Script generated by TTT

Title: Petter: Compiler Construction (04.06.2020) - LR(k) Grammars

Date: Tue May 26 15:14:27 CEST 2020

Duration: 33:36 min

Pages: 20

LR(0)-Parser

Attention:
Unfortunately, the $LR(0)$-parser is in general non-deterministic.

We identify two reasons for a state $q \in Q$:

Reduce-Reduce-Conflict:

q: $A \to \gamma \bullet$, $A' \to \gamma' \bullet$ with $A \neq A' \vee \gamma \neq \gamma'$

Shift-Reduce-Conflict:

q: $A \to \gamma \bullet$, $A' \to \alpha \bullet a \beta$ with $a \in T$

Those states are called $LR(0)$-unsuited.

Revisiting the Conflicts of the LR(0)-Automaton

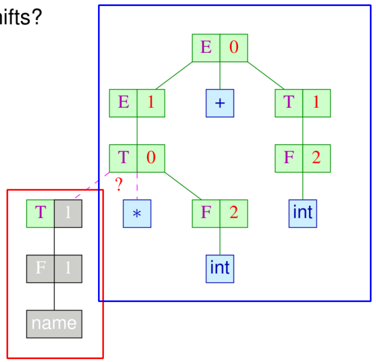
What differenciates the particular Reductions and Shifts?

Input:

$* \; 2 + 40$

Pushdown:

$(q_0 \; T)$



$$
\begin{array}{lcl}
E & \to & E + T \mid T \\
T & \to & T * F \mid F \\
F & \to & (\, E \,) \mid \text{int}
\end{array}
$$

Revisiting the Conflicts of the LR(0)-Automaton

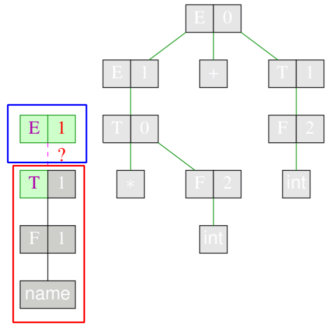
What differenciates the particular Reductions and Shifts?

Input:

$* \; 2 + 40$

Pushdown:

$(q_0 \; T)$



$$
\begin{array}{lcl}
E & \to & E + T \mid T \\
T & \to & T * F \mid F \\
F & \to & (\, E \,) \mid \text{int}
\end{array}
$$

Revisiting the Conflicts of the LR(0)-Automaton

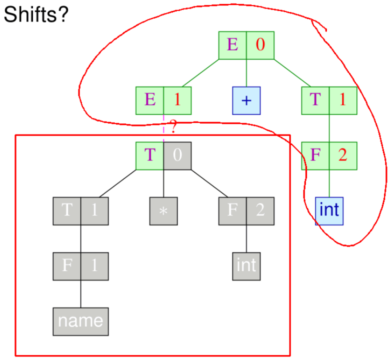
What differenciates the particular Reductions and Shifts?

Input:

$+ \; 40$

Pushdown:

$(\; q_0 \; T \;)$



$$
\begin{array}{lcll}
E & \rightarrow & E+T & | \; T \\
T & \rightarrow & T*F & | \; F \\
F & \rightarrow & (\,E\,) & | \; \text{int}
\end{array}
$$

Revisiting the Conflicts of the LR(0)-Automaton

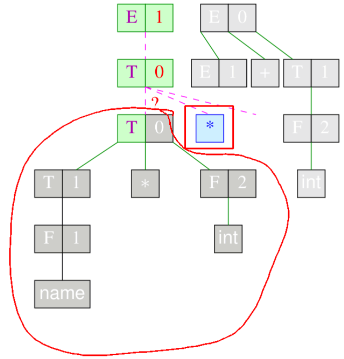
What differenciates the particular Reductions and Shifts?

Input:

$+ \; 40$

Pushdown:

$(\; q_0 \; T \;)$



$$
\begin{array}{lcll}
E & \rightarrow & E+T & | \; T \\
T & \rightarrow & T*F & | \; F \\
F & \rightarrow & (\,E\,) & | \; \text{int}
\end{array}
$$

Revisiting the Conflicts of the LR(0)-Automaton

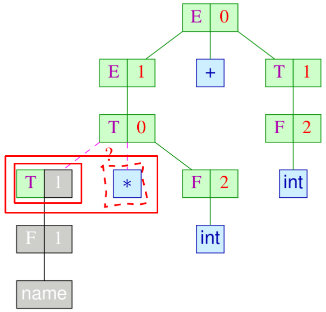
Idea: In reverse rightmost derivations, *right context* determines derivations!

