Emil Navrátil

Context-free grammars with regular conditions

Kybernetika, Vol. 6 (1970), No. 2, (118)--126

Persistent URL: http://dml.cz/dmlcz/124905

KYBERNETIKA ČÍSLO 2, ROČNÍK 6/1970

Context-Free Grammars with Regular Conditions

EMIL NAVRÁTIL

It is shown that the ε-free CF-grammars with regular conditions generate exactly the class of context-sensitive languages.

If one generates a language by a context-free grammar, the rewriting rules of the form $u \to v$ (u and v denote a symbol and a string resp.) can be applied to whatever string containing u, without any other restrictions.

However, it is naturally to limit this application by some condition imposed to the string in consideration. For example, in [1], the application of the rule $u \to v$ is allowed only for such strings in which some prescribed substrings appear and the others do not. One can show that such conditions can be well expressed by regular sets.

In this paper, we shall study therefore such context-free grammars, where each rule $u \to v$ is labelled by the pair of regular sets M_1, M_2 and the rule can be applied only to that strings, whose initial subword to the left from the rewriting symbol u is a string of the set M_1 and the terminal subword to the right from u is an element of M_2.

Now, formal definitions follow.

Definition 1. *A context-free grammar with conditions* is a system $G = \langle N, T, P, S \rangle$, where N and T are disjoint finite nonempty sets (nonterminal and terminal alphabets), $S \in N$ (the start symbol), P is a nonempty finite set of rules of the form $u \to v[M_1, M_2]$, where $u \in N$, $v \in (N \cup T)^*$ and M_1, M_2 are nonempty subsets of the set $(N \cup T)^*$.

Definition 2. Let α, β be two strings, $\alpha \in (N \cup T)^* N(N \cup T)^*$, $\beta \in (N \cup T)^*$. We say that α *can be immediately rewritten to* β (we denote $\alpha \underset{G}{\Rightarrow} \beta$), if there exists a rule $u \to v[M_1, M_2] \in P$ and two such strings $\omega_1 \in M_1$, $\omega_2 \in M_2$ that $\alpha = \omega_1 u \omega_2$, $\beta = \omega_1 v \omega_2$. We say that α *can be rewritten to* β ($\alpha \underset{G}{\Rightarrow}^* \beta$), if either $\beta = \alpha$ or there exists a finite sequence of strings $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ that $\alpha_0 = \alpha$, $\alpha_i \underset{G}{\Rightarrow} \alpha_{i+1}$ for each $i = 0, 1, \ldots, n-1$ and $\alpha_n = \beta$. If $\alpha_0 = S$, the sequence $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ is called *derivation over G of length n.*

Definition 3. *Language $L(G)$ generated by grammar G* is the set $\{y \in T^*/S \underset{G}{\Rightarrow}^* y\}$.

Notation. We shall write shortly $u \to v$ instead of $u \to v[(N \cup T)^*, (N \cup T)^*]$.

Let us introduce the following denotation:

Let T' be a set, $|T'| = |T|$,* $T' \cap T = \emptyset$, $T' \cap N = \emptyset$. Let φ be one-to-one mapping from T to T'. Let us define $x' = x$ for $x \in N$, $x' = \varphi(x)$ for $x \in T$, $\varepsilon' = \varepsilon$ (ε denotes the empty word), $(\beta x)' = \beta' x'$ ($\beta \in (N \cup T)^*, x \in N \cup T$) and for $M \subset (N \cup T)^*$ $M' = \{\alpha' / \alpha \in M\}$.

Lemma 1. *Let G be a context-free grammar with conditions, $G = \langle N, T, P, S\rangle$. Let $G' = \langle N', T, P', S\rangle$ be the following grammar with conditions: $N' = N \cup T'$, $P' = \{t' \to t / t \in T\} \cup \{u' \to v'[M_1', M_2'] / u \to v[M_1, M_2] \in P\}$. Then $L(G') = L(G)$.*

(We construct P' therefore in the following way: we prescribe a new set T' to the set T, the set of "nonterminal doubles of terminal symbols". Further, in all rules of P the occurrences of terminal symbols are substituted by their doubles and the rules $t' \to t$ for each symbol $t \in T$ are added.)

