Scattered Context 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

Scattered Context Grammar

Scattered Context Grammar

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

V is a finite alphabet

T is a set of terminals, $T \subset V$

S is the start symbol, $S \in V - T$

P is a finite set of productions of the form

$$(A_1, \ldots, A_n) \to (x_1, \ldots, x_n),$$

where $A_1, \ldots, A_n \in V - T$, $x_1, \ldots, x_n \in V^*$

Propagating Scattered Context Grammar

- each $(A_1, \ldots, A_n) \to (x_1, \ldots, x_n)$ satisfies $x_1, \ldots, x_n \in V^+$

Derivation Step

Derivation Step

For $(A_1, \dots, A_n) \rightarrow (x_1, \dots, x_n) \in P$ and

$$u = u_1 A_1 \dots u_n A_n u_{n+1}$$
$$v = u_1 x_1 \dots u_n x_n u_{n+1}$$

we write $u \Rightarrow v \, [(A_1, \dots, A_n) \rightarrow (x_1, \dots, x_n)]$

Generated Language

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

Generative Power

- $\mathscr{L}(SC) = \mathscr{L}(RE)$
- $\mathscr{L}(CF) \subset \mathscr{L}(PSC) \subseteq \mathscr{L}(CS)$

Example I

Example

Propagating scattered context grammar

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

with

$$\begin{aligned} P = \{&(S) \rightarrow (ABC), \\ &(A, B, C) \rightarrow (aA, bB, cC), \\ &(A, B, C) \rightarrow (a, b, c)\} \end{aligned}$$

Example of derivation

$$S \Rightarrow ABC \Rightarrow aAbBcC \Rightarrow aaAbbBccC \Rightarrow aaabbbccc$$

Generated language

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

Example II

Example

Propagating scattered context grammar

$$G = (\{S, W, X, Y, Z, A, a\}, \{a\}, P, S),$$

where

$$\begin{aligned} P = \{1 &: (S) \rightarrow (a), \\ 2 &: (S) \rightarrow (aa), \\ 3 &: (S) \rightarrow (WAXY), \\ 4 &: (W, A, X, Y) \rightarrow (a, W, X, AAY), \\ 5 &: (W, X, Y) \rightarrow (a, W, AXY), \\ 6 &: (W, X, Y) \rightarrow (Z, Z, a), \\ 7 &: (Z, A, Z) \rightarrow (Z, a, Z), \\ 8 &: (Z, Z) \rightarrow (a, a)\} \end{aligned}$$

$$L = \{a^{2^n} \,:\, n \geq 0\}$$

$$\begin{aligned} S &\Rightarrow WAXY && [3] \\ &\Rightarrow aWXA^2Y && [4] \\ &\Rightarrow a^2WAA^2XY && [5] \\ &\Rightarrow a^3WAAXA^2Y && [4] \\ &\Rightarrow a^4WAXA^4Y && [4] \\ &\Rightarrow a^5WXA^6Y && [4] \\ &\Rightarrow a^6WA^7XY && [5] \\ &\Rightarrow a^6ZA^7Za && [6] \\ &\Rightarrow^7 a^{13}ZZa && [7^7] \\ &\Rightarrow a^{16} && [8] \end{aligned}$$

Reduction – Definitions

Production length

- $\text{len}((A_1, \ldots, A_n) \to (x_1, \ldots, x_n)) = |A_1 \ldots A_n| = n$

Definitions

nonterminal complexity is the number of nonterminals in G

degree of context-sensitivity $\text{dcs}(G)$ is the number of context-sensitive productions in G

maximum context sensitivity $\text{mcs}(G)$ is the greatest number in

$$\{\text{len}(p_i) - 1 \,:\, 1 \leq i \leq |P|\}$$

overall context sensitivity $\text{ocs}(G)$ is the sum of all members in

$$\{\text{len}(p_i) - 1 \,:\, 1 \leq i \leq |P|\}$$

Reduction – Results I

Lemma

There exists a scattered context grammar G such that G defines a non-context-free language and $\mathrm{dcs}(G) = \mathrm{mcs}(G) = \mathrm{ocs}(G) = 1$.

Proof

Consider a scattered context grammar

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

with

$$\begin{aligned} P = \{&(S) \rightarrow (AC), \\ &(A) \rightarrow (aAbB), \\ &(A) \rightarrow (\varepsilon), \\ &(C) \rightarrow (cCD), \\ &(C) \rightarrow (\varepsilon), \\ &(B, D) \rightarrow (\varepsilon, \varepsilon)\} \end{aligned}$$

$$L(G) = \{a^n b^n c^n : n \geq 0\}$$

$$\mathrm{dcs}(G) = \mathrm{mcs}(G) = \mathrm{ocs}(G) = 1$$

Reduction – Results II

Theorem

There are context-sensitive languages which cannot be described by a scattered context grammar $G = (V, T, P, S)$ *satisfying* $|V - T| = 1$.

