Monday, August 22, 2022 11:41 AM

last time: predictive parsing by example

Today. FIRST and FOLLOW sets

Predictive parsing with FIRST and FOLLOW.

Definitions.

FIRST set: FIRST sets will be calculated for terminals and non-terminals & ε

Given a grammar G, with terminals set T, FIRST sets can contain elements of T or ε

1. Terminal $a \in$ FIRST(A) where A is a terminal or non-terminal iff

in 0 or more steps

$$A \overset{*}{\Rightarrow} a\, w$$

at the start

where w is a sequence of terminals/non-terminals

2. $\varepsilon \in$ FIRST(A) iff

$$A \overset{*}{\Rightarrow} \varepsilon$$

FOLLOW sets We will compute FOLLOW sets for non-terminals.

FOLLOW sets can contain terminals or $\$$ = EOF

a, which is a terminal or $\$$ is in FOLLOW(A) iff.

$$S\$ \overset{*}{\Rightarrow} x\, A\, a\, y$$

where x and y are sequences of terminals/non-terminals

NOTE ε cannot be part of a FOLLOW set.

$\$$ cannot be part of a FIRST set.

NOTE → ε cannot be part of a FOLLOW set.

$ cannot be part of a FIRST set.

Example

$$
\begin{aligned}
S &\to AB \mid C \\
A &\to aA \mid BC \\
B &\to bB \mid \varepsilon \\
C &\to cC \mid D \\
D &\to d
\end{aligned}
$$

$a \in FIRST(A)$: $A \Rightarrow aA$

$b \in FIRST(B)$: $B \Rightarrow bB$

$b \in FIRST(A)$: $A \Rightarrow BC \Rightarrow bBC$

$c \in FIRST(A)$: $A \Rightarrow BC \Rightarrow C \Rightarrow cC$

$d \in FIRST(A)$: $A \Rightarrow BC \Rightarrow C \Rightarrow D \Rightarrow d$

$\varepsilon \notin FIRST(A)$

$FIRST(C) = \{ \cancel{a}, \cancel{b}, c ✓, d ✓, \cancel{\varepsilon} \}$

$$
\begin{aligned}
S &\to AB \mid C \\
A &\to aA \mid BC \\
B &\to bB \mid \varepsilon \\
C &\to cC \mid D \\
D &\to d
\end{aligned}
$$

FOLLOW

$S\$ \overset{*}{\Rightarrow} \ldots Aa \ldots$

FOLLOW(S) $\quad S\$ \overset{*}{\Rightarrow} S\$$

$b \in FOLLOW(A)$ $\quad S\$ \overset{*}{\Rightarrow} AB\$ \Rightarrow AbB\$$

$\$ \in FOLLOW(A)$ $\quad S\$ \overset{*}{\Rightarrow} AB\$ \Rightarrow A\$$

$\$ \in FOLLOW(B)$ $\quad S\$ \Rightarrow AB\$$

$c \in FOLLOW(B)$ $\quad S\$ \Rightarrow AB\$ \Rightarrow BCB\$$

$\Rightarrow BcCB\$$

$d \in FOLLOW(B)$ $\quad S\$ \Rightarrow AB\$ \Rightarrow BCB\$$

$\Rightarrow BdB\$$

$b \notin FOLLOW(B)$? $\quad S\$ \Rightarrow AB\$ \Rightarrow BCB\$$

$b \notin FOLLOW(B)$? $S\$ \Rightarrow AB\$ \Rightarrow BCB\$$

$$
\begin{aligned}
S &\rightarrow AB \mid C \\
A &\rightarrow aA \mid BC \\
B &\rightarrow bB \mid \varepsilon \\
C &\rightarrow cC \mid D \\
D &\rightarrow d
\end{aligned}
$$

Conditions for Predictive Parsing with one token lookahead.

1. For every non-terminal A and any two rules $A \rightarrow \alpha$ and $A \rightarrow \beta$,
   we should have $FIRST(\alpha) \cap FIRST(\beta) = \phi$

Note. $FIRST(A_1 \ldots A_k)$?
If we introduce a new symbol A' and rule

$$A' \rightarrow A_1 \ldots A_k .$$

$$FIRST(A') = FIRST(A_1 \ldots A_k)$$

2. If $A \overset{*}{\Rightarrow} \varepsilon$, we should have
   $FIRST(A) \cap FOLLOW(A) = \phi$

Example.

$$
\begin{aligned}
S &\rightarrow AB \\
A &\rightarrow bAc \mid \varepsilon \\
B &\rightarrow b
\end{aligned}
$$

```
parse-A()
{     t = peek(1);
      if (t.token_type == b_type)
      {       expect(b) ?
              parse(A)
```

$\underbrace{bc}_{A}\ \underbrace{b}_{B}$

$\underbrace{}_{A}\ \underbrace{b}_{B}$

{ expect (b)
parse (A)
expect (c) }
}

$A \to \varepsilon$

Example.

$S \to AeB \mid gC$

$A \to aA \mid BC \mid fA$

$B \to bB \mid \varepsilon$

$C \to cC \mid D \mid \varepsilon$

$D \to d$

$FIRST(S) = \{a, b, c, d, f, e, g\}$

$FIRST(A) = \{a, b, c, d, \varepsilon, f\}$

$FIRST(B) = \{b, \varepsilon\}$

$FIRST(C) = \{c, d, \varepsilon\}$

$FIRT(D) = \{d\}$

$FOLLOW(S) = \{\$$

$FOLLOW(A) = \{e,$

$FOLLOW(B) = \{c, d, e, \$$

$FOLLOW(C) = \{e, \$$

$FOLLOW(D) = \{e, \$$

$S\$ \Rightarrow AeB\$ \Rightarrow \cancel{B}CeB\$ \Rightarrow \cancel{B}eB\$$

$\Rightarrow BDeB\$$

$S\$ \Rightarrow gC\$ \Rightarrow gD\$$