Archivum Mathematicum

Miroslav Novotný

On the role of configurations in the theory of grammars

Archivum Mathematicum, Vol. 6 (1970), No. 3, 171--184

Persistent URL: http://dml.cz/dmlcz/104720

ON THE ROLE OF CONFIGURATIONS IN THE THEORY OF GRAMMARS

MIROSLAV NOVOTNÝ, Brno

(Received May 19, 1969)

INTRODUCTION

In [1] I found a new characterization of context-free languages: Every context-free language is the intersection of a full language and a trace of a language of strong depth 1. A full language over the set V is a language containing all strings over V. Languages of strong depth 1 can be defined by means of strong configurations of order 1: They are languages over a finite vocabulary for which the set of strings containing no strong configuration of order 1 is finite and the set of all so called simple strong configurations of order 1 is finite, too. The trace of a language in a free monoid U^* is the language which can be obtained by cancelling all symbols not belonging to U in each string of the given language.

The aim of the present paper is to find a similar characterization of languages of the type 0 of the classification of Chomsky. We prove that every language of the type 0 is the intersection of a full language and a trace of a language finitely generated. A finitely generated language is a language over a finite vocabulary for which a natural number n exists such that each string of this language of length $> n$ contains a weak configuration of order 1 of length $\leqq n$.

Thus we see that finitely generated languages form a kernel from which the class of all languages of the type 0 can be obtained by means of two operations: trace and intersection. In the theory of finitely generated languages the main concept is that of a weak configuration of order 1 which appeared in the literature earlier; this concept was studied in [2] under the name of the configuration of order 1. The idea of using configurations to construct grammars is due to Gladkij [3].

1. GENERALIZED GRAMMARS AND GRAMMARS

If V is a set, we denote by V^* the free monoid over V, i.e. the set of all finite sequences of elements of V in which the operation of concatenation is defined; we suppose that the empty sequence Λ is an element of V^*, too. We identify one-element-sequences with elements of V; thus we have $V \subseteq V^*$ and for every natural number k and for $x_1, x_2, \ldots, x_k$ we write $x_1x_2 \ldots x_k$ instead of $(x_1, x_2, \ldots, x_k)$. The elements of V are called symbols, the elements of V^* strings.

We put $|\Lambda| = 0$. If $x \in V^*$, $x = x_1x_2 \dots x_n$ where n is a natural number and $x_i \in V$ for $i = 1, 2, \dots, n$, then we put $|x| = n$.

If n is a natural number and $A_i \subseteq V^*$ for $i = 1, 2, \dots, n$, then we put $A_1A_2 \dots A_n = \{a_1a_2 \dots a_n;\ a_i \in A_i,\ i = 1, 2, \dots, n\}$.

1.1. Definition. Let V, U be sets, let f be a mapping of the set V into U^*. We put $f_*(\Lambda) = \Lambda$; if $x = x_1x_2 \dots x_n$ where n is a natural number and $x_i \in V$ for $i = 1, 2, \dots, n$, then we put $f_*(x) = f(x_1)f(x_2) \dots \dots f(x_n)$.

1.2. Remark. If V, U are sets and f a mapping of V into U^*, then $f_*(xy) = f_*(x)f_*(y)$ for every $x, y \in V^*$.

1.3. Definition. Let V be a set, $L \subseteq V^*$; then the pair (V, L) is called a *language*.

1.4. Definition. Let V be a set. Then the language (V, V^*) is called *full*.

1.5. Definition. Let (V, L), (U, M) be languages. Then the language $(V \cap U, L \cap M)$ is called the *intersection of the languages* (V, L), (U, M).

1.6. Definition. Let V, V_T, S, R be sets with the properties $V_T \subseteq V$, $S \subseteq V^*$, $R \subseteq V^* \times V^*$. Then the quadruple $G = \langle V, V_T, S, R\rangle$ is called a *generalized grammar*. The elements of R are called *rules*.

1.7. Definition. Let $G = \langle V, V_T, S, R\rangle$ be a generalized grammar. We write, for $x, y \in V^*$, $x \to y\ (G)$ instead of $(x, y) \in R$. For $x, y \in V^*$, we write $x \Rightarrow y\ (G)$ iff there exist such strings $u, v, t, z \in V^*$ that $x = utv$, $uzv = y$, $t \to z\ (G)$. For $x, y \in V^*$ we write $x \overset{*}{\Rightarrow} y\ (G)$ iff there exist a non-negative integer p and some strings $t_0, t_1, \dots, t_p$ of V^* such that $x = t_0$, $t_p = y$ and $t_{i-1} \Rightarrow t_i(G)$ for $i = 1, 2, \dots, p$. The sequence $t_0, t_1, \dots, t_p$ is called an *x-derivation of y in G*, p is called the *length of* this *x-derivation*.

We put

$\mathscr{L}(G) = \{x;\ x \in V_T^*$ and there exists some $s \in S$ with the property $s \overset{*}{\Rightarrow} x\ (G)\}$.

The language $(V_T, \mathscr{L}(G))$ is called the *language generated by G*.

1.8. Definition. Let $G = \langle V, V_T, S, R\rangle$ be a generalized grammar. If $V = V_T$, then this generalized grammar is called a *special generalized grammar*. We write $\langle V, S, R\rangle$ instead of $\langle V, V, S, R\rangle$ if $\langle V, V, S, R\rangle$ is a special generalized grammar.

1.9. Definition. Let $G = \langle V, V_T, S, R\rangle$ be a generalized grammar. Then G is called a *grammar* iff the sets V, S, R are finite.

1.10. Remark. From the above definitions it is clear what is meant by a special grammar.

1.11. Definition. Let (V, L) be a language. This language is called a *special language* iff there exists a special grammar generating (V, L).

1.12. Definition. Let $G = \langle V, V_T, S, R\rangle$ be a grammar. This grammar is called a *phrase structure grammar* iff the following conditions are satisfied: (1) There exists such an element $\sigma \in V - V_T$ that $S = \{\sigma\}$. (2) If $(x, y) \in R$ then $\Lambda \neq x \in (V - V_T)^*$ (see [4], [5]).

1.13. Definition. Let (V, L) be a language. This language is called a *language of the type 0* iff there exists a phrase structure grammar $G = \langle W, V, \{\sigma\}, R\rangle$ which generates (V, L).

1.14. Definition. Let $G = \langle V, V_T, \{\sigma\}, R\rangle$ be a phrase structure grammar. Then G is called a *grammar having the standard form* iff the following condition is satisfied: If $(x, y) \in R$ then $|x| \leqq |y| + 1$.

1.15. Lemma. *Let (V, L) be a language of the type 0. Then there exists a grammar $\langle W, V, \{\sigma\}, R\rangle$ having the standard form which generates (V, L).*

Proof. Let $G = \langle U, V, \{\sigma\}, Q\rangle$ be a phrase structure grammar generating (V, L). If G has not the standard form then there exists a rule $(x, y) \in Q$ having the following property (S): $|x| > |y| + 1$. We put $|x| = m$, $|y| = n$, $x = x_1x_2 \ldots x_m$, $y = y_1y_2 \ldots y_n$ where $x_i \in U - V$ for $i = 1, 2, \ldots, m$, $y_j \in U$ for $j = 1, 2, \ldots, n$. We have $m > n + 1$; we put $q = m - n - 1$. We take new and mutually distinct elements $c_{1,1}, c_{1,2}, \ldots, c_{1,m-1}, c_{2,1}, c_{2,2}, \ldots, c_{2,m-2}, \ldots, c_{q,1}, c_{q,2}, \ldots, c_{q,n+1}$ and we define $W_1 = U \cup \{c_{i,j};\ i = 1, 2, \ldots, q,\ j = 1, 2, \ldots, m - i\}$, $Q_0 = \{(x, c_{1,1}c_{1,2} \ldots c_{1,m-1}), (c_{1,1}c_{1,2} \ldots c_{1,m-1}, c_{2,1}c_{2,2} \ldots c_{2,m-2}), \ldots, (c_{q,1}c_{q,2} \ldots c_{q,n+1}, y)\}$, $Q_1 = (Q - \{(x, y)\}) \cup Q_0$, $G_1 = \langle W_1, V, \{\sigma\}, Q_1\rangle$.

