

Aufgabe 1) Kontextfreie Grammatiken

Punkte: 7

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

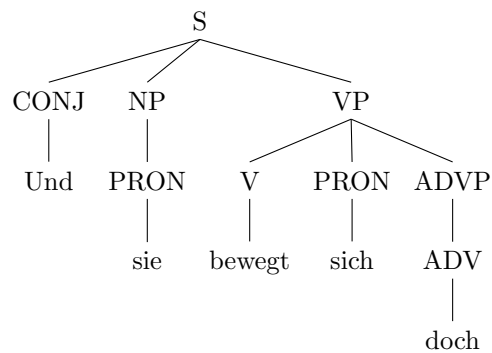
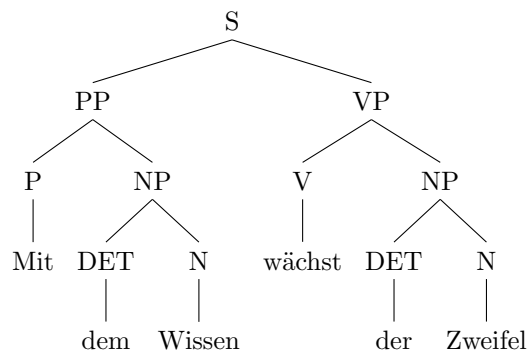
- Mit/P dem/DET Wissen/N wächst/V der/DET Zweifel/N
- Und/CONJ sie/PRON bewegt/V sich/PRON doch/ADV
- Alt/ADJ werden/V ist/V nichts/PRON für/P Feiglinge/N
- Könige/N und/CONJ Bettler/N brauchen/V keine/DET Empfehlungen/N

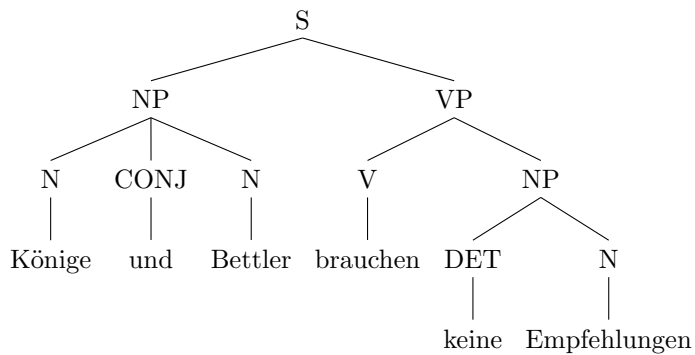
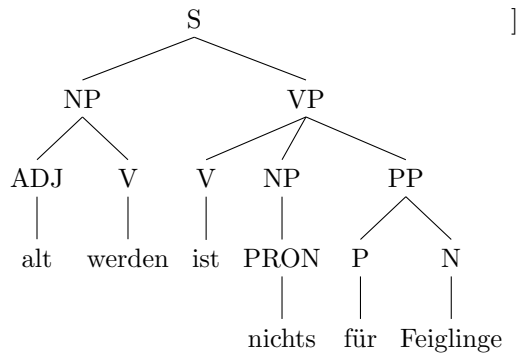
a) Parsebäume erstellen

Punkte: 4

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

Lösung zu 1a)





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).

Lösung zu 1b)

- S → PP VP
- S → NP VP
- S → CONJ NP VP
- PP → P NP
- PP → P N
- VP → V NP
- VP → V NP PP
- VP → V PRON ADVP
- ADVP → ADV
- NP → DET N
- NP → ADJ V
- NP → PRON
- NP → N CONJ N

c) Probleme einer simplen Grammatik

Punkte: 1

Konstruieren Sie aus Ihrer Grammatik und dem gegebenen Vokabular einen Satz (mit Syn-

taxbaum), welcher Ihrem Sprachverständnis nach syntaktisch nicht korrekt ist. Markieren Sie, welcher Teil des Satzes nicht korrekt ist.

Lösung zu 1c)

Viele Möglichkeiten, bei denen Agreement nicht funktioniert. Zum Beispiel “Mit der Könige bewegt nichts für Empfehlungen” Durch die simple Grammatik kann keine Kongruenz von Kasus, Genus und Numerus sichergestellt werden.

d) Erweiterung der Grammatik

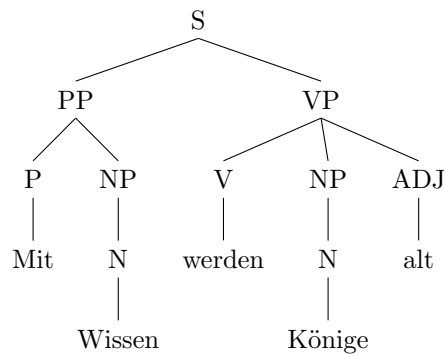
Punkte: 1

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)

Lösung zu 1d)



VP → V NP ADJ

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

Parzen 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).

