KYBERNETIKA — VOLUME 22 (1986), NUMBER 6

A NOTE ON EXPONENTIAL DENSITY OF ETOL LANGUAGES

ALEXANDR MEDUNA

It is shown that any infinite ETOL language is exponentially dense.

0. INTRODUCTION

In the last years much attention has been paid to the study of the properties of ETOL languages (see e.g. Chap. V. 2. in [2]). This topic is also the subject of the present note investigating whether or not every infinite ETOL language is exponentially dense. Our answer is in the affirmative.

Since the family of infinite ETOL languages is quite a large subset of the family of infinite context sensitive languages and, moreover, because the family of infinite context sensitive languages and that of exponentially dense languages are known to be incomparable (see the conclusion of the present note) the fact that every infinite ETOL language is exponentially dense is of some interest. This result can help us in the first place when we want to prove the given language not to be an ETOL language. In such a case it will namely do to show that it is not exponentially dense.

1. PRELIMINARIES

We shall assume that the reader is familiar with the theory of parallel rewriting systems. Items not defined explicitly are standard in language theory, see e.g. [2]. We recall now some terminology and notation.

Let V be an alphabet and let $x = a_1 \ldots a_n$, $n \geqq 0$, be a word over V, i.e. $x \in V^*$. We use $|x|$ to denote the length of x, i.e. $|x| = n$. If $n = 0$ then the word is called the empty word, denoted as Λ. The alphabet of x, denoted $alph(x)$, is the set of all symbols from V that appear at least once in x. For a finite alphabet Δ, $\#_\Delta(x)$ denotes the number of occurrences of symbols from Δ in x.

An *ETOL system* is a construct $G = (V, P, S, \Sigma)$ where V is the alphabet of G, Σ is a subset of V, called the terminal alphabet of G, S is a symbol from $V \setminus \Sigma$, called the axiom and P is a finite set of finite substitutions on V, called tables. Usually, a table $T \in P$ in an ETOL system $G = (V, P, S, \Sigma)$ will be specified by its set of productions, i.e. $T = \{a \to \alpha : a \in V, \alpha \in T(a)\}$. We also write $a \to^T \alpha$ to indicate that $a \to \alpha$ is a production in T.

Let $G = (V, P, S, \Sigma)$ be an ETOL system. Let x and y be words over V. We say that x *directly derives* y (in G), denoted as $x \Rightarrow_G y$ if $y \in T(x)$ for some table $T \in P$. Let $\Rightarrow_G^*$ be the reflexive and transitive closure of $\Rightarrow_G$. Let x and y be words over V. We say that x *derives* y if $x \Rightarrow_G^* y$. The *language* of G, denoted $L(G)$, is the set of all terminal words that can be derived from the axiom, i.e. $L(G) = \{x \in \Sigma^* : S \Rightarrow_G^* x\}$. The relations $\Rightarrow_G$ and $\Rightarrow_G^*$ are often denoted as $\Rightarrow$ and $\Rightarrow^*$ if G is understood. If we want to express that x directly derives y using the table $T \in P$ then we write $x \Rightarrow_G^T y$ or $x \Rightarrow^T y$ if G is understood. We say that G is a *propagating* ETOL system, abbreviated as EPTOL system, if every table T in P is Λ-free. Let σ be a word in P^* and let x and y be words over V. We say that x derives y using σ denoted as $x \Rightarrow_G^\sigma y$ (or $x \Rightarrow^\sigma y$ if G is understood) if either $\sigma = \Lambda$ and $x = y$ or there exist words $x_1, \ldots, x_{|\sigma|} \in V^*$ such that $x \Rightarrow_G^{\sigma(1)} x_1 \Rightarrow_G^{\sigma(2)} \ldots \Rightarrow_G^{\sigma(|\sigma|)} x_{|\sigma|} = y$. A symbol a from V is called *active* (in G) if there exists a table T in P and a word α in V^* different from a such that $a \to^T \alpha$. We use $A(G)$ to denote the set of all active symbols in G. Let k be a positive integer. We say that G is of *index* k if for every word w in $L(G)$ there exists a derivation $D: S = x_0 \Rightarrow_G x_1 \Rightarrow_G \ldots \Rightarrow_G x_n = w$ of w such that for $0 \leqq i \leqq n$, $\#_{A(G)}(x_i) \leqq k$. We say that G is of *finite index* if G is of index k for some $k \geqq 1$. A symbol A from V is called *lasting actively recursive* (in G), abbreviated *LA*-recursive, if there exist x_1, x_2 in V^*, w in Σ^* and ϱ in P^* such that

(i) $S \Rightarrow^* x_1 A x_2 \Rightarrow^* w \in L(G)$,

(ii) $A \Rightarrow^\varrho \alpha A \beta$, $x_1 \Rightarrow^\varrho \bar{x}_1$, $x_2 \Rightarrow^\varrho \bar{x}_2$ where $\#_{A(G)}\, \alpha\beta \geqq 1$, $alph(\bar{x}_1 \alpha A \beta \bar{x}_2) \subseteq$ $\subseteq alph(x_1 A x_2)$,

(iii) there exists an active symbol B such that $B \in alph(\alpha\beta)$ and $B \Rightarrow^\varrho \gamma_1 B \gamma_2$, where $alph(\gamma_1\gamma_2) \subseteq alph(x_1 A x_2)$.