Clearly $x \overset{*}{\Rightarrow} y\ (G_1)$. It follows $\mathscr{L}(G) \subseteq \mathscr{L}(G_1)$.

Let us suppose $\sigma \overset{*}{\Rightarrow} z\ (G_1)$, $z \in V^*$. There exists a σ-derivation $\sigma = t_0, t_1, \ldots, t_p = z$ of z in G_1. If no rule of the set Q_0 has been applied then we have $\sigma \overset{*}{\Rightarrow} z\ (G)$. Let us suppose that a rule of Q_0 has been applied in the above derivation; let us suppose that r, $0 < r \leqq p$ is the least index such that t_r has been derived from t_{r-1} by means of a rule of Q_0. Then $t_{r-1} = uxv$, $t_r = uc_{1,1}c_{1,2} \ldots c_{1,m-1}v$ for some $u, v \in U^*$. It is easy to see that we can construct such a σ-derivation $\sigma = t'_0, t'_1, \ldots, t'_p = z$ of z in G_1 that $t'_i = t_i$ for $i = 0, 1, \ldots, r$, $t'_{r-1} = uxv$, $t'_r = uc_{1,1}c_{1,2} \ldots c_{1,m-1}v$, $t'_{r+1} = uc_{2,1}c_{2,2} \ldots c_{2,m-2}v$, $\ldots$, $t'_{r+q-1} = uc_{q,1}c_{q,2} \ldots c_{q,n+1}v$, $t'_{r+q} = uyv$. Thus, $\sigma \overset{*}{\Rightarrow} t'_{r+q}(G)$. By repeating this argument we prove $\sigma \overset{*}{\Rightarrow} z\ (G)$. Thus $\mathscr{L}(G_1) \subseteq \mathscr{L}(G)$.

We have proved $\mathscr{L}(G) = \mathscr{L}(G_1)$. The number of rules in Q_1 having the property (S) is less than the number of such rules in Q. By repeating this procedure we construct, after a finite number of steps, a phrase structure grammar $H = \langle W, V, \{\sigma\}, R\rangle$ such that $\mathscr{L}(H) = \mathscr{L}(G)$ and that no rule of H has the property (S), i.e. H has the standard form.

1.16. Lemma. *Let (V, L) be a language. This language is of the type 0 iff it is the intersection of a special and a full language.*

Proof. If (V, L) is a language of the type 0, then there exists a phrase structure grammar $G = \langle W, V, \{\sigma\}, R\rangle$ generating (V, L). We put $H = \langle W, \{\sigma\}, R\rangle$; then H is a special grammar and $\mathscr{L}(G) = \mathscr{L}(H) \cap V^*$. Thus, $(V, L) = (V, \mathscr{L}(G)) = (W \cap V, \mathscr{L}(H) \cap V^*)$. Thus (V, L) is the intersection of the special language $(W, \mathscr{L}(H))$ and the full language (V, V^*).

Let (V, L) be the intersection of the special language (U, M) and the full language (W, W^*). We can clearly suppose that $W \subseteq U$; therefore $V = W \cap U = W \subseteq U$. There exists a special grammar $G = \langle U, S, R \rangle$ generating (U, M). It follows that $H = \langle U, V, S, R \rangle$ is a grammar generating (V, L). It is well known (cf. e.g. [2], p. 95) that for every grammar there exists a phrase structure grammar which generates the same language. Thus (V, L) can be generated by a phrase structure grammar and therefore (V, L) is of the type 0.

1.17. Lemma. *Let $G = \langle V, V_T, \{\sigma\}, R \rangle$ be a phrase structure grammar, U an arbitrary set. Then there exists such a phrase structure grammar $H = \langle W, V_T, \{\tau\}, P \rangle$ that $\mathscr{L}(G) = \mathscr{L}(H)$* and $(W - V_T) \cap U = \varnothing$.

Proof. Let A be an arbitrary set equivalent with $V - V_T$ with the properties $A \cap U = \varnothing$, $A \cap V_T = \varnothing$, b a bijection of $V - V_T$ onto A; we put $b(t) = t$ for each $t \in V_T$. Thus b is a bijection of V onto $A \cup V_T$. We put $W = A \cup V_T$, $\tau = b(\sigma)$ and $P = \{(b_*(x), b_*(y));\ (x, y) \in R\}$ where b_* is defined according to 1.1. We define $H = \langle W, V_T, \{\tau\}, P \rangle$. Clearly, $\mathscr{L}(G) = \mathscr{L}(H)$ and $(W - V_T) \cap U = A \cap U = \varnothing$.

1.18. Remark. Let (V, L) be a language of the type 0, V_1 an arbitrary finite set. Then $(V \cup V_1, L)$ is a language of the type 0. — Indeed there exists a phrase structure grammar $G = \langle W, V, \{\sigma\}, R \rangle$ such that $\mathscr{L}(G) = L$. We can suppose $(W - V) \cap V_1 = \varnothing$ according to 1.17. We define $H = \langle W \cup V_1, V \cup V_1, \{\sigma\}, R \rangle$. Clearly, $\mathscr{L}(G) \subseteq \mathscr{L}(H)$. Let us have $x \in \mathscr{L}(H)$. Then $x \in (V \cup V_1)^*$, $\sigma \overset{*}{\Rightarrow} x\ (H)$. It implies $x \in (V \cup V_1)^*$, $\sigma \overset{*}{\Rightarrow} x\ (G)$, thus $x \in W^*$. Therefore, $x \in W^* \cap (V \cup V_1)^* = ((W - V) \cup \cup V)^* \cap (V \cup V_1)^* = V^*$ and $x \in \mathscr{L}(G)$. Thus, $L = \mathscr{L}(G) = \mathscr{L}(H)$ and $(V \cup V_1, \mathscr{L}(H)) = (V \cup V_1, L)$ is a language of the type 0.

1.19. Remark: The intersection of a language of the type 0 with a full language is a language of the type 0.

Indeed, if (V, L) is a language of the type 0, then there exists a special language (U, M) and a full language (W, W^*) such that (V, L) is the intersection of (U, M) and (W, W^*) according to 1.16. Let (Z, Z^*) be a full language. Then the intersection $(W \cap Z, W^* \cap Z^*)$ is the full language $(W \cap Z, (W \cap Z)^*)$ and we have $(V \cap Z, L \cap Z^*) = (U \cap W \cap Z, M \cap W^* \cap Z^*) = (U \cap (W \cap Z), M \cap (W \cap Z)^*)$ which is the intersection of the special language (U, M) and the full language $(W \cap Z, (W \cap Z)^*)$. Thus the intersection of (V, L) and (Z, Z^*) is a language of the type 0 according to 1.16.

