INFORMATION AND CONTROL 44, 134–163 (1980)

A Study in Parallel Rewriting Systems

H. C. M. KLEIJN AND G. ROZENBERG

Institute of Applied Mathematics and Computer Science, University of Leiden, Wassenaarseweg 80, Leiden, The Netherlands

In this paper we study systematically three basic classes of grammars incorporating parallel rewriting: Indian parallel grammars, Russian parallel grammars and L systems. In particular by extracting basic characteristics of these systems and combining them we introduce new classes of rewriting systems ($ET0L_{[k]}$ systems, $ET0LIP$ systems and $ET0LRP$ systems) Among others, some results on the combinatorial structure of Indian parallel languages and on the combinatorial structures of the new classes of languages are proved. As far as $ET0L$ systems are concerned we prove that every $ET0L$ language can be generated with a fixed (equal to 8) bounded degree of parallelism.

INTRODUCTION

The study of parallel rewriting systems constitutes a central trend in formal language theory. The parallel rewriting in its most "pure" form is present in L systems (Rozenberg and Salomaa). Also quite a number of rewriting systems were investigated in the literature which form a "bridge" between pure sequential rewriting systems (as, e.g. context free grammars) and L systems; among those Indian parallel grammars and Russian parallel grammars form two very interesting cases see, e.g., Siromoney and Krithivasan (1974), Levitina (1972), Skyum (1974), Dassow (1979) and Salomaa (1974).

This paper studies Indian parallel, Russian parallel and $ET0L$ ways of rewriting. We study those systems (and languages they generate) as well as by combining them we introduce new classes of rewriting systems. In this way this paper directly continues the work begun in Salomaa (1974). We believe that such a comparative study sheds light on both the nature of parallel rewriting and the nature of sequential rewriting. Understanding each of those kinds of rewriting separately, and understanding the differences and similarities between them is, in our opinion, one of the important research areas of formal language theory.

The paper is organized as follows.

In Section I we introduce some basic notation for our paper.

In Section II we investigate Indian parallel grammars. In particular we prove a result on the combinatorial structure of Indian parallel languages that is analogous to the pumping theorem for context free languages.

0019-9958/80/020134-30$02.00/0

In Section III we combine Indian parallel and $ET0L$ ways of rewriting. This results in a new kind of rewriting systems called $ET0LIP$ systems. We investigate the language generating power of those systems. We also formalize the notion of the deterministic part of an $ET0L$ language and then characterize it. We believe that in this way we contribute to the understanding of the notion of determinism in grammars.

In Section IV we introduce 1-restricted $ET0L$ systems which, within the frame work of $ET0L$ systems, form a "nondeterministic" counterpart of Indian parallel grammars. We prove a theorem on the combinatorial structure of languages generated by those systems.

In Section V we extend the notion of a 1-restricted $ET0L$ system to a k-restricted $ET0L$ system; that is an $ET0L$ system which uses only tables such that each of them has productions different from identity for no more than k symbols. A very natural question is whether or not with the growth of k one gets larger classes of languages. We prove a rather surprising fact that there exists a k_0 such that the k_0-restricted $ET0L$ systems generate all $ET0L$ languages.

In Section VI we combine the Russian parallel mechanism of rewriting with $ET0L$ systems and introduce the so called $ET0LRP$ systems. We investigate the language generating power of those systems. Also we investigate the relationship between $E0LRP$ systems and $ET0L$ systems. We provide a normal form for $E0LRP$ systems which indicates that computationally those systems resemble a restricted class of controlled two-table $ET0L$ systems. Also we show how every $ET0L$ language can be represented using an $E0LRP$ language, a regular language and a homomorphism.

In the last section we provide a diagram of relationships between the different classes of languages considered in this paper.

I. Preliminaries

We assume the reader to be familiar with the theory of parallel rewriting systems, e.g. in the scope of Salomaa (1974), Rozenberg and Salomaa, chapters II and V. Perhaps the following notational matters require an additional explanation.

(1) For a finite set V, $\#V$ denotes its cardinality.

(2) For a word x, $|x|$ denotes its length and $\text{alph}(x)$ denotes the set of letters occurring in x. For a letter b, $\#_b x$ denotes the number of occurrences of b in x. Λ denotes the empty word.

(3) Given an alphabet Σ (we consider finite alphabets only!) we will often use its barred version $\bar{\Sigma} = \{\bar{a} \mid a \in \Sigma\}$. Then for a word $\alpha \in \Sigma^+$, $\alpha = a_1 \cdots a_n$, $a_1, \ldots, a_n \in \Sigma$ we use $\bar{\alpha}$ to denote the word $\bar{a}_1 \cdots \bar{a}_n$. Also $\bar{\Lambda} = \Lambda$. A homomorphism h on Σ^* is called *weak identity* if for every $b \in \Sigma$ either $h(b) = b$ or $h(b) = \Lambda$.

(4) All the rewriting systems that we will consider use context free productions, that is productions of the form $A \to \alpha$ where A is a letter and α is a word; then A is referred to as the left-hand side of the production and α as its right-hand side. Given a set of productions P, $LH(P)$ denotes the set of all left-hand sides of productions in P. For a rewriting system G, maxr(G) denotes the maximal length of the right-hand sides of all productions in G. As usual, $\Rightarrow_G$, $\Rightarrow_G^+$ and $\overset{*}{\Rightarrow}_G$ will be used for denoting the direct derivation relation, the "real" derivation relation and the derivation relation in G, respectively; we will also use $\Rightarrow$, $\Rightarrow^+$ and $\overset{*}{\Rightarrow}$ whenever G is clear from the context. Also $\Rightarrow_G^n$, $\Rightarrow_G^{\leqslant n}$ and $\Rightarrow_G^{\geqslant n}$ will denote the relations "derives in n steps", derives in no more than n steps" and "derives in no less than n steps", respectively.

(5) Given a class X of rewriting systems, $\mathscr{L}(X)$ denotes the family of all languages generated by systems in X. Also if a system is of type X (e.g. $ET0L$) then the language it generates is also referred to as a type X language. We use $\mathscr{L}(REG)$ and $\mathscr{L}(CF)$ to denote the classes of regular and context free languages respectively.

II. Indian Parallel Grammars

In this section we will investigate Indian parallel grammars and in particular we will prove a result on the structure of Indian parallel languages which corresponds to the pumping lemma for context free languages. This result will allow us to provide examples of languages that are not Indian parallel.

We start by recalling the definition of an Indian parallel grammar and language.

Definition. (1) An *Indian parallel grammar*, abbreviated an *IP grammar*, is a construct $G = (\Sigma, P, S, \Delta)$, where Σ is a nonempty alphabet, Δ a nonempty subset of Σ (the elements of Δ are referred to as *terminals*), $S \in \Sigma \backslash \Delta$ (the *axiom*) and P is a finite nonempty set of *productions* each of which is of the form $A \to \alpha$, where $A \in \Sigma \backslash \Delta$ and $\alpha \in \Sigma^*$. The elements of $\Sigma \backslash \Delta$ are called *nonterminals*.

(2) Let $x \in \Sigma^*$ and $y \in \Sigma^*$. We say that *x directly derives y in G*, denoted as $x \Rightarrow_G y$, if there exists a production $A \to \alpha$ in G such that $x = x_0 A x_1 A \cdots A x_k$, $y = x_0 \alpha x_1 \alpha \cdots \alpha x_k$, $k \geqslant 1$ and $A \notin \text{alph}(x_0 x_1 \cdots x_k)$.

(3) As usual $\overset{*}{\Rightarrow}_G$ is defined as the transitive and the reflexive closure of the relation $\Rightarrow_G$. If $x \overset{*}{\Rightarrow}_G y$ then we say that *x derives y in G*.

(4) The language of G, denoted $L(G)$, is defined by

$$L(G) = \{x \in \Delta^* \mid S \overset{*}{\underset{G}{\Rightarrow}} x\};$$

we say that $L(G)$ is an *Indian parallel language* or *IP language*.

The notions of a derivation and of a derivation tree in an *IP* grammar are defined analogously to the case of a context free (*CF*) grammar. Given a derivation D in an *IP* grammar G, leading from x to y, one can assign to it the unique sequence τ of productions applied (in this order) in D. This sequence τ is called the *control sequence of* D and we also write $\tau(x) = y$ (thus we view τ as both the sequence of productions and as a function); moreover the sequence of words $x = x_0, x_1, ..., x_n = y$, corresponding to applications of productions from τ (in this order) is called the *trace of* τ (*on* x). In the same way we can assign a control sequence to a derivation tree T by first taking a derivation D corresponding to T and then taking the control sequence of D. Given a control sequence τ and a nonterminal symbol A we use $\tau\backslash A$ to denote the sequence of productions resulting from τ by omitting in τ all productions with A as the left-hand side (the so called *A-productions*).

Analogously to the case of *CF* grammars we term an *IP* grammar $G = (\Sigma, P, S, \Delta)$ *reduced* if every nonterminal A is reachable (that is $S \stackrel{*}{\Rightarrow}_G x_0 A x_1$ for some $x_0, x_1 \in \Sigma^*$) and productive (that is $A \stackrel{*}{\Rightarrow}_G w$ for a word $w \in \Delta^*$).

The following notion will be useful in the proof of the main theorem of this section.

Definition. Let $G = (\Sigma, P, S, \Delta)$ be an *IP* grammar and let $A \in \Sigma \backslash \Delta$. Let D be a derivation leading from A to a terminal word x and let τ be the control sequence of D. We say that D is *composed* if $\tau = \mu \circ \rho$ where $\rho(A) = \alpha A \beta$ for $\alpha, \beta \in \Sigma^*$, $\alpha\beta \neq \Lambda$, $\mu(\alpha) = \bar{\alpha}$, $\mu(\beta) = \bar{\beta}$ and $\bar{\alpha}\bar{\beta} \neq \Lambda$. We also say that A is a *composed letter* and that the derivation tree corresponding to D is *composed*.

We will define now a new kind of rewriting systems. They will turn out to be useful in investigating the structure of *IP* languages.

Definition. (1) An *embracing grammar* G is a construct $(\Sigma; x_0, ..., x_m; w)$, where $m \geq 1$, Σ is a nonempty alphabet and $x_0, ..., x_m, w \in \Sigma^*$. The *sequence of* G, denoted $E(G)$, is defined by $E(G) = w_0, w_1, ...$, where $w_0 = w$ and $w_{i+1} = x_0 w_i x_1 \cdots x_{m-1} w_i x_m$ for $i \geq 0$. The *language of* G, denoted $L(G)$, is defined by $L(G) = \{w_0, w_1, ...\}$.

(2) A *Λ-augmented embracing grammar* G is either an embracing grammar or it is a construct $(\Sigma; x_0, ..., x_m; w, \Lambda)$, where $U(G) = (\Sigma; x_0, ..., x_m; w)$ is an embracing grammar. (If G is an embracing grammar, then we set $U(G) = G$). If G is an embracing grammar, then its sequence and language are defined as above. In the case that G is not an embracing grammar, then its sequence $E(G)$ is defined by $E(G) = \Lambda, w_0, w_1, ...$, where $E(U(G)) = w_0, w_1, ...$, and its language is defined by $L(G) = L(U(G)) \cup \{\Lambda\}$. G is called *nontrivial* if $L(G)$ is infinite.

The following obvious result characterizing Λ-augmented embracing grammars is given without a proof.

LEMMA II.1. *Let G be a Λ-augmented embracing grammar with $U(G) = (\Sigma; x_0, \ldots, x_m; w)$. Then G is nontrivial if and only if either $m = 1$ and $x_0x_1 \neq \Lambda$, or $m \geqslant 2$ and $w \neq \Lambda$, or $m \geqslant 2$ and $x_i \neq \Lambda$ for some $i \in \{0, \ldots, m\}$.*

Our next result is the main theorem of this section and it concerns the combinatorial structure of *IP* languages. It is analogous to the celebrated pumping theorem for context free languages. (The existence of such a result is hinted at at the end of Siromoney and Krithivasan (1974)).

THEOREM II.1. *For every infinite IP language L there exist positive integers n, l and nontrivial Λ-augmented embracing grammars $H_1, \ldots, H_l$, such that for every word x in L the following holds: if $|x| > n$, then there exist positive integers r, t, $1 \leqslant r \leqslant l$ and words $x_0, \ldots, x_t$, such that*

$$x = x_0\mathbf{0}(E(H_r))\, x_1\mathbf{0}(E(H_r))\, x_2 \cdots \mathbf{0}(E(H_r))\, x_t$$

and for every positive integer m $x_0\mathbf{m}(E(H_r))\, x_1\mathbf{m}(E(H_r))\, x_2 \cdots \mathbf{m}(E(H_r))\, x_t \in L$, where $\mathbf{m}(E(H_r))$ denotes the mth element of $E(H_r)$.

Proof. Let L be an infinite *IP* language and let $G = (\Sigma, P, S, \Delta)$ be a reduced *IP* system, generating L.

(1) There exists a positive integer n_0, such that for each word $z \in L$, where $|z| > n_0$, there exists a derivation tree for z containing a composed subtree of height smaller than n_0. This is seen as follows.

(i) If z is a word of length greater than $\bar{m} = (\mathrm{maxr}(G))^{\# V_N}$, where $\mathrm{maxr}(G) = \max\{|\alpha| \mid A \rightarrow \alpha \in P\}$, then, clearly every derivation tree of z in G must have a composed subtree.

