Reduction of Simple Semi-Conditional Grammars with Respect to the Number of Conditional Productions

Alexander Meduna and Martin Švec*

Abstract

The present paper discusses the descriptional complexity of simple semi-conditional grammars with respect to the number of conditional productions. More specifically, it demonstrates that for every phrase-structure grammar, there exists an equivalent simple semi-conditional grammar that has no more than twelve conditional productions.

Keywords: descriptional complexity, simple semi-conditional grammars

1 Introduction

To describe languages as economically and succintly as possible, formal language theory has recently intensively investigated how to reduce grammars without any decrease of their power (see [1], [4], and [5]). Continuing with this vivid investigation, the present paper discusses the reduction of simple semi-conditional grammars, which characterize the family of recursively enumerable languages (see [2]).

More specifically, besides ordinary context-free productions, simple semi-conditional grammars may have some conditional productions which have an attached string representing a forbidding condition or a permitting condition. This paper concentrates its discussion on the reduction of simple semi-conditional grammars with respect to the number of conditional productions. It demonstrates that for every recursively enumerable language, there exists an equivalent simple semi-conditional grammar that has no more than twelve conditional productions.

2 Definitions

This paper assumes that the reader is familiar with the language theory (see [3]).

Let V be an alphabet. V^* denotes the free monoid generated by V under the operation of concatenation where ε denotes the unit of V^*. Let $V^+ = V^* -$

*Department of Computer Science and Engineering, Brno University of Technology, Božetěchova 2, Brno 61266, Czech Republic

$\{\varepsilon\}$. Given a word, $w \in V^*$, $|w|$ represents the length of w. We set $\mathrm{sub}(w) = \{y : y \text{ is a subword of } w\}$. Given a symbol, $a \in V$, $\#_a w$ denotes the number of occurences of a in w. For $w \in V^+$, $\mathrm{first}(w)$ denotes the leftmost symbol of w.

A *semi-conditional grammar* (an *sc*-grammar for short) is a quadruple, $G = (V, T, P, S)$, where V, T and S are the total alphabet, the terminal alphabet ($T \subset V$), and the axiom ($S \in V - T$), respectively, and P is a finite set of productions of the form $(A \to x, \alpha, \beta)$ with $A \in V - T$, $x \in V^*$, $\alpha \in V^+ \cup \{0\}$ and $\beta \in V^+ \cup \{0\}$, where 0 is a special symbol, $0 \notin V$ (intuitively, 0 means that the production's condition is missing). Production $(A \to x, \alpha, \beta) \in P$ is said to be conditional, if $\alpha \neq 0$ or $\beta \neq 0$. G has degree (i, j), where i and j are two natural numbers, if for every $(A \to x, \alpha, \beta) \in P$, $\alpha \in V^+$ implies $|\alpha| \leq i$, and $\beta \in V^+$ implies $|\beta| \leq j$. Let $u, v \in V^*$, and $(A \to x, \alpha, \beta) \in P$. Then, u directly derives v according to $(A \to x, \alpha, \beta)$ in G, denoted by

$$u \Rightarrow_G v \ [(A \to x, \alpha, \beta)]$$

provided for some $u_1, u_2 \in V^*$, the following conditions (a) through (d) hold

(a) $u = u_1 A u_2$,

(b) $v = u_1 x u_2$,

(c) $\alpha \neq 0$ implies $\alpha \in \mathrm{sub}(u)$,

(d) $\beta \neq 0$ implies $\beta \notin \mathrm{sub}(u)$.

When no confusion exists, we simply write $u \Rightarrow_G v$. As usual, we extend $\Rightarrow_G$ to $\Rightarrow_G^i$ (where $i \geq 0$), $\Rightarrow_G^+$, and $\Rightarrow_G^*$. The language of G, denoted by $L(G)$, is defined as $L(G) = \{w \in T^* : S \Rightarrow_G^* w\}$.

