Theoretical Computer Science 1 (1975) 59–75. © North-Holland Publishing Company

SUBWORD COMPLEXITIES OF VARIOUS CLASSES OF DETERMINISTIC DEVELOPMENTAL LANGUAGES WITHOUT INTERACTIONS

A. EHRENFEUCHT

Department of Computer Science, University of Colorado, Boulder, Colo. 80302, *USA*

K. P. LEE[1]

Department of Computer Science, State University of New York, Amherst, N.Y. 14226, *USA*

and

G. ROZENBERG

Institute of Mathematics, University of Utrecht, Utrecht, The Netherlands

Communicated by A. Salomaa
Received 2 April 1974
Revised 30 July 1974

1. Introduction

This paper investigates the role of the deterministic restriction in the so called *developmental systems without interactions*, abbreviated as 0L systems. (They were introduced in [8] and further studied in, for example, [9, 14].)

A 0L system is a triple $G = \langle \Sigma, \delta, \omega \rangle$, where Σ is a finite alphabet, δ is a finite substitution from Σ into 2^{Σ^*} and ω is a nonempty word over Σ. The language of G consists of all strings over Σ (ω included) which can be obtained from ω by the iterative application of δ. G is called a *deterministic* 0L *system* (abbreviated as a D0L *system*) if δ is such that it maps each element of Σ into a singleton.

The study of D0L systems and the languages they generate (called D0L *languages*) constitutes one of the very well motivated and very vigorously pursued research topics in developmental systems theory. (See, e.g., [2, 6, 10, 11, 12, 13, 15, 16, 17].)

One of the possible ways of investigating the role of the deterministic restriction in D0L systems is to search for a general property which is intrinsic for D0L languages. Such a research was initiated in [3]. The property introduced there is the number of subwords of a given length which occur in a language. A corollary of the main

[1] This paper is based on part of this author's Ph.D. Thesis.

result of [3] is that if L is a D0L language over an alphabet containing at least two letters, then the ratio, referred to as r in the following, of the number of different subwords of a given length k occurring in the words of L to the number of all possible words of length k tends to zero as k increases. We believe that this result is a fundamental one for characterizing D0L languages (for a discussion see [3]).

This paper explores further the subword point of view on the deterministic restriction in 0L systems. In particular we demonstrate that this way of looking at D0L systems and languages possesses one very pleasant and desirable feature. It is "very sensitive" to various structural changes imposed on the class of D0L systems. In fact we will be able to classify a number of subclasses of the class of D0L systems according to their subword generating ability.

First of all we show that an arbitrary D0L system cannot generate more than $C \cdot k^2$, for some constant C, subwords of length k. Hence the ratio r tends to 0 like k^2/n^k (where n is the number of letters in the alphabet). We also show that $C \cdot k^2$ is the best bound.

Then we investigate how the bound is changed by adding various restrictions on D0L systems. It should be noted that these restrictions are not introduced for the purpose of this paper, but have been studied earlier in the theory of developmental systems.

If a D0L system $G = \langle \Sigma, \delta, \omega \rangle$ is such that δ maps each letter of Σ into a word in Σ^* of length at least 2, then it is called an *everywhere growing* D0L system (abbreviated as a GD0L system). Such systems were studied in [4], where they were called growing systems. For GD0L systems, we show that no more than $C \cdot k \cdot \log k$, where C is a constant, subwords of length k can be generated. Hence the ratio r tends to 0 like $(k \cdot \log k)/n^k$. Again it is shown that $C \cdot k \cdot \log k$ is the best bound.

If a GD0L system $G = \langle \Sigma, \delta, \omega \rangle$ is such that δ maps all letters of Σ into words in Σ^* of the same length, say t, where $t \geqslant 2$, then it is called a *uniform* D0L system (abbreviated as a UD0L system). Such systems were studied in [1, 7]. For UD0L systems we show that no more than $C \cdot k$, where C is a constant, subwords of length k can be generated. Hence the ratio r tends to 0 as k/n^k. Also shown is that $C \cdot k$ is the best bound.

Since the theory of developmental systems and languages is still a young branch of formal language theory, there are few useful techniques for attacking problems within the theory. This paper presents some techniques for counting (characterizing) subwords, which hopefully will be useful in future investigations.

2. Notations and definitions

Throughout this paper N will denote the set of natural numbers and $\mathrm{N}^+ = \mathrm{N} - \{0\}$. If x is a real number, then $\log x$ denotes $\log_2 x$ and $\lceil x \rceil$ denotes the smallest natural number greater than or equal to x. If A is a finite set, then $\# A$ denotes the cardinality of A.

Let Σ be a finite non-empty set (called an *alphabet*). Every finite sequence of elements in Σ (possibly with repetitions) is called a *word* or a *string over* Σ. The empty sequence (word) is denoted by Λ. The set of all nonempty words over Σ is denoted by Σ^+ and $\Sigma^+ = \Sigma^* \cup \{\Lambda\}$. $|x|$ denotes the length of a string. If $x, y \in \Sigma^*$ and there exist words $\xi_1, \xi_2 \in \Sigma^*$ such that $x = \xi_1 y \xi_2$, then y is called a *subword* of x. $\mathrm{Sub}_k(x)$ denotes the set of all subwords of x of length k. $\mathrm{Sub}(x) = \bigcup_{k \in N} \mathrm{Sub}_k(x)$. If $l \in \mathrm{N}^+$, $x \in \Sigma^+$, $x = a_1 \dots a_m$ with $a_i \in \Sigma$ for $1 \leqslant i \leqslant m$, then the *prefix* of length l of x (denoted as $\mathrm{Pref}_l(x)$) is defined as

$$\mathrm{Pref}_l(x) = \begin{cases} x & \text{if } l \geqslant m, \\ a_1 \dots a_l & \text{if } l < m. \end{cases}$$

Similarly, the *suffix* of length l of x (denoted as $\mathrm{Suf}_l(x)$) is

$$\mathrm{Suf}_l(x) = \begin{cases} x & \text{if } l \geqslant m, \\ a_{m-(l-1)} \dots a_m & \text{if } l < m. \end{cases}$$

If $L \subseteq \Sigma^*$, then L is called a *language (over* Σ*)*. $\mathrm{Sub}_k(L) = \bigcup_{x \in L} \mathrm{Sub}_k(x)$ and $\mathrm{Sub}(L) = \bigcup_{k \in N} \mathrm{Sub}_k(L)$. $\pi_k(L)$ denotes the number of elements of $\mathrm{Sub}_k(L)$.

If $s = \omega_1, \omega_2, \dots$ is a sequence of words over Σ, then $U(s) = \bigcup_{i \in N} \{\omega_i\}$ and $U(s)$ is called the *language generated by s*. If s is an infinite sequence but $U(s)$ is finite, then s is called *singly infinite*. Otherwise s is called *doubly infinite*.