Language K is called *exponentially dense* if there exist positive constants c_1 and c_2 having the following property: for any $n \geqq 1$ there exists a string x in L such that $c_1\, \mathrm{e}^{(n-1)c_2} \leqq |x| < c_1\, \mathrm{e}^{nc_2}$.

2. EXPONENTIAL DENSITY OF ETOL LANGUAGES

In this section we show that each infinite ETOL language is exponentially dense.

Theorem. Every infinite ETOL language is exponentially dense.

Proof. I. Let $G = (V, P, S, \Sigma)$ be an ETOL system which is not of finite index and let $L(G)$ be infinite. The standard construction of proof of Theorem 0.2 in [3] to produce an EPTOL system

$$G' = (V', P', S, \Sigma)$$

equivalent to the given G preserves the index. (Thus G' is not of finite index either.) Immediately from Corollary 1.3. in [3] it follows that V' in G' contains at least one LA-recursive symbol A. From the definition of the LA-rescursive symbol it follows that in G' there exists an infinite sequence of derivations of the following form:

$$S \Rightarrow^{\pi} x_1 A y_1 \Rightarrow^{\sigma} w_1$$

$$S \Rightarrow^{\pi} x_1 A y_1 \Rightarrow^{\varrho} \bar{x}_1 \alpha A \beta \bar{y}_1 = x_2 A y_2 \Rightarrow^{\sigma} w_2$$

$$\vdots$$

$$S \Rightarrow^{\pi} x_1 A y_1 \Rightarrow^{\varrho} x_2 A y_2 \Rightarrow^{\varrho} \dots x_{i-1} A y_{i-1} \Rightarrow^{\varrho} \bar{x}_{i-1} \alpha A \beta \bar{y}_{i-1} = x_i A y_i \Rightarrow^{\sigma} w_i$$

$$S \Rightarrow^{\pi} x_1 A y_1 \Rightarrow^{\varrho} x_2 A y_2 \Rightarrow^{\varrho} \dots x_{i-1} A y_{i-1} \Rightarrow^{\varrho} x_i A y_i \Rightarrow^{\varrho} \bar{x}_i \alpha A \beta \bar{y}_i = x_{i+1} A y_{i+1} \Rightarrow^{\sigma} w_{i+1}$$

$$\vdots$$

for each $i \geqq 1$ where $w_i \in L(G')$ and $A \in V, x_1, \dots, x_{i+1}, y_1, \dots, y_{i+1} \in V^*$, $\pi, \varrho, \sigma \in$ $\in P^*$.

Since $\#_{A(G')} \alpha\beta \geqq 1$ and G' is propagating we have:

$$|w_i| < |w_{i+1}| \quad \text{for each} \quad i \geqq 1\,.$$

Let

$$r = \max\{|T(a)| : \text{for some } T \in P' \text{ and } a \in V'\},$$
$$s = |\varrho| + |\sigma|$$
$$d = r^s\,.$$

Clearly

$$d > 1\,.$$

Let

$$c_1 = |w_1|\,, \quad c_2 = \ln d\,.$$

Now using a very simple procedure we will prove that $L(G')$ is exponentially dense: Since $d > 1$ we have for $n = 1$: $c_1 \leqq |w_1| < c_1\,\mathrm{e}^{c_2} = c_1\,\mathrm{e}^{\ln d} = c_1 d$. Let n be an arbitrary fixed integer, $n > 1$. Then there exist w_i, w_{i+1} for some $i \geqq 1$ (see above) such that: $|w_i| < c_1\,\mathrm{e}^{nc_2} \leqq |w_{i+1}|$. Since $c_1\,\mathrm{e}^{nc_2} \leqq |w_{i+1}|$ and $|w_{i+1}|/d \leqq w_i$ we have:

$$c_1\,\mathrm{e}^{(n-1)c_2} = \frac{c_1\,\mathrm{e}^{(n-1)c_2}\,d}{d} = \frac{c_1\,\mathrm{e}^{(n-1)c_2}\,\mathrm{e}^{\ln d}}{d} = \frac{c_1\,\mathrm{e}^{(n-1)c_2}\,\mathrm{e}^{c_2}}{d} = \frac{c_1\,\mathrm{e}^{nc_2}}{d} \leqq \frac{|w_{i+1}|}{d} \leqq |w_i|$$

