JUHANI KARHUMÄKI

The decidability of the equivalence problem for polynomially bounded DOL sequences

RAIRO. Informatique théorique, tome 11, nº 1 (1977), p. 17-28
<http://www.numdam.org/item?id=ITA_1977__11_1_17_0>

R.A.I.R.O. Informatique théorique/Theoretical Computer Science
(vol. 11, n° 1, 1977, p. 17 à 28)

THE DECIDABILITY OF THE EQUIVALENCE PROBLEM FOR POLYNOMIALLY BOUNDED DOL SEQUENCES (*)

by Juhani KARHUMÄKI (1)

Communicated by J. BERSTEL

Abstract. — *It is proved that the equivalence problem for polynomially bounded DOL sequences is decidable.*

1. INTRODUCTION

The problem of whether two DOL systems generate the same sequence, i.e. the DOL equivalence problem, was introduced (for propagating systems) in [7]. Since then many attempts have been made to solve the problem, and its decidability has been established in certain special cases. For instance Culik [1] and Valiant [9] have shown that the problem is decidable for so-called smooth families of DOL systems, and using this result Culik deduced the decidability for so-called simple DOL sequences and Valiant for binary DOL sequences. Moreover, it is known to Ehrenfeucht and Rozenberg that the problem is decidable for polynomially bounded DOL sequences, *see* [8].

In this paper we intend to give a proof for the result of Ehrenfeucht and Rozenberg. Our proof is rather technical and the resulting algorithm is not at all practical. On the other hand, our considerations show that the family of polynomially bounded DOL systems forms a smooth family of DOL systems. Thus, the result of Culik and Valiant gives another algorithm to check the equality of two polynomially bounded DOL sequences. Unfortunately to prove the smoothness we need almost everything presented in this paper. So the result of Culik and Valiant does not shorten our considerations.

(*) Received April 1976. Revised June 1976.
(1) Department of Mathematics, University of Turku, Finland.

Our proof is based on the following ideas. We are throughout working with equivalent polynomially bounded DOL sequences. So we obtain certain necessary conditions for the equality of two sequences, the most essential being that a certain finite number of DOL sequences with a lower growth order must be equivalent. Moreover these conditions turn out to be sufficient, too. So we may conclude inductively the existence of the algorithm.

2. PRELIMINARIES

Let $G = \langle \Sigma, \omega, \delta \rangle$ be a (reduced) DOL system, *see* [4]. G is said to be *polynomially bounded* iff there exists *a* polynomial p such that

$$|\delta^n(\omega)| \leq p(n) \qquad \text{for} \quad n \geq 0,$$

where the vertical bars denote the length of a word. Assume now that G is polynomially bounded. Then each letter a in Σ satisfies

$$|\delta^n(a)| \leq q(n), \qquad n \geq 0, \tag{1}$$

for some polynomial q. If (1) is valid for a polynomial of degree N but is not valid for any polynomial of degree $N - 1$, then a is said to be of *growth order N*. Obviously we may also talk about the growth order of a word or a DOL system. By a *linear DOL system* we mean a DOL system having the growth order 1.

Denote

$$\Sigma_i = \{ a \in \Sigma \mid a \text{ is of growth order } i \}.$$

Then it is easy to show that G has the following structural properties, *see* [2]. First, for each $i \geq 1$ and for each letter $a \in \Sigma_i$ the system $\langle \Sigma_i, a, \delta|\Sigma_i \rangle$ is λ-free and it generates a finite language. (Here the value of the homomorphism $\delta|\Sigma_i : \Sigma_i^* \longrightarrow \Sigma_i^*$ on b is obtained from $\delta(b)$ by erasing from it all letters in $\Sigma - \Sigma_i$.) So each letter in Σ_i produces in each step at least one letter of Σ_i. Secondly, each letter in Σ_i derives in a number of steps at least one letter of Σ_j for all $j < i$. These facts are utilized later several times without further mention.

Let P be any word in Σ^* and $V \subseteq \Sigma$. Then we define

$\mathrm{sw}_V(P)$ = the maximal subword of P belonging to $\{ \lambda \} \cup V \cup V\Sigma^* V$

and

$$\#_V(P) = \sum_{a \in V} \#_a(P),$$

where $\#_a(P)$ denotes the number of a's in P. Moreover, by min (P) we mean a set of letters occurring in P.

