LALR Table Computation

(Ref: *Efficient Computation of LALR (1) Look-Ahead Sets*, by F. DeRemer and T. Pennello, ACM TOPLAS, Oct. 1982)

To compute:

$LA(q, A \to w) =$

set of terminals t that are lookaheads for $A \to w\cdot$ in state q

Define:

follow$(p, A) =$

$\{t \in \Sigma | S \stackrel{*}{\Rightarrow} \alpha A t z$ and α takes the parser to state $p\}$

Then:

$LA(q, A \to w) =$

$\bigcup$ { follow (p, A) | there is a transition for A in state p and a sequence of transitions $p \to \stackrel{w}{\to} \to q$ }

To compute follow (p, A):

follow $(p, A) =$ read $(p, A) \cup$
$\bigcup$ { follow (p', B) | (p, A) includes (p', B) }
$= \bigcup$ { $DR(r, C)$ | (p, A) includes* $\circ$ reads* (r, C) }

where

I. $DR(r, C)$ (direct read) is computed by inspecting the $LR(0)$ parser.

For any $t \epsilon \Sigma$, $t \epsilon DR(r, C)$ if there is a transition for nonterminal C in state r to state r' and an action for t in state r'

$$r \xrightarrow{C} r' \xrightarrow{t}$$

II. (p, A) reads (r, C) if there is a transition for nonterminal A in state p to state r and a transition for nonterminal C in state r, where $C \stackrel{*}{\Rightarrow} \epsilon$.

(i.e., whatever terminals t follow C starting from state r also follow A starting from state p)

III. For any $t \epsilon \Sigma$, $t \epsilon$ read (p, A) if there is a sequence of transitions (p, A) reads (r_o, C_1), (r_o, C_1) reads (r_1, C_2), $\ldots (r_{n-2}, C_{n-1})$ reads (r_{n-1}, C_n) and $t \epsilon DR(r_{n-1}, C_n)$, where for $1 \leq i \leq n$, $C_i \stackrel{*}{\Rightarrow} \epsilon$

$$p \xrightarrow{A} r_o \xrightarrow{C_1 \stackrel{*}{\Rightarrow} \epsilon} r_1 \xrightarrow{C_2 \stackrel{*}{\Rightarrow} \epsilon} \ldots \longrightarrow r_{n-1} \xrightarrow{C_n \stackrel{*}{\Rightarrow} \epsilon} r_n \xrightarrow{t}$$

The Dragon book calls these t's *propagated lookaheads*.

IV. If the grammar has a rule $B \to \beta A \gamma$ and derivations $\gamma \stackrel{*}{\Rightarrow} \epsilon$, where there is a sequence of transitions starting in state p' and reaching state p by shifting over the symbols of β, there is a transition for B from state p', and there is a transition for A from state p, then
(p, A) includes (p', B)
(i.e., follow $(p', B) \subseteq$ follow (p, A))

The Dragon book calls these t's *spontaneous lookaheads*.

Summary of LALR Computation

compute	*NULL*
compute	*DR*
compute	*reads*
compute	*reads* and then read*
compute	*includes*
compute	*includes* and then follow*
compute	*LA sets from follow sets*

check states for conflicts