Parallel Grammars

Alexander Meduna
meduna@dcse.fee.vutbr.cz

Faculty of Information Technologies
Brno University of Technology
Bozetechova 2, Brno 61266, Czech Republic

Introduction

Partially Parallel Grammars

- Scattered Context Grammars
- Multigrammars

Totally Parallel Grammars

- L systems

Conclusion and Discussion

Introduction

Fundamental Ideas:

- grammatical model of parallel computation
- application of several productions during a single derivation step
- increase of generative power

Main Topic: Descriptional Complexity

- reduction of the number of grammatical components
  - reduction of the number of nonterminals
  - reduction of the number of productions
  - economical transformations

Classification of Parallel Grammars:

Partially parallel grammars rewrite **some** symbols while leaving the others unrewritten

- scattered context grammars
- multigrammars

Totally parallel grammars rewrite **all** symbols of the sentential form during a single derivation step

- *L* systems

Scattered Context Grammars (SCGs)

Concept:

- sequences of context-free productions
- several nonterminals are rewritten in parallel

Definition:
Scattered context grammar (SCG)

- $G = (N, T, P, S)$
- N, T, and S as in a CFG
- P is a finite set of productions of the form
  $(A_1, A_2, ..., A_n) \rightarrow (x_1, x_2, ..., x_n)$
  where $A_i \in N$ and $x_i \in V^*$ with $V = N \cup T$

direct derivation:

$u_1A_1u_2A_2\ u_3\ ...\ u_nA_nu_{n+1} \Rightarrow u_1x_1u_2x_2\ u_3\ ...\ u_nx_nu_{n+1}$ if
$(A_1, A_2, ..., A_n) \rightarrow (x_1, x_2, ..., x_n)$

generated language:

$L(G) = \{ w: S \Rightarrow^* w \text{ and } w \in T^* \}$

Example (SCG)

Productions:
$(S) \rightarrow (AA), (A, A) \rightarrow (aA, bAc), (A, A) \rightarrow (\varepsilon, \varepsilon)$

Derivation:
$S \Rightarrow AA \Rightarrow aAbAc \Rightarrow aaAbbAcc \Rightarrow aabbcc$

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

Language Families

CS—the context sensitive languages

RE—the recursively enumerable languages

SC = { $L(G)$: G is a SCG}

for every $n \geq 1$,

***SC*(*n*)** = { $L(G)$: G is a SCG with no more than n nonterminals}

Reduction of the Number of Nonterminals

Theorem 1. $\boldsymbol{RE} = \boldsymbol{SC}(3)$

Theorem 2. $\boldsymbol{CS} \not\subset \boldsymbol{SC}(1)$

Proof (Sketch): Consider $L = \{ a^{2^n} : n \geq 1\}$. Assume that $L = L(G)$, where $G = (\{S\}, \{a\}, P, S)$ is a SCG. In G,

$$S \Rightarrow^* a^i S a^j \Rightarrow^* a^i a^k a^j$$

for some $i, j \geq 0$ such that $i + j, k \geq 1$. Thus,

$$S \Rightarrow^* a^{in} S a^{jn} \Rightarrow^* a^{in} a^k a^{jn}$$

for every n ≥ 0. As $a^i a^k a^j \in L$, $|a^i a^k a^j| = i + k + j = 2^m$. Consider $v = a^{2i} a^k a^{2j} \in L$. Then, $2^m < |v| = 2^m + i + j < 2^{m+1}$, so $v \notin L$—contradiction.

QED

Corollary 3. $\boldsymbol{SC}(1) \subset \boldsymbol{RE}$

Open Problem: $\boldsymbol{RE} = \boldsymbol{SC}(2)$?

Reduction of the Number of Context Productions

Context production:
$(A_1, A_2, ..., A_n) \rightarrow (x_1, x_2, ..., x_n)$ with $n \geq 2$

Theorem 4. Every language in ***RE*** is generated by a scattered context grammar with only these three context productions :

$(\$, 0, 0, \$) \rightarrow (\varepsilon , \$, \$, \varepsilon)$
$(\$, 1, 1, \$) \rightarrow (\varepsilon , \$, \$, \varepsilon)$
$(\$, \$) \rightarrow (\varepsilon , \varepsilon)$

Proof (Basic Idea):

I. Left-Extended Queue Grammar:

$$Q = (V, T, W, F, s, R)$$

R—finite set of productions of the form (a, q, z, r). Every generation of $h \in L(Q)$ has this form