Definition 1. A *deterministic L-system without interactions* (abbreviated as a D0L system) is an ordered triple $G = \langle \Sigma, P, \omega \rangle$ such that

(i) Σ is a finite non-empty set, called the *alphabet of G*;

(ii) P is a finite non-empty binary relation, $P \subset \Sigma \times \Sigma^*$, such that for every $a \in \Sigma$, there exists exactly one $\alpha \in \Sigma^*$ such that $\langle a, \alpha \rangle \in P$. P is called the *set of productions of G*. (If $\langle a, \alpha \rangle \in P$, then we write $a \to \alpha \in P$).

(iii) $\omega \in \Sigma^+$, called the *axiom of G*.

Remark. The set of productions P of a D0L system $G = \langle \Sigma, P, \omega \rangle$ may be regarded as a homomorphism from Σ into Σ^*. In such a case we shall write $G = \langle \Sigma, \delta, \omega \rangle$ where $\delta : \Sigma \to \Sigma^*$ is a homomorphism such that $\delta(a) = \alpha$ if and only if $a \to \alpha \in P$.

Definition 2. Let $G = \langle \Sigma, P, \omega \rangle$ be a D0L system. Let $x \in \Sigma^+$, $y \in \Sigma^*$. We say that *x directly derives y in G* (denoted as $x \underset{G}{\Rightarrow} y$) if $x = a_1 \dots a_n, y = \alpha_1 \dots \alpha_n$ for some $n \geqslant 1, a_1, \dots, a_n \in \Sigma$, $\alpha_1, \dots, \alpha_n \in \Sigma^*$ such that $a_i \to \alpha_i \in P$ for $1 \leqslant i \leqslant n$. Thus $\underset{G}{\Rightarrow}$ is a binary relation on $\Sigma^+ \times \Sigma^*$. $\overset{+}{\underset{G}{\Rightarrow}}$ and $\overset{*}{\underset{G}{\Rightarrow}}$ denote the transitive and reflexive-transitive closure of $\underset{G}{\Rightarrow}$, respectively. If $x \overset{*}{\underset{G}{\Rightarrow}} y$, then we say that *x derives y in G*. If $x \underset{G}{\Rightarrow} x_1, x_1 \underset{G}{\Rightarrow} x_2, \dots, x_{k-1} \underset{G}{\Rightarrow} y$ for some $k \geqslant 1, x, x_1, \dots, y \in \Sigma^*$, then we say that *x derives y in G in k steps* and denote it as $x \overset{k}{\underset{G}{\Rightarrow}} y$. Whenever it does not lead to confusion we shall write $x \overset{+}{\Rightarrow} y$, $x \overset{k}{\Rightarrow} y$, $x \overset{*}{\Rightarrow} y$, $x \Rightarrow y$ for $x \underset{G}{\overset{+}{\Rightarrow}} y$, $x \underset{G}{\overset{k}{\Rightarrow}} y$, $x \underset{G}{\overset{*}{\Rightarrow}} y$ and $x \underset{G}{\Rightarrow} y$, respectively.

Definition 3. Let $G = \langle \Sigma, P, \omega \rangle$ be a D0L system.

(i) A letter $a \in \Sigma$ is called *non-propagating* if there exists a $k \geq 1$ such that $a \overset{k}{\Rightarrow} \Lambda$.

(ii) A letter $a \in \Sigma$ is called *propagating* if it is not non-propagating.

(iii) $\hat{\Sigma}(G)$ denotes the set of all propagating letters in G.

(iv) A letter $a \in \Sigma$ is called *growing* if for every $n \in \mathrm{N}^+$, there exists $x \in \Sigma^*$ such that $a \overset{*}{\Rightarrow} x$ and $|x| \geq n$.

(v) The *sequence generated by G* (denoted as $\mathcal{E}(G)$) is a sequence of strings $\omega_0, \omega_1, \ldots$ such that $\omega_0 = \omega$ and $\omega_i \underset{G}{\Rightarrow} \omega_{i+1}$ for every $i \geq 0$.

(vi) The *language generated by G* (denoted as $\mathcal{L}(G)$) is defined as $U(\mathcal{E}(G))$ or, equivalently, $\mathcal{L}(G) = \{x \in \Sigma^* | \omega \underset{G}{\overset{*}{\Rightarrow}} x\}$.

Definition 4. Let $G = \langle \Sigma, P, \omega \rangle$ be a D0L system.

(i) G is called a *propagating* D0L system (abbreviated PD0L system) if for every α such that $a \to \alpha \in P$, $\alpha \neq \Lambda$.

(ii) G is called an *everywhere growing* D0L system (abbreviated GD0L system) if for every α such that $a \to \alpha \in P$, $|\alpha| \geq 2$.

(iii) G is called a *uniform* D0L system (abbreviated UD0L system) if there exists $t \geq 2$ such that for every α such that $a \to \alpha \in P$, $|\alpha| = t$.

Definition 5. Let Σ be an alphabet. Let $X \in \{\text{D0L, PD0L, GD0L, UD0L}\}$.

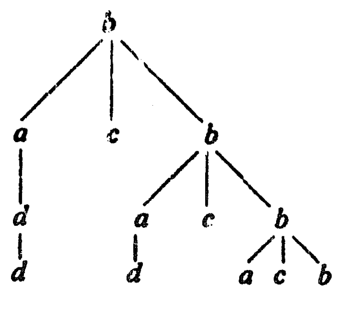
(i) A sequence s of words over Σ is called an *X sequence* if and only if there exists an X system G such that $\mathcal{E}(G) = s$.

(ii) A subset L of Σ^* is called an *X language* if and only if there exists an X system G such that $\mathcal{L}(G) = L$.

A number of other notions are best explained by an example. As usual in formal language theory it is convenient to use *derivation trees* or *forests*. For example, let

$$G = \langle \{a, b, c, d\}, \{a \to d, b \to acb, c \to \Lambda, d \to d\}, b \rangle.$$

Then the *derivation* $b \underset{G}{\Rightarrow} acb \underset{G}{\Rightarrow} dacb \underset{G}{\Rightarrow} ddacb$ will be represented by the following self-explanatory derivation tree:



We say that, for example, the first occurrence of d in the fourth string is a *direct descendant* of the occurrence of d in the third string. Conversely, the occurrence of a in the third string is the *direct ancestor* of the second occurrence of d in the fourth string. Thus the occurrence of b in the second string is an *ancestor* of the second occurrence of d in the fourth string, and the occurrence of a in the third string is a *descendant* of the occurrence of b in the first string.

Definition 6. Let $G = \langle \Sigma, P, \omega \rangle$ be a DOL system. Let V denote the set of all letters which occur in $\mathcal{L}(G)$. The *reduced version* of G is a DOL system $H = \langle V, R, \omega \rangle$, where $R = P \cap (V \times V^*)$.

