

Parsing Beyond CFG

Homework 8: LCFRS — CYK and Earley Parsing

Laura Kallmeyer, Tatiana Bladier

Sommersemester 2018

Question 1 (CYK parsing)

Consider the following LCFRS G in sRCG format and the input word **aabcbaaaac**.

$$G = \langle \{S, A, B, C, D\}, \{a, b, c\}, \{T, U, V, W, X, Y, Z\}, P, S \rangle$$

where

$$P = \left\{ \begin{array}{l} S(UTVZ) \rightarrow A(U, V)D(T, Z), \\ D(WYX, Z) \rightarrow B(W, X)C(Y, Z), \\ A(a, aa) \rightarrow \epsilon, \\ A(aX, aaY) \rightarrow A(X, Y), \\ B(b, b) \rightarrow \epsilon, \\ B(bX, bY) \rightarrow B(X, Y), \\ C(c, c) \rightarrow \epsilon, \\ C(cX, cY) \rightarrow C(X, Y) \end{array} \right\}$$

Give the **complete** CYK trace considering that G is **ordered**. The first lines of the trace are already provided below:

	Item	Rule
1	$[A, \langle \langle 0, 1 \rangle, \langle 5, 7 \rangle \rangle]$	scan
2	$[A, \langle \langle 0, 1 \rangle, \langle 6, 8 \rangle \rangle]$	scan
3	$[A, \langle \langle 0, 1 \rangle, \langle 7, 9 \rangle \rangle]$	scan

Solution:

	Item	Rule
1	$[A, \langle \langle 0, 1 \rangle, \langle 5, 7 \rangle \rangle]$	scan
2	$[A, \langle \langle 0, 1 \rangle, \langle 6, 8 \rangle \rangle]$	scan
3	$[A, \langle \langle 0, 1 \rangle, \langle 7, 9 \rangle \rangle]$	scan
4	$[A, \langle \langle 1, 2 \rangle, \langle 5, 7 \rangle \rangle]$	scan
5	$[A, \langle \langle 1, 2 \rangle, \langle 6, 8 \rangle \rangle]$	scan
6	$[A, \langle \langle 1, 2 \rangle, \langle 7, 9 \rangle \rangle]$	scan
7	$[A, \langle \langle 5, 6 \rangle, \langle 6, 8 \rangle \rangle]$	scan
8	$[A, \langle \langle 5, 6 \rangle, \langle 7, 9 \rangle \rangle]$	scan
9	$[A, \langle \langle 6, 7 \rangle, \langle 7, 9 \rangle \rangle]$	scan
10	$[B, \langle \langle 2, 3 \rangle, \langle 4, 5 \rangle \rangle]$	scan
11	$[C, \langle \langle 3, 4 \rangle, \langle 9, 10 \rangle \rangle]$	scan
12	$[A, \langle \langle 0, 2 \rangle, \langle 5, 9 \rangle \rangle]$	complete with 1 and 6
13	$[D, \langle \langle 2, 5 \rangle, \langle 9, 10 \rangle \rangle]$	complete with 10 and 11
14	$[S, \langle \langle 0, 10 \rangle \rangle]$	complete with 12 and 13

Question 2 (Earley parsing)

Consider the following LCFRS G in sRCG format and the input word **abcbaac**.

$$G = \langle \{S, A, B, C, D\}, \{a, b, c\}, \{T, U, V, W, X, Y, Z\}, P, S \rangle$$

where

$$P = \left\{ \begin{array}{l} S(UTVZ) \rightarrow A(U, V)D(T, Z), \\ D(WYX, Z) \rightarrow B(W, X)C(Y, Z), \\ A(a, aa) \rightarrow \epsilon, \\ A(aX, aaY) \rightarrow A(X, Y), \\ B(b, b) \rightarrow \epsilon, \\ C(c, c) \rightarrow \epsilon \end{array} \right\}$$

Give the Earley trace for this LCFRS. The first lines of the trace are already provided below:

	pos.	item	$\vec{\rho}$ (bindings)	rule
1	0	$S(\bullet UTVZ) \rightarrow A(U, V)D(T, Z)$	$(?, ?, ?, ?)$	axiom
2	0	$A(\bullet a, aa) \rightarrow \epsilon$	$(?, ?, ?)$	predict, 1
3	0	$A(\bullet aX, aaY) \rightarrow A(X, Y)$	$(?, ?, ?, ?, ?)$	predict, 1