$\#a_0q_0$

$\Rightarrow\ a_0\#x_0q_1 \qquad [(a_0, q_0, z_0, q_1)]$

$\Rightarrow\ a_0a_1\#x_1q_2 \qquad [(a_1, q_1, z_1, q_2)]$

…

$\Rightarrow\ a_0a_1..a_k\# x_kq_{k+1}$

$\Rightarrow\ a_0a_1..a_ka_{k+1}\# x_{k+1}y_1q_{k+2} \qquad [(a_{k+1}, q_{k+1}, y_1, q_{k+2})]$

…

$\Rightarrow\ a_0a_1..a_ka_{k+1}..a_{k+m-1}\# x_{k+m-1}y_1..y_{m-1}q_{k+m}$

$[(a_{k+m-1}, q_{k+m-1}, y_{m-1}, q_{k+m})]$

$\Rightarrow\ a_0a_1..a_ka_{k+1}..a_{k+m}\# y_1..y_mq_{k+m+1}$

$[(a_{k+m}, q_{k+m}, y_m, q_{k+m+1})]$

where $h = y_1..y_m$ with $q_{k+m+1} \in F$

II. Substitutions

For some $n \geq 1$, define an injection, f, from R to $(\{1\}^n(\{0, 1\}^n - \{1\}^n))$.
Extend the definition of f from R^* to $(\{1\}^n(\{0, 1\}^n - \{1\}^n))^*$.

Define the substitution, g, from V^* to $(\{1\}^n(\{0, 1\}^n - 1^n))^*$ as $g(a) = \{ f(r): r = (a, q, z, p) \in R\}$ for every $a \in V$.

Define the substitution, h, from W^* to $(\{0, 1\}^n - 1^n)\{1\}^n)^*$ as $h(q) = \{$ reversal$(f(r)): r = (a, q, z, p) \in R\}$ for every $q \in W$

III. Introduction of SCG:

$G = (N, T, CF \cup Context, S)$

$$Context = \{ (\$, 0, 0, \$) \rightarrow (\varepsilon, \$, \$, \varepsilon),$$
$$(\$, 1, 1, \$) \rightarrow (\varepsilon, \$, \$, \varepsilon),$$
$$(\$, \$) \rightarrow (\varepsilon, \varepsilon)\}$$

IV. *CF* used to generate

$$\$g(a_0a_1..a_ka_{k+1}..a_{k+m})y_1..y_mh(q_{k+m}..q_{k+1}q_k..q_1q_0)\$$

V. *Context* used to verify

$$g(a_0a_1..a_ka_{k+1}..a_{k+m}) = h(q_0q_1..q_kq_{k+1}..q_{k+m})$$

let $g(a_0a_1\ldots \quad a_ka_{k+1}\ldots a_{k+m}) = c_0c_1..c_{(k+m)2n}$

$h(q_0q_1..q_kq_{k+1}..q_{k+m}) = d_0d_1..d_{(k+m)2n}$

where each $c_i, d_i \in \{0, 1\}$

By using (\$, 0, 0, \$) $\rightarrow$(ε , \$, \$, ε),
(\$, 1, 1, \$) $\rightarrow$(ε , \$, \$, ε), and (\$, \$) $\rightarrow$(ε , ε),
G makes

$\$c_0c_1c_2..c_{(k+m)2n}y_1..y_m\ d_{(k+m)2n}..d_2d_1d_0\$$

$\$c_1c_2..c_{(k+m)2n}y_1..y_m\ d_{(k+m)2n}..d_2d_1\$$

$\$c_2..c_{(k+m)2n}y_1..y_m\ d_{(k+m)2n}..d_2\$$

M

$\$y_1..y_m\$$

$y_1..y_m$

QED

Corollary 5. The *SCGs* with three context productions characterize ***RE.***

Economical Transformations

General idea:

transformation of formal models in an economical way

Theorem 6. There exists an algorithm that converts any type-0 grammar, *H*, in *PNF* to an equivalent *SCG* that has the same nonterminals and context-free productions as H and, in addition,

- five new nonterminals
- four new context productions
- one new context-free production

Note on ε-free SCGs

ε-*free SCG* has every production $(A_1, A_2, ..., A_n) \rightarrow (x_1, x_2, ..., x_n)$ with $x_i \neq \varepsilon$ for all $i = 1, 2, ..., n$

ε-free SCGs are less powerful than SCGs because every language they generate is context-sensitive

Open Problem: Do ε-free SCGs characterize ***CS***?

ε-*free* $\boldsymbol{SC}(n)$ = { $L(G)$: G is an ε-free SCG with no more than n nonterminals} for every $n \geq 1$

Characterization of *RE* by operations over ε*-free SC*(5)

Illustration:
Almost identity, ι, is a weak identity that maps one letter to ε and all the other letters to themselves.