1.20. Definition. Let P be a linearly ordered set, V a set, let us suppose $V_T \subseteq V$, $S \subseteq V^*$, $R_\lambda \subseteq V^* \times V^*$ for each $\lambda \in P$. Let us suppose that the sets R_λ are mutually disjoint. Then the quadruple $G = \langle V, V_T, S, (R_\lambda)_{\lambda \in P} \rangle$ is called a *generalized grammar with a linearly ordered decomposition on the set of rules* (*a generalized o-grammar*). If the sets V, S, $\bigcup_{\lambda \in P} R_\lambda$

are finite, then the quadruple $\langle V, V_T, S, (R_\lambda)_{\lambda\in P}\rangle$ is called a *grammar with a linearly ordered decomposition on the set of rules* (*an o-grammar*). The pairs $(x, y) \in R_\lambda$ are called *rules*.

1.21. Definition. Let $G = \langle V, V_T, S, (R_\lambda)_{\lambda\in P}\rangle$ be a generalized o-grammar. Let us have $\lambda_0 \in P$. Then for $x, y \in V^*$ we put $x \underset{R_{\lambda_0}}{\rightarrow} y$ iff $(x, y) \in R_{\lambda_0}$. For $x, y \in V^*$ we put $x \underset{R_{\lambda_0}}{\Rightarrow} y$ iff there exist such strings $u, v, t, z \in V^*$ that $x = utv$, $uzv = y$, $t \underset{R_{\lambda_0}}{\rightarrow} z$. For $x, y \in V^*$ we put $x \overset{*}{\underset{R_{\lambda_0}}{\Rightarrow}} y$ iff there exist an integer $q \geqq 0$ and some strings $t_0, t_1, \ldots, t_q$ in V^* such that $x = t_0$, $t_q = y$ and $t_{i-1} \underset{R_{\lambda_0}}{\Rightarrow} t_i$ for $i = 1, 2, \ldots, q$.

For $x, y \in V^*$ we define $x \overset{*}{\underset{(R_\lambda)_{\lambda\in P}}{\Rightarrow}} y$ iff there exist such an integer $p \geqq 0$, such a finite increasing sequence $\lambda_1 < \lambda_2 < \ldots < \lambda_p$ of elements of P and such elements $t_0, t_1, \ldots, t_p$ of V^* that $x = t_0$, $t_p = y$ and $t_{i-1} \overset{*}{\underset{R_{\lambda_i}}{\Rightarrow}} t_i$ for $i = 1, 2, \ldots, p$.

We put

$$\mathscr{L}(G) = \{x;\ x \in V_T^* \text{ and there exists such } s \in S \text{ that } s \overset{*}{\underset{(R_\lambda)_{\lambda\in P}}{\Rightarrow}} x\}.$$

The language $(V_T, \mathscr{L}(G))$ is called the *language generated by the generalized o-grammar G*.

1.22. Definition. Let $G = \langle V, V_T, S, (R_\lambda)_{\lambda\in P}\rangle$ be a generalized o-grammar. This generalized o-grammar is called *special* iff $V = V_T$. We write $\langle V, S, (R_\lambda)_{\lambda\in P}\rangle$ instead of $\langle V, V, S, (R_\lambda)_{\lambda\in P}\rangle$ if $\langle V, V, S, (R_\lambda)_{\lambda\in P}\rangle$ is a special generalized o-grammar.

1.23 Remark. It is clear what is meant by a special o-grammar.

In [6] we proved the following theorem:

1.24. Theorem: *Let G be a special o-grammar. Then there exists such grammar H that* $\mathscr{L}(G) = \mathscr{L}(H)$.

2. TRACES OF LANGUAGES AND GRAMMARS

2.1. Definition. Let V, U be sets. For each $v \in V$ we put $\mathbf{t}^U(v) = v$ if $v \in U$ and $\mathbf{t}^U(v) = \Lambda$ if $v \in V - U$. According to 1.1 we define the mapping $\mathbf{t}^U_*$ of V^* into U^*. If $x \in V^*$ is a string, then $\mathbf{t}^U_*(x)$ is called the *trace of x in* U^*.

2.2. Lemma. *Let V, U be sets. Then the mapping* $\mathbf{t}^U_*$ *has the following properties:*

(A) *For each* $x, y \in V^*$ *it holds true* $\mathbf{t}^U_*(xy) = \mathbf{t}^U_*(x)\, \mathbf{t}^U_*(y)$.

(B) *If* $\mathbf{t}^U_*(u) = x'y'$ *for some* $u \in V^*$, $x', y' \in U^*$, *then there exist such strings* $x, y \in V^*$ *that* $\mathbf{t}^U_*(x) = x'$, $\mathbf{t}^U_*(y) = y'$, $xy = u$.

(C) *For each* $x \in V^*$ *we have* $\mathbf{t}_*^U(\mathbf{t}_*^U(x)) = \mathbf{t}_*^U(x)$.

The proof can be found in [1].

2.3. Definition. Let (V, L) be a language, U a set. We put $\mathbf{t}_*^U(L) = \{\mathbf{t}_*^U(x); x \in L\}$; the language $(U, \mathbf{t}_*^U(L))$ is called the *trace of the language* (V, L) *in* U^*.

2.4. Definition. Let $G = \langle V, S, R\rangle$ be a special generalized grammar, U a set. We put $S_1 = \mathbf{t}_*^U(S)$, $R_1 = \{(\mathbf{t}_*^U(x), \mathbf{t}_*^U(y)); (x, y) \in R\}$.
Then the special generalized grammar $\langle U, S_1, R_1\rangle$ is called the *trace of the special generalized grammar* G *in* U^*.

2.5. Theorem. *Let* (V, L) *be a language of the type 0,* U *an arbitrary finite set. Then the trace of the language* (V, L) *in* U^* *is a language of the type 0.*

Proof. There exists a phrase structure grammar $G = \langle W, V, \{\sigma\}, R\rangle$ such that $\mathscr{L}(G) = L$. We can suppose, without loss of generality, that $(W - V) \cap U = \emptyset$ according to 1.17. We define the o-grammar $H = \langle W, U, \{\sigma\}, (R_\lambda)_{\lambda \in P}\rangle$ where $P = \{1, 2\}$ with the natural ordering, $R_1 = R$, $R_2 = \{(a, \Lambda); a \in V - U\}$.

Let us have $x \in \mathscr{L}(H)$. Then $x \in U^*$ and one of the following three possibilities occurs:

(a) There exists a string $t_0 \in W^*$ such that $\sigma = t_0 = x$. But $\sigma \in W - U$ and $x \in U$ wich is a contradiction. Thus, the first possibility cannot occur.

(b) There exist a natural number k_1 and some elements $t_0, t_1, \ldots, t_{k_1} \in W^*$ such that $\sigma = t_0$, $t_{k_1} = x$ and $t_{i-1} \underset{R_1}{\Rightarrow} t_i$ for $i = 1, 2, \ldots, k_1$ or $t_{i-1} \underset{R_2}{\Rightarrow} t_i$ for $i = 1, 2, \ldots, k_1$. In the first case, we have $x \in U^*$ and $x \in \mathscr{L}(G)$ which implies $x = \mathbf{t}_*^U(x) \in \mathbf{t}_*^U(\mathscr{L}(G))$. In the second case we have $t_0 = \sigma$ and $t_0 \underset{R_2}{\Rightarrow} t_1$, thus $\sigma \underset{R_2}{\rightarrow} t_1$ which is a contradiction as $\sigma \in W - V$ and $t \underset{R_2}{\rightarrow} z$ implies $t \in V - U$. Thus, the second case cannot occur.