Based upon the concept of *sc*-grammars, Meduna and Gopalaratnam [2] have defined a *simple semi-conditional grammar* (an *ssc*-grammar for short) as an *sc*-grammar in which every production has no more than one condition. Formally, let $G = (V, T, P, S)$ be an *sc*-grammar. G is a simple semi-conditional grammar if $(A \to x, \alpha, \beta) \in P$ implies $\{0\} \subseteq \{\alpha, \beta\}$.

3 Results

Theorem 1 *Every recursively enumerable language can be defined by a simple semi-conditional grammar of degree (2,1) with no more than 12 conditional productions.*

Proof. Let L be a recursively enumerable language. By Geffert [1], we can assume that L is generated by a grammar G of the form

$$G = (V, T, P \cup \{AB \to \varepsilon, CD \to \varepsilon\}, S)$$

such that P contains only context-free productions and

$$V - T = \{S, A, B, C, D\}.$$

We construct an *ssc*-grammar G' of degree (2,1) as follows:

$$G' = (V', T, P', S), \quad \text{where}$$
$$V' = V \cup W,$$
$$W = \{\widetilde{A}, \widetilde{B}, \langle\varepsilon_A\rangle, \$, \widetilde{C}, \widetilde{D}, \langle\varepsilon_C\rangle, \#\},\ V \cap W = \emptyset.$$

The set of productions P' is defined in the following way:

1. if $H \to \alpha \in P$, $H \in V - T$, $\alpha \in V^*$, then add $(H \to \alpha, 0, 0)$ to P';
2. add the following six productions to P':

$$(A \to \widetilde{A}, 0, \widetilde{A}),$$
$$(B \to \widetilde{B}, 0, \widetilde{B}),$$
$$(\widetilde{A} \to \langle\varepsilon_A\rangle, \widetilde{A}\widetilde{B}, 0),$$
$$(\widetilde{B} \to \$, \langle\varepsilon_A\rangle\widetilde{B}, 0),$$
$$(\langle\varepsilon_A\rangle \to \varepsilon, 0, \widetilde{B}),$$
$$(\$ \to \varepsilon, 0, \langle\varepsilon_A\rangle);$$

3. add the following six productions to P':

$$(C \to \widetilde{C}, 0, \widetilde{C}),$$
$$(D \to \widetilde{D}, 0, \widetilde{D}),$$
$$(\widetilde{C} \to \langle\varepsilon_C\rangle, \widetilde{C}\widetilde{D}, 0),$$
$$(\widetilde{D} \to \#, \langle\varepsilon_C\rangle\widetilde{D}, 0),$$
$$(\langle\varepsilon_C\rangle \to \varepsilon, 0, \widetilde{D}),$$
$$(\# \to \varepsilon, 0, \langle\varepsilon_C\rangle).$$

Next, we prove that $L(G') = L(G)$.

Basic idea: Notice that G' has degree (2,1) and contains only 12 conditional productions. The productions of (2) simulate the application of $AB \to \varepsilon$ in G' and the productions of (3) simulate the application of $CD \to \varepsilon$ in G'.

Let us describe the simulation of $AB \to \varepsilon$. First, one occurence of A and one occurence of B are rewritten to $\widetilde{A}$ and $\widetilde{B}$, respectively (no more than one $\widetilde{A}$ and one $\widetilde{B}$ appear in any sentential form). The right neighbor of $\widetilde{A}$ is checked to be $\widetilde{B}$ and $\widetilde{A}$ is rewritten to $\langle\varepsilon_A\rangle$. Then, analogously, the left neighbor of $\widetilde{B}$ is checked to be $\langle\varepsilon_A\rangle$ and $\widetilde{B}$ is rewritten to \$. Finally, $\langle\varepsilon_A\rangle$ and \$ are erased. The simulation of $CD \to \varepsilon$ is analogous.

To establish $L(G) = L(G')$, we first prove the following two claims.

