Parsing

Homework 4 (Top-Down Parsing), due 03 May 2022

Laura Kallmeyer

SS 2022, Heinrich-Heine-Universität Düsseldorf

Question 1 (Unger recognition) Consider the CFG G with $N = \{S, A, B\}, T = \{a, b\}$, start symbol S and productions

$$S \to aAB \mid AB \quad A \to a \mid Aa \quad B \to b$$

and the input w = aab.

How does the resulting chart like when we do Unger recognition with tabulation along the lines of the example on slide 18, assuming that (i) we make a prediction of a span of length 1 for a terminal a in a righthand side only if the terminal in the input that this span consists of is actually a and (ii) we only predict spans of length ≥ 1 .

Question 2 (Top-Down parsing) Consider the CFG G with $N = \{S, X\}, T = \{a\}$, start symbol S and productions

$$S \to aSa \,|\, X, X \to aX \,|\, \varepsilon$$

and the input w = aa.

- 1. Give all parse trees for this input.
- 2. Is G ambiguous?
- 3. You can get all the 24 pairs of prediction stack, remaining input that arise in a top-down parsing for this input by performing the parsing using the CL-Taskbox linked on the course web page. Which of the 24 items are the successful ones for the parses of this input?

Note that a pair of stack and remaining input is notated as $[\Gamma, i]$ where Γ is the stack and i is the position in the input up to which the input has already been processed.

- 4. Give the resulting leftmost derivations.
- 5. Does a filter that avoids generating pairs where the stack is longer than the remaining input make sense, given this grammar?