Theorem

Every recursively enumerable language is generated by a scattered context grammar $G = (V, T, P, S)$ *satisfying*

$$|V - T| = 3,\ \mathrm{dcs}(G) = \infty,\ \mathrm{mcs}(G) = \infty,\ \mathrm{ocs}(G) = \infty.$$

Theorem

Every recursively enumerable language is generated by a scattered context grammar $G = (V, T, P, S)$ *satisfying*

$$|V - T| = 5,\ \mathrm{dcs}(G) = 2,\ \mathrm{mcs}(G) = 3,\ \mathrm{ocs}(G) = 6.$$

Reduction – Results III

Theorem

Every recursively enumerable language is generated by a scattered context grammar $G = (V, T, P, S)$ satisfying

$$|V - T| = 8,\ \mathrm{dcs}(G) = 6,\ \mathrm{mcs}(G) = 1,\ \mathrm{ocs}(G) = 6.$$

Theorem

Every recursively enumerable language is generated by a scattered context grammar $G = (V, T, P, S)$ satisfying

$$|V - T| = 4,\ \mathrm{dcs}(G) = 4,\ \mathrm{mcs}(G) = 5,\ \mathrm{ocs}(G) = 20.$$

Economical Transformations

Context-Free and Context-Sensitive Productions

For a scattered context production p, if $\text{len}(p)$

$= 1$ then the production is context-free

≥ 2 then the production is context-sensitive

Theorem

Let $H = (M, T, R, S)$ be a phrase-structure grammar in Kuroda normal form. Then, there exists a scattered context grammar, $G = (V, T, P, E)$, that satisfies

1. $L(G) = L(H)$,
2. $|M| = |V| + 5$,
3. *P contains 4 new context productions,*
4. *P contains 1 new context-free production.*

Leftmost Derivations

Leftmost Derivation Step

For $(A_1, \ldots, A_n) \rightarrow (x_1, \ldots, x_n) \in P$ and

$$\begin{aligned} u &= u_1 A_1 \ldots u_n A_n u_{n+1} \\ v &= u_1 x_1 \ldots u_n x_n u_{n+1}, \end{aligned}$$

where $A_i \notin \text{alph}(u_i)$ for all $1 \leq i \leq n$, we write

$$u \; {}_{lm}\!\!\Rightarrow v \; [(A_1, \ldots, A_n) \rightarrow (x_1, \ldots, x_n)]$$

Theorem

Every context-sensitive language can be generated by a propagating scattered context grammar which uses only leftmost derivations.

Extended Propagating Scattered Context Grammars

Extended Propagating Scattered Context Grammar

An extended propagating scattered context grammar is a scattered context grammar

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

in which every

$$(A_1, \ldots, A_n) \to (x_1, \ldots, x_n) \in P$$

satisfies $|x_1 \ldots x_n| \geq n$

Theorem

Every context-sensitive language can be generated by an extended propagating scattered context grammar.

Unordered Scattered Context Grammar

Unordered Scattered Context Grammar

- scattered context grammar in which the order of context-free productions in a scattered context production is unimportant
- for $(A_1, \ldots, A_n) \rightarrow (x_1, \ldots, x_n) \in P$, a permuation $\pi : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$, and

$$u = u_1 A_{\pi(1)} \ldots u_n A_{\pi(n)} u_{n+1}$$
$$v = u_1 x_{\pi(1)} \ldots u_n x_{\pi(n)} u_{n+1}$$

we write $u \Rightarrow v \; [(A_1, \ldots, A_n) \rightarrow (x_1, \ldots, x_n)]$

Generative Power

- $\mathscr{L}(USC) = \mathscr{L}(P, \varepsilon)$
- $\mathscr{L}(PUSC) = \mathscr{L}(P) \subset \mathscr{L}(PSC)$

Open Problems

Open Problem

Are propagating scattered context grammars powerful enough to characterize all context-sensitive languages?

Open Problem

Can every recursively enumerable language be described by a scattered context grammar containing only two nonterminals?

Bibliography I

H. Fernau.
Scattered context grammars with regulation.
Annals of Bucharest University, Mathematics-Informatics Series, 45(1):41–49, 1996.

J. Gonczarowski and M. K. Warmuth.
Scattered versus context-sensitive rewriting.
Acta Informatica, 27:81–95, 1989.

S. Greibach and J. Hopcroft.
Scattered context grammars.
Journal of Computer and System Sciences, 3:233–247, 1969.

T. Masopust.
Regulated Formal Models and Their Reductions.
PhD thesis, Faculty of Information Technology, Brno University of Technology, Brno, Czech Republic, 2007.

Bibliography II

A. Meduna.
Economical transformation of phrase-structure grammars to scattered context grammars.
Acta Cybernetica, 13:225–242, 1998.

A. Meduna.
Generative power of three-nonterminal scattered context grammars.
Theoretical Computer Science, 246:276–284, 2000.

A. Meduna and J. Techet.
Maximal and minimal scattered context rewriting.
In *FCT 2007 Proceedings*, volume 2007, pages 412–423, Budapest, 2007. Springer Verlag.

Bibliography III

J. Techet.
A note on scattered context grammars with non-context-free components.
In *MEMICS 2007 Proceedings*, pages 225–232, Znojmo, 2007. Ing. Zdenek Novotny, CSc.

Gy. Vaszil.
On the descriptional complexity of some rewriting mechanisms regulated by context conditions.
Theoretical Computer Science, 330:361–373, 2005.

V. Virkkunen.
On scattered context grammars.
Acta Universitatis Ouluensis, Series A, Mathematica 6:75–82, 1973.