4.2.2 Forbidding ET0L Grammars

In this section, we discuss forbidding ET0L grammars (see [138]). First, we define forbidding ET0L grammars. Then, we establish their generative power.

Definition 17. Let $G = (V, T, P_1, \ldots, P_t, S)$ be a CET0L grammar. If every $p = (a \to x, Per, For) \in P_i$, where $i = 1, \ldots, t$, satisfies $Per = \emptyset$, then G is said to be *forbidding ET0L grammar* (an *FET0L grammar* for short). If G is a propagating FET0L grammar, than G is said to be an *FEPT0L grammar*. If $t = 1$, G is called an *FE0L grammar*. If G is a propagating FE0L grammar, G is called an *FEP0L grammar*.

Convention 4. Let $G = (V, T, P_1, \ldots, P_t, S)$ be an FET0L grammar of degree (r, s). Clearly, $(a \to x, Per, For) \in P_i$ implies $Per = \emptyset$ for all $i = 1, \ldots, t$. By analogy with sequential forbidding grammars, we thus omit the empty set in the productions. For simplicity, we also say that G's degree is s instead of (r, s).

The families of languages defined by FE0L grammars, FEP0L grammars, FET0L grammars, and FEPT0L grammars of degree s are denoted by $\mathbf{FE0L}(s)$, $\mathbf{FEP0L}(s)$, $\mathbf{FET0L}(s)$, and $\mathbf{FEPT0L}(s)$, respectively. Moreover,

$$\mathbf{FEPT0L} = \bigcup_{s=0}^{\infty} \mathbf{FEPT0L}(s), \qquad \mathbf{FET0L} = \bigcup_{s=0}^{\infty} \mathbf{FET0L}(s),$$

$$\mathbf{FEP0L} = \bigcup_{s=0}^{\infty} \mathbf{FEP0L}(s), \qquad \mathbf{FE0L} = \bigcup_{s=0}^{\infty} \mathbf{FE0L}(s).$$

Example 8. Let

$$G = (\{S, A, B, C, a, \bar{a}, b\}, \{a, b\}, P, S)$$

be an FEP0L grammar, where

$$\begin{aligned} P = \{ & (S \to ABA, \emptyset), \\ & (A \to aA, \{\bar{a}\}), \\ & (B \to bB, \emptyset), \\ & (A \to \bar{a}, \{\bar{a}\}), \\ & (\bar{a} \to a, \emptyset), \\ & (B \to C, \emptyset), \\ & (C \to bC, \{A\}), \\ & (C \to b, \{A\}), \\ & (a \to a, \emptyset), \\ & (b \to b, \emptyset)\}. \end{aligned}$$

Obviously, G is an FEP0L grammar of degree 1. Observe that for every word from $L(G)$, there exists a derivation of the form

$$\begin{aligned} S & \Rightarrow_G ABA \\ & \Rightarrow_G aAbBaA \\ & \Rightarrow_G^+ a^{m-1}Ab^{m-1}Ba^{m-1}A \\ & \Rightarrow_G a^{m-1}\bar{a}b^{m-1}Ca^{m-1}\bar{a} \\ & \Rightarrow_G a^m b^m C a^m \\ & \Rightarrow_G^+ a^m b^{n-1} C a^m \\ & \Rightarrow_G a^m b^n a^m, \end{aligned}$$

with $1 \leq m \leq n$. Hence,

$$L(G) = \{a^m b^n a^m : \ 1 \leq m \leq n\}.$$

Note that $L(G) \notin \mathbf{E0L}$ (see page 268 in Volume 1 of [158]); however, $L(G) \in \mathbf{FEP0L}(1)$. As a result, FEP0L grammars (of degree 1) are more powerful than ordinary E0L grammars.

Next, we investigate the generative power of FET0L grammars of all degrees.

Theorem 36. *$\mathbf{FEPT0L}(0) = \mathbf{EPT0L}$, $\mathbf{FET0L}(0) = \mathbf{ET0L}$, $\mathbf{FEP0L}(0) = \mathbf{EP0L}$, and $\mathbf{FE0L}(0) = \mathbf{E0L}$.*

Proof. It follows from the definition of FET0L grammars. □

Lemmas 13, 14, 15, and 16 inspect the generative power of forbidding ET0L grammars of degree 1. As a conclusion, in Theorem 37, we demonstrate that both FEPT0L(1) and FET0L(1) grammars generate precisely the family of ET0L languages.

Lemma 13. *$\mathbf{EPT0L} \subseteq \mathbf{FEP0L}(1)$.*

Proof. Let

$$G = (V, T, P_1, \ldots, P_t, S)$$

be an EPT0L grammar, where $t \geq 1$. Set

$$W = \{\langle a, i\rangle : a \in V, \ i = 1, \ldots, t\}$$

and

$$F(i) = \{\langle a, j\rangle \in W : j \neq i\}.$$

Then, construct an FEP0L grammar of degree 1,

$$G' = (V', T, P', S),$$

where

$$V' = V \cup W, \ (V \cap W = \emptyset),$$

and the set of productions P' is defined as follows:

1. for each $a \in V$ and $i = 1, \ldots, t$, add $(a \to \langle a, i\rangle, \emptyset)$ to P';
2. if $a \to z \in P_i$ for some $i \in \{1, \ldots, t\}$, $a \in V$, $z \in V^+$, add $(\langle a, i\rangle \to z, F(i))$ to P'.

Let us demonstrate that $L(G) = L(G')$.

Claim 23. *For each derivation $S \Rightarrow^n_{G'} x$, $n \geq 0$,*

(I) *if $n = 2k + 1$ for some $k \geq 0$, $x \in W^+$;*

(II) *if $n = 2k$ for some $k \geq 0$, $x \in V^+$.*

Proof. The claim follows from the definition of P'. Indeed, every production in P' is either of the form $(a \rightarrow \langle a, i\rangle, \emptyset)$ or $(\langle a, i\rangle \rightarrow z, F(i))$, where $a \in V$, $\langle a, i\rangle \in W$, $z \in V^+$, $i \in \{1, \dots, t\}$. Since $S \in V$,

$$S \Rightarrow_{G'}^{2k+1} x \quad \text{implies} \quad x \in W^+$$

and

$$S \Rightarrow_{G'}^{2k} x \quad \text{implies} \quad x \in V^+;$$

thus, the claim holds. □

Define the finite substitution g from V^* to $(V')^*$ such that for every $a \in V$,

$$g(a) = \{a\} \cup \{\langle a, i\rangle \in W : i = 1, \dots, t\}.$$

Claim 24. *$S \Rightarrow_G^* x$ if and only if $S \Rightarrow_{G'}^* x'$ for some $x' \in g(x)$, $x \in V^+$, $x' \in (V')^+$.*

Proof.

Only If: By induction on $n \geq 0$, we show that for all $x \in V^+$,

$$S \Rightarrow_G^n x \quad \text{implies} \quad S \Rightarrow_{G'}^{2n} x.$$

Basis: Let $n = 0$. Then, the only x is S; therefore, $S \Rightarrow_G^0 S$ and also $S \Rightarrow_{G'}^0 S$.

Induction Hypothesis: Suppose that

$$S \Rightarrow_G^n x \quad \text{implies} \quad S \Rightarrow_{G'}^{2n} x$$

for all derivations of length n or less, for some $n \geq 0$.

