an occurrence of) a subword in  $\omega_r$  then we use  $contr_z(u)$  to denote the contribution of  $u$  to  $z$ .

*In order not to overburden the (already involved) notation*

(1) *we will often not distinguish notationally between a (sub)word and its occurrence, and*

(2) *we will often not distinguish in our notation between nodes and their labels;*

*as the precise meaning should be clear from the context, these conventions should not lead to a confusion.*

We will recall now two useful notions concerning DOL systems. Let  $G = (\Sigma, h, \omega)$  be a DOL system.

A letter  $a \in \Sigma$  has *rank* 0 (in  $G$ ) (see, e.g., [ER2]) if  $L(G_a)$  is finite, where  $G_a = (\Sigma, h, a)$ . Let for  $i \geq 1, \Sigma_{(i)} = \Sigma - \{a \in \Sigma : a \text{ is of rank smaller than } i\}$  and let  $f_{(i)}$  be the homomorphism of  $\Sigma^*$  defined by :  $f_{(i)}(a) = a$  for  $a \in \Sigma_{(i)}$  and  $f_{(i)}(a) = \Lambda$  for  $a \in \Sigma - \Sigma_{(i)}$ . Then let  $h_{(i)}$  be the homomorphism of  $\Sigma_{(i)}^*$  defined by  $h_{(i)}(a) = f_{(i)}(h(a))$ . If a letter  $a \in \Sigma_{(i)}$  is such that the language of the DOL system  $(\Sigma_{(i)}, h_{(i)}, a)$  is finite then  $a$  has *rank*  $i$  (in  $G$ ). For  $i \geq 0$ , we use  $\Sigma_i$  to denote the set of all letters from  $\Sigma$  of rank  $i$ .

Let  $G = (\Sigma, h, \omega)$  and  $\bar{G} = (\bar{\Sigma}, \bar{h}, \bar{\omega})$  be DOL systems.  $\bar{G}$  is called a *simplification* of  $G$  if  $\#\bar{\Sigma} < \#\Sigma$  and there exist homomorphisms  $f : \Sigma^* \rightarrow \bar{\Sigma}^*$ ,  $g : \bar{\Sigma}^* \rightarrow \Sigma^*$  such that  $h = g f$ ,  $\bar{h} = f g$  and  $\bar{\omega} = f(\omega)$ . If  $G$  does not have a simplification it is called *elementary*. It is known ([ER1]) that if  $G$  is elementary then  $h$  is injective. If  $G_0, G_1, \dots, G_n, n \geq 0$ , is the sequence of DOL systems such that  $G_0 = G$ ,  $G_i$  is a simplification of  $G_{i-1}$  for  $1 \leq i \leq n$  and  $G_n$  is elementary, then  $G_n$  is called an *elementary version* of  $G$ .

## 2. BASIC DEFINITIONS AND RESULTS

In this section we define some basic notions (and some basic results concerning them) to be investigated in this paper. These include the main notion of (strong) repetitiveness of a language as well as several more technical notions which will be useful for proving the main results of this paper.

*Definition.* Let  $K$  be a language,  $K \subseteq \Sigma^*$ .

- (1)  $K$  is *repetitive* if for each  $n \in \mathbb{N}^+$  there exists a word  $w \in \Sigma^+$  such that  $w^n \in \text{sub}(K)$ .

(2)  $K$  is *strongly repetitive* if there exists a word  $w \in \Sigma^+$  such that  $w^n \in \text{sub}(K)$  for each  $n \in \mathbf{N}^+$ . ■

Obviously, if  $K$  is strongly repetitive then  $K$  is repetitive, but there exist repetitive languages, that are not strongly repetitive. Consider, e.g., the language  $K_0 \subseteq \{a, b, c, d\}^+$  defined by

$$K_0 = \{(wd)^n : n \in \mathbf{N}^+, w \in \{a, b, c\}^+, |w| = n \text{ and for no } x, y \in \{a, b, c\}^*, z \in \{a, b, c\}^+, w = xzzy\}.$$

Clearly  $K_0$  is repetitive but not strongly repetitive language (notice that  $K$  is a context-sensitive language).

