

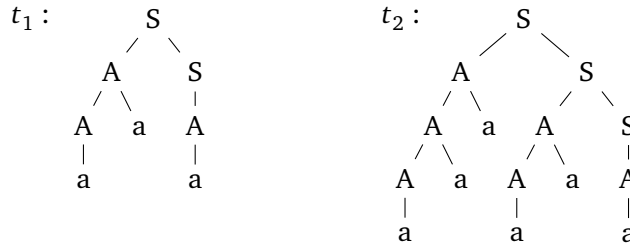
Maschinelles Übersetzen natürlicher Sprachen

6. Übungsblatt

2014-11-20

Aufgabe 1

Consider the following trees.



Let C be a corpus with $C(t_1) = 2$ and $C(t_2) = 1$. Train a pcfg on C .

Aufgabe 2

Let (G, p) be a pcfg with the following rules and probability assignment.

$$S \rightarrow S \quad \# 0.25$$

$$S \rightarrow SS \quad \# 0.25$$

$$S \rightarrow a \quad \# 0.25$$

$$S \rightarrow \varepsilon \quad \# 0.25$$

Determine the reduct $a \triangleleft G$; you may ignore rules which are not useful. Approximate the probability of a .

Aufgabe 3

Supervised training of pcfg results in proper and consistent pcfgs. Give intuitive arguments why these pcfgs are consistent.

Aufgabe 4

Let (G, p) be a pcfg with $G = (N, \Sigma, S, R)$. Let in be defined as in the lecture.

$$in: N \rightarrow \mathbb{R}_{\infty}^{\geq 0}: A \mapsto \sum_{d \in D_G(A, \Sigma^*)} P(d)$$

Show that the following equation holds, as stated in the lecture.

$$\forall A \in N: in(A) = \sum_{\substack{k \in \mathbb{N}, w_0, \dots, w_k \in \Sigma^*, \\ A_1, \dots, A_k \in N: \\ r = A \rightarrow w_0 A_1 w_1 \dots A_k w_k \in R_A}} p(r) \cdot \prod_{i=1}^k in(A_i)$$

Aufgabe 5

Let (G, p) be a proper pcfg with $G = (N, \Sigma, S, R)$. Using your knowledge about in , show the following.

$$\forall A \in N: \sum_{d \in D_G(A, \Sigma^*)} P(d) \leq 1$$