3. THE RESULT

In this section we prove

THEOREM 1 : *The equivalence problem for polynomially bounded DOL sequences is decidable.*

To prove the Theorem we state an induction hypothesis : The equivalence problem for polynomially bounded DOL sequences with a growth order less than N, where $N \geq 2$, is decidable.

Assume now that $G = \langle \Sigma, \omega, \delta \rangle$ and $G' = \langle \Sigma, \omega, \delta' \rangle$ are two equivalent polynomially bounded DOL systems with growth order N, where $N \geq 2$. Let Σ_i, for $i = 0, \ldots, N$, be as in the previous section and further denote

$$\Sigma_i^c = \{ a \in \Sigma_i \mid \#_{\Sigma_i}(\delta^k(a)) = 1 \quad \text{for all } k, \text{ and } a \in \min(\delta^k(a))$$

for some $k \geq 1$ },

$$\Sigma_i^n = \Sigma_i - \Sigma_i^c,$$

$$\Sigma^c = \bigcup_{i=0}^{N} \Sigma_i^c, \qquad \Sigma^n = \bigcup_{i=0}^{N} \Sigma_i^n,$$

$$V_0 = \{ a \in \Sigma \mid (\#_a(\delta^k(\omega)))_{k \geqslant 0} \text{ is bounded} \},$$

$$V_1 = \{ a \in \Sigma - V_0 \mid (\tfrac{1}{k} \#_a(\delta^k(\omega)))_{k \geqslant 0} \text{ is bounded} \},$$

and similarly $\Sigma_i', \ldots, V_1'$ for G'. Let us call letters in Σ^c, Σ^n, V_0 and V_1 *cyclic, noncyclic, finite* and *linear letters* with respect to G. Obviously the equivalence of G and G' implies

$$(i) \qquad V_0 = V_0' \quad \text{and} \quad V_1 = V_1'.$$

Now we characterize V_0 in terms of so-called maximal letters. A letter a in Σ_i is called *maximal* (or V_0-*maximal*) with respect to G iff it is noncyclic and not derived from any letter in Σ^c or it is cyclic and not derived from any letter in $\bigcup_{j=i+1}^{N} \Sigma_j^c$. Let W_0 be the set of V_0-*maximal* letters. Then obviously W_0 contains all letters from Σ_N and, in addition, it may contain some other letters, too. Moreover, we have

$$W_0 \subseteq V_0,$$

or more precisely

$$V_0 = W_0 \cup \{ a \in \Sigma^n \mid \text{there exists a derivation path from a letter } b \text{ in } W_0 \text{ to } a,$$
$$\text{while there exists no such path containing a letter in } \Sigma^c - W_0 \}.$$

Let P be any word in $\Sigma^*(W_0 \cap \Sigma^c)\Sigma^*$. We say that P is *bounded* (or *V_0-bounded*) with respect to G iff the sequence

$$\Delta(P) = (\Delta_n(P))_{n \geq 0} = (sw_{V_0}(\delta^n(P)))_{n \geq 0}$$

is ultimately periodic (i. e. the corresponding length sequence is bounded). Otherwise P is said to be *unbounded* (or *V_0-unbounded*).

It should be clear, by the characterization of V_0, that there exists an n_0 such that, for all $n \geq n_0$, we can write

$$\delta^n(\omega) = \alpha_0 P_1 \alpha_1 \ldots \alpha_{s-1} P_s \alpha_s, \tag{2}$$

where $\alpha_0 \alpha_1 \ldots \alpha_s \notin \Sigma^* V_0 \Sigma^*$, $P_i's$ are V_0-bounded and the words $P_i \alpha_i P_{i+1}$ are V_0-unbounded. Note that here s is independent of n because every V_0-bounded (resp. V_0-unbounded) word derives according to G a V_0-bounded (resp. V_0-unbounded) word. Furthermore, the representation (2) is unique if n_0 is great enough (so that $\alpha_1, \ldots, \alpha_{s-1}$ can be chosen "long") and all $P_i's$ are chosen to be cyclic in the sense that $sw_{V_0}(\delta^k(P_i)) = P_i$ for some $k \geq 1$. Using this unique representation we may describe the direct derivation, for $n \geq n_0$, as follows