Claim 1 *$S \Rightarrow^*_{G'} x'$ implies $\#_{\widetilde{X}} x' \leq 1$ for all $\widetilde{X} \in \{\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}\}$ and some $x' \in (V')^*$.*

Proof. By inspection of productions in P', the only production that can generate $\widetilde{X}$ is of the form $(X \to \widetilde{X}, 0, \widetilde{X})$. This production can be applied only when no $\widetilde{X}$ occurs in the rewritten sentential form. Thus, it is not possible to derive x' from S such that $\#_{\widetilde{X}} x' \geq 2$. □

Informally, next claim says that every occurence of $\langle \varepsilon_A \rangle$ in derivations from S is always followed either by $\widetilde{B}$ or \$, and every occurence of $\langle \varepsilon_C \rangle$ is always followed either by $\widetilde{D}$ or #.

Claim 2 *It holds that*

A) $S \Rightarrow^*_{G'} y'_1 \langle \varepsilon_A \rangle y'_2$ *implies* $y'_2 \in (V')^+ \wedge \mathrm{first}(y'_2) \in \{\widetilde{B}, \$\}$ *for any* $y'_1 \in (V')^*$*;*

B) $S \Rightarrow^*_{G'} y'_1 \langle \varepsilon_C \rangle y'_2$ *implies* $y'_2 \in (V')^+ \wedge \mathrm{first}(y'_2) \in \{\widetilde{D}, \#\}$ *for any* $y'_1 \in (V')^*$*.*

Proof. We establish the proof by the examination of all possible forms of derivations that may occur when deriving a sentential form containing $\langle \varepsilon_A \rangle$ or $\langle \varepsilon_C \rangle$.

A) By the definition of P', the only production that can generate $\langle \varepsilon_A \rangle$ is $p = (\widetilde{A} \to \langle \varepsilon_A \rangle, \widetilde{A}\widetilde{B}, 0)$. This production has the permitting condition $\widetilde{A}\widetilde{B}$, so it can be used provided that $\widetilde{A}\widetilde{B}$ occurs in a sentential form. Furthermore, by Claim 1, no other occurence of $\widetilde{A}$ or $\widetilde{B}$ can appear in the given sentential form. Consequently, we obtain a derivation

$$S \Rightarrow^*_{G'} u'_1 \widetilde{A}\widetilde{B} u'_2 \Rightarrow_{G'} u'_1 \langle \varepsilon_A \rangle \widetilde{B} u'_2 \ [p]$$

for some $u'_1, u'_2 \in (V')^*$, $\widetilde{A}, \widetilde{B} \notin \mathrm{sub}(u'_1 u'_2)$, which represents the only way how to get $\langle \varepsilon_A \rangle$. Obviously, $\langle \varepsilon_A \rangle$ is always followed by $\widetilde{B}$ in $u'_1 \langle \varepsilon_A \rangle \widetilde{B} u'_2$.

Next, we discuss how G' can rewrite the subword $\langle \varepsilon_A \rangle \widetilde{B}$ in $u'_1 \langle \varepsilon_A \rangle \widetilde{B} u'_2$. There are only two productions having the nonterminals $\langle \varepsilon_A \rangle$ or $\widetilde{B}$ on their left-hand side—$p_1 = (\widetilde{B} \to \$, \langle \varepsilon_A \rangle \widetilde{B}, 0)$ and $p_2 = (\langle \varepsilon_A \rangle \to \varepsilon, 0, \widetilde{B})$. G' cannot use p_2 to erase $\langle \varepsilon_A \rangle$ in $u'_1 \langle \varepsilon_A \rangle \widetilde{B} u'_2$ because p_2 forbids an occurence of $\widetilde{B}$ in the string to be rewritten. Production p_1 has also a context condition, but $\langle \varepsilon_A \rangle \widetilde{B} \in \mathrm{sub}(u'_1 \langle \varepsilon_A \rangle \widetilde{B} u'_2)$ and thus p_1 can be used to rewrite $\widetilde{B}$ with \$. Hence, we obtain a derivation of the form

$$\begin{array}{lll} S & \Rightarrow^*_{G'} u'_1 \widetilde{A}\widetilde{B} u'_2 & \Rightarrow_{G'} u'_1 \langle \varepsilon_A \rangle \widetilde{B} u'_2 \quad [p] \\ & \Rightarrow^*_{G'} v'_1 \langle \varepsilon_A \rangle \widetilde{B} v'_2 & \Rightarrow_{G'} v'_1 \langle \varepsilon_A \rangle \$ v'_2 \quad [p_1]. \end{array}$$

Notice that during this derivation, G' may rewrite u'_1 and u'_2 to some v'_1 and v'_2, respectively ($v'_1, v'_2 \in (V')^*$); however, $\langle \varepsilon_A \rangle \widetilde{B}$ remains unchanged after this rewriting.

