Maschinelles Übersetzen natürlicher Sprachen

2. Übungsblatt

2016-11-03

Wir bearbeiten noch Aufgaben 6–8 des ersten Übungsblatts.

Aufgabe 1

1. Extend the bigram model from the lecture to a trigram model, i.e. a language model where the probability of a word in a sentence depends on two preceding words.
2. Define a general n-gram model.

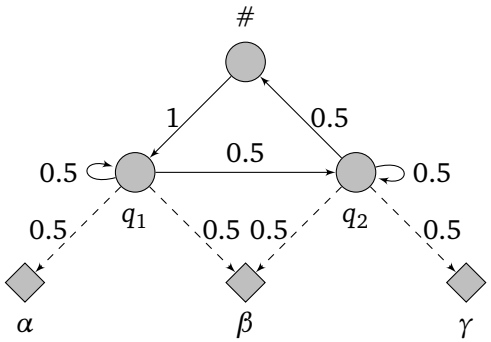
Aufgabe 2 (Bigrams)
Let $E = \{\text{du}, \text{su}, \text{ur}, \text{fur}, \text{mu}\}$ and $F = \{\text{kra}, \text{ban}, \text{las}, \text{gha}, \text{ra}\}$. Consider the following dictionary:

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

		e'					
$b(e' \mid e)$		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 3
Consider the hidden Markov model $H = (Q, V, \#, t, e)$ where $Q = \{q_1, q_2\}$, $V = \{\alpha, \beta, \gamma\}$, and t and e are given by the following graphic.



Calculate the following probabilities.

$$P(\mathbb{Q} = q_1 q_1 q_2, \mathbb{V} = \alpha\beta\gamma)$$
$$P(\mathbb{V} = \alpha\beta\gamma)$$
$$P(\mathbb{V} = \alpha\beta^n\gamma) \text{ where } n \in \mathbb{N}$$
$$P(\mathbb{Q} = q_1 q_1 q_2)$$
$$P(\mathbb{V}_2 = \beta)$$