Proof of Lemma 1. a) Suppose that $\gamma \in L(G)$, i.e. there exists a derivation over G $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ such that $\alpha_n = \gamma$ ($\gamma \in T^*$). We state that then for each $i = 0, 1, \ldots, n$ $S \underset{G'}{\Rightarrow}^* \alpha_i'$ holds. Proof by induction:

1. For $i = 0$ the statement holds obviously.

2. For each i $(0 \leqq i \leqq n-1)$ the relation $S \underset{G'}{\Rightarrow}^* \alpha_i'$ implies $S \underset{G'}{\Rightarrow}^* \alpha_{i+1}'$. To prove this it is sufficient to show that $\alpha_i' \underset{G'}{\Rightarrow} \alpha_{i+1}'$. Since $\alpha_i \underset{G'}{\Rightarrow} \alpha_{i+1}$, there exist according to Definition 2 the rule $u \to v[M_1, M_2] \in P$ and two such strings $\omega_1 \in M_1$, $\omega_2 \in M_2$ that $\alpha_i = \omega_1 u \omega_2$, $\alpha_{i+1} = \omega_1 v \omega_2$. Therefore $u' \to v'[M_1', M_2'] \in P'$, $\omega_1' \in M_1'$, $\omega_2' \in M_2'$ and $\alpha_i' \underset{G'}{\Rightarrow} \alpha_{i+1}'$ for each i $(0 \leqq i \leqq n-1)$.

We have proved $S \underset{G'}{\Rightarrow}^* \alpha_n'$, it remains to show $S \underset{G'}{\Rightarrow}^* \alpha_n$. If $\alpha_n = \varepsilon$, it is obvious. Otherwise $\alpha_n = x_1 x_2 \ldots x_r$, $x_j \in T$ for $1, 2, \ldots, r$. Then $\alpha_n' = x_1' x_2' \ldots x_r'$ and obviously $\alpha_n' \underset{G'}{\Rightarrow}^* \alpha_n$, therefore $\gamma \in L(G')$.

b) Suppose that $\gamma \in L(G')$, i.e. there exists a derivation $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ over G' for which $\alpha_n = \gamma$, $\alpha_i \in (N \cup T' \cup T)^*$ for $i = 0, 1, \ldots, n-1$.

Let us introduce the following denotation:

$x^0 = \varphi^{-1}(x)$ for $x \in T'$, $x^0 = x$ for $x \in N \cup T$, $\varepsilon^0 = \varepsilon$, $(\beta x)^0 = \beta^0 x^0$ ($\beta \in (N \cup T' \cup T)^*$, $x \in N \cup T' \cup T$), $M^0 = \{\alpha^0 / \alpha \in M\}$ for $M \subset (N \cup T' \cup T)^*$.

* If $\mathfrak{M}$ is a finite set, then $|\mathfrak{M}|$ denotes the number of its elements.

We shall prove $S \underset{G}{\Rightarrow} \alpha_i^0$ for $i = 0, 1, \ldots, n$. It is sufficient to prove that for each i $(0 \leqq i \leqq n-1)$ is either $\alpha_i^0 = \alpha_{i+1}^0$ or $\alpha_i^0 \underset{G}{\Rightarrow} \alpha_{i+1}^0$. As $\alpha_i \underset{G'}{\Rightarrow} \alpha_{i+1}$, it is either $\alpha_i = \omega_1 t' \omega_2$ and $\alpha_{i+1} = \omega_1 t \omega_2$ $(t \in T)$ and therefore $\alpha_i^0 = \alpha_{i+1}^0$, or $\alpha_i = \omega_1' u' \omega_2'$, $\alpha_{i+1} = \omega_1' v' \omega_2'$, $u' \to v'[M_1', M_2'] \in P'$, $\omega_1' \in M_1'$, $\omega_2' \in M_2'$. Then $\alpha_i^0 = \omega_1 u \omega_2$, $\alpha_{i+1}^0 = \omega_1 v \omega_2$, where $u \to v[M_1, M_2] \in P$, $\omega_1 \in M_1$, $\omega_2 \in M_2$; therefore $\alpha_i^0 \underset{G}{\Rightarrow} \alpha_{i+1}^0$.

We have proved $S \underset{G}{\Rightarrow}^* \alpha_n^0$, and since $\alpha_n^0 = \gamma^0 = \gamma$, $\gamma \in L(G)$ holds. Q.E.D.

Corollary 1. *If $\gamma \in L(G')$, then $S \underset{G'}{\Rightarrow}^* \gamma'$.*

Proof. If $\gamma \in L(G')$, then $\gamma \in L(G)$, and according to a) of the proof of the Lemma 1 $S \underset{G'}{\Rightarrow}^* \gamma'$ holds. Q.E.D.