(c) There exist such integers $0 < k_1 < k_2$ and such strings $t_0, t_1, \ldots, t_{k_2} \in W^*$ that $\sigma = t_0$, $t_{k_2} = x$ and $t_{i-1} \underset{R_1}{\Rightarrow} t_i$ for $i = 1, 2, \ldots, k_1$ and $t_{i-1} \underset{R_2}{\Rightarrow} t_i$ for $i = k_1 + 1, \ldots, k_2$. Thus, we have $\sigma \overset{*}{\Rightarrow} t_{k_1}(G)$. We have $t_{k_1} \in V^*$; indeed, if $t_{k_1} \in W^* - V^*$, then at least one of the symbols of t_{k_1} belongs to $W - V$. But this symbol occurs in t_{k_2} as the symbols of $W - V$ cannot be removed by means of the rules of the set R_2. Thus, $t_{k_1} \in W^* - V^*$ implies $x = t_{k_2} \in W^* - U^*$, which is a contradiction. Therefore $t_{k_1} \in V^*$ and $\sigma \overset{*}{\Rightarrow} t_{k_1}(G)$ which implies $t_{k_1} \in \mathscr{L}(G)$. The rules of R_2 transform all symbols of t_{k_1} belonging to $V - U$ into Λ; thus $x = t_{k_2} = \mathbf{t}_*^U(t_{k_1})$ which implies $x \in \mathbf{t}_*^U(\mathscr{L}(G))$.

We have proved $\mathscr{L}(H) \subseteq \mathbf{t}_*^U(\mathscr{L}(G))$.

Let us have $x \in \mathsf{t}^U_*(\mathscr{L}(G))$. Then there exists such a string $y \in \mathscr{L}(G)$ that $\mathsf{t}^U_*(y) = x$. Thus we have $\sigma \overset{*}{\Rightarrow} y\ (G)$. It implies $\sigma \underset{R_1}{\overset{*}{\Rightarrow}} y$. Clearly, $y \underset{R_2}{\overset{*}{\Rightarrow}} x$. Thus, $\sigma \underset{(R\lambda)\lambda \in P}{\overset{*}{\Rightarrow}} x$, $x \in U^*$. It follows $x \in \mathscr{L}(H)$.

We have proved $\mathsf{t}^U_*(\mathscr{L}(G)) \subseteq \mathscr{L}(H)$.

Thus we have $\mathsf{t}^U_*(\mathscr{L}(G)) = \mathscr{L}(H)$.

We put $H_1 = \langle W, \{\sigma\}, (R_\lambda)_{\lambda \in P}\rangle$. Clearly, $\mathscr{L}(H) = \mathscr{L}(H_1) \cap U^*$. According to 1.24 there exists such a grammar $G' = \langle V', V'_T, S', R'\rangle$ that $\mathscr{L}(G') = \mathscr{L}(H_1)$. It implies that $(V'_T, \mathscr{L}(H_1))$ is a language of the type 0. It follows that $(V'_T \cap U, \mathscr{L}(H_1) \cap U^*) = (V'_T \cap U, \mathscr{L}(H))$ is a language of the type 0 according to 1.19. It implies that $(U, \mathscr{L}(H))$ is a language of the type 0 according to 1.18. Thus the trace of $(V, \mathscr{L}(G))$ in U^*, namely the language $(U, \mathscr{L}(H)) = (U, \mathsf{t}^U_*(\mathscr{L}(G)))$, is a language of the type 0.

3. FINITELY GENERATED LANGUAGES

In the following definitions (V, L) is an arbitrary language.

3.1. Definition. For $x \in V^*$ we write $xv(V, L)$ iff there exist such strings $u, v \in V^*$ that $uxv \in L$.

3.2. Definition. For $x, y \in V^*$ we write $x > y\ (V, L)$ iff, for every $u, v \in V^*$, $uxv \in L$ implies $uyv \in L$.

3.3. Definition. For $x, y \in V^*$ we write $x \equiv y\ (V, L)$ iff $x > y\ (V, L)$ and $y > x\ (V, L)$.

3.4. Definition. Let $x, y \in V^*$ be strings. The string x is called a *weak configuration of order 1 of the language* (V, L) *with the result* y iff the following conditions are satisfied: $xv(V, L)$, $x \equiv y\ (V, L)$, $|x| > |y|$.

For the sake of brevity we say "configuration" instead of "weak configuration of order 1" as no other configurations will be studied in the present paper.

By $C(V, L)$ we denote the set of all configurations of the language (V, L); we put $E(V, L) = \{(y, x);\ x \in C(V, L), y \text{ a result of } x\}$, $B(V, L) = L - V^*C(V, L)\,V^*$.

3.5. Definition. Let (V, L) be a language. This language will be called a *language with bounded configurations* iff there exists such an integer n that, for each string $w \in L$ with the property $|w| > n$, there exist such strings $x, y, u, v \in V^*$ that $w = uxv$, $(y, x) \in E(V, L)$ and $|x| \leqq n$.

If (V, L) is a language with bounded configurations, then we denote by $i(V, L)$ the least integer n with the above mentioned property. We put $D(V, L) = \{(y, x);\ (y, x) \in E(V, L),\ |x| \leqq i(V, L)\}$, $K(V, L) = \langle V, B(V, L), D(V, L)\rangle$. Then $K(V, L)$ is called the *generalized bounded configurational grammar of depth 1.*

3.6. Theorem. *Let* (V, L) *be a language with bounded configurations,*

$K(V, L)$ its generalized bounded configurational grammar of depth 1. *Then* $\mathscr{L}(K(V, L)) = L$.

Proof. 1. Let $E(n)$ be the following assertion: If $x \in \mathscr{L}(K(V, L))$ and $|x| = n$, then $x \in L$.

Clearly, $E(0)$ is valid as $x \in \mathscr{L}(K(V, L))$ and $|x| = 0$ implies $x \in B(V, L) \subseteq L$.

Let $m > 0$ be an integer and suppose that $E(0), E(1), \ldots, E(m-1)$ are valid. Let us have $x \in \mathscr{L}(K(V, L))$, $|x| = m$. Then there exist an integer $p \geq 0$ and such elements $s_0, s_1, \ldots, s_p$ of V^* that $s_0 \in B(V, L)$, $s_p = x$ and $s_{i-1} \Rightarrow s_i(K(V, L))$ for $i = 1, 2, \ldots, p$. If $p = 0$, then $x = s_0 \in B(V, L) \subseteq L$. If $p > 0$, then we have $s_{p-1} \Rightarrow x\,(K(V, L))$ which implies the existence of such strings $u, v, t, z \in V^*$ that $utv = s_{p-1}$, $x = uzv$ and $(t, z) \in D(V, L)$. Thus, $|s_{p-1}| = |utv| < |uzv| = |x| = m$ and $s_{p-1} \in \mathscr{L}(K(V, L))$. According to $E(0)$ or $E(1)$ or $\ldots$ or $E(m-1)$ we have $s_{p-1} \in L$. We have $t \equiv z\ (V, L)$ as $(t, z) \in D(V, L)$. Thus $utv = s_{p-1} \in L$ implies $x = uzv \in L$.

We have proved the validity of $E(m)$.

Thus $E(n)$ is valid for every integer $n \geq 0$. It follows $\mathscr{L}(K(V, L)) \subseteq L$.

2. Let $F(n)$ be the following assertion: If $x \in L$ and $|x| = n$, then $x \in \mathscr{L}(K(V, L))$.

Clearly, $F(0)$ is valid as $x \in L$ and $|x| = 0$ implies $x \in B(V, L) \subseteq \mathscr{L}(K(V, L))$.

Let $m > 0$ be an integer and suppose that $F(0), F(1), \ldots, F(m-1)$ are valid. Let us have $x \in L$, $|x| = m$. If $x \in B(V, L)$, then $x \in \mathscr{L}(K(V, L))$.

Let us have $x \in V^*C(V, L)\,V^*$. Then two possibilities can occur:

(α) If $m \leq i(V, L)$ then there exist such strings $u, v, z \in V^*$ that $x = uzv$, $z \in C(V, L)$. We have $|z| \leq |uzv| = |x| = m \leq i(V, L)$.

