

Übungsblatt 4

Anmerkung: Dieses Übungsblatt kann komplett in einem Jupyter Notebook bearbeitet werden. Da die Grammatikregeln dort im Python-Code und auf diesem Übungsblatt in separaten Dateien definiert werden, weicht der Code in Aufgabe 1 jeweils leicht ab. Man sollte sich also für Übungsblatt-Code oder Notebook-Code entscheiden und dann konsequent dabei bleiben.

Hinweis: Unter den folgenden Links finden Sie einige nützliche Funktionen für den Umgang mit Grammatiken und Syntaxbäumen im NLTK:

- <http://www.nltk.org/howto/generate.html>
- <http://www.nltk.org/howto/tree.html>

Präsenzaufgaben

Aufgabe 1 Eine erste (Phrasenstruktur-)Grammatik

Laden Sie sich die Grammatik `04-grammatik.cfg` und das Test-Skript `04-skript.py` von der Kurs-Website herunter und führen Sie letzteres aus.

Werfen Sie einen Blick in die Datei `04-grammatik.cfg` und erklären Sie die Funktionsweise dieser sehr einfachen Grammatik:

```
1 | S → NP VP
2 | NP → DET N
3 | DET → "der" | "die" | "das"
4 | N → "Mann" | "Frau" | "Buch"
5 | VP → V NP NP
6 | V → "gibt" | "schenkt"
```

Aufgabe 2 Ein- und zweistellige Verben

Bis jetzt akzeptiert die Grammatik in `04-grammatik.cfg` lediglich dreistellige Verben. Erweitern Sie sie so, dass auch Verben mit weniger als zwei Objekten korrekte Verbalphrasen bilden können.

Folgende Sätze sollten akzeptiert werden:

- (1) der Mann schläft
- (2) das Buch gefällt der Frau
- (3) die Frau kennt das Buch

Aufgabe 3 Beliebige lange Phrasen

Erweitern Sie die Grammatik nun derart, dass Nominalphrasen auch Adjektive enthalten dürfen – und zwar **beliebig viele**.

Beispiele:

- (4) die kluge schöne Frau kennt das Buch
- (5) der schöne kluge Mann gibt der Frau das dicke Buch
- (6) das dicke schöne kluge Buch schläft

Hausaufgaben

Aufgabe 4 Eigennamen

Eigennamen können auch ohne Artikel eine Nominalphrase bilden. Erweitern Sie die Grammatik `04-grammatik.cfg` entsprechend.

Folgende Sätze sollten korrekt analysiert werden:

- (7) der Mann kennt Chomsky
- (8) Marie gibt Fritz das Buch

Folgende Sätze sollten NICHT akzeptiert werden:

- (9) * Mann gibt Frau Buch
- (10) * Mann schläft

Aufgabe 5 Präpositionalphrasen

Erweitern Sie die Grammatik nun derart, dass **sowohl Nominalphrasen als auch Verbalphrasen** durch Präpositionalphrasen modifiziert werden können. Erkennen Sie die Ambiguität von Übungsblatt 2 wieder?

Folgende Sätze sollten syntaktisch analysiert werden können:

- (11) der Mann schläft neben dem Buch
- (12) die Frau kennt das dicke Buch über Chomsky

(13) die Frau schenkt dem Mann das Buch auf dem Tisch

Beachten Sie, dass die Form des Artikels *dem* und das Nomen *Tisch* auch noch als lexikalische Regeln ergänzt werden müssen.

Stellen Sie sicher, dass zu Satz (12) und (13) **zwei** Analysen generiert werden!

Achten Sie außerdem darauf, dass Ihre Erweiterung nicht die Analyse zuvor akzeptierter grammatischer Sätze verhindert!

Aufgabe 6 Fragen zu NLTK-Kapitel 8.3

Lesen Sie das NLTK-Teilkapitel 8.3 ('Context Free Grammar'): <http://www.nltk.org/book/ch08.html>.

Beantworten Sie insbesondere folgende Fragen:

- (a) Wie ist im NLTK vorzugehen, wenn ein Satz von einer Grammatik nicht erkannt wird?
- (b) Wie können im NLTK die Regeln einer Grammatik ausgegeben werden?
- (c) Geben Sie 2 NP-Phrasenstrukturregeln mit indirekter Rekursion an.
- (d) Geben Sie ein Beispiel für eine links- sowie eine rechtsrekursive NP-Phrasenstrukturregel an.

*Aufgabe 7 Recursive Descent Parser

Die Demo-Applikation `nltk.app.rdparsing`() veranschaulicht die Funktionsweise eines Recursive Descent Parsers. Spielen Sie ein wenig mit der Demo, nehmen Sie Änderungen an der Grammatik vor und stellen Sie sicher, dass Sie verstehen wie jeder einzelne Schritt zustandekommt.