In this derivation we obtained the second symbol, \$, that can appear as the right neighbor of $\langle \varepsilon_A \rangle$. It suffices to show that there is no other symbol that could appear immediately after $\langle \varepsilon_A \rangle$. By inspection of P', only $(\$ \to \varepsilon, 0, \langle \varepsilon_A \rangle)$ can rewrite \$. However, this production cannot be applied when $\langle \varepsilon_A \rangle$ occurs in the given sentential form. In other words, the occurence of \$ in the subword $\langle \varepsilon_A \rangle \$$

cannot be rewritten before $\langle\varepsilon_A\rangle$ is erased by the production p_2. Hence, $\langle\varepsilon_A\rangle$ is always followed either by $\widetilde{B}$ or \$ and thus the first part of Claim 2 holds.

B) By inspection of productions simulating $AB \to \varepsilon$ and $CD \to \varepsilon$ in G' (see (2) and (3) in the definition of P'), these two sets of productions work analogously. Thus, part B of Claim 2 can be proven by analogy with part A. □

Let us return to the main part of the proof. Let g be a finite substitution from $(V')^*$ to V^* defined as follows:

1. for all $X \in V: \; g(X) = \{X\}$;
2. $g(\widetilde{A}) = \{A\}, \; g(\widetilde{B}) = \{B\}, g(\langle\varepsilon_A\rangle) = \{A\}, \; g(\$) = \{B, AB\}$;
3. $g(\widetilde{C}) = \{C\}, \; g(\widetilde{D}) = \{D\}, g(\langle\varepsilon_C\rangle) = \{C\}, \; g(\#) = \{C, CD\}$.

Having this substitution, we can now prove the following claim:

Claim 3 *$S \Rightarrow_G^* x$ if and only if $S \Rightarrow_{G'}^* x'$ for some $x \in g(x')$, $x \in V^*$, $x' \in (V')^*$.*

Proof. The claim is proven by induction on the length of derivations.
Only if: We show that

$$S \Rightarrow_G^m x \quad \text{implies} \quad S \Rightarrow_{G'}^* x,$$

where $m \geq 0$, $x \in V^*$; clearly $x \in g(x)$. This is established by induction on m.
Basis: Let $m = 0$. That is, $S \Rightarrow_G^0 S$. Clearly, $S \Rightarrow_{G'}^0 S$.
Induction Hypothesis: Suppose that the claim holds for all derivations of length m or less, for some $m \geq 0$.
Induction Step: Let us consider a derivation $S \Rightarrow_G^{m+1} x, \; x \in V^*$. Since $m+1 \geq 1$, there is some $y \in V^+$ and $p \in P \cup \{AB \to \varepsilon, CD \to \varepsilon\}$ such that $S \Rightarrow_G^m y \Rightarrow_G x \; [p]$. By the induction hypothesis, there is a derivation $S \Rightarrow_{G'}^* y$.

There are three cases that cover all possible forms of the production p:

(i) $p = H \to y_2 \in P, \; H \in V - T, \; y_2 \in V^*$. Then, $y = y_1 H y_3$ and $x = y_1 y_2 y_3$, $y_1, y_3 \in V^*$. Because we have $(H \to y_2, 0, 0) \in P'$, $S \Rightarrow_{G'}^* y_1 H y_3 \Rightarrow_{G'} y_1 y_2 y_3 \; [(H \to y_2, 0, 0)]$ and $y_1 y_2 y_3 = x$.