Definition 7. Let $G = \langle \Sigma, P, \omega \rangle$ be a DOL system such that ω contains at least one propagating letter. A *spine version* of G is a DOL system H such that H is the reduced version of the DOL system $H_1 = \langle V, R, \sigma \rangle$, where:

(i) $V = \Sigma \cup \{\bar{a}: a \in \Sigma\}$ (elements of $\{\bar{a}: a \in \Sigma\}$ are called *spine letters*);

(ii) $\sigma = \alpha_1 \bar{a} \alpha_2$ for some $\alpha_1, \alpha_2 \in \Sigma^*$, $a \in \hat{\Sigma}(G)$ such that $\omega = \alpha_1 a \alpha_2$;

(iii) $R = P \cup \{\bar{a} \to \beta_1 \bar{b} \beta_2$ for some $\beta_1, \beta_2 \in \Sigma^*, b \in \hat{\Sigma}(G)$ such that $a \to \beta_1 b \beta_2 \in P\}$.

Note that as H_1 is a DOL system, for each $\bar{a}$, we may have one production only.

The motivation for considering spine versions of DOL systems can be found in [12]. Here we shall only give an example.

Example 1. Let G be the DOL system defined as

$$G = \langle \{a, b, c\}, \{a \to ab, b \to abc, c \to \Lambda\}, ac \rangle.$$

The following DOL system H is a spine version of G:

$$H = \langle \{a, b, c, \bar{a}, \bar{b}\}, \{a \to ab, b \to abc, c \to \Lambda, \bar{a} \to a\bar{b}, \bar{b} \to \bar{a}bc\}, \bar{a}c \rangle.$$

The beginning of the derivation in H looks as follows:

$$\begin{array}{c} \bar{a}\,c \\ a\,\bar{b} \\ a\,b\,\bar{a}\,b\,c \\ a\,b\,a\,b\,c\,a\,\bar{b}\,a\,b\,c \\ a\,b\,a\,b\,c\,a\,b\,a\,b\,c\,a\,b\,\bar{a}\,b\,c\,a\,b\,a\,b\,c. \end{array}$$

Definition 8. Let $G = \langle \Sigma, P, \omega \rangle$ be a DOL system such that $\mathcal{E}(G) = \omega_0, \omega_1, \ldots$ is infinite and let m be a positive integer. The *m-decomposition of* G is the set $\{H_0, \ldots, H_{m-1}\}$ where, for $0 \leqslant i \leqslant m-1$, H_i is the reduced version of the DOL system $\langle \Sigma, R, \omega_i \rangle$ and R consists of the following productions: if $a \overset{k}{\underset{G}{\Rightarrow}} \Lambda$ for some $k < m$, then $a \to \Lambda$ is in R; otherwise $a \to \alpha$ is in R where $a \overset{m}{\underset{G}{\Rightarrow}} \alpha$. A DOL system which belongs to the m-decomposition of G is called a *component system* (*of* G).

3. Periodicities in D0L sequences

In [12] it is proved that in a D0L sequence the sequence of prefixes (suffixes) of length k, for any $k \geqslant 1$, is ultimately periodic with a period which is independent of k. In this section we shall prove two results which strengthen the above theorem.

Our first result says that the point from which the sequence of prefixes (suffixes) of length k becomes periodic is at a distance which is bounded by a constant multiple of k from the beginning of the sequence.

Remark. Throughout this section, lemmas and theorems will be stated for prefixes and suffixes but proofs will be for prefixes only. In all cases the proofs for suffixes will be similar and hence will be omitted.

First we consider PD0L sequences.

Lemma 1.[1] *Let $\omega_0, \omega_1, \ldots$ be a doubly infinite PD0L sequence. There exist constants f and C such that for every $k \in \mathrm{N}^+$ there exists N_k such that $N_k \leqslant C \cdot k$ and for every $j \geqslant N_k$ and $m \in \mathrm{N}$,*

$$\mathrm{Pref}_k(\omega_j) = \mathrm{Pref}_k(\omega_{j+mf})$$

and

$$\mathrm{Suf}_k(\omega_j) = \mathrm{Suf}_k(\omega_{j+mf}).$$

Proof. Let $s = \omega_0, \omega_1, \ldots$ be an arbitrary doubly infinite PD0L sequence. Let $G = \langle \Sigma, P, \omega \rangle$ be a PD0L system such that $\mathcal{E}(G) = s$. If $p = \# \Sigma$, put $f = p!$ and $C = 2p$. For $g \geqslant 0$, let $\omega_g = b_{g,1} b_{g,2} \ldots b_{g,v_g}$ for some $v_g \geqslant 1$, where $b_{g,i} \in \Sigma$ for $1 \leqslant i \leqslant v_g$.

We shall prove the lemma by induction on k.

(1) $k = 1$.

Obviously for some i_1 and i_2 such that $0 \leqslant i_1 < i_2 \leqslant p$, $b_{i_1,1} = b_{i_2,1}$ and $b_{i_1,1}$ is an ancestor of $b_{i_2,1}$. Let $N_1 = i_1 \leqslant C \cdot 1$. Then since $i_2 - i_1$ divides f, $b_{i_1,1} = b_{(i_1+mf),1}$ for every $m \in \mathrm{N}$. Since s is a PD0L sequence, $b_{i_1+l,1} = b_{(i_1+l+mf),1}$ for every $l \in \mathrm{N}$ and $m \in \mathrm{N}$. Hence if $j \geqslant N_1$,

$$\mathrm{Pref}_1(\omega_j) = \mathrm{Pref}_1(\omega_{j+mf}).$$

(2) Let us assume that the lemma is true for $1, \ldots, k$.

(3) Let $g = 2pk+p$. Note that ω_g contains at least $k+1$ letters.

(a) Suppose the direct ancestor of $b_{g,k+1}$ is one of the first k letters of ω_{g-1}. Note that $g-1 = 2pk+p-1 \geqslant 2pk = C \cdot k \geqslant N_k$. So for every r such that $r \geqslant g-1$ and $m \in \mathrm{N}$,

$$\mathrm{Pref}_k(\omega_r) = \mathrm{Pref}_k(\omega_{r+mf}).$$

[1] To avoid excessively cumbersome wordings in proofs, we shall not, in the rest of this paper, distinguish very carefully between occurrences of letters and the letters themselves. However this should not lead to confusion as it will always be clear from the context what is intended.

Hence by induction hypothesis

$$b_{g,k+1} = b_{g+f,k+1}$$
$$b_{g+1,k+1} = b_{g+f+1,k+1}$$
$$\vdots$$
$$b_{g+f-1,k+1} = b_{g+2f-1,k+1}.$$

Let $N_{k+1} = g \leqslant C\cdot(k+1)$. The above equalities together with the induction hypothesis imply that

$$\mathrm{Pref}_{k+1}(\omega_j) = \mathrm{Pref}_{k+1}(\omega_{j+mf})$$

for every $j \geqslant N_{k+1}$ and $m \in \mathbf{N}$.

