Parsing Beyond Context-Free Grammars
Abschlussklausur

17.7.2013

Laura Kallmeyer, Patrick Hommers

SS 2013, Heinrich-Heine-Universität Düsseldorf

Klausurdauer: 90 Minuten.

Hilfmittel: Sämtliche Unterrichtsmaterialien und Notizen in nicht-elektronischer Form.

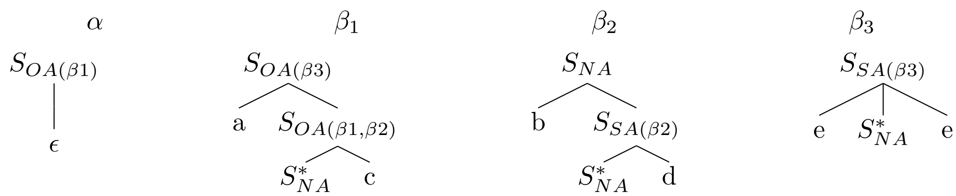
Aufgabe 1

Geben Sie für die Sprache

$$L = \{e^k a^n b^m c^n d^m e^k\}$$

eine TAG an, die diese Sprache generiert. Verwenden Sie dabei die adjunction constraints und geben Sie in den adjunction constraints die Menge der Elementarbäume an, für die diese Beschränkungen gelten (z.B. $OA_{\beta 1}$ an einem Knoten, an dem Adjunktion von $\beta 1$ obligatorisch ist).

Lösung:



Aufgabe 2 *Betrachten Sie folgende TAG:*

Geben Sie die trace des Early Parsers für die Eingabe $w = aaa$ an. Geben Sie nur die Items an, die zum korrekten parse führen. Erklären Sie für jedes Item aus welcher Regel es generiert wurde und welches Antezedens dafür verwendet wurde (die Items werden über die Nummern in der ersten Spalte identifiziert).

Id	*Item*	*Operation*	*Antezedens-Items*
1	$[\alpha, \varepsilon, la, 0, -, -, 0, 0]$	*Initialize*	–
2	$[\beta, \varepsilon, la, 0, -, -, 0, 0]$	...	*1*
3	$[\beta, \varepsilon, lb, 0, -, -, 0, 0]$	...	*2*
4	$[\beta, 1, la, 0, -, -, 0, 0]$	...	*3*

Lösung:

Id	Item	Operation	Antezedens-Items
1	$[\alpha, \epsilon, la, 0, -, -, 0, 0]$	Initialize	–
2	...	...	...

still pending

Aufgabe 3

Geben Sie für die Sprache

$$L = \{e^m b^n a^{n+1} g c^{n+1} d^n f^m | n, m \geq 0\}$$

eine LCFRS an, die diese Sprache generiert.

Lösung:

$A(a, c) \to \epsilon$, $B(\varepsilon, \varepsilon) \to \epsilon$, $B(e, f) \to \epsilon$

$A(bXa, cYd) \to (X, Y)$, $B(eX, fY) \to B(X, Y)$

$C(g) \to \epsilon$

$S(X, V, W, Z, Y) \to B(V.Z)A(X, Y)C(W)$

Aufgabe 4

Betrachten Sie die LCFRS $G = \langle\{A, S, B, C\}, \{a, b, c, d, e, f\}, \{X, Y, Z, U\}, S, P\rangle$, wobei P folgende Regeln enthält:

$S(XVZWYf) \to A(X, Y)C(V, W)B(Z)$

$A(a, b) \to \epsilon$	$B(\varepsilon) \to \epsilon$
$A(aX, bY) \to (X, Y)$	$B(cd) \to \epsilon$
$C(c, e) \to \epsilon$	$B(cdX) \to X$

1. *Geben Sie die Sprache L an, die von dieser Grammatik erzeugt wird.*
2. *Geben Sie eine äquivalente binäre LCFRS an.*

Lösung:

1. $L = \{a^n c(cd)^m e b^n f | n \geq 1, m \geq 0\}$
2. Die erste Regel wird durch $S(XVYf) \to A(X, Y)D(V)$ und $D(VZW) \to C(V, W)B(Z)$ ersetzt.