(β) If $m > i(V, L)$, then there exist such strings $u, v, z \in V^*$ that $x = uzv$, $z \in C(V, L)$ and $|z| \leq i(V, L)$ according to the definition of $i(V, L)$.

Thus, in both cases, there exist such strings $u, v, z \in V^*$ that $x = uzv$, $z \in C(V, L)$ and $|z| \leq i(V, L)$. Let t be an arbitrary result of z. Then $(t, z) \in D(V, L)$ which means $t \equiv z\ (V, L)$, $|t| < |z|$. Thus, $uzv = x \in L$ implies $utv \in L$; moreover, we have $|utv| < |uzv| = m$. According to $F(0)$ or $F(1)$ or $\ldots$ or $F(m-1)$ we have $utv \in \mathscr{L}(K(V, L))$ which implies the existence of such a string $s \in B(V, L)$ that $s \overset{*}{\Rightarrow} utv\,(K(V, L))$. As $utv \Rightarrow uzv\,(K(V, L))$ we have $s \overset{*}{\Rightarrow} uzv\,(K(V, L))$ and $x = uzv \in \mathscr{L}(K(V, L))$.

We have proved the validity of $F(m)$.

Thus, $F(n)$ is valid for every integer $n \geq 0$. It follows $L \subseteq \mathscr{L}(K(V, L))$.

3. We have proved $L = \mathscr{L}(K(V, L))$.

3.7. Definition. Let (V, L) be such a language with bounded configurations that V is a finite set. Then (V, L) is called a *finitely generated language.*

3.8. Lemma. *Let (V, L) be a finitely generated language. Then the sets $B(V, L)$, $D(V, L)$ are finite.*

Proof. The finiteness of $D(V, L)$ follows from the fact that there exists only a finite number of strings $x \in V^*$ with the property $|x| \leqq i(V, L)$. Thus, there exists only a finite number of ordered pairs of such strings; all rules of the set $D(V, L)$ are contained in the finite set of such pairs.

It follows from the definition of $B(V, L)$ that $s \in B(V, L)$ implies $|s| \leqq i(V, L)$. Thus, $B(V, L)$ is finite.

3.9. Theorem. *Every finitely generated language is a language of the type 0.*

Proof. It follows from 3.8 that $K(V, L)$ is a grammar for the finitely generated language (V, L).

4. CHARACTERIZATION OF LANGUAGES OF THE TYPE 0

4.1. Lemma *Let $G = \langle V, V_T, \{\sigma\}, R\rangle$ be an arbitrary grammar having the standard form such that $\mathscr{L}(G) \neq \varnothing$. We put $G_1 = \langle V, \{\sigma\}, R\rangle$. Let Z_1, Z_2, Z_3 be sets with the following properties: there exists a bijection f_1 of R onto Z_1, a bijection f_2 of R onto Z_2, Z_3 has precisely two elements $:$ and $—$ and the sets V, Z_1, Z_2, Z_3 are mutually disjoint. We put $f_1(r) = [_r$, $f_2(r) =]_r$ for each $r \in R$, $V_0 = V \cup Z_1 \cup Z_2 \cup Z_3$, $R_1 = \{(x, [_r y]_r);\ r = (x, y) \in R\}$, $R_2 = \{(a[_r, : [_r — a);\ r \in R, a \in V\}$, $R_3 = \{(a]_r, :]_r — a);\ r \in R,\ a \in V\}$, $R_4 = \{(a :, : : a);\ a \in V\}$, $R_5 = \{(a —, — — a);\ a \in V\}$, $R_0 = R_1 \cup R_2 \cup R_3 \cup R_4 \cup R_5$, $G_0 = \langle V_0, \{\sigma\}, R_0\rangle$.*

Then the following assertions hold true:

(i) If $x, y \in V_0^*$, $x \overset{*}{\Rightarrow} y\ (G_0)$, then $|x| \leqq |y|$. If $\sigma \overset{*}{\Rightarrow} y\ (G_0)$ and $|y| = 1$, then $y = \sigma$; if $\sigma \Rightarrow y\ (G_0)$ and $|y| = 2$, then $(\sigma, y) \in R_1$.

(ii) The language $(V, \mathscr{L}(G_1))$ is the trace of $(V_0, \mathscr{L}(G_0))$ in V^*.

(iii) If $(x, y) \in R_0$, $(t, z) \in R_0$ and $u, v, p, q \in V_0^*$ are such elements that $uyv = pzq$, then either $u = pzq_1$, $q_1q_2 = q$, $q_2 = yv$ for suitable strings $q_1, q_2 \in V_0^*$ or $uy = p_1$, $p_1p_2 = p$, $p_2zq = v$ for suitable strings $p_1, p_2 \in V_0^*$ or $u = p$, $y = z$, $v = q$.

(iv) If $(x, y) \in R_0$, then $y > x\ (V_0, \mathscr{L}(G_0))$.

(v) $(V_0, \mathscr{L}(G_0))$ is a finitely generated language.

Proof of (*i*). Clearly, $(x, y) \in R$ implies $|x| \leqq |y| + 1$, thus, $(t, z) \in R_0$ implies $|t| < |z|$. It follows the first assertion. As $|z| \geqq 2$ for each $(t, z) \in R_0$, then $\sigma \overset{*}{\Rightarrow} y\ (G_0)$ and $|y| = 1$ imply $y = \sigma$ and $\sigma \overset{*}{\Rightarrow} y\ (G_0)$ and $|y| = 2$ imply $(\sigma, y) \in R_1$.

Proof of (*ii*). 1. Let $C(n)$ be the following assertion: If $w \in \mathscr{L}(G_0)$ and $|w| = n$, then $\mathrm{t}_*^V(w) \in \mathscr{L}(G_1)$.

Clearly, $w \in \mathscr{L}(G_0)$ implies $|w| \geqq 1$ according to (*i*).

$C(1)$ is valid, as $w \in \mathscr{L}(G_0)$, $|w| = 1$ imply $w = \sigma$ according to (*i*) and $\mathbf{t}_*^V(w) = \mathbf{t}_*^V(\sigma) = \sigma \in \mathscr{L}(G_1)$ according to 2.1.

Let $m > 1$ be an integer, let us suppose that $C(1), C(2), \ldots, C(m-1)$ are valid. Let us have $w \in \mathscr{L}(G_0)$, $|w| = m$. Then there exists a σ-derivation $s_0, s_1, \ldots, s_n$ of w in G_0 and $n \geqq 1$. Especially, we have $s_{n-1} \Rightarrow w\,(G_0)$. Thus some strings $u, v \in V_0^*$ and a rule $(t, z) \in R_0$ exist such that $s_{n-1} = utv$, $uzv = w$. We have $|utv| < |uzv| = |w| = m$, $utv = s_{n-1} \in \mathscr{L}(G_0)$ according to (*i*). According to $C(1)$ or $C(2)$ or ... or $C(m-1)$ we have $\mathbf{t}_*^V(utv) \in \mathscr{L}(G_1)$, thus — according to 2.2 (A) — $\mathbf{t}_*^V(u)\,\mathbf{t}_*^V(t)\,\mathbf{t}_*^V(v) \in \mathscr{L}(G_1)$.

If $(t, z) \in R_1$, then $t \in V^*$ which implies $\mathbf{t}_*^V(t) = t$ and $z = [_r w']_r$ for a suitable $w' \in V^*$ where $r = (t, w') \in R$; it follows $\mathbf{t}_*^V(z) = w'$. Thus we have $\mathbf{t}_*^V(u)\,t\,\mathbf{t}_*^V(v) \in \mathscr{L}(G_1)$ which implies $\mathbf{t}_*^V(w) = \mathbf{t}_*^V(uzv) = \mathbf{t}_*^V(u)\,\mathbf{t}_*^V(z)\,\mathbf{t}_*^V(v) = \mathbf{t}_*^V(u)\,w'\,\mathbf{t}_*^V(v) \in \mathscr{L}(G_1)$.

