ON RELATIONS ON PRODUCTIONS FOR COOPERATIVE DISTRIBUTED GRAMMAR SYSTEMS

Filip Goldefus
Postgraduate Programme (3), FIT BUT
E-mail: igoldefu@fit.vutbr.cz

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

ABSTRACT

The present paper introduces cooperative distributed grammar systems with ordered grammars as components. These grammars have a ordering on productions, which leads to a increase of the generative power compared to a cooperative distributed grammar systems with context-free grammars as components. The cooperating mode =2 is investigated and proved that cooperative distributed grammar systems with ordered grammars as components are as powerful as programmed grammars with appearance checking containing erasing productions.

1 INTRODUCTION

In the formal language theory, cooperative distributed grammar systems are based on context-free productions, or more precisely context-free grammars. The present paper introduces ordered grammars as components of cooperative distributed grammar systems and investigates their generative power.

The ordered grammars ([3]), as their name indicates, has an ordering on productions, which limits the nondeterminism on derivations, such that not every production is applicable on a sentential form, compared to the context-free grammars with same productions and sentential form.

This paper proves that for every programmed grammar with appearance checking consisting erasing productions ([1]), there exists a cooperative distributed grammar system working in mode =2 generating the same language. The class of languages generated by programmed grammars with appearance checking is equal to the class of recursively enumerable languages from Chomsky hierarchy ([2]).

2 PRELIMINARIES AND DEFINITIONS

We assume that reader is familiar with the language theory (see [2]). A *context-free grammar* is a quadruple, $G = (N, T, S, P)$, where N is a finite set of nonterminal symbols, T is a finite set of terminal symbols, $S \in N$ is the starting nonterminal (axiom), and P is a finite set of productions of the form $p : A \to \alpha$, with $A \in N, \alpha \in (N \cup T)^*$ and p is unique label. For $p : A \to v$ and $x, y \in V^*$, we say that x directly derives y, written as $x = uAw \Rightarrow uvw = y$ $[p]$ or, simply, $x \Rightarrow y$.

In the standard manner, extend $\Rightarrow$ to $\Rightarrow^n$, where $n \geq 0$; then, based on $\Rightarrow^n$, define $\Rightarrow^+$ and $\Rightarrow^*$. The language of G, $L(G)$, is defined as $L(G) = \{w \in T^* | S \Rightarrow^* w\}$.

A *programmed grammar with appearance checking* is a triple, $H = (G,R,F)$, where $G = (N,T,S,P)$ is a context-free grammar, and R,F are finite relations on P. If $p : A \rightarrow v \in P$, $R(p) = W$, and $F(p) = X$, we write $(p : A \rightarrow x, W, X)$, where W and X are success and failure fields, respectively. For $(x,p),(y,q) \in (N \cup T)^* \times P$, $(x,p) \Rightarrow (y,q)$ in H if either $x \Rightarrow y$ $[p]$ in G and $q \in R(p)$, or $x = y$, $q \in F(p)$, p is not applicable to x. The language of H, $L(H)$, is defined as $L(H) = \{w \in T^* | (S,p) \Rightarrow^* (w,q), p,q \in P\}$. For every programmed grammar with appearance checking, $H = (G,R,F)$, where $G = (N,T,S,P)$, there exists a *well-formed programmed grammar with appearance checking* $M = (G',R',F')$, with $G = (N,T,S,P')$, such that $L(H) = L(M)$ and for every production $p \in P$', $R'(p) \neq \emptyset$ and $F'(p) \neq \emptyset$. The proof is left to reader.

An *ordered grammar* is a quadruple $G = (N,T,S,P)$ where N,T and S are specified as in a context-free grammar and P is a finite partially ordered set of context-free productions, the ordering relation is transitive, denoted by $<$. For $x,y \in (N \cup T)^*$, $x \Rightarrow y$, iff there is a production $p : A \rightarrow w$ such that $x = x'Ax''$, $y = x'wx''$ and there is no production $q : B \rightarrow v \in P$ such that $q < p$ and B occurs in x, we say p is greater than q.

A *ordered cooperating distributed (OCD) grammar system of degree n* is an $(n+3)$-tuple $\Gamma = (N,T,S,\ P_1,\dots,P_n)$, where for all $i = 1,\dots,n$, each component $G_i = (N,T,P_i,S)$ is a ordered grammar, for $n \geq 1$.

Let $\Gamma = (N,T,S,P_1,\dots,P_n)$ be a OCD grammar system of degree n, for $1 \leq i \leq n$, the *k-steps* (=*k*-mode) *derivation* of *i*-th component denoted $\Rightarrow_i^{=k}$, is defined by $x \Rightarrow_i^{=k} y$ for $x,y \in (N \cup T)^*$ iff there are $x_1,\dots,x_k \in (N \cup T)^*$such that $x = x_1$, $y = x_{k+1}$ and $x_j \Rightarrow x_{j+1}$ for each $1 \leq j \leq k$ in a ordered grammar $G_i = (N,T,P_i,S)$. The $\Rightarrow^{=k*}$ denotes the reflexive and transitive closure of the relation $\Rightarrow_i^{=k}$. The language of Γ in =*k*-mode is defined as $L_{=k}(\Gamma) = \{w \in T^* | S \Rightarrow_{i_1}^{=k} w_1 \Rightarrow_{i_2}^{=k} \dots \Rightarrow_{i_n}^{=k} w_n = w, n \geq 1, i_j \in \{1,\dots,n\}, 1 \leq j \leq n\}$.

The families of languages generated by programmed grammars with appearance checking consisting erasing productions, ordered grammars, cooperating grammars of degree n in mode =*k* and ordered cooperating grammars of degree n in mode =k, respectively, are denoted by P_{ac}^{ε}, OR, $CD_{=k}(n)$, $OCD_{=k}(n)$.

3 MAIN RESULTS

This section proves that cooperative distributed grammar systems with ordered grammars as components are as powerful as programmed grammars with appearance checking.

Theorem 1. $P_{ac}^{\varepsilon} = \bigcup_{n=1}^{\infty} OCD_{=1}(n)$.

Proof. Let $H = (G,R,F)$ be a well-formed programmed grammar with appearance checking and $G = (N,T,S,P)$, construct a OCD grammar system of degree $2|P|+1$, $\Gamma = (\mathcal{N},T,S_\Gamma,\ P_0,\dots,\ P_{2|P|})$, such that $\mathcal{N} = \{S_\Gamma\} \cup N_{\langle\rangle} \cup \overline{N}_{\langle\rangle}$ with $N_{\langle\rangle} = \{\langle X;p\rangle | X \in N \cup T, p : A \rightarrow v \in P\}$ and $\overline{N}_{\langle\rangle} = \{\overline{X} | X \in N_{\langle\rangle}\} \cup \{\overline{\langle\varepsilon;p\rangle} | p : A \rightarrow v \in P\}$. The sets of productions are defined as follows.

Let $p : X \rightarrow X_1X_2\dots X_n \in P$, $X_i \in (N \cup T), 1 \leq i \leq n, q \in R(p)$, and $r \in F(p)$, create a set P_k such that k is unique, $1 \leq k \leq |P|$, and $P_k = P_k^1 \cup P_k^2 \cup P_k^3$, where

1. $P_k^1 = \{X \to X | X \in \overline{N}_{\langle\rangle}\}$,
2. If $n \geq 1$, then $P_k^2 = \{\langle X;p\rangle \to \overline{\langle X_1;q\rangle}\langle X_2;p\rangle \dots \langle X_n;p\rangle\}$, else $P_k^2 = \{\langle X;p\rangle \to \overline{\langle \varepsilon;q\rangle}\}$
3. $P_k^3 = \{\langle Y;p\rangle \to \overline{\langle Y;r\rangle} | Y \in (N \cup T)\}$.

The following inequations hold, for all $p \in P_k^1$, $q \in P_k^2$, $r \in P_k^3$, $p < q$, $p < r$ and $q < r$.

Create a set $P_k = P_k^1 \cup P_k^2$ corresponding to a production $p : X \to v \in P$, with unique k, $|P| + 1 \leq k \leq 2|P|$ such that

1. $P_k^1 = \{\langle X;q\rangle \to \langle X;p\rangle | X \in (N \cup T), q : Y \to z \in P - \{p : X \to v\}\}$,
2. If $|v| \geq 1$, then $P_k^2 = \{\overline{\langle X;p\rangle} \to \langle X;p\rangle | X \in N \cup T\}$, else $P_k^2 = \{\overline{\langle \varepsilon;p\rangle} \to \varepsilon\}$,
3. $P_k^3 = \{X \to X | X \in \mathcal{N}\}$.

Following inequation holds, for all $p \in P_k^1$, and for all $q \in P_k^2$, $p < q$.

The set P_0 is constructed as follows, $P_0 = P_0^1 \cup P_0^2 \cup P_0^3$, with

1. $P_0^1 = \{X \to X | X \in \overline{N}_{\langle\rangle}\}$,
2. $P_0^2 = \{\langle a;p\rangle \to a | a \in T,\ p : X \to v \in P\}$,
3. $P_0^3 = \{S_\Gamma \to \langle S;p\rangle | p : S \to v \in P\} \cup \{X \to X | X \in \mathcal{N}\}$.

The following inequation holds, for all $p \in P_k^1$, and for all $q \in P_k^2$, $p < q$.

The cooperative distributed grammar system Γ simulates derivation steps of the programmed grammar with appearance checking H. A typical sentential form of Γ is of the form

$$\langle X_1;p\rangle\langle X_2;p\rangle \dots \langle X_n;p\rangle.$$

This form corresponds to the configuration $(X_1X_2\dots X_n, p)$ of H. Grammar Γ simulates one derivation step of grammar H by a sequence of derivation steps. If a sentential form of Γ contains a nonterminal $\overline{\langle X_k;q\rangle} \in \overline{N}_{\langle\rangle}$ then remaining nonterminals in sentential form are synchronized by productions from a set P_k, $|P| + 1 \leq k \leq 2|P|$, corresponding to the production labeled by q, to the form $\langle X_i;q\rangle \in N_{\langle\rangle}$, where the second component of nonterminal has to be the label of production $q : X_j \to \beta \in P$.

Every set of production P_k, $1 \leq k \leq 2|P|$, corresponds to a production from programmed grammar H. Some sets of productions contain productions of the form $X \to X$, ensuring that a sentential form keeps unchanged in case that it contains the nonterminal X.

To prove that $L(H) \subseteq L(\Gamma)$, consider a derivation $(S,r) \Rightarrow^* (A_1A_2\dots A_i\dots A_n, q) \Rightarrow (\beta, p)$ in H using a production $p : X \to B_1\dots B_m \in P, r \in Q, R(p) \neq \emptyset$ and $F(p) \neq \emptyset$. For $i = 1,\dots,n$, $A_i \in (N \cup T)$.

Sentential form of Γ is of the form

$$\alpha = \langle A_1;q\rangle\langle A_2;q\rangle \dots \langle A_{j-1};q\rangle\overline{\langle A_j;p\rangle}\langle A_{j+1};q\rangle \dots \langle A_n;q\rangle$$

then there exist k, such that $|P|+1 \leq k \leq 2|P|$ corresponding to the production labeled with p. If $A_j \in (N \cup T)$,

$$\alpha \Rightarrow_k^{=2} \langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_{j-1};p\rangle\langle A_j;p\rangle\langle A_{j+1};p\rangle\ldots\langle A_n;p\rangle,$$

else $A_j = \varepsilon$ and

$$\alpha \Rightarrow_k^{=2} \langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_{j-1};p\rangle\langle A_{j+1};p\rangle\ldots\langle A_n;p\rangle$$

in Γ by multiple application of productions from the set of productions P_k.

Now, there exists P_k, with $1 \leq k \leq |P|$ corresponding to the production labeled with p, and a nonterminal $A_i = X$ in the sentential form, so for a production $s: X \to X_1 \ldots X_o \in R(p)$. If $m \geq 1$, then

$$\begin{array}{ll}\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_i;p\rangle\ldots\langle A_n;p\rangle & \Rightarrow_k^{=2} \\ \langle A_1;p\rangle\langle A_2;p\rangle\ldots\overline{\langle B_1;s\rangle}\langle B_2;p\rangle\ldots\langle B_m;p\rangle\ldots\langle A_n;p\rangle, & \end{array}$$

else

$$\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_i;p\rangle\ldots\langle A_n;p\rangle \Rightarrow_k^{=2} \langle A_1;p\rangle\langle A_2;p\rangle\ldots\overline{\langle \varepsilon;s\rangle}\ldots\langle A_n;p\rangle$$

in Γ. Finally, consider that a nonterminal $\langle X;p\rangle$ is not present in the sentential form and $r \in R(p)$, thus

$$\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_i;p\rangle\ldots\langle A_n;p\rangle \Rightarrow_k^{=2} \langle A_1;p\rangle\langle A_2;p\rangle\ldots\overline{\langle A_i;r\rangle}\ldots\langle A_n;p\rangle$$

in Γ and the derivation proceeds by induction.

Let $\langle a_1;p\rangle\langle a_2;p\rangle\ldots\langle a_n;p\rangle$ be a sentential form of Γ and $a_i \in T$ for all $1 \leq i \leq n$, then only productions form the set P_0 are applicable and $\langle a_1;p\rangle\langle a_2;p\rangle\ldots\langle a_n;p\rangle \Rightarrow_0^{=2*} a_1\ldots a_n$. To prove that $L(\Gamma) \subseteq L(H)$, consider a shortest derivation of the form

$$S_\Gamma \Rightarrow_0^{=2} \ldots \Rightarrow^{=2*} \langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_n;p\rangle \Rightarrow_k^{=2*} \langle B_1;q\rangle\langle B_2;q\rangle\ldots\langle B_m;q\rangle$$

in Γ. Without any loss of generality productions from the set P_0 are applied on

$$\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_n;p\rangle,$$

for $A_1\ldots A_n \in T^+$. Consider k, $1 \leq k \leq |P|$, if set P_k corresponds to a production $t: Y \to \alpha \in P$, if $t \neq p$, then there is no production applicable on the sentential form $\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_n;p\rangle$. If $p = t$, $p: X \to D_1\ldots D_s$, $A_i = X$ for some $1 \leq i \leq n$ and $q \in R(p)$, then for $s \geq 1$

$$\begin{array}{ll}\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_i;p\rangle\ldots\langle A_n;p\rangle & \Rightarrow_k^{=2} \\ \langle A_1;p\rangle\langle A_2;p\rangle\ldots\overline{\langle D_1;q\rangle}\langle D_2;p\rangle\ldots\langle D_s;p\rangle\ldots\langle A_n;p\rangle & \end{array}$$

and for $s = 0$, $\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_i;p\rangle\ldots\langle A_n;p\rangle \Rightarrow_k^{=2} \langle A_1;p\rangle\langle A_2;p\rangle\ldots\overline{\langle \varepsilon;q\rangle}\ldots\langle A_n;p\rangle$ in Γ.

Now, assume that $\langle X;p\rangle$ is not present in the sentential form and $q \in F(p)$, then

$$\begin{array}{ll}\langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_i;p\rangle\ldots\langle A_n;p\rangle & \Rightarrow_k^{=} \\ \langle A_1;p\rangle\langle A_2;p\rangle\ldots\langle A_{i-1};p\rangle\overline{\langle A_i;q\rangle}\langle A_{i+1};p\rangle\ldots\langle A_n;p\rangle & \end{array}$$

in Γ for some $1 \leq i \leq n$.

Let sentential form is of the form $\langle B_1;p\rangle\langle B_2;p\rangle\ldots\langle B_{i-1};p\rangle\overline{\langle B_i;q\rangle}\langle B_{i+1};p\rangle\ldots\langle B_m;p\rangle$. All sets of productions except P_k, $|P|+1 \leq k \leq 2|P|$, corresponding to a production $q : Z \to \beta \in P$, contains productions $X \to X$ for $X \in \overline{N}_{\langle\rangle}$. The set of productions P_k ensures, that nonterminals $\langle B_j;p\rangle \in N_{\langle\rangle}$ will be rewritten on $\langle B_j;q\rangle$, $j = \{1,\ldots,m\} - \{i\}$, and consequently for $B_i \in (N \cup T)$,

$$\begin{array}{ll}\langle B_1;q\rangle\langle B_2;q\rangle\ldots\langle B_{i-1};q\rangle\overline{\langle B_i;q\rangle}\langle B_{i+1};q\rangle\ldots\langle B_m;q\rangle & \Rightarrow_k^{=2} \\ \langle B_1;q\rangle\langle B_2;q\rangle\ldots\langle B_{i-1};q\rangle\langle B_i;q\rangle\langle B_{i+1};q\rangle\ldots\langle B_m;q\rangle, & \end{array}$$

and for $B_i = \varepsilon$

$$\begin{array}{ll}\langle B_1;q\rangle\langle B_2;q\rangle\ldots\langle B_{i-1};q\rangle\overline{\langle \varepsilon_i;q\rangle}\langle B_{i+1};q\rangle\ldots\langle B_m;q\rangle & \Rightarrow_k^{=2} \\ \langle B_1;q\rangle\langle B_2;q\rangle\ldots\langle B_{i-1};q\rangle\langle B_{i+1};q\rangle\ldots\langle B_m;q\rangle, & \end{array}$$

in Γ. The proof now proceeds by induction.

As any derivation of Γ finishes by using productions from P_0 when $b_1\ldots b_m \in T^+$, so

$$\langle b_1;q\rangle\langle b_2;q\rangle\ldots\langle b_m;q\rangle \Rightarrow_0^{=2*} b_1b_2\ldots b_m.$$

By Church's thesis, $P_{ac}^{\varepsilon} = RE$, so $P_{ac}^{\varepsilon} = \bigcup_{n=1}^{\infty} OCD_{=2}(n)$. □

4 CONCLUSIONS

We denote by CF the class of context-free languages, FOR denotes the class of languages generated by forbidding grammars and CS denotes the class of context sensitive languages. Recall that it is well-known (see [4]) that $CF = \bigcup_{n=1}^{\infty} CD_{=1}(n)$, $FOR = OR \subset CS$, $P_{ac}^{\varepsilon} = RE$. Previous section proved that $RE = P_{ac}^{\varepsilon} = \bigcup_{n=1}^{\infty} OCD_{=2}(n)$.

REFERENCES

[1] Jurgen Dassow and Gheorge Paun. *Regulated Rewriting in Formal Language Theory*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1990.

[2] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages, and Computation (3rd Edition)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2006.

[3] Carlos Martín-Vide, Victor Mitrana, and Gheorghe Paun. *Formal Languages and Applications*. Springer-Verlag, London, UK, 2004.

[4] Grzegorz Rozenberg and Arto Salomaa, editors. *Handbook of Formal Languages, vol. 2: Linear Modeling: Background and Application*. Springer-Verlag New York, Inc., New York, NY, USA, 1997.