Matrix Grammars

Jiří Techet   Tomáš Masopust   Alexander Meduna

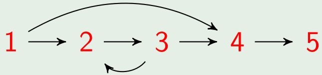
Department of Information Systems
Faculty of Information Technology
Brno University of Technology
Božetěchova 2, Brno 61266, Czech Republic

Modern Formal Language Theory, 2007

Regulated Rewriting

Example

$1 : S \rightarrow AB$
$2 : A \rightarrow aA$
$3 : B \rightarrow bBc$
$4 : A \rightarrow a$
$5 : B \rightarrow bc$



Example of derivation

$$\begin{aligned} S \Rightarrow AB\ [1] &\Rightarrow aAB\ [2] \Rightarrow aAbBc\ [3] \\ &\Rightarrow aaAbBc\ [2] \Rightarrow aaAbbBcc\ [3] \\ &\Rightarrow aaabbBcc\ [4] \Rightarrow aaabbbccc\ [5] \end{aligned}$$

Generated language

$$L(G) = \{a^n b^n c^n \ : \ n \geq 1\}$$

Matrix Grammar

Matrix Grammar

A matrix grammar is a pair

$$H = (G, M),$$

where

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

Notation

- Let $N = \{A_1, \dots, A_m\}$ for some $m \geq 1$
- For some $m_i = p_{i_1} \dots p_{i_j} \dots p_{i_{k_i}} \in M$,

$$p_{i_j} : A_{i_j} \rightarrow x_{i_j}$$

Generated Language

Derivation Step

For $x, y \in (N \cup T)^*$, $m \in M$,

$$x \Rightarrow y\ [m]$$

in H if there are $x_0, \ldots, x_n$ such that $x = x_0$, $x_n = y$, and

1. $x_0 \Rightarrow x_1\ [p_1] \Rightarrow x_2\ [p_2] \Rightarrow\ \ldots\ \Rightarrow x_n\ [p_n]$ in G, and
2. $m = p_1 \ldots p_n$

Generated Language

$$L(H) = \{x \in T^* \,:\, S \Rightarrow^* x\}$$

Example I

Example

$$H = (G, M),$$

where

- $G = (N, T, P, S)$, where

$$\begin{aligned} N &= \{S, A, B\} \\ T &= \{a, b, c\} \\ P &= \{1 : S \rightarrow AB, \\ &\quad\ 2 : A \rightarrow aA, \\ &\quad\ 3 : B \rightarrow bBc, \\ &\quad\ 4 : A \rightarrow a, \\ &\quad\ 5 : B \rightarrow bc\} \end{aligned}$$

- $M = \{1, 23, 45\}$

Example II

Example

1. In G,

$$\begin{aligned} aAbBc &\Rightarrow aaAbBc \quad [2] \\ &\Rightarrow aaAbbBcc\ [3] \end{aligned}$$

As $23 \in M$,

$$aAbBc \Rightarrow aaAbbBcc\ [23]$$

in H

2. In G,

$$\begin{aligned} aAbBc &\Rightarrow aaAbBc\ [2] \\ &\Rightarrow aaAbbcc\ [5] \end{aligned}$$

As $25 \notin M$

$$aAbBc \not\Rightarrow aaAbbcc\ [25]$$

in H

Example III

Example

$S \Rightarrow AB$ [1]	$S \Rightarrow AB$ [1]
	$\Rightarrow aAB$ [2]
$\Rightarrow aAbBc$ [23]	$\Rightarrow aAbBc$ [3]
	$\Rightarrow aaAbBc$ [2]
$\Rightarrow aaAbbBcc$ [23]	$\Rightarrow aaAbbBcc$ [3]
	$\Rightarrow aaabbBcc$ [4]
$\Rightarrow aaabbbccc$ [45]	$\Rightarrow aaabbbccc$ [5]
in H	in G

Example IV

Example

By using $23 \in M$ n-times, $n \geq 0$

$$
\begin{aligned}
S &\Rightarrow AB && [1] \\
&\Rightarrow aAbBc && [23] \\
&\Rightarrow aaAbbBcc && [23] \\
&\quad \vdots \\
&\Rightarrow a^n A b^n B c^n && [23] \\
&\Rightarrow a^{n+1} b^{n+1} c^{n+1} && [45]
\end{aligned}
$$

Generated language

$$L(H) = \{a^m b^m c^m : m \geq 1\}$$

Example V

Example

Claim A

If $AB \Rightarrow^{n} x$, where $n \geq 1$, then $x \in \{a^n b^n c^n, a^n A b^n B c^n\}$.

Proof by Induction on $n \geq 1$

- Basis: $n = 1$.

$$AB \Rightarrow aAbBc$$
$$AB \Rightarrow abc$$

- Induction Hypothesis:
  Assume Claim A holds for all $n = 1, \ldots, k$, where k is a positive integer.