(ii) Clearly the number of different words in Σ^* that can be derived from a word in Σ^* without introducing a composed subtree in the derivation tree is smaller than some positive integer $\bar{n}$ dependent on G only.

(iii) We will demonstrate now that $n_0 = \max\{\bar{n}, \bar{m}\}$ satisfies the statement of our claim.

Assume that $z \in L$, where $|z| > n_0$, and let T be a derivation tree of z in G. Since $|z| > \bar{m}$, (i) implies that T has a composed subtree. If no composed subtree of T is of height smaller than n_0, then (ii) implies that among the last $\bar{n}$ words of the trace of T there are two identical words. Thus T can be shortened to yield a derivation tree $T^{(1)}$ of z in G which is of height smaller than the height of T. If no composed subtree of $T^{(1)}$ is of height smaller than n_0, then we iterate the above procedure which yields then the sequence $T, T^{(1)}, T^{(2)}, \ldots$ of derivation trees of z in G such that each next tree in the sequence is of height smaller than the previous one. Thus for some $i \geqslant 1$, $T^{(i)}$ must be a derivation tree of z in G

such that it contains a composed subtree of height smaller than n_0. Hence our claim holds.

(2) For every nonterminal A let Term(A) denote the set of all words $w \in \Delta^*$ such that A can derive w in G in no more than n_0 steps. Since G is reduced Term(A) is nonempty and (1) implies that if A is composed then Term(A) contains a nonempty word.

(3) Now with every composed letter A and every element δ of Term(A) we associate a fixed nontrivial embracing grammar $G_{A,\delta}$ as follows. Let A be a composed letter and let T_A be a fixed composed tree for A. Let τ_A be a fixed control sequence of T_A and let $A, z_1, \ldots, z_q$ be the trace of τ_A on A. Let p be the largest integer such that $A \in \text{alph}(z_p)$ and then let μ_A, ν_A be the decomposition of τ_A such that ν_A leads from z_p to z_q (hence $\tau_A = \nu_A \circ \mu_A$). Thus we have $\mu_A(A) = \alpha A \beta$, $\nu_A(A) = \gamma$, $\nu_A(\alpha\beta) = \bar{\alpha}\bar{\beta}$ for some $\alpha, \beta \in \Sigma^*$ and $\bar{\alpha}, \bar{\beta}, \gamma \in \Delta^*$ with $\alpha\beta \neq \Lambda$ and $\bar{\alpha}\bar{\beta} \neq \Lambda$. Hence $(\nu_A \backslash A)(\alpha A \beta) = w_0 A w_1 A \cdots A w_k$, where $w_0, w_1, \ldots, w_k \in \Delta^*$ and either $k = 1$ and $w_0 w_1 \neq \Lambda$ or $k \geqslant 2$. Let us consider now an arbitrary element δ from Term(A). We have two cases to consider.

(i) $\delta \neq \Lambda$.

Then, by Lemma II.1, $G_{A,\delta} = (V; w_0, \ldots, w_k; \delta)$ is a nontrivial embracing grammar, where $V = \text{alph}(w_0 \cdots w_k \delta)$.

(ii) $\delta = \Lambda$.

(ii)(1) If for some $i \in \{0, \ldots, k\}$, $w_i \neq \Lambda$, then by Lemma II.1, $G_{A,\delta} = (V; w_0, \ldots, w_k; \Lambda)$ is a nontrivial embracing grammar, where $V = \text{alph}(w_0 \cdots w_k)$.

(ii)(2) If $w_i = \Lambda$ for every $i \in \{0, \ldots, k\}$, then it must be that $k \geqslant 2$ and, by Lemma II.1, $G_{A,\delta} = (V; w_0, \ldots, w_k; \bar{\alpha}\gamma\bar{\beta}, \Lambda)$ is a nontrivial Λ-augmented embracing grammar, where $V = \text{alph}(w_0 \cdots w_k \bar{\alpha}\gamma\bar{\beta})$.

(4) Now we complete the proof of the theorem as follows. Let $x \in L$ and $|x| > n_0$. By (1) there is a derivation tree T of x in G such that T contains a composed subtree of height smaller than n_0. Let A be the label of the root of such a subtree, let τ be a fixed control sequence of T and let $S, u_1, \ldots, u_g = x$ be the trace of τ on S. Let f be the largest integer such that $A \in \text{alph}(u_f)$ and let ρ, π be the decomposition of τ such that ρ leads from S to u_f and π leads from u_f to u_g (hence $\tau = \pi \circ \rho$). Let $\rho(S) = y_0 A y_1 A \cdots A y_t$, where $t \geqslant 1$ and $A \notin \text{alph}(y_1 \cdots y_t)$. Let $(\pi \backslash A)(y_0 A y_1 A \cdots A y_t) = x_0 A x_1 A \cdots A x_t$, where $x_1, \ldots, x_t \in \Delta^*$ and let $\pi(A) = \delta$; obviously $\delta \in \text{Term}(A)$. Thus $x = \tau(S) = x_0 \delta x_1 \delta \cdots x_t$ and moreover, for every $m \geqslant 1$, $x_0 \theta_m(A) x_1 \theta_m(A) \cdots \theta_m(A) x_t \in L$, where $\theta_m = \pi \circ ((\nu_A \backslash A) \circ \mu_A)^m$ if $\delta \neq \Lambda$, $\theta_m = \nu_A \circ \mu_A \circ ((\nu_A \backslash A) \circ \mu_A)^{m-1}$ if $\delta = \Lambda$, and μ_A and ν_A are the fixed control sequences from (3). In other words $x = x_0 \mathbf{0}(E(G_{A,\delta}))\, x_1 \mathbf{0}(E(G_{A,\delta})) \cdots \mathbf{0}(E(G_{A,\delta})) x_t$ and for every $m \geqslant 1$

$$x_0 \mathbf{m}(E(G_{A,\delta}))\, x_1 \mathbf{m}(E(G_{A,\delta})) \cdots \mathbf{m}(E(G_{A,\delta})) x_t \in L.$$

Thus if we set $n = n_0$ and $\{H_1, ..., H_l\}$ to be the set of all A-augmented embracing grammars $G_{A,\delta}$ as defined in (3) (A is a composed letter and $\delta \in \text{Term}(A)$) then the theorem holds. ∎

Before we state our next result we need the following notion. Let $x = a_1 \cdots a_n$, $n \geqslant 2$, be a word over Σ, where $a_1, ..., a_n$ are occurrences of letters from Σ in x, and let $\{\Sigma_1, \Sigma_2\}$ be a nonempty partition of Σ. Then we say that an occurrence a_i, $1 \leqslant i \leqslant n - 1$, is a $\{\Sigma_1, \Sigma_2\}$-*switch* if a_i is an occurrence of a letter of Σ_1 and a_{i+1} is an occurrence of a letter from Σ_2.

It is well known that the length set of an infinite *CF* language contains an infinite arithmetic progression. Thus if the length set of an infinite *IP* language does not contain an infinite arithmetic progression the language is not *CF*; for example $\{a^{2^n} \mid n \geqslant 1\}$ is in $\mathscr{L}(IP)\setminus\mathscr{L}(CF)$. (At the same time it should be observed that an infinite *IP* language the length set of which contains an infinite arithmetic progression does not have to be *CF*; $\{ww \mid w \in \{0, 1\}^*\}$ is an example.) The following theorem allows us to provide examples of infinite languages such that their length sets do not contain an infinite arithmetic progression and the languages are not in $\mathscr{L}(IP)$.

THEOREM II.2. *Let L be an infinite IP language over an alphabet Σ and let Σ_1, Σ_2 be a nonempty partition of Σ, then either* (1) *the length set of L contains an infinite arithmetic progression, or* (2) *there exists a positive integer k_1 such that infinitely many words of L have no more than k_1 occurrences of symbols of Σ_1, or* (3) *there exists a positive integer k_2 such that infinitely many words of L have no more than k_2 occurrences of symbols of Σ_2, or* (4) *for every nonnegative integer n, there exists a word z in L, such that z has at least n $\{\Sigma_1, \Sigma_2\}$-switches.*

Proof. Let n be as in the statement of Theorem II.1, and let x in L be such that $|x| > n$. Let H_r be as in the statement of Theorem II.1 and let $U(H_r) = (\Sigma; w_0, ..., w_k; z)$. If $k = 1$, then (1) holds, if $w_0, ..., w_k, z \in \Sigma_2^*$, then (2) holds, if $w_0, ..., w_k, z \in \Sigma_1^*$, then (3) holds and if the word $w_0 \cdots w_k z$ contains occurrences of letters both from Σ_1 and Σ_2 then (4) holds. ∎

As an example of the application of the above theorem we get the following result.

COROLLARY II.1. $\{a^n b^{2^n} \mid n \geqslant 0\} \notin \mathscr{L}(IP)$.

This can be generalized to the following result.

COROLLARY II.2. *Let Σ be a finite nonempty alphabet, let Σ_1, Σ_2 be a nonempty partition of Σ and let K_1 and K_2 be infinite languages over Σ_1 and Σ_2 respectively. If $f: K_1 \to K_2$ is an injective function and the length set of $K = \{xf(x) \mid x \in K_1\}$ does not contain an infinite arithmetic progression, then $K \notin \mathscr{L}(IP)$.*

Proof. Since the length set of K does not contain an infinite arithmetic progression the case (1) from the statement of the previous theorem does not hold. Cases (2) and (3) from the statement of Theorem II.2 cannot hold, because f is injective. Case (4) from the statement of Theorem II.2 cannot hold, because each word of K has at most one $\{\Sigma_1, \Sigma_2\}$-switch. ∎

III. *ET0LIP* Systems

In this section we will combine the *IP* mechanism and the *ET0L* mechanism of rewriting; the resulting construct is an *ET0LIP* system. In Siromoney and Siromoney (1975–1976) the *IP* mechanism was combined with *0L* systems. However the results stated there are not very helpful in establishing the properties of *ET0LIP* systems; the use of nonterminals changes the situation completely.

Definition. An *ET0LIP* system is a construct $G = (\Sigma, \mathscr{P}, S, \Delta)$ where $\Sigma, \mathscr{P}, S, \Delta$ are as in *ET0L* systems. Given $x \in \Sigma^+$ and $y \in \Sigma^*$ we say that x *directly derives* y *in* G, denoted $x \Rightarrow_G y$, if $x = x_1 \cdots x_n$ with $n \geqslant 1$, $x_1, ..., x_n \in \Sigma$, $y = y_1 \cdots y_n$ with $y_1, ..., y_n \in \Sigma^*$ and there exists a $P \in \mathscr{P}$ such that $x_i \rightarrow y_i$ is production of P for each $i \in \{1, ..., n\}$ where $y_k = y_j$, whenever $x_k = x_j$, $1 \leqslant k, j \leqslant n$. The relation $\stackrel{*}{\Rightarrow}_G$ is defined as the transitive and the reflexive closure of $\Rightarrow_G$; if $x \stackrel{*}{\Rightarrow}_G y$ then we say that x *derives* y *in* G. The *language of* G is defined by $L(G) = \{x \in \Delta^* \mid S \stackrel{*}{\Rightarrow}_G x\}$.

The notation and the terminology concerning *ET0L* systems and languages are carried over to *ET0LIP* systems. In particular an *E0LIP* system is an *ET0LIP* system $(\Sigma, \mathscr{P}, S, \Delta)$ where $\# \mathscr{P} = 1$.

First of all we compare the language generating power of *E0LIP* systems and *IP* grammars.

Theorem III.1. $\mathscr{L}(IP) \subsetneqq \mathscr{L}(E0LIP)$.

Proof. Let $G = (\Sigma, P, S, \Delta)$ be an *IP* grammar. If $\Sigma \setminus \Delta = \{S = A_1, A_2, ..., A_n\}$, with $n \geqslant 1$, then let, for $j = 1, ..., n$, $\Sigma^{(j)}$ denote the set $\{A^{(j)} \mid A \in \Sigma \setminus \Delta\}$, such that $\Sigma, \Sigma^{(i)}, \Sigma^{(j)}$, for $i \neq j$, $1 \leqslant i, j \leqslant n$, are pairwise disjoint, and let for $j = 1, ..., n$ f_j be a homomorphism defined by $f_j(A_i) = A_i^{(j+1)}$ for $1 \leqslant j < n$, $1 \leqslant i \leqslant n$, $f_n(A_i) = A_i^{(1)}$ for $1 \leqslant i \leqslant n$ and $f_j(a) = a$ for $1 \leqslant j \leqslant n$ and $a \in \Delta$. Let $\bar{G} = (\bar{\Sigma}, \bar{P}, \bar{S}, \Delta) \in E0LIP$, where $\bar{\Sigma} = \bigcup_{j=1}^{n} \Sigma^{(j)} \cup \Delta$, $\bar{S} = A_1^{(1)}$ and $\bar{P}$ the following set of productions.

$$\bar{P} = \{A_i^{(i)} \rightarrow f_i(w) \mid A_i \rightarrow w \in P,\ i = 1, ..., n\}$$

$$\cup \{A_j^{(i)} \rightarrow f_i(A_i) \mid 1 \leqslant i, j \leqslant n\} \cup \{a \rightarrow a \mid a \in \Delta\}.$$

From this construction it is clear that $S \overset{*}{\Rightarrow}_G x$ if and only if there exists an integer $j \in \{1,..., n\}$, such that $\bar{S} \overset{*}{\Rightarrow}_{\bar{G}} f_j(x)$. So $L(\bar{G}) = L(G)$. Thus for every *IP* grammar G there exists an *E0LIP* system $\bar{G}$, such that $L(\bar{G}) = L(G)$ and so $\mathscr{L}(IP) \subset \mathscr{L}(E0LIP)$. Since $\{a^n b^{2^n} \mid n \geqslant 0\} \in \mathscr{L}(E0LIP) \backslash \mathscr{L}(IP)$, see Corollary II.1, it follows that this inclusion is strict. ∎

Remark. The inclusion $\mathscr{L}(IP) \subset \mathscr{L}(E0LIP)$ is stated in Siromoney and Siromoney (1975–1976), Theorem 3.2 (in a different formulation). Since its proof there seems to be incorrect, we provided the full proof of Theorem III.1.

