

Aufgabe 1) Kontextfreie Grammatiken**Punkte: 7**

Betrachten Sie folgendes Korpus, das schon mit POS Tags versehen ist:

- Mit/P dem/DET Wissen/N wächst/V der/DET Zweifel/N
- Und/CONJ sie/PRON bewegt/V sich/PRON doch/ADV
- Alt/ADJ werden/V ist/V nichts/PRON für/P Feiglinge/N
- Könige/N und/CONJ Bettler/N brauchen/V keine/DET Empfehlungen/N

a) Parsebäume erstellen*Punkte: 4*

Erstellen Sie für jeden Satz einen Parsebaum mit sinnvoller Phrasengruppierung. Chomsky Normalform ist nicht nötig.

b) Grammatik erstellen*Punkte: 1*

Lesen Sie aus der nun erstellten kleinen Baumbank eine Grammatik heraus. Es ist ausreichend, die Phrasenregeln anzugeben (POS Regeln werden also nicht gebraucht).

c) Probleme einer simplen Grammatik*Punkte: 1*

Konstruieren Sie aus Ihrer Grammatik und dem gegebenen Vokabular einen Satz (mit Syntaxbaum), welcher Ihrem Sprachverständnis nach syntaktisch nicht korrekt ist. Markieren Sie, welcher Teil des Satzes nicht korrekt ist.

d) Erweiterung der Grammatik*Punkte: 1*

Passen Sie die Grammatik mit zusätzlichen Regeln so an, dass der folgende Satz geparkt werden kann:

Mit/CONJ Wissen/N werden/V Könige/N alt/ADJ

(Groß-/Kleinschreibung von "alt" können Sie ignorieren)

Aufgabe 2) CYK-Algorithmus

Punkte: 7

Gegeben sind die folgenden Regeln einer Grammatik:

$S \rightarrow NP VP$
 $S \rightarrow N VP$
 $NP \rightarrow N N$
 $NP \rightarrow NP N$
 $VP \rightarrow V S$
 $VP \rightarrow V NP$
 $VP \rightarrow V N$
 $N \rightarrow \{\text{landmine,claims,dog,arms,company}\}$
 $V \rightarrow \{\text{claims,arms}\}$

Zusätzlich ist der folgende Satz (im Überschriftenstil) gegeben:

`landmine claims dog arms company`

Parsen Sie den Satz mit Hilfe des CYK-Algorithmus. Sie sollen am Ende die Parses herauslesen. Leider ist die Chart recht voll, so dass es verwirrend sein wird, alle Pfeile einzumalen. Indizierung der wichtigsten Konstituenten ist eine Alternative. Wenn Pfeile, dann bitte farbig. (4 Punkte für die Chart).

Zeichnen Sie dann alle Bäume, die für den Satz laut CYK erstellt werden und paraphrasieren Sie die Lesart (auf Deutsch oder Englisch). (3 Punkte).

Aufgabe 3) Grammatiken schreiben

Punkte: 6

1. Gegeben ist die Menge der Terminale Σ als $\{0, 1\}$. Die Sprache L ist gegeben als die Menge aller Zeichenketten ("Sätze") über den Terminalen mit genausovielen Nullen wie Einsen. Schreiben Sie eine kontextfreie Grammatik $G = \{NT, \Sigma, R, S\}$ mit $L(G) = L$. (Hinweis: Sie brauchen außer dem Startsymbol kein weiteres Nichtterminal.) (3 Punkte)
2. Gegeben ist die Menge der Terminale Σ als $\{0, 1\}$. Die Sprache L ist gegeben als die Menge aller Zeichenketten ("Sätze") über den Terminalen, die mit 00 enden. Schreiben Sie eine kontextfreie Grammatik $G = \{NT, \Sigma, R, S\}$ mit $L(G) = L$. Könnten Sie stattdessen auch einen regulären Ausdruck schreiben? (3 Punkte).