

How can we know if a  $G(CFG)$  is  $LL(1)$ ?

First, we should separate the

$$FIRST(X), FOLLOW(X), SELECT(X \rightarrow \alpha)$$

sets of this  $G(CFG)$ .

Tips: all the  $FIRST(X)$  is the most left terminal state of the productions of  $X$ .

We use the symbol  $\alpha$  as the production of  $X$ .

To separate the  $FIRST(X)$  sets, there are 3 rules:

1. If  $X$  the production of  $X$  is a terminal state, then  $FIRST(X) = \alpha$ .
2. If  $\alpha$  is a nonterminal state, and  $FIRST(X) \rightarrow \alpha_1\alpha_2\alpha_3...\alpha_k$  ( $k \geq 1$ )
  - 2.1 If  $\alpha_k$  not include  $\varepsilon$ , then  $FIRST(X) = FIRST(\alpha)$ , until the next nonterminal state  $\alpha$  has a production of terminal state.
  - 2.2 If  $\alpha_k$  include  $\varepsilon$ , then  $FIRST(X) = FIRST(\alpha_k) - \{\varepsilon\} + FIRST(\alpha_{k+1})$ , until the next nonterminal state has a production of terminal state or  $\alpha_{k+1}$  is a terminal state.
3. If  $FIRST(X) \rightarrow \varepsilon$ , add the  $\{\varepsilon\}$  into the  $FIRST(X)$ .

For example, in the  $G(A)$ :

$$A \rightarrow BCc \mid gDB$$

$$B \rightarrow bCDE \mid \varepsilon$$

$$C \rightarrow DaB \mid ca$$

$$D \rightarrow dD \mid \varepsilon$$

$$E \rightarrow gAf \mid c$$

$$FIRST(A) = FIRST(B) - \{\varepsilon\} + FIRST(C) + g = \{a, b, c, d, g\}$$

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

$$FIRST(C) = FIRST(D) - \{\varepsilon\} + a + c = \{a, c, d\}$$

$$FIRST(D) = \{d, \varepsilon\}$$

$$FIRST(E) = \{c, g\}$$

To separate the  $FOLLOW(X)$  sets, there are also 5 rules:

1. If  $X$  is the initial state, add the terminal symbol  $Z_0$  into  $FOLLOW(X)$ .
2. Find all the  $X$ s in the productions (right side)
  - 2.1 If  $X$  is in the end of the production, add  $FOLLOW(A)$  into  $FOLLOW(X)$
  - 2.2 If the nonterminal state  $\alpha$  next to  $X$  not include  $\varepsilon$ , add  $FIRST(\alpha)$  into  $FOLLOW(X)$ .
  - 2.3 If the nonterminal state  $\alpha$  next to  $X$  include  $\varepsilon$ , add  $FIRST(\alpha_k) - \{\varepsilon\}$  into  $FOLLOW(X)$  until there is no more nonterminal states or reach a terminal state. If it reaches the end of all the productions, you have to also add the  $FOLLOW(A)$  into  $FOLLOW(X)$  as the rule in 2.1.

For example, in the  $G(A)$ :

$$A \rightarrow BCc \mid gDB$$

$$B \rightarrow bCDE \mid \varepsilon$$

$$C \rightarrow DaB \mid ca$$

$$D \rightarrow dD \mid \varepsilon$$

$$E \rightarrow gAf \mid c$$

$$FOLLOW(A) = f + Z_0 = \{f, Z_0\}$$

$$FOLLOW(B) = FIRST(C) + FOLLOW(A) + FOLLOW(C) = \{a, c, d, g, f, Z_0\}$$

$$FOLLOW(C) = c + FIRST(D) - \{\varepsilon\} + FIRST(E) = \{c, d, g\}$$

$$FOLLOW(D) = FIRST(B) - \{\varepsilon\} + FOLLOW(A) + a + FIRST(E) = \{a, b, c, g, f, Z_0\}$$

$$FOLLOW(E) = FOLLOW(B) = \{a, c, d, g, f, Z_0\}$$

To get the  $SELECT(A \rightarrow X)$ , we should get all the  $FIRST(X)$ s and  $FOLLOW(X)$ s from the last step, there are 2 rules to get the  $SELECT(A \rightarrow X)$  sets:

1. Suppose we have two productions of a nonterminal state:

$$A \rightarrow \alpha$$

$$A \rightarrow \beta$$

One of the necessary and sufficient conditions to prove a  $LL(1)$  grammar is, to two productions of a nonterminal state,  $A \rightarrow \alpha$ ,  $A \rightarrow \beta$ , must be met

$$SELECT(A \rightarrow \alpha) \cap SELECT(A \rightarrow \beta) = \emptyset$$

Tips: the productions  $\alpha$  and  $\beta$  can't be both  $\varepsilon$ .

2. Calculate each  $FIRST(\alpha)$  of each production:

in the  $G(A)$

$$A \rightarrow BCc \mid gDB$$

$$B \rightarrow bCDE \mid \varepsilon$$

$$C \rightarrow DaB \mid ca$$

$$D \rightarrow dD \mid \varepsilon$$

$$E \rightarrow gAf \mid c$$

$$FIRST(BCc) = FIRST(B) - \{\varepsilon\} + FIRST(C) = \{a, b, c, d\}$$

$$FIRST(gDB) = \{g\}$$

$$FIRST(bCDE) = \{b\}$$

$$FIRST(\varepsilon) = \{\varepsilon\}$$

$$FIRST(DaB) = FIRST(D) - \{\varepsilon\} + a = \{a, d\}$$

$$FIRST(ca) = \{c\}$$

$$FIRST(dD) = \{d\}$$

$$FIRST(gAf) = \{g\}$$

$$FIRST(c) = \{c\}$$

3. Get the  $SELECT(A \rightarrow \alpha)$  sets

- 3.1 If  $\alpha \neq \varepsilon$  then

$$SELECT(A \rightarrow \alpha) = FIRST(\alpha)$$

3.2 If  $\alpha \rightarrow \varepsilon$  then

$$SELECT(A \rightarrow \alpha) = (FIRST(\alpha) - \{\varepsilon\}) \cup FOLLOW(A)$$

For example, in the  $G(A)$ :

$$A \rightarrow BCc \mid gDB$$

$$B \rightarrow bCDE \mid \varepsilon$$

$$C \rightarrow DaB \mid ca$$

$$D \rightarrow dD \mid \varepsilon$$

$$E \rightarrow gAf \mid c$$

$$SELECT(A \rightarrow BCc) = \{a, b, c, d\}$$

$$SELECT(A \rightarrow gDB) = \{g\}$$

$$SELECT(B \rightarrow bCDE) = \{b\}$$

$$SELECT(B \rightarrow \varepsilon) = (FIRST(\varepsilon) - \{\varepsilon\}) \cup FOLLOW(A) = \{a, c, d, g, f, Z_0\}$$

$$SELECT(C \rightarrow DaB) = \{a, d\}$$

$$SELECT(C \rightarrow ca) = \{c\}$$

$$SELECT(D \rightarrow dD) = \{d\}$$

$$SELECT(D \rightarrow \varepsilon) = (FIRST(\varepsilon) - \{\varepsilon\}) \cup FOLLOW(D) = \{a, b, c, g, f, Z_0\}$$

$$SELECT(E \rightarrow gAf) = \{g\}$$

$$SELECT(E \rightarrow c) = \{c