

# Parsing

## Homework 12 (A\* parsing), due 12 July 2021

Laura Kallmeyer

SS 2021, Heinrich-Heine-Universität Düsseldorf

### Question 1 (A\* parsing inside and outside viterbi estimates)

Consider the PCFG  $G$  with  $N = \{S, A\}$ ,  $T = \{a\}$ , start symbol  $S$  and productions

$$\begin{array}{lll} 0.5 & S \rightarrow SS & 0.125 & S \rightarrow AS & 0.25 & S \rightarrow SA \\ 0.125 & S \rightarrow a & 1 & A \rightarrow a & & \end{array}$$

For weights, use  $|\log_2(p)|$ .

Compute the inside viterbi estimates for lengths  $1 \leq l \leq 4$  and the outside  $SX$  estimates for length  $n = 4$ .

### Question 2 (A\* parsing)

Consider the PCFG given in the example on slides 14 (A\* slides) and the outside scores computed on the subsequent slides.

As input consider “nice red house bike”.

1. Show the weighted deductive CYK-Parsing with chart and agenda using this grammar and input with weights as described on slide 18 (incorporating the viterbi inside score and the  $SX$  outside estimate).

Write each weight as a pair  $(in, out)$  where  $in$  is the inside viterbi score and  $out$  the outside estimate (using  $|\log(p)|$  instead of  $p$ ).

Concerning the chart column, it is enough to list only new items in each row. (This is different from the agenda where items are not only added but also removed and reordering depending on weights takes place.)

2. The log used here is  $\log_{10}$ . Compute the probability of the best parse tree from the weight of the goal item.