If $(t, z) \in R_2 \cup R_3 \cup R_4 \cup R_5$, then, clearly $\mathbf{t}_*^V(t) = \mathbf{t}_*^V(z)$. Thus $\mathbf{t}_*^V(w) = \mathbf{t}_*^V(uzv) = \mathbf{t}_*^V(u)\,\mathbf{t}_*^V(z)\,\mathbf{t}_*^V(v) = \mathbf{t}_*^V(u)\,\mathbf{t}_*^V(t)\,\mathbf{t}_*^V(v) = \mathbf{t}_*^V(utv) \in \mathscr{L}(G_1)$.

We have proved $C(m)$.

Thus $C(n)$ is valid for $n = 1, 2, \ldots$. It implies $\mathbf{t}_*^V(\mathscr{L}(G_0)) \subseteq \mathscr{L}(G_1)$.

2. Let $D(n)$ be the following assertion: If $x \in \mathscr{L}(G_1)$ and if there exists an σ-derivation $\sigma = s_0, s_1, \ldots, s_n = x$ of x in G_1 of length n, then there exists an element $w \in \mathscr{L}(G_0)$ such that $\mathbf{t}_*^V(w) = x$.

$D(0)$ is valid as $x \in \mathscr{L}(G_1)$ and $n = 0$ imply $\sigma = s_0 = x$ and $\mathbf{t}_*^V(\sigma) = \sigma = x$.

Let $m > 1$ be an integer, let us suppose that $D(0), D(1), \ldots, D(m-1)$ are valid. Let $x \in \mathscr{L}(G_1)$ and let us have an σ-derivation $\sigma = s_0, s_1, \ldots, s_m = x$. Then $s_{m-1} \in \mathscr{L}(G_1)$ and $\sigma = s_0, s_1, \ldots, s_{m-1}$ is an σ-derivation of s_{m-1} of length $m-1$. According to $D(m-1)$ there exists an element $w' \in \mathscr{L}(G_0)$ such that $\mathbf{t}_*^V(w') = s_{m-1}$. As we have $s_{m-1} \Rightarrow x\,(G_1)$, there exist a rule $r = (t, z) \in R$ and some strings $u, v \in V^*$ such that $s_{m-1} = utv$, $uzv = x$. It follows the existence of $u', t', v' \in V_0^*$ such that $u't'v' = w'$, $\mathbf{t}_*^V(u') = u$, $\mathbf{t}_*^V(t') = t$, $\mathbf{t}_*^V(v') = v$ according to 2.2 (B). If $t' \neq t$, then applying some rules of $R_2 \cup R_3 \cup R_4 \cup R_5$ we get $t' \overset{*}{\Rightarrow} yt\,(G_0)$ where $y \in (V_0 - V)^*$. It follows $\mathbf{t}_*^V(y) = \Lambda$. If $t' = t$, then $t' \overset{*}{\Rightarrow} t\,(G_0)$ holds trivially. Thus, in both cases there exists a string $y \in (V_0 - V)^*$ such that $t' \overset{*}{\Rightarrow} yt\,(G_0)$. It implies $u't'v' \overset{*}{\Rightarrow} u'ytv'(G_0)$. As $(t, [_r z]_r) \in R_0$ we have $u't'v' \overset{*}{\Rightarrow} u'y[_r z]_r\,v'(G_0)$. The fact that $u't'v' = w' \in \mathscr{L}(G_0)$ implies $u'y[_r z]_r\,v' \in \mathscr{L}(G_0)$. We have $\mathbf{t}_*^V(u'y[_r z]_r\,v') = \mathbf{t}_*^V(u')\,z\,\mathbf{t}_*^V(v') = uzv = x$.

We have proved $D(m)$.

Thus $D(n)$ is valid for $n = 0, 1, 2, \ldots$. It implies $\mathscr{L}(G_1) \subseteq \mathbf{t}_*^V(\mathscr{L}(G_0))$.

3. We have $\mathscr{L}(G_1) = \mathbf{t}_*^V(\mathscr{L}(G_0))$, $V \subseteq V_0$. Thus, $(V, \mathscr{L}(G_1)) = (V, \mathbf{t}_*^V(\mathscr{L}(G_0)))$ is the trace of $(V_0, \mathscr{L}(G_0))$ in V^*.

Proof of (*iii*). Let us have $(x, y) \in R_0$, $(t, z) \in R_0$, $u, v, p, q \in V_0^*$,

$uyv \Rightarrow pzq$. Let us suppose that there exist such strings $y_1, y_2, z_1, z_2 \in V_0^*$, $c \in V_0$ that $y = y_1cy_2$, $z = z_1cz_2$, $uy_1 = pz_1$, $y_2v = z_2q$.

(a) If $(x, y) \in R_1$, $(t, z) \in R_1$, then $y = [_rw]_r$, $z = [_{r'}w']_{r'}$ for suitable $r, r' \in R$, $w, w' \in V^*$. It follows $r = r'$, $w = w'$ as $[_r, [_{r'},]_r,]_{r'} \in V_0 - V$ It implies $y = [_rw]_r = [_{r'}w']_{r'} = z$, $u = p$, $v = q$.

(b) If $(x, y) \in R_1$, $(t, z) \in R_2$, then the above mentioned situation is impossible as $y = y_1cy_2$, $z = z_1cz_2$ imply $c \in V$ or $c = [_r$ for a suitable $r \in R$. In the first case we have $y_1 \neq \Lambda \neq y_2$, $z_1 = \,:[_r -$, $z_2 = \Lambda$. From the fact that $uy_1 = pz_1$ it follows that the last symbol of y_1 is —, which is a contradiction. In the second case we have $y_1 = \Lambda$, $y_2 \neq \Lambda$, $z_1 = \,:$, $z_2 = -a$ for a suitable $a \in V$. From the fact that $y_2v = z_2q$ it follows that the first symbol of y_2 is —, which is a contradiction.

(c) If $(x, y) \in R_1$, $(t, z) \in R_3$, then the above mentioned situation is impossible as $y = y_1cy_2$, $z = z_1cz_2$ imply $c \in V$ or $c =]_r$ for a suitable $r \in R$. In the first case we have $y_1 \neq \Lambda \neq y_2$, $z_1 = \,:]_r -$, $z_2 = \Lambda$. From the fact that $uy_1 = pz_1$ it follows that the last symbol of y_1 is —, which is a contradiction. In the second case we have $y_1 \neq \Lambda$, $y_2 = \Lambda$, $z_1 = \,:$, $z_2 = -a$ for a suitable $a \in V$. From the fact that $uy_1 = pz_1$ it follows that the last symbol of y_1 is :, which is a contradiction.

(d) If $(x, y) \in R_1$, $(t, z) \in R_4$, then the above mentioned situation is impossible as $y = y_1cy_2$, $z = z_1cz_2$ imply $c \in V$. Thus, $c \in V$, $y_1 \neq \Lambda \neq y_2$, $z_1 = \,::$, $z_2 = \Lambda$. From the fact that $uy_1 = pz_1$ it follows that the last symbol of y_1 is :, which is a contradiction.

(e) If $(x, y) \in R_1$, $(t, z) \in R_5$, then the above mentioned situation is impossible; it can be proved similarly as in case (d).

(f) If $(x, y) \in R_2$, $(t, z) \in R_2$, then clearly $y = z$, $u = p$, $v = q$.

(g) If $(x, y) \in R_2$, $(t, z) \in R_3$, then $c \in V$ or $c = \,:$ or $c = -$. If $c \in V$, then $y_1 = \,:[_r -$, $y_2 = \Lambda$, $z_1 = \,:]_{r'} -$, $z_2 = \Lambda$ for suitable $r, r' \in R$. It follows from $uy_1 = pz_1$ that $[_r =]_{r'}$, which is a contradiction. If $c = \,:$, then $y_1 = \Lambda$, $y_2 = [_r - a$, $z_1 = \Lambda$, $z_2 =]_{r'} - a$ for suitable $r, r' \in R$, $a, a' \in V$. From the fact that $y_2v = z_2q$ it follows that $[_r =]_{r'}$, which is a contradiction. If $c = -$, then $y_1 = \,:[_r$, $y_2 = a \in V$, $z_1 = \,:]_{r'}$, $z_2 = = a \in V$ for suitable $r, r' \in R$, $a, a' \in V$. From the fact that $uy_1 = pz_1$ it follows that $[_r =]_{r'}$, which is a contradiction. Thus the above mentioned situation is impossible.

(h) If $(x, y) \in R_2$, $(t, z) \in R_4$, then $c \in V$ or $c = \,:$. In the first case we have $y_1 = \,:[_r -$, $y_2 = \Lambda$, $z_1 = \,::$, $z_2 = \Lambda$ for a suitable $r \in R$. From the fact that $uy_1 = pz_1$ it follows $- = \,:$, which is impossible. In the second case we have $y_1 = \Lambda$, $y_2 = [_r - a$, $z_1 = \Lambda$, $z_2 = \,:a'$ or $z_1 = \,:$, $z_2 = a'$ for suitable $r \in R$, $a, a' \in V$. Thus $[_r = \,:$ or $[_r = a' \in V$, which is impossible. Thus the above mentioned situation is impossible.

(j) If $(x, y) \in R_2$, $(t, z) \in R_5$, then the above mentioned situation is impossible; it can be proved similarly as in the case (h).

(k) If $(x, y) \in R_3$, $(t, z) \in R_3$, then clearly $y = z$, $u = p$, $v = q$.

(l) If $(x, y) \in R_3$, $(t, z) \in R_4$, then $c \in V$ or $c = \,:$. In the first case we have $y_1 = \,:\,]_r —$, $y_2 = \Lambda$, $z_1 = \,::$, $z_2 = \Lambda$. From the fact that $uy_1 = pz_1$ it follows $— = \,:$, which is impossible. In the second case we have $y_1 = \Lambda$, $y_2 = \,]_r — a$, $z_1 = \Lambda$, $z_2 = \,: a'$ or $z_1 = \,:$, $z_2 = a'$ for suitable $a, a' \in V$. From the fact that $y_2 v = z_2 q$ it follows $]_r = \,:$ or $]_r = a' \in V$, which is impossible. Thus the above mentioned situation cannot occur.

(m) If $(x, y) \in R_3$, $(t, z) \in R_5$, then $c \in V$ or $c = —$. In the first case we have $y_1 = \,:\,]_r —$, $y_2 = \Lambda$, $z_1 = ——$, $z_2 = \Lambda$ for a suitable $r \in R$. From the fact that $uy_1 = pz_1$ it follows $]_r = —$, which is impossible. In the second case we have $y_1 = \,:\,]_r$, $y_2 = a$ and simultaneously $z_1 = \Lambda$, $z_2 = — a'$ or $z_1 = —$, $z_2 = a'$ for suitable $r \in R$, $a, a' \in V$. The first possibility cannot occur as $y_2 v = z_2 q$ implies $— = a \in V$, which is impossible; similarly the second possibility cannot occur as $uy_1 = pz_1$ implies $]_r = —$, which is impossible.

Thus the above mentioned situation cannot occur.

(n) If $(x, y) \in R_4$, $(t, z) \in R_4$, then clearly $y = z$, $u = p$, $v = q$.

(o) If $(x, y) \in R_4$, $(t, z) \in R_5$, then $c \in V$. It implies $y_1 = \,::$, $y_2 = \Lambda$, $z_1 = ——$, $z_2 = \Lambda$. Thus $uy_1 = pz_1$ implies $: = —$, which is impossible. Thus the above mentioned situation cannot occur.

(p) If $(x, y) \in R_5$, $(t, z) \in R_5$, then clearly $y = z$, $u = p$, $v = q$.

From the above analysis it follows that $(x, y) \in R_0$, $(t, z) \in R_0$ and $u, v, p, q \in V_0^*$ and $uyv = pzq$ imply that either z is a substring of u and y a substring of q or z is a substring of v and y a substring of p or $u = p$, $y = z$, $v = q$.

Proof of (*iv*). Let $E(n)$ be the following assertion: If $(x, y) \in R_0$, $u, v \in V_0^*$, $uyv \in \mathscr{L}(G_0)$, $|uyv| = n$, then $uxv \in \mathscr{L}(G_0)$. We have $|y| \geqq 2$, thus $|uyv| \geqq 2$.

We prove $E(2)$. Let us have $(x, y) \in R_0$, $u, v \in V_0^*$, $uyv \in \mathscr{L}(G_0)$, $|uyv| = 2$. As $|y| \geqq 2$, we have $u = \Lambda = v$. We have $\sigma \overset{*}{\Rightarrow} uyv\,(G_0)$, thus $\sigma \overset{*}{\Rightarrow} y\,(G_0)$. According to (*i*) we have $(\tilde{\sigma}, y) \in R_1$. It implies the existence of an element $r = (\sigma, w) \in R$ such that $y = [_r w]_r$. As $|y| = 2$, we have $w = \Lambda$ and $(\sigma, y) = (\sigma, [_r]_r)$. Now, $(x, y) \in R_0$, $|y| = 2$ implies $(x, y) \in R_1$. Thus an element $r' = (x, w') \in R$ exists such that $y = [_{r'} w']_{r'}$. But $[_r]_r = y = [_{r'} w']_{r'}$ implies $r = r'$, $w' = \Lambda$. It follows $(x, \Lambda) = r' = r = (\sigma, \Lambda)$ and $x = \sigma$. Thus $\sigma \overset{*}{\Rightarrow} \sigma\,(G_0)$ and $\sigma = uxv$. It implies $uxv \in \mathscr{L}(G_0)$.

Let $m > 2$ be a natural number. Suppose that $E(2), \ldots, E(m-1)$ are valid.

Let us have $(x, y) \in R_0$, $u, v \in V_0^*$, $uyv \in \mathscr{L}(G_0)$, $|uyv| = m$. We have $\sigma \overset{*}{\Rightarrow} uyv\,(G_0)$. Thus an σ-derivation $s_0, s_1, \ldots, s_l$ of uyv in G_0 exists and we have $l \geqq 1$. Especially we have $s_{l-1} \in \mathscr{L}(G_0)$; $s_{l-1} \Rightarrow uyv\,(G_0)$. Thus such strings $p, q \in V_0^*$, $(t, z) \in R_0$ exist that $s_{l-1} = ptq$, $pzq = uyv$. According to (*iii*) the following three cases can occur:

(a) There exist such strings $q_1, q_2 \in V_0^*$ that $u = pzq_1$, $q_1q_2 = q$, $q_2 = yv$. Then $(x, y) \in R_0$, $ptq_1 \in V_0^*$, $v \in V_0^*$, $ptq_1yv = ptq_1q_2 = ptq = s_{l-1} \in \mathscr{L}(G_0)$, $|ptq_1yv| < |pzq_1yv| = |pzq| = |uyv| = m$ according to (*i*). According to $E(2)$ or ... or $E(m-1)$ we have $ptq_1xv \in \mathscr{L}(G_0)$, which implies $uxv = pzq_1xv \in \mathscr{L}(G_0)$.