Lösung zu 2)

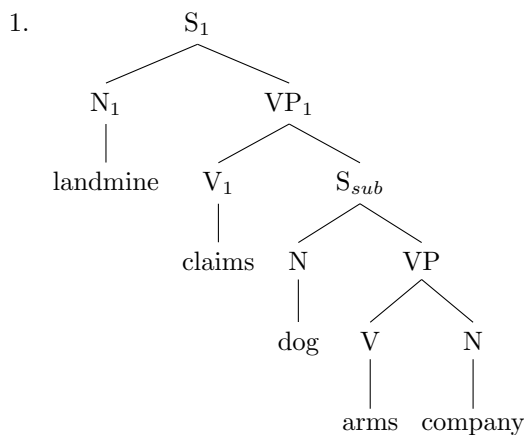
0	<i>landmine</i>	1	<i>claims</i>	2	<i>dog</i>	3	<i>arms</i>	4	<i>company</i>	5
	N_1 [0, 1]	NP [0, 2]	NP_3, S [0, 3]	S, NP [0, 4]	S_1, S_2, S_3, NP [0, 5]					
		N, V_1 [1, 2]	NP, VP [1, 3]	VP, NP [1, 4]	VP_1, VP_2, S, NP [1, 5]					
			N [2, 3]	NP [2, 4]	S_{sub}, NP_2 [2, 5]					
				N, V [3, 4]	NP, VP_3 [3, 5]					
					N [4, 5]					

Nur die wichtigsten Konstituenten sind indiziert. Hierbei kommt in der letzten Spalte VP_3 von der VP Interpretation von *arms company*.

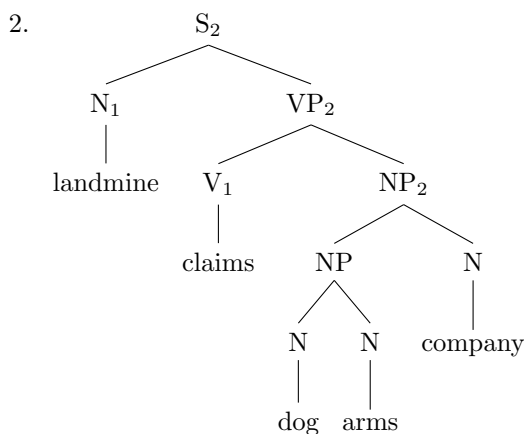
Des weiteren kommt in der letzten Spalte (Zelle [1, 5]) VP_1 von der Zellenkombination [1, 2] mit [2, 5] aus der V_1 und S_{sub} Kombination (interpretation *claims that a dog arms a company*). Dagegen kommt VP_2 aus der gleichen Zellenkombination [1, 2] mit [2, 5] aus der V_1 und NP_2 Kombination (*claims a dog-arms-company*). Die NP in dieser Zelle kommt aus der Zellenkombination [1, 4] mit [4, 5] aus der NP und N Kombination und entspricht einem sehr langem Nominalkompositum *claims dog arms company*.

In der letzten Zelle [0, 5] erhält man dann drei Sätze. S_1 ergibt sich aus $N_1 = \textit{landmine}$ (Zelle [0, 1]) mit VP_1 . S_2 ergibt sich aus $N_1 = \textit{landmine}$ (Zelle [0, 1]) mit VP_2 . S_3 ergibt sich aus der Kombination von NP_3 mit VP_3 (ein *landmine-claims-dog* bewaffnet eine Firma).

Die Sätze unten geben die drei entsprechenden Bäume an.

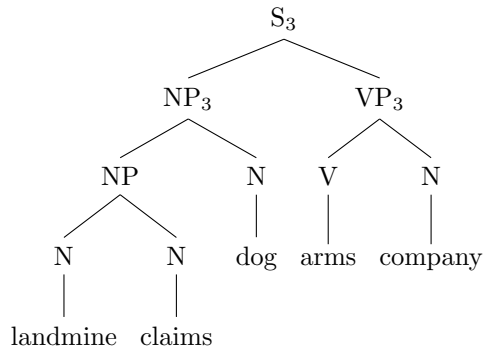


Bedeutung: *Landmine behauptet, dass Hund eine Firma bewaffnet*



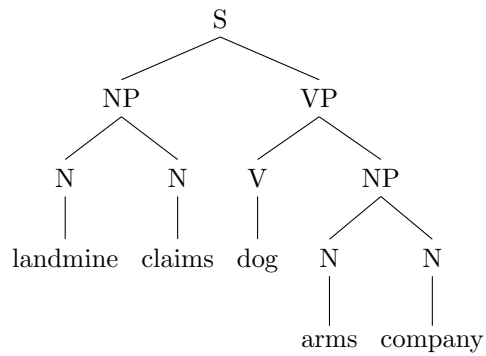
Bedeutung: *Landmine nimmt für sich eine Hundenwaffenfirma in Anspruch.*

3.



Bedeutung: *Ein Landminenansprüchehund (?) bewaffnet eine Firma.*

Eine vierte Lesart wäre möglich, wenn *dog* auch als Verb erlaubt wäre (was im Englischen geht, aber nicht in der Grammatik vorgesehen ist).



Dies ist eigentlich die sinnvollste (*Ansprüche aus Landminenschäden plagen Waffenfirma*).

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).

Lösung zu 3)

1. Sie brauchen die folgenden Regeln.

- $S \rightarrow \{01, 10\}$
- $S \rightarrow SS$
- $S \rightarrow 0S1$
- $S \rightarrow 1S0$
- $S \rightarrow \epsilon$

Diese Sprache ist nicht regulär.

2. Sie brauchen außer S ein weiteres Nichtterminal A und die folgenden Regeln:

- $S \rightarrow A00$
- $A \rightarrow \epsilon$
- $A \rightarrow A0$
- $A \rightarrow A1$

Man könnte auch leicht einen regulären Ausdruck schreiben. Diese Sprache ist regulär.