*Definition.* A DOL system  $G$  is called (*strongly*) *repetitive* if  $L(G)$  is (*strongly*) *repetitive*. ■

The following special subclass of DOL systems will be useful in the considerations of the next section.

*Definition.* A DOL system  $G = (\Sigma, h, \omega)$  is *pushy* if  $\text{sub}(L(G)) \cap \Sigma_0^*$  is infinite; otherwise  $G$  is *not pushy*. ■

If a DOL system  $G$  is not pushy then  $q(G)$  denotes  $\max\{|w| : w \in \text{sub}(L(G)) \cap \Sigma_0^*\}$ .

*Lemma 2.1.*

- (1) It is decidable whether or not an arbitrary DOL system is pushy.
- (2) If a DOL system  $G = (\Sigma, h, \omega)$  is not pushy then  $\Sigma_i = \phi$  for all  $i > 0$ .
- (3) If a DOL system  $G$  is not pushy then  $q(G)$  is effectively computable. ■

Our next notion is the fundamental technical notion of this paper.

*Definition.* A DOL system  $G = (\Sigma, h, \omega)$  is called *special*, abbreviated a SDOL system, if it satisfies the following conditions.

- (0)  $G$  is reduced.

- (1)  $G$  is sliced meaning that
- (1.1) for each  $a \in \Sigma$ , and each  $n \in \mathbf{N}^+$ ,  $\text{alph}(h^n(a)) = \text{alph}(h(a))$ ,
- (1.2) for each  $a \in \Sigma$ , the length sequence  $\{|h^n(a)|\}_{n \geq 0}$  is either strictly increasing or constant and
- (1.3)  $\omega \in \Sigma$ .
- (2)  $G$  is strongly growing meaning that
- (2.1)  $G$  is propagating and
- (2.2) no letter in  $G$  has a rank (including the zero rank).
- (3)  $G$  is elementary. ■

The next few results bind the notion of repetitiveness with several subclasses of DOL systems as well as they indicate how this notion carries over through some operations on languages and DOL systems.

*Lemma 2.2.* Let  $G$  be a DOL system.

- (1) If  $G$  is pushy then  $G$  is strongly repetitive.
- (2) If  $G$  is finite then  $G$  is not repetitive. ■

*Definition.* Let  $K$  be a language and let  $(K_1, \dots, K_n)$ ,  $n \geq 1$ , be a  $n$ -tuple of languages. Then  $K < (K_1, \dots, K_n)$  if  $K \subseteq K_1 K_2 \dots K_n$  and  $K_i \subseteq \text{sub}(K)$  for each  $1 \leq i \leq n$ . ■

*Lemma 2.3.* Let  $K, K_1, \dots, K_n$ ,  $n \geq 1$ , be languages.

- (1) Let  $K = \bigcup_{i=1}^n K_i$ . Then

$K$  is (strongly) repetitive if and only if there exists a  $1 \leq i \leq n$  such that  $K_i$  is (strongly) repetitive.

- (2) Let  $K < (K_1, \dots, K_n)$ . Then

$K$  is (strongly) repetitive if and only if there exists a  $1 \leq i \leq n$  such that  $K_i$  is (strongly) repetitive. ■



*Lemma 2.4.* Let  $G$  be a DOL system and let  $G'$  be its simplification. Then  $G$  is (strongly) repetitive if and only if  $G'$  is (strongly repetitive). ■

### 3. MAIN RESULTS

In this section we state two main results of this paper and indicate the strategy of their proofs.

The following two results are the main results of this paper.

