Formal Languages

Philippe de Groote

2018-2019

Outline

Definition

Let $G = \langle N, \Sigma, P, S \rangle$ be a context-free grammar.

A *dotted rule* is a rule of the form $A \rightarrow \alpha \bullet \beta$ where:

- $A \in N$;
- $\alpha, \beta \in (N \cup \Sigma)^*$;
- $(A \rightarrow \alpha\beta) \in P$.

Given a word to be parsed, $\alpha = a_1 a_2 \ldots a_n$, an *item* is an expression of the form $<A \rightarrow \alpha \bullet \beta, i, j>$ where:

- $A \rightarrow \alpha \bullet \beta$ is a dotted rule;
- $0 \leq i \leq j \leq n$.

CKY-algorthm

Init

$$<A \to \bullet\alpha, i, i>$$

Scan

$$\frac{<A \to \alpha \bullet a\beta, i, j>}{<A \to \alpha a \bullet \beta, i, j+1>} \quad a = a_{j+1}$$

Complete

$$\frac{<A \to \alpha \bullet B\beta, i, j> \quad <B \to \gamma\bullet, j, k>}{<A \to \alpha B \bullet \beta, i, k>}$$

Earley algorithm

Init

$$<S \to \bullet\alpha, 0, 0>$$

Scan

$$\frac{<A \to \alpha \bullet a\beta, i, j>}{<A \to \alpha a \bullet \beta, i, j+1>} \quad a = a_{j+1}$$

Complete

$$\frac{<A \to \alpha \bullet B\beta, i, j> \quad <B \to \gamma\bullet, j, k>}{<A \to \alpha B \bullet \beta, i, k>}$$

Predict

$$\frac{<A \to \alpha \bullet B\beta, i, j>}{<B \to \bullet\gamma, j, j>}$$