

## 6. Übung zur Vorlesung “NLP – Analyse des Wissensrohstoffes Text” im Sommersemester 2008

Dr. Andreas Hotho, Dipl.-Inform. Dominik Benz, Wi.-Inf. Beate Krause

25. Juni 2008

### 1 Parse Tree Derivations

- Verschiedene Derivationen eines Parsebaumes machen die Berechnung der Wahrscheinlichkeit eines Parsebaumes unnötig kompliziert. Erklären Sie den Begriff Derivation in diesem Kontext. Was wird angewandt, um die unnötige Komplexität zu reduzieren?

### 2 Evaluation

- Das Beurteilen der Ähnlichkeit zweier Parsebäume stellt eine fundamentale Aufgabe dar. Welche Bewertungsmaße kennen Sie dafür?
- Betrachten Sie die verschiedenen Parsebäume auf Seite 410, Abb. 12.2. Parsebaum (a) sei der “Gold-Standard”. Berechnen Sie die Ähnlichkeit der Parsebäume (b) und (c) mit (a) mittels mindestens 3 verschiedener Maße.

### 3 Independence Assumption

- Worin besteht die Independence Assumption bei PCFGs?
- Erklären Sie einige Phänomene natürlicher Sprache, die der Independence Assumption widersprechen. Kennen Sie Ansätze, die diesen Problemen begegnen?