*Theorem 1.* It is decidable whether or not an arbitrary DOL system  $G$  is repetitive. ■

*Theorem 2.* Every repetitive DOL system is strongly repetitive. ■

In order to prove these results we will prove the following two (more technical) theorems. They allow us to concentrate on SDOL systems (rather than consider arbitrary DOL systems).

*Theorem 3.*

- (1) It is decidable whether or not an arbitrary DOL system is repetitive if and only if it is decidable whether or not an arbitrary SDOL system is repetitive.
- (2) If every repetitive SDOL system is strongly repetitive, then every repetitive DOL system is strongly repetitive. ■

*Theorem 4.*

- (1) It is decidable whether or not an arbitrary SDOL system is repetitive.
- (2) Every repetitive SDOL system is strongly repetitive. ■

Clearly Theorem 3 and Theorem 4 together imply Theorem 1 and Theorem 2. Thus the rest of this paper is devoted to proofs of Theorem 3 and Theorem 4.

In the next section we prove Theorem 3. In Section 4 we consider closed and strongly closed subalphabets of the alphabet of a SDOL system. Considerations of this section form important technical tools for Section 5 where Theorem 4 is proved.

#### 4. PROOF OF THEOREM 3

In this section Theorem 3 is proved.

*Theorem 3.*

- (i) It is decidable whether or not an arbitrary DOL system is repetitive if and only if it is decidable whether or not an arbitrary SDOL system is repetitive.
- (ii) If every repetitive SDOL system is strongly repetitive then every repetitive DOL system is strongly repetitive. ■

*Proof.* (i) Clearly it suffices to prove the if part of the statement only. To this aim we proceed as follows.

Let  $G = (\Sigma, h, \omega)$  be an arbitrary DOL system.

First we decide whether or not  $G$  is finite (it is well known that finiteness is decidable for DOL systems). If  $G$  is finite then (see Lemma 2.2(2))  $G$  is not repetitive and we are done. If  $G$  is infinite then (see Lemma 2.1(1)) we decide whether or not  $G$  is pushy. If it is, then (by Lemma 2.2(1))  $G$  is strongly repetitive and we are done.

Thus let us assume that  $G$  is not pushy.

Let  $G^c$  be the "coded version of  $G$ " defined as follows.  $G^c = (\Sigma^c, h^c, \omega^c)$  where

$$\Sigma^c = \{ (\alpha, x, \beta) : x \in \Sigma - \Sigma_0, \alpha, \beta \in \Sigma_0^* \text{ and } |\alpha|, |\beta| \leq q(G) \},$$

$$\omega^c = (\alpha_1, y_1, \alpha_2)(\alpha_2, y_2, \alpha_3) \dots (\alpha_{n-1}, y_{n-1}, \alpha_n) \text{ where}$$

$$\omega = \alpha_1 y_1 \alpha_2 y_2 \dots \alpha_{n-1} y_{n-1} \alpha_n, y_i \in \Sigma - \Sigma_0 \text{ and } \alpha_j \in \Sigma_0^* \text{ for } 1 \leq i \leq n-1$$

and  $1 \leq j \leq n$ ,

for  $(\alpha, x, \beta) \in \Sigma^c$ ,

$$h^c((\alpha, x, \beta)) = (h(\alpha)\alpha_1, y_1, \alpha_2)(\alpha_2, y_2, \alpha_3) \dots (\alpha_{n-1}, y_{n-1}, \alpha_n h(\beta))$$

where

$$h(x) = \alpha_1 y_1 \alpha_2 y_2 \dots \alpha_{n-1} y_{n-1} \alpha_n, y_i \in \Sigma - \Sigma_0 \text{ and } \alpha_j \in \Sigma_0^* \text{ for } 1 \leq i \leq n-1$$

and  $1 \leq j \leq n$ .

By Lemma 2.1  $G^c$  is effectively constructible.