and thus

$$c_1\,\mathrm{e}^{(n-1)c_2} \leqq |w_i| < c_1\,\mathrm{e}^{nc_2}\,.$$

II. Let $G = (V, P, S, \Sigma)$ be an ETOL system of index k (for some $k \geqq 1$) and let $L(G)$ be infinite. By Theorem 1.10. in [3] we can assume that there exist positive integers z and $\bar{z}$ such that, for every word w in $L(G)$ that is longer than z, there exists a positive integer $t \leqq 2k$ such that w can be written in the form $w = y_0 \alpha_1 y_1 \alpha_2 \dots \alpha_t y_t$ with $|\alpha_i| < \bar{z}$ for $1 \leqq i \leqq t$, $\alpha_1 \alpha_2 \dots \alpha_t \neq \Lambda$ and for every positive integer m, the word $y_0 \alpha_1^m y_1 \alpha_2^m \dots \alpha_t^m y_t \in L(G)$. Without loss of generality we can assume that we have some such w from $L(G)$ for which it moreover holds that

$$|w| \geqq 2 .$$

Let

$$c_1 = 1 , \quad c_2 = \ln d .$$

where d is an arbitrary fixed integer, such that $2 \leqq |w| < d$. If $n = 1$ then

$$c_1 \leqq |w| < c_1 \mathrm{e}^{c_2} = \mathrm{e}^{\ln d} = d .$$

Let n be an arbitrary fixed integer, such that $n > 1$. Since $L(G)$ is infinite, we can assume that there exist $x_{i-1}, x_i \in L(G)$ such that $|x_{i-1}| < c_1 \mathrm{e}^{nc_2} \leqq |x_i|$ and no word $u \in L(G)$ for which it would hold that $|x_{i-1}| < |u| < |x_i|$.
Since

$$d > |w| \geqq |\alpha_1 \dots \alpha_t|$$

we have

$$|x_i| - |x_{i-1}| \leqq d$$

and thus

$$\frac{|x_i|}{d} \leqq \frac{|x_{i-1}|}{d} + 1 .$$

Using relation $|x_i| \geqq c_1 \mathrm{e}^{nc_2}$ we get

$$\frac{c_1 \mathrm{e}^{nc_2}}{d} = \frac{c_1 \mathrm{e}^{(n-1)c_2} \mathrm{e}^{c_2}}{d} = \frac{c_1 \mathrm{e}^{(n-1)c_2} \mathrm{e}^{\ln d}}{d} = \frac{c_1 \mathrm{e}^{(n-1)c_2} d}{d} =$$

$$= c_1 \mathrm{e}^{(n-1)c_2} \leqq \frac{|x_i|}{d} \leqq \frac{|x_{i-1}|}{d} + 1 .$$

Since $2 < d$ and clearly $|x_{i-1}| \geqq 2$ we obtain

$$c_1 \mathrm{e}^{(n-1)c_2} \leqq \frac{|x_{i-1}|}{d} + 1 \leqq |x_{i-1}|$$

and thus

$$c_1 \mathrm{e}^{(n-1)c_2} \leqq |x_{i-1}| < c_1 \mathrm{e}^{nc_2} .$$

Hence, any infinite ETOL language is exponentially dense. □

We note that immediately from the proof of Theorem 10 in [1] it follows that the family of infinite context sensitive languages and that of exponentially dense languages are incomparable. But the family of infinite ETOL languages is a subset of both the family of exponentially dense languages (as demonstrated above) and the family of infinite context sensitive languages. Thus no infinite context sensitive language which is not exponentially dense can be an ETOL language. And this is a fact that can often be of help by proving (e.g. the context sensitive language $\{a^{2^{2^n}}: n \geqq 0\}$ is not an ETOL language for not being exponentially dense).

ACKNOWLEDGEMENT

The author thanks doc. RNDr. Jiří Kopřiva, CSc. for useful discussions during the preparation of this paper.

(Received February 26, 1985.)

REFERENCES

[1] K. Culik, II and J. Opatrny: Context in parallel rewriting. In: *L* Systems (G. Rozenberg and A. Salomaa eds.), Springer-Verlag, New York 1974.

[2] G. Rozenberg and A. Salomaa: The Mathematical Theory of *L* Systems. Academic Press, London 1980.

[3] D. Vermeir: On Structural Restriction of ETOL Systems. Ph. D. Thesis, Univ. of Antwerp 1978.

RNDr. Alexandr Meduna, Laboratoř výpočetní techniky University Palackého (Laboratory of Computational Technique — Palacký University), Leninova 26, 771 46 Olomouc. Present address: *Oblastní výpočetní centrum VUT (Computing Centre of the Technical University), Obránců míru 21, 602 00 Brno. Czechoslovakia.*