Note. From the definition of grammar G' it follows that in each derivation $(\alpha_0, \alpha_1, \ldots \ldots, \alpha_n)$ over G' no rule of the form $u' \to v'[M_1', M_2']$ can be used after any rule of the form $t' \to t$. (Since $M_i' \subset (N \cup T')^*$ $(i = 1, 2)$, the rule $u' \to v'[M_1', M_2']$ can be applied only to the strings containing no terminal symbol.) Moreover, not the rule $u' \to v'$ but the "conditional" rule $u' \to v'[(N \cup T')^*, (N \cup T')^*]$ corresponds to the unconditional rule $u \to v$; the note can be applied even for such rules.

With the above introduced notation for grammars G and G' we define grammar $G^\#$ in the following way: $G^\# = \langle N^\#, T, P^\#, S^\# \rangle$, where $N^\# = N' \cup \{S^\#\} \cup \{\#\}$ ($S^\#$ and $\#$ are symbols different from all symbols of $N' \cup T$),

$$P^\# = \{S^\# \to \# S \#,\ \# \to \varepsilon\} \cup \{t' \leqq t \mid t \in T\} \cup$$
$$\cup \{u' \to v'[\{\#\} . M_1', M_2' . \{\#\}] / u' \to v'[M_1', M_2'] \in P'\}.$$

Then the following lemma holds:

Lemma 2. $L(G^\#) = L(G')$.

Proof. a) Let $\gamma \in L(G')$, i.e. there exists such a derivation of length n $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ over G' that $\alpha_n = \gamma$. But then $(S^\#, \#\alpha_0\#, \#\alpha_1\#, \ldots, \#\alpha_n\#, \#\alpha_n, \alpha_n)$ is the derivation over G of length $n + 3$, therefore $\gamma \in L(G)^\#)$.

b) Let $\gamma \in L(G^\#)$, i.e. there exists such a derivation of length $n + 3$ $(n \geqq 1)$ $(\beta_0, \beta_1, \beta_2, \ldots, \beta_{n+3})$ over $G^\#$ that $\beta_{n+3} = \gamma$. Obviously $\beta_0 = S^\#$, $\beta_1 = \# S \#$, $\beta_k = \widetilde{\#}_{k,1} \gamma_k \widetilde{\#}_{k,2}$ $(1 \leqq k \leqq n + 3)$, where $\gamma_k \in (N \cup T' \cup T)^*$ and $\widetilde{\#}_{k,1}$, $\widetilde{\#}_{k,2}$ mean either $\#$ or ε. We shall show that for each k $(1 \leqq k \leqq n + 2)$ is either $\gamma_k = \gamma_{k+1}$ or $\gamma_k \underset{G'}{\Rightarrow} \gamma_{k+1}$. As $\beta_k \underset{G^\#}{\Rightarrow} \beta_{k+1}$, then just one of the following four possibilities takes place:

1. there exist a rule $u' \to v'[\{\#\} . M_1', M_2' . \{\#\}] \in P^\#$ and strings $\omega_1' \in M_1'$, $\omega_2' \in M_2'$ such that $\omega_k = \# \omega_1' u' \omega_2' \#$, $\beta_{k+1} = \# \omega_1' v' \omega_2' \#$.
2. there exists a rule $t' \to t$ $(t \in T)$ such that $\beta_k = \chi_1 t' \chi_2$, $\beta_{k+1} = \chi_1 t \chi_2$.
3. $\widetilde{\#}_{k,1} = \#$, $\widetilde{\#}_{k+1,1} = \varepsilon$.
4. $\widetilde{\#}_{k,2} = \widetilde{\#}$, $\#_{k+1,2} = \varepsilon$.

In cases 1 and 2 obviously $\gamma_k \underset{G'}{\Rightarrow} \gamma_{k+1}$, in cases 3 and 4 $\gamma_k = \gamma_{k+1}$. From this it follows $\gamma_1 \underset{G'}{\Rightarrow}^* \gamma_k$ for $1 \leqq k \leqq n + 3$. Since $\gamma_1 = S$, $\gamma_{n+3} = \beta_{n+3} = \gamma$, we get $S \underset{G'}{\Rightarrow}^* \gamma$, therefore $\gamma \in L(G')$. Q.E.D.

Corollary 2. *If* $\gamma \in L(G^{\#})$ *then* $S^{\#} \underset{G^{\#}}{\Rightarrow}^* \#\gamma'\#$.

Proof. Let $\gamma \in L(G^{\#})$, i.e. $\gamma \in L(G')$ and according to Corollary 1 $S \underset{G'}{\Rightarrow}^* \gamma'$ holds. Then, of course, $S^{\#} \underset{G^{\#}}{\Rightarrow}^* \#\gamma'\#$. Q.E.D.