$$\begin{array}{ccc} \delta^n(\omega) & = & \alpha_0 P_1 \alpha_1 \ldots \alpha_{s-1} P_s \alpha_s \\ \Big\downarrow & & \Big\downarrow \\ \delta^{n+1}(\omega) & = & \bar{\alpha}_0 \bar{P}_1 \bar{\alpha}_1 \ldots \bar{\alpha}_{s-1} \bar{P}_s \bar{\alpha}_s \end{array} \tag{3}$$

Observe now that G and G' are equivalent not only in Σ but also in V_0. So the assumption that $\alpha_1, \ldots, \alpha_{s-1}$ are "long" implies

$$sw_{V_0}(\delta'(P_i)) = \bar{P}_i$$

showing that the representations (2) and (3) are valid both for G and for G'.

Now, clearly, the equivalence of G and G' implies (and is implied together with (i) by)

(ii) $\qquad \delta^n(\omega) = \delta'^n(\omega) \qquad$ for $\quad n \leq n_0$,

(iii) $\qquad \Delta(P_i \alpha_i P_{i+1}) = \Delta'(P_i \alpha_i P_{i+1}) \qquad$ for $\quad i = 1, \ldots, s-1$,

(iv) $\qquad \Delta^p(\alpha_0 P_1) = \Delta'^p(\alpha_0 P_1) \qquad$ and $\qquad \Delta^s(P_s \alpha_s) = \Delta'^s(P_s \alpha_s)$,

where $\delta^{n_0}(\omega) = \alpha_0 P_1 \alpha_1 \ldots \alpha_{s-1} P_s \alpha_s$ and Δ^p (resp. Δ^s) is defined as Δ with the difference that "maximal subword" in the definition of sw_V is replaced by "maximal prefix" (resp. "maximal suffix").

From this on we consider the condition (iii) for a fixed (i). The reader will have no difficulties in converting our considerations to cover also the sequences in (iv). Further to simplify notations we denote $P_i = P$, $P_{i+1} = Q$ and $\alpha_i = \alpha$. Observe that the sequences $\Delta(P\alpha Q)$ and $\Delta'(P\alpha Q)$ are "almost DOL sequences", and it is easy to convert them to DOL sequences without affecting the equality of the original sequences. (This is done by assuming that $P's$ and $Q's$ are in disjoint alphabets and by introducing, if necessary, some new copies of certain letters). So assume that $\Delta(P\alpha Q)$ and $\Delta'(P\alpha Q)$ are DOL sequences generated by G and G', respectively. If these sequences are of a growth order less than N then their equality can be solved by the induction hypothesis. So let us assume that they are of growth order N.

We continue by characterizing V_1, the set of linear letters. For this purpose we need a notion of a V_1-maximal letter. A letter a is called *V_1-maximal* with respect to G iff it is maximal in the previous sense with respect to the system $\langle \Sigma - V_0, \nu, \delta \mid \Sigma - V_0 \rangle$, where ν is any word in $(\Sigma - V_0)^+$. So all V_1-maximal letters are in $\Sigma - V_0$ and furthermore each of them is derived from a letter in V_0. The characterization of V_1 is now very similar to that of finite letters : V_1 contains :

a) W_1 = the set of V_1-maximal letters, and

b) each noncyclic letter a having the property : there exists a derivation path from a letter in W_1 to a, while there exists no such path containing a letter in $\Sigma^c - W_1$.

As before we may also talk about V_1-bounded and V_1-unbounded words with respect to G or G'. Moreover, by a V_1-cyclic word we mean a word P satisfying $sw_{V_1}(\delta^k(P)) = P$ for some $k \geq 1$. Using these notations we write

$$\Delta_{n_0}(P\alpha Q) = P_{n_0}\beta_0 R_1 \beta_1 \ldots \beta_{p-1} R_p \beta_p Q_{n_0}, \tag{4}$$

where $\beta_i's$ are "long" words of $(\Sigma - (V_0 \cup V_1))^+$, $R_i's$ are both V_1-cyclic and V_1-bounded, and the words $R_i \beta_i R_{i+1}$ are V_1-unbounded. If $\Delta_{n_0}(P\alpha Q)$ has no such representation choose a greater n_0 having one. This is possible because G is of a growth order of at least 2.