Theorem 7. ***RE*** **=** ι**(**ε**-free** *SC***(5))**

Multigrammars

Concept:

- context-free-like productions
- regular selectors consisted of activated and non-activated symbols
- activated symbols are rewritten while non-activated symbols are not

Two types:

- multisequential grammars
- multicontinuous grammars

Multisequential grammars

Definition:

Multisequential grammar

- $G = (N, T, P, S, K)$
- N, T, S same as in a SCG
- P is a finite set of productions of the form
  $a \rightarrow x$, where $a \in V$ and $x \in V^*$ with $V = N \cup T$
- K is a finite set of selectors of the form
  $X_1\textbf{active}(Y_1)\, X_2 \ldots X_n\textbf{active}(Y_n)\, X_{n+1}$
  each $X_i \in \{Z^*: Z \subseteq V\}$
  each $Y_j \in \{Z: Z \subseteq V \text{ and } Z \neq \emptyset\}$

direct derivation:

$u_1a_1u_2a_2\, u_3 \ldots u_na_nu_{n+1} \Rightarrow u_1x_1u_2x_2\, u_3 \ldots u_nx_nu_{n+1}$ **if**

- $X_1\textbf{active}\,(Y_1)\, X_2 \ldots X_n\textbf{active}(Y_n)\, X_{n+1} \in K$
- $u_i \in X_i$ for all $i = 1, \ldots, n+1$
- $a_j \in Y_j$ and $a_j \rightarrow x_j \in P$ for all $j = 1, \ldots, n$

generated language:

$L(G) = \{ w: S \Rightarrow^* w \text{ and } w \in T\}$

Example (Multisequential grammar)

Productions:

$S \rightarrow AA$, $S \rightarrow BB$, $A \rightarrow Aa$, $A \rightarrow Ba$, $B \rightarrow Ab$, $B \rightarrow Bb$, $A \rightarrow \varepsilon$, $B \rightarrow \varepsilon$

Selectors:

active($\{S\}$)
active($\{A\}$)$\{$a, b $\}^*$**active**($\{A\}$)$\{$a, $b\}^*$
active($\{B\}$)$\{$a, b $\}^*$**active**($\{B\}$)$\{$a, $b\}^*$

Derivation:

$S \Rightarrow BB \Rightarrow AbAb \Rightarrow AabAab \Rightarrow abab$

Generated Language:

$L(G) = \{ ww: w \in \{a, b\}^* \}$

Theorem 8. The six-nonterminal multisequential grammars characterize ***RE***.

Multicontinues grammars

Concept:
Analogue to multisequetial grammars except that selectors have the form

X_1**active**(Y_1) X_2... X_n**active**(Y_n) X_{n+1}
each $X_i \in \quad \{Z^*: Z \subseteq V\}$
each $Y_j \in \quad \{Z^+: Z \subseteq V$ and $Z \neq \varnothing\}$

Example (Multicontinuous grammar)

Productions:

$S \rightarrow bcb$, $b \rightarrow bb$, $c \rightarrow cc$,

Selectors:

active($\{S\}$)

active($\{b\}^+$)$\{c\}^*$**active**($\{b\}^+$)

active($\{b\}^+$)**active**($\{c\}^+$)**active**($\{b\}^+$)

Derivation:

$S \Rightarrow bcb \Rightarrow bbcbb \Rightarrow bbbbccbbbb$

Generated Language:

$L(G) = \{ b^j c^k b^j : j = 2^i \text{ and } k = 2^r \text{ so that } i \geq r \geq 0\}$

Theorem 9. The six-nonterminal multicontinuous grammars characterize ***RE***.

Totally Parallel Grammars: *EIL* systems

Concept:

- productions of the form
  $(x, a, y) \rightarrow z$

- every symbol is rewritten during a derivation step

Example: $E(1, 0)L$ system

Axiom: ***abc***

Productions:
$(a, b, 0) \rightarrow (abb)$, $(b, c, 0) \rightarrow (cc)$

(in all other cases, the symbols are changed to themselves)

$(a, a, 0) \rightarrow (a)$, $(0, a, 0) \rightarrow (a)$,
$(b, b, 0) \rightarrow (b)$, $(c, c, 0) \rightarrow (c)$ …

Derivation:
$abc \Rightarrow aabbcc \Rightarrow aaabbbccc \Rightarrow \ldots$

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

Theorem 10. The three-nonterminal *EIL* grammars characterize ***RE***.

Open Problem Areas

I. Can we improve any of the above characterizations of ***RE***?

Specific questions:

- Do the two-nonterminal SCGs characterize ***RE***?
- Do the two-context-production SCGs characterize ***RE***?

II. Can ***RE*** be characterized by scattered context grammars with a reduced number of both nonterminals and context productions?

III. What is the descriptional complexity of other partially parallel grammars?