Solution:

	pos.	item and dot position	$\vec{\rho}$ (bindings)	rule
1	0	$S(\bullet UTVZ) \rightarrow A(U, V)D(T, Z)$	$(?, ?, ?, ?)$	axiom
2	0	$A(\bullet a, aa) \rightarrow \epsilon$	$(?, ?, ?)$	predict, 1
3	0	$A(\bullet aX, aaY) \rightarrow A(X, Y)$	$(?, ?, ?, ?, ?)$	predict, 1
4	1	$A(a\bullet, aa) \rightarrow \epsilon$	$(\langle 0, 1 \rangle, ?, ?)$	scan, 2
5	1	$A(a\bullet X, aaY) \rightarrow A(X, Y)$	$(\langle 0, 1 \rangle, ?, ?, ?, ?)$	scan, 3
6	1	$S(U\bullet TVZ) \rightarrow A(U, V)D(T, Z)$	$(\langle 0, 1 \rangle, ?, ?, ?)$	suspend, 4 and 1
7	1	$A(\bullet a, aa) \rightarrow \epsilon$	$(?, ?, ?)$	predict, 5
8	1	$A(\bullet aX, aaY) \rightarrow A(X, Y)$	$(?, ?, ?, ?, ?)$	predict, 5
9	1	$D(\bullet WYX, Z) \rightarrow B(W, X)C(Y, Z)$	$(?, ?, ?, ?)$	predict, 6
10	1	$B(\bullet b, b) \rightarrow \epsilon$	$(?, ?)$	predict, 9
11	2	$B(b\bullet, b) \rightarrow \epsilon$	$(\langle 1, 2 \rangle, ?)$	scan, 10
12	2	$D(W\bullet YX, Z) \rightarrow B(W, X)C(Y, Z)$	$(\langle 1, 2 \rangle, ?, ?, ?)$	suspend, 11 and 9
13	2	$C(\bullet c, c) \rightarrow \epsilon$	$(?, ?)$	predict, 12
14	3	$C(c\bullet, c) \rightarrow \epsilon$	$(\langle 2, 3 \rangle, ?)$	scan, 13
15	3	$D(WY\bullet X, Z) \rightarrow B(W, X)C(Y, Z)$	$(\langle 1, 2 \rangle, \langle 2, 3 \rangle, ?, ?)$	suspend, 14 and 12
16	3	$B(b, \bullet b) \rightarrow \epsilon$	$(\langle 1, 2 \rangle, ?)$	resume, 15 and 11
17	4	$B(b, b\bullet) \rightarrow \epsilon$	$(\langle 1, 2 \rangle, \langle 3, 4 \rangle)$	scan, 16
18	4	$B(\langle 1, 2 \rangle, \langle 3, 4 \rangle)$		convert, 17
19	4	$D(WYX\bullet, Z) \rightarrow B(W, X)C(Y, Z)$	$(\langle 1, 2 \rangle, \langle 2, 3 \rangle, \langle 3, 4 \rangle, ?)$	complete, 18 and 15
20	4	$S(UT\bullet VZ) \rightarrow A(U, V)D(T, Z)$	$(\langle 0, 1 \rangle, \langle 1, 4 \rangle, ?, ?)$	suspend, 19 and 6
21	4	$A(a, \bullet aa) \rightarrow \epsilon$	$(\langle 0, 1 \rangle, ?, ?)$	resume, 20 and 4
22	5	$A(a, a\bullet a) \rightarrow \epsilon$	$(\langle 0, 1 \rangle, \langle 4, 5 \rangle, ?)$	scan, 21
23	6	$A(a, aa\bullet) \rightarrow \epsilon$	$(\langle 0, 1 \rangle, \langle 4, 5 \rangle, \langle 5, 6 \rangle)$	scan, 22
24	6	$A(\langle 0, 1 \rangle, \langle 4, 6 \rangle)$		convert, 23

	pos.	item and dot position	$\vec{\rho}$ (bindings)	rule
25	6	$S(UTV \bullet Z) \longrightarrow A(U, V)D(T, Z)$	$(\langle 0, 1 \rangle, \langle 1, 4 \rangle, \langle 4, 6 \rangle, ?)$	complete, 25 and 20
26	6	$D(WYX, \bullet Z) \longrightarrow B(W, X)C(Y, Z)$	$(\langle 1, 2 \rangle, \langle 2, 3 \rangle, \langle 3, 4 \rangle, ?)$	resume, 25 and 19
27	6	$C(c, \bullet c) \longrightarrow \varepsilon$	$(\langle 2, 3 \rangle, ?)$	resume, 26 and 14
28	7	$C(c, c \bullet) \longrightarrow \varepsilon$	$(\langle 2, 3 \rangle, \langle 6, 7 \rangle)$	scan, 27
29	7	$C(\langle 2, 3 \rangle, \langle 6, 7 \rangle)$		convert, 28
30	7	$D(WYX, Z \bullet) \longrightarrow B(W, X)C(Y, Z)$	$(\langle 1, 4 \rangle, \langle 6, 7 \rangle)$	complete, 29 and 26
31	7	$D(\langle 1, 4 \rangle, \langle 6, 7 \rangle)$		convert, 30
32	7	$S(UTVZ \bullet) \longrightarrow A(U, V)D(T, Z)$	$(\langle 0, 1 \rangle, \langle 1, 4 \rangle, \langle 4, 6 \rangle, \langle 6, 7 \rangle)$	complete, 31 and 25
33	7	$S(\langle 0, 7 \rangle)$		convert, 32