Observe now that the words $sw_{V_1}(\delta(R_i))$ are both V_1-cyclic and V_1-bounded, while the words $sw_{V_1}(\delta(R_i \beta_i R_{i+1}))$ are V_1-unbounded. So $\Delta_{n_0+1}(P\alpha Q)$ can be written also in the form (4) possibly with a greater p, because it is quite possible that P_{n_0} (or Q_{n_0}) produces some new V_1-bounded words. Of course one must ensure here that all the words of the form $R_i \beta_i R_{i+1}$ in $\Delta_{n_0+1}(P\alpha Q)$ are V_1-unbounded and that all β_i-words are "long". But these demands do not cause any problems.

The above considerations show that we can write, for all $n \geq n_0$,

$$\Delta_n(P\alpha Q) = P_n \beta_0 R_1 \beta_1 \ldots \beta_{p-1} R_p \beta_p Q_n, \tag{5}$$

where β_i- and R_i-words are as in (4). Similarly, for all $n \geq n_0$, we obtain

$$\Delta'_n(P\alpha Q) = P'_n \beta'_0 R'_1 \beta'_1 \ldots \beta'_{q-1} R'_q \beta'_q Q'_n, \tag{6}$$

where the words β'_i and R'_i are defined as above but now with respect to G'. In addition, the words P_n and P'_n (resp. Q_n and Q'_n) are supposed to be in $\Sigma^* V_1$ (resp. $V_1 \Sigma^*$).

Since β_i-and β'_i-words are "long" and R_i- and R'_i-words are "short" (by their cyclicity) it follows that if $|P_m| > |P'_m|$, then $P_m = P'_m \beta'_0 \ldots R'_{i_0}$ for some i_0. But the representation $P_m \beta'_{i_0} R'_{i_0+1} \ldots Q'_m$ is of the form (6) showing that a situation where $P_n = P'_n$ for all $n \geq n_0$ is easily achieved. Similarly we may assume that $Q_n = Q'_n$ for all $n \geq n_0$, and moreover the sequences $(P_n)_{n \geq n_0}$ and $(Q_n)_{n \geq n_0}$ may be supposed to be periodic. A similar argument based on the different lengths of β_i- and β'_i-words compared with R_i- and R'_i-words shows further that $R_1 = R'_1$. So, finally, it follows that $p = q$ and that $R_i = R'_i$ for $i = 1, \ldots, p$. Let us denote

$$L = \text{the set of } R_i\text{- (or } R'_i\text{-)words.} \tag{7}$$

Thus we have proved that the representation (5) is valid both for G and for G'. Moreover, the above shows that the direct derivation according to G looks as

$$\begin{array}{c} \Delta_n(P \alpha Q) = P_n \beta_0 R_1 \beta_1 \ldots\ldots\ldots \beta_{p-1} R_p \beta_p Q_n \\ \Delta_{n+1}(P \alpha Q) = P_{n+1} \overline{\beta}_0 \overline{R}_1 \ldots \overline{R}_k \overline{\beta}_k \overline{R}_{k+1} \ldots \overline{R}_{k+p} \overline{\beta}_{k+p} \overline{R}_{k+p+1} \ldots \overline{R}_{k+p+m} \overline{\beta}_{k+p+m} Q_{n+1} \end{array} \tag{8}$$

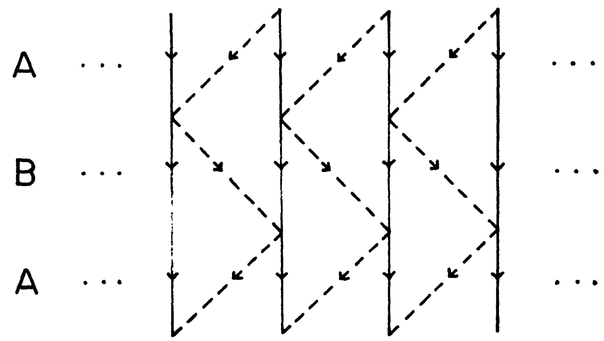
Here we have omitted some arrows compared with formula (3). Furthermore, in this illustration k and m (dependent on n) are assumed to be nonnegative. However, after we have chosen longer prefixes we can no longer be sure of the fact that $\delta(P_n)$ is a prefix of P_{n+1}. So we must allow the case where k or m is negative (for some n), meaning that some R_i-words may produce a part from P_{n+1} (or Q_{n+1}) and thus in a sense "disappear". But this does not cause any further problems.