Aufgabe 5 *Geben Sie für die LCFRS $G = \langle\{A, S, B\}, \{a, b, c\}, \{X, Y, Z\}, S, P\rangle$ mit den Regeln*

$S(XYZ) \to A(X, Y, Z)$	$A(aX, bY, cZ) \to A(X, Y, Z)$	$B(aX, cY) \to B(X, Y)$
$S(XbbZ) \to B(X, Z)$	$A(a, b, c) \to \epsilon$	$B(a, c) \to \epsilon$

alle Items an, die sich für das CYK-LCFRS-Parsing mit dem Input $w = aabbcc$ ergeben. Bei jedem Item soll angegeben werden, mit welcher Regel es aus welchen Antezedensitems hergeleitet wurde (die Items werden über die Nummern in der ersten Spalte identifiziert).

Id	*Item*	*Operation*	*Antezedens-Items*
1	$[A, \langle\langle 0, 1\rangle\langle 2, 3\rangle\langle 4, 5\rangle\rangle]$	*Axiom*	–
2	$[B, \langle\langle 0, 1\rangle\langle 4, 5\rangle\rangle]$	*Axiom*	–

Lösung:

Id	Item	Operation	Antezedens-Items
1	$[A, \langle\langle 0,1\rangle\langle 2,3\rangle\langle 4,5\rangle\rangle]$	Axiom	–
2	$[B, \langle\langle 0,1\rangle\langle 4,5\rangle\rangle]$	Axiom	–
3	$[A, \langle\langle 0,1\rangle\langle 2,3\rangle\langle 5,6\rangle\rangle]$	Axiom	–
4	$[A, \langle\langle 0,1\rangle\langle 3,4\rangle\langle 5,6\rangle\rangle]$	Axiom	–
5	$[A, \langle\langle 1,2\rangle\langle 3,4\rangle\langle 5,6\rangle\rangle]$	Axiom	–
6	$[A, \langle\langle 1,2\rangle\langle 2,3\rangle\langle 5,6\rangle\rangle]$	Axiom	–
7	$[A, \langle\langle 1,2\rangle\langle 2,3\rangle\langle 4,5\rangle\rangle]$	Axiom	–
8	$[B, \langle\langle 0,1\rangle\langle 5,6\rangle\rangle]$	Axiom	–
9	$[B, \langle\langle 1,2\rangle\langle 4,5\rangle\rangle]$	Axiom	–
10	$[B, \langle\langle 1,2\rangle\langle 5,6\rangle\rangle]$	Axiom	–
11	$[B, \langle\langle 0,2\rangle\langle 4,6\rangle\rangle]$	Complete	aus 10
12	$[S, \langle\langle 0,6\rangle\rangle]$	Complete	aus 11
13	$[A, \langle\langle 0,2\rangle\langle 2,4\rangle\langle 4,6\rangle\rangle]$	Complete	aus 5
14	$[S, \langle\langle 0,6\rangle\rangle]$	Complete	aus 13

Aufgabe 6

1. *Wir haben schon gezeigt, dass $L = \{a^n b^n c^n d^n e^n \mid n \geq 0\}$ keine Tree Adjoining Language ist.*
   Zeigen Sie, dass auch $L' = \{a^{2n} c b^n (gh)^n d^n h^n \mid n \geq 0\}$ keine Tree Adjoining Language ist.
2. *Wir haben schon gezeigt, dass $L = \{(a^m b^m)^5 \mid m \geq 1\}$ keine 4-MCFL ist.*
   Zeigen Sie, dass auch $L' = \{w^5 \mid w \in \{a, b\}\}$ keine 4-MCFL ist.

Tipp: Nutzen Sie die Abgeschlossenheitseigenschaften der jeweiligen Sprachklassen aus.

Lösung:

1. Homomorphismus f mit $f(aa) \to a$, $f(c) \to \varepsilon$, $f(b) \to b$, $f(gh) \to c$, $f(d) \to d$, $f(h) \to e$ bildet L' auf L ab. D.h., wenn L' eine TAL ist, dann muss auch L eine TAL sein. Widerspruch.
2. $L' \cap L(a^+b^+a^+b^+a^+b^+a^+b^+a^+b^+) = L$. . D.h., wenn L' eine 4-MCFL ist, dann muss auch L eine sein. Widerspruch.