(b) Suppose the direct ancestor of $b_{g,k+1}$ is not one of the first k letters of ω_{g-1}. Then there are two cases.

(b.1) For some ξ such that $2pk+p < \xi \leqslant 2pk+2p$, the direct ancestor of $b_{\xi,k+1}$ is one of the first k letters of $\omega_{\xi-1}$. Now we can take $N_{k+1} = \xi \leqslant C\cdot(k+1)$. By a similar argument as in (a), we can prove the lemma for $k+1$.

(b.2) For every ξ such that $2pk+p \leqslant \xi \leqslant 2pk+2p$, $b_{\xi-1,k+1}$ is the direct ancestor of $b_{\xi,k+1}$. Then for some ξ_1 and ξ_2 such that $2pk+p \leqslant \xi_1 < \xi_2 \leqslant 2pk+2p$, $b_{\xi_1,k+1} = b_{\xi_2,k+1}$. Now we can take $N_{k+1} = \xi_1 \leqslant C\cdot(k+1)$. By a similar argument as for the case $k = 1$, we can prove the lemma for $k+1$.

Thus Lemma 1 holds.

Next we turn to arbitrary D0L sequences.

Lemma 2. *Let $\omega_0, \omega_1, \ldots$ be a doubly infinite D0L sequence. There exist constants f and C such that for every $k \in \mathbf{N}^+$ there exists N_k such that $N_k \leqslant C\cdot k$ and for every $j \geqslant N_k$ and every $m \in \mathbf{N}$,*

$$\mathrm{Pref}_k(\omega_j) = \mathrm{Pref}_k(\omega_{j+mf})$$

and

$$\mathrm{Suf}_k(\omega_j) = \mathrm{Suf}_k(\omega_{j+mf}).$$

Proof. Let $s = \omega_0, \omega_1, \ldots$ be an arbitrary doubly infinite D0L sequence. Let $G = \langle \Sigma, P, \omega\rangle$ be a D0L system such that $\mathcal{E}(G) = s$. Let $\# \Sigma = p$ and $f = p!$. We shall let, for $g \geqslant 0$, $\omega_g = b_{g,1} b_{g,2} \ldots b_{g,v_g}$ for some $v_g \geqslant 1$, where $b_{g,i} \in \Sigma$ for $1 \leqslant i \leqslant v_g$.

Define $\phi: \Sigma \to \Sigma \cup \{\Lambda\}$ as follows

$$\phi(a) = \begin{cases} a & \text{if } a \text{ is propagating in } G,\\ \Lambda & \text{otherwise.}\end{cases}$$

ϕ can be extended to strings over Σ in the usual way.

Let $\Sigma' = \{a \in \Sigma \mid a \text{ is propagating in } G\}$,

$P' = \{a \to \phi(\alpha) \mid a \in \Sigma' \text{ and } a \to \alpha \in P\}$,

$\omega' = \phi(\omega)$.

Then $G' = \langle \Sigma', P', \omega' \rangle$ is a PD0L system such that $\mathcal{E}(G') = \phi(\omega_0), \phi(\omega_1), \ldots$.

By Lemma 1, there are constants $f'(= (\# \Sigma')!)$ and $C'(= 2\cdot(\# \Sigma'))$ such that for any $k \in \mathbf{N}^+$ there exists N'_k such that $N'_k \leqslant C'\cdot k$ and for every $j \geqslant N'_k$ and $m \in \mathbf{N}$,

$$\mathrm{Pref}_k(\phi(\omega_j)) = \mathrm{Pref}_k(\phi(\omega_{j+mf'}))$$

and

$$\mathrm{Suf}_k(\phi(\omega_j)) = \mathrm{Suf}_k(\phi(\omega_{j+mf'})).$$

We know that f' divides f and so for every $j \geqslant N'_k$ and $m \in \mathbf{N}$,

$$\mathrm{Pref}_k(\phi(\omega_j)) = \mathrm{Pref}_k(\phi(\omega_{j+mf}))$$

and

$$\mathrm{Suf}_k(\phi(\omega_j)) = \mathrm{Suf}_k(\phi(\omega_{j+mf})).$$

Let $C = 3p \geqslant C'+p$. For $k \in \mathbf{N}^+$, let $N_k = N'_k+p \leqslant C\cdot k$. For any $l \in \mathbf{N}$, the ancestors of $\mathrm{Pref}_k(\omega_{N_k+l})$ in $\omega_{N'_k+l}$ must occur in $\mathrm{Pref}_k(\phi(\omega_{N'_k+l}))$. More precisely, if η_k denotes the subword of $\mathrm{Pref}_k(\phi(\omega_{N'_k+l}))$ which consists of all ancestors of $\mathrm{Pref}_k(\omega_{N_k+l})$ in $\omega_{N'_k+l}$, then

$$\eta_k = \mathrm{Pref}_t(\mathrm{Pref}_k(\phi(\omega_{N'_k+l})))$$

for some $t \leqslant k$. From the equality

$$\mathrm{Pref}_k(\phi(\omega_{N'_k+l})) = \mathrm{Pref}_k(\phi(\omega_{N'_k+l+mf})),$$

we have

$$\mathrm{Pref}_k(\omega_{N_k+l}) = \mathrm{Pref}_k(\omega_{N_k+l+mf}).$$

It follows that for any $j \geqslant N_k$ and any $m \in \mathbf{N}$,

$$\mathrm{Pref}_k(\omega_j) = \mathrm{Pref}_k(\omega_{j+mf}).$$

Thus Lemma 2 holds.

Now we shall state our result in a more general form.

Theorem 1. *Let G be a D0L system and H be a spine version of G. Let $\mathcal{E}_1(H)$ be a sequence of prefixes of consecutive elements of $\mathcal{E}(H)$ such that each element in $\mathcal{E}_1(H)$ ends with a spine letter. Let $\mathcal{E}_2(H)$ be a sequence of suffixes of consecutive elements of $\mathcal{E}(H)$ such that each element in $\mathcal{E}_2(H)$ starts with a spine letter. For $i \in \{1, 2\}$, if $\mathcal{E}_i(H)$ is doubly infinite, then there exist positive integers f and C such that for every $k \in \mathbf{N}^+$ there exists N_k such that $N_k \leqslant C\cdot k$ and for every $j \geqslant N_k$ and $m \in \mathbf{N}$,*

$$\mathrm{Pref}_k(\omega_j) = \mathrm{Pref}_k(\omega_{j+mf})$$

and

$$\mathrm{Suf}_k(\omega_j) = \mathrm{Suf}_k(\omega_{j+mf}),$$

where $\mathcal{E}_i(H) = \omega_0, \omega_1, \ldots$.

Proof. (Outline). We shall prove the theorem for $i = 1$, leaving the almost identical case of $i = 2$ to the reader.

