

# Statistisches Parsing

Detlef Prescher

13. Juli 2007

## Abschlussklausur (Montag, 16. Juli 2007)

Für die Klausur sind fünf der sechs Aufgaben zu lösen. Wird für jede der sechs Aufgaben ein Lösungsversuch abgegeben, so wird der Lösungsversuch der Aufgabe (vi) nicht gewertet. Viel Erfolg!

(i) Gib die Aktionen des Shift-Reduce-, des Top-Down- und des Left-Corner-Parsers an. Denk Dir eine geeignete Analyse des Satzes "Jill saw Jim with glasses" aus und gib die Aktionen des Top-Down-Parsers für diese Analyse an.

(ii) Schreibe den CKY-, Count-, Viterbi- und Semiring-Algorithmus in Pseudocode auf. Verzichte hierbei auf die Wiedergabe identischer Zeilen! Gib den Semiring des Viterbi-Algorithmus an.

(iii) Gib eine Formel für die Regel-Wahrscheinlichkeiten beim Baumbank-Training an. Denke Dir nun eine Baumbank aus und wende die Formel auf eine hinreichend interessante Regel an. Wofür benutzt man eigentlich probabilistische Grammatiken? Motiviere Deine Antwort mit drei Beispielen!

(iv) Beschreibe in Stichpunkten die Gemeinsamkeiten und Unterschiede der kopflexikalierten PCFGs von Charniak und Collins. Denke Dir eine kontextfreie Regel mit drei Töchtern aus, versehe sie mit geeigneten Köpfen, und berechne die Wahrscheinlichkeit dieses Gebildes mit den Modellen von Charniak und Collins.

(v) Beschreibe das Wahrscheinlichkeitsmodell von DOP. Zeige, dass die wahrscheinlichste DOP-Derivation eines Satzes genauso effizient berechenbar ist wie der wahrscheinlichste PCFG-Parse dieses Satzes.

(vi) Fasse das in der Vorlesung besprochene Papier von Klein und Manning in maximal drei Sätzen zusammen. Was verstehen die Autoren unter horizontaler und vertikaler Markovisierung, und was unter interner und externer Annotation? Gib für Deine Antworten jeweils ein Beispiel an.

## Abschlussklausur (Ende)