Obviously the direct derivation according to G' has a representation similar to (8), the difference being that the number of new R_i-words introduced by P_n (or Q_n) may be different. In what follows we show that this is not possible.

Our first observation is that in each step the number of new R_i-words created by P_n and Q_n together according to G is the same as that according to G'. But we can say even more: The number of new R_i-words introduced by $P'_n s$ (or $Q'_n s$) in a common period of $(P_n)_{n \geq n_0}$ and $(Q_n)_{n \geq n_0}$ is the same according to both systems. This follows because otherwise the longest β_i-words in $\Delta_n(P \alpha Q)$ and $\Delta'_n(P \alpha Q)$ would be in different places, and so the sequences would not be equivalent.

From now on we suppose that both $P'_n s$ and $Q'_n s$ produce V_1-bounded words according to G and G'. (The argument above shows that both the systems behave in the same way; so the only other possibility is that only $P'_n s$ (or $Q'_n s$) produce V_1-bounded words. But this is essentially the same case as (iv)).

Now the problem is: Can P_n for some n introduce more V_1-bounded words according to G than according to G'? Assume that this is the case, and let words P_{n_1+jt}, for $j \geq 0$, possess the property. Now take from the derivation trees $\Delta(P\alpha Q)$ and $\Delta'(P\alpha Q)$ those levels indexed on numbers $n_1 + jt$ and $n_1 + 1 + jt$ with $j \geq 0$. The sequences thus obtained must of course be equivalent. Moreover, they may be regarded as DOL sequences because it is easy to convert them to DOL sequences without affecting the equality of the original sequences. (The construction can be carried out with (slightly generalized) decompositions of DOL systems, *see* [5]). So we may assume that the DOL systems G and G' generate the derivation trees described above. These trees restricted to V_1-bounded words look as



(9)

Figure 1.

where the nodes represent elements of L, where $\downarrow$ means the direct derivation according to G and $\swarrow$ or $\searrow$ the direct derivation according to G'. Observe that levels A are those where $P'_n s$ produce more elements of L according to G than according to G', while on levels B the situation is the opposite. (In our illustration the above word "more" must be read as meaning "exactly one more").

We continue by writing

$$\Delta_n(P\alpha Q) = P_n \gamma_n S_n \varepsilon_n Q_n, \tag{10}$$

with $S_n \in ((\Sigma - V_1)^* L)^K (\Sigma - V_1)^*$, where $S'_n s$ are chosen in such a way that all $R'_i s$ on the left (resp. right) hand side of S_n are created by $P'_n s$ (resp. $Q'_n s$) *according to G*, and K is chosen as small as possible. Note here that K is independent of n. Denote further

$$O^i_n = \text{the } i\text{th occurence of } \varepsilon_n \text{ belonging to } L(\Sigma - V_1)^* L,$$

and

V_n^i = the ith occurence of γ_n from the right belonging to $L(\Sigma - V_1)^* L$.

Next we are going to show that there exist i_1, i_2, n_1 and n_2 satisfying

$$V_{n_1}^{i_1} = O_{n_2}^{i_2}. \tag{11}$$

For this purpose let us consider a fixed V_m^j. Define

$$\bar{\delta} = \begin{cases} \delta \text{ on the levels } A \text{ of (9)}, \\ \delta' \text{ on the levels } B \text{ of (9)}. \end{cases}$$

Then obviously for a sufficiently large k:

$$sw_{V_1}(\bar{\delta}^k(V_m^j)) = O_{n_2}^{i_2}$$

for some i_2 and n_2. Remember now that the sequence $(P_n)_{n \geq n_0}$ is periodic. So it follows that

$$(V_n^j)_{n \geq m} = (V_n^{j+s})_{n \geq m'} \tag{12}$$

for a suitably chosen s. Of course we may choose s arbitrarily large which implies that

$$sw_{V_1}(\bar{\delta}^k(V_m^j)) = sw_{V_1}(\bar{\delta}^k(V_{m'}^{j+s})) = V_{n_1}^{i_1}$$

for some i_1 and n_1. So the equation (11) follows.