Let G, $H = \langle \Sigma, P, \omega \rangle$, $\mathcal{E}_1(H)$ satisfy the conditions of the theorem. Let H_1 be the reduced version of a D0L system $H_2 = \langle \Sigma, R, \sigma \rangle$, where:

(i) $\sigma = \alpha_1 \bar{a}$ for $\alpha_1 \in \Sigma^*$, $\bar{a} \in \Sigma$ such that $\bar{a}$ is a spine letter and $\omega = \alpha_1 \bar{a} \alpha_2$ for some $\alpha_2 \in \Sigma^*$.

(ii) R consists of the following productions:

(a) if $a \to \alpha \in P$ and a is not a spine letter, then $a \to \alpha \in R$.

(b) if $a \to \alpha \in P$ and $\bar{a}$ is a spine letter, then $\bar{a} \to \beta_1 \bar{b} \in R$, where $\alpha = \beta_1 \bar{b} \beta_2$ for some $\beta_1, \beta_2 \in \Sigma^*$ and $\bar{b}$ is a spine letter.

From the construction of H_1, it is obvious that $\mathcal{E}(H_1) = \mathcal{E}_1(H)$ and so the theorem follows from Lemma 2.

Our second result shows that for a GD0L sequence, Theorem 1 can be improved to read that the point from which the sequence of prefixes (suffixes) of length k becomes periodic is at a distance which is bounded by a constant multiple of $\log k$ from the beginning of the sequence.

Lemma 3. *Let $\omega_0, \omega_1, \ldots$ be a GD0L sequence. There exist constants f and C such that for every $k \geqslant 2$ there exists N_k such that $N_k \leqslant C \cdot \log k$ and for every $j \geqslant N_k$ and $m \in \mathbb{N}$,*

$$\mathrm{Pref}_k(\omega_j) = \mathrm{Pref}_k(\omega_{j+mf})$$

and

$$\mathrm{Suf}_k(\omega_j) = \mathrm{Suf}_k(\omega_{j+mf}).$$

Proof. Let $s = \omega_0, \omega_1, \ldots$ be an arbitrary GD0L sequence. Let $G = \langle \Sigma, P, \omega \rangle$ be a GD0L system such that $\mathcal{E}(G) = s$. If $p = \# \Sigma$, put $f = p!$ and $C = p+1$. For $g \geqslant 0$, let $\omega_g = b_{g,1} b_{g,2} \ldots b_{g,v_g}$ for some $v_g \geqslant 1$, where $b_{g,i} \in \Sigma$ for $1 \leqslant i \leqslant v_g$. For $k \geqslant 2$, we define $N_k = p-1+\lceil \log k \rceil$. Note that $N_k = p-1+\lceil \log k \rceil \leqslant p+\log k \leqslant C \cdot \log k$.

We shall now show by induction on k that N_k has the required property.

(1) $k = 2$.

Obviously for some i_1 and i_2 such that $0 \leqslant i_1 < i_2 \leqslant p$, $b_{i_1,1} = b_{i_2,1}$ and $b_{i_1,1}$ is an ancestor of $b_{i_2,1}$. Since $i_2 - i_1$ divides f, $b_{i_1,1} = b_{(i_1+mf),1}$ for every $m \in \mathbb{N}$. From this and from the fact that G is a GD0L system, it follows easily that for $j \geqslant N_2 = p$ and $m \in \mathbb{N}$,

$$\mathrm{Pref}_2(\omega_j) = \mathrm{Pref}_2(\omega_{j+mf}).$$

(2) Let us assume that the lemma is true for $2, \ldots, k$.

(3) Consider now $k+1$. If $k+1$ is of the form 2^r+1, for some r, then $N_{k+1} = N_k + 1$; while if $k+1$ is not of the form 2^r+1, then $N_{k+1} = N_k$.

In either case, $\omega_{N_{k+1}}$ has at least $k+1$ letters. Furthermore the direct ancestor of $b_{N_{k+1},k+1}$ is $b_{N_{k+1}-1,t}$, for some $t \leqslant \frac{1}{2}(k+1)$. But it is easy to see that $N_{k+1} \geqslant N_t$. This, together with the induction assumption, shows that for $j \geqslant N_{k+1}$ and $m \in \mathbb{N}$,

$$\mathrm{Pref}_{k+1}(\omega_j) = \mathrm{Pref}_{k+1}(\omega_{j+mf}).$$

Thus Lemma 3 holds.

4. The bound for the number of subwords of a given length in an arbitrary D0L language

In sections 4 and 5 we show that the number of possible subwords of length k that can occur in a D0L language (respectively a GD0L language, a UD0L language) is bounded by a constant multiple of k^2 (respectively $k \cdot \log k$, k). Furthermore it is shown that in each case the bound obtained is the best possible.

First we need a definition.

Definition 9. Let $G = \langle \Sigma, \delta, \omega \rangle$ be a D0L system and $u, v \in \Sigma^*$, u is said to *cover* v (*in* G) if $u \overset{*}{\underset{G}{\Rightarrow}} w$ and v is a subword of w.

Remark. All the proofs in this section will be presented rather informally. However, they can be easily formalized (e.g. using the formalism in [5]). We feel that presenting the proofs of the theorems in this section in a formal way would result in losing the intuition in formal details. In addition, the proofs themselves will be considerably lengthened.

Our first result concerns D0L languages.

Theorem 2. *Let L be a D0L language. There exists a constant C such that for every $k \in \mathbb{N}^+$, $\pi_k(L) \leqslant C \cdot k^2$.*

Proof. Let L be a D0L language and $G = \langle \Sigma, \delta, \omega \rangle$ be a D0L system such that $L = \mathcal{L}(G)$. Since the result is obvious if L is finite, we shall assume in the sequel that L is infinite. Denote $\# \Sigma$ by p. Let $\mathcal{D} = \{H_1, H_2, ..., H_p\}$ be the p-decomposition of G.

Consider a component system $H = \langle \Sigma, \delta', \omega' \rangle$. By adding an extra symbol if necessary, we may assume without loss of generality that ω' is a single letter. Let $\mathcal{E}(H) = \omega_0, \omega_1, ...$ For notational convenience, we shall denote δ' and ω' by δ and ω respectively. This will not lead to confusion since we shall be dealing exclusively with component systems. First we note that since H is a component system in the p-decomposition of G, the following hold:

(1) If a is a growing letter in G, $|\delta^i(a)| \geqslant i+1$ for every $i \in \mathbb{N}$.

(2) If a is a non-propagating letter in G, $|\delta(a)| = 0$.

Next we make use of the known fact (see [12]) that there exists a constant K such that, for any $u \in \mathrm{Sub}(L)$, if $|u| \geqslant K$, then u contains at least one occurrence of a propagating letter in G. Also we shall let R be the maximum length of the right hand side of any production of H. Consider now any k such that $k > \max\{K, R\}$ and let $u \in \mathrm{Sub}_k(L)$ be such that u occurs as a subword in $\mathcal{E}(H)$. We may assume that u is a subword of ω_r, for some $r \geqslant 1$. Now we define a sequence of subwords inductively as follows:

(i) $u_0 = u$.