(b) There exist such strings $p_1, p_2 \in V_0^*$ that $uy = p_1$, $p_1p_2 = p$, $p_2zq = v$. We prove $uxv \in \mathscr{L}(G_0)$ similarly as in the case (a).

(c) We have $u = p$, $y = z$, $v = q$. Then there exists a natural number i, $1 \leqq i \leqq 5$, such that $(x, y) \in R_i$, $(t, z) \in R_i$. If $i = 1$, then there exist such elements $r, r' \in R$, $w, w' \in V^*$ that $r = (x, w)$, $r' = (t, w')$, $y = [_r w]_r$, $z = [_{r'} w']_{r'}$. From $y = z$ it follows that $r = r'$, which implies $x = t$. If $i = 2$ or 3 or 4 or 5, then clearly $y = z$ implies $x = t$.

Thus $uxv = ptq = s_{l-1} \in \mathscr{L}(G_0)$.

We have proved that $E(n)$ holds true for $n = 2, 3, \ldots$. Thus, $(x, y) \in R_0$, $u, v \in V_0^*$, $uyv \in \mathscr{L}(G_0)$ implies $uxv \in \mathscr{L}(G_0)$. Therefore, $(x, y) \in R_0$ implies $y > x\,(V_0, \mathscr{L}(G_0))$.

Proof of (*v*). As V is finite and G is a grammar, then V_0 is finite, too. We have to demonstrate that $(V_0, \mathscr{L}(G_0))$ is a language with bounded configurations.

Let P be the set of all $(x, y) \in R_0$ with the following property: There exist such strings $u, v \in V_0^*$ that $\sigma \overset{*}{\Rightarrow} uxv\,(G_0)$. Clearly $P \subseteq R_0$ and P is a finite set.

We have $x > y\,(V_0, \mathscr{L}(G_0))$ for each $(x, y) \in R_0$. According to (*iv*) we have $y > x\,(V_0, \mathscr{L}(G_0))$ for each $(x, y) \in R_0$. Thus, $x \equiv y\,(V_0, \mathscr{L}(G_0))$, $|x| < |y|$ for each $(x, y) \in R_0$. If $(x, y) \in P$, then $x\nu(V_0, \mathscr{L}(G_0))$, which implies $y\nu(V_0, \mathscr{L}(G_0))$. Thus $(x, y) \in P$ implies $(x, y) \in E(V_0, \mathscr{L}(G_0))$ and we have $P \subseteq E(V_0, \mathscr{L}(G_0))$.

If $P = \varnothing$, then $\sigma \overset{*}{\Rightarrow} uxv\,(G_0)$ for no $u, v \in V_0^*$ and no $(x, y) \in R_0$. Thus $\mathscr{L}(G_0) = \{\sigma\}$, which implies $\mathscr{L}(G_1) = \{\sigma\}$ according to (*ii*), and $\mathscr{L}(G) = \mathscr{L}(G_1) \cap V_T^* = \varnothing$, which is a contradiction. Thus $P \neq \varnothing$. We put $n = \max\{|y|;\ (x, y) \in P\}$. Clearly $|y| \geqq 2$ for each $(x, y) \in P$, thus $n \geqq 2$. Let us have $w \in \mathscr{L}(G_0)$, $|w| > n$. Thus $|w| > 2$ and $\sigma \overset{*}{\Rightarrow} w\,(G_0)$. According to (*i*) there exists a σ-derivation $s_0, s_1, \ldots, s_p$ of w in G_0 with the property $p \geqq 1$. We have $s_{p-1} \Rightarrow w\,(G_0)$. Thus there exist such strings $u, v \in V_0^*$ and such rule $(t, z) \in R_0$ that $s_{p-1} = utv$, $uzv = w$. As $\sigma \overset{*}{\Rightarrow} utv\,(G_0)$, we have $(t, z) \in P$, thus $|z| \leqq n$. We have thus found such an integer n that for each string $w \in \mathscr{L}(G_0)$ with the property $|w| > n$ there exist such strings $t, z, u, v \in V_0^*$ that $w = uzv$, $(t, z) \in E(V_0, \mathscr{L}(G_0))$ and $|z| \leqq n$.

We have proved that $(V_0, \mathscr{L}(G_0))$ is a language with bounded configurations; thus it is finitely generated.

4.2. Theorem. *Let (U, L) be a language. Then the following two assertions are equivalent:*

(A) *(U, L) is a language of the type 0.*

(B) *There exist such a finitely generated language (V_0, L_0), such a finite set V and such a set W that (U, L) is the intersection of the trace of (V_0, L_0) in the set V^* with the full language (W, W^*).*

Proof. Let (A) be satisfied. If $L = \emptyset$, then $(U, \emptyset)$ is a finitely generated language in a trivial way; we put $V_0 = V = W = U$, $L_0 = \emptyset$. Then $(V_0, L_0) = (U, \emptyset)$ is finitely generated, $(V, \emptyset) = (U, \emptyset)$ is the trace of (V_0, L_0) in V^* and $(U, \emptyset)$ is the intersection of $(V, \emptyset) = (U, \emptyset)$ with $(W, W^*) = (U, U^*)$.

We can suppose $L \neq \emptyset$. According to 1.15 there exists such a grammar $H = \langle V, U, \{\sigma\}, R\rangle$ having the standard form that $\emptyset \neq L = \mathscr{L}(H)$. We put $G_1 = \langle V, \{\sigma\}, R\rangle$ and we define $G_0 = \langle V_0, \{\sigma\}, R_0\rangle$ according to 4.1. Then $(V_0, \mathscr{L}(G_0))$ is a finitely generated language and $(V, \mathscr{L}(G_1))$ is the trace of $(V_0, \mathscr{L}(G_0))$ in V^* according to 4.1 (*v*) and (*ii*). Clearly, $\mathscr{L}(H) = \mathscr{L}(G_1) \cap U^*$, thus $(U, \mathscr{L}(H)) = (V \cap U, \mathscr{L}(G_1) \cap U^*)$ is the intersection of $(V, \mathscr{L}(G_1))$ with the full language (U, U^*).

We have proved that (A) implies (B).

Let (B) be satisfied. Then (V_0, L_0) is a language of the type 0 according to 3.9, the trace (V, L_1) of (V_0, L_0) in the set V^* is a language of the type 0 according to 2.5 and the intersection (U, L) of (V, L_1) with the full language (W, W^*) is a language of the type 0 according to 1.19.

We have proved that (B) implies (A).

BIBLIOGRAPHY

[1] Novotný, M.: *On a class of languages.* Archivum Matematicum Brno **6** (1970), 155—170.

[2] Novotný, M.: *Algebraic structures of mathematical linguistics.* Bull. Math. de la Soc. Sci. Math. de la R. S. de Roumanie **12** (**60**) (1969), 87—101.

[3] Gladkij, A. V.: *Konfiguracionnyje charakteristiki jazykov.* Problemy kibernetiki **10** (1963), 251—260.

[4] Ginsburg, S.: *The mathematical theory of context-free languages.* Mc Graw Hill Book Company 1966.

[5] Chomsky, N.: *On certain formal properties of grammars.* Information and Control **2** (1959), 137—167.

[6] Novotný, M.: *On configurations in the sense of Kulagina.* Publ. Fac. Sci. Univ. J. E. Purkyně Brno No 507 (1969), 301—316.

Department of Mathematics
J. E. Purkyně University, Brno
Czechoslovakia