Now take a k_0 in such a way that

$$\left.\begin{aligned} V_{m_1}^{j_1} &\overset{\text{def}}{=} sw_{V_1}(\bar{\delta}^{k_0}(V_{n_1}^{i_1})) \in (V_n^{i_1})_{n \geq n'}, \\ O_{m_2} &\overset{\text{def}}{=} sw_{V_1}(\bar{\delta}^{k_0}(O_{n_2}^{i_2})) \in (O_n^{i_2})_{n \geq n''}. \end{aligned}\right\} \tag{13}$$

Of course this is possible the argument being the same as in establishing (12). Next we consider the following four sequences

$$\left.\begin{aligned} \text{I} &= (V_n^{i_1})_{n \geq n_1} = \omega_{\text{I}}, \ldots, \\ \text{II} &= (V_n^{j_1})_{n \geq m_1} = \omega_{\text{II}}, \ldots, \\ \text{III} &= (O_n^{i_2})_{n \geq n_2} = \omega_{\text{III}}, \ldots, \\ \text{IV} &= (O_n^{j_2})_{n \geq m_2} = \omega_{\text{IV}}, \ldots \end{aligned}\right\} \tag{14}$$

Using the above notations we can visualize the situation as follows.

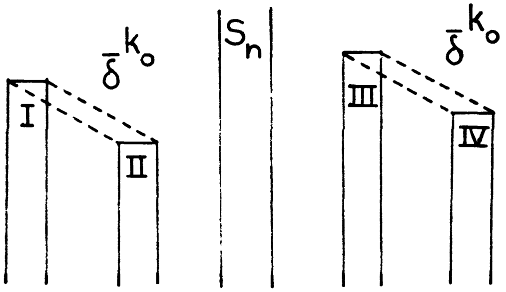


Figure 2.

It follows from (11) and (14) that I = III and II = IV. Moreover, by the above figure and (13), we have

$$sw_{V_1}(\delta^{k_1}(\omega_{\mathrm{I}})) = \omega_{\mathrm{II}} \quad \text{for some } k_1 > k_0$$

and

$$sw_{V_1}(\delta^{k_2}(\omega_{\mathrm{III}})) = \omega_{\mathrm{IV}} \quad \text{for some } k_2 < k_0.$$

Note that here k_2 may be negative, meaning that $\omega_{\mathrm{III}} = sw_{V_1}(\delta^{-k_2}(\omega_{\mathrm{IV}}))$. Now we obtain

$$sw_{V_1}(\delta^{k_1}(\omega_{\mathrm{I}})) = \omega_{\mathrm{II}} = \omega_{\mathrm{IV}} = sw_{V_1}(\delta^{k_2}(\omega_{\mathrm{III}})) = sw_{V_1}(\delta^{k_2}(\omega_{\mathrm{I}}))$$

with $k_1 > k_2$. But this means that the sequence I is V_1-bounded according to G, which is a contradiction.

What we have proved on the last few pages is that the formula (8) with the same $k's$ and $m's$ describes the direct derivation both according to G and G'. In particular, it follows that, for all n, k and m in (8) are nonnegative, if only $P_n's$ (resp. $Q_n's$) are chosen, as is natural, in such a way that either $\delta(P_n)$ (resp. $\delta(Q_n)$ or $\delta'(P_n)$ (resp. $\delta'(Q_n)$) is an initial subword of P_{n+1} (resp. a final subword of Q_{n+1}). Moreover we conclude that the equality of the sequences $\Delta(P\alpha Q)$ and $\Delta'(P\alpha Q)$ implies (and is implied by)

$$\text{(v)} \quad \begin{cases} sw_{V_0 \cup V_1}(\delta(P_n\beta_0 R_1)) = sw_{V_0 \cup V_1}(\delta'(P_n\beta_0 R_1)) & \text{for } n \geq n_0, \\ sw_{V_0 \cup V_1}(\delta(R_p\beta_p Q_n)) = sw_{V_0 \cup V_1}(\delta'(R_p\beta_p Q_n)) & \text{for } n \geq n_0, \end{cases}$$