(ii) Let $i \leqslant r$. u_i is the minimal subword in ω_{r-i} which covers u_{i-1}, by which we mean that u_i covers u_{i-1} and no other subword of u_i covers u_{i-1}. (See Fig. 1).

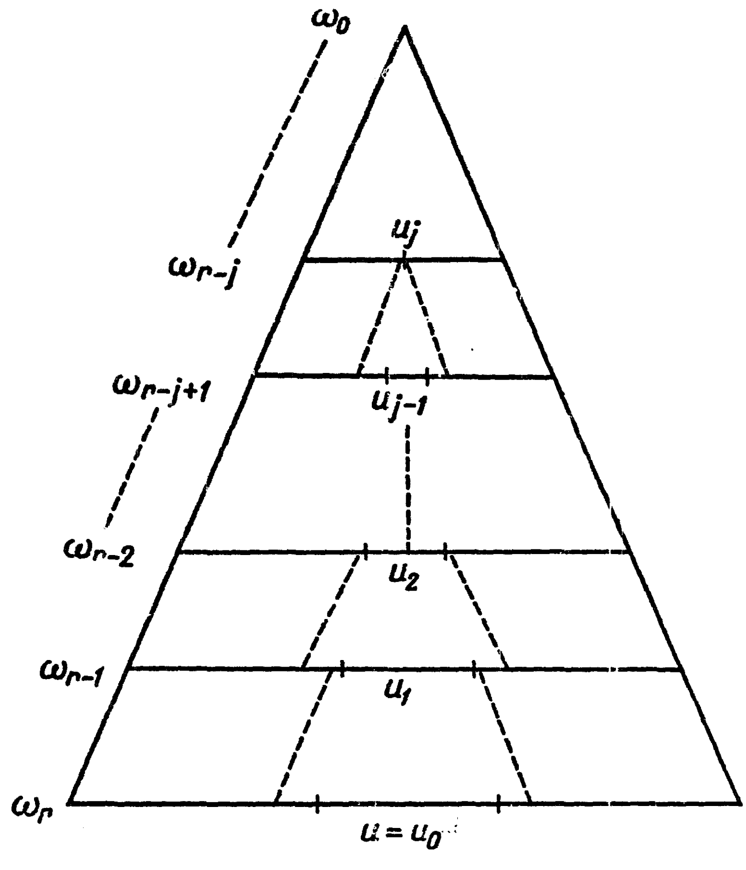


Fig. 1

Let j be the smallest number such that $u_j \in \Sigma$. By our assumption on the size of k, $j \geqslant 2$. Furthermore u_{j-1} is a subword of $\delta(u_j)$. It is clear that for each $u \in \mathrm{Sub}_k(L)$ which occurs as a subword in $B(H)$, we can obtain such a u_{j-1}.

Let B denote the set of all words $\beta \in \Sigma^*$ such that $|\beta| > 1$ and $\beta \in \mathrm{Sub}(\delta(a))$ for some $a \in \Sigma$. For each $\beta \in B$, we define a set $L^k_{\beta,H}$ as follows. $u \in L^k_{\beta,H}$ if and only if there is an occurrence of $u \in \mathrm{Sub}_k(L)$ in $\mathcal{E}(H)$ such that by the procedure described above the u_{j-1} obtained is β. (Note that a u may belong to several $L^k_{\beta,H}$.) Now we estimate the size of $L^k_{\beta,H}$. Let $u \in L^k_{\beta,H}$. There are three cases, exhausting all possibilities.

(a) β contains no growing letter. Since there is M such that $|\delta^i(\beta)| \leqslant M$ for every $i \in \mathrm{N}$, $\bigcup_{i \in \mathrm{N}} \mathrm{Sub}(\delta^i(\beta))$ is finite and its cardinality is independent of k. Hence also the cardinality of $L^k_{\beta,H}$ cannot exceed a constant which is independent of k.

(b) $\beta = \alpha_1 a \alpha_2$ where a is a growing letter and $\alpha_1, \alpha_2 \in \Sigma^+$. In this case, since β is the minimal subword in ω_{r-j+1} covering u and both α_1 and α_2 are nonempty, $\delta^{j-1}(a)$ is a subword of u. We now make the following two observations:

(b.1) $|\delta^{j-1}(a)| \geqslant j$,

(b.2) $k = |u| > |\delta^{j-1}(a)| \geqslant j$.

It follows that all elements of $L^k_{\beta,H}$ must be derived from β in at most k steps. But for $0 < i < k$, there are at most k elements in Sub $(\delta^i(\beta))$ which have $\delta^i(a)$ as a subword. Hence $L^k_{\beta,H}$ has at most k^2 elements.

(c) $\beta = a\alpha$, where a is a growing letter and $\alpha \in \Sigma^+$. (The case where $\beta = \alpha a$ can be treated in a similar way.) Since β is the minimal subword in ω_{r-j+1} covering u and $|\beta| \geqslant 2$, α contains at least one propagating letter. Let us write $\alpha = \alpha' b$, where b is a propagating letter. Consider now the D0L systems $K_1 = \langle \Sigma, \delta, a\rangle$ and $K_2 = \langle \Sigma, \delta, b\rangle$. Let $\mathcal{E}(K_1) = \eta_0^{(1)}, \eta_1^{(1)}, \ldots$ and $\mathcal{E}(K_2) = \eta_0^{(2)}, \eta_1^{(2)}, \ldots$ It is an easy consequence of Theorem 1 that there are constants C_1 and f, independent of k, such that there exists $n \in \mathbf{N}^+$ with the following property:

(c.1) $n \leqslant C_1(k-1)$

(c.2) for every $i \geqslant n$ and $m \in \mathbf{N}$,

$$\mathrm{Suf}_{k-1}(\eta_i^{(1)}) = \mathrm{Suf}_{k-1}(\eta_{i+mf}^{(1)}),$$

and

$$\mathrm{Pref}_{k-1}(\eta_i^{(2)}) = \mathrm{Pref}_{k-1}(\eta_{i+mf}^{(2)}).$$

Remembering that if $u \in L^k_{\beta,H}$, then u must contain a descendant of a and a descendant of b, we see that for any $i \in \mathbf{N}$, there are at most k subwords of $\delta^i(\beta)$ which can be in $L^k_{\beta,H}$. This, together with the fact that after $n \leqslant C_1(k-1)$ steps, the suffixes and prefixes of length $k-1$ of $\mathcal{E}(K_1)$ and $\mathcal{E}(K_2)$ respectively become periodic, shows that there is a constant C_2 such that $\# L^k_{\beta,H} \leqslant C_2 k^2$.

Finally we note the following equality:

$$\mathrm{Sub}_k(L) = \bigcup_{\beta\in B} \bigcup_{H\in\mathcal{D}} L^k_{\beta,H} \cup F,$$

where F is a finite set, whose size is independent of k, consisting of all subwords x of length k such that either

(i) x consists of non-propagating letters only; or

(ii) $x \in \mathrm{Sub}(\delta(a))$ for some $a \in \Sigma$.