Definition 4. A CF-grammar with conditions $G = \langle N, T, P, S \rangle$ is *ε-free*, if for each its rule $u \to v[M_1, M_2] \in P$ $v \neq \varepsilon$ holds.

Definition 5. A CF-grammar with conditions $G = \langle N, T, P, S \rangle$ is called *CF-grammar with regular conditions*, if for each rule $u \to v[M_1, M_2] \in P$ the sets M_1, M_2 are regular over the alphabet $N \cup T$.

Theorem 1. *Let* $G = \langle N, T, P, S \rangle$ *be an ε-free CF-grammar with regular conditions. Then* $L(G)$ *is a context-sensitive language.*

Proof. According to Lemmas 1 and 2 $L(G) = L(G^{\#})$. For each rule p of the form $u \to v[M_1, M_2] \in P$ is $u' \to v'[\{\#\} . M_1', M_2' . \{\#\}] \in P^{\#}$, where the sets M_1', M_2' are regular over the alphabet $N \cup T'$. Therefore, there exist finite automata $\mathfrak{A}_{(1,p)}$, $\mathfrak{A}_{(2,p)}$, which accept sets M_1', M_2'. For each pair* $(i, p) \in \{1, 2\} \times P$, $\mathfrak{A}_{(i,p)} = \langle K_{(i,p)}, \delta_{(i,p)}, s^0_{(i,p)}, F_{(i,p)} \rangle$, where the used symbols have the following meaning: $K_{(i,p)}$ – finite nonempty set (of states), $\delta_{(i,p)}$ – mapping from $K_{(i,p)} \times (N \cup T')$ to $K_{(i,p)}$ (the transition function), $s^0_{(i,p)} \in K_{(i,p)}$ (the start state), $F_{(i,p)} \subset K_{(i,p)}$ (the set of final states).

The sets $K_{(i,p)}$ can be choosen so that $K_{(i,p)} \cap K_{(j,q)} = \emptyset$ for each two pairs $(i, p) \neq (j, q)$, $(i, p), (j, q) \in \{1, 2\} \times P$. Further**

$$\delta^*_{(i,p)}(s^0_{(i,p)}, \omega_i') \in F_{(i,p)} \Leftrightarrow \omega_i' \in M_i' .$$

Let us define the grammar $\bar{G} = \langle \bar{N}, T, \bar{P}, S^{\#} \rangle$ in the following way:

$$\bar{N} = N^{\#} \cup \bigcup_{(i,p)\in\{1,2\}\times P} (N \cup T' \cup \{\#\}) \times K_{(i,p)} \times \{1, 2, 3, 4\} ,$$

$$\bar{P} = \bigcup_{i=1}^{14} Qi .$$

* The symbol $\times$ denotes the operation of cartesian product. The elements of a cartesian product $A \times B \times C$ will be denoted by $[a, b, c]$.

** $\delta^*_{(i,p)}$ is the mapping from $K_{(i,p)} \times (N \cup T')^*$ to $K_{(i,p)}$ defined by the relations $\delta^*_{(i,p)} \cdot (s_{(i,p)}, \varepsilon) = s_{(i,p)}$, $\delta^*_{(i,p)}(s_{(i,p)}, ax) = \delta_{(i,p)}(\delta^*_{(i,p)}(s_{(i,p)}, a), x)$, which hold for whatever $s_{(i,p)} \in K_{(i,p)}$, $a \in (N \cup T')^*$, $x \in N \cup T'$.

 where

$Q_1 = \{S^\# \to \#S\#, \# \to \varepsilon\}$,

$Q_2 = \{t' \to t / t \in T\}$,

$Q_3 = \{\# \to [\#, s^0_{(1,p)}, 1] / p \in P\}$,

$Q_4 = \{[\alpha, s_{(1,p)}, 1] . \beta \to [\alpha, s_{(1,p)}, 1] . [\beta, \delta_{(1,p)} s_{(1,p)}, \beta), 1] /$
$/ \alpha \in N \cup T' \cup \{\#\}, \beta \in N \cup T', s_{(1,p)} \in K_{(1,p)}, p \in P\}$,

$Q_5 = \{[\alpha, s_{(1,p)}, 1] . u' \to [\alpha, s_{(1,p)}, 1] . [u', s^0_{(2,p)}, 2] /$
$/ \alpha \in N \cup T' \cup \{\#\}, s_{(1,p)} \in F_{(1,p)}, p = u \to v[M_1, M_2] \in P\}$,