where $P_n\beta_0 R_1$ (resp. $R_p\beta_p Q_n$) denotes the minimal prefix (resp. suffix) of $\Delta_n(P\alpha Q)$ belonging to $P_n\Sigma^* L$ (resp. $L\Sigma^* Q_n$), and

$$\text{(vi)} \qquad (sw_{V_1}(\delta^n(Y)))_{n\geq 0} = (sw_{V_1}(\delta'^n(Y)))_{n\geq 0}$$

for any word $Y \in L(\Sigma - V_1)^* L$ which is a subword of one of the following: $\Delta_{n_0}(P\alpha Q)$, $\delta(P_n\beta_0 R_1)$ for $n \geq n_0$ or $\delta(R_p\beta_p Q_n)$ for $n \geq n_0$.

Up to now we have proved that the equality of our original sequences is equivalent to the conditions (i) and (ii) together with a finite number of conditions of type (v) and (vi). The validities of (i) and (ii) are easy to decide. Similarly, by the periodicity argument, to check (v) requires only a small amount of work. Moreover, one can effectively find all the Y-words appearing in the conditions (vi). So to test the equivalence of G and G' it suffices to ascertain whether (vi) holds true for a given word Y. But the sequences in (vi) are of growth order N-1. Thus, we have proved the following: The equivalence problem for polynomially bounded DOL sequences with a growth order $N(\geq 2)$ reduces to the equivalence problem for polynomially bounded DOL sequences with a growth order N-1.

Now Theorem 1 follows from:

LEMMA 1: *The equivalence problem for linear DOL sequences is decidable.*

Proof: In this special case the languages in (iii) and (iv) are bounded context-free languages (*see* the proof of Lemma 2). Thus, the decidability of the language equivalence problem for linear DOL systems follows, *see* [3]. So the Lemma is true by a result of Nielsen, *see* [6].

On the other hand, it is not difficult to give a direct proof for Lemma 1.

REMARK: Everything presented above is constructive (as it must be to guarantee the existence of the algorithm). The reason for not paying attention to this fact is that we can thus avoid some inessential and long explanations. However, we believe that the reader will find no difficulties in convincing himself of the constructivity.

4. THE SMOOTHNESS

Here we prove that the family of polynomially bounded DOL systems forms a smooth family in the sense of Culik. Let us recall what this means. Assume that $\mathscr{F}$ is a family of DOL systems and let $G = \langle \Sigma, \omega, \delta \rangle$ and $G' = \langle \Sigma, \omega, \delta' \rangle$ be any two sequence equivalent systems in $\mathscr{F}$. Then $\mathscr{F}$ is said to be *smooth* iff the pair (G, G') has a *bounded balance*, i. e. there exists a constant K (dependent on G and G') such that for any prefix P of a word in $\{\delta^n(\omega) \mid n \geq 0\}$:

$$\big|\, |\delta(P)| - |\delta'(P)| \,\big| < K.$$

Next we establish

LEMMA 2: *The family of linear DOL systems is smooth.*

Proof: Clearly, it suffices to prove the Lemma for linear DOL systems generating sequences of the form (iii) in Section 3. So assume that $G = \langle \Sigma, \omega, \delta \rangle$ and $G' = \langle \Sigma, \omega, \delta' \rangle$ are sequence equivalent systems of such a kind.

An easy consideration shows the existence of numbers t and p such that

$$\left.\begin{aligned} \delta^{t+j+ip}(\omega) &= A_j(x_j)^i B_j(y_j)^i C_j \\ \text{and} \qquad & \\ \delta'^{t+j+ip}(\omega) &= A'_j(x'_j)^i B'_j(y'_j)^i C'_j \end{aligned}\right\} \tag{15}$$

for $i \geq 0$ and $0 \leq j < p$. Moreover, we may choose the words x_j and x'_j in such a way that

$$\left.\begin{aligned} \delta(x_j) &= x_{j+1} \quad \text{for} \quad 0 \leq j < p-1, \\ \delta(x_{p-1}) &= x_0, \\ \delta'(x'_j) &= x'_{j+1} \quad \text{for} \quad 0 \leq j < p-1, \\ \delta'(x'_{p-1}) &= x'_0, \end{aligned}\right\} \tag{16}$$

and the same hold also for the y_j- and y'_j-words.