*Claim 4.1.*

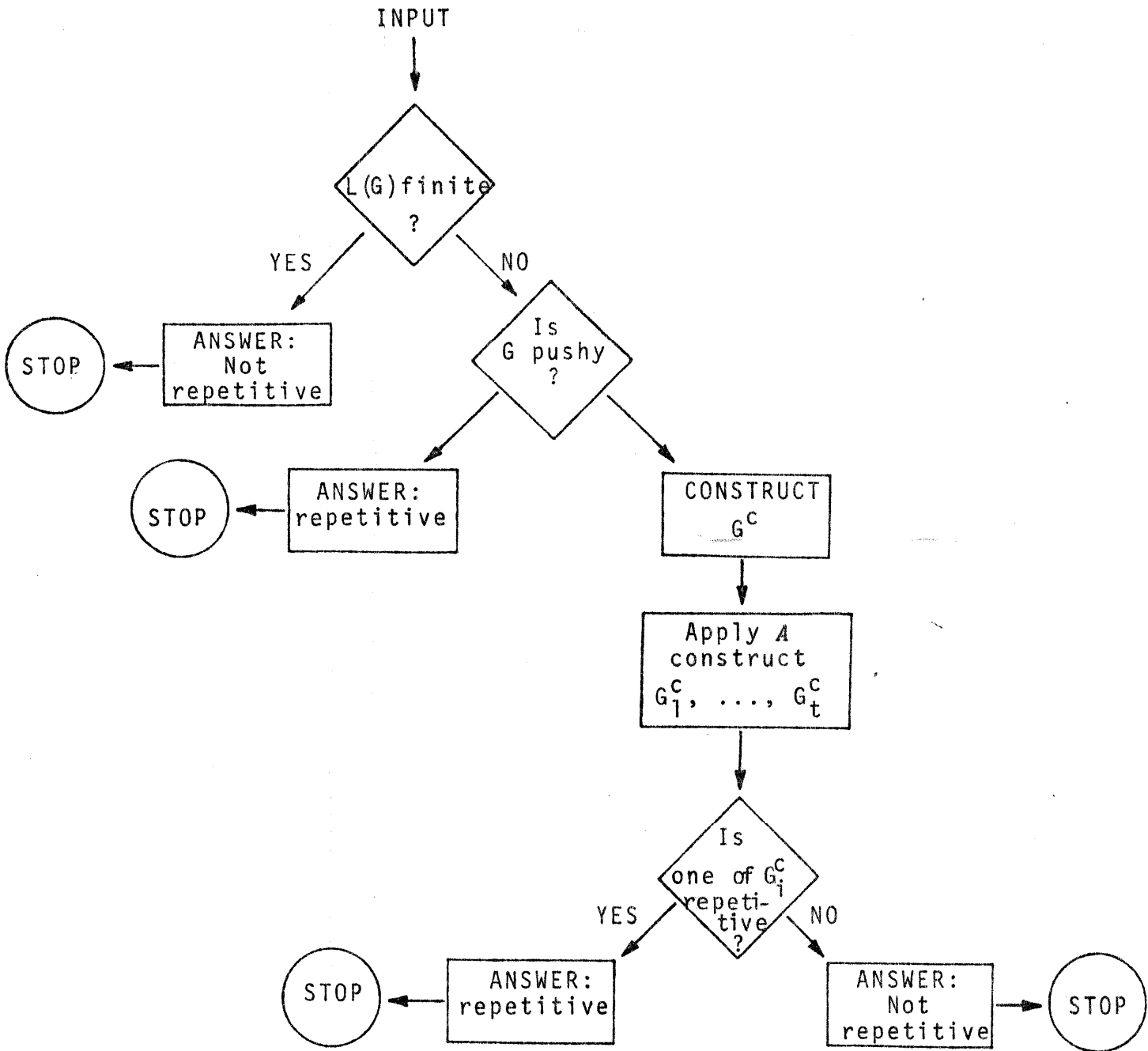
- (1)  $G^c$  is (strongly) repetitive if and only if  $G$  is (strongly) repetitive.
- (2)  $G^c$  is strongly growing. ■

*Claim 4.2.* There exists an algorithm which given a strongly growing DOL system  $H$  produces a finite set  $H_1, \dots, H_t, t \geq 1$ , of DOL systems such that

- (1)  $H$  is (strongly) repetitive if and only if  $H_i$  is (strongly) repetitive for some  $1 \leq i \leq t$ .
- (2)  $H_i$  is special for each  $1 \leq i \leq t$ . ■

Now we complete the proof of Theorem 3.(i) as follows.