$Q_6 = \{[u', s^0_{(2,p)}, 2] . \beta \to [u', s^0_{(2,p)}, 2] . [\beta, \delta_{(2,p)}(s^0_{(2,p)}, \beta), 1] /$
$/ \beta \in N \cup T', p = u \to v[M_1, M_2] \in P\}$,

$Q_7 = \{[\alpha, s_{(2,p)}, 1] . \beta \to [\alpha, s_{(2,p)}, 1] . [\beta, \delta_{(2,p)}(s_{(2,p)}, \beta), 1] /$
$/ \alpha \in N \cup T', \beta \in N \cup T', s_{(2,p)} \in K_{(2,p)}, p \in P\}$,

$Q_8 = \{[\alpha, s_{(2,p)}, 1] . \# \to [\alpha, s_{(2,p)}, 3] . \# / \alpha \in N \cup T', s_{(2,p)} \in F_{(2,p)}, p \in P\}$,

$Q_9 = \{[\alpha, s_{(2,p)}, 1] . [\beta, s'_{(2,p)}, 3] \to [\alpha, s_{(2,p)}, 3] . \beta /$
$/ \alpha, \beta \in N \cup T', s_{(2,p)}, s'_{(2,p)} \in K_{(2,p)}, p \in P\}$,

$Q_{10} = \{[u', s^0_{(2,p)}, 2] . [\beta, s_{(2,p)}, 3] \to [u', s^0_{(2,p)}, 4] . \beta /$
$/ \beta \in N \cup T', s_{(2,p)} \varepsilon K_{(2,p)}, p = u \to v[M_1, M_2] \varepsilon P\}$,

$Q_{11} = \{[u', s^0_{(2,p)}, 2] . \# \to [u', s^0_{(2,p)}, 4] . \# /$
$/ p = u \to v[M_1, M_2] \in P, s^0_{(2,p)} \in F_{(2,p)}\}$,

$Q_{12} = \{[\alpha, s_{(1,p)}, 1] . [u', s^0_{(2,p)}, 4] \to [\alpha, s_{(1,p)}, 3] . v' /$
$/ \alpha \in N \cup T' \cup \{\#\}, s_{(1,p)} \in K_{(1,p)}, p = u \to v[M_1, M_2] \in P\}$,

$Q_{13} = \{[\alpha, s_{(1,p)}, 1] . [\beta, s'_{(1,p)}, 3] \to [\alpha, s_{(1,p)}, 3] . \beta /$
$/ \alpha \in N \cup T' \cup \{\#\}, s_{(1,p)}, s'_{(1,p)} \in K_{(1,p)}, \beta \in N \cup T', p \in P\}$,

$Q_{14} = \{[\#, s^0_{(1,p)}, 3] \to \# / p \varepsilon P\}$.

The grammar $\bar{G}$ contains one length-shorting rule, namely the rule $\# \to \varepsilon$. However, in whatever derivation over $\bar{G}$ of arbitrary word $\gamma \in L(\bar{G})$ the rule $\# \to \varepsilon$ is used exactly twice and according to [2] (Theorem 3.1), $L(\bar{G})$ is context-sensitive language. By showing $L(\bar{G}) = L(G^\#)$, we prove that $L(G)$ is context-sensitive language.

a) Let $\gamma \in L(G^\#)$. According to Corollary 2 $S^\# \underset{G^\#}{\Rightarrow^*} \# \gamma' \#$ i.e. there exists a derivation $(\beta_0, \beta_1, \ldots, \beta_{n+1})$ over $G^\#$ such that $\beta_0 = S^\#$, $\beta_1 = \#S\#$, $\beta_{n+1} = \#\gamma'\#$ $(n \geqq 1)$. For each i $(1 \leqq i \leqq n)$ is therefore $\beta_i = \#\omega'_1 u' \omega'_2 \#$, $\beta_{i+1} = \#\omega'_1 v' \omega'_2 \#$, where $u' \to v'[\{\#\} . M'_1, M'_2 . \{\#\}] \in P^\#$, $\omega'_1 \in M'_1$, $\omega'_2 \in M'_2$.

We shall show that $\beta_i \underset{\bar{G}}{\Rightarrow}^* \beta_{i+1}$.

