Script generated by TTT

Title: Petter: Compiler Construction (04.06.2020) - Canonical LR(1)

Date: Tue May 26 16:32:20 CEST 2020

Duration: 15:19 min

Pages: 7

LR(1)-Parsing

Idea: Let's equip items with 1-lookahead

Definition LR(1)-Item

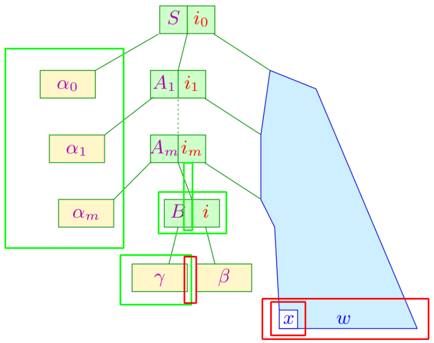
An $LR(1)$-item is a pair $[B \to \alpha \bullet \beta, x]$ with

$$x \in \text{Follow}_1(B) = \bigcup \{\text{First}_1(\nu) \mid S \to^* \mu B \nu\}$$

Admissible LR(1)-Items

The $LR(1)$-Item $[B \to \gamma \bullet \beta, x]$ is *admissable* for $\alpha\,\gamma$ if:

$$S \to_R^* \alpha\, B\, w \qquad \text{with} \qquad \{x\} = \text{First}_1(w)$$



... with $\alpha_0 \dots \alpha_m = \alpha$

The Characteristic LR(1)-Automaton

The set of admissible $LR(1)$-items for viable prefixes is again computed with the help of the finite automaton $c(G, 1)$.

The automaton $c(G, 1)$:

States: $LR(1)$-items

Start state: $[S' \to \bullet S, \$]$

Final states: $\{[B \to \gamma \bullet, x] \mid B \to \gamma \in P, x \in \text{Follow}_1(B)\}$

Transitions:
(1) $([A \to \alpha \bullet X \beta, x], X, [A \to \alpha X \bullet \beta, x]), \quad X \in (N \cup T)$
(2) $([A \to \alpha \bullet B \beta, x], \epsilon, [B \to \bullet \gamma, x']), \quad A \to \alpha B \beta, B \to \gamma \in P,$
$x' \in \text{First}_1(\beta) \odot_1 \{x\}$

The Characteristic LR(1)-Automaton

The set of admissible $LR(1)$-items for viable prefixes is again computed with the help of the finite automaton $c(G,1)$.

The automaton $c(G,1)$:

States: $LR(1)$-items

Start state: $[S' \to \bullet\, S,\ \$]$

Final states: $\{[B \to \gamma \bullet,\ x] \mid B \to \gamma \in P, x \in \text{Follow}_1(B)\}$

Transitions:

(1) $([A \to \alpha \bullet X\,\beta,\ x], X, [A \to \alpha\, X \bullet \beta,\ x]),\quad X \in (N \cup T)$

(2) $([A \to \alpha \bullet B\,\beta,\ x], \epsilon,\ [B \to \bullet\, \gamma,\ x']),\quad A \to \alpha\, B\, \beta\,,\ B \to \gamma \in P,$ $x' \in \text{First}_1(\beta) \odot_1 \{x\}$

This automaton works like $c(G)$ — but additionally manages a 1-prefix from Follow_1 of the left-hand sides.

The Canonical LR(1)-Automaton

The canonical $LR(1)$-automaton $LR(G,1)$ is created from $c(G,1)$, by performing arbitrarily many ϵ-transitions and then making the resulting automaton deterministic ...

The Canonical LR(1)-Automaton

The canonical $LR(1)$-automaton $LR(G,1)$ is created from $c(G,1)$, by performing arbitrarily many ϵ-transitions and then making the resulting automaton deterministic ...

But again, it can be constructed directly from the grammar; analoguously to $LR(0)$, we need the ϵ-closure δ_ϵ^* as a helper function:

$$\delta_\epsilon^*(q) = q \cup \{[C \to \bullet\, \gamma,\ x] \mid [A \to \alpha \bullet B\,\beta',\ x'] \in q,\quad B \to^* C\,\beta,\quad C \to \gamma \in P,\ x \in \text{First}_1(\beta\,\beta') \odot_1 \{x'\}\}$$

The Canonical LR(1)-Automaton

The canonical $LR(1)$-automaton $LR(G,1)$ is created from $c(G,1)$, by performing arbitrarily many ϵ-transitions and then making the resulting automaton deterministic ...

But again, it can be constructed directly from the grammar; analoguously to $LR(0)$, we need the ϵ-closure δ_ϵ^* as a helper function:

$$\delta_\epsilon^*(q) = q \cup \{[C \to \bullet\, \gamma,\ x] \mid [A \to \alpha \bullet B\,\beta',\ x'] \in q,\quad B \to^* C\,\beta,\quad C \to \gamma \in P,\ x \in \text{First}_1(\beta\,\beta') \odot_1 \{x'\}\}$$

Then, we define:

States: Sets of $LR(1)$-items;

Start state: $\delta_\epsilon^*\,\{[S' \to \bullet\, S,\ \$]\}$

Final states: $\{q \mid [A \to \alpha \bullet,\ x] \in q\}$

Transitions: $\delta(q, X) = \delta_\epsilon^*\{[A \to \alpha\, X \bullet \beta,\ x] \mid [A \to \alpha \bullet X\,\beta,\ x] \in q\}$