Let us consider the algorithm **R** given by the following diagram.



Clearly, if it is decidable whether or not an arbitrary SDOL system is repetitive, then (from Claim 4.1 and Claim 4.2 it follows that) the algorithm **R** decides whether or not an arbitrary DOL system is repetitive.

Hence (i) holds.

(ii) To prove (ii) let us assume that every repetitive SDOL system is strongly repetitive. Let us analyze the algorithm **R** and in particular the cases when it decides that a DOL system in question is repetitive. There are two such cases.

(1) The answer "repetitive" given on the exit YES from the test "Is  $G$  pushy?". In this case, by Lemma 2.2(1),  $G$  is also strongly repetitive.

(2) The answer "repetitive" given on the exit YES from the test "Is one of  $G_i^c$  repetitive?". In this case we know that (at least) one of the "component systems"  $G_1^c, \dots, G_t^c$  is repetitive; since all these systems are special, our assumption implies that (at least) one of the systems  $G_1^c, \dots, G_t^c$  is strongly repetitive. Then, by Lemma 2.3 and Lemma 2.4,  $G$  is strongly repetitive.

Hence, whenever  $G$  is repetitive it is also strongly repetitive and (ii) holds.

Consequently Theorem 1 holds.  $\square$

## 5. CLOSED AND STRONGLY CLOSED SUBSETS OF $\Sigma$

In this section we define and investigate closed and strongly closed subsets of (the alphabet of) a DOL system. The results of this section form an important tool in proving Theorem 4 in the next section.

Let  $G = (\Sigma, h, S)$  be a DOL system and let  $\Theta$  be a nonempty subset of  $\Sigma$ . We say that  $\Theta$  is *closed* (with respect to  $G$ ) if  $h(a) \in \Theta^*$  for each  $a \in \Theta$  and we say that  $\Theta$  is *strongly closed* (with respect to  $G$ ) if  $\text{alph}(h(a)) = \Theta$  for each  $a \in \Theta$ . Note that if  $\Theta$  is strongly closed then it is also closed.

Let  $w \in \text{sub}(L(G))$ , let  $\Theta \subseteq \Sigma$  and let  $u \in \Theta^+$ . We say that  $u$  is a  $\Theta$ -block (of  $w$ ) if  $w = \alpha a u b \beta$  where  $a, b \in \Sigma - \Theta$ . A  $\Theta$ -block  $u$  is *maximal* in  $w$  if no other  $\Theta$ -block in  $w$  is longer than  $u$ ; then  $B_{\Theta}(w)$  denotes the number of different maximal  $\Theta$ -blocks in  $w$ . (E.g., if  $w = a^3 c a^2 c^2 a^2 c a c$  and  $\Theta = \{a\}$  then  $B_{\Theta}(w) = 2$ ).

Now let  $G = (\Sigma, h, \omega)$  be an arbitrary (but fixed) special DOL system with  $E(G) = \omega_0, \omega_1, \dots$  and let  $m = \text{maxr}(G)$ . We will investigate several useful properties of closed and strongly closed subsets of  $\Sigma$ .

*Lemma 5.5*  $B_{\Theta}(w) \leq \#\Sigma m^4$  for each closed subset  $\Theta$  of  $\Sigma$  and each  $w \in \text{sub}(L(G))$ . ■

Now we move to investigate strongly closed subsets of  $\Sigma$ . We start by noting the following property.