Input:

$* \, 2 + 40$

Pushdown:

$(\; q_0 \; T \;)$



$$
\begin{array}{lcll}
E & \rightarrow & E+T & | \; T \\
T & \rightarrow & T*F & | \; F \\
F & \rightarrow & (\,E\,) & | \; \text{int}
\end{array}
$$

Revisiting the Conflicts of the LR(0)-Automaton

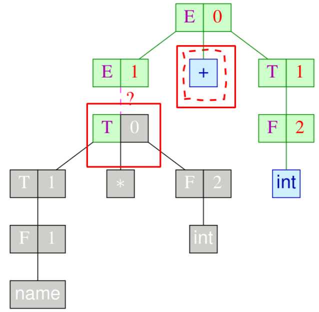
Idea: In reverse rightmost derivations, *right context* determines derivations!

Input:

$+ \, 40$

Pushdown:

$(\; q_0 \; T \;)$



$$
\begin{array}{lcll}
E & \rightarrow & E+T & | \; T \\
T & \rightarrow & T*F & | \; F \\
F & \rightarrow & (\,E\,) & | \; \text{int}
\end{array}
$$

LR(k)-Grammars

Idea: Consider k-lookahead in conflict situations.

Definition:
The reduced contextfree grammar G is called $LR(k)$-grammar, if
$\alpha\,\beta\,w|_{|\alpha\beta|+k} = \alpha'\,\beta'\,w'|_{|\alpha\beta|+k}$ with:

$$\left.\begin{array}{lllll} S & \to_R^* & \alpha\,A\,w & \to & \alpha\,\beta\,w \\ S & \to_R^* & \alpha'\,A'\,w' & \to & \alpha'\,\beta'\,w' \end{array}\right\} \text{follows: } \alpha = \alpha' \;\wedge\; \beta = \beta' \;\wedge\; A = A'$$

LR(k)-Grammars

Idea: Consider k-lookahead in conflict situations.

Definition:
The reduced contextfree grammar G is called $LR(k)$-grammar, if
$\alpha\,\beta\,w|_{|\alpha\beta|+k} = \alpha'\,\beta'\,w'|_{|\alpha\beta|+k}$ with:

$$\left.\begin{array}{lllll} S & \to_R^* & \alpha\,A\,w & \to & \alpha\,\beta\,w \\ S & \to_R^* & \alpha'\,A'\,w' & \to & \alpha'\,\beta'\,w' \end{array}\right\} \text{follows: } \alpha = \alpha' \;\wedge\; \beta = \beta' \;\wedge\; A = A'$$

Strategy for testing Grammars for $LR(k)$-property

1. Focus iteratively on all rightmost derivations $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$
2. Iterate over $k \geq 0$
   1. For each $\gamma = \alpha\,\beta w|_{|\alpha\beta|+k}$ (*handle with k-lookahead*) check if there exists a differently right-derivable $\alpha'\,\beta'\,w'$ for which $\gamma = \alpha'\,\beta'\,w'|_{|\alpha\beta|+k}$
   2. if there is none, we have found no objection against k being enough lookahead to disambiguate $\alpha\beta w$ from other rightmost derivations

LR(k)-Grammars

for example:

(1) $\quad S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

LR(k)-Grammars

for example:

(1) $\quad S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... is not $LL(k)$ for any k :

Let $\;S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w\;$. Then $\;\alpha\,\underline{\beta}\;$ is of one of these forms:

LR(k)-Grammars

for example:

(1) $S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... is not $LL(k)$ for any k :

Let $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Then $\alpha\,\underline{\beta}\,w$ is of one of these forms:

$$\underline{A}\,,\ \underline{B}\,,\ a^n\,\underline{a\,A\,b}\,,\ a^n\,\underline{a\,B\,b\,b}\,,\ a^n\,\underline{0}\,,\ a^n\,\underline{1} \qquad (n \geq 0)$$

LR(k)-Grammars

for example:

(1) $S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... is not $LL(k)$ for any k — but $LR(0)$:

Let $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Then $\alpha\,\underline{\beta}$ is of one of these forms:

$$\underline{A}\,,\ \underline{B}\,,\ a^n\,\underline{a\,A\,b}\,,\ a^n\,\underline{a\,B\,b\,b}\,,\ a^n\,\underline{0}\,,\ a^n\,\underline{1} \qquad (n \geq 0)$$

LR(k)-Grammars

for example:

(1) $S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... is not $LL(k)$ for any k — but $LR(0)$:

Let $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Then $\alpha\,\underline{\beta}$ is of one of these forms:

$$\underline{A}\,,\ \underline{B}\,,\ a^n\,\underline{a\,A\,b}\,,\ a^n\,\underline{a\,B\,b\,b}\,,\ a^n\,\underline{0}\,,\ a^n\,\underline{1} \qquad (n \geq 0)$$

(2) $S \to a\,A\,c \qquad A \to A\,b\,b \mid b$

... is also not $LL(k)$ for any k :

Let $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Then $\alpha\,\underline{\beta}$ is of one of these forms:

LR(k)-Grammars

for example:

(1) $S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... is not $LL(k)$ for any k — but $LR(0)$:

Let $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Then $\alpha\,\underline{\beta}$ is of one of these forms:

$$\underline{A}\,,\ \underline{B}\,,\ a^n\,\underline{a\,A\,b}\,,\ a^n\,\underline{a\,B\,b\,b}\,,\ a^n\,\underline{0}\,,\ a^n\,\underline{1} \qquad (n \geq 0)$$

(2) $S \to a\,A\,c \qquad A \to A\,b\,b \mid b$

... is also not $LL(k)$ for any k :

Let $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Then $\alpha\,\underline{\beta}$ is of one of these forms:

$$a\,\underline{b}\,,\ a\,\underline{A\,b\,b}\,,\ \underline{a\,A\,c}$$

LR(k)-Grammars

for example:

(1) $S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... is not $LL(k)$ for any k — but $LR(0)$:

Let $S \to^*_R \alpha\, X\, w \to \alpha\, \beta\, w$. Then $\alpha\, \underline{\beta}$ is of one of these forms:

$$\underline{A}\,,\ \underline{B}\,,\ a^n\,\underline{a\,A\,b}\,,\ a^n\,\underline{a\,B\,b\,b}\,,\ a^n\,\underline{0}\,,\ a^n\,\underline{1} \qquad (n \geq 0)$$

(2) $S \to a\,A\,c \qquad A \to A\,b\,b \mid b$

... is also not $LL(k)$ for any k — but again $LR(0)$:

Let $S \to^*_R \alpha\, X\, w \to \alpha\, \beta\, w$. Then $\alpha\, \underline{\beta}$ is of one of these forms:

$$a\,\underline{b}\,,\ a\,\underline{A\,b\,b}\,,\ \underline{a\,A\,c}$$

LR(k)-Grammars

for example:

(3) $S \to a\,A\,c \qquad A \to b\,b\,A \mid b$

LR(k)-Grammars

for example:

(3) $S \to a\,A\,c \qquad A \to b\,b\,A \mid b$

Let $S \to^*_R \alpha\, X\, w \to \alpha\, \beta\, w$ with $\{y\} = \text{First}_k(w)$ then $\alpha\, \underline{\beta}\, y$ is of one of these forms:

$$a\,b^{2n}\,\underline{b}\,c\,,\ a\,b^{2n}\,\underline{b\,b\,A}\,c\,,\ \underline{a\,A\,c}$$

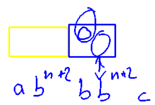
LR(k)-Grammars

for example:

(3) $S \to a\,A\,c \qquad A \to b\,b\,A \mid b$ ... is not $LR(0)$, but $LR(1)$:

Let $S \to^*_R \alpha\, X\, w \to \alpha\, \beta\, w$ with $\{y\} = \text{First}_k(w)$ then $\alpha\, \underline{\beta}\, y$ is of one of these forms:

$$a\,b^{2n}\,\underline{b}\,c\,,\ a\,b^{2n}\,\underline{b\,b\,A}\,c\,,\ \underline{a\,A\,c}$$



for example:

(3) $S \to a\,A\,c \qquad A \to b\,b\,A \;|\; b$ ... is not $LR(0)$, but $LR(1)$:

Let $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$ with $\{y\} = \text{First}_k(w)$ then $\alpha\,\underline{\beta}\,y$ is of one of these forms:

$$a\,b^{2n}\,\underline{b}\,c\;,\;a\,b^{2n}\,\underline{b\,b\,A}\,c\;,\;\underline{a\,A\,c}$$

(4) $S \to a\,A\,c \qquad A \to b\,A\,b \;|\; b$ ... is not $LR(k)$ for any $k \geq 0$:

Consider the rightmost derivations:

$$S \to_R^* a\,b^n\,A\,b^n\,c \to a\,b^n\,\underline{b}\,b^n\,c$$