Here is another way of arriving at *ET0LIP* languages.

For a set P of context free productions *on* Σ, that is productions of the form $A \rightarrow \alpha$, $A \in \Sigma$, $\alpha \in \Sigma^*$, where P contains a production for *each* element of Σ, we use $\det(P)$ to denote the family of all sets of productions R such that $R \subset P$ and R contains *exactly* one production for each element of Σ.

DEFINITION. Let G be an *ET0L* system, $G = (\Sigma, \mathscr{P}, S, \Delta)$. The *combinatorially complete (cc) version of* G, denoted G_{cc}, is the *EDT0L* system $G_{cc} = (\Sigma, \mathscr{R}, S, \Delta)$, where $\mathscr{R} = \bigcup_{P \in \mathscr{P}} \det(P)$. G_{cc} is referred to as an $ET0L_{cc}$ *system* (or $E0L_{cc}$ *system* if G is an *E0L* system).

We use $ET0L_{cc}$ and $E0L_{cc}$ to denote the classes of $ET0L_{cc}$ and $E0L_{cc}$systems respectively.

Directly from the above definitions we get the following results.

LEMMA III.1. *If G is an ET0L system then* $L(G_{cc}) \subset L(G)$.

LEMMA III.2. (1) $\mathscr{L}(ET0LIP) = \mathscr{L}(ET0L_{cc}) = \mathscr{L}(EDT0L)$,

(2) $\mathscr{L}(E0LIP) = \mathscr{L}(E0L_{cc})$.

We compare now the classes of *E0L* and $E0L_{cc}$ languages.

THEOREM III.2. *$\mathscr{L}(E0L)$ and $\mathscr{L}(E0L_{cc})$ are incomparable but not disjoint.*

Proof. Since in Ehrenfeucht and Rozenberg (1977) it is proved that $\mathscr{L}(CF) \not\subset \mathscr{L}(EDT0L)$ and by definition $\mathscr{L}(E0L_{cc}) \subset \mathscr{L}(EDT0L)$, it is clear that $\mathscr{L}(CF) \not\subset \mathscr{L}(E0L_{cc})$. But it is well known that $\mathscr{L}(CF) \subsetneq \mathscr{L}(E0L)$ (see, e.g., Rozenberg and Salomaa) and so $\mathscr{L}(E0L) \not\subset \mathscr{L}(E0L_{cc})$. On the other hand $L = \{¢w¢w¢w \mid w \in \{0, 1\}^*\} \notin \mathscr{L}(E0L)$ see, e.g., Rozenberg and Salomaa, while L is generated by the $E0L_{cc}$ system $G = (\{S, ¢, 0, 1\}, \{T_1, T_2\}, S, \{¢, 0, 1\})$, where

$$T_1 = \{S \rightarrow ¢¢¢,\ ¢ \rightarrow ¢0,\ 0 \rightarrow 0,\ 1 \rightarrow 1\}$$

and

$$T_2 = \{S \rightarrow ¢¢¢,\ ¢ \rightarrow ¢1,\ 0 \rightarrow 0,\ 1 \rightarrow 1\}.$$

Thus $\mathscr{L}(E0L_{cc}) \not\subset \mathscr{L}(E0L)$. It is clear that $\mathscr{L}(REG) \subset \mathscr{L}(E0L_{cc}) \cap \mathscr{L}(E0L)$ and so the theorem holds. ▮

The following result is useful in establishing the relationship between the language generated by an $ET0L$ system and the language generated by the cc version of the same system.

THEOREM III.3. *Let G be an $ET0L$ system. For every word $w \in L(G) \backslash L(G_{cc})$ there exist words w_1, w_2, w_3, α_1 and α_2, where $\alpha_1 \neq \alpha_2$, such that $w = w_1\alpha_1 w_2 \alpha_2 w_3$ and $w_{1,1} = w_1\alpha_1 w_2\alpha_1 w_3$, $w_{2,2} = w_1\alpha_2 w_2 \alpha_2 w_3$, $w_{2,1} = w_1 \alpha_2 w_2 \alpha_1 w_3 \in L(G)$.*

Proof. Let $G = (\Sigma, \mathscr{P}, S, \Delta)$ be an $ET0L$ system. Let $w \in L(G)\backslash L(G_{cc})$.

(1) First of all we may assume that there exists a derivation tree T of w with the following property. If v_1 and v_2 are two different nodes on the same level of T such that both have the same label and the same contribution to w, then the subtrees rooted at v_1 and v_2 are identical.

This is seen as follows. Take an arbitrary derivation tree $\bar{T}$ of w and proceed to "clean up" $\bar{T}$ top-down as follows: on each level of $\bar{T}$ replace all subtrees rooted at nodes with the same label and contributing the same result to w by one of those subtrees. Once this procedure ends the resulting tree T satisfies the above conditions.

(2) Let T be a derivation tree of w satisfying (1). Since $w \in L(G)\backslash L(G_{cc})$ there must be a level in T on which two different occurrences of the same symbol (say, A) have different contributions to w (say, α_1 and α_2). Then $w = w_1\alpha_1 w_2 \alpha_2 w_3$ for some $w_1, w_2, w_3 \in \Delta^*$ and obviously also $w_1\alpha_1 w_2 \alpha_1 w_3$, $w_1\alpha_2 w_2\alpha_2 w_3$, $w_1\alpha_2 w_2\alpha_1 w_3 \in L(G)$. Thus the result holds. ▮

Remark. Observe that words w, $w_{1,1}$ and $w_{2,2}$ as stated in the above theorem are all different words, but that it is possible for w and $w_{2,1}$ to be the same word. Also if $w_{2,1}$ can only be obtained as in the proof above, then $w_{2,1} \in L(G)\backslash L(G_{cc})$.

As an example of an application of the above result we present the following corollary.

COROLLARY III.1. *If $K \subset \{a^n b^n \mid n \geq 0\}$ and $K \in \mathscr{L}(ET0L)$ then $K \in \mathscr{L}(EDT0L)$.*

Proof. If $K = \{\Lambda\}$ then $K \in L(EDT0L)$. Thus assume that $K \neq \{\Lambda\}$. Let $G \in ET0L$ generate K. Let us assume that there exists a word w in $L(G)\backslash L(G_{cc})$. By Theorem III.3 there exist w_1, w_2, w_3, α_1 and α_2, where $\alpha_1 \neq \alpha_2$, such that $w = w_1\alpha_1 w_2\alpha_2 w_3$ and $w_{1,1} = w_1\alpha_1 w_2 \alpha_1 w_3$, $w_{2,2} = w_1\alpha_2 w_2\alpha_2 w_3$, $w_{2,1} = w_1\alpha_2 w_2 \alpha_1 w_3 \in K$. Let us consider all possible "distributions" of α_1 and α_2 in w.

(1) If either $\alpha_1, \alpha_2 \in \{a\}^*$ or $\alpha_1, \alpha_2 \in \{b^*\}$, then $w_{1,1} \notin K$; a contradiction.

(2) If one of α_1, α_2 is in $\{a^*\}$ and the other one is in $\{b\}^*$, then $w_{2,2} \notin K$; a contradiction.

(3) If one of α_1, α_2 contains occurrences of both a and b, then either $w_{1,1} \notin K$ or $w_{2,2} \notin K$; a contradiction.

Thus we get a contradiction in each case and consequently $L(G)\backslash L(G_{cc}) = \varnothing$. Hence $K = L(G_{cc})$ and the corollary holds. ∎

Theorem III.3 leads naturally to the following notions.

DEFINITION. Let $K \in \mathscr{L}(ET0L)$. The *deterministic core* of K, denoted $\text{dcor}(K)$, is defined by $\text{dcor}(K) = \bigcap_G \{x \mid L(G) = K \wedge x \in L(G_{cc})\}$.

DEFINITION. (1) If $K \in \mathscr{L}(ET0L)$, then a word w in K is called a *social word of* K if there exist w_1, w_2, w_3, α_1 and α_2, with $\alpha_1 \neq \alpha_2$, such that $w = w_1\alpha_1w_2\alpha_2w_3$ and $w_{1,1} = w_1\alpha_1w_2\alpha_1w_3$, $w_{2,2} = w_1\alpha_2w_2\alpha_2w_3$, $w_{2,1} = w_1\alpha_2w_2\alpha_1w_3$ are elements of K.

(2) If a word w of K is not a social word of K, then it is called an *isolated word of* K. The set of isolated words of K is denoted $\text{isol}(K)$.

We are able now to characterize the deterministic core of an $ET0L$ language by isolated words.

THEOREM III.4. *Let* $K \in \mathscr{L}(ET0L)$. *Then* $\text{dcor}(K) = \text{isol}(K)$.

Proof. (1) Let $w \in \text{isol}(K)$. From Theorem III.3 it follows that there exists no $ET0L$ system G, such that $K = L(G)$ and $w \in L(G)\backslash L(G_{cc})$. Hence $w \in \text{dcor}(K)$.

(2) Let w be a social word of K. Then there exist w_1, w_2, w_3, α_1 and α_2 as in the statement of Theorem III.3 such that $w_{1,1} = w_1\alpha_1w_2\alpha_1w_3$, $w_{2,2} = w_1\alpha_2w_2\alpha_2w_3$ and $w_{2,1} = w_1\alpha_2w_2\alpha_1w_3$ are words in K. Let $M = \{w, w_{1,1}, w_{2,2}, w_{2,1}\}$. Obviously there exists an $ET0L$ system H such that $L(H) = K\backslash M$. (This is seen as follows: let G be an $ET0L$ system over a terminal alphabet Δ, generating K and let R denote the regular language $\Delta^*\backslash M$; since $\mathscr{L}(ET0L)$ is closed under intersection with regular languages, see, e.g., Rozenberg and Salomaa, there exists an $ET0L$ system generating $K \cap R = K\backslash M$). Let $H = (\Sigma, \mathscr{P}, S, \Delta)$ and let $\hat{H} = (\hat{\Sigma}, \hat{\mathscr{P}}, A, \Delta)$ be the $ET0L$ system constructed as follows: $\hat{\Sigma}_1 = \Sigma \cup \{A, B, F\}$, where $A, B, F \notin \Sigma$, and

$$\hat{\mathscr{P}} = \bigcup_{P \in \mathscr{P}} \{P \cup \{A \to F, B \to F, F \to F\}\} \cup P_{in},$$

where

$$P_{in} = \{A \to S, A \to w_1 B w_2 B w_3, B \to \alpha_1, B \to \alpha_2\} \cup \{x \to x ; x \in \Sigma \cup \{F\}\}.$$

Hence $\hat{H} \in ET0L$ and it is easy to see that $L(\hat{H}) = L(H) \cup M = K$ and that w and $w_{2,1} \notin L(\hat{H}_{cc})$. So $w \notin \text{dcor}(K)$.

From (1) and (2) the theorem follows. ∎

Remark. Notice that the above theorem implies that if $K \in \mathscr{L}(ET0L)$ and $A \in K$, then $A \in \text{dcor}(K)$.

We conclude this section with the following two applications of Theorem III.3.

COROLLARY III.2. *If $K \in \mathscr{L}(ET0L)$, $K \subset \{a\}^*$ and the length set of K does not contain an arithmetic progression involving three or more elements then $K \in \mathscr{L}(EDT0L)$.*

Proof. Let G be an $ET0L$ system such that $L(G) = K$. If $L(G) \setminus L(G_{cc}) \neq \varnothing$, then Theorem III.3 implies that the length set of K contains an arithmetic progression involving three elements. Hence $L(G) \setminus L(G_{cc}) = \varnothing$ and $K = L(G_{cc})$. ∎

COROLLARY III.3. *If $K \in \mathscr{L}(ET0L)$ and the length set of K is thin (meaning that for each n in the length set of K there exist at most two elements x, y of K such that $|x| = |y| = n$) and it does not contain an arithmetic progression involving three or more elements then $K \in \mathscr{L}(EDT0L)$.*

Proof. Similar to the proof of Corollary III.2. ∎

IV. 1-RESTRICTED $ET0L$ SYSTEMS

Except for the fact that terminal symbols cannot be rewritten, an IP system is an $EDT0L$ system such that in each table of it at most one symbol is rewritten into something else than the symbol itself. A very natural step at this stage is to consider the "nondeterministic version" of those systems: that is to consider the class of $ET0L$ systems such that in each table of a system from this class at most one symbol can be rewritten into something else than the symbol itself. The difference is that, while as before, in a single derivationstep one chooses one symbol to rewrite, different occurrences of this symbol in a string can be rewritten, in different ways.

Those systems are termed *1-restricted ET0L systems* and they will be considered now.

DEFINITION. A *1-restricted ET0L system*, abbreviated *$ET0L_{[1]}$ system*, is an $ET0L$ system $G = (\Sigma, \mathscr{P}, S, \Delta)$ such that for every $P \in \mathscr{P}$ there exists a letter b in Σ such that if $c \in \Sigma \setminus \{b\}$ and $c \to_P \alpha$ then $\alpha = c$.

Hence in an $ET0L_{[1]}$ system each table can rewrite at most one symbol into something else than the symbol itself.

All notation and terminology concerning $ET0L$ systems are carried over to $ET0L_{[1]}$ systems. Also we term an $ET0L_{[1]}$ system $G = (\Sigma, \mathscr{P}, S, \Delta)$ *reduced* if every nonterminal A from Σ is reachable (that is $S \overset{*}{\Rightarrow}_G x_0 A x_1$ for some $x_0, x_1 \in \Sigma^*$) and productive (that is $A \overset{*}{\Rightarrow}_G w$ for a word $w \in \Delta^+$). We will consider reduced $ET0L_{[1]}$ systems only.

Before we prove our first technical result we need the following notion.

DEFINITION. Let $G = (\Sigma, \mathscr{P}, S, \Delta)$ be an $ET0L_{[1]}$ system.

(1) For every element $\sigma \in \Sigma$, the set of all productions $\sigma \to \alpha_1 \cdots \alpha_n$ such that $\sigma \notin \text{alph}(\alpha_1 \cdots \alpha_n)$, is denoted Π_σ^∞.

(2) For every element $\sigma \in \Sigma$, the *order* of σ, denoted $\rho(\sigma)$, is defined by $\rho(\sigma) = 0$ if $\sigma \in \Delta$, and $\rho(\sigma) = \min_{\pi \in \pi_\sigma}\{\max\{\rho(a) \mid a \text{ occurs at the right-hand side of } \pi\}\} + 1$ if $\sigma \notin \Delta$.

(3) For every word $w \in \Sigma^+$, where $w = \sigma_1 \cdots \sigma_n$, with $n \geqslant 1$ and $\sigma_1, \ldots, \sigma_n \in \Sigma$, the *order* of w, denoted $\rho(w)$, is defined by

$$\rho(w) = \max\{\rho(\sigma_i) \mid i = 1, \ldots, n\}.$$

Since we consider reduced $ET0L_{[1]}$ systems only, ρ is a well defined function on Σ^+.

The following technical result will be quite useful in proving the main theorem of this section.

LEMMA IV.1. *Let $G = (\Sigma, \mathscr{P}, S, \Delta)$ be an $ET0L_{[1]}$ system. There exists a nonnegative integer l such that every word $w \in \Sigma^+$ derives a nonempty word in Δ^* in no more than l steps.*

Proof. As a matter of fact we will prove that: every word $w \in \Sigma^+$ derives a nonempty word in Δ^* in no more than $(\#\Sigma) . \rho(w)$ steps. Since obviously $\rho(w) \leqslant \#\Sigma$ for every $w \in \Sigma^+$, the lemma follows from the above claim. The claim is proved by induction on $\rho(w)$ as follows.

(1) If $\rho(w) = 0$ then the claim trivially holds.

(2) Let us assume that the claim holds for all $w \in \Sigma^+$ such that $\rho(w) \leqslant k$.

(3) Let $w \in \Sigma^+$ be such that $\rho(w) = k + 1$. Clearly w contains an occurrence of a letter σ such that $\rho(\sigma) = k + 1$. If we rewrite each letter σ from $\text{alph}(w)$ with $\rho(\sigma) = k + 1$ using a production of the form $\sigma \to \alpha$ where $\rho(\alpha) = k$ then in less than $\#\Sigma$ steps w derives a word $\bar{w}$ containing no letter of order higher than k. Thus, by the induction hypothesis, in no more than $\#\Sigma + \#\Sigma \cdot k = \#\Sigma \cdot (k + 1)$ steps w derives a nonempty terminal word in Δ^*.

Thus the induction is completed and the claim holds.

Consequently the lemma holds. ▌

To state our result on the combinatorial structure of $ET0L_{[1]}$ languages we need the following two definitions.

DEFINITION. Let K be an infinite language over an alphabet Σ and let $b \in \Sigma$. We say that K is *logarithmically b-clustered* if there exists a positive integer C such that for every word w in K if $b \in \text{alph}(w)$, then $w = w_0 b w_1 \cdots b w_n$, $n \geqslant 1$, $w_0, \ldots, w_n \in \Sigma^*$, $b \notin \text{alph}(w_0 \cdots w_n)$ and $|w_j| \leqslant C \log n$ for $j = 0, \ldots, n$.

We say that K is *logarithmically clustered* if there exists a letter b in Σ, such that K is logarithmically b-clustered.

DEFINITION. Let K be a language over an alphabet Σ. We say that K is *pump-generated* if there exist positive integers r, q and words $x_0, x_1, \ldots, x_r$, $u, w, z \in \Sigma^*$ with $|uz| \neq \Lambda$ and $|uz| < q$ such that

$$K = \bigcup_{i \geqslant 0} x_0 u^i w z^i x_1 \cdots u^i w z^i x_r .$$

THEOREM IV.1. *If K is an infinite $ET0L_{[1]}$ language then either K contains an infinite logarithmically clustered language or K contains a pump-generated language.*

Proof. Let K be an infinite language generated by an $ET0L_{[1]}$ system $G = (\Sigma, \mathscr{P}, S, \Delta)$.

(1) Since K is infinite there is a symbol (say A) in Σ such that $A \overset{*}{\Rightarrow}_G \gamma A \delta$ for some $\gamma, \delta \in \Sigma^*$, where $\gamma\delta \neq \Lambda$. Let $\mathscr{D}_A$ denote the set of all derivations D leading from A to a word of the form $\gamma A \delta$ with $\gamma\delta \in \Sigma^+$ and such that at each step of D all occurrences of the letter under rewriting are rewritten by the same production. Since K is infinite $\mathscr{D}_A$ is not empty. Now let D_0 be a fixed element of $\mathscr{D}$ such that no derivation in $\mathscr{D}_A$ is shorter than D_0. Let D_0 lead from A to $\alpha A \beta$ where $\alpha\beta \in \Sigma^+$, and let $\nu_0, \ldots, \nu_k$ be the productions used by D_0 (in this order). Together $\tau = (\nu_0, \ldots, \nu_k)$ forms the *control sequence* of D_0 and we can consider τ and each of $\nu_0, \ldots, \nu_k$ also as a transformation from Σ^* into Σ^*. Note that since D_0 was "the shortest" element of $\mathscr{D}_A$, each occurrence rewritten in D_0 must

contribute at least one occurrence of A to $\alpha A\beta$. Consequently if $\nu_j = A_j \to x_j$ then $\tau(A_j)$ contains an occurrence of A. Let $R = \{A = A_0, A_1, ..., A_k\}$.

(2) Assume that $R_A \cap \text{alph}(\alpha\beta) \neq \varnothing$. Let $z \in \Sigma^*$ be such that $S \overset{*}{\Rightarrow}_G z$ and $A \in \text{alph}(z)$; since G is reduced such a z exists and moreover we can choose a z which can be derived from S in no more than $\#\Sigma$ steps. Let $\tau(z) = z_1$ and $\tau^2(z) = \tau(z_1) = z_2$. Since $R_A \cap \text{alph}(\alpha\beta) \neq \varnothing$, z_2 contains at least two occurrences of A and consequently for each $n \geqslant 1$, $\tau^{2n}(z) = z_{2n}$ contains at least 2^n occurrences of A. Let q be the maximal distance between two occurrences from R_A in z (a distance between two occurrences c_1, c_2 from R_A in z is the number of occurrences between c_1 and c_2 in z; if z contains only one occurrence c from R_A then the maximal distance is determined by the largest of the two distances: from c to the leftmost occurrence in z and from c to the rightmost occurrence in z). Note that q is bounded by $(\text{maxr}(G))^{\#\Sigma}$.

Then the maximal distance between two occurrences from R_A in z_2 is bounded by $2 \cdot 2(\text{maxr}(G))^k + q$, and in general, for $n \geqslant 1$, the maximal distance between two occurrences from R_A in z_{2n} is bounded by $2n \cdot 2(\text{maxr}(G))^k + q$. Let then $z_{2n} = u_0Au_1A \cdots Au_m$ where $u_0, ..., u_m \in \Sigma^*$ and $A \notin \text{alph}(u_0 \cdots u_m)$; we know that $m \geqslant 2^n$. By lemma IV.1 we know that there exist a constant l and a word $w_A \in \Delta^*$ such that A derives w_A in G in less than l steps; moreover we can obviously assume that in rewriting A into w_A an occurrence of A will never be introduced. Let b be a fixed letter from $\text{alph}(w_A)$. Hence $z_{2n} = u_0Au_1A \cdots Au_m$ derives in less than l steps the word $\bar{z}_{2n} = y_0w_Ay_1w_A \cdots w_Ay_m$, $y_0, ..., y_m \in \Sigma^*$, which derives in less than l steps the word $\hat{z}_{2n} = x_0w_Ax_1w_A \cdots w_Ax_m \in \Delta^+$ where the maximal distance between two occurrences of b in $\hat{z}_{2n}$ is bounded by

$$\begin{aligned}(4n(\text{maxr}(G))^k + q) \cdot (\text{maxr}(G))^{2l} + 2(\text{maxr}(G))^{2l} \\ = 4n(\text{maxr}(G))^k \cdot (\text{maxr}(G))^{2l} + (\text{maxr}(G))^{2l} \cdot (q + 2) \leqslant r \cdot n + s\end{aligned}$$

where $r = 4(\text{maxr}(G))^k \cdot (\text{maxr}(G))^{2l}$ and $s = (\text{maxr}(G))^{2l} \cdot (2 + (\text{maxr}(G))^{\#\Sigma})$. Since $m \geqslant 2^n$, $K = \{\hat{z}_{2n} \mid n \geqslant 1\}$ is an infinite logarithmically b-clustered language contained in K.

(3) Assume that $R \cap \text{alph}(\alpha\beta) = \varnothing$. Since G is reduced, there exists a word $z = y_0Ay_1A \cdots Ay_k$ with $k \geqslant 1$ and $A \notin \text{alph}(y_0 \cdots y_k)$ such that $S \overset{*}{\Rightarrow}_G z$. By Lemma IV.1 there exists a derivation leading from z to a terminal word; fix one such derivation and change it in such a way that each time A is introduced it is not rewritten anymore. In this way we get $z \overset{*}{\Rightarrow}_G x_0Ax_1A \cdots Ax_p$ where $p \geqslant 1$, $A \notin \text{alph}(x_0x_1 \cdots x_p)$ and $x_0, ..., x_p \in \Delta^*$. Now for $n \geqslant 0$, $\tau^n(A) = \alpha^nA\beta^n$ and so $\tau^n(x_0Ax_1A \cdots Ax_p) = x_0\alpha^nA\beta^nx_1\alpha^nA\beta^n \cdots \alpha^nA\beta^nx_p$. By Lemma IV.1 there exists a derivation leading from $\alpha A\beta$ to a terminal word. Let us fix one such derivation and let the control sequence of this derivation be such that it leads from A to a terminal word w, it leads from α to a terminal word u and it leads from β to a terminal word t; by Lemma IV.1 we can assume that $ut \neq \Lambda$. Thus

for each $n \geqslant 0$, $x_0u^nwt^nx_1u^nwt^n \cdots u^nwt^nx_p \in K$. Consequently K contains an infinite pump-generated language $\{x_0u^nwt^nx_1u^nwt^n \cdots u^nwt^nx_p \mid n \geqslant 0\}$.

(4) Since either $R_A \cap \text{alph}(\alpha\beta) \neq \varnothing$ or $R_A \cap \text{alph}(\alpha\beta) = \varnothing$ the theorem follows from (2) and (3). ∎

As an example of applications of the above theorem we provide now two examples of languages not in $\mathscr{L}(ET0L_{[1]})$.

EXAMPLE IV.1 $K = \{a^{2^n}b^{2^n} \mid n \geqslant 0\} \notin \mathscr{L}(ET0L_{[1]})$. This is seen as follows. Since obviously the length set of K does not contain an infinite arithmetic progresssion, K does not contain a pump-generated language. However it is easily seen that K does not contain an infinite logarithmically clustered language. Hence Theorem IV.1 implies that K is not an $ET0L_{[1]}$ language. ∎

EXAMPLE IV.2 $K = \{a^nb^nc^n \mid n \geqslant 0\} \notin \mathscr{L}(ET0L_{[1]})$. This is seen as follows. First of all it is obvious that K does not contain an infinite logarithmically clustered language. Secondly it is easily seen that K does not contain a pump-generated language.
Thus Theorem IV.1 implies that $K \notin \mathscr{L}(ET0L_{[1]})$. ∎

It is instructive to notice at this point that $\{a^nb^n \mid n \geqslant 0\} \in \mathscr{L}(ET0L_{[1]})$.

V. k-RESTRICTED $ET0L$ SYSTEMS

In the previous section we have seen that 1-restricted $ET0L$ systems are weaker in their language generating power than $ET0L$ systems in general. Hence it is natural to consider now k-restricted $ET0L$ systems; that is $ET0L$ systems which use only tables such that each of them has productions different from identity for no more than k symbols. The question is whether or not with the growth of k one gets larger classes of languages generated by k-restricted $ET0L$ systems. Answering this question is certainly important for understanding the way that $ET0L$ systems work; it certainly sheds light on the nature of parallel rewriting in general. Intuitively it is clear that the considerable language generating power of $ET0L$ systems comes from the fact that in rewriting a string x an $ET0L$ system G can "force" different sorts of letters to behave synchronously. For example, if occurrences of a letter b in x are rewritten by elements of a set B then at the same time occurrences of a letter c must be rewritten by elements of a set C, occurrences of a letter d must be rewritten by elements of a set D, etc. (Think, e.g., of the simplest way to generate $\{a_1^{2^n}a_2^{2^n} \cdots a_k^{2^n} \mid n \geqslant 0\}$ where k is a fixed integer, $k \geqslant 2$). Hence, intuitively, it seems conceivable that if more letters can be forced to behave synchronously then the language generating power increases.
In this section we disprove this conjecture by showing a rather surprising fact

that there exists a k_0 such that k_0-restricted $ET0L$ systems generate all $ET0L$ languages.

Formally k-restricted $ET0L$ systems are defined as follows.

DEFINITION. Let $G = (\Sigma, \mathscr{P}, S, \Delta)$ be an $ET0L$ sytem and let k be a positive integer.

(1) A table $P \in \mathscr{P}$ is said to be *k-restricted* if there exists a subset $\bar{\Sigma}$ of Σ such that $\#\bar{\Sigma} \leqslant k$ and if $b \to \beta$ is in P for $b \in \Sigma \backslash \bar{\Sigma}$, then $\beta = b$.

(2) G is said to be *k-restricted* if each table of G is k-restricted; we also say that G is an *$ET0L_{[k]}$ system.*

First of all we have the following result.

THEOREM V.1. *For every ET0L system G there exists an equivalent ET0L system $\bar{G}$ with three tables and such that two tables of $\bar{G}$ are 2-restricted.*

Proof. Let $G = (\Sigma, \mathscr{P}, S, \Delta)$ be an $ET0L$ system. It is well known (Rozenberg and Salomaa) that every $ET0L$ language may be generated by an $ET0L$ system containing two tables only. Hence we can assume that $\mathscr{P} = \{T_1, T_2\}$. Let f and g be homomorphisms on Σ, such that $f(a) = ¢a$ and $g(a) = \$a$, where $¢, \$ \notin \Sigma$. Let F be a symbol not in $\Sigma \cup \{¢, \$\}$. Let $\bar{G} = (\bar{\Sigma}, \bar{\mathscr{P}}, S, \Delta)$ be an $ET0L$ system, where $\bar{\Sigma} = \Sigma \cup \{¢, \$, F\}$ and $\bar{\mathscr{P}} = \{P_1, P_2, P_3\}$ with

$$P_1 = \{a \to a \mid a \in \{¢, \$, F\}\} \cup \{a \to f(w) \mid a \to w \in T_1\} \cup \{a \to g(w) \mid a \to w \in T_2\},$$
$$P_2 = \{a \to a \mid a \in \Sigma \cup \{F\}\} \cup \{¢ \to \Lambda\} \cup \{\$ \to F\}$$

and

$$P_3 = \{a \to a \mid a \in \Sigma \cup \{F\}\} \cup \{¢ \to F\} \cup (\$ \to \Lambda\}.$$

Clearly $L(\bar{G}) = L(G)$ and P_2 and P_3 are 2-restricted. Hence the theorem holds. ▌

We move now to investigate the influence of increasing the parameter k onto the language generating power of $ET0L_{[k]}$ systems. We start by observing the following.

LEMMA V.1. $\mathscr{L}(ET0L_{[1]}) \subsetneq \mathscr{L}(ET0L_{[2]})$.

Proof. Let $K = \{a^{2^n}b^{2^n} \mid n \geqslant 0\}$. By Example IV.1 $K \notin \mathscr{L}(ET0L_{[1]})$ whereas K is generated by the $ET0L_{[2]}$ system

$$G = (\{S, a, b\}, \{\{S \to ab, a \to a, b \to b\}, \{S \to S, a \to a^2, b \to b^2\}\}, S, \{a, b\}).$$

Thus the result holds. ▌

In the rest of this section we will demonstrate that the above result is not typical for the situation when one transits from k to $k+1$. We start by defining a construction which is very essential for the proof of the main result of this section.

The Carrier Construction

Let n be a fixed positive integer.

(1) Let $V = \{a_1, \ldots, a_n, b_1, \ldots, b_n, A, B, F\}$ and let $U_1, \ldots, U_n, T_1, \ldots, T_n$ be the following sets of productions.

$$U_1 = \{A \to B, B \to F, a_1 \to b_1, b_n \to F\} \cup \{v \to v \mid v \in V \setminus \{A, B, a_1, b_n\}\},$$
$$T_1 = \{A \to F, B \to A, a_n \to F, b_1 \to a_1\} \cup \{v \to v \mid v \in V \setminus \{A, B, a_n, b_1\}\}$$

and for $k = 2, \ldots, n$

$$U_k = \{A \to F, a_k \to b_k, a_{k-1} \to F, b_k \to F\} \cup \{v \to v \mid v \in V \setminus \{A, a_k, a_{k-1}, b_k\}\}$$

and

$$T_k = \{B \to F, a_k \to F, b_k \to a_k, b_{k-1} \to F\} \cup \{v \to v \mid v \in V \setminus \{B, a_k, b_k, b_{k-1}\}\}.$$

Then the construct $C = (V; U_1, \ldots, U_n, T_1, \ldots, T_n; a_1 \cdots a_n A)$ is called a *general carrier*.

The reader should note the following. Let

$$P = \{S \to a_1 \cdots a_n A\} \cup \{v \to v \mid v \in V\},$$

$\Delta_v = \{a_1, \ldots, a_n, A\}$ and $G = (V, \mathscr{R}, S, \Delta_v)$ be the $ET0L$ system, where $\mathscr{R} = \bigcup_{k=1}^{n} U_k \cup \bigcup_{k=1}^{n} T_k \cup P$. Then $L(G) = \{a_1 \cdots a_n A\}$ and the only "real way" to derive $a_1 \cdots a_n A$ in G (that is we consider only those sequences of tables that indeed rewrite the current word into something else than itself) is to start with P and then repeat any number of times the cycle $U_1 \cdots U_n T_1 \cdots T_n$.

(2) Let $C = (V; U_1, \ldots, U_n, T_1, \ldots, T_n; a_1 \cdots a_n A)$ be a general carrier, then the 0, 1-*extended carrier of* C is a construct $(V \cup \{0, 1\}; U_1^0, \ldots, U_n^0, U_1^1, \ldots, U_n^1, T_1^0, \ldots, T_n^0, T_1^1, \ldots, T_n^1; a_1 \cdots a_n A0)$ where $V \cap \{0, 1\} = \varnothing$ and $U_1^0 = U_1 \cup \{0 \to 0, 1 \to 1\}$, $U_1^1 = U_1 \cup \{0 \to 1, 1 \to 1\}$, for $2 \leqslant k \leqslant n$ $U_k^0 = U_k \cup \{0 \to 0, 1 \to F\}$ and $U_k^1 = U_k \cup \{0 \to F, 1 \to 1\}$, for $1 \leqslant k \leqslant n-1$ $T_k^0 = T_k \cup \{0 \to 0, 1 \to F\}$ and $T_k^1 = T_k \cup \{0 \to F, 1 \to 1\}$, $T_n^0 = T_n \cup \{0 \to 0, 1 \to F\}$ and $T_n^1 = T_n \cup \{1 \to 0, 0 \to F\}$.

THEOREM V.2. $\mathscr{L}(ET0L_{[8]}) = \mathscr{L}(ET0L)$.

Proof. Obviously $\mathscr{L}(ET0L_{[8]}) \subset \mathscr{L}(ET0L)$. To prove the converse inclusion we proceed as follows.

Let $K \in \mathscr{L}(ET0L)$. Since $\mathscr{L}(ET0L)$ is closed under intersection with regular languages (Rozenberg and Salomaa), $K = \bigcup_{i=1}^{s} K_i$, $s \geqslant 1$ where each K_i is an $ET0L$ language such that if $x, y \in K_i$ then $\text{alph}(x) = \text{alph}(y)$.

(1) Let us consider a fixed language K_i, $1 \leqslant i \leqslant s$, as above (say $K_i = L$). Let $H = (\Sigma, \mathscr{P}, S, \Delta)$ be an $ET0L$ system generating L. It is well known (Rozenberg and Salomaa) that we can assume that $\#\mathscr{P} = 2$ (say $\mathscr{P} = \{P_0, P_1\}$) and clearly we can assume that there exists a nonterminal (say N) such that N occurs in every intermediate word in every successful derivation in H.

Let t be a fixed terminal symbol occurring in every word of $L(H)$.

Let $\bar{\Sigma} = \{\bar{\sigma} \mid \sigma \in \Sigma\}$ and $\hat{\Sigma} = \{\hat{\sigma} \mid \sigma \in \Sigma\}$, where $\Sigma, \bar{\Sigma}, \hat{\Sigma}$ are pairwise disjoint.

Let $(V \cup \{0, 1\}; U_1^0, \ldots, U_n^0, U_1^1, \ldots, U_n^1, T_1^0, \ldots, T_n^0, T_1^1, \ldots, T_n^1; a_1 \cdots a_n A0)$ be a $\{0, 1\}$-extended carrier, where $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$ and V, $\{0, 1\}$, $\Sigma, \bar{\Sigma}, \hat{\Sigma}$ are pairwise disjoint.

Let $\mathscr{P}'$ be the following set of tables over the alphabet $\Sigma' = \Sigma \cup \bar{\Sigma} \cup \hat{\Sigma} \cup V \cup \{0, 1\} \cup \{S'\}$, S' being a new symbol. (In every of the tables given below we list only productions different from the identity productions.)

(i) $P_{in} = \{S' \to \bar{S}a_1 \cdots a_n A0\}$.

(ii) For $i = 1, \ldots, n$,

$$R_i^0 = U_i^0 \cup \{\bar{\sigma}_i \to \bar{\delta} \mid \sigma_i \to \delta \in P_0\} \cup \{t \to F\}$$
$$R_i^1 = U_i^1 \cup \{\bar{\sigma}_i \to \bar{\delta} \mid \sigma_i \to \delta \in P_1\} \cup \{t \to F\}$$
$$P_i^0 = T_i^0 \cup \{\hat{\sigma}_i \to \bar{\sigma}_i\} \cup \{t \to F\},$$

and

$$P_i^1 = T_i^1 \cup \{\hat{\sigma}_i \to \bar{\sigma}_i\} \cup \{t \to F\}.$$

(iii) $I_1 = \{A \to A, B \to F, b_1 \to F, b_n \to F, t \to t, \bar{N} \to F\}$.

(iv) For all $\delta \in \Delta$, $\delta \neq t$,

$$T^\delta = \{A \to F, B \to F, N \to F, \bar{\delta} \to \delta\}.$$

(v) For $i = 1, \ldots, n$,

$$T^i = \{A \to F, B \to F, N \to F, a_i \to A\}.$$

(vi) $T^0 = \{A \to F, B \to F, N \to F, 0 \to A\}$.

From the construction it follows that $H' = (\Sigma', \mathscr{P}', S', \Delta)$ is an 8-restricted $ET0L$ system. That $L(H') = L(H)$ is seen as follows. $S' \Rightarrow_{H'} \bar{S}a_1 \cdots a_n A0$ and then R_1^0 or R_1^1 have to be chosen, which is equivalent with a choice of P_0 or P_1 in H. The choice of R_1^0 (R_1^1 respectively) implies that next $R_2^0, \ldots, R_n^0$, $P_1^0, \ldots, P_n^0$($R_2^1, \ldots, R_n^1$, $P_1^1, \ldots, P_n^1$, respectively) have to be used in this order, thus deriving $\bar{w}a_1 \cdots a_n A0$ if $S \to w \in P_0$ ($S \to w \in P_1$ respectively). This way of simulating in H' direct derivation steps from H is iterated. Hence every derivation $S \overset{*}{\Rightarrow}_H z$ corresponds to $S' \Rightarrow_{H'} \bar{S}a_1 \cdots a_n A0 \overset{*}{\Rightarrow}_{H'} \bar{z}a_1 \cdots a_n A0$ in such a way that when, in H, P_0 (P_1 respectively) is used, then, in H', the cycle $R_1{}^0,..., R_n{}^0$, $P_1{}^0,..., P_n{}^0$ ($R_1{}^1,..., R_n{}^1$, $P_1{}^1,..., P_n{}^1$, respectively) is used. Since iterating those cycles is the only way to get a derivation that does not introduce the rejection symbol F, we may conclude that $S \overset{*}{\Rightarrow}_H z$ if and only if $S' \Rightarrow_{H'} \bar{z}a_1 \cdots a_n A0$.

The only way to get a terminal word in H' from $\bar{z}a_1 \cdots a_n A0$ is to use tables of type (iii) through (vi) on the condition that $z \in \Delta^*$; moreover the table I_1 must be used first.

Since I_1 rewrites $\bar{t}$ in t, $\bar{t} \in \text{alph}(\bar{z})$ and $t \to F$ is in every of the tables $R_i{}^j$ and $T_i{}^j$ for $j = 1, 2$ and $i = 1,..., n$, those tables cannot be used anymore.

Now using tables of type (iv) through (vi) we get the word z.

(2) Now let us return to the language K. We have $K = \bigcup_{i=1}^{s} K_i$.

Let each K_i be generated by an 8-restricted $ET0L$ system G_i constructed in the same way as H' was constructed for L in (1).

Let $G_i = (\Sigma^{(i)}, \mathscr{P}^{(i)}, S^{(i)}, \Delta^{(i)})$ for $i = 1,..., s$.

Let $H_i = (\Sigma_{(i)}, \mathscr{P}_{(i)}, S_{(i)}, \Delta_{(i)})$ result from G_i, $1 \leqslant i \leqslant s$, by renaming all symbols in G_i except for the special symbols A, B, F and N from (1) in such a way that each symbol δ from $\Delta^{(i)}$ in G_i becomes now $\delta_{(i)}$ in H_i and $\Sigma_{(i)} \cap \Sigma_{(j)} = \{A, B, F, N\}$ for $i \neq j$.

Finally let $G = (\Sigma_{(1)} \cup \cdots \cup \Sigma_{(s)} \cup \Delta \cup \{S\}, \mathscr{R}, S, \Delta)$ where

$$S \notin \Sigma_{(1)} \cup \cdots \cup \Sigma_{(s)} \cup \Delta$$

and $\mathscr{R}$ consists of the following tables.

(I) $\{S \to S_{(i)}\}$ is in $\mathscr{R}$ for $i = 1,..., s$.

(II) $\mathscr{P}_{(1)},..., \mathscr{P}_{(s)}$ are in $\mathscr{R}$.

(III) For every $1 \leqslant i \leqslant s$ and every $\delta_{(i)} \in \Delta_{(i)}$ $\{A \to F, B \to F, N \to F$ $\delta_{(i)} \to \delta\}$ is in $\mathscr{R}$.

Clearly $L(G) = K$ and G is an 8-restricted $ET0L$ system.

Hence the theorem holds. ∎

The above result is, in our opinion, an instructive result on the nature of parallel rewriting. It says that a parallel rewriting process (in the scope modelled by $ET0L$ systems) requires a bounded amount of "cooperation" between different symbols. That is, very $ET0L$ language can be generated by an $ET0L$ system in which in each rewriting step it suffices to rewrite only a bounded number of different symbols—not more than 8 of them. It is an interesting open problem to find out the lower bound on the amount of coorperation needed to generate the whole class of $ET0L$ languages. In Lemma IV.1 it is shown that to set this parameter equal to 1 is a real restriction, hence 1 is not the lower bound.

One should notice at this point that $ET0L_{[k]}$ systems form in a sense a generalization of $ET0L$ systems of index k, see Rozenberg and Vermeir (1975). Hence it is instructive to compare the above result with the result about $ET0L$ systems of index k, which says that increasing the index k leads to an infinite hierarchy of classes of languages, see Rozenberg and Vermeir (1975).

VI. *ET0LRP* Systems

In this section we study the effect of combining the mechanism of Russian parallel rewriting (see, e.g., Levitina (1972) and Salomaa (1974)) with the mechanism of $ET0L$ rewriting, in a fashion analogous to Section III where we have combined Indian parallel and $ET0L$ ways of rewriting.

We start by recalling the notion of a Russian parallel grammar.

DEFINITION. (1) A *composed set of productions over* an alphabet Σ is an ordered pair $P = (P_1, P_2)$ such that both P_1 and P_2 are finite sets of productions of the form $A \to \alpha$, $A \in \Sigma$, $\alpha \in \Sigma^*$ ($LH(P_1)$ and $LH(P_2)$ do not have to be disjoint). We refer to P_1 as the bounded part of P, denoted $\mathbf{bnd}(P)$, and to P_2 as the free part of P, denoted $\mathbf{fr}(P)$.

(2) A *Russian parallel grammar*, abbreviated *RP grammar*, is a construct $G = (\Sigma, P, S, \Delta)$, where Σ, P, S, Δ are as in the definition of a CF grammar (that is the total alphabet, the set of productions, the axiom and the terminal alphabet of G, respectively), except that P is a composed set of productions over $\Sigma\backslash\Delta$.

(3) Let $x \in \Sigma^*$ and $y \in \Sigma^*$. We say that x *directly derives* y *in* G, denoted $x \Rightarrow_G y$ if $x = x_0Ax_1 \cdots Ax_n$, where $A \in \Sigma\backslash\Delta$, $n \geqslant 1$, $x_0, \ldots, x_n \in \Sigma$ and $A \notin \mathrm{alph}(x_0 \cdots x_n)$, and either $y = x_0Ax_1 \cdots Ax_j\alpha x_{j+1}A \cdots Ax_n$ for some j, $0 \leqslant j \leqslant n-1$, and $A \to \alpha \in \mathbf{fr}(P)$ or $y = x_0\alpha x_1 \cdots \alpha x_n$ and $A \to \alpha \in \mathbf{bnd}(P)$.

(4) The relation $\overset{*}{\Rightarrow}_G$ is defined as the transitive and the reflexive closure of $\Rightarrow_G$. If $x \overset{*}{\Rightarrow}_G y$, then we say that x *derives* y *in* G.

(5) The *language of* G is defined by $L(G) = \{x \in \Delta^* \mid S \overset{*}{\Rightarrow} x\}$.

Combining the Russian parallel rewriting mechanism with $ET0L$ systems, we get the following construct.

DEFINITION. (1) Let Σ be an alphabet. A *composed table over* Σ is an ordered pair $P = (P_1, P_2)$ such that both P_1 and P_2 are finite sets of *productions* of the form $A \to \alpha$, $A \in \Sigma$, $\alpha \in \Sigma^*$ where $LH(P_1) \cup LH(P_2) = \Sigma$ (but $LH(P_1)$ and $LH(P_2)$ do not have to be disjoint). We refer to P_1 as the *bounded component of* P, denoted $\mathbf{bnd}(P)$, and to P_2 as the *free component of* P, denoted $\mathbf{fr}(P)$.

(2) A *Russian parallel* $ET0L$ *system*, abbreviated $ET0LRP$ *system*, is a

construct $G = (\Sigma, \mathscr{P}, S, \Delta)$ where $\Sigma, \mathscr{P}, S, \Delta$ are as in the definition of an $ET0L$ system except that P is a finite set of composed tables over Σ.

(3) Let $x \in \Sigma^+$ and $y \in \Sigma^*$. We say that x *directly derives* y *in* G, denoted $x \Rightarrow_G y$, if $x = x_1 \cdots x_n$ with $n \geq 1$ and $x_1, \ldots, x_n \in \Sigma$, $y = y_1 \cdots y_n$ with $y_1, \ldots, y_n \in \Sigma^*$, and there exist $P \in \mathscr{P}$, $P_1 \subset \mathbf{bnd}(P)$, $P_2 \subset \mathbf{fr}(P)$ with $LH(P_1) \cap LH(P_2) = \varnothing$ such that $x_i \to y_i \in P_1 \cup P_2$ for $1 \leq i \leq n$ and whenever $x_i \to y_i \in P_1$, $x_i = x_j$, $1 \leq i, j \leq n$ then $y_i = y_j$. The relation $\overset{*}{\Rightarrow}_G$ is defined as the transitive and the reflexive closure of $\Rightarrow_G$; if $x \overset{*}{\Rightarrow}_G y$ then we say that x *derives* y *in* G.

(4) The *language of* G, denoted $L(G)$, is defined by

$$L(G) = \{x \in \Delta^* \mid S \overset{*}{\Rightarrow} x\}.$$

Thus in an $ET0LRP$ system G a single rewriting step is performed as follows. Given a word x to be rewritten, one chooses *first* a composed table P, *then* one decides on letters in x all occurrences of which will be rewritten by productions in $\mathbf{bnd}(P)$ (hence in the "Indian parallel way") and *then* the other (occurrences of) letters in x will be rewritten by productions from $\mathbf{fr}(P)$ (hence in "normal $E0L$ fashion"). In this way in the framework of $ET0L$ systems, $ET0LRP$ systems play the role that RP grammars play in the framework of CF grammars.

All notations and terminology concerning $ET0L$ systems are carried over to $ET0LRP$ systems. Thus, e.g., an *$E0LRP$ system* is an $ET0LRP$ system $(\Sigma, \mathscr{P}, S, \Delta)$ where $\#\mathscr{P} = 1$. Also when we deal with an $ET0LRP$ system we will use the term "table" to refer to a composed table, this however should not lead to confusion.

First of all we demonstrate that augmenting $ET0L$ systems with the Russian parallel mechanism yields a class of rewriting systems generating precisely the class of $ET0L$ languages.

THEOREM VI.1. $\mathscr{L}(ET0L) = \mathscr{L}(ET0LRP)$.

Proof. (1) $\mathscr{L}(ET0L) \subset \mathscr{L}(ET0LRP)$. This is easily seen. Given an $ET0L$ system $G = (\Sigma, \mathscr{P}, S, \Delta)$ one constructs an $ET0LRP$ system $\bar{G}$ by taking for every table $P \in \mathscr{P}$ a composed table $\bar{P}$ to $\bar{G}$ where $\mathbf{bnd}(\bar{P}) = \varnothing$ and $\mathbf{fr}(\bar{P}) = P$. Clearly $L(\bar{G}) = L(G)$.

(2) To see that $\mathscr{L}(ET0LRP) \subset \mathscr{L}(ET0L)$ we proceed as follows. Let $G = (\Sigma, \mathscr{P}, S, \Delta)$ be an $ET0LRP$ system. For each $P \in \mathscr{P}$ let $Z(P)$ be the set of all composed tables of the form (T_1, T_2) where $T_1 \subset \mathbf{bnd}(P)$, T_1 is deterministic, $T_2 = \mathbf{fr}(P) \setminus \{A \to \alpha \mid A \to \alpha \in \mathbf{fr}(P)$ and $A \in LH(T_1)\}$ and $LH(T_1) \cup LH(T_2) = \Sigma$. Then let $\bar{G} = (\Sigma, \bar{\mathscr{P}}, S, \Delta)$ be the $ET0LRP$ system where $\bar{\mathscr{P}} = \bigcup_{P \in \mathscr{P}} Z(P)$.

Clearly $L(\bar{G}) = L(G)$ but $\bar{G}$ has the pleasant feature that, for every table T of $\bar{G}$, $\{LH(\mathbf{bnd}(T)), LH(\mathbf{fr}\ T))\}$ forms a partition of Σ. Now let $\bar{\Sigma} = \{\bar{a} \mid a \in \Sigma\}$, $\dot{\Sigma} = \{\dot{a} \mid a \in \Sigma\}$ and $\hat{\Sigma} = \{\hat{a} \mid a \in \Sigma\}$, where Σ, $\bar{\Sigma}$, $\dot{\Sigma}$ and $\hat{\Sigma}$ are pairwise disjoint.

Let $F \notin \Sigma \cup \bar{\Sigma} \cup \dot{\Sigma} \cup \hat{\Sigma}$ and $\Sigma' = \Sigma \cup \bar{\Sigma} \cup \dot{\Sigma} \cup \hat{\Sigma} \cup \{F\}$. Let $T_c = \{\bar{a} \to \dot{a} \mid a \in \Sigma\} \cup \{\bar{a} \to \hat{a} \mid a \in \Sigma\} \cup \{\sigma \to F \mid \sigma \in \Sigma' \setminus \bar{\Sigma}\}$. $T_{\text{fin}} = \{\bar{a} \to a \mid a \in \Delta\} \cup \{\sigma \to F \mid \sigma \in \Sigma' \setminus \Delta\}$, and for every $T \in \bar{\mathscr{P}}$, $R_T = \{\dot{a} \to \bar{\alpha} \mid a \to \alpha \in \mathbf{bnd}\ T)\} \cup \{\hat{a} \to \bar{\alpha} \mid a \to \alpha \in \mathbf{fr}(T)\} \cup \{\sigma \to F \mid \sigma \in \Sigma' \setminus (\dot{\Sigma} \cup \hat{\Sigma})\}$.

Finally let $H = (\Sigma', \mathscr{P}, \bar{S}, \Delta)$ be the $ET0L$ system with $\mathscr{R} = \{T_c, T_{\text{fin}}\} \cup U_{T \in \bar{\mathscr{P}}} R_n$. Note that each seccessful derivation from $\bar{G}$ is simulated in H in such a way that a single derivation step from $\bar{G}$ corresponding to an application of a table T is simulated by two derivation steps in H. The first step is an application of the "coordination table" T_c which divides letters in a string into those to be rewritten, in $\bar{G}$, by $\mathbf{bnd}(T)$ (they become elements of $\dot{\Sigma}$) and those to be rewritten, in $\bar{G}$, by $\mathbf{fr}(T)$ (thy become elements of $\hat{\Sigma}$). The second step rewrites elements from $\dot{\Sigma}$ by productions corresponding to $\mathbf{bnd}(T)$ and elements from $\hat{\Sigma}$ by productions corresponding to $\mathbf{fr}(T)$. Each successful derivation in H ends by an application of T_{fin} thus using the standard synchronization method. Hence clearly $L(G) = L(\bar{G}) = L(H)$.

(3) The theorem follows from (1) and (2). ▮

However the situation is different on the level of $E0L$ systems; that is, augmenting $E0L$ systems with the Russian parallel mechanism of rewriting yields a class of systems generating a class of languages strictly containing $\mathscr{L}(E0L)$.

THEOREM VI.2. $\mathscr{L}(E0L) \subsetneq \mathscr{L}(E0LRP)$.

Proof. The inclusion $\mathscr{L}(E0L) \subset \mathscr{L}(E0LRP)$ is obvious. It is well known that $L = \{w \in \{a, b\}^* \mid \#_a w = 2^n, n \geqslant 0\}$ is not an $E0L$ language (Ehrenfeucht and Rozenberg, 1974). However L is generated by the $E0LRP$ system $G = (\{S, a, b\}, P, S, \{a, b\})$, where $\mathbf{bnd}(P) = \{a \to aa\}$ and $\mathbf{fr}(P) = \{S \to a, a \to ab, a \to ba, a \to a, b \to b\}$. Thus the theorem holds. ▮

Before we proceed further in our investigation of $E0LRP$ languages, we notice the following about the class of RP languages.