(ii) $p = AB \to \varepsilon$. Then, $y = y_1 AB y_3$ and $x = y_1 y_3, \; y_1, y_3 \in V^*$. In this case, there is the following derivation:

$$\begin{array}{llll}
S & \Rightarrow_{G'}^* & y_1 AB y_3 & \\
 & \Rightarrow_{G'} & y_1 \widetilde{A} B y_3 & [(A \to \widetilde{A}, 0, \widetilde{A})] \\
 & \Rightarrow_{G'} & y_1 \widetilde{A}\widetilde{B} y_3 & [(B \to \widetilde{B}, 0, \widetilde{B})] \\
 & \Rightarrow_{G'} & y_1 \langle\varepsilon_A\rangle \widetilde{B} y_3 & [(\widetilde{A} \to \langle\varepsilon_A\rangle, \widetilde{A}\widetilde{B}, 0)] \\
 & \Rightarrow_{G'} & y_1 \langle\varepsilon_A\rangle \$ y_3 & [(\widetilde{B} \to \$, \langle\varepsilon_A\rangle\widetilde{B}, 0)] \\
 & \Rightarrow_{G'} & y_1 \$ y_3 & [(\langle\varepsilon_A\rangle \to \varepsilon, 0, \widetilde{B})] \\
 & \Rightarrow_{G'} & y_1 y_3 & [(\$ \to \varepsilon, 0, \langle\varepsilon_A\rangle)].
\end{array}$$

(iii) $p = CD \to \varepsilon$. Then, $y = y_1CDy_3$ and $x = y_1y_3$, $y_1, y_3 \in V^*$. By analogy with (ii), there exists the derivation $S \Rightarrow^*_{G'} y_1CDy_3 \Rightarrow^6_{G'} y_1y_3$.

If: By induction on the length n of derivations in G', we prove that

$$S \Rightarrow^n_{G'} x' \quad \text{implies} \quad S \Rightarrow^*_G x$$

for some $x \in g(x')$, $x \in V^*$, $x' \in (V')^*$.
Basis: Let $n = 0$. That is, $S \Rightarrow^0_{G'} S$. It is obvious that $S \Rightarrow^0_G S$ and $S \in g(S)$.
Induction Hypothesis: Assume that the claim holds for all derivations of length n or less, for some $n \geq 0$.
Induction Step: Consider a derivation $S \Rightarrow^{n+1}_{G'} x'$, $x' \in (V')^*$. Since $n + 1 \geq 1$, there is some $y' \in (V')^+$ and $p' \in P'$ such that $S \Rightarrow^n_{G'} y' \Rightarrow_{G'} x'$ $[p']$, and by the induction hypothesis, there is also a derivation $S \Rightarrow^*_G y$ such that $y \in g(y')$.

By inspection of P', the following cases (i) through (xiii) cover all possible forms of p':

(i) $p' = (H \to y_2, 0, 0) \in P'$, $H \in V - T$, $y_2 \in V^*$. Then, $y' = y_1'Hy_3'$, $x' = y_1'y_2y_3'$, $y_1', y_3' \in (V')^*$ and y has the form $y = y_1Zy_3$, where $y_1 \in g(y_1')$, $y_3 \in g(y_3')$ and $Z \in g(H)$. Because for all $X \in V - T : g(X) = \{X\}$, the only Z is H and thus $y = y_1Hy_3$. By the definition of P' (see (1)), there exists a production $p = H \to y_2$ in P and we can construct the derivation $S \Rightarrow^*_G y_1Hy_3 \Rightarrow_G y_1y_2y_3$ $[p]$ such that $y_1y_2y_3 = x$, $x \in g(x')$.

(ii) $p' = (A \to \widetilde{A}, 0, \widetilde{A})$. Then, $y' = y_1'Ay_3'$, $x' = y_1'\widetilde{A}y_3'$, $y_1', y_3' \in (V')^*$ and $y = y_1Zy_3$, where $y_1 \in g(y_1')$, $y_3 \in g(y_3')$ and $Z \in g(A)$. Because $g(A) = \{A\}$, the only Z is A, so we can express $y = y_1Ay_3$. Having the derivation $S \Rightarrow^*_G y$ such that $y \in g(y')$, it is easy to see that also $y \in g(x')$ because $A \in g(\widetilde{A})$.

