

Grundlegende Parsingalgorithmen

Chart-Parsing: Earley-Parser

Kurt Eberle

k.eberle@lingenio.de

(Viele Folien, Teile von Folien, Materialien von **Helmut Schmid's**
Parsing-Kurs WS14 Tübingen, u.a.)

02. August, 2018

Überblick

Earley-Parser

Earley-Erkenner

Aktionen des Earley-Erkenners

Earley-Parsing mit Look-Ahead

Überblick

Earley-Parser

Earley-Erkennen

Aktionen des Earley-Erkenners

Earley-Parsing mit Look-Ahead

Chart-Parser

- ▶ Alle Parsing-Zwischenergebnisse (d.h. schon gefundene Konstituenten) werden in eine Tabelle (die Chart) eingetragen
- ▶ mit Feldern $Chart[i, j]$ (mit Zeilen i und Spalten j) wobei
- ▶ $A \in Chart[i, j]$ gdw $A \xRightarrow{*} w_{i+1} \dots w_j$
- ▶ Varianten: CYK, Earley, Left-Corner

Earley-Parser

- ▶ Top-down Chart-Parser
- ▶ Chart enthält Punktregeln
 $A \rightarrow \alpha \cdot \beta \in \text{Chart}[i, k],$
gdw $\alpha \xRightarrow{*} w_{i+1} \dots w_k$ und $S \xRightarrow{*} w_1 \dots w_i A \gamma.$
- ▶ Arten von Punktregeln
 - ▶ aktive Regel: $A \rightarrow \alpha \cdot \beta$ mit $\beta \neq \varepsilon$
 - ▶ passive Regel: $A \rightarrow \alpha \cdot$
 - ▶ vorhergesagte Regel: $A \rightarrow \cdot \alpha$

Aktionen des Earley-Parsers

► Aktionen

1. Scan:

If a is the k -th input symbol,
then store $A \rightarrow \alpha a \cdot \beta$ in $Chart[i, k]$
for all $A \rightarrow \alpha \cdot a\beta \in Chart[i, k - 1]$

2. Predict:

If some $A \rightarrow \alpha \cdot B\beta$ is added to $Chart[i, k]$,
then store $B \rightarrow \cdot \gamma$ in $Chart[k, k]$ for all $B \rightarrow \gamma \in P$.

3. Complete:

If $B \rightarrow \gamma \cdot$ is added to $Chart[j, k]$,
then store $A \rightarrow \alpha B \cdot \beta$ in $Chart[i, k]$ for all
 $A \rightarrow \alpha \cdot B\beta \in Chart[i, j]$

- Start: Für alle $S \rightarrow \alpha \in P$, speichere $S \rightarrow \cdot \alpha$ in $Chart[0, 0]$
- Ende: Akzeptiere den String, wenn $S \rightarrow \alpha \cdot \in Chart[0, n]$ für ein bel. α

Überblick

Earley-Parser

Earley-Erkenner

Aktionen des Earley-Erkenners

Earley-Parsing mit Look-Ahead

Earley-Erkenner

Chart-Repräsentation: Liste t_k für jede Input-Position k .
Die Liste enthält Punktregeln der Form $\langle A \rightarrow \alpha \cdot \beta, i, k \rangle$.

Recognize($x_1, \dots, x_n$)

$t_0 := \{\}$

Predict($S, 0$)

For $j := 1$ to n

$t_j := \{\}$

For all $\langle A \rightarrow \alpha \cdot x_j \beta, i, j-1 \rangle \in t_{j-1}$

Store($A \rightarrow \alpha x_j \cdot \beta, i, j$)

If $\langle S \rightarrow \alpha \cdot, 0, n \rangle \in t_n$ for some α then

accept

else

report error

Überblick

Earley-Parser

Earley-Erkenner

Aktionen des Earley-Erkenners

Earley-Parsing mit Look-Ahead

Aktionen des Earley-Erkenners

Store($A \rightarrow \alpha \cdot \beta, i, k$)

If $\langle A \rightarrow \alpha \cdot \beta, i, k \rangle \notin t_k$ then

$t_k := t_k \cup \{ \langle A \rightarrow \alpha \cdot \beta, i, k \rangle \}$

If $\beta = \varepsilon$ then

Complete(A, i, k)

else if $\beta = B\gamma$ for some B, γ then

If $\langle B \rightarrow \delta \cdot, k, k \rangle \in t_k$ then

Store($A \rightarrow \alpha B \cdot \gamma, i, k$)

else

Predict(B, k)

Aktionen des Earley-Erkenners

Predict(A, i)

If $\text{called}[A, i] = \text{false}$ then

$\text{called}[A, i] := \text{true}$

 For all $A \rightarrow \alpha \in P$

Store($A \rightarrow \cdot \alpha, i, i$)

Complete(A, j, k)

If $\text{called}[A, j, k] = \text{false}$ then

$\text{called}[A, j, k] := \text{true}$

 For all $\langle B \rightarrow \beta \cdot A\gamma, i, j \rangle \in t_j$

Store($B \rightarrow \beta A \cdot \gamma, i, k$)

Überblick

Earley-Parser

Earley-Erkenner

Aktionen des Earley-Erkenners

Earley-Parsing mit Look-Ahead

Earley-Parsing mit Look-Ahead

- ▶ Problem: Die Predikt-Funktion führt viele unnötige Punktregeln in die Chart ein.
- ▶ Lösung Look-Ahead Funktion L (analog zum SLR Parsing)
 $L(A \rightarrow \alpha \cdot \beta, a) = \text{true}$ if
 1. $a \in \text{First}(\beta)$ or
 2. $\varepsilon \in \text{First}(\beta)$ and $a \in \text{Follow}(A)$else $L(A \rightarrow \alpha \cdot \beta, a) = \text{false}$
- ▶ Neue Store-Funktion:
Store $(A \rightarrow \alpha \cdot \beta, i, k)$
If $L(A \rightarrow \alpha \cdot \beta, x_{k+1})$ then
 $\text{Store}_{old}(A \rightarrow \alpha \cdot \beta, i, k)$