

# Parsing Beyond CFG

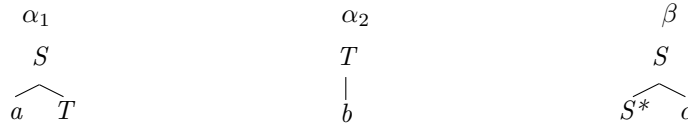
## Homework 5: TAG Earley Parsing, Abgabe 29.05.2013

Laura Kallmeyer, Patrick Hommers

SS 2013, Heinrich-Heine-Universität Düsseldorf


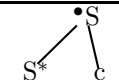
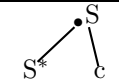
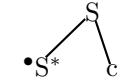
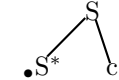
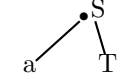
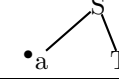
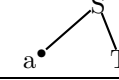
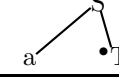
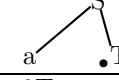
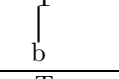
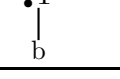
### Question 1 (TAG Earley parsing)

Consider again the TAG consisting of the two trees  $\alpha$  and  $\beta$ :



Give the trace of the Earley parse (the version from the course slides) of  $w = abc$ . Give only those items that lead to the correct parse. Explain for each of these items, by which operation it is obtained and from which antecedent item(s).

Solution:

	Item	dotted tree	rule
1.	$[\alpha_1, \epsilon, la, 0, -, -, 0, 0]$		Initialize
2.	$[\beta, \epsilon, la, 0, -, -, 0, 0]$		PredictAdjoinable
3.	$[\beta, \epsilon, lb, 0, -, -, 0, 0]$		PredictNoAdj
4.	$[\beta, 1, la, 0, -, -, 0, 0]$		MoveDown
5.	$[\beta, 1, lb, 0, -, -, 0, 0]$		PredictNoAdj
6.	$[\alpha_1, \epsilon, lb, 0, -, -, 0, 0]$		PredictAdjoined
7.	$[\alpha_1, 1, la, 0, -, -, 0, 0]$		MoveDown
8.	$[\alpha_1, 1, ra, 0, -, -, 1, 0]$		ScanTerm
9.	$[\alpha_1, 2, la, 0, -, -, 1, 0]$		MoveRight
10.	$[\alpha_1, 2, lb, 1, -, -, 1, 0]$		PredictNoAdj
11.	$[\alpha_2, \epsilon, la, 1, -, -, 1, 0]$		PredictSubs
12.	$[\alpha_2, \epsilon, lb, 1, -, -, 1, 0]$		PredictNoAdj

13.	$[\alpha_2, 1, la, 1, -, -, 1, 0]$	$\begin{array}{c} T \\ \bullet \\ b \end{array}$	MoveDown
14.	$[\alpha_2, 1, ra, 1, -, -, 2, 0]$	$\begin{array}{c} T \\   \\ b \bullet \end{array}$	ScanTerm
15.	$[\alpha_2, \epsilon, rb, 1, -, -, 2, 0]$	$\begin{array}{c} T \bullet \\   \\ b \end{array}$	MoveUp
16.	$[\alpha_2, \epsilon, ra, 1, -, -, 2, 0]$	$\begin{array}{c} T \bullet \\   \\ b \end{array}$	CompleteNode
17.	$[\alpha_1, 2, rb, 1, -, -, 2, 0]$	$\begin{array}{c} S \\ / \quad \backslash \\ a \quad T \bullet \end{array}$	Substitute 16. and 10.
18.	$[\alpha_1, 2, ra, 0, -, -, 2, 0]$	$\begin{array}{c} S \\ / \quad \backslash \\ a \quad T \bullet \end{array}$	CompleteNode
19.	$[\alpha_1, \epsilon, rb, 0, -, -, 2, 0]$	$\begin{array}{c} S \bullet \\ / \quad \backslash \\ a \quad T \end{array}$	MoveUp
20.	$[\beta, 1, rb, 0, 0, 2, 2, 0]$	$\begin{array}{c} S \\ / \quad \backslash \\ S \bullet^* \quad c \end{array}$	CompleteNode
21.	$[\beta, 1, ra, 0, 0, 2, 2, 0]$	$\begin{array}{c} S \\ / \quad \backslash \\ S \bullet^* \quad c \end{array}$	CompleteNode
22.	$[\beta, 2, la, 0, 0, 2, 2, 0]$	$\begin{array}{c} S \\ / \quad \backslash \\ S^* \quad c \bullet \end{array}$	MoveRight
23.	$[\beta, 2, ra, 0, 0, 2, 3, 0]$	$\begin{array}{c} S \\ / \quad \backslash \\ S^* \quad c \bullet \end{array}$	ScanTerm
24.	$[\beta, \epsilon, rb, 0, 0, 2, 3, 0]$	$\begin{array}{c} S \bullet \\ / \quad \backslash \\ S^* \quad c \end{array}$	MoveUp
25.	$[\beta, \epsilon, ra, 0, 0, 2, 3, 0]$	$\begin{array}{c} S \bullet \\ / \quad \backslash \\ S^* \quad c \end{array}$	CompleteNode
26.	$[\alpha_1, \epsilon, rb, 0, -, -, 3, 1]$	$\begin{array}{c} S \bullet \\ / \quad \backslash \\ a \quad T \end{array}$	Adjoin 25. and 19.
27.	$[\alpha_1, \epsilon, ra, 0, -, -, 3, 0]$	$\begin{array}{c} S \bullet \\ / \quad \backslash \\ a \quad T \end{array}$	CompleteNode

**Question 2** Consider the rule **Substitute** of the Earley TAG parser from the course slides. Give an alternative rule **CompleteSubstitution** that moves also from the tree  $\alpha$  back into  $\gamma$  and that checks, in addition, that there exists a  $\gamma$  item that has triggered the prediction of the substitution of  $\alpha$ .

Solution:

**CompleteSubstitute:**  $\frac{[\gamma, p, lb, i, -, -, i, 0], [\alpha, \epsilon, ra, i, -, -, j, 0]}{[\gamma, p, rb, i, -, -, j, 0]}$   $\gamma(p)$  a substitution node,  
 $\alpha \in I, l(\gamma, p) = l(\alpha, \epsilon)$

**Question 3** What is the time complexity (in the length  $n$  of the input word) of the Earley TAG parser given in the course slides?

Solution:

The rules have at most 6 different indices ranging from 0 to  $n$ . Consequently, the complexity is  $\mathcal{O}(n^6)$ .