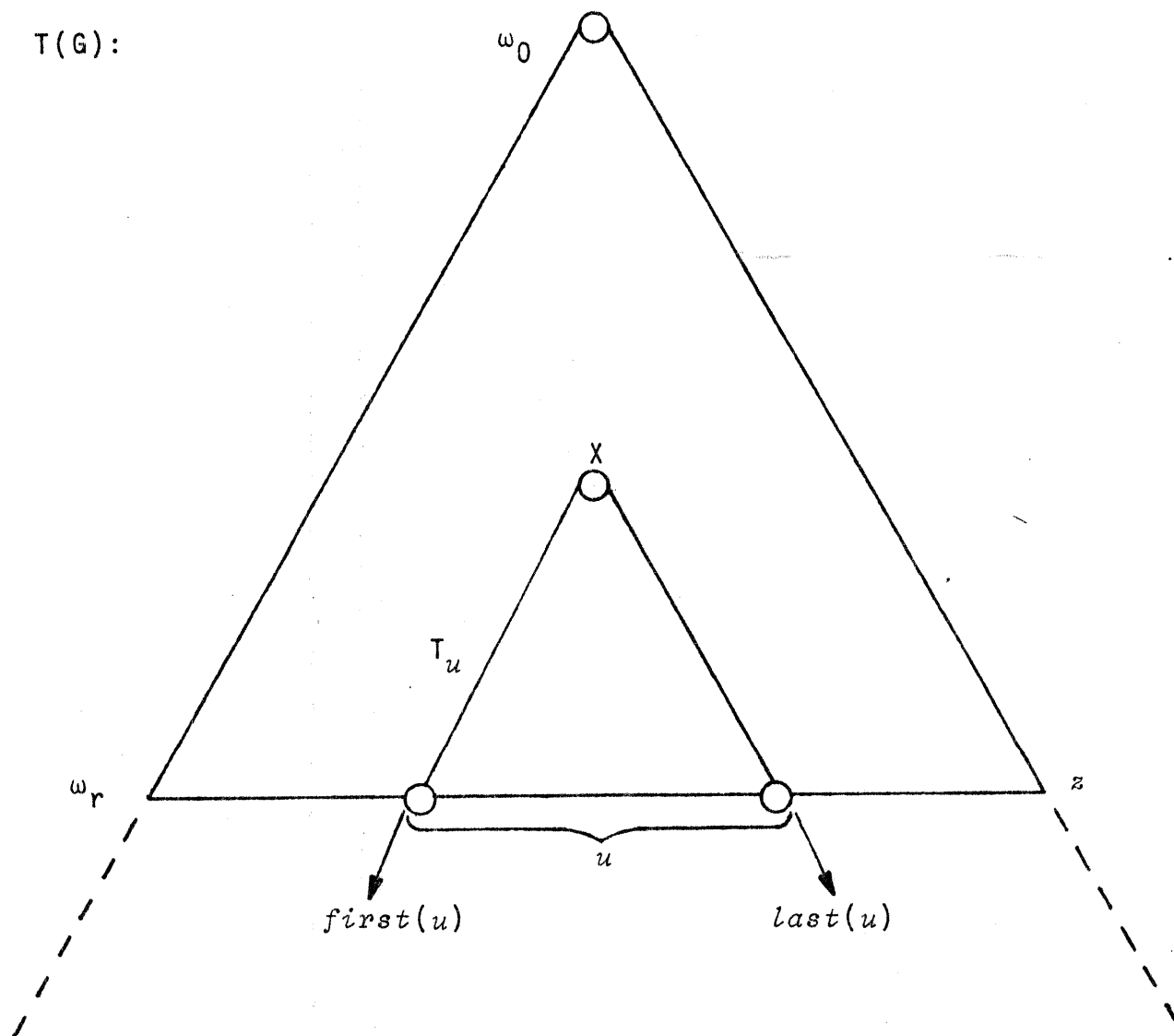
*Lemma 5.6* Let  $\Theta$  be a strongly closed subset of  $\Sigma$ . For every  $n \in \mathbf{N}^+$  and every  $a, b \in \Theta$ ,  $|h^n(b)| \leq m \cdot |h^n(a)|$ . ■

The relevance of strongly closed subsets of  $\Sigma$  to the investigation of repetitive properties of  $G$  stems from the following result.

