

Übungsblatt 2

24.10.18

Präsenzaufgaben

Aufgabe 1 Eine erste syntaktische Analyse

Kreieren Sie Ihre ersten Syntaxbäume, indem Sie das folgende Python-Skript ausführen!

- (a) Warum gibt die `parse`-Funktion in Zeile 16 des Skripts vermutlich eine Sequenz zurück? Warum ist die Alternative `tree = parser.parse(sent)` weniger wünschenswert?
- (b) Worin liegen die Unterschiede zwischen den ausgegebenen Bäumen?
- (c) Warum ist der gegebene Satz „I shot an elephant in my pajamas“ syntaktisch ambig?

```
1 import nltk
2
3 grammar = nltk.CFG.fromstring("""
4     S -> NP VP
5     PP -> P NP
6     NP -> Det N | Det N PP | 'I'
7     VP -> V NP | VP PP
8     Det -> 'an' | 'my'
9     N -> 'elephant' | 'pajamas'
10    V -> 'shot'
11    P -> 'in'
12 """)
13 parser = nltk.ChartParser(grammar)
14
15 sent = "I shot an elephant in my pajamas".split()
16 for tree in parser.parse(sent):
17     print(tree)
18     tree.draw()
```

Aufgabe 2 Was ist Grammatik?

- (a) Erläutern Sie die drei Grammatikbegriffe, die den folgenden Sätzen zugrunde liegen.
- (b) Zeigen Sie am Beispiel von Satz (1) den Unterschied zwischen Grammatik und Syntax auf.

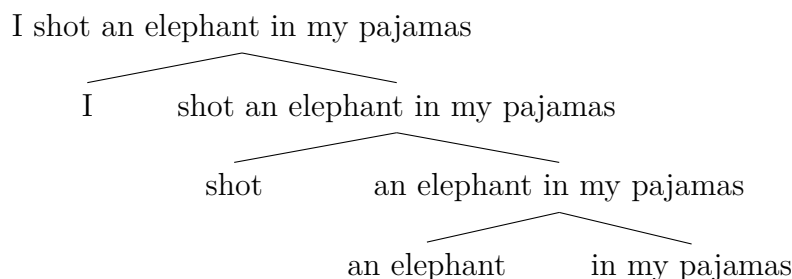
- (1) Die Grammatik des Lateinischen ist schwer. Vor allem die Syntax bereitet Schwierigkeiten.
- (2) Der Duden ist eine Grammatik.
- (3) Der Begründer der Dependenzgrammatik ist Lucien Tesnière.

Aufgabe 3 Ebenen linguistischer Strukturbeschreibung

- (a) Welche Hierarchie-Ebenen lassen sich in der Beschreibung des Aufbaus sprachlicher Strukturen unterscheiden?
- (b) Erklären Sie den Unterschied zwischen dem Wortbegriff einer natürlichen Sprache und dem Wortbegriff einer formalen Sprache (was sind die Blätter des Syntaxbaums einer Ableitung eines natürlichsprachlichen Satzes in einer die Sprache modellierenden formalen Grammatik?)

Aufgabe 4 Darstellung syntaktischer Strukturen

- (a) Welche formalen Repräsentationen der syntaktischen Struktur eines natürlichsprachlichen Satzes haben Sie in der Vorlesung kennengelernt?
- (b) Wie kann eine solche Struktur dargestellt werden? (vgl. `print(tree)` vs `tree.draw()` im NLTK)
- (c) Geben Sie zu folgendem Syntaxbaum den Klammerausdruck an:



- (d) Geben Sie für zu folgendem Klammerausdruck den Syntaxbaum an:

(4) [[I] [[[shot] [an elephant]] [in my pajamas]]]

Hausaufgaben

Aufgabe 5 Ambiguität

Sehen Sie sich die folgenden englischen Sätze an. Können Sie sich jeweils zwei unterschiedliche Interpretationen vorstellen?

- (6) Fighting animals could be dangerous.

- (7) Visiting relatives can be tiresome.

Beantworten Sie insbesondere die Frage nach dem Ursprung der Ambiguität.

- (a) Lexikalische Ambiguität der einzelnen Wörter?
- (b) Morphologisch ambige Formen?
- (c) Andere Gründe?

Aufgabe 6 Fragen zu NLTK-Kapitel 8.1

Lesen Sie das Vorwort von NLTK-Kapitel 8 sowie das Kapitel 8.1 ('Some Grammatical Dilemmas'): <http://www.nltk.org/book/ch08.html>.

Beantworten Sie insbesondere folgende Fragen:

- (a) Hinsichtlich welcher beiden syntaktischen Grundprinzipien werden natürliche Sprachen mit formalen Methoden analysiert?
- (b) Welche beiden Mittel werden (in Kombination) für eine solche Analyse verwendet? (Hinweis: s. Aufgabe 1 dieses Übungsblattes)
- (c) Welche ja/nein-Entscheidung bzgl. des Satzes wird bei dieser Analyse unweigerlich getroffen?
- (d) Welchen Vorteil hat die Modellierung der Satzstruktur mit Hilfe einer formalen Grammatik?