First assume that $|x_j| = |x'_j|$ for all $j < p$. Then also $|y_j| = |y'_j|$ for all $j < p$, since otherwise the sequences would not be equivalent. So it follows immediately from (15) and (16) that the pair (G, G') has a bounded balance.

Secondly assume that there exists a k such that $|x_k| > |x'_k|$. Furthermore suppose that all the words x_j, x'_j, y_j and y'_j are nonempty. Let q be the least common multiple of the numbers $|x_k|$ and $|y'_k|$, and denote by $\bar{x}_k$ the initial subword of x_k^q having the length q. Now we choose an i_0 in such a way that

$$|A_k(\bar{x}_k)^r| > |A'_k(x'_k)^{i_0} B'_k(y'_k)^q|,$$

where $r = [q^{-1} i_0 \, |x_k|]$. So the fact that $\delta^{t+k+i_0 p}(\omega) = \delta'^{t+k+i_0 p}(\omega)$ together with the choice of q implies that for some s the following holds

$$\delta^{t+k+i_0 p}(\omega) = A_k(\bar{x}_k)^s \bar{C}_k \tag{17}$$

with $|\bar{C}_k| \leq q + |C_k|$. From this we conclude that

$$|\delta^j(\bar{x}_k)| = |\delta'^j(\bar{x}_k)| \qquad \text{for} \quad j = 0, \ldots, p,$$

since otherwise the words $\delta^m(\omega)$ and $\delta'^m(\omega)$, with $m = t + j + k + i_0 p$, would be of different length (at least when i_0 and thus also s is chosen great enough). So the fact that the pair (G, G') has a bounded balance follows.

If in our second case at least one of the words x_j, x'_j, y_j or y'_j is empty (for some j), then we directly get a representation similar to (17) for some word in $\{\delta^n(\omega) \mid n \geq 0\}$. Thus, as seen before, this implies that the pair (G, G') has a bounded balance in this case, too. Hence our proof for Lemma 2 is complete.

THEOREM 2: *The family of polynomially bounded DOL systems is smooth.*

Proof: The Theorem follows immediately from Lemma 2 and from the considerations of Section 3. Namely, Lemma 2 together with formulas (3) and (8) guarantee that any pair of sequence equivalent DOL systems with a growth order 2 has a bounded balance. Repeating the argument we may conclude that the same holds true for DOL systems with an arbitrary growth order.

REFERENCES

1. K. CULIK, II, *On the Decidability of the Sequence Equivalence Problem for DOL-Systems*, Manuscript, 1975.
2. A. EHRENFEUCHT and G. ROZENBERG, *DOL-Systems with Rank*, Lecture Notes in Computer Science, Springer, Vol. 15, 1974, pp. 136-141.
3. S. GINSBURG, *The Mathematical Theory of Context-Free Languages*, McGraw-Hill, New York, 1966.
4. G. HERMAN and G. ROZENBERG, *Developmental Systems and Languages*, North-Holland, Amsterdam, 1975.
5. K. P. LEE and G. ROZENBERG, *The Length Sets of DOL Languages Are Uniformely Bounded*, Information Processing Letters, Vol. 2, 1974, pp. 185-188.
6. M. NIELSEN, *On the Decidability of Some Equivalence Problems For DOL-Systems*, Information and Control, Vol. 25, 1974, pp. 166-193.
7. *Problem book. Unusual Automata Theory*, January 1972, Dept. of Computer Science, Univ. of Aarhus, Tech. Report DAIMI PB-15, 1973, pp. 14-26.
8. G. ROZENBERG and A. SALOMAA, *The Mathematical Theory of L Systems*, in J. T. TOU (ed.), *Advances in Information Systems Science*, Plenum Press, New York, Vol. 6, 1976, pp. 161-206.
9. L. G. VALIANT, *The Equivalence Problem for DOL Systems and its Decidability for Binary Alphabets*, Manuscript, 1975.