Induction Step: Consider $S \Rightarrow_G^{n+1} x$. Because $n + 1 \geq 1$, we can express

$$S \Rightarrow_G^{n+1} x$$

as

$$S \Rightarrow_G^n y \Rightarrow_G x \ [p_1, p_2, \dots, p_q]$$

such that $y \in V^+$, $q = |y|$, and $p_j \in P_i$ for all $j = 1, \dots, q$ and some $i \in \{1, \dots, t\}$. By the induction hypothesis,

$$S \Rightarrow_{G'}^{2n} y.$$

Suppose that $y = a_1 a_2 \dots a_q$, $a_j \in V$. Let G' make the derivation

$$\begin{array}{lll} S & \Rightarrow_{G'}^{2n} & a_1 a_2 \dots a_q \\ & \Rightarrow_{G'} & \langle a_1, i\rangle\langle a_2, i\rangle \dots \langle a_q, i\rangle \quad [p'_1, p'_2, \dots, p'_q] \\ & \Rightarrow_{G'} & z_1 z_2 \dots z_q \quad [p''_1, p''_2, \dots, p''_q] \end{array}$$

where $p'_j = (a_j \rightarrow \langle a_j, i\rangle, \emptyset)$ and $p''_j = (\langle a_j, i\rangle \rightarrow z_j, F(i))$ such that $p_j = a_j \rightarrow z_j$, $z_j \in V^+$, for all $j = 1, \dots, q$. Then, $z_1 z_2 \dots z_q = x$ and, therefore,

$$S \Rightarrow_{G'}^{2(n+1)} x.$$

If: The converse implication is established by induction on the length of derivations in G'. We prove that

$$S \Rightarrow_{G'}^{n} x' \quad \text{implies} \quad S \Rightarrow_{G}^{*} x$$

for some $x' \in g(x)$, $n \geq 0$.

Basis: For $n = 0$, $S \Rightarrow_{G'}^{0} S$ and $S \Rightarrow_{G}^{0} S$; clearly, $S \in g(S)$.

Induction Hypothesis: Assume that there exists a natural number m such that the claim holds for every $0 \leq n \leq m$.

Induction Step: Let

$$S \Rightarrow_{G'}^{m+1} x'.$$

Express this derivation as

$$S \Rightarrow_{G'}^{m} y' \Rightarrow_{G'} x' \ [p'_1, p'_2, \ldots, p'_q],$$

where $y' \in (V')^+$, $q = |y'|$, and $p'_1, p'_2, \ldots, p'_q$ is a sequence of productions from P'. By the induction hypothesis,

$$S \Rightarrow_{G}^{*} y,$$

where $y \in V^+$, $y' \in g(y)$. Claim 23 says that there exist the following two cases:

(i) Let $m = 2k$ for some $k \geq 0$. Then, $y' \in V^+$, $x' \in W^+$, and every production

$$p'_j = (a_j \rightarrow \langle a_j, i\rangle, \emptyset),$$

where $a_j \in V$, $\langle a_j, i\rangle \in W$, $i \in \{1, \ldots, t\}$. In this case, $\langle a_j, i\rangle \in g(a_j)$ for every a_j and any i (see the definition of g); hence, $x' \in g(y)$ as well.

(ii) Let $m = 2k + 1$. Then, $y' \in W^+$, $x' \in V^+$, and each p'_j is of the form

$$p'_j = (\langle a_j, i\rangle \rightarrow z_j, F(i)),$$

where $\langle a_j, i\rangle \in W$, $z_j \in V^+$. Moreover, according to the forbidding conditions of p'_j, all $\langle a_j, i\rangle$ in y' have the same i. Thus, $y' = \langle a_1, i\rangle\langle a_2, i\rangle \ldots \langle a_q, i\rangle$ for some $i \in \{1, \ldots, t\}$, $y = g^{-1}(y') = a_1 a_2 \ldots a_q$, and $x' = z_1 z_2 \ldots z_q$. By the definition of P',

$$(\langle a_j, i\rangle \rightarrow z_j, F(i)) \in P' \quad \text{implies} \quad a_j \rightarrow z_j \in P_i.$$

Therefore,

$$S \Rightarrow_{G}^{*} a_1 a_2 \ldots a_q \Rightarrow_{G} z_1 z_2 \ldots z_q \ [p_1, p_2, \ldots, p_q],$$

where $p_j = a_j \rightarrow z_j \in P_i$ such that $p'_j = (\langle a_j, i\rangle \rightarrow z_j, F(i))$. Obviously, $z_1 z_2 \ldots z_q = x = x'$.

This completes the induction and establishes Claim 24. □

By Claim 24, for any $x \in T^+$,

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

Therefore, $L(G) = L(G')$, so the lemma holds. □

In order to simplify the notation in the following lemma, for a set of productions

$$P \subseteq \{(a \to z, F): \ a \in V, \ z \in V^*, \ F \subseteq V\},$$

define

$$\mathrm{left}(P) = \{a: \ (a \to z, F) \in P\}.$$

Informally, left(P) denotes the set of left-hand sides of all productions in P.

Lemma 14. FEPT0L$(1) \subseteq$ **EPT0L**.

Proof. Let

$$G = (V, T, P_1, \ldots, P_t, S)$$

be an FEPT0L grammar of degree 1, $t \geq 1$. Let Q be the set of all subsets $O \subseteq P_i$, $1 \leq i \leq t$, such that every $(a \to z, F) \in O$, $a \in V$, $z \in V^+$, $F \subseteq V$, satisfies $F \cap \mathrm{left}(O) = \emptyset$. Create a new set, Q', so that for each $O \in Q$, add

$$\{a \to z: \ (a \to z, F) \in O\}$$

to Q'. Express

$$Q' = \{Q'_1, \ldots, Q'_m\},$$

where m is the cardinality of Q'. Then, construct the EPT0L grammar

$$G' = (V, T, Q'_1, \ldots, Q'_m, S).$$

Basic Idea. To see the basic idea behind the construction of G', consider a pair of productions $p_1 = (a_1 \to z_1, F_1)$ and $p_2 = (a_2 \to z_2, F_2)$ from P_i, for some $i \in \{1, \ldots, t\}$. During a single derivation step, p_1 and p_2 can concurrently rewrite a_1 and a_2 provided that $a_2 \notin F_1$ and $a_1 \notin F_2$, respectively. Consider any $O \subseteq P_i$ containing no pair of productions $(a_1 \to z_1, F_1)$ and $(a_2 \to z_2, F_2)$ such that $a_1 \in F_2$ or $a_2 \in F_1$. Observe that for any derivation step based on O, no production from O is blocked by its forbidding conditions; thus, the conditions can be omitted. Formal proof is given next.

Claim 25. *$S \Rightarrow_G^n x$ if and only if $S \Rightarrow_{G'}^n x$, $x \in V^*$, $n \geq 0$.*

Proof. The claim is proven by induction on the length of derivations.

Only If: By induction on n, $n \geq 0$, we prove that

$$S \Rightarrow_G^n x \quad \text{implies} \quad S \Rightarrow_{G'}^n x$$

for all $x \in V^*$.

Basis: Let $n = 0$. Then, $S \Rightarrow_G^0 S$ and $S \Rightarrow_{G'}^0 S$.

Induction Hypothesis: Suppose that the claim holds for all derivations of length n or less, for some $n \geq 0$.

Induction Step: Consider a derivation $S \Rightarrow_G^{n+1} x$. Because $n+1 \geq 1$, there exists $y \in V^+$, $q = |y|$, and a sequence $p_1, \ldots, p_q$, where $p_j \in P_i$ for all $j = 1, \ldots, q$ and some $i \in \{1, \ldots, t\}$, such that

$$S \Rightarrow_G^n y \Rightarrow_G x \ [p_1, \ldots, p_q].$$

By the induction hypothesis, $S \Rightarrow^n_{G'} y$. Let $O = \{p_j : 1 \leq j \leq q\}$. Observe that $y \Rightarrow_G x\ [p_1, \ldots, p_q]$ implies $\text{alph}(y) = \text{left}(O)$. Moreover, every $p_j = (a \to z, F) \in O$, $a \in V$, $z \in V^+$, $F \subseteq V$, statisfies $F \cap \text{alph}(y) = \emptyset$. Hence, $(a \to z, F) \in O$ implies $F \cap \text{left}(O) = \emptyset$. Inspect the definition of G' to see that there exists

$$Q'_r = \{a \to z :\ (a \to z, F) \in O\}$$

for some r, $1 \leq r \leq m$. Therefore,

$$S \Rightarrow^n_{G'} y \Rightarrow_{G'} x\ [p'_1, \ldots, p'_q],$$

where $p'_j = a \to z \in Q'_r$ such that $p_j = (a \to z, F) \in O$, for all $j = 1, \ldots, q$.

