

Prof. Dr. Andreas Hotho,
M.Sc. Janna Omeliyanenko
Lecture Chair X for Data Science, Universität Würzburg

7. Exercise for “Sprachverarbeitung und Text Mining”

17.12.2021

1 Knowledge Questions

1. State the main characteristics in which the CKY and Earley parser differ from each other.
2. Briefly describe the characteristic procedures of the Earley Parser.
3. Name and describe the data structure used in the lecture to represent the dependencies of a sentence.

What are the three important properties of this structure?

4. What is the role of an oracle in Shift-Reduce dependency parsing?
How is an oracle provided for this task?

2 CKY

Use the given grammar to parse the following sentence with the CKY Algorithm:
Much to learn you still have.
You don't have to calculate probabilities in this exercise, since we focus on how to apply the rules here.

$ADV \rightarrow Much$
 $APPR \rightarrow to$
 $V \rightarrow learn$
 $PRO \rightarrow you$
 $ADV \rightarrow still$
 $V \rightarrow have$
 $S \rightarrow X_0 VP$
 $X_0 \rightarrow VP PRO$
 $VP \rightarrow APPR V$
 $VP \rightarrow ADV V$
 $VP \rightarrow ADV VP$

3 Early Parsing

Check if the sentence below can be parsed with the given grammar by doing the “forward parse” of the Earley Algorithm. In other words: write down all state sets as seen in the lecture and check if the last one contains an item that satisfies the conditions for a successful parse.

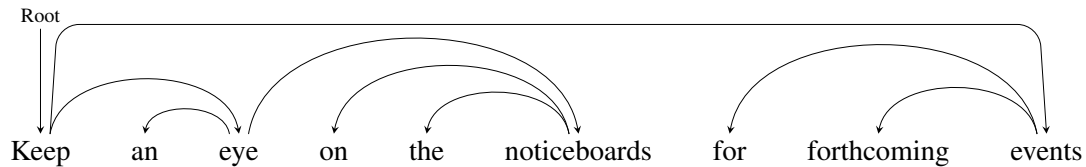
Teacher and students philosophize

The following grammar is used:

$S \rightarrow NP VP$
 $NP \rightarrow NN KONJ NN$
 $VP \rightarrow V$
 $NN \rightarrow students$
 $KONJ \rightarrow and$
 $NN \rightarrow Teacher$
 $V \rightarrow philosophize$

4 Shift-Reduce Parsing

Given is the following dependency tree:



In order to learn the state transitions of Shift-Reduce parsing, data labelled through dependency trees must be converted into Shift-Reduce state transitions.

Perform this conversion by decomposing the graph into a sequence of states during Shift-Reduce parsing. Initially, root is on the stack and the input buffer contains the entire set.

For each state, specify the contents of the stack, the input buffer and the selected operation in tabular form (as shown below). The column *Operation* contains the action that will be performed to move from the current line to the next - so start by adding the first operation.

Stack	Buffer	Operation
root	Keep, an, eye, on, the, noticeboards, for, forthcoming, events	???