The rewriting β_i to β_{i+1} in $\bar{G}$ will be provided by succesive application of the following rules:

the rule of Q_3,

$|\omega'_1|$-times the rules of Q_4, *

the rule of Q_6 (since $\omega'_1 \in M'_1$, $s_{(1,p)} \in F_{(1,p)}$ holds and the rule of Q_5 can be applied),

the rule of Q_5 (if $\omega'_2 \neq \varepsilon$), (if $\omega'_2 = \varepsilon$ the rule of Q_{11} is applied),

$|\omega'_2| - 1$-times the rules of Q_7,

the rule of Q_8 (if $\omega'_2 \neq \varepsilon$),

$|\omega'_2| - 1$-times the rules of Q_9,

the rule of Q_{10} (if $\omega'_2 \neq \varepsilon$),

the rule of Q_{12},

$|\omega'_1|$-times the rules of Q_{13} and finally

the rule of Q_{14}.

Thus the required string β_{i+1} is obtained. This completes the proof that $S^\# \underset{\bar{G}}{\Rightarrow}^* \underset{\bar{G}}{\Rightarrow}^* \#\gamma'\#$ and therefore $\gamma \in L(\bar{G})$.

b) Let $\gamma \in L(\bar{G})$. We shall show first of all that $S^\# \underset{\bar{G}}{\Rightarrow}^* \#\gamma'\#$.

Let us introduce the homomorphism τ (cf. [3]) by the relation $\tau(x) = x'$ for $x \in T$, $\tau(x) = x$ for $x \in \bar{N}$. According to the assumption there exists such a derivation $(w_0, w_1, \ldots, w_n)$ over $\bar{G}$ that $w_0 = S^\#$, $w_n = \gamma$. From the relation $w_i \underset{\bar{G}}{\Rightarrow} w_{i+1}$ it follows that either $\tau(w_i) = \tau(w_{i+1})$ or $\tau(w_i) \underset{\bar{G}}{\Rightarrow} \tau(w_{i+1})$. Hence, $w_0 = \tau(w_0) \underset{\bar{G}}{\Rightarrow}^* \tau(w_n) = \gamma'$.

Therefore there exists a derivation $(\psi_0, \psi_1, \ldots, \psi_m, \ldots, \psi_s)$ over $\bar{G}$ such that $\psi_0 = S\#$, $\psi_1 = \#S\#$, $\psi_m = \#\chi_m$, $\psi_{m+1} = \chi_m$, $\psi_s = \gamma'$.

Let us denote $\tilde{\psi}_l = \#\psi_l$ for $l \geq m+1$. Then $(\psi_0, \psi_1, \ldots, \psi_m, \tilde{\psi}_{m+2}, \ldots, \tilde{\psi}_s)$ is the derivation over $\bar{G}$. Therefore $S^\# \underset{\bar{G}}{\Rightarrow}^* \#\gamma'$. We can prove similarly that $S^* \underset{\bar{G}}{\Rightarrow}^* \#\gamma'\#$.

Let us take the derivation $(\alpha_0, \alpha_1, \ldots, \alpha_w)$ over $\bar{G}$, where $\alpha_0 = S^\#$, $\alpha_1 = \#S\#, \ldots$ $\ldots, \alpha_w = \#\gamma'\#$. Obviously $S^\# \underset{G^\#}{\Rightarrow} \alpha_1$. It is sufficient to prove that for each k ($1 \leq k \leq$ $\leq w - 1$) the following statement holds: if $S^\# \underset{G^\#}{\Rightarrow}^* \alpha_k$, then there exists $l > k$ ($2 \leq l \leq w$) such that $\alpha_k \underset{G^\#}{\Rightarrow} \alpha_l$. In this case $S^\# \underset{G^\#}{\Rightarrow}^* \alpha_w$, further $\alpha_w \underset{G^\#}{\Rightarrow}^* \gamma$, hence $\gamma \in L(G^\#)$.

Therefore, let $S^\# \underset{G^\#}{\Rightarrow}^* \alpha_k$. Then $\alpha_k = \#\eta'_k\#$, where $\eta'_k \in (N \cup T')^*$. Further, there must exist a rule $\# \to [\#, s^0_{(1,p)}, 1]$ such that $\alpha_{k+1} = [\#, s^0_{(1,p)}, 1)\, \eta'_k\#$. By deriving from α_{k+1} to α_w any rule of Q_1 and Q_2 cannot be obviously applied.

* $|\omega|$ denotes the length of the string ω.

