Ergänzungen zum maschinellen Übersetzen natürlicher Sprachen

5. Übungsblatt

2016-06-07

Exercise 1

Imagine the following game: Two coins are thrown and you win, if both coins land on the same side. You are only told, if you won. Assume that one coin is rather thick and may land on the edge. We represent the possible events with the set $Y = \{M_1, M_2, M_3\} \times \{N_1, N_2\}$. You win with the events (M_1, N_1) and (M_2, N_2).

After playing the game several times, you won 6 times and lost 18 times. Instantiate the corpus-based EM algorithm with this scenario and calculate one EM step. Start with the probability of 2/5 for M_1 and M_2, and 1/3 for N_1.

Exercise 2

Imagine another game: One of two coins is thrown and you win, if the coin lands on the head or on the edge. Assume that second coin is thin and never lands on the edge. You are only told, if you won. We represent the possible events with the set $Y = \{M_1, M_2, M_3, N_1, N_2\}$, i.e., you win with the events M_1, M_3, and N_1.

After playing the game several times, you won 8 times and lost 12 times. Instantiate the simple counting EM algorithm with this scenario and calculate one EM step. Start with the probability of 2/5 for M_1 and M_2, 1/3 for N_1, and 1/3 for choosing the first coin.

Assure yourself that $\mathrm{cmle}_q(c\langle \omega, \kappa \rangle) = \mathrm{cmle}_p(c\langle \omega, p \rangle)$ in this particular instance of simple counting EM.