

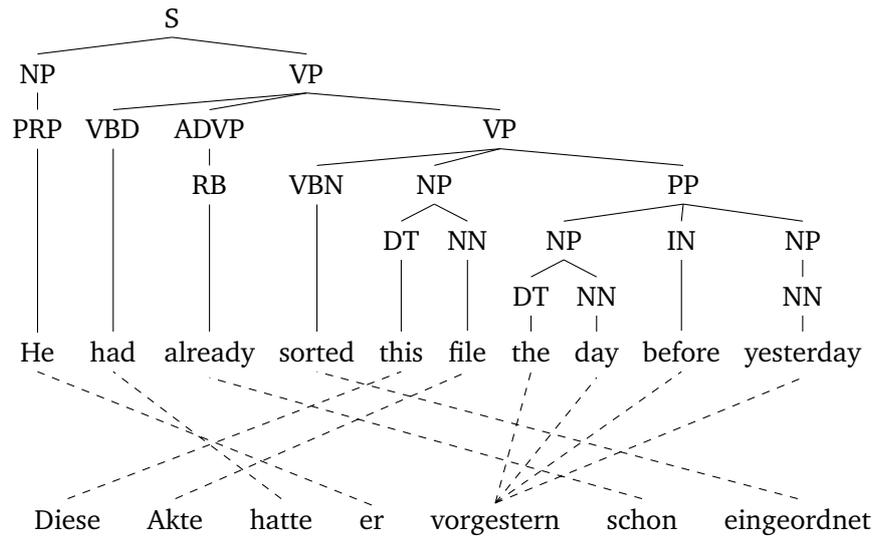
Maschinelles Übersetzen natürlicher Sprachen

9. Übungsblatt

2015-01-22

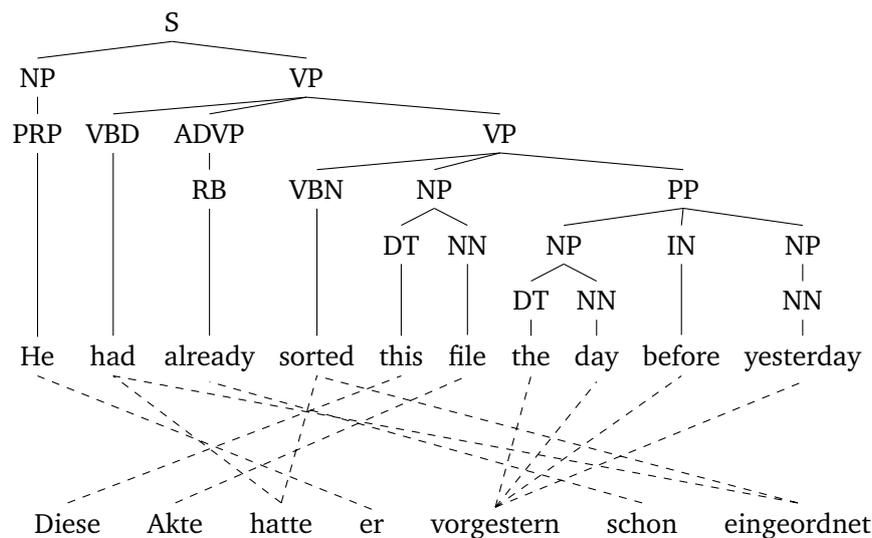
Aufgabe 1 (Regelextraktion für yXTT)

Gegeben ist der folgende Alignment-Graph $G(\xi, f, A)$:



Nutzen Sie die in der Vorlesung vorgestellte Technik, um Regeln für einen yXTT zu erzeugen. Kennzeichnen Sie dabei die *frontier nodes* des Graphen!

Zusatzaufgabe: Inwiefern verändern sich die Regeln, wenn wir zum Extrahieren den folgenden Alignment-Graphen $G(\xi, f, A')$ verwenden? Welche Regeln versprechen eine bessere Übersetzung?



Aufgabe 2 (Konstruktion von RTG)

Gegeben sei das Alphabet $\Sigma = \{\sigma, \tau, \gamma, \alpha\}$. Konstruieren Sie eine RTG \mathcal{G} , so dass $\xi \in L(\mathcal{G})$ genau dann, wenn

- die Symbole σ und τ nur mit Rang 2 in ξ vorkommen, γ nur mit Rang 1 und α mit Rang 0, und außerdem
- das Symbol τ genau einmal in ξ enthalten ist.

Geben Sie für den Baum $\sigma(\alpha, \tau(\alpha, \gamma(\alpha)))$ eine Ableitung von \mathcal{G} an!

Aufgabe 3 (Top-Down-Determinismus)

Wir nehmen das Alphabet $\Sigma = \{\sigma, \alpha, \beta\}$ an. Warum existiert für die Baumsprache $L = \{\sigma(\alpha, \beta), \sigma(\beta, \alpha)\}$ keine top-down-deterministische RTG \mathcal{G} mit $L = L(\mathcal{G})$?

Wie stehen die Sprachklassen REG und REG_{det} daher also in Beziehung, wenn REG die Klasse aller durch reguläre Baumgrammatiken, und REG_{det} die Klasse aller durch top-down-deterministische reguläre Baumgrammatiken erzeugten Baumsprachen bezeichnet?

Welchen Vorteil bieten deterministische Grammatiken jedoch im Vergleich zu nicht deterministischen?

Aufgabe 4 (Input- and Output-Product)

Sei $\Sigma = \{\sigma, \gamma, \alpha, \beta\}$. Wir nehmen einen XTT $\mathcal{M} = (\{q, p\}, \Sigma, \Sigma, R)$ an mit den Regeln

$$\begin{aligned} q(\sigma(x_1, x_2)) &\rightarrow \sigma(q(x_2), p(x_1)) & q(\alpha) &\rightarrow \alpha \\ p(\gamma(x_1)) &\rightarrow \gamma(\gamma(\gamma(p(x_1)))) & p(\alpha) &\rightarrow \beta, \end{aligned}$$

sowie RTGs $\mathcal{G}_1 = (\{S, A\}, \Sigma, S, R_1)$ und $\mathcal{G}_2 = (\{S, B, C\}, \Sigma, S, R_2)$, wobei die Regeln aus R_1 durch

$$\begin{aligned} S &\rightarrow \sigma(A, S) & A &\rightarrow \gamma(A) \\ S &\rightarrow \alpha & A &\rightarrow \alpha \end{aligned}$$

und die aus R_2 durch

$$\begin{aligned} S &\rightarrow \sigma(S, B) & S &\rightarrow \alpha \\ B &\rightarrow \gamma(C) & B &\rightarrow \beta \\ C &\rightarrow \gamma(B) \end{aligned}$$

gegeben sind.

Bilden Sie mittels der in der Vorlesung vorgestellten Konstruktion das *Input- and Output-Product* $\text{Prod}(\mathcal{G}_1, \mathcal{M}, \mathcal{G}_2)$.

Fleißaufgabe: Berechnen Sie $\text{Prod}(\mathcal{G}'_1, \mathcal{M}, \mathcal{G}_2)$ für $\mathcal{G}'_1 = (\{S, A, B, C\}, \Sigma, S, R'_1)$ mit den Regeln

$$\begin{aligned} S &\rightarrow \sigma(A, S) & S &\rightarrow \alpha & A &\rightarrow \alpha \\ A &\rightarrow \gamma(B) & B &\rightarrow \gamma(C) & C &\rightarrow \gamma(A). \end{aligned}$$