

Einführung in die statistische Sprachverarbeitung

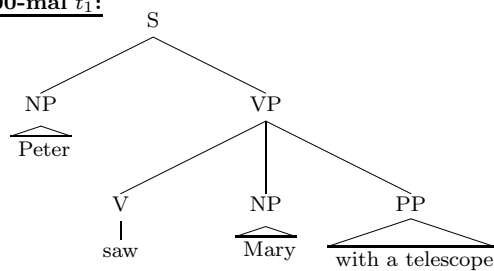
Detlef Prescher

May 15, 2007

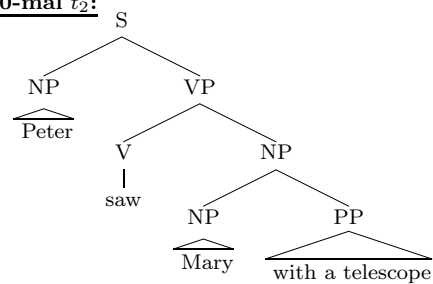
Hausaufgabe 6

(i) Es sei die folgende Baumbank gegeben:

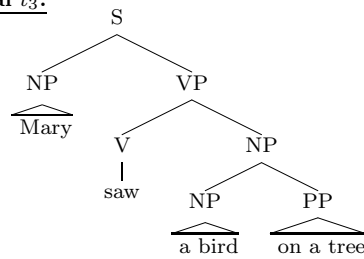
200-mal t_1 :



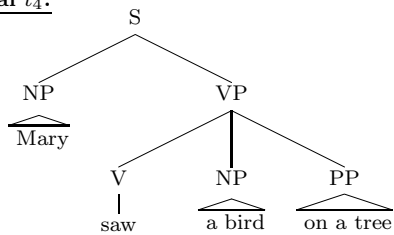
10-mal t_2 :



200-mal t_3 :



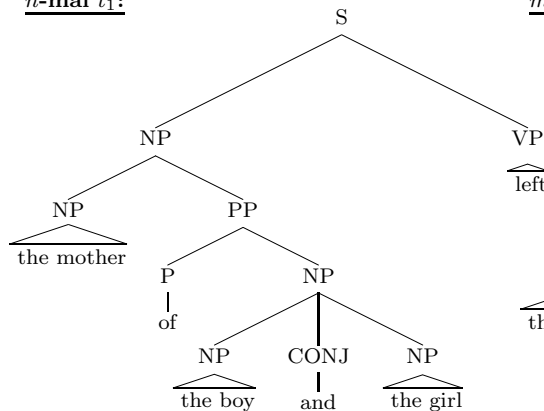
5-mal t_4 :



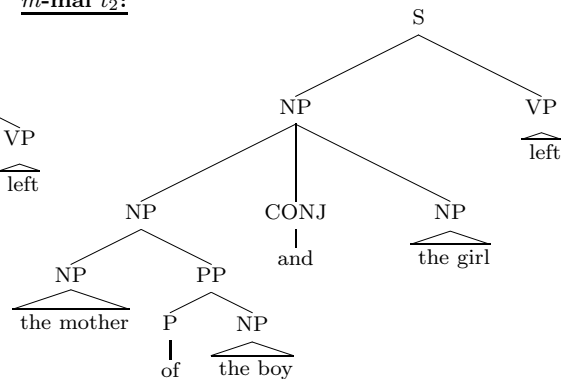
Lexikalisiere diese Baumbank (nur die PP-Regeln), lies die Baumbankgrammatik ab, und zeige, dass sie gute Disambiguierungsergebnisse liefert.

(ii) Es sei die folgende Baumbank gegeben:

n -mal t_1 :



m -mal t_2 :



Lies die Baumbankgrammatik von dieser Baumbank ab. Für welche Werte von n und m wird der Baum t_1 dem Baum t_2 vorgezogen? Führe Parent-Encoding durch und beantworte die Frage nochmals!

(iii)-(v) Gib drei weitere Ambiguitätstypen für kontextfreie Grammatiken an, und diskutiere die Frage, ob der jeweilige Ambiguitätstyp durch eine probabilistische Grammatik auflösbar ist. Wenn nein, diskutiere die Frage, ob dies eventuell nach einer Grammatiktransformation möglich ist.

Abgabetermin: Dienstag, 29. Mai (früh morgens!!!)

Besprechung: Mittwoch, 30. Mai