In Salomaa (1974) the following is stated (Theorem 5). Assume that k_i, $i = 1, 2,...$ is a sequence of natural numbers, such that the set $\{a^{k_i} \mid i \geqslant 1\}$ is not regular. Then the language $L_k = \{a^{k_i} b^{k_i} \mid i \geqslant 1\}$ is not in $\mathscr{L}(RP)$. Consequently L_k is not in $\mathscr{L}(IP)$.

This theorem can be slightly generalized yielding the following result.

THEOREM VI.3. *Let* $\tau = (k_1, k_2, ...)$ *and* $\rho = (l_1, l_2, ...)$ *be infinite sequences of natural numbers, such that there exists a bijective function* f *from* $\{x \mid x$ *occurs in* $\tau\}$ *onto* $\{x \mid x$ *occurs in* $\rho\}$ *such that* $f(k_i) = l_i$ *for* $i = 1, 2, ...$ *. If either* $\{a^{k_i} \mid i \geqslant 1\}$ *or* $\{b^{l_i} \mid i \geqslant 1\}$ *is not regular, then* $L_{k,l} = \{a^{k_i} b^{l_i} \mid i \geqslant 1\}$ *is not in* $\mathscr{L}(RP)$.

Proof. Assume the contrary. Let $G = (\Sigma, P, S, \Delta) \in RP$ generate $L_{k,l}$.

Since $\{a^{k_i} \mid i \geq 1\}$ or $\{b^{l_i} \mid i \geq 1\}$ is not regular, $L_{k,l}$ is not context free. Then there is at least one nonterminal A in G with the properties

(i) $S \Rightarrow z_1Az_2Az_3$ for some $z_1, z_2, z_3 \in \Sigma^*$ and

(ii) there exist $x_1, x_2 \in \Delta^*$, $x_1 \neq x_2$, such that $A \overset{*}{\Rightarrow} x_1$ and $A \overset{*}{\Rightarrow} x_2$. (If such an A does not exist, then $L_{k,l} \in \mathscr{L}(CF)$).

Continue the rewriting from $z_1Az_2Az_3$ eliminating all nonterminals except A. Since we may assume that all nonterminals generate some terminal word, $z_1Az_2Az_3$ derives $y_1Ay_2A \cdots Ay_m$ in G, with $m \geq 3$ and $y_j \in \Delta^*$ for $j = 1, \dots, m$. Then both of the words $y_1x_1y_2x_1 \cdots x_1y_m$ and $y_1x_2y_2x_2 \cdots x_2y_m$ are in $L_{k,l}$.

Since $m \geq 3$, both x_1 and x_2 are words over a one letter alphabet and $\text{alph}(x_1) = \text{alph}(x_2)$.

Consequently $x_1 = x_2$; a contradiction. Thus the result holds. ∎

As a direct application of the above theorem we get the following example of a language that is not Russian parallel.

EXAMPLE VI.1. $L = \{a^nb^{2^n} \mid n \geq 0\} \notin \mathscr{L}(RP)$.

It is instructive at this point to contrast Theorem VI.3 with Corollary II.2 about IP languages.

We show now that the language generating power of $E0LRP$ systems is stronger than the language generating power of either RP grammars or $E0LIP$ systems.

THEOREM VI.4. $\mathscr{L}(RP) \subsetneq \mathscr{L}(E0LRP)$.

Proof. Let $G = (\Sigma, P, S, \Delta)$ be a RP system. Let $\Sigma\backslash\Delta = \{A_1, \dots, A_n\}$. Then $\Sigma^{(j)} = \{A^{(j)} \mid A \in \Sigma\backslash\Delta\}$ for $j = 1, \dots, n$ and Σ, $\Sigma^{(i)}$ and $\Sigma^{(j)}$ are pairwise disjoint if $i \neq j$, $1 \leq i, j \leq n$.

Let f_j for $j = 1, \dots, n$ be a homomorphism on Σ, defined by $f_j(A) = A^{(j+1)}$, $j = 1, \dots, n-1$, $f_n(A) = A^1$ for $A \in \Sigma\backslash\Delta$ and $f_j(a) = a$ if $a \in \Delta$, $1 \leq j \leq n$.

Let P' be a composed table of productions over $\Sigma' = \Delta \cup \bigcup_{j=1}^n \Sigma^{(j)}$, defined by

$$\mathbf{bnd}(P') = \{A_j^{(j)} \to f_j(w) \mid A_j \to w \in \mathbf{bnd}(P), j = 1, \dots, n\}$$

and

$$\mathbf{fr}(P') = \{A_j^{(j)} \to f_j(w) \mid A_j \to w \in \mathbf{fr}(P), j = 1, \dots, n\}$$
$$\cup \{A_i^{(j)} \to f_j(A_i) \mid 1 \leq i, j \leq n\} \cup \{a \to a \mid a \in \Delta\}.$$

It is easily seen that $H = (\Sigma', P', S', \Delta)$, where $S' = S^{(1)}$, is an $E0LRP$ system, which generates $L(G)$.

By example VI.1 $L = \{a^nb^{2^n} \mid n \geq 0\}$ is not a RP language. However it is easily seen that $L \in \mathscr{L}(E0L)$, and so by Theorem VI.2 L is an $E0LRP$ language.

Hence the theorem holds. ∎

THEOREM VI.5. $\mathscr{L}(E0LIP) \subsetneqq \mathscr{L}(E0LRP)$.

Proof. The inclusion follows immediately from the definitions of *E0LIP* systems and *E0LRP* systems.

It is well known that $\mathscr{L}(CF) \not\subset \mathscr{L}(EDT0L)$ (Ehrenfeucht and Rozenberg, 1977) and $\mathscr{L}(CF) \subsetneqq \mathscr{L}(E0L)$ (see, e.g. Rozenberg and Salomaa). Since $\mathscr{L}(E0LIP) \subset \mathscr{L}(EDT0L)$ (see Section III), it follows that $\mathscr{L}(E0LIP \subsetneqq \mathscr{L}(E0LRP)$. ∎

We move now to compare *E0LRP* systems with *EDT0L* systems. Our first result tells us that one can generate very *E0LRP* language by an *E0LRP* system in which all successful "computations" are organized in a way that reminds "computations" in an *ET0L* system with two tables.

DEFINITION. An *E0LRP* system $G = (\Sigma, P, S, \Delta)$ is said to be in *strong disjoint normal form* if $LH(\mathbf{bnd}(P)) \cap LH(\mathbf{fr}(P)) = \varnothing$ and each successful derivation D in G is such that at each step of D either only productions from $\mathbf{bnd}(P)$ are used or only productions from $\mathbf{fr}(P)$ are used and moreover applications of $\mathbf{bnd}(P)$ and of $\mathbf{fr}(P)$ alternate in D.

THEOREM VI.6. *For every E0LRP system G there exists an equivalent E0LRP system H in strong disjoint normal form.*

Proof. Let $G = (\Sigma, P, S, \Delta)$ be an *E0LRP* system. Let $\bar{\Sigma} = \{\bar{a} \mid a \in \Sigma\}$, $\hat{\Sigma} = \{\hat{a} \mid a \in \Sigma\}$, where Σ, $\bar{\Sigma}$ and $\hat{\Sigma}$ are pairwise disjoint, and let $\Sigma' = \Sigma \cup \bar{\Sigma} \cup \hat{\Sigma}$. Let P' be the composed table with $\mathbf{bnd}(P') = \{a \to \alpha \mid a \to \alpha \in \mathbf{bnd}(P)\} \cup \{a \to \hat{a} \mid a \in LH(\mathbf{fr}(P))\}$ and $\mathbf{fr}(P') = \{\hat{a} \to \alpha \mid a \to \alpha \in \mathbf{fr}(P)\} \cup \{\bar{a} \to a \mid a \in \Sigma\}$.

Let $H = (\Sigma', P', S, \Delta)$.

Clearly $L(H) = L(G)$ and H is in strong disjoint normal form. Hence the theorem holds. ∎

It is instructive to compare *E0LRP* systems in strong disjoint normal form with *ET0L* systems. An *E0LRP* system in strong disjoint normal form can be considered as an *ET0L* system with two tables one of which (the bounded part) is deterministic. It is well known that (see, e.g., Rozenberg and Salomaa) every *ET0L* language can be generated by an *ET0L* system with two tables only, one of which is deterministic. However an *E0LRP* system in strong disjoint normal form is using its "tables" in a very special (restrictive) way. In each successful derivation the application of the two tables must alternate. Although one can show that for every *ET0L* language K one can find a positive integer k and an *ET0L* system G with two tables T_1, T_2 (one of which is deterministic, T_1 say) such that G generates K and each successful derivation in G uses the tables T_1, T_2 in the fashion $T_1^{l_1} T_2 T_1^{l_2} T_2 \cdots T_1^{l_n} T_2$, where $n \geqslant 1$, $1 \leqslant l_1, ..., l_n \leqslant k$, it is not known whether or not one can set in the above $k = 1$ (we conjecture that not). If one can set $k = 1$ in the above, then we would get that $\mathscr{L}(E0LRP) = \mathscr{L}(ET0L)$; otherwise we would get $\mathscr{L}(E0LRP) \subsetneqq \mathscr{L}(ET0L)$.

Anyhow, we are not able to prove or to disprove the equation $\mathscr{L}(E0LRP) = \mathscr{L}(ET0L)$; we conjecture that $\mathscr{L}(E0LRP) \subsetneqq \mathscr{L}(ET0L)$. However we will demonstrate now that the class $\mathscr{L}(E0LRP)$ provides a quite elegant representation of the class $\mathscr{L}(ET0L)$.

THEOREM VI.7. *For every ET0L language K there exists an E0LRP language $\bar{K}$, a regular language R and a weak identity ϕ, such that $K = \phi(\bar{K} \cap R)$.*

Proof. Let $K \in \mathscr{L}(ET0L)$. We can assume that there exists an $ET0L$ system $G = (\Sigma, \mathscr{P}, S, \Delta)$ generating K such that $\mathscr{P} = \{T_1, T_2\}$ where T_1 is a deterministic table. Let $\bar{\Sigma} = \{\bar{a} \mid a \in \Sigma\}$ and $\hat{\Sigma} = \{\hat{a} \mid a \in \Sigma\}$ where Σ, $\bar{\Sigma}$ and $\hat{\Sigma}$ are pairwise disjoint, $\Sigma' = \Sigma \cup \bar{\Sigma} \cup \hat{\Sigma} \cup \{*, ¢, 1, 2, \$, £, F, S'\}$ where $\{*, ¢, 1, 2, \$, £, F, S'\} \cap (\Sigma \cup \bar{\Sigma} \cup \hat{\Sigma}) = \varnothing$, and let $\Delta' = \Delta \cup \hat{\Sigma} \cup \{*, 1, 2, \$, £\}$. Let P be a composed table such that its bounded component and its free component are defined by:

$$\begin{aligned}
\mathbf{bnd}(P) &= \{\bar{a} \to 1\$\bar{\alpha}\$1 \mid a \to \alpha \in T_1\}\\
\mathbf{fr}(P) &= \{\bar{a} \to 2\$\bar{\alpha}\$2 \mid a \to \alpha \in T_2\} \cup \{S' \to ¢\bar{S}\} \cup \{¢ \to \mathbf{per}(\bar{\Sigma})¢\}\\
&\cup \{\bar{a} \to a \mid a \in \Delta\} \cup \{\bar{a} \to \hat{a} \mid a \in \Sigma\} \cup \{a \to F \mid a \in \Sigma\}\\
&\cup \{\hat{a} \to F \mid \hat{a} \in \hat{\Sigma}\} \cup \{¢ \to £\} \cup \{£ \to F\} \cup \{\sigma \to \sigma \mid \sigma \in \{*, \$, 1, 2, F\}\},
\end{aligned}$$

where $\mathbf{per}(\bar{\Sigma})$ denotes the word $*\bar{a}_1 \cdots \bar{a}_n * \bar{a}_2\bar{a}_1 \cdots \bar{a}_n * \cdots *\bar{a}_n \cdots \bar{a}_1*$, which consists of all permutations of the elements of $\bar{\Sigma}$ separated by $*$.

Let $H = (\Sigma', P, S', \Delta')$ be an $E0LRP$ system.

(1) Note that if x is a word such that $¢\bar{S} \Rightarrow_H^+ x$ and $£ \notin \text{alph}(x)$ then $\mathbf{per}(\bar{\Sigma})$ is a subword of x, and so for all $\bar{a}, \bar{b} \in \bar{\Sigma}$, where $\bar{a} \neq \bar{b}$, both $\bar{a}\bar{b}$ and $\bar{b}\bar{a}$ are subword of x. But $\bar{a}\bar{b} \Rightarrow_H 1\$\bar{\alpha}\$12\$\bar{\beta}\$2$ if and only if $a \to \alpha \in T_1$ and $b \to \beta \in T_2$. Thus if $S' \Rightarrow_H ¢\bar{S} \Rightarrow_H^{\geqslant 2} w \in L(H)$ then w contains the subword 12 or 21 if and only if in one direct derivation step of such a derivation a rule of the form $\bar{a} \to 1\$\bar{\alpha}\$1$ and a rule of the form $\bar{b} \to 2\$\bar{\beta}\$2$ have been applied.