If: The if-part demonstrates for every $n \geq 0$,

$$S \Rightarrow^n_{G'} x \ \text{ implies } \ S \Rightarrow^n_G x,$$

where $x \in V^*$.

Basis: Suppose that $n = 0$. Then, $S \Rightarrow^0_{G'} S$ and $S \Rightarrow^0_G S$.

Induction Hypothesis: Assume that the claim holds for all derivations of length n or less, for some $n \geq 0$.

Induction Step: Let

$$S \Rightarrow^{n+1}_{G'} x.$$

As $n + 1 \geq 1$, there exists a derivation

$$S \Rightarrow^n_{G'} y \Rightarrow_{G'} x\ [p'_1, \ldots, p'_q]$$

such that $y \in V^+$, $q = |y|$, each $p'_i \in Q'_r$ for some $r \in \{1, \ldots, m\}$, and, by the induction hypothesis,

$$S \Rightarrow^n_G y.$$

Then, by the definition of Q'_r, there exists P_i and $O \subseteq P_i$ such that every $(a \to z, F) \in O$, $a \in V$, $z \in V^+$, $F \subseteq V$, statisfies $a \to z \in Q'_r$ and $F \cap \text{left}(O) = \emptyset$. Since $\text{alph}(y) \subseteq \text{left}(O)$, $(a \to z, F) \in O$ implies $F \cap \text{alph}(y) = \emptyset$. Hence,

$$S \Rightarrow^n_G y \Rightarrow_G x\ [p_1, \ldots, p_q],$$

where $p_j = (a \to z, F) \in O$ for all $j = 1, \ldots, q$. □

From the above claim,

$$S \Rightarrow^*_G x \quad \text{if and only if} \quad S \Rightarrow^*_{G'} x$$

for all $x \in T^*$. Consequently, $L(G) = L(G')$. □

The following two lemmas can be proven by analogy with Lemmas 13 and 14. The details are left to the reader.

Lemma 15. ET0L $\subseteq$ **FE0L**(1).

Lemma 16. $\mathbf{FET0L}(1) \subseteq \mathbf{ET0L}$.

Theorem 37. $\mathbf{FEP0L}(1) = \mathbf{FEPT0L}(1) = \mathbf{FE0L}(1) = \mathbf{FET0L}(1) = \mathbf{EPT0L} = \mathbf{ET0L}$.

Proof. By Lemmas 13 and 14, $\mathbf{EPT0L} \subseteq \mathbf{FEP0L}(1)$ and $\mathbf{FEPT0L}(1) \subseteq \mathbf{EPT0L}$, respectively. Since $\mathbf{FEP0L}(1) \subseteq \mathbf{FEPT0L}(1)$, $\mathbf{FEP0L}(1) = \mathbf{FEPT0L}(1) = \mathbf{EPT0L}$. Analogously, from Lemmas 15 and 16, $\mathbf{FE0L}(1) = \mathbf{FET0L}(1) = \mathbf{ET0L}$. However, $\mathbf{EPT0L} = \mathbf{ET0L}$ (see Theorem V.1.6 on page 239 in [156]). Therefore,

$$\mathbf{FEP0L}(1) = \mathbf{FEPT0L}(1) = \mathbf{FE0L}(1) = \mathbf{FET0L}(1) = \mathbf{EPT0L} = \mathbf{ET0L};$$

thus, the theorem holds. □

Next, we investigate the generative power of FEPT0L grammars of degree 2. The following lemma establishes a normal form for context-sensitive grammars so that the grammars satisfying this form generate only sentential forms containing no nonterminal from N_{CS} as the leftmost symbol of the string. We make use of this normal form in Lemma 18.

Lemma 17. *Every context-sensitive language, $L \in \mathbf{CS}$, can be generated by a context-sensitive grammar, $G = (N_1 \cup N_{CF} \cup N_{CS} \cup T, T, P, S_1)$, where N_1, N_{CF}, N_{CS}, and T are pairwise disjoint alphabets, $S_1 \in N_1$, and every production in P has one of the following forms:*

(i) $AB \to AC$, *where* $A \in (N_1 \cup N_{CF})$, $B \in N_{CS}$, $C \in N_{CF}$;

(ii) $A \to B$, *where* $A \in N_{CF}$, $B \in N_{CS}$;

(iii) $A \to a$, *where* $A \in (N_1 \cup N_{CF})$, $a \in T$;

(iv) $A \to C$, *where* $A, C \in N_{CF}$;

(v) $A_1 \to C_1$, *where* $A_1, C_1 \in N_1$;

(vi) $A \to DE$, *where* $A, D, E \in N_{CF}$;

(vii) $A_1 \to D_1E$, *where* $A_1, D_1 \in N_1$, $E \in N_{CF}$.

Proof. Let

$$G' = (N_{CF} \cup N_{CS} \cup T, T, P', S)$$

be a context-sensitive grammar of the form defined in Lemma 4. From this grammar, we construct a grammar

$$G = (N_1 \cup N_{CF} \cup N_{CS} \cup T, T, P, S_1),$$

where

$$\begin{aligned} N_1 &= \{X_1 : X \in N_{CF}\}, \\ P &= P' \cup \{A_1B \to A_1C : AB \to AC \in P', A, C \in N_{CF}, B \in N_{CS}, A_1 \in N_1\} \\ &\quad \cup \{A_1 \to a : A \to a \in P', A \in N_{CF}, A_1 \in N_1, a \in T\} \\ &\quad \cup \{A_1 \to C_1 : A \to C \in P', A, C \in N_{CF}, A_1, C_1 \in N_1\} \\ &\quad \cup \{A_1 \to D_1E : A \to DE \in P', A, D, E \in N_{CF}, A_1, D_1 \in N_1\}. \end{aligned}$$

Basic Idea. G works by analogy with G' except that in G' every sentential form starts with a symbol from $N_1 \cup T$ followed by symbols that are not in N_1. Notice, however, that by $AB \to AC$, G' can never rewrite the leftmost symbol of any sentential form. Based on these observations, it is rather easy to see that $L(G) = L(G')$; a formal proof of this identity is left to the reader. As G is of the required form, Lemma 17 holds. $\square$

Lemma 18. CS $\subseteq$ FEP0L(2).

Proof. Let L be a context-sensitive language generated by a grammar

$$G = (N_1 \cup N_{CF} \cup N_{CS} \cup T, T, P, S_1)$$

of the form of Lemma 17. Let

$$\begin{aligned} V &= N_1 \cup N_{CF} \cup N_{CS} \cup T, \\ P_{CS} &= \{AB \to AC : AB \to AC \in P, A \in (N_1 \cup N_{CF}), B \in N_{CS}, C \in N_{CF}\}, \\ P_{CF} &= P - P_{CS}. \end{aligned}$$

Informally, P_{CS} and P_{CF} are the sets of context-sensitive and context-free productions in P, respectively, and V denotes the total alphabet of G.

Let f be an arbitrary bijection from V to $\{1, \ldots, m\}$, where m is the cardinality of V, and let f^{-1} be the inverse of f.

Construct an FEP0L grammar of degree 2,

$$G' = (V', T, P', S_1),$$

with V' defined as

$$\begin{aligned} W_0 &= \{\langle A, B, C\rangle : AB \to AC \in P_{CS}\}, \\ W_S &= \{\langle A, B, C, j\rangle : AB \to AC \in P_{CS}, 1 \leq j \leq m+1\}, \\ W &= W_0 \cup W_S, \\ V' &= V \cup W. \end{aligned}$$

where V, W_0, and W_S are pairwise disjoint alphabets. The set of productions P' is defined as follows:

1. for every $X \in V$, add $(X \to X, \emptyset)$ to P';
2. for every $A \to u \in P_{CF}$, add $(A \to u, W)$ to P';
3. for every $AB \to AC \in P_{CS}$, add the following productions to P':
   (a) $(B \to \langle A, B, C\rangle, W)$;
   (b) $(\langle A, B, C\rangle \to \langle A, B, C, 1\rangle, W - \{\langle A, B, C\rangle\})$;
   (c) $(\langle A, B, C, j\rangle \to \langle A, B, C, j+1\rangle, \{f^{-1}(j)\langle A, B, C, j\rangle\})$ for all $1 \leq j \leq m$ such that $f(A) \neq j$;
   (d) $(\langle A, B, C, f(A)\rangle \to \langle A, B, C, f(A)+1\rangle, \emptyset)$;
   (e) $(\langle A, B, C, m+1\rangle \to C, \{\langle A, B, C, m+1\rangle^2\})$.

Basic Idea. Let us informally explain how G' simulates the non-context-free productions of the form $AB \to AC$ (see productions of (3) in the construction of P'). First, chosen occurences of B are rewritten with $\langle A, B, C\rangle$ by $(B \to \langle A, B, C\rangle, W)$. The forbidding condition of this production guarantees that there is no simulation already in process. After that, left neighbors of all occurences of $\langle A, B, C\rangle$ are checked not to be any symbols from $V - \{A\}$. In more detail, G' rewrites $\langle A, B, C\rangle$ with $\langle A, B, C, i\rangle$ for $i = 1$. Then, in every $\langle A, B, C, i\rangle$, G' increments i by one as long as i is less or equal to the cardinality of V; simultaneously, it verifies that the left neighbor of every $\langle A, B, C, i\rangle$ differs from the symbol that f maps to i except for the case when $f(A) = i$. Finally, G' checks that there are no two adjoining symbols $\langle A, B, C, m + 1\rangle$. At this point, the left neighbors of $\langle A, B, C, m + 1\rangle$ are necessarily equal to A, so every occurence of $\langle A, B, C, m + 1\rangle$ is rewritten to C.

Observe that the other symbols remain unchanged during the simulation. Indeed, by the forbidding conditions, the only productions that can rewrite symbols $X \notin W$ are of the form $(X \to X, \emptyset)$. Moreover, the forbidding condition of $(\langle A, B, C\rangle \to \langle A, B, C, 1\rangle, W - \{\langle A, B, C\rangle\})$ implies that it is not possible to simulate two different non-context-free productions at the same time.

To establish the identity of languages generated by G and G', we first prove Claims 26 through 30.

Claim 26. *$S_1 \Rightarrow^n_{G'} x'$ implies* $\text{first}(x') \in (N_1 \cup T)$ *for every $n \geq 0$, $x' \in (V')^*$.*

Proof. The claim is proven by induction on n.

Basis: Let $n = 0$. Then, $S_1 \Rightarrow^0_{G'} S_1$ and $S_1 \in N_1$.

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

$$S_1 \Rightarrow^{n+1}_{G'} x',$$

where $x' \in (V')^*$. Because $n + 1 \geq 1$, there is a derivation

$$S_1 \Rightarrow^n_{G'} y' \Rightarrow_{G'} x' \; [p_1, \ldots, p_q],$$

$y' \in (V')^*$, $q = |y'|$, and, by the induction hypothesis, $\text{first}(y') \in (N_1 \cup T)$. Inspect P' to see that the production p_1 that rewrites the leftmost symbol of y' is one of the following forms: $(A_1 \to A_1, \emptyset)$, $(a \to a, \emptyset)$, $(A_1 \to a, W)$, $(A_1 \to C_1, W)$, or $(A_1 \to D_1E, W)$, where $A_1, C_1, D_1 \in N_1$, $a \in T$, $E \in N_{CF}$ (see (1) and (2) in the definition of P' and Lemma 17). It is obvious that the leftmost symbols of the right-hand sides of these productions belong to $(N_1 \cup T)$. Hence,

$$\text{first}(x') \in (N_1 \cup T),$$

so the claim holds. □

Claim 27. *$S_1 \Rightarrow^n_{G'} y'_1 X y'_3$, $X \in W_S$, implies $y'_1 \in (V')^+$ for any $y'_3 \in (V')^*$.*

Proof. Informally, the claim says that every occurence of a symbol from W_S has always a left neighbor. Clearly, this claim follows from the statement of Claim 26. Since $W_S \cap (N_1 \cup T) = \emptyset$, X cannot be the leftmost symbol in a sentential form and the claim holds. □

Claim 28. *$S_1 \Rightarrow^n_{G'} x'$, $n \geq 0$, implies that x' has one of the following three forms:*

(I) $x' \in V^*$;

(II) $x' \in (V \cup W_0)^*$ *and* $\#_{W_0} x' > 0$;

(III) $x' \in (V \cup \{\langle A, B, C, j\rangle\})^*$, $\#_{\{\langle A,B,C,j\rangle\}} x' > 0$, *and* $\{f^{-1}(k)\langle A, B, C, j\rangle : 1 \leq k < j, k \neq f(A)\} \cap \mathrm{sub}(x') = \emptyset$, *where* $\langle A, B, C, j\rangle \in W_S$, $A \in (N_1 \cup N_{CF})$, $B \in N_{CS}$, $C \in N_{CF}$, $1 \leq j \leq m + 1$.

Proof. We prove the claim by the induction on $n \geq 0$.

Basis: Let $n = 0$. Clearly, $S_1 \Rightarrow^0_{G'} S_1$ and S_1 is of type (I).

Induction Hypothesis: Suppose that the claim holds for all derivations of length n or less, for some $n \geq 0$.

Induction Step: Let us consider any derivation of the form

$$S_1 \Rightarrow^{n+1}_{G'} x'.$$

Because $n+1 \geq 1$, there exists $y' \in (V')^*$ and a sequence of productions $p_1, \ldots, p_q$, where $p_i \in P'$, $1 \leq i \leq q$, $q = |y'|$, such that

$$S_1 \Rightarrow^n_{G'} y' \Rightarrow_{G'} x' \; [p_1, \ldots, p_q].$$

Let $y' = a_1 a_2 \ldots a_q$, $a_i \in V'$.

By the induction hypothesis, y' can only be of forms (I) through (III). Thus, the following three cases cover all possible forms of y':

(i) Let $y' \in V^*$ (form (I)). In this case, every production p_i can be either of the form $(a_i \to a_i, \emptyset)$, $a_i \in V$, or $(a_i \to u, W)$ such that $a_i \to u \in P_{CF}$, or $(a_i \to \langle A, a_i, C\rangle, W)$, $a_i \in N_{CS}$, $\langle A, a_i, C\rangle \in W_0$ (see the definition of P').

Suppose that for every $i \in \{1, \ldots, q\}$, p_i has one of the first two listed forms. According to the right-hand sides of these productions, we obtain $x' \in V^*$; that is, x' is of form (I).

If there exists i such that $p_i = (a_i \to \langle A, a_i, C\rangle, W)$ for some $A \in (N_1 \cup N_{CF})$, $a_i \in N_{CS}$, $C \in N_{CF}$, $\langle A, a_i, C\rangle \in W_0$, we get $x' \in (V \cup W_0)^*$ with $\#_{W_0} x' > 0$. Thus, x' belongs to (II).

(ii) Let $y' \in (V \cup W_0)^*$ and $\#_{W_0} y' > 0$ (form (II)). At this point, p_i is either $(a_i \to a_i, \emptyset)$ (rewriting $a_i \in V$ to itself) or $(\langle A, B, C\rangle \to \langle A, B, C, 1\rangle, W - \{\langle A, B, C\rangle\})$ rewriting $a_i = \langle A, B, C\rangle \in W_0$ to $\langle A, B, C, 1\rangle \in W_S$, where $A \in (N_1 \cup N_{CF})$, $B \in N_{CS}$, $C \in N_{CF}$. Since $\#_{W_0} y' > 0$, there exists at least one i such that $a_i = \langle A, B, C\rangle \in W_0$. The corresponding production p_i can be used provided that $\#_{(W - \{\langle A,B,C\rangle\})} y' = 0$. Therefore, $y' \in (V \cup \{\langle A, B, C\rangle\})^*$ and hence $x' \in (V \cup \{\langle A, B, C, 1\rangle\})^*$, $\#_{\{\langle A,B,C,1\rangle\}} x' > 0$; that is, x' is of type (III).

(iii) Assume that $y' \in (V \cup \{\langle A, B, C, j\rangle\})^*$, $\#_{\{\langle A,B,C,j\rangle\}} y' > 0$, and

$$\mathrm{sub}(y') \cap \{f^{-1}(k)\langle A, B, C, j\rangle : \; 1 \leq k < j, k \neq f(A)\} = \emptyset,$$

where $\langle A,B,C,j\rangle \in W_S$, $A \in (N_1 \cup N_{CF})$, $B \in N_{CS}$, $C \in N_{CF}$, $1 \le j \le m+1$ (form (III)). By inspection of P', we see that the following four forms of productions can be used to rewrite y' to x':

(a) $(a_i \to a_i, \emptyset)$, $a_i \in V$;
(b) $(\langle A,B,C,j\rangle \to \langle A,B,C,j+1\rangle, \{f^{-1}(j)\langle A,B,C,j\rangle\})$, $1 \le j \le m$, $j \ne f(A)$;
(c) $(\langle A,B,C,f(A)\rangle \to \langle A,B,C,f(A)+1\rangle, \emptyset)$;
(d) $(\langle A,B,C,m+1\rangle \to C, \{\langle A,B,C,m+1\rangle^2\})$.

Let $1 \le j \le m$, $j \ne f(A)$. Then, symbols from V are rewritten to themselves (case (a)) and every occurence of $\langle A,B,C,j\rangle$ is rewritten to $\langle A,B,C,j+1\rangle$ by (b). Clearly, we obtain $x' \in (V \cup \{\langle A,B,C,j+1\rangle\})^*$ such that $\#_{\{\langle A,B,C,j+1\rangle\}} x' > 0$. Furthermore, (b) can be used only when $f^{-1}(j)\langle A,B,C,j\rangle \notin \mathrm{sub}(y')$. As

$$\mathrm{sub}(y') \cap \{f^{-1}(k)\langle A,B,C,j\rangle :\ 1 \le k < j,\ k \ne f(A)\} = \emptyset,$$

it holds that

$$\mathrm{sub}(y') \cap \{f^{-1}(k)\langle A,B,C,j\rangle :\ 1 \le k \le j,\ k \ne f(A)\} = \emptyset.$$

Since every occurence of $\langle A,B,C,j\rangle$ is rewritten to $\langle A,B,C,j+1\rangle$ and other symbols are unchanged,

$$\mathrm{sub}(x') \cap \{f^{-1}(k)\langle A,B,C,j+1\rangle :\ 1 \le k < j+1,\ k \ne f(A)\} = \emptyset;$$

therefore, x' is of form (III).

Assume that $j = f(A)$. Then, all occurences of $\langle A,B,C,j\rangle$ are rewritten to $\langle A,B,C,j+1\rangle$ by (c) and symbols from V are rewritten to themselves. As before, we obtain $x' \in (V \cup \{\langle A,B,C,j+1\rangle\})^*$ and $\#_{\{\langle A,B,C,j+1\rangle\}} x' > 0$. Moreover, because

$$\mathrm{sub}(y') \cap \{f^{-1}(k)\langle A,B,C,j\rangle :\ 1 \le k < j,\ k \ne f(A)\} = \emptyset$$

and j is just $f(A)$,

$$\mathrm{sub}(x') \cap \{f^{-1}(k)\langle A,B,C,j+1\rangle :\ 1 \le k < j+1,\ k \ne f(A)\} = \emptyset$$

and x' belongs to (III) as well.

Finally, let $j = m+1$. Then, every occurence of $\langle A,B,C,j\rangle$ is rewritten to C (case (d)) and, therefore, $x' \in V^*$; that is, x' has form (I).

In (i), (ii), and (iii), we have considered all derivations that rewrite y' to x', and in each of these cases, we have shown that x' has one of the requested forms. Therefore, Claim 28 holds. □

To prove the following claims, we need a finite letter-to-letters substitution g from V^* into $(V')^*$ defined as

$$\begin{aligned} g(X) = {} & \{X\} \cup \{\langle A,X,C\rangle :\ \langle A,X,C\rangle \in W_0\} \\ & \cup \{\langle A,X,C,j\rangle :\ \langle A,X,C,j\rangle \in W_S, 1 \le j \le m+1\} \end{aligned}$$

for all $X \in V$, $A \in (N_1 \cup N_{CF})$, $C \in N_{CF}$. Let g^{-1} be the inverse of g.

Claim 29. *Let* $y' = a_1 a_2 \dots a_q$*,* $a_i \in V'$*,* $q = |y'|$*, and* $g^{-1}(a_i) \Rightarrow_G^{h_i} g^{-1}(u_i)$ *for all* $i \in \{1, \dots, q\}$ *and some* $h_i \in \{0, 1\}$*,* $u_i \in (V')^+$*. Then,* $g^{-1}(y') \Rightarrow_G^r g^{-1}(x')$ *such that* $x' = u_1 u_2 \dots u_q$*,* $r = \sum_{i=1}^{q} h_i$*,* $r \le q$*.*

Proof. First, consider a derivation $g^{-1}(X) \Rightarrow_G^h g^{-1}(u)$, $X \in V'$, $u \in (V')^+$, $h \in \{0, 1\}$. If $h = 0$ then $g^{-1}(X) = g^{-1}(u)$. Let $h = 1$. Then, there surely exists a production $p = g^{-1}(X) \to g^{-1}(u) \in P$ such that $g^{-1}(X) \Rightarrow_G g^{-1}(u)$ $[p]$.

Return to the statement of this claim. We can construct a derivation

$$\begin{aligned} g^{-1}(a_1)g^{-1}(a_2)\dots g^{-1}(a_q) &\Rightarrow_G^{h_1} g^{-1}(u_1)g^{-1}(a_2)\dots g^{-1}(a_q) \\ &\Rightarrow_G^{h_2} g^{-1}(u_1)g^{-1}(u_2)\dots g^{-1}(a_q) \\ &\ \vdots \\ &\Rightarrow_G^{h_q} g^{-1}(u_1)g^{-1}(u_2)\dots g^{-1}(u_q) \end{aligned}$$

where $g^{-1}(y') = g^{-1}(a_1)\dots g^{-1}(a_q)$ and $g^{-1}(u_1)\dots g^{-1}(u_q) = g^{-1}(u_1 \dots u_q) = g^{-1}(x')$. In such a derivation, each $g^{-1}(a_i)$ is either left unchanged (if $h_i = 0$) or rewritten to $g^{-1}(u_i)$ by the corresponding production $g^{-1}(a_i) \to g^{-1}(u_i)$. Obviously, the length of this derivation is $\sum_{i=1}^{q} h_i$. □

Claim 30. $S_1 \Rightarrow_G^* x$ *if and only if* $S_1 \Rightarrow_{G'}^* x'$*, where* $x \in V^*$*,* $x' \in (V')^*$*,* $x' \in g(x)$*.*

Proof.

Only if: The only-if part is established by induction on the length of derivations in G. That is, we show that

$$S_1 \Rightarrow_G^n x \quad \text{implies} \quad S_1 \Rightarrow_{G'}^* x$$

where $x \in V^*$, for $n \ge 0$.

Basis: Let $n = 0$. Then, $S_1 \Rightarrow_G^0 S_1$ and $S_1 \Rightarrow_{G'}^0 S_1$ as well.

Induction Hypothesis: Assume that the claim holds for all derivations of length n or less, for some $n \ge 0$.

Induction Step: Consider a derivation

$$S_1 \Rightarrow_G^{n+1} x.$$

Because $n + 1 > 0$, there exists $y \in V^*$ and $p \in P$ such that

$$S_1 \Rightarrow_G^n y \Rightarrow_G x \ [p],$$

and, by the induction hypothesis, there is also a derivation

$$S_1 \Rightarrow_{G'}^* y.$$

Let $y = a_1 a_2 \dots a_q$, $a_i \in V$, $1 \le i \le q$, $q = |y|$. The following cases (i) and (ii) cover all possible forms of p.

(i) $p = A \rightarrow u \in P_{CF}$, $A \in (N_1 \cup N_{CF})$, $u \in V^*$. Then, $y = y_1 A y_3$ and $x = y_1 u y_3$, $y_1, y_3 \in V^*$. Let $s = |y_1| + 1$. Since we have $(A \rightarrow u, W) \in P'$, we can construct a derivation

$$S_1 \Rightarrow^*_{G'} y \Rightarrow_{G'} x\ [p_1, \ldots, p_q]$$

such that $p_s = (A \rightarrow u, W)$ and $p_i = (a_i \rightarrow a_i, \emptyset)$ for all $i \in \{1, \ldots, q\}$, $i \neq s$.

(ii) $p = AB \rightarrow AC \in P_{CS}$, $A \in (N_1 \cup N_{CF})$, $B \in N_{CS}$, $C \in N_{CF}$. Then, $y = y_1 A B y_3$ and $x = y_1 A C y_3$, $y_1, y_3 \in V^*$. Let $s = |y_1| + 2$. In this case, there is the following derivation:

$$\begin{array}{lll}
S_1 & \Rightarrow^*_{G'} y_1 A B y_3 & \\
 & \Rightarrow_{G'} y_1 A \langle A, B, C\rangle y_3 & [p_s = (B \rightarrow \langle A, B, C\rangle, W)] \\
 & \Rightarrow_{G'} y_1 A \langle A, B, C, 1\rangle y_3 & [p_s = (\langle A, B, C\rangle \rightarrow \langle A, B, C, 1\rangle, \\
 & & W - \{\langle A, B, C\rangle\})] \\
 & \Rightarrow_{G'} y_1 A \langle A, B, C, 2\rangle y_3 & [p_s = (\langle A, B, C, 1\rangle \rightarrow \langle A, B, C, 2\rangle, \\
 & & \{f^{-1}(1)\langle A, B, C, j\rangle\})] \\
 & \vdots & \\
 & \Rightarrow_{G'} y_1 A \langle A, B, C, f(A)\rangle y_3 & [p_s = (\langle A, B, C, f(A) - 1\rangle \rightarrow \langle A, B, C, f(A)\rangle, \\
 & & \{f^{-1}(f(A) - 1)\langle A, B, C, f(A) - 1\rangle\})] \\
 & \Rightarrow_{G'} y_1 A \langle A, B, C, f(A) + 1\rangle y_3 & [p_s = (\langle A, B, C, f(A)\rangle \rightarrow \langle A, B, C, f(A) + 1\rangle, \emptyset)] \\
 & \Rightarrow_{G'} y_1 A \langle A, B, C, f(A) + 2\rangle y_3 & [p_s = (\langle A, B, C, f(A) + 1\rangle \rightarrow \langle A, B, C, f(A) + 2\rangle, \\
 & & \{f^{-1}(f(A) + 1)\langle A, B, C, f(A) + 1\rangle\})] \\
 & \vdots & \\
 & \Rightarrow_{G'} y_1 A \langle A, B, C, m + 1\rangle y_3 & [p_s = (\langle A, B, C, m\rangle \rightarrow \langle A, B, C, m + 1\rangle, \\
 & & \{f^{-1}(m)\langle A, B, C, m\rangle\})] \\
 & \Rightarrow_{G'} y_1 A C y_3 & [p_s = (\langle A, B, C, m + 1\rangle \rightarrow C, \\
 & & \{\langle A, B, C, m + 1\rangle^2\})]
\end{array}$$

such that $p_i = (a_i \rightarrow a_i, \emptyset)$ for all $i \in \{1, \ldots, q\}$, $i \neq s$.

If: By induction on n, we prove that

$$S_1 \Rightarrow^n_{G'} x' \quad \text{implies} \quad S_1 \Rightarrow^*_G x,$$

where $x' \in (V')^*$, $x \in V^*$ and $x' \in g(x)$.

Basis: Let $n = 0$. The only x' is S_1 because $S_1 \Rightarrow^0_{G'} S_1$. Obviously, $S_1 \Rightarrow^0_G S_1$ and $S_1 \in g(S_1)$.

Induction Hypothesis: Suppose that the claim holds for any derivation of length n or less, for some $n \geq 0$.

Induction Step: Consider a derivation of the form

$$S_1 \Rightarrow^{n+1}_{G'} x'.$$

Since $n + 1 \geq 1$, there exists $y' \in (V')^*$ and a sequence of productions $p_1, \ldots, p_q$ from P', $q = |x'|$, such that

$$S_1 \Rightarrow^n_{G'} y' \Rightarrow_{G'} x'\ [p_1, \ldots, p_q].$$

Let $y' = a_1 a_2 \dots a_q$, $a_i \in V'$, $1 \leq i \leq q$. By the induction hypothesis, we have

$$S_1 \Rightarrow_G^* y,$$

where $y \in V^*$, such that $y' \in g(y)$.

From Claim 28, y' can have one of the following forms:

(i) Let $y' \in (V')^*$ (see (I) in Claim 28). Inspect P' to see that there are three forms of productions rewriting symbols a_i in y':

(a) $p_i = (a_i \to a_i, \emptyset) \in P'$, $a_i \in V$. In this case,

$$g^{-1}(a_i) \Rightarrow_G^0 g^{-1}(a_i).$$

(b) $p_i = (a_i \to u_i, W) \in P'$ such that $a_i \to u_i \in P_{CF}$. Because $a_i = g^{-1}(a_i)$, $u_i = g^{-1}(u_i)$ and $a_i \to u_i \in P$,

$$g^{-1}(a_i) \Rightarrow_G g^{-1}(u_i)\ [a_i \to u_i].$$

(c) $p_i = (a_i \to \langle A, a_i, C\rangle, W) \in P'$, $a_i \in N_{CS}$, $A \in (N_1 \cup N_{CF})$, $C \in N_{CF}$. Since $g^{-1}(a_i) = g^{-1}(\langle A, a_i, C\rangle)$, we have

$$g^{-1}(a_i) \Rightarrow_G^0 g^{-1}(\langle A, a_i, C\rangle).$$

We see that for all a_i, there exists a derivation

$$g^{-1}(a_i) \Rightarrow_G^{h_i} g^{-1}(z_i)$$

for some $h_i \in \{0, 1\}$, where $z_i \in (V')^+$, $x' = z_1 z_2 \dots z_q$. Therefore, by Claim 29, we can construct

$$S_1 \Rightarrow_G^* y \Rightarrow_G^r x,$$

where $0 \leq r \leq q$, $x = g^{-1}(x')$.

(ii) Let $y' \in (V \cup W_0)^*$ and $\#_{W_0} y' > 0$ (see (II)). At this point, the following two forms of productions can be used to rewrite a_i in y':

(a) $p_i = (a_i \to a_i, \emptyset) \in P'$, $a_i \in V$. As in case (i.a),

$$g^{-1}(a_i) \Rightarrow_G^0 g^{-1}(a_i).$$

(b) $p_i = (\langle A, B, C\rangle \to \langle A, B, C, 1\rangle, W - \{\langle A, B, C\rangle\})$, $a_i = \langle A, B, C\rangle \in W_0$, $A \in (N_1 \cup N_{CF})$, $B \in N_{CS}$, $C \in N_{CF}$. Because $g^{-1}(\langle A, B, C\rangle) = g^{-1}(\langle A, B, C, 1\rangle)$,

$$g^{-1}(\langle A, B, C\rangle) \Rightarrow_G^0 g^{-1}(\langle A, B, C, 1\rangle).$$

Thus, there exists a derivation

$$S_1 \Rightarrow_G^* y \Rightarrow_G^0 x,$$

where $x = g^{-1}(x')$.

(iii) Let $y' \in (V \cup \{\langle A, B, C, j\rangle\})^*$, $\#_{\{\langle A,B,C,j\rangle\}} y' > 0$, and

$$\mathrm{sub}(y') \cap \{f^{-1}(k)\langle A, B, C, j\rangle : \ 1 \leq k < j, \ k \neq f(A)\} = \emptyset,$$

where $\langle A, B, C, j\rangle \in W_S$, $A \in (N_1 \cup N_{CF})$, $B \in N_{CS}$, $C \in N_{CF}$, $1 \leq j \leq m+1$ (see (III)). By inspection of P', the following four forms of productions can be used to rewrite y' to x':

(a) $p_i = (a_i \to a_i, \emptyset)$, $a_i \in V$;

(b) $p_i = (\langle A, B, C, j\rangle \to \langle A, B, C, j+1\rangle, \{f^{-1}(j)\langle A, B, C, j\rangle\})$, $1 \leq j \leq m$, $j \neq f(A)$;

(c) $p_i = (\langle A, B, C, f(A)\rangle \to \langle A, B, C, f(A)+1\rangle, \emptyset)$;

(d) $p_i = (\langle A, B, C, m+1\rangle \to C, \{\langle A, B, C, m+1\rangle^2\})$.

Let $1 \leq j \leq m$. G' can rewrite such y' using only the productions (a) through (c). Because $g^{-1}(\langle A, B, C, j\rangle) = g^{-1}(\langle A, B, C, j+1\rangle)$ and $g^{-1}(a_i) = g^{-1}(a_i)$, by analogy with (ii), we obtain a derivation

$$S_1 \Rightarrow_G^* y \Rightarrow_G^0 x$$

such that $x = g^{-1}(x')$.

Let $j = m+1$. In this case, only the productions (a) and (d) can be used. Since $\#_{\{\langle A,B,C,j\rangle\}} y' > 0$, there is at least one occurence of $\langle A, B, C, m+1\rangle$ in y' and, by the forbidding condition of the production (c), $\langle A, B, C, m+1\rangle^2 \notin \mathrm{sub}(y')$. Observe that for $j = m+1$,

$$\{f^{-1}(k)\langle A, B, C, m+1\rangle : \ 1 \leq k < j, \ k \neq f(A)\} = \\ \{X\langle A, B, C, m+1\rangle : \ X \in V, \ X \neq A\}$$

and, thus,

$$\mathrm{sub}(y') \cap \{X\langle A, B, C, m+1\rangle : \ X \in V, \ X \neq A\} = \emptyset.$$

According to Claim 27, $\langle A, B, C, m+1\rangle$ has always a left neighbor in y'. As a result, the left neighbor of every occurence of $\langle A, B, C, m+1\rangle$ is A. Therefore, we can express:

$$\begin{aligned} y' &= y_1 A\langle A, B, C, m+1\rangle y_2 A\langle A, B, C, m+1\rangle y_3 \ldots y_r A\langle A, B, C, m+1\rangle y_{r+1}, \\ y &= g^{-1}(y_1) AB g^{-1}(y_2) AB g^{-1}(y_3) \ldots g^{-1}(y_r) AB g^{-1}(y_{r+1}), \\ x' &= y_1 AC y_2 AC y_3 \ldots y_r AC y_{r+1}, \end{aligned}$$

where $r \geq 1$, $y_s \in V^*$, $1 \leq s \leq r+1$. Since we have $p = AB \to AC \in P$, there is a derivation:

$$\begin{aligned} S_1 &\Rightarrow_G^* g^{-1}(y_1) AB g^{-1}(y_2) AB g^{-1}(y_3) \ldots g^{-1}(y_r) AB g^{-1}(y_{r+1}) \\ &\Rightarrow_G g^{-1}(y_1) AC g^{-1}(y_2) AB g^{-1}(y_3) \ldots g^{-1}(y_r) AB g^{-1}(y_{r+1}) \quad [p] \\ &\Rightarrow_G g^{-1}(y_1) AC g^{-1}(y_2) AC g^{-1}(y_3) \ldots g^{-1}(y_r) AB g^{-1}(y_{r+1}) \quad [p] \\ &\vdots \\ &\Rightarrow_G g^{-1}(y_1) AC g^{-1}(y_2) AC g^{-1}(y_3) \ldots g^{-1}(y_r) AC g^{-1}(y_{r+1}) \quad [p] \end{aligned}$$

where $g^{-1}(y_1) AC g^{-1}(y_2) AC g^{-1}(y_3) \ldots g^{-1}(y_r) AC g^{-1}(y_{r+1}) = g^{-1}(x') = x$.

Because cases (i), (ii) and (iii) cover all possible forms of y', we have completed the induction and established Claim 30. $\square$

The equivalence of G and G' follows from Claim 30. Indeed, observe that by the definition of g, we have $g(a) = \{a\}$ for all $a \in T$. Therefore, by Claim 30, we have for any $x \in T^*$:

$$S_1 \Rightarrow_G^* x \quad \text{if and only if} \quad S_1 \Rightarrow_{G'}^* x.$$

Thus, $L(G) = L(G')$ and the lemma holds. $\square$

Theorem 38. $\mathbf{CS} = \mathbf{FEP0L}(2) = \mathbf{FEPT0L}(2) = \mathbf{FEP0L} = \mathbf{FEPT0L}$.

Proof. By Lemma 18, $\mathbf{CS} \subseteq \mathbf{FEP0L}(2) \subseteq \mathbf{FEPT0L}(2) \subseteq \mathbf{FEPT0L}$. From Lemma 11 and the definition of FET0L grammars, it follows that $\mathbf{FEPT0L}(s) \subseteq \mathbf{FEPT0L} \subseteq \mathbf{CEPT0L} \subseteq \mathbf{CS}$ for any $s \geq 0$. Moreover, $\mathbf{FEP0L}(s) \subseteq \mathbf{FEP0L} \subseteq \mathbf{FEPT0L}$. Thus, $\mathbf{CS} = \mathbf{FEP0L}(2) = \mathbf{FEPT0L}(2) = \mathbf{FEP0L} = \mathbf{FEPT0L}$, and the theorem holds. $\square$

Return to the proof of Lemma 18. Observe that the productions of the FEP0L grammar G' are of restricted forms. This observation gives rise to the next corollary.

