
Vorlesung “Grammatikformalismen”

2. Übung (12.05.2017)

Sommersemester 2017 – Prof. Dr. Alexander Koller

1 Installation von OpenCCG

Installieren Sie den OpenCCG-Parser von der Kurs-Homepage. Parsen Sie mit der mitgelieferten englischen Grammatik (im Verzeichnis `$OPENCCG_HOME/ccg-format-grammars/tiny`) die Sätze “the teachers buy a book” und “he rented her a DVD”. Geben Sie als Lösung dieser Aufgabe die möglichen Ableitungen (ohne Semantik und Features) an, die der CCG-Parser findet. Warum ist es korrekt, dass die Grammatik mehrere Ableitungen für “the teachers buy a book” zulässt?

2 Grammatikfragment

Schreiben Sie eine CCG-Grammatik für einfache deutsche Verbletztsätze und Relativsätze. Ihre Grammatik soll die folgenden Ausdrücke korrekt analysieren (und fehlerhafte Varianten natürlich ablehnen):

- (1) der Mann die Frau liebt
- (2) das Kind schläft
- (3) der Mann der die Frau liebt schläft
- (4) die Frau die der Mann liebt schläft

Geben Sie zusammen mit Ihrer Grammatik mindestens zehn grammatische und fünf ungrammatische Sätze an, die Ihre Grammatik korrekt verarbeitet. Gehen Sie dabei folgendermaßen vor:

- a) Analysieren Sie zunächst Verbletztsätze ohne Relativsätze wie (1) und (2). Sie können sich auf kanonische Wortstellung (erst Subjekt, dann Objekt, dann Verb) sowie definite Singular-NPs einschränken, aber Ihre Grammatik muss intransitive und transitive Verben zulassen und Kasus und Genus korrekt verarbeiten.
- b) Erweitern Sie dann Ihre Grammatik um Subjektrelativsätze wie (3). Überlegen Sie dazu, mit was für Kategorien sich ein Relativpronomen

auf der rechten und linken Seite verbinden möchte. Sie können davon ausgehen, dass der Satz kein Komma enthält. Beachten Sie: OpenCCG hat einen Bug, wegen dem jedes Wort nur mit einer Wortart in der Grammatik vorkommen kann. Sie müssen in dieser Aufgabe deshalb das Relativpronomen anders schreiben als den gleichlautenden Artikel, z.B. als “_der”.

- c) Erweitern Sie schließlich Ihre Grammatik um Objektrelativsätze wie (4), und zwar ohne zusätzliche Lexikoneinträge für Verben einzuführen. Die Regeln für Typeraising und crossed composition könnten hier hilfreich sein.

3 Nicht kontextfreie Sprachen

In dieser Aufgabe schreiben Sie eine OpenCCG-Grammatik für eine nicht-kontextfreie Sprache. Versuchen Sie dafür zunächst, eine Grammatik für die Sprache $\text{COUNT}(3) = \{a^n b^n c^n \mid n \geq 1\}$ zu schreiben. Das wird Ihnen nicht gelingen (ich erkläre in der Übung, warum; Sie können ja schon mal drüber nachdenken).

Schreiben Sie daher stattdessen eine (übergenerierende) Grammatik G , so dass (a) $\text{COUNT}(3) \subseteq L(G)$ und (b) alle anderen Strings in $L(G)$ ebenfalls gleich viele a 's, b 's und c 's enthalten. Geben Sie einen String an, der in der Sprache $L(G)$ Ihrer Grammatik, aber nicht in $\text{COUNT}(3)$ ist.

Begründen Sie dann präzise, warum $L(G)$ keine kontextfreie Sprache ist. Denken Sie daran, dass kontextfreie Sprachen unter Schnitt mit regulären Sprachen abgeschlossen sind; wenn also eine Sprache L_2 regulär ist und $L_1 = L_2 \cap L_3$ nicht kontextfrei, dann kann auch L_3 nicht kontextfrei sein.