

Ergänzungen zum maschinellen Übersetzen natürlicher Sprachen

2. Übungsblatt

2015-05-12

Aufgabe 1

By throwing a die $m \geq 1$ times you win a prize, if you obtain every number $i \in \{1, \dots, 6\}$ exactly m_i times, where $m = m_1 + \dots + m_6$. Assume you can produce your own die subject to the following condition: opposing sides must have the same probability. Define the underlying probability model. Which probability distribution would you chose?

Aufgabe 2

Let $n \in \mathbb{N}$. Let $\Omega_1, \dots, \Omega_n$ and X_1, \dots, X_n be non-empty, finite sets. For every $i \in \{1, \dots, n\}$, let $p^{(i)}: \Omega_i \rightarrow \mathcal{M}(X_i)$ be a Ω_i -probability model. Let $\Omega = \Omega_1 \times \dots \times \Omega_n$ and $X = X_1 \times \dots \times X_n$. We define the Ω -probability model $p: \Omega \rightarrow \mathcal{M}(X)$ by

$$p_{(\omega_1, \dots, \omega_n)}(x_1, \dots, x_n) = p_{\omega_1}^{(1)}(x_1) \cdot \dots \cdot p_{\omega_n}^{(n)}(x_n)$$

for every $\omega_i \in \Omega_i$, $x_i \in X_i$, $i \in \{1, \dots, n\}$.

Let c be an X -corpus. For every $i \in \{1, \dots, n\}$, let c_i be an X_i -corpus such that $c_i(x') = \sum_{x=(x_1, \dots, x_n) \in X: x_i=x'} c(x)$. Show that $\text{mle}_p(c) = \text{mle}_{p^{(1)}}(c_1) \times \dots \times \text{mle}_{p^{(n)}}(c_n)$.

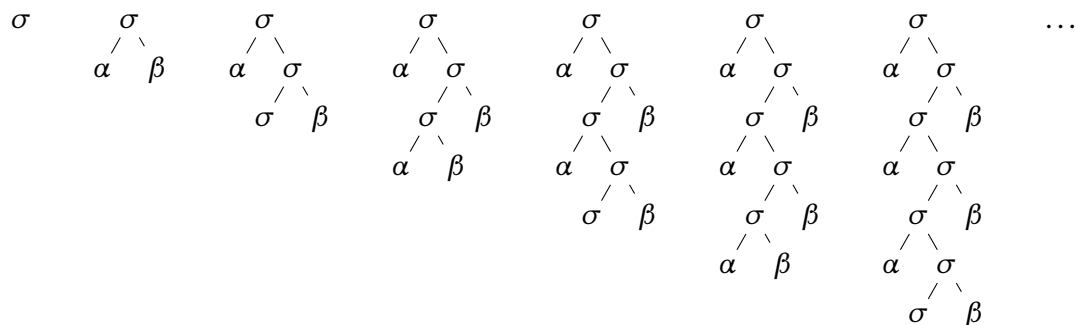
Aufgabe 3

Let Σ be an alphabet, U a set, $k \in \mathbb{N}$, and $t, t', t'_1, \dots, t'_k \in U_\Sigma(V)$. Formally define the following notions from the lecture:

- pos ,
- $t(w)$ for $w \in \text{pos}(t)$,
- $t|_w$ for $w \in \text{pos}(t)$,
- $t[t'_i]_w$ for $w \in \text{pos}(t)$,
- $c[t'_1, \dots, t'_k]$ for $k \in \mathbb{N}$ and $c \in C_\Sigma^k$.

Aufgabe 4

Construct a regular tree grammar G , such that $\llbracket G \rrbracket$ contains exactly the trees of the following form.



Note that every second tree has a σ leaf. Give some elements of $D^q(G)$ for some states q of G . Is your G unambiguous? Is it deterministic? If not, can you find a deterministic grammar?