We shall show that there exists $z \geqq k+1$ such that the rewriting $\alpha_z \underset{\bar{G}}{\Rightarrow} \alpha_{z+1}$ is realized by the rule $[\alpha, s_{(1,p)}, 1] . u' \to [\alpha, s_{(1,p)}, 1] . [u', s^0_{(2,p)}, 2]$ of Q_5. Suppose the contrary is valid. Let $|\eta'_k| = \lambda$. Then fo reach i $(1 \leqq i \leqq \lambda)$ the rewriting $\alpha_{k+i} \underset{\bar{G}}{\Rightarrow} \alpha_{k+i+1}$ is realized by the rule of Q_4. (No rule of Q_3 can be applied to the symbol $\#$ at the end of the string α_{k+i}, because another rewriting of the formed symbol would be impossible.) To the string $\alpha_{k+\lambda+1}$ there can be applied only the rule of Q_3 and no rule can be applied after it, which gives the contradiction.

Let us take the smallest $z \geqq k+1$ with above mentioned property. Then $\alpha_k = \# \omega'_1 u' \omega'_2 \#$, where $\omega'_1 = c_1 c_2 \dots c_{z-(k+1)}$ $(c_i \in N \cup T', i = 1, 2, \dots, z-(k+1))$, $\omega'_1 = \varepsilon$, if $z = k+1$; $\omega'_2 \in (N \cup T')^*$. Further $\alpha_{k+1+i} = \tilde{c}_0 \tilde{c}_1 \dots \tilde{c}_i c_{i+1} \dots \dots c_{z-(k+1)} u' \omega'_2 \#$, where $\tilde{c}_0 = [\#, s^0_{(1,p)}, 1]$ and for each j $(1 \leqq j \leqq i)$ $\tilde{c}_j = [c_j, s^j_{(1,p)}, 1]$, where $s^j_{(1,p)} = \delta_{(1,p)}(s^{j-1}_{(1,p)}, c_j)$, $i = 0, 1, 2, \dots, z-(k+1)$. Further $\alpha_{z+1} = \tilde{c}_0 \dots \tilde{c}_{z-(k+1)} [u', s^0_{(2,p)}, 2] . \omega'_2 \#$. Since the rule of Q_5 was applied, $s^{z-(k+1)}_{(1,p)} = \delta^*_{(1,p)}(s^0_{(1,p)}, \omega'_1) \in F_{(1,p)}$, hence $\omega'_1 \in M'_1$ (the rule p is fixed).

We shall show that $\omega'_2 \in M'_2$ and that in mentioned derivation there appears a member α_q $(q \geqq z+2)$ such that $\alpha_q = \tilde{c}_0 \dots \tilde{c}_{z-(k+1)} . [u', s^0_{(2,p)}, 4] . \omega'_2 \#$. If $\omega'_2 = \varepsilon$, then rewriting $\alpha_{z+1} \underset{\bar{G}}{\Rightarrow} \alpha_{z+2}$ is necessarily realized by the rule of Q_{11} and therefore $s^0_{(2,p)} \in F_{(2,p)}$, i.e. $\varepsilon \in M'_2$ and the statement holds $(q = z+2)$. Let thus $\omega'_2 \neq \varepsilon$, $\omega'_2 = d_1 d_2 \dots d_f$ $(d_i \in N \cup T', i = 1, \dots, f; f \geqq 1)$. Then

$$\alpha_{z+2} = \tilde{c}_0 \dots \tilde{c}_{z-(k+1)} . [u', s^0_{(2,p)}, 2] . [d_1, \delta_{(2,p)}(s^0_{(2,p)}, d_1), 1] . d_2 \dots d_f \#$$

(applying the rule of Q_6), and obviously

$$\alpha_{z+1+f} = \tilde{c}_0 \dots \tilde{c}_{z-(k+1)} [u', s^0_{(2,p)}, 2] . \tilde{d}_1 \dots \tilde{d}_f \# ,$$

where

$$\tilde{d}_i = [d_i, s^i_{(2,p)}, 1] \quad (i = 1, \dots, f), \quad s^i_{(2,p)} = \delta_{(2,p)}(s^{i-1}_{(2,p)}, d_i)$$

(using the rules of Q_7). Then the derivation $\alpha_{z+1+f} \underset{\bar{G}}{\Rightarrow} \alpha_{z+2+f}$ is necessarily realized by the rule of Q_8, hence $s^f_{(2,p)} \in F_{(2,p)}$ and since $s^f_{(2,p)} = \delta^*_{(2,p)}(s^0_{(2,p)}, \omega'_2)$, $\omega'_2 \in M'_2$ holds. We can show in a similar way that in the mentioned derivation over $\bar{G}$ there follows $(f-1)$-times applying of the rules of Q_9, then one using of the rule of Q_{10}, by which the required string α_q is generated $(q = z+2+2f)$.