(2) Now let R be the regular language defined by $R = (\hat{\Sigma} \cup \{*, 1, 2, \$\})^* £(\Delta \cup \{1, 2, \$\})^* \setminus \{w \mid$ either w contains the subword 12 or w contains the subword 21$\}$ and let ϕ be a weak identity on Δ' defined by $\phi(a) = a$ if $a \in \Delta$ and $\phi(a) = \Lambda$ if $a \in \Delta' \setminus \Delta$. Then (1) implies that $L(G) = \phi(L(H) \cap R)$ and so the theorem holds. ▮

VII. THE RELATIONSHIP DIAGRAM

The aim of this section is to establish the relationship diagram between various classes of languages considered in this paper.

First of all to construct the relationship diagram we can use the following known results.

LEMMA VII.1. (1) $\mathscr{L}(CF_{\text{fin}}) \subsetneqq \mathscr{L}(CF)$ (see, e.g., Salomaa, 1973),

(2) $\mathscr{L}(CF_{\text{fin}}) \subsetneqq \mathscr{L}(IP)$ (see, e.g., Skyum, 1974),

(3) $\mathscr{L}(CF)$ and $\mathscr{L}(IP)$ are incomparable but not disjoint, (see, e.g., Skyum, 1974),

(4) $\mathscr{L}(CF) \subsetneqq \mathscr{L}(RP)$ (see, e.g., Levitina, 1972),

(5) $\mathscr{L}(CF) \subsetneqq \mathscr{L}(E0L)$ (see, e.g., Rozenberg and Salomaa),

(6) $\mathscr{L}(CF) \not\subset \mathscr{L}(EDT0L)$ (see, e.g., Ehrenfeucht and Rozenberg, 1977),

(7) $\mathscr{L}(EDT0L) \subsetneqq \mathscr{L}(ET0L)$ (see, e.g., Rozenberg and Salomaa),

(8) $\mathscr{L}(E0L)$ and $\mathscr{L}(EDT0L)$ are incomparable but not disjoint (see, e.g., Rozenberg and Salomaa).

Then in addition to results established in previous sections we also need the following results.

LEMMA VII.2. $\mathscr{L}(IP) \subsetneqq \mathscr{L}(RP)$.

Proof. The inclusion $\mathscr{L}(IP) \subset \mathscr{L}(RP)$ is an immediate consequence of the definitions of IP grammars and RP grammars.
That it is strict follows from Lemma VII.1 points (3) and (4). ∎

LEMMA VII.3. (1) *$\mathscr{L}(IP)$ and $\mathscr{L}(E0L)$ are incomparable but not disjoint.*

(2) *$\mathscr{L}(RP)$ and $\mathscr{L}(E0L)$ are incomparable but not disjoint.*

Proof. (1) It is known that $L = \{¢w¢w¢w \mid w \in \{0, 1\}^*\}$ is not an $E0L$ language (Rozenberg and Salomaa). Since the IP grammar $(\{S, C, ¢, 0, 1\}, \{S \to CCC, C \to C0, C \to C1, C \to ¢\}, S, \{¢, 0, 1\})$ generates L, it is clear that $\mathscr{L}(IP)$ is not contained in $\mathscr{L}(E0L)$. The first part of the lemma follows then from Lemma VII.1 points (1), (2), (3) and (5).

(2) From (1) and Lemma VII.2 it follows that $\mathscr{L}(RP)$ is not contained in $\mathscr{L}(E0L)$. Since $\{a^n b^n c^n \mid n \geq 0\}$ is an $E0L$ language and not a RP language (Levitina, 1972) and $\mathscr{L}(CF_{\text{fin}}) \subset \mathscr{L}(E0L) \cap \mathscr{L}(RP)$ the second statement of the lemma holds. ∎

LEMMA VII.4. (1) $\mathscr{L}(CF) \subsetneqq \mathscr{L}(ET0L_{[1]})$.

(2) $\mathscr{L}(RP) \not\subset \mathscr{L}(EDT0L_{[1]})$.

(3) $\mathscr{L}(EDT0L_{[1]} \subsetneqq \mathscr{L}(ET0L_{[1]})$.

Proof. (1) Clearly

$$\mathscr{L}(CF) \subset \mathscr{L}(ET0L_{[1]}) \quad \text{and} \quad \{a^{2^n} \mid n \geq 0\} \in \mathscr{L}(ET0L_{[1]}) \backslash \mathscr{L}(CF).$$

(2) and (3) From (1) and Lemma VII.1 points (4) and (6) the second and third statement of this lemma follow immediately. ▮

LEMMA VII.5. (1) $\mathscr{L}(IP) \subset \mathscr{L}(EDT0L_{[1]})$.

(2) *$\mathscr{L}(E0L)$ and $\mathscr{L}(EDT0L_{[1]})$ are incomparable but disjoint.*

Proof. (1) and (2). Obvious. ▮

LEMMA VII.6. *$\mathscr{L}(E0L)$ and $\mathscr{L}(ET0L_{[1]})$ are incomparable but not disjoint.*

Proof. It is proved in Section IV that $L = \{a^{2^n}b^{2^n} \mid n \geqslant 0\} \notin \mathscr{L}(ET0L_{[1]})$. It is easy to see that $L \in \mathscr{L}(E0L)$, so $\mathscr{L}(E0L)$ is not contained in $\mathscr{L}(ET0L_{[1]})$. On the other hand from the previous result it follows that $\mathscr{L}(ET0L_{[1]})$ is not contained in $\mathscr{L}(E0L)$.

Hence the result holds. ▮

LEMMA VII.7. $\mathscr{L}(RP) \subset \mathscr{L}(ET0L_{[1]})$.

Proof. Let $G = (\Sigma, P, S, \Delta)$ be a RP grammar. Let P be given by the following tables.

If $A \to w \in \mathbf{bnd}(P)$ then $\mathscr{P}$ contains a table $\{A \to w\} \cup \{a \to a \mid a \in \Sigma \setminus \{A\}\}$ and if $A \to w \in \mathbf{fr}(P)$ then $\mathscr{P}$ contains a table $\{A \to w, A \to A\} \cup \{a \to a \mid a \in \Sigma \setminus \{A\}\}$. ($\mathscr{P}$ consists of these tables only.) Clearly the $ET0L_{[1]}$ system $(\Sigma, \mathscr{P}, S, \Delta)$ generates $L(G)$. ▮

LEMMA VII.8. *The following pairs of families of languages are incomparable but not disjoint.*

(1) *$\mathscr{L}(RP)$ and $\mathscr{L}(E0LIP)$,*

(2) *$\mathscr{L}(RP)$ and $\mathscr{L}(EDT0L)$,*

(3) *$\mathscr{L}(ET0L_{[1]})$ and $L(E0LIP)$,*

(4) *$\mathscr{L}(ET0L_{[1]})$ and $\mathscr{L}(EDT0L)$.*

Proof. This follows from $\mathscr{L}(E0LIP) \subset \mathscr{L}(EDT0L)$, Lemma VII.1 points (4) and (6) the previous lemma and Examples IV.1 and VI.1. ▮

LEMMA VII.9. $\mathscr{L}(EDT0L_{[1]}) \subsetneqq \mathscr{L}(E0LIP)$.

Proof. Let $G = (\Sigma, \mathscr{P}, S, \Delta)$ be an $EDT0L_{[1]}$ system, where $\mathscr{P} = \{T_1,..., T_n\}$. Let $\Sigma^{(j)} = \{a^{(j)} \mid a \in \Sigma\}$ for $j = 1,..., n$ where Σ, $\Sigma^{(i)}$ and $\Sigma^{(j)}$ are pairwise disjoint if $i \neq j$, $1 \leqslant i, j \leqslant n$. Let f_j, $j = 1,..., n$ be a homomorphism on Σ, defined by $f_j(a) = a^{(j+1)}$ if $1 \leqslant j \leqslant n - 1$ and $f_n(a) = a^{(1)}$. Let P be the following table of productions over $\Delta \cup \bigcup_{j=1}^{n} \Sigma^{(j)} \cup \{F\}$, where $F \notin \Sigma \cup \bigcup_{j=1}^{n} \Sigma^{(j)}$:

$$P = \{a^{(j)} \to f_j(w) \mid a \to w \in T_j, j = 1,..., n\} \cup \{a^{(j)} \to f_j(a) \mid a \in \Sigma, j = 1,..., n\} \cup \{a^{(j)} \to a \mid a \in \Delta, j = 1,..., n\} \cup \{a \to F \mid a \in \Delta \cup \{F\}\}.$$

Let $H = (\Sigma', P, S', \Delta)$ be the $E0LIP$ system, where $\Sigma' = \Delta \cup \bigcup_{j=1}^{n} \Sigma^{(j)} \cup \{F\}$ and $S' = S^{(1)} \in \Sigma^{(1)}$.

Clearly H generates $L(G)$.

Since $\{a^{2^n}b^{2^n} \mid n \geqslant 0\} \in \mathscr{L}(E0LIP) \setminus \mathscr{L}(EDT0L_{[1]})$ the lemma holds. ∎

LEMMA VII.10. $\mathscr{L}(ET0L_{[1]}) \subsetneqq \mathscr{L}(E0LRP)$.

Proof. The proof of the inclusion is analogous to the proof of the previous theorem, except that now we set

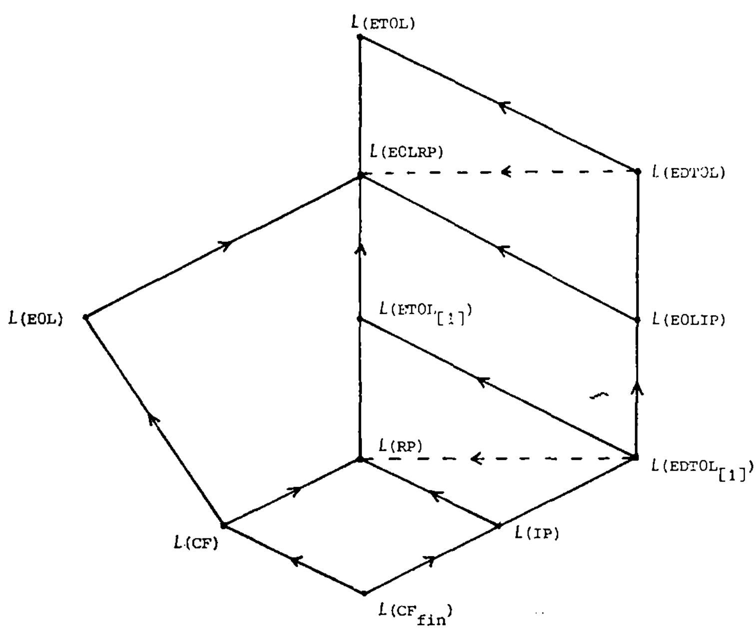
$$\mathbf{fr}(P) = \{a^{(j)} \to f_j(w) \mid a \to w \in T_j, j = 1,..., n\} \text{ and}$$

$$\mathbf{bnd}(P) = \{a^{(j)} \to f_j(a) \mid a \in \Sigma, j = 1,..., n\} \cup \{a^{(j)} \to a \mid a \in \Delta, j = 1,..., n\} \cup \{a \to F \mid a \in \Delta \cup \{F\}\}.$$

Since $\{a^{2^n}b^{2^n} \mid n \geqslant 0\} \in \mathscr{L}(E0LRP) \setminus \mathscr{L}(ET0L_{[1]})$ the lemma holds. ∎

Finally, combining all those comparison results we have, we get the following theorem.

THEOREM VII.1. *The following diagram holds*:



(If there is a directed chain of edges in the diagram leading from a class X to a class Y then $X \subsetneqq Y$, an undirected chain means that we do not know whether the inclusion is proper. A dotted directed edge leading from a class X to a class Y means that we do not know whether $X \subset Y$, but we do know that $Y \not\subset X$. Otherwise X and Y are incomparable but not disjoint.)

References

Dassow, D. (1979), On some extensions of Indian parallel context free grammars, Department of Mathematics, Wilhelm-Pieck-University, Rostock, GDR.

Ehrenfeucht, A., and Rozenberg, G. (1974), The number of occurrences of letters versus their distribution in some EOL languages, *Inform. Contr.* **26**, 256–271.

Ehrenfeucht, A., and Rozenberg, G. (1977), On some context free languages that are not deterministic ETOL languages, *R.A.I.R.O. Inform. Théor.* **11**, 273–291.

Levitina, M. K. (1972), O nekotoryh grammatikah s pravilami globalnoi podstanovki, *Automat. Perevoda Tekstov Ser. 2*, No. 3, 32–36 (In Russian).

Rozenberg, G., and Salomaa, A., "The Mathematical Theory of L Systems," Academic Press, New York, in press.

Rozenberg, G., and Vermeir, D. (1975), "On ETOL Systems of Finite Index," Dept. of Math., Univ. of Antwerp, U.I.A., Tech. Rept. 75-13.

Salomaa, A. (1973), "Formal Languages," Academic Press, New York/London.

Salomaa, A. (1974), Parallelism in rewriting systems, *in* "Automata, Languages and Programming," Lecture Notes in Computer Science No. 14, pp. 223–233, Springer-Verlag, Berlin/New York.

Skyum, S. (1974), Parallel context free languages. *Inform. Contr.* **26**, 280–285.

Siromoney, R., and Krithivasan, K. (1974), Parallel context-free languages, *Inform. Contr.* 24, 155–162.

Siromoney, G., and Siromoney, R. (1975–1976), Parallel 0L languages, *Internat. J. Comput. Math.* **5**, 109–123.