(iii) $p' = (B \to \widetilde{B}, 0, \widetilde{B})$. By analogy with (ii), $y' = y_1'By_3'$, $x' = y_1'\widetilde{B}y_3'$, $y = y_1By_3$, where $y_1', y_3' \in (V')^*$, $y_1 \in g(y_1')$, $y_3 \in g(y_3')$ and thus $y \in g(x')$ because $B \in g(\widetilde{B})$.

(iv) $p' = (\widetilde{A} \to \langle\varepsilon_A\rangle, \widetilde{A}\widetilde{B}, 0)$. By the permitting condition of this production, $\widetilde{A}\widetilde{B}$ surely occurs in y'. By Claim 1, no more than one $\widetilde{A}$ can occur in y'. Therefore, y' must be of the form $y' = y_1'\widetilde{A}\widetilde{B}y_3'$, where $y_1', y_3' \in (V')^*$ and $\widetilde{A} \notin \text{sub}(y_1'y_3')$. Then, $x' = y_1'\langle\varepsilon_A\rangle\widetilde{B}y_3'$ and y is of the form $y = y_1Zy_3$, where $y_1 \in g(y_1')$, $y_3 \in g(y_3')$ and $Z \in g(\widetilde{A}\widetilde{B})$. Because $g(\widetilde{A}\widetilde{B}) = \{AB\}$, the only Z is AB; thus, we obtain $y = y_1ABy_3$. By the induction hypothesis, we have a derivation $S \Rightarrow^*_G y$ such that $y \in g(y')$. According to the definition of g, $y \in g(x')$ as well because $A \in g(\langle\varepsilon_A\rangle)$ and $B \in g(\widetilde{B})$.

(v) $p' = (\widetilde{B} \to \$, \langle\varepsilon_A\rangle\widetilde{B}, 0)$. This production can be applied provided that $\langle\varepsilon_A\rangle\widetilde{B} \in \text{sub}(y')$. Moreover, by Claim 1, $\#_{\widetilde{B}}y' \leq 1$. Hence, we can express $y' = y_1'\langle\varepsilon_A\rangle\widetilde{B}y_3'$, where $y_1', y_3' \in (V')^*$ and $\widetilde{B} \notin \text{sub}(y_1'y_3')$. Then, $x' = y_1'\langle\varepsilon_A\rangle\$y_3'$ and $y = y_1Zy_3$, where $y_1 \in g(y_1')$, $y_3 \in g(y_3')$ and

$Z \in g(\langle \varepsilon_A \rangle \widetilde{B})$. By the definition of g, $g(\langle \varepsilon_A \rangle \widetilde{B}) = \{AB\}$, so $Z = AB$ and $y = y_1 AB y_3$. By the induction hypothesis, we have a derivation $S \Rightarrow_G^* y$ such that $y \in g(y')$. Because $A \in g(\langle \varepsilon_A \rangle)$ and $B \in g(\$)$, $y \in g(x')$ as well.

(vi) $p' = (\langle \varepsilon_A \rangle \rightarrow \varepsilon, 0, \widetilde{B})$. Application of $(\langle \varepsilon_A \rangle \rightarrow \varepsilon, 0, \widetilde{B})$ implies that $\langle \varepsilon_A \rangle$ occurs in y'. Claim 2 says that $\langle \varepsilon_A \rangle$ has either $\widetilde{B}$ or \$ as its right neighbor. Since the forbidding condition of p' forbids an occurence of $\widetilde{B}$ in y', the right neighbor of $\langle \varepsilon_A \rangle$ must be \$. As a result, we obtain $y' = y_1' \langle \varepsilon_A \rangle \$ y_3'$ where $y_1', y_3' \in (V')^*$. Then, $x' = y_1' \$ y_3'$ and y is of the form $y = y_1 Z y_3$, where $y_1 \in g(y_1')$, $y_3 \in g(y_3')$ and $Z \in g(\langle \varepsilon_A \rangle \$)$. By the definition of g, $g(\langle \varepsilon_A \rangle \$) = \{AB, AAB\}$. If $Z = AB$, $y = y_1 AB y_3$. Having the derivation $S \Rightarrow_G^* y$, it holds that $y \in g(x')$ because $AB \in g(\$)$.

(vii) $p' = (\$ \rightarrow \varepsilon, 0, \langle \varepsilon_A \rangle)$. Then, $y' = y_1' \$ y_3'$ and $x' = y_1' y_3'$, where $y_1', y_3' \in (V')^*$. Express $y = y_1 Z y_3$ so that $y_1 \in g(y_1')$, $y_3 \in g(y_3')$ and $Z \in g(\$)$, where $g(\$) = \{B, AB\}$. Let $Z = AB$. Then, $y = y_1 AB y_3$ and there exists the derivation $S \Rightarrow_G^* y_1 AB y_3 \Rightarrow_G y_1 y_3$ $[AB \rightarrow \varepsilon]$, where $y_1 y_3 = x$, $x \in g(x')$.

In cases (ii) through (vii) we discussed all six productions simulating the application of $AB \rightarrow \varepsilon$ in G' (see (2) in the definition of P'). Cases (viii) – (xiii) should cover productions simulating the application of $CD \rightarrow \varepsilon$ in G' (see (3)). However, by inspection of these two sets of productions, it is easy to see that they work analogously. Therefore, we leave this part of the proof to the reader (it can be established by analogy with (ii) – (vii) by replacing nonterminals $A, B, \widetilde{A}, \widetilde{B}, \langle \varepsilon_A \rangle$ and \$ with $C, D, \widetilde{C}, \widetilde{D}, \langle \varepsilon_C \rangle$ and #).

We have completed the proof and established Claim 3 by the principle of induction. □

Observe that $L(G) = L(G')$ follows from Claim 3. Indeed, according to the definition of g, we have $g(a) = \{a\}$ for all $a \in T$. Thus, from Claim 3, we have for any $x \in T^*$:

$$S \Rightarrow_G^* x \quad \text{if and only if} \quad S \Rightarrow_{G'}^* x.$$

Consequently, $L(G) = L(G')$ and the theorem holds. ■

In fact, the previous proof established more than stated in Theorem 1. Indeed, it also reduced the number of nonterminals as the next corollary says.

Corollary 1 *Every recursively enumerable language can be generated by a simple semi-conditional grammar of degree (2,1) with no more than 12 conditional productions and 13 nonterminals.*

Proof. Observe that G' has 13 nonterminals in the proof of Theorem 1. □

The above corollary tells us that besides the number of conditional productions, we have also reduced the semi-conditional grammars with respect to the number of nonterminals. In addition, we were able to establish this result for semi-conditional grammars of degree (2,1). This result gives rise to a question of whether we can further reduce the number of conditional productions in the semi-conditional grammars of any degree. In other words, consider the semi-conditional grammars with productions having context conditions of any length. Can they generate any recursively enumerable language with fewer than 12 conditional productions?

Acknowledgment

The authors thank the anonymous referee for several useful comments regarding the first version of this paper.

References

[1] Geffert, V.: How to generate languages using only two pairs of parentheses. J. Inform. Process. Cybernet, EIK27, 781–786 (1990).

[2] Meduna, A., and Gopalaratnam, A.: On Semi-Conditional Grammars with Productions Having either Forbidding or Permitting Conditions, Acta Cybernetica 11 (1994), 307–323.

[3] Meduna, A.: Automata and Languages: Theory and Applications. Springer, London, 2000.

[4] Proceedings of the International Workshop on Descriptional Complexity of Automata, Grammars and Related Structures, July 20–23, 1999, Ed. by J. Dassow and D. Wotschke, Otto-von-Guericke University, Magdeburg, Germany, 1999.

[5] Proceedings of the Second Workshop on Descriptional Complexity of Automata, Grammars and Related Structures, July 27–29, 2000, London, Canada.

Received February, 2001