*Lemma 5.7.* If  $n > \#\Sigma m^4 + 4$  then for each  $w \neq \Lambda$  holds: if  $w^n \in \text{sub}(L(G))$  then  $\text{alph}(w)$  is strongly closed. ■

We define now a concept important for our further considerations.

Let  $z \in L(G)$ ,  $\Theta$  be a strongly closed subset of  $\Sigma$  and let  $u \in \text{sub}(z) \cap \Theta^+$ . Thus we have the following situation.

$T(G):$ 

where  $X$  is the first (bottom-up) common ancestor of  $first(u)$  and  $last(u)$ ;  $T_u$  is a subtree of  $T(G)$  rooted at  $X$  with  $u$  being its frontier.

The *cover* of  $u$ , denoted  $cov(u)$ , is the subgraph of  $T_u$  spanned on all nodes of  $T_u$  the contribution of which to  $\omega_r$  is totally included in  $u$  (this includes also nodes from  $u$ ). The *surface* of  $cov(u)$ , denoted  $sur(u)$ , consists of all nodes of  $cov(u)$  such that their direct ancestor in  $T(G)$  is not in  $cov(u)$ . Let  $s < r$  be the smallest integer such that some nodes of  $cov(u)$  are on the level  $s$  of  $T(G)$ . All nodes from  $cov(u)$  on the level  $s$  form the *level 0* of  $cov(u)$  - their set is denoted by  $cov_0(u)$ ; all nodes from  $cov(u)$  which are on level  $s+1$  of  $T(G)$  form *level 1* of  $cov(u)$  - their set is denoted by  $cov_1(u)$ , and so on up to  $i = (s-r)$  where  $(s-r)$  is called the *height* of  $cov(u)$  and denoted by  $ht(u)$ .

*Lemma 5.8.* The number of surface nodes on each level of  $cov(u)$  is bounded by  $2m^2$ . ■

*Lemma 5.9* For each node  $b$  of  $cov_0(u)$ ,  $\frac{|contr_u(b)|}{|u|} > \frac{1}{4m^3}$ . ■

*Lemma 5.10.* Let  $0 \leq l \leq ht(u)-1$  and let  $a \in cov_l(u)$ . Then  $\frac{|contr_u(a)|}{|u|} > \frac{1}{4m^{4+l}}$ .

## 6. PROOF OF THEOREM 4

In this section we provide a proof of Theorem 4. We start by introducing the following useful notion.

Let  $\Theta$  be a strongly closed subset of  $\Sigma$ ; we assume some fixed order of elements of  $\Theta$ . Let  $\pi$  be a cyclic permutation of  $\Theta$ . We say that  $h$  is  $(\Theta, \pi)$ -cyclic if the following two conditions are satisfied:



- (1) for each  $x \in \Theta$ , if  $h(x) = x_1 \cdots x_m$  where  $x_1, \dots, x_m \in \Theta$ , then  $x_{i+1} = \pi(x_i)$  for each  $i \leq i \leq m-1$ ,
- (2) for each  $x, y \in \Theta$ , if  $\pi(x) = y$  then  $\pi(\text{last}(h(x))) = \text{first}(h(y))$ .