From what has been proved so far, it is clear that there is a constant C such that

$$\pi_k(L) \leqslant C\cdot k^2.$$

Hence Theorem 2 holds.

Now we shall show that the bound obtained in Theorem 2 is the best possible. To do this we need a lemma.

Lemma 4. *Let G be the following PD0L system:*

$$G = \langle \{a, b, c\}, \{a \to ac, b \to bb, c \to c\}, bab\rangle$$

and let $L = \mathcal{L}(G)$. Then for sufficiently large k, $\pi_k(L) \geqslant \frac{1}{2} k^2$.

Proof. Let $\mathcal{E}(G) = \omega_0, \omega_1, \ldots$. Then we see that $\omega_l = b^{2^l} ac^l b^{2^l}$. We shall fix our attention on the subset B_k of $\mathrm{Sub}_k(L)$ consisting of all subwords of length k of the form $b^f ac^u b^g$, for some $f, g \geqslant 0$, $u \geqslant 1$. Let us now estimate how many elements each ω_l contributes to B_k. There are two cases:

(1) If $l \geqslant \log k$, then all possible subwords of the required form with $u = l$ will be found in ω_l.

(2) If $l < \log k$, then we may not have all possible subwords of the required form because the number of b's is too small. But in any case, for each such l, no more than k of such subwords will be missed. So at most $k \log k$ of subwords of the required form will be excluded from B_k.

However, the number of subwords of length k of the form $b^f ac^u b^g$ is easily calculated to be $\frac{1}{2} k(k-1)$. So

$$\pi_k(L) \geqslant \# B_k \geqslant \frac{k(k-1)}{2} - k \log k = k^2 \left(\frac{1}{2} - \frac{1}{2k} - \frac{\log k}{k}\right)$$
$$\to \frac{k^2}{2} \text{ as } k \to \infty.$$

Hence Lemma 4 holds.

Theorem 3. *Given any constant l, there exists a PD0L language L such that $\pi_k(L) \geqslant l \cdot k^2$ for infinitely many k.*

Proof. Let $m = [2l]$. Intuitively speaking, we obtain our result by simply catenating m copies of the G in Lemma 4. Formally, let $\Sigma = \{a_i, b_i, c_i \mid 1 \leqslant i \leqslant m\}$ be an alphabet of $3m$ elements; $P = \{a_i \to a_i c_i, b_i \to b_i b_i, c_i \to c_i \mid 1 \leqslant i \leqslant m\}$; and $\omega = b_1 a_1 b_1 b_2 a_2 b_2 \ldots b_m a_m b_m$. Consider the PD0L system $H = \langle \Sigma, P, \omega \rangle$ and let $L = \mathcal{L}(H)$. For $1 \leqslant i \leqslant m$, let B_k^i denote the subset of $\mathrm{Sub}_k(L)$ consisting of all subwords of length k of the form $b_i^f a_i c_i^u b_i^g$, for some $f, g \geqslant 0$, $u \geqslant 1$. Then the B_k^i are pairwise disjoint and $\mathrm{Sub}_k(L) \supset \bigcup_{i=1}^{m} B_k^i$. Clearly it follows from Lemma 4 that $\pi_k(L) \geqslant \sum_{i=1}^{m} \# B_k^i \geqslant m \frac{k^2}{2} \geqslant lk^2$ for all suffiiciently large k.

Hence Theorem 3 holds.

5. Bounds for the number of subwords of a given length in GD0L and UD0L languages

Theorem 4. *Let L be a GD0L language. There exists a constant C such that for every $k \geqslant 2$, $\pi_k(L) \leqslant C \cdot k \cdot \log k$.*

Proof. (Outline). The proof of Theorem 4 can be carried out similarly to that of Theorem 2 with the following changes.

(I) Since we are dealing with a GD0L and hence PD0L system, there is no need to form the p-decomposition of G but we can instead let $\mathcal{E}(G) = \omega_0, \omega_1, \ldots$.

(II) Case (a) will not arise since L is a GD0L language.

(III) In (b), instead of $|\delta^{j-1}(a)| \geqslant j$, we have $|\delta^{j-1}(a)| \geqslant 2^{j-1}$. From this it follows that all elements of $L^k_{\beta,G}$ must be derived from β in at most $\lceil \log k \rceil$ steps. Hence $\# L^k_{\beta,G} \leqslant k \lceil \log k \rceil \leqslant 2k \cdot \log k$.

(IV) In (c), instead of using Theorem 1, we may use Lemma 3 to conclude that there is a constant C_2 such that $\# L^k_{\beta,G} \leqslant C_2 k \cdot \log k$. These observations show that if L is a GD0L language, the bound can be improved to read, for $k \geqslant 2$,

$$\pi_k(L) \leqslant C \cdot k \cdot \log k.$$

Hence Theorem 4 holds.

Next we shall also show that the bound obtained in Theorem 4 is the best possible. First we prove the following lemma which will be needed in proving Theorem 5.

Lemma 5. *Let G be the following GD0L system:*

$$G = \langle \{a, b\}, \{a \to a^2, b \to b^{16}\}, bab \rangle$$

and let $L = \mathcal{L}(G)$. Then for sufficiently large k, $\pi_k(L) \geqslant \frac{1}{4} k \cdot \log k$.

Proof. Let $\mathcal{E}(G) = \omega_0, \omega_1, \ldots$. Then we see that $\omega_l = b^{16^l} a^{2^l} b^{16^l}$. We shall fix our attention on the subset B_k of $\mathrm{Sub}_k(L)$ consisting of all subwords of length k of the form $b^f a^{2^u} b^g$ for some $f, g \geqslant 0$ and u such that $2^u \leqslant \frac{1}{2} k$, or equivalently, $u \leqslant \log k) - 1$. Let us now estimate how many elements each ω_l contributes to B_k. There are two cases:

(1) If $16^l \geqslant k$, then all possible subwords of the required form with $u = l$ will be found in ω_l.

(2) If $16^l < k$, i.e. $l < \frac{1}{4} \log k$, then we may not have all possible subwords of the required form because the number of b's is too small. But in any case no more than k of such subwords will be missed. So at most $\frac{1}{4} k \cdot \log k$ subwords of the required form will be excluded from B_k.

However the number of subwords of length k of the form $b^f a^{2^u} b^g$ is easily calculated to be at least $((\log k) - 1) k/2$. So

$$\pi_k(L) \geqslant \# B_k \geqslant \frac{k((\log k) - 1)}{2} - \frac{k \log k}{4}$$

$$= \frac{k \cdot \log k}{2} \left(1 - \frac{1}{\log k} - \frac{1}{2}\right)$$

$$\to \frac{k \cdot \log k}{4} \text{ as } k \to \infty.$$

Hence Lemma 5 holds.