Then necessarily $\alpha_{q+1} = \tilde{c}_0 \dots \tilde{c}_{z-l-2} . [c_{z-(l+1)}, s^{z-(k+1)}_{(1,p)}, 3] . v' \omega'_2 \#$ (the rule of Q_{12}), after this $z-(l+1)$-times the rules of Q_{13} and finally the rule $[\#, S^0_{(1,p)}, 3] \to \#$ must be applied. In this way the member $\alpha_t = \# \omega'_1 v' \omega'_2 \#$ appears in the derivation. However, since $\alpha_t = \# \omega'_1 u' \omega'_2 \#$, $\omega'_1 \in M'_1$, $\omega'_2 \in M'_2$ and $u' \to v'[M'_1, M'_2] \in P'$, $\alpha_k \underset{G^\#}{\Rightarrow} \alpha_s$ holds and the Theorem is established.

Theorem 2. *Each context-sensitive language can be generated by an ε-free CF-grammar with regular conditions.*

Proof. Let G be a CS-grammar $G = \langle N, T, P, S\rangle$, where P is finite set of rules of the form $z_1 x z_2 \to z_1 y z_2$ ($x \in N$; $y, z_1, z_2 \in (N \cup T)^*$, $y \neq \varepsilon$). Let us form the CF-grammar with regular conditions $G_1 = \langle N, T, P_1, S\rangle$, where $P_1 = \{x \to y\, [(N \cup T)^* . \{z_1\}, \{z_2\} . (N \cup T)^*] / z_1 x z_2 \to z_1 y z_2 \in P\}$. The grammar G_1 is obviously ε-free and the sets $(N \cup T)^* \{z_1\}$, $\{z_2\} (N \cup T)^*$ are regular over the alphabet $N \cup T$. For each pair of strings u, v the relation $u \underset{G}{\Rightarrow} v$ holds if and only if $u \underset{G_1}{\Rightarrow} v$. Therefore $L(G) = L(G_1)$. Q.E.D.

Theorem 3. *Each type 0 language can be generated by a CF-grammar with regular conditions (generally not ε-free).*

Proof. It is known (cf. [1]) that each type 0 language can be generated by context-sensitive grammar with erasing rules. Let L be arbitrary type 0 language and G be a CS-grammar with erasing rules which generates L. G has the same form as in the proof of Theorem 2 with the distinction that also $y = \varepsilon$ is admissible. Forming the CF-grammar with regular conditions G_1 in the same way as in the proof of Theorem 2, we get $L(G_1) = L(G) = L$. Q.E.D.

Note. It has come to author's knowledge after finishing the paper that some modification of Theorem 1 has been proved in paper [4]. (The proof is, of course, different).

(Received June 23th, 1969.)

REFERENCES

[1] K. Čulík: Mathematical theory of languages. (To appear.)

[2] S. Ginsburg and S. Greibach: Mappings which preserve context-sensitive languages. Inf. and Control. *9* (1966), 563–582.

[3] S. Ginsburg: The mathematical theory of context-free languages. McGraw-Hill, 1966.

[4] I. Friš: Grammars with partial ordering of rules. Inf. and Control, *12* (1968), 415–425.

Bezkontextové gramatiky s regulárními podmínkami

Emil Navrátil

Bezkontextová gramatika $G = \langle N, T, P, S\rangle$ s regulárními podmínkami je bezkontextová gramatika, jejíž každé pravidlo je opatřeno dvojicí množin M_1, M_2, které jsou regulární nad abecedou $N \cup T$. Řetězec α lze bezprostředně přepsat na β ($\alpha \underset{G}{\Rightarrow} \beta$), jestliže existuje pravidlo $u \to v[M_1 . M_2] \in P$ a řetězce $\omega_1 \in M_1$, $\omega_2 \in M_2$ takové, že $\alpha = \omega_1 u \omega_2$, $\beta = \omega_1 v \omega_2$. Jazyk generovaný gramatikou G je množina všech terminálních produkcí odvozených z výchozího symbolu S.

Je dokázáno, že množina všech jazyků generovatelných ε-free gramatikami uvedeného typu je rovna třídě všech kontextových jazyků. Vypustíme-li předpoklad, že gramatika je ε-free, je možno generovat libovolný jazyk typu 0.

Ing. Emil Navrátil, Fakulta jaderná a fyzikálně inženýrská, Břehová 7, Praha 1.