

Maschinelles Übersetzen natürlicher Sprachen

3. Übungsblatt

2014-10-30

Aufgabe 1 (Bigrams)

Let $E = \{\text{du, su, ur, fur, mu}\}$ and $F = \{\text{kra, ban, las, gha, ra}\}$. Consider the following dictionary:

$t(f e)$	f				
	kra	ban	las	gha	ra
du	1	0	0	0	0
su	0	1	0	0	0
e ur	0	0	1	0	0
fur	0	0	0	1	0
mu	0	0	0	0	1

Let $\varepsilon(m | l) = 1$ if $l = m$ and $\varepsilon(m | l) = 0$ otherwise. Decode the sentence “kra ban las gha ra” using the following bigram model:

$b(e' e)$	e'					
	du	su	ur	fur	mu	#
du	0	0.2	0.1	0.3	0.2	0.2
su	0.3	0	0	0.3	0.3	0.1
e ur	0.1	0	0	0.7	0	0.2
fur	0.3	0.2	0.2	0	0	0.3
mu	0.2	0.5	0	0	0	0.3
#	1	0	0	0	0	0

Aufgabe 2

Let $V_E = \{a, b\}$ and $V_F = \{\alpha, \beta, \gamma\}$ be an English and French vocabulary, respectively. Consider the following bigram model, length model, and dictionary.

$b(\rightarrow \downarrow)$	#	a	b	$\varepsilon(m l) = \begin{cases} 2^{-1} & \text{if } m = l \\ 2^{-2} & \text{if } m - l = 1 \\ 0 & \text{otherwise} \end{cases}$	$t(\rightarrow \downarrow)$	α	β	γ	
#	2^{-1}	2^{-1}	0			a	2^{-1}	0	2^{-1}
a	0	2^{-1}	2^{-1}			b	0	2^{-1}	2^{-1}
b	2^{-1}	0	2^{-1}						

Decode the sentence $\beta\beta$ using the algorithm from the lecture. Annotate the hypotheses with a target length to make the calculation feasible.

Zusatzaufgabe 1

Write a small program to solve the previous task.