*Lemma 6.11.* If  $h$  is  $(\Theta, \pi)$ -cyclic, then for every  $x \in \Theta$  and every  $n \in \mathbf{N}^+$  there exists a  $w \in \Theta^+$  such that  $|w| = \#\Theta$  and  $w^n \in \text{sub}(h^n(x))$ . ■

*Lemma 6.12.* There exists a  $\rho \in \mathbf{N}^+$  such that, for each  $w \neq \Lambda$  and each  $n \geq \rho$ , if  $w^n \in \text{sub}(L(G))$ , then  $\text{alph}(w)$  is strongly closed and there exists a permutation  $\pi$  of  $\text{alph}(w)$  such that  $h$  is  $(\text{alph}(w), \pi)$ -cyclic. ■

Now Theorem 4 is proved as follows.

*Claim 6.11.*  $G$  is repetitive if and only if there exists a strongly closed  $\Theta \subseteq \Sigma$  and a permutation  $\pi$  of  $\Theta$  such that  $h$  is  $(\Theta, \pi)$ -cyclic. ■

Since (obviously) it is decidable whether or not there exists a strongly closed  $\Theta \subseteq \Sigma$  and a permutation  $\pi$  of  $\Theta$  such that  $h$  is  $(\Theta, \pi)$ -closed, Claim 6.11 implies (a) of Theorem 2.

Part (b) of Theorem 4 is seen as follows.

By Lemma 6.12, if  $G$  is repetitive then there exist a strongly closed  $\Theta \subseteq \Sigma$  and a permutation  $\pi$  of  $\Theta$  such that  $h$  is  $(\Theta, \pi)$ -cyclic. Then by Lemma 6.11, for each  $a \in \Theta$  and each  $n \in \mathbf{N}^+$   $(a\pi(a)\pi^2(a) \cdots \pi^{\#\Theta-1}(a))^n \in \text{sub}(L(G))$ . Consequently  $G$  is strongly repetitive.

Thus Theorem 4 holds. ■

#### ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of NSF grant MSC 79-03838. They are indebted to R. Verraedt for useful comments concerning the first draft of this paper.

## REFERENCES

- [BEM] D. R. Bean, A. Ehrenfeucht and G. F. McNulty, Avoidable patterns in strings of symbols, 1979, *Pacific Journal of Mathematics*, v. 85, no. 2, 261-293.
- [B] J. Berstel, Sur les mots sans carre definis par un morphisme, 1979, *Springer Lecture Notes in Computer Science*, v. 71, 16-25.
- [C] A. Cobham, Uniform tag sequences, *Mathematical Systems Theory*, 1972, v. 6, n. 2, 164-191.
- [ER1] A. Ehrenfeucht and G. Rozenberg, Simplifications of homomorphisms, *Information and Control*, 1978, v. 38, n. 3, 298-309.
- [ER2] A. Ehrenfeucht and G. Rozenberg, On the structure of polynomially bounded DOL systems, *Fundamenta Informaticae*, 1979, v. 2, 187-197.
- [ER3] A. Ehrenfeucht and G. Rozenberg, On the subword complexity of square-free DOL languages, 1981, *Theoretical Computer Science*, v. 16, 25-32.
- [H] M. Harrison, *Introduction to formal language theory*, 1978, Addison-Wesley, Reading, Massachusetts.
- [K] J. Karhumaki, On cubic-free  $\omega$ -words generated by binary morphisms, *Discrete Applied Mathematics*, to appear.
- [MH] M. Morse and G. Hedlund, Unending chess, symbolic dynamics and a problem of semigroups, 1944, *Duke Mathematical Journal*, v. 11, 1-7.
- [RS] G. Rozenberg and A. Salomaa, *The mathematical theory of L systems*, 1980, Academic-Press, London, New York.
- [S1] A. Salomaa, Morphisms on free monoids and language theory, in R. V. Book, ed., *Formal language theory, perspectives and open problems*, 1980, Academic Press, London, New York, 141-166.

- [S2] A. Salomaa, *Jewels of formal language theory*, 1981, Computer Science Press.
- [T] A. Thue, Über unendliche Zeichenreihen, 1906, *Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiania*, v. 7, 1-22.