Ergänzungen zum maschinellen Übersetzen natürlicher Sprachen

2. Übungsblatt

2016-04-19

Exercise 1
Let $X = \{h, t\}$ and c be an X-corpus such that $c(h) = 4$ and $c(t) = 6$. Moreover, let p be a probability distribution of X such that $p(h) = 0.3$ and $p(t) = 0.7$. Then

$$p(c) =$$

Determine $\widetilde{c}$:

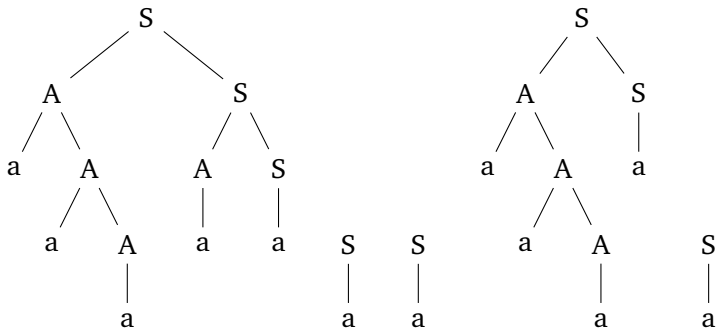
Let $p \in \mathcal{M}(X)$. Determine $\overline{p}(c)$:

Exercise 2
Let $\mathbb{N}_{\geq 1}$ be the set of positive integers. The corpora $c_1, c_2, c_3 \colon \mathbb{N}_{\geq 1} \to \mathbb{R}_{\geq 0}$ are defined as follows:

$$c_1(1) = 5, \quad c_1(2) = 10, \quad c_1(3) = 5, \quad \text{and 0 otherwise;}$$
$$c_2(n) = 2^{-n};$$
$$c_3(n) = \frac{1}{n}.$$

Determine $\overline{p}(c_i)$ for every $i \in \{1, 2, 3\}$ and arbitrary $p \in \mathcal{M}(\mathbb{N}_{\geq 1})$.

Exercise 3
Consider the context-free grammar $G = (Z, \Sigma, S, P)$ with nonterminal symbols $Z = \{S, A\}$, terminals $\Sigma = \{a, b\}$, start symbol S, and a set of productions $P = \{z \to x \mid z \in Z, x \in X\}$ where $X = \{AS, a, aA, b\}$. We observe the following sequence of parse trees:



a) Specify an $X \times Y$-corpus c that reflects this observation.

Compute $c_S(AS) =$, $c_S(a) =$, $c_A(aA) =$, $c_A(a) =$.

b) Let $p \in \mathcal{M}(X|Z)$ be such that

$$p(AS \mid S) = 0.3 \qquad p(aA \mid A) = 0.5$$
$$p(\ \ a \mid S) = 0.7 \qquad p(\ \ a \mid A) = 0.5.$$

Compute $p(c)$:

$p(c) =$

c) Let $p \in \mathcal{M}(X \mid Z)$. Then

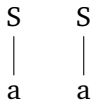
$\overline{p}(c)(S) =$ $|c_S| =$

$\widetilde{c_S}(AS) =$ $\widetilde{c_S}(a) =$

$\overline{p}(c)(A) =$ $|c_A| =$

$\widetilde{c_A}(aA) =$ $\widetilde{c_A}(a) =$.

d) Now assume, that we observe the following sequence of parse trees.



Specify an $X \times Y$-corpus c' that reflects this observation and repeat tasks b) and c) with c'.

e) Let $\Omega = \{(u, v) \mid u, v \in [0, 1]\}$ and $p \colon \Omega \to \mathcal{M}(X \mid Z)$ such that

$$\big(p(u,v)\big)(AS \mid S) = u \qquad \big(p(u,v)\big)(aA \mid A) = v$$
$$\big(p(u,v)\big)(\ \ a \mid S) = 1-u \qquad \big(p(u,v)\big)(\ \ a \mid A) = 1-v.$$

Instantiate the optimization problem $\mathrm{cmle}_p(c)$ for p and c.

$\mathrm{cmle}_p(c) =$