

## Aufgabe 8

### Parsing mit neuronalen Netzen: Einlesen der Daten

In den folgenden Übungen werden Sie einen Shift-Reduce-Parser auf Basis von neuronalen Netzwerken implementieren.

In der ersten Übung geht es darum, Parsebäume einzulesen und die entsprechende Folge von Aktionen des Shift-Reduce-Parser für jeden Parsebaum zu generieren. Diese wird dann vom Shift-Reduce-Parser für das Training verwendet werden. Die Daten finden Sie in <http://www.cis.uni-muenchen.de/~schmid/lehre/data/PennTreebank.zip>.

Beispiel:

Für die Eingabe

(TOP(S(NP(NNP Ms.)(NNP Haag))(VP(VBZ plays)(NP(NNP Elianti)))(. .)))

soll zurückgegeben werden:

- die Wortliste: 'Ms.', 'Haag', 'plays', 'Elianti', '.'
- die Aktionsliste:  
 'Shift', 'NNP 1', 'Shift', 'NNP 1', 'NP 2', 'Shift', 'VBZ 1', 'Shift', 'NNP 1',  
 'NP 1', 'VP 2', 'Shift', '. 1', 'S 3', 'TOP 1'  
 'VP 2' bezeichnet hier eine Reduce-Aktion, welche die beiden obersten Stack-  
 symbole durch "VP" ersetzt.

Zusätzlich sollen Sie eine Funktion implementieren, welche aus der Wortliste und der Liste der Aktionen wieder den originalen Parsebaum generiert und ausgibt.

### Vorüberlegungen

- In welche Teilaufgaben lässt sich die Aufgabe gliedern?
- Wie kann das Programm sinnvoll strukturiert werden?