Theorem 5. *Given any constant l, there exists a GD0L language L such that $\pi_k(L) \geqslant l \cdot k \cdot \log k$ for infinitely many k.*

Proof. (Outline) We need only remark that we can take $m = \lceil 4l \rceil$. Then Theorem 5 follows from Lemma 5 in the same way as Theorem 3 from Lemma 4.

The final result is concerned with UD0L languages.

Theorem 6. *Let L be a UD0L language. There exists a constant C such that for $k \in N^+$, $\pi_k(L) \leqslant C \cdot k$.*

Proof. Let L be a UD0L language and $G = \langle \Sigma, \delta, \omega \rangle$ be a UD0L system such that $\mathcal{L}(G) = L$. Since the result holds trivially if L is finite, we assume that L is infinite. Let $t > 1$ be a number such that if $a \in \Sigma$, then $|\delta(a)| = t$, and let $\mathcal{E}(G) = \omega_0, \omega_1, \ldots$

Let $u \in \mathrm{Sub}_k(L)$. We may assume that it occurs in ω_r for r sufficiently large such that the following argument holds. A sequence of subwords is defined inductively as follows:

(i) $u_0 = u$.

(ii) If $i \leqslant r$, then u_i is the minimal subword in ω_{r-i} covering u_{i-1}.

Now we shall obtain a bound on the length of u_i.

$$|u_1| \leqslant \frac{|u_0|+2(t-1)}{t} = \frac{k}{t} + \frac{2(t-1)}{t},$$

$$|u_2| \leqslant \frac{|u_1|+2(t-1)}{t} \leqslant \frac{\frac{k}{t} + \frac{2(t-1)}{t} + 2(t-1)}{t} = \frac{k}{t^2} + \frac{2(t-1)}{t^2} + \frac{2(t-1)}{t}.$$

In general,

$$\begin{aligned} |u_l| &\leqslant \frac{k}{t^l} + \frac{2(t-1)}{t^l} + \frac{2(t-1)}{t^{l-1}} + \ldots + \frac{2(t-1)}{t} \\ &= \frac{k}{t^l} + (2t-2)\left[\frac{1}{t^l} + \frac{1}{t^{l-1}} + \ldots + \frac{1}{t}\right] \\ &\leqslant \frac{k}{t^l} + (2t-2)\frac{1}{t-1} \\ &= \frac{k}{t^l} + 2. \end{aligned}$$

So for $l = \lceil \log_t k \rceil$,

$$|u_l| \leqslant 1+2 = 3.$$

This means that any word of length k must be derived in at most $\lceil \log_t k \rceil$ steps from a word of length 3. But the number of different subwords of length 3 in $\mathcal{E}(G)$ is finite and they can all be found among $\omega_0, \omega_1, \ldots, \omega_{l_0}$, for some l_0. So all subwords of length k can be found among $s = \omega_0, \omega_1, \ldots, \omega_m$, where $m = l_0 + \lceil \log_t k \rceil$. But the number of subwords of any length that can be found among the elements of s is bounded by the total number of occurrences of letters in s, which is

$$a + at + \ldots + at^m < at^{m+1}$$

where $|\omega_0| = a$. Hence

$$\pi_k(L) \leqslant at^{l_0 + \lceil \log_t k \rceil + 1} \leqslant at^{l_0} t^2 k = C \cdot k$$

where $C = at^{2+l_0}$. Thus Theorem 6 holds.

Finally we show that the bound obtained in Theorem 6 is the best possible. We need the following lemma.

Lemma 6. *Let G be the following UD0L system:*

$$G = \langle \{a, b\}, \{a \to a^2, b \to b^2\}, ab \rangle$$

and let $L = \mathcal{L}(G)$. *Then for all* $k \geqslant 1$, $\pi_k(L) > k$.

Proof. Let $\mathcal{E}(G) = \omega_0, \omega_1, \ldots$. Then we see that $\omega_l = a^{2^l} b^{2^l}$. It is clear that $\mathrm{Sub}_k(L) = \{a^l b^m \mid l, m \geqslant 0,\ l + m = k\}$ and so $\pi_k(L) = k + 1 > k$. Hence Lemma 6 holds.

Theorem 7. *For any constant l, there exists a UD0L language L such that for all* $k \geqslant 1$, $\pi_k(L) \geqslant l \cdot k$.

Proof. (Outline) We need only remark that we can take $m = \lceil l \rceil$. Then Theorem 7 follows from Lemma 6 in the same way as Theorem 3 follows from Lemma 4.

References

[1] J. Berstel, Une remarque sur certains langages de Lindenmayer, unpublished manuscript.
[2] P. Doucet, On the membership question in some Lindenmayer systems, Indag. Math. 34 (1972) 45–52.
[3] A. Ehrenfeucht and G. Rozenberg, A limit theorem for sets of subwords in deterministic T0L systems, Inf. Proc. Lett. 2 (1973) 70–73.
[4] G. T. Herman, K. P. Lee, J. van Leeuwen and G. Rozenberg, Characterization of unary developmental languages, Discrete Math. 6 (1973) 235–247.
[5] G. T. Herman and G. Rozenberg, Developmental Systems and Languages (North-Holland, Amsterdam, 1975).
[6] K. P. Lee and G. Rozenberg, The length sets of D0L languages are uniformly bounded, Inf. Proc. Lett. 2 (1974) 185–188.
[7] J. van Leeuwen, A note on uniform Lindenmayer-languages, unpublished manuscript.

[8] A. Lindenmayer, Mathematical models for cellular interactions in development, Parts I & II, J. Theoret. Biol. 18 (1968) 280–315.

[9] A. Lindenmayer, Developmental systems without cellular interactions, their languages and grammars, J. Theoret. Biol. 30 (1971) 455–484.

[10] A. Lindenmayer and G. Rozenberg, Developmental systems and languages, Proc. of the Fourth ACM Symposium on Theory of Computing (1972) 214–221.

[11] A. Paz and A. Salomaa, Integral sequential word functions and growth equivalence of Lindenmayer systems, Information and Control 23 (1973) 313–343.

[12] G. Rozenberg, D0L sequences, Discrete Math. 7 (1974) 323–347.

[13] G. Rozenberg, Circularities in D0L sequences, Revue Roum. de Math. Pures et Appl. 19 (9) (1974) 1131–1152.

[14] G. Rozenberg and P. Doucet, On 0L-languages, Information and Control 19 (1971) 302–318.

[15] G. Rozenberg and A. Lindenmayer, Developmental systems with locally-catenative formulas, Acta Informatica 2 (1973) 214–248.

[16] A. Salomaa, On exponential growth in Lindenmayer systems, Indag. Math. 35 (1973) 23–30.

[17] P. Vitanyi, D0L-languages and a feasible solution for a word problem, technical report No. MR 138/72, Mathematisch Centrum, Amsterdam (1972).