LEFT-FORBIDDING COOPERATING DISTRIBUTED GRAMMAR SYSTEMS

Filip Goldefus
Doctoral degree programme (1), FIT BUT
E-mail: igoldefu@fit.vutbr.cz

Supervised by: Alexander Meduna
E-mail: meduna@fit.vutbr.cz

ABSTRACT

In a left-forbidding grammar, a set of nonterminals is attached to every context-free production, and such a production can rewrite a nonterminal if no symbol from the attached set occurs to the left of the rewritten nonterminal in the sentential form. The present paper discusses left-forbidding cooperating distributed grammar systems that work in the t-mode and have left-forbidding grammars as their components. It demonstrates that with two components, these systems generate the family of recursively enumerable languages.

1 INTRODUCTION

The present paper discusses cooperating distributed grammar systems working in the t-mode (see [1] for details). In a left-forbidding grammar, a set of nonterminals is attached to every context-free production. Such a production can rewrite a nonterminal provided that no symbol from its attached set occurs to the left of the rewritten nonterminal in the sentential form.

We study the generative power of left-forbidding cooperating distributed grammar systems, whose components are left-forbidding grammars. In what follows with two components, they generate the family of recursively enumerable languages. This main result of the present paper is of some interest because two-component cooperating distributed grammar systems whose components are ordinary context-free grammars generate only the family of context-free languages [1].

2 PRELIMINARIES AND DEFINITIONS

This paper assumes that the reader is familiar with formal language theory (see [4]). For an alphabet V, V^* represents the free monoid generated by V. The unit of V^* is denoted by ε. Set $V^+ = V^* - \{\varepsilon\}$. For $w \in V^*$, $|w|$ denotes the length of w, $alph(w)$ denotes the set of letters occurring in w, and for all $i = 1, \ldots, |w|$, w_i denotes the ith symbol of w. Denote the family of recursively enumerable languages by $\mathscr{L}_{RE}$.

A *state grammar* (see [3]) is a sextuple $G = (N, T, Q, P, S, q_0)$, where N is a nonterminal alphabet, T is a terminal alphabet, $V = N \cup T$, Q is a finite set of states, N, T, Q are pairwise disjoint, $S \in N$ is the start symbol, $q_0 \in Q$ is the start state, and P is a finite set of productions of the form $(A,p) \rightarrow (x,q)$, where $p,q \in Q$, $A \in N$, and $x \in V^*$. Let $lhs((A,p) \rightarrow (x,q))$ and $rhs((A,p) \rightarrow (x,q))$ denote (A,p) and (x,q), respectively.

For $u,v \in V^*$, $u \Rightarrow v$ provided that $u = (rAs,p)$, $v = (rxs,q)$, for some $r,s \in V^*$, $(A,p) \rightarrow (x,q) \in P$, and for every $(B,p) \rightarrow (y,t) \in P$, $B \notin alph(r)$.

In the standard manner, extend $\Rightarrow$ to $\Rightarrow^n$, for $n \geq 0$, $\Rightarrow^+$, and $\Rightarrow^*$. The language generated by G is defined as $L(G) = \{w \in T^* : (S,q_0) \Rightarrow^* (w,q)$ for some $q \in Q\}$. Denote the family of languages generated by state grammars as $\mathscr{L}_{ST}$. And $\mathscr{L}_{ST} = \mathscr{L}_{RE}$ (see [2]).

A *left-forbidding grammar* is a quadruple $G = (N,T,P,S)$, where N is a nonterminal alphabet, T is a terminal alphabet such that $N \cap T = \emptyset$, $V = N \cup T$, $S \in N$ is the start symbol, and P is a finite set of productions of the form $(A \rightarrow x, W)$, where $A \in N$, $x \in V^*$, and $W \subseteq N$.

For $u,v \in V^*$ and $(A \rightarrow x,W) \in P$, $uAv \Rightarrow uxv$ provided that $alph(u) \cap W = \emptyset$. In the standard manner, extend $\Rightarrow$ to $\Rightarrow^n$, for $n \geq 0$, $\Rightarrow^+$, and $\Rightarrow^*$. The language generated by G is defined as $L(G) = \{w \in T^* : S \Rightarrow^* w\}$.

Let G be a left-forbidding grammar. Write $u \Rightarrow_t v$ in G if $u \Rightarrow^* v$ in G and for no $w \in V^*$, $v \Rightarrow w$ in G.

Let $n \geq 1$. A *left-forbidding cooperating distributed grammar system* is an $(n+3)$-tuple $\Gamma = (N,T,P_1,\ldots,P_n,S)$, where for $i = 1,\ldots,n$, $G_i = (N,T,P_i,S)$ is a left-forbidding grammar. For $u,v \in V^*$, $u \Rightarrow_i v$ denotes a derivation step made by a production from P_i.

Γ *t-generates* $z \in T^*$ if and only if, for some $l \geq 1$, there are $\alpha_i \in V^*$, for $i = 1,\ldots,l$, such that $\alpha_i \Rightarrow_t \alpha_{i+1}$ in H_i, $H_i \in \{G_1,\ldots,G_n\}$, $\alpha_1 = S$ and $\alpha_l = z$. Symbolically written as $S \Rightarrow^t z$.

The *t-language* generated by Γ is defined as $L(\Gamma,t) = \{w \in T^* : S \Rightarrow^t w\}$. Denote the family of languages t-generated by left-forbidding cooperating distributed grammar systems as $\mathscr{L}_t$. For $u,v \in V^*$, $u \Rightarrow_i^t v$ if $u \Rightarrow_t v$ in G_i.

3 MAIN RESULT

Theorem 1. $\mathscr{L}_t = \mathscr{L}_{RE}$.

Proof. Clearly, by Church's thesis, $\mathscr{L}_t \subseteq \mathscr{L}_{RE}$.

To prove the other inclusion, let L be a recursively enumerable language. There is a state grammar $G = (N,T,Q,P,S,q_0)$ such that $L(G) = L$. Construct a left-forbidding cooperating distributed grammar system $\Gamma = (N_\Gamma,T,P_1,P_2,S')$ with $N_\Gamma = N \cup N_1 \cup N_2 \cup N_3$, where

$$\begin{aligned}
N_1 &= \{[x,p,q,i] : x \in V \cup \{\varepsilon\},\ p,q \in Q,\ i \in \{1,2,3\}\}, \\
N_2 &= \{\langle w\rangle, [\langle w\rangle,p,q,i] : (X,p) \rightarrow (w,q) \in P\}, \\
N_3 &= \{\widehat{x} : x \in V \cup \{\varepsilon\}\} \cup \{\widehat{\langle w\rangle} : (X,p) \rightarrow (w,q) \in P\} \cup \{\langle BLOCK\rangle\}.
\end{aligned}$$

P_1 is constructed as follows:

1. For all $r \in Q$, add $(S' \rightarrow [S,q_0,r,3],\emptyset)$ to P_1.

2. For $[B,p,q,1] \in N_1$, if there is no $(B,p) \rightarrow (w,q) \in P$, add $([B,p,q,1] \rightarrow [B,p,q,2],\emptyset)$ to P_1.

3. For $(B,q) \rightarrow (w,h) \in P$, add $(B \rightarrow \langle w\rangle,W)$ and $(\widehat{B} \rightarrow \widehat{\langle w\rangle},W)$ to P_1, where

$W = \{[x,r,s,i],[\langle w\rangle,r,s,i],\langle w\rangle \in N_\Gamma : i=1\} \cup \{[x,r,s,i] \in N_1 : r \neq q \text{ or } s \neq h, \text{ and } i=2\} \cup \{X \in N : (X,q) \to (w,r) \in P\}$.

4. For $(B,q) \to (w,h) \in P$, add $([B,q,h,1] \to [\langle w\rangle,q,h,1],\emptyset)$ and $([B,q,h,3] \to [\langle w\rangle,q,h,3],\emptyset)$ to P_1.

5. For $a \in T \cup \{\varepsilon\}$, $q,h \in Q$, $i = 1,3$, add $([a,q,h,i] \to a,\emptyset)$ to P_1.

6. For $a \in T \cup \{\varepsilon\}$ add $(\widehat{a} \to a, N_\Gamma)$ to P_1.

7. For $a \in V \cup \{\varepsilon\}$, add $(\widehat{a} \to \langle BLOCK\rangle, \{[\langle w\rangle,q,h,i],\langle w\rangle \in N_2 : i=1\}$ to P_1.

P_2 is constructed as follows:

1'. For all $p,q,r \in Q$ and $x \in V \cup \{\varepsilon\}$, add $([x,p,q,2] \to [x,q,r,1],\emptyset)$ to P_2.

2'. For $(B,q) \to (w,h) \in P$, add $(\langle w\rangle \to w, \{[x,r,s,i] \in N_1 : i=2\})$ to P_2.

3'. For $(B,q) \to (w,h) \in P$ and $r \in Q$, add

a) for $|w| \geq 2$, $([\langle w\rangle,q,h,1] \to [w_1,h,r,1]w_2\dots w_{|w|-1}\widehat{w_{|w|}},\emptyset)$ to P_2;

b) for $|w| \leq 1$, $([\langle w\rangle,q,h,1] \to [w,h,r,1],\emptyset)$ to P_2.

4'. For $(B,q) \to (w,h) \in P$ and $r \in Q$, add

a) for $|w| \geq 2$, $([\langle w\rangle,q,h,3] \to [w_1,h,r,1]w_2\dots w_{|w|-1}\widehat{w_{|w|}},\emptyset)$ to P_2;

b) for $|w| \leq 1$, $([\langle w\rangle,q,h,3] \to [w,h,r,3],\emptyset)$ to P_2.

5'. For $(B,q) \to (w,h) \in P$, add

a) for $|w| \geq 2$, $(\widehat{\langle w\rangle} \to w_1\dots w_{|w|-1}\widehat{w_{|w|}}, \{[x,r,s,i] \in N_1 : i=2\})$ to P_2;

b) for $|w| \leq 1$, $(\widehat{\langle w\rangle} \to \widehat{w}, \{[x,r,s,i] \in N_1 : i=2\})$ to P_2.

6'. For all $\widehat{x} \in N_3$, add $(\widehat{x} \to \langle BLOCK\rangle, \{[y,r,s,i] \in N_1 : i \in \{1,2\}\})$ to P_2.

To prove that $L(G) \subseteq L(\Gamma)$, consider a derivation step $(\alpha',q) \Rightarrow (\beta',h)$ in G. Let $\alpha' = a_1a_2\dots a_n$ and $\beta' = b_1b_2\dots b_m$, where for $i = 1,\dots,n$, $j = 1,\dots,m$, $a_i,b_j \in V$. We prove that $\alpha \Rightarrow^+ \beta$ in Γ, for $\alpha = [x_1,q,r,1]x_2\dots x_{l-1}\widehat{x_l}$, or $\alpha = [a_1,q,r,3]$ if $|\alpha'| = 1$, where $l \geq 2$, $x_1,x_l \in V \cup \{\varepsilon\}$, $x_i \in V$, for $i = 2,\dots,l-1$, and for some $r \in Q$, such that $x_1x_2\dots x_l = a_1a_2\dots a_n$.

Assume that $\alpha' \Rightarrow \beta'$ by a production $(a_i,q) \to (w,h) \in P$. Then, $(a_1\dots a_{i-1}a_ia_{i+1}\dots a_n,q) \Rightarrow (a_1\dots a_{i-1}wa_{i+1}\dots a_n,h)$ in G.

(1) If $i=1$ and $|w| \geq 2$, then $[a_1,q,h,3] \Rightarrow_1 [\langle w\rangle,q,h,3] \Rightarrow_2 [w_1,h,r,1]w_2\dots w_{|w|-1}\widehat{w_{|w|}}$ in Γ by productions (4) and (4'a), for some $r \in Q$.

(2) If $i=1$ and $|w| \leq 1$, then $[a_1,q,h,3] \Rightarrow_1 [\langle w\rangle,q,h,3] \Rightarrow_2 [w,h,r,3]$ in Γ by productions (4) and (4'b), for some $r \in Q$.

(3) If $i=1$ and $x_1 = a_1$, then $[x_1,q,h,1]x_2\dots x_{l-1}\widehat{x_l} \Rightarrow_1 [\langle w\rangle,q,h,1]x_2\dots x_{l-1}\widehat{x_l}$
$\Rightarrow_2 [w_1,h,r,1]w_2\dots w_{|w|}x_2\dots x_{l-1}\widehat{x_l}$ in Γ by productions (4) and (3'), for some $r \in Q$, where for $w=\varepsilon$ we have $w_1 = \varepsilon$.

(4) If $1 < i < n$, or $i=1$ and $x_1 \neq a_1$, or $i=n$ and $x_l \neq a_n$, then

$$[x_1,q,h,1]x_2\ldots x_{l-1}\widehat{x_l} \Rightarrow_1 [x_1,q,h,2]x_2\ldots x_{i-1}x_ix_{i+1}\ldots x_{l-1}\widehat{x_l} \Rightarrow_1$$
$$[x_1,q,h,2]x_2\ldots x_{i-1}\langle w\rangle x_{i+1}\ldots x_{l-1}\widehat{x_l} \Rightarrow_2 [x_1,h,r,1]x_2\ldots x_{i-1}\langle w\rangle x_{i+1}\ldots x_{l-1}\widehat{x_l} \Rightarrow_2$$
$$[x_1,h,r,1]x_2\ldots x_{i-1}wx_{i+1}\ldots x_{l-1}\widehat{x_l}$$

in Γ by productions (2), (3), (1'), and (2'), for some $r \in Q$. Recall that G rewrites the leftmost nonterminal rewritable in the current sentential form.

(5) If $i = n$ and $x_l = a_n$, then

$$[x_1,q,h,1]x_2\ldots x_{l-1}\widehat{x_l} \Rightarrow_1 [x_1,q,h,2]x_2\ldots x_{l-1}\widehat{x_l} \Rightarrow_1$$
$$[x_1,q,h,2]x_2\ldots x_{l-1}\widehat{\langle w\rangle} \Rightarrow_2 [x_1,h,r,1]x_2\ldots x_{l-1}\widehat{\langle w\rangle} \Rightarrow_2$$
$$[x_1,h,r,1]x_2\ldots x_{l-1}w_1\ldots w_{|w|-1}\widehat{w_{|w|}}$$

in Γ by productions (2), (3), (1'), and (5'), for some $r \in Q$, where for $w = \varepsilon$ we have $\widehat{w_1} = \widehat{\varepsilon}$. The proof then follows by induction.

Assume that $a_1\ldots a_n \in T^*$. Then, $[a_1,q,h,3] \Rightarrow_1 a_1$ by a production constructed in (5), and $[x_1,q,h,1]x_2\ldots x_{l-1}\widehat{x_l} \Rightarrow_1^t x_1x_2\ldots x_{l-1}x_l$ by productions constructed in (5) and (6).

Clearly, Γ simulates a derivation of G so that it starts by a production constructed in (1) and then it continues as shown above.

To prove that $L(\Gamma) \subseteq L(G)$, consider a terminating derivation in Γ. Such a derivation is of the form $S \Rightarrow_1^t \ldots \Rightarrow_2^t x_0 \Rightarrow_1^t x_1 \Rightarrow_2^t x_2 \Rightarrow_1^t \ldots \Rightarrow_2^t w' \Rightarrow_1^t w$, for some $w \in T^*$. Consider a subderivation, $x_0 \Rightarrow_1^t x_1 \Rightarrow_2^t x_2$, and examine the forms of this subderivation.

Assume that $x_0 = [a,p,q,3]$. If $a \in N$, then only a production constructed in (4), corresponding to $(a,p) \to (u,q) \in P$, is applicable.

$|u| \geq 2$**:** Then, after the production constructed in (4), no production from P_1 is applicable and, only a production constructed in (4'a) is applicable. Thus, $[a,p,q,3] \Rightarrow_1^t [\langle u\rangle,p,q,3] \Rightarrow_2^t [u_1,q,h,1]u_2\ldots u_{|u|-1}\widehat{u_{|u|}}$ in Γ.

$|u| \leq 1$**:** Then, after the production constructed in (4), no production from P_1 is applicable, and only a production constructed in (4'b) is applicable. Thus, $[a,p,q,3] \Rightarrow_1^t [\langle u\rangle,p,q,3] \Rightarrow_2^t [u,q,h,3]$ in Γ.

Clearly, $(a,p) \Rightarrow (u,q)$ in G. If $a \in T \cup \{\varepsilon\}$, then only a production constructed in (5) is applicable. Thus, $[a,p,q,3] \Rightarrow_1^t a$.

Assume that $x_0 = [a_1,p,q,1]a_2\ldots a_{n-1}\widehat{a_n}$. If $a_1 \in N$, then:

1. Assume that $(a_1,p) \to (u,q) \in P$. Then, no productions constructed in (2) and (3) are applicable because all nonterminals of the form $[x,r,s,1] \in N_\Gamma$ are included in the forbidding set of productions constructed in (3). Thus, only a production constructed in (4), corresponding to $(a_1,p) \to (u,q)$, is applicable. Then, only a production constructed in (3') is applicable. Thus, $[a_1,p,q,1]a_2\ldots a_{n-1}\widehat{a_n} \Rightarrow_1^t [\langle u\rangle,p,q,1]a_2\ldots a_{n-1}\widehat{a_n} \Rightarrow_2^t [u_1,q,h,1]u_2\ldots u_{|u|}a_2\ldots a_{n-1}\widehat{a_n}$, where $u_1 = \varepsilon$ if $|u| = 0$. Again, $(a_1\ldots a_n,p) \Rightarrow (ua_2\ldots a_n,q)$ in G.

2. Assume that there is no production $(a_1,p) \to (v,q)$ in P. Then, only a production constructed in (2) is applicable. Then, only a production constructed in (3) is applicable, corresponding to $(a_j,p) \to (u,q) \in P$, for some $1 < j \leq n$, and such that there is no applicable production $(a_k,p) \to (w,t) \in P$, for all $k < j$; of course, if there is no such production, then a production constructed in (7) blocks the derivation. Then, only a production constructed in (1') is applica-

ble. Thus, $[a_1,p,q,1]a_2\dots a_{j-1}a_ja_{j+1}\dots a_{n-1}\widehat{a_n} \Rightarrow_1 [a_1,p,q,2]a_2\dots a_{j-1}a_ja_{j+1}\dots a_{n-1}\widehat{a_n} \Rightarrow_1$ $[a_1,p,q,2]a_2\dots a_{j-1}\langle u\rangle a_{j+1}\dots a_{n-1}\widehat{a_n} \Rightarrow_2 [a_1,q,h,1]a_2\dots a_{j-1}\langle u\rangle a_{j+1}\dots a_{n-1}\widehat{a_n}$.

(i) Assume that $j < n$. Then, only a production constructed in (2') is applicable. Thus,

$$[a_1,q,h,1]a_2\dots a_{j-1}\langle u\rangle a_{j+1}\dots a_{n-1}\widehat{a_n} \quad \Rightarrow_2^t \quad [a_1,q,h,1]a_2\dots a_{j-1}ua_{j+1}\dots a_{n-1}\widehat{a_n}.$$

Again, $(a_1\dots a_n,p) \Rightarrow (a_1\dots a_{j-1}ua_{j+1}\dots a_n,q)$ in G.

(ii) Assume that $j = n$. Then, only a production constructed in (5') is applicable. Thus,

$$[a_1,q,h,1]a_2\dots a_{n-1}\widehat{\langle u\rangle} \quad \Rightarrow_2^t \quad [a_1,q,h,1]a_2\dots u_1\dots u_{|u|-1}\widehat{u_{|u|}}.$$

In case $|u| = 0$, we have $u_{|u|} = \varepsilon$. Again, $(a_1\dots a_{n-1}a_n,p) \Rightarrow (a_1\dots a_{n-1}u,q)$ in G.

If $a_1 \in T\cup\{\varepsilon\}$, then only productions constructed in (2), (5) and (7) are applicable.

1. Let, for all $i = 2,\dots,n$, $a_i \in T\cup\{\varepsilon\}$. Consider a production constructed in (2), then

$$[a_1,p,q,1]a_2\dots a_{n-1}\widehat{a_n} \quad \Rightarrow_1 \quad [a_1,p,q,2]a_2\dots a_{n-1}\widehat{a_n}$$

and only a production constructed in (7) is applicable. However, these productions block the derivation. Thus, only a production constructed in (5) is now applicable in the terminal derivation, i.e., $[a_1,p,q,1]a_2\dots a_{n-1}\widehat{a_n} \Rightarrow_1 a_1a_2\dots a_{n-1}\widehat{a_n}$. Then, only productions constructed in (6) and (7) are applicable. Again, (7) blocks the derivation, thus consider a production constructed in (6), i.e. $a_1a_2\dots a_{n-1}\widehat{a_n} \Rightarrow_1^t a_1a_2\dots a_{n-1}a_n$, which finishes the derivation.

2. Let there be $i \in \{2,\dots,n\}$ such that $a_i \in N$. Consider a production constructed in (2) is applied first, then the derivation continues as in 2 above because there is no production $(a_1,p) \to (v,q)$ in P.

Consider a production constructed in (5) is applied first, then only productions constructed in (3) and (7) are applicable. Again, (7) blocks the derivation, thus consider a production constructed in (3). Then, $a_1\dots a_k\dots a_{n-1}\widehat{a_n} \Rightarrow_1^t a_1\dots\langle w\rangle\dots a_{n-1}\widehat{a_n}$ and some other productions constructed in (3) are applicable. After this, no production from P_1 is applicable and only productions constructed in (2'), (5'), and (6') are applicable. In all cases, production (6') has to be applied—the derivation is blocked. The proof then follows by induction.

Any derivation of Γ starts by a production constructed in (1), $S' \Rightarrow_1 [S,q_0,r,3]$, for some $r \in Q$, and then continues as proved above. □

REFERENCES

[1] Jürgen Dassow, Gheorghe Păun, and Grzegorz Rozenberg. Grammar systems. In *Handbook of formal languages, Vol. 2*, pages 155–213. Springer-Verlag, Berlin, 1997.

[2] G. Horvat and A. Meduna. On state grammars. *Acta Cybernetica*, pages 4–14, 1988.

[3] T. Kasai. An hierarchy between context-free and context-sensitive languages. *Journal of Computer and System Sciences*, 4(5):492–508, 1970.

[4] A. Meduna. *Automata and Languages: Theory and Applications*. Springer-Verlag, London, 2005.