

Sprachtechnologie

5. Übung

Prof. Tobias Scheffer
Uwe Dick

Sommer 2015

Ausgabe am: 18.05.15
Besprechung am: 01.06.15

Aufgabe 1

Finite State Transducer

Gegeben Sei folgende synchrone reguläre Grammatik für Sätze in deutscher und blergischer Sprache:

Regel	Wahrscheinlichkeit
$S_1 \rightarrow \text{Die/schmurp}, S_2$	0,3
$S_1 \rightarrow \text{Die/flump}, S_3$	0,7
$S_2 \rightarrow \text{Vorlesung/grink}, S_3$	0,5
$S_2 \rightarrow \text{Vorlesung/blip}, S_4$	0,5
$S_3 \rightarrow \text{Vorlesung/frunz}, S_3$	0,1
$S_3 \rightarrow \text{Vorlesung/mirx}, S_4$	0,4
$S_3 \rightarrow \text{macht/zwup}, S_4$	0,2
$S_3 \rightarrow \text{Spaß/klonk}, S_e$	0,3
$S_4 \rightarrow \text{macht/schmotter}, S_3$	0,5
$S_4 \rightarrow \text{Spaß/zeng}, S_e$	0,5

Übersetzen Sie mit Hilfe der Grammatik den Satz “Die Vorlesung macht Spaß” vom Deutschen in die Blerg-Sprache, indem Sie zuerst die wahrscheinlichste Regelfolge bestimmen.

Aufgabe 2

Synchrone PCFG

Gegeben Sei folgende synchrone PCFG für Sätze in deutscher und gnarfischer Sprache, wobei S das Startsymbol ist:

Regel	Wahrscheinlichkeit
$S \rightarrow A_{[1]}A_{[2]}/A_{[2]}A_{[1]}$	0,5
$S \rightarrow C_{[1]}B_{[2]}/C_{[1]}B_{[2]}$	0,5
$A \rightarrow C_{[1]}A_{[2]}/A_{[2]}C_{[1]}$	0,2
$A \rightarrow D_{[1]}A_{[2]}/A_{[2]}D_{[1]}$	0,3
$A \rightarrow \text{Übung/flarg}$	0,3
$A \rightarrow \text{Spaß/yarg}$	0,2
$B \rightarrow B_{[1]}D_{[2]}/D_{[2]}B_{[1]}$	0,3
$B \rightarrow \text{Übung/schnarg}$	0,4
$B \rightarrow \text{Spaß/hjarb}$	0,3
$C \rightarrow C_{[1]}B_{[2]}/C_{[1]}B_{[2]}$	0,3
$C \rightarrow \text{Die/warb}$	0,7
$D \rightarrow \text{macht/bralb}$	1

Übersetzen Sie mit Hilfe der Grammatik den Satz “Die Übung macht Spaß” vom Deutschen in die Gnarf-Sprache, indem Sie zuerst den wahrscheinlichsten Parsebaum bestimmen.

Aufgabe 3

Phrasen

Gegeben das nachstehende Wort-zu-Wort-Alignment, extrahieren Sie alle Phrasenpaare, und dann alle Regeln für die hierarchische Phrasenübersetzung.