Corollary 15. *Every context-sensitive language can be generated by an FEP0L grammar $G = (V, T, P, S)$ of degree 2 such that every production from P has one of the following forms:*

- *(i)* $(a \to a, \emptyset)$, $a \in V$;
- *(ii)* $(X \to x, F)$, $X \in V - T$, $|x| \in \{1, 2\}$, $\max(F) = 1$;
- *(iii)* $(X \to Y, \{z\})$, $X, Y \in V - T$, $z \in V^2$.

Next, we demonstrate that the family of recursively enumerable languages is generated by the forbidding E0L grammars of degree 2.

Lemma 19. $\mathbf{RE} \subseteq \mathbf{FE0L}(2)$.

Proof. Let L be a recursively enumerable language generated by a phrase structure grammar

$$G = (V, T, P, S)$$

having the form defined in Lemma 5, where

$$\begin{array}{lcl} V & = & N_{CF} \cup N_{CS} \cup T, \\ P_{CS} & = & \{AB \to AC \in P : \ A, C \in N_{CF}, B \in N_{CS}\}, \\ P_{CF} & = & P - P_{CS}. \end{array}$$

Let \$ be a new symbol and m be the cardinality of $V \cup \{\$\}$. Furthermore, let f be an arbitrary bijection from $V \cup \{\$\}$ onto $\{1, \ldots, m\}$, and let f^{-1} be the inverse of f.

Then, we define an FE0L grammar

$$G' = (V', T, P', S')$$

of degree 2 as follows:

$$\begin{array}{rcl} W_0 & = & \{\langle A,B,C\rangle : AB \to AC \in P\}, \\ W_S & = & \{\langle A,B,C,j\rangle : AB \to AC \in P, 1 \leq j \leq m\}, \\ W & = & W_0 \cup W_S, \\ V' & = & V \cup W \cup \{S', \$\}, \end{array}$$

where $A, C \in N_{CF}, B \in N_{CS}$, and V, W_0, W_S, and $\{S', \$\}$ are pairwise disjoint alphabets. The set of productions P' is defined in the following way:

1. add $(S' \to \$S, \emptyset)$, $(\$ \to \$, \emptyset)$ and $(\$ \to \varepsilon, V' - T - \{\$\})$ to P';
2. for all $X \in V$, add $(X \to X, \emptyset)$ to P';
3. for all $A \to u \in P_{CF}$, $A \in N_{CF}$, $u \in \{\varepsilon\} \cup N_{CS} \cup T \cup (\bigcup_{i=1}^{2} N_{CF}^i)$, add $(A \to u, W)$ to P';
4. if $AB \to AC \in P_{CS}$, $A, C \in N_{CF}, B \in N_{CS}$, then add the following productions to P':
   (a) $(B \to \langle A,B,C\rangle, W)$;
   (b) $(\langle A,B,C\rangle \to \langle A,B,C,1\rangle, W - \{\langle A,B,C\rangle\})$;
   (c) $(\langle A,B,C,j\rangle \to \langle A,B,C,j+1\rangle, \{f^{-1}(j)\langle A,B,C,j\rangle\})$ for all $1 \leq j \leq m$ such that $f(A) \neq j$;
   (d) $(\langle A,B,C,f(A)\rangle \to \langle A,B,C,f(A)+1\rangle, \emptyset)$;
   (e) $(\langle A,B,C,m+1\rangle \to C, \{\langle A,B,C,m+1\rangle^2\})$.

Basic Idea. Let us only sketch the proof that $L(G) = L(G')$. The above construction resembles the construction in Lemma 18 very much. Indeed, to simulate the non-context-free productions $AB \to AC$ in FE0L grammars, we use the same technique as in FEP0L grammars from Lemma 18. We only need to guarantee that no sentential form begins with a symbol from N_{CS}. This is solved by an auxiliary nonterminal \$ in the definition of G'. The symbol is always generated in the first derivation step by $(S' \to \$S, \emptyset)$ (see (1) in the definition of P'). After that, it appears as the leftmost symbol of all sentential forms containing some nonterminals. The only production that can erase it is $(\$ \to \varepsilon, V' - T - \{\$\})$.

Therefore, by analogy with the technique used in Lemma 18, we can establish

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

such that $x \in V^*$, $x' \in (V' - \{S', \$\})^*$, $x' \in g(x)$, where g is a finite substitution from V^* into $(V' - \{S', \$\})^*$ defined as

$$\begin{array}{rcl} g(X) & = & \{X\} \cup \{\langle A,X,C\rangle :\ \langle A,X,C\rangle \in W_0\} \\ & & \cup\ \{\langle A,X,C,j\rangle :\ \langle A,X,C,j\rangle \in W_S, 1 \leq j \leq m+1\} \end{array}$$

for all $X \in V$, $A, C \in N_{CF}$. The details are left to the reader.

As in Lemma 18, we have $g(a) = \{a\}$ for all $a \in T$; hence, for all $x \in T^*$:

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

Since

$$\$x \Rightarrow_{G'} x \ [(\$ \to \varepsilon, V' - T - \{\$\})],$$

we obtain

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

Consequently, $L(G) = L(G')$; thus, $\mathbf{RE} \subseteq \mathbf{FE0L}(2)$. □

Theorem 39. $\mathbf{RE} = \mathbf{FE0L}(2) = \mathbf{FET0L}(2) = \mathbf{FE0L} = \mathbf{FET0L}$.

Proof. By Lemma 19, we have $\mathbf{RE} \subseteq \mathbf{FE0L}(2) \subseteq \mathbf{FET0L}(2) \subseteq \mathbf{FET0L}$. From Lemma 12, it follows that $\mathbf{FET0L}(s) \subseteq \mathbf{FET0L} \subseteq \mathbf{CET0L} \subseteq \mathbf{RE}$, for any $s \geq 0$. Therefore, $\mathbf{RE} = \mathbf{FE0L}(2) = \mathbf{FET0L}(2) = \mathbf{FE0L} = \mathbf{FET0L}$, so the theorem holds. □

By analogy with Corollary 15, we obtain the following normal form.

Corollary 16. *Every recursively enumerable language can be generated by an FE0L grammar $G = (V, T, P, S)$ of degree 2 such that every production from P has one of the following forms:*

(i) $(a \to a, \emptyset)$, $a \in V$;

(ii) $(X \to x, F)$, $X \in V - T$, $|x| \leq 2$, *and* $F \neq \emptyset$ *implies* $\max(F) = 1$;

(iii) $(X \to Y, \{z\})$, $X, Y \in V - T$, $z \in V^2$.

Theorems 36, 37, 38, and 39 imply the following relationships of FET0L language families:

Corollary 17.

$$\begin{gathered}
\mathbf{CF} \\
\subset \\
\mathbf{FEP0L}(0) = \mathbf{FE0L}(0) = \mathbf{EP0L} = \mathbf{E0L} \\
\subset \\
\mathbf{FEP0L}(1) = \mathbf{FEPT0L}(1) = \mathbf{FE0L}(1) = \mathbf{FET0L}(1) = \\
\mathbf{FEPT0L}(0) = \mathbf{FET0L}(0) = \mathbf{EPT0L} = \mathbf{ET0L} \\
\subset \\
\mathbf{FEP0L}(2) = \mathbf{FEPT0L}(2) = \mathbf{FEP0L} = \mathbf{FEPT0L} = \mathbf{CS} \\
\subset \\
\mathbf{FE0L}(2) = \mathbf{FET0L}(2) = \mathbf{FE0L} = \mathbf{FET0L} = \mathbf{RE}.
\end{gathered}$$