
Vorlesung “Computerlinguistische Techniken”

3. Übung (21.11.2023)

Wintersemester 2023/2024 – Prof. Dr. David Schlangen

1 Earley-Parser ausprobieren - 30

Wenden Sie (von Hand) den Earley-Parser an, um den Satz “Hans beobachtet den Mann mit dem Fernrohr” mit der kontextfreien Grammatik aus Aufgabe 2 der 1. Übung zu parsen. Geben Sie dabei an, in welcher Reihenfolge Sie die einzelnen Parse-Items berechnen und mit welcher Regel jedes einzelne Item begründet ist.

Erklären Sie dann, wie man systematisch die Parsebäume des Satzes aus den berechneten Items rekonstruieren kann. Benutzen Sie Ihren Algorithmus, um die beiden Parsebäume des angegebenen Satzes auszurechnen.

Geben Sie eine pdf-Datei ab.

2 Schemabasierte Implementierung von Shift-Reduce - 35

In der Vorlesung haben wir ein Parsingschema für den Shift-Reduce-Erkenner gesehen. Implementieren Sie dieses Parsingschema in Python. Sie können `earley.py`, die schemabasierte Implementierung eines Earley-Erkenners, auf Moodle als Grundlage verwenden.

Geben Sie eine Python-Datei `shift_reduce.py` ab, die eine Funktion namens `parse(grammar, sentence)` enthält (wobei `sentence` eine Liste von Strings und `grammar` vom Typ `nlk.grammar.CFG` ist).

3 Merkmalsstrukturen - 35

Eine minimale Grammatik, die Merkmalsstrukturen verwendet, können Sie im NLTK wie folgt spezifizieren:

```
fsgram = nltk.grammar.FeatureGrammar.fromstring("""
% start S
S -> NP[NUM=?n] VP[NUM=?n]
NP[NUM=?n] -> Det[NUM=?n] N[NUM=?n]
Det[NUM=sg] -> 'der'
Det[NUM=pl] -> 'die'
N[NUM=sg] -> 'Mann'
N[NUM=pl] -> 'Maenner'
VP[NUM=sg] -> 'lacht'
VP[NUM=pl] -> 'lachen'
""")
```

Einen Satz können Sie damit wie folgt parsen:

```
parser = nltk.FeatureChartParser(fsgram, trace=1)
print(parser.parse("die Maenner lachen".split()))
```

Konstruieren Sie einige Merkmalsstrukturen (nicht notwendigerweise aus dem linguistischen Bereich, sie können auch andere Objekte auf diese Weise modellieren). Bilden Sie dabei Beispiele für Strukturen, die sich unifizieren lassen, und solche, die sich nicht unifizieren lassen (und zeigen Sie dies jeweils).

Erweitern Sie die Grammatik von oben so, dass diese auch weitere Kongruenzphänomene abbildet. So soll die Grammatik z.B. “Die Mann lieben dem Frauen” abweisen – aber natürlich entsprechende grammatikalisch korrekte Versionen annehmen.

Geben Sie eine Datei mit Namen `feature_structures.py` ab!