

Parsing

Homework 12 (Weighted deductive parsing), due 13 July 2020

Laura Kallmeyer

SS 2020, Heinrich-Heine-Universität Düsseldorf

Question 1 (Weighted Deductive Parsing)

Betrachten Sie das Earley Parsing Deduktionssystem, versehen mit zusätzlichen Gewichten:

$$\text{Axioms: } \frac{}{\langle |\log(p)|, |\log(p)| \rangle : [S \rightarrow \bullet\alpha, 0, 0]} \quad p : S \rightarrow \alpha \in P, w_1 \in \text{First}(\alpha)$$

$$\text{Scan: } \frac{\langle x, y \rangle : [A \rightarrow \alpha \bullet a\beta, i, j]}{\langle x, y \rangle : [A \rightarrow \alpha a \bullet \beta, i, j + 1]} \quad w_{j+1} = a$$

$$\text{Predict: } \frac{\langle x, y \rangle : [A \rightarrow \alpha \bullet B\beta, i, j]}{\langle x + |\log(p)|, |\log(p)| \rangle : [B \rightarrow \bullet\gamma, j, j]} \quad p : B \rightarrow \gamma \in P, w_{j+1} \in \text{First}(\gamma)$$

$$\text{Complete: } \frac{\langle x_1, y_1 \rangle : [A \rightarrow \alpha \bullet B\beta, i, j], \langle x_2, y_2 \rangle : [B \rightarrow \gamma \bullet, j, k]}{\langle x_1 + y_2, y_1 + y_2 \rangle : [A \rightarrow \alpha B \bullet \beta, i, k]}$$

Goal item: $[S, 1, n]$.

Die Gewichte sind wie folgt geordnet:

$\langle x_1, y_1 \rangle < \langle x_2, y_2 \rangle$ genau dann wenn entweder $x_1 < x_2$ oder $x_1 = x_2$ und $y_1 < y_2$.

Betrachten Sie die PCFG $\langle N, T, P, S \rangle$ mit

$N = \{S\}$, $T = \{a, b, c\}$, Startsymbol S , Produktionen

$$0.6 : S \rightarrow aSb, \quad 0.1 : S \rightarrow aSc, \quad 0.1 : S \rightarrow a, \quad 0.2 : S \rightarrow c$$

Es gilt (gerundet) $|\log(0.6)| = 0.22$, $|\log(0.1)| = 1$ und $|\log(0.2)| = 0.7$.

Zeigen Sie in Form einer Tabelle, wie sich Chart und Agenda schrittweise entwickeln, wenn mit dieser Grammatik und dem angegebenen gewichteten Earley Deduktionssystem die Eingabe $w = acb$ geparkt wird. (Ohne Backpointer, also nur Recognition.)

Beachten Sie dabei die Bedingungen bzgl. der First Menge dessen, was rechts vom Dot steht und des nächsten Eingabesymbols.

Es genügt im Prinzip, die schrittweise Entwicklung der Agenda zu zeigen, da sich die Chart unmittelbar hieraus ergibt. (Die Chart enthält alles, was im Laufe des Parsens aus der Agenda entfernt wird.)

Lösung:

