Regulated Rewriting – The Hierarchy

Jiří Techet Tomáš Masopust Alexander Meduna

Department of Information Systems
Faculty of Information Technology
Brno University of Technology
Božetěchova 2, Brno 61266, Czech Republic

Modern Formal Language Theory, 2007

Hierarchy of Language Families – Description

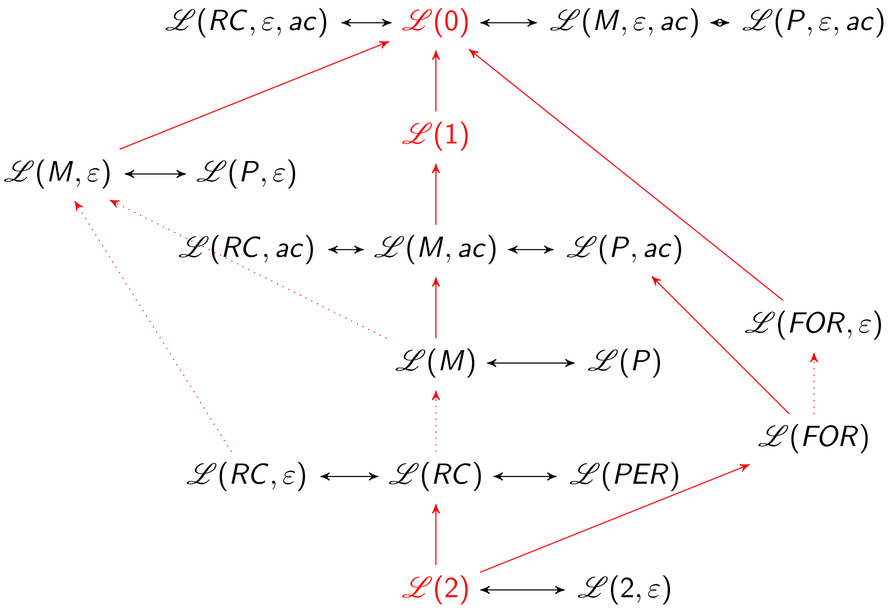
Notation

- ac appearance checking
- ε with erasing productions
- M matrix grammars
- P programmed grammars
- RC random context grammars
- PER permitting grammars
- FOR forbidding grammars
- 0, 1, 2 type-0, type-1, type-2 grammars, respectively

Example

$\mathscr{L}(RC, \varepsilon, ac)$ – the family of languages generated by random context grammars with erasing productions in appearance checking mode

Regulated Rewriting – The Hierarchy



Proof of $\mathscr{L}(M) = \mathscr{L}(P)$

Theorem

$\mathscr{L}(M) = \mathscr{L}(P)$.

Proof–Basic Idea

1. $\mathscr{L}(M) \subseteq \mathscr{L}(P)$: Transform any matrix grammar to an equivalent programmed grammar
2. $\mathscr{L}(P) \subseteq \mathscr{L}(M)$: Transform any programmed grammar to an equivalent matrix grammar
3. $\mathscr{L}(M) = \mathscr{L}(P)$

Proof of $\mathscr{L}(M) \subseteq \mathscr{L}(P)$ I

Let

$$H = (G, M)$$

be a matrix grammar, where

- $G = (N, T, P, S)$ is a context-free grammar
- M is a finite language over P

Express M as

$$M = \{m_1, \ldots, m_r\}$$

for some $r \geq 1$ and

$$m_i = p_{i_1} \ldots p_{i_{k_i}}$$

with $k_i = |m_i|$. Set

$$N' = \{A' : A \in N\}$$

Define the homomorphism h from $(T \cup N)^*$ to $(T \cup N')^*$ as

- $h(a) = a$ for every $a \in T$
- $h(A) = A'$ for every $A \in N$

Proof of $\mathscr{L}(M) \subseteq \mathscr{L}(P)$ II

1. For every $p_{i_j} : A_{i_j} \rightarrow x_{i_j}$ such that $\text{alph}(x_{i_j}) \cap N \neq \emptyset$ (some nonterminals occur in x_{i_j}) and $j < |m_i|$ (not the last production in m_i), introduce this programmed production:

$$(\langle i, j \rangle : A_{i_j} \rightarrow x_{i_j}, \{\langle i, j+1 \rangle\})$$

2. For every $p_{i_j} : A_{i_j} \rightarrow x_{i_j}$ such that $\text{alph}(x_{i_j}) \cap N = \emptyset$ (no nonterminal occurs in x_{i_j}) and $j < |m_i|$ (not the last production in m_i), introduce
   - a $(\langle i, j \rangle : A_{i_j} \rightarrow h(A_{i_j}), \{\langle i, j, 1 \rangle, \ldots, \langle i, j, m \rangle\})$
   - b $(\langle i, j, q \rangle : A_q \rightarrow A_q, \{\langle i, j \rangle'\})$ for $q = 1, \ldots, m$ (make sure there is a nonterminal)
   - c $(\langle i, j \rangle' : h(A_{i_j}) \rightarrow x_{i_j}, \{\langle i, j+1 \rangle\})$

   provided that $N = \{A_1, \ldots, A_m\}$
3. For every $p_{i_{k_i}} : A_{i_{k_i}} \rightarrow x_{i_{k_i}}$ (simulate the last production of m_i and start simulating a new matrix), introduce

$$(\langle i, k_i \rangle : A_{i_{k_i}} \rightarrow x_{i_{k_i}}, \{\langle 1, 1 \rangle, \ldots, \langle r, 1 \rangle\})$$

Proof of $\mathscr{L}(M) \subseteq \mathscr{L}(P)$ III

Let P' consist of all rules constructed above and

$$(\$: S' \rightarrow S, \{\langle 1, 1\rangle, \ldots, \langle r, 1\rangle\})$$

Define the programmed grammar

$$G' = (N \cup N' \cup \{S'\}, T, P', S')$$

Observe that

$$L(H) = L(G')$$

$\square$

Proof of $\mathscr{L}(P) \subseteq \mathscr{L}(M)$ I

Let

$$H = (G, R)$$

be a programmed grammar. Consider the simplified description of H which uses productions of the form

$$(p : A \rightarrow x, R(p))$$

1. If $(p : A \rightarrow x, Q)$, $(r : B \rightarrow y, R) \in P$ and $r \in Q$, then introduce the matrix

$$(\langle A \rightarrow x \rangle \rightarrow A, A \rightarrow x, B \rightarrow \langle B \rightarrow y \rangle)$$

   (simulation will continue)

2. If $(p : A \rightarrow x, Q) \in P$, then introduce

$$(\langle A \rightarrow x \rangle \rightarrow A, A \rightarrow x)$$

   (simulation ends)

Proof of $\mathscr{L}(P) \subseteq \mathscr{L}(M)$ II

3. If $(p : S \to x, Q) \in P$, introduce

$$(S' \to \langle S \to x \rangle)$$

(simulation starts)

Set

$$\begin{aligned} N' = N \\ \cup \{\langle A \to x\rangle : A \to x \in P\} \\ \cup \{S'\} \end{aligned}$$

Define the matrix grammar

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

where P' consists of all the above matrices. Observe that

$$L(H) = L(G')$$

□

Bibliography I

S. Abraham.
Some questions of phrase-structure grammars.
Computational Linguistics, 4:61–70, 1965.

J. Dassow and Gh. Păun.
Regulated Rewriting in Formal Language Theory.
Akademie-Verlag, Berlin, 1989.

H. Fernau and H. Bordihn.
Accepting grammars and systems via context condition grammars.
Journal of Automata, Languages and Combinatorics, 1(2):97–112, 1996.

D. J. Rosenkrantz.
Programmed grammars and classes of formal languages.
Journal of the ACM, 16:107–131, 1969.

Bibliography II

G. Rozenberg and A. Salomaa.
Handbook of Formal Languages, volume 1–3.
Springer, Berlin, 1997.

A. P. J. van der Walt.
Random context grammars.
In *Proceedings of Symposium on Formal Languages*, 1970.

G. Zetzsche.
On erasing productions in random context grammars.
In S. Abramsky et al., editor, *Automata, Languages and Programming*, volume 6199 of *LNCS*, pages 175–186. Springer, Berlin, 2010.