Example VI

Example

Proof by Induction on $n \geq 1$

- Induction Step:

$$AB \Rightarrow^{k+1} x$$

can be rewritten as

$$AB \Rightarrow^{k} y \Rightarrow x$$

By Induction Hypothesis, $y \in \{a^k A b^k B c^k, a^k b^k c^k\}$. As $y \Rightarrow x$, $y = a^k A b^k B c^k$,

$$y \Rightarrow x\ [23] \text{ and } x = a^{k+1} A b^{k+1} B c^{k+1}$$
$$y \Rightarrow x\ [45] \text{ and } x = a^{k+1} b^{k+1} c^{k+1}$$

so $x \in \{a^{k+1} A b^{k+1} B c^{k+1}, a^{k+1} b^{k+1} c^{k+1}\}$ □

Example VII

Example

Claim B

$L(H) \subseteq \{a^k b^k c^k : k \geq 1\}$.

Proof

$L(H) = \{x \in T^* : S \Rightarrow^* x\}$. Every $S \Rightarrow^* x$ with $x \in T^*$ has the form

$$S \Rightarrow AB \Rightarrow^* x$$

From Claim A,

$$L(H) \subseteq \{a^k b^k c^k : k \geq 1\}$$

□

Example VIII

Example

Claim C

For every $a^n b^n c^n$, where $n \geq 1$, $S \Rightarrow^* a^n b^n c^n$.

Proof by Induction on $n \geq 1$

- Basis: $n = 1$, $abc = x$.

$$\begin{aligned} S &\Rightarrow AB \ [1] \\ &\Rightarrow abc \ [45] \end{aligned}$$

- Induction Hypothesis:
  Assume Claim C holds for all $n = 1, \ldots, k$, where k is a positive integer.

Example IX

Example

Proof by Induction on $n \geq 1$

- Induction Step:

$$x = a^{k+1}b^{k+1}c^{k+1}$$

Consider $a^k b^k c^k$. By Induction Hypothesis, $S \Rightarrow^* a^k b^k c^k$. Express this derivation as

$$\begin{aligned} S &\Rightarrow^* a^{k-1}Ab^{k-1}Bc^{k-1} \\ &\Rightarrow a^k b^k c^k \qquad [45] \end{aligned}$$

Then,

$$\begin{aligned} S &\Rightarrow^* a^{k-1}Ab^{k-1}Bc^{k-1} \\ &\Rightarrow a^k Ab^k Bc^k \qquad [23] \\ &\Rightarrow a^{k+1}b^{k+1}c^{k+1} = x \end{aligned}$$

Example X

Example

From Claim B,

$$L(H) \subseteq \{a^k b^k c^k \ : \ k \geq 1\}$$

From Claim C,

$$\{a^k b^k c^k \ : \ k \geq 1\} \subseteq L(H)$$

Thus,

$$L(H) = \{a^k b^k c^k \ : \ k \geq 1\}$$

Matrix Grammar with Appearance Checking

A matrix grammar with appearance checking is a pair

$$H = (G, M)$$

where

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

Derivation Step

Derivation Step

For $x, y \in (N \cup T)^*$, $m = (p_1, q_1) \dots (p_n, q_n) \in M$, $p_i \in P$, $q_i \in \{-, +\}$, $i = 1, \dots, n$,

$$x \Rightarrow y\ [m]$$

in H if there are $x_0, \dots, x_n$ such that $x = x_0$, $y = x_n$, and for $i = 1, \dots, n$

- either $x_{i-1} \Rightarrow x_i\ [p_i]$ in G
- or $q_i = +$, $x_{i-1} = x_i$, and p_i is not applicable to x_{i-1}

Example I

Example

$1: S \to a$

$2: S \to aa$

$3: S \to AB$

$4: A \to A, B \to CC$

$5: A \to A'C, \underline{B \to X}$

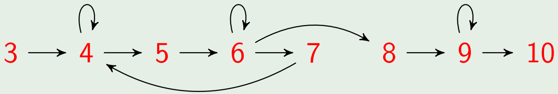
$6: A' \to A', C \to B$

$7: A' \to A, \underline{C \to X}$

$8: A' \to A'', \underline{C \to X}$

$9: A'' \to A'', B \to a$

$10: A'' \to a$



Notes:

Underlined productions are in appearance checking mode $(+)$

X is a "block" symbol

Example II

Example

Derivation example

$$\ldots\ ABBB \Rightarrow ABCCB\ [4] \Rightarrow ACCCCB\ [4] \Rightarrow ACCCCCC\ [4]$$
$$\Rightarrow A'C^7\ [5] \Rightarrow A'CCBCCCC\ [6] \Rightarrow^6 A'B^7\ [6\ldots6] \Rightarrow AB^7\ [7]$$
$$\Rightarrow \ldots$$
$$\Rightarrow A''BBBBBBB\ [8] \Rightarrow^7 A''aaaaaaa\ [9\ldots9] \Rightarrow aaaaaaaa\ [10]$$

The generated language is $L(G) = \{a^{2^i} : i \geq 0\}$

- for $i \geq 2$, the derivation can be expressed as

$$S\ {}_3\!\Rightarrow AB\ {}_{4,5}\!\Rightarrow^* A'C^3\ {}_{6,7}\!\Rightarrow^* AB^3 \Rightarrow^* AC^7$$
$$\vdots$$
$$\Rightarrow A'B^{2^i-1}\ {}_8\!\Rightarrow A''B^{2^i-1}\ {}_{9,10}\!\Rightarrow^* a^{2^i}$$

- for $i = 1, 2$, $S \Rightarrow a\ [1]$ and $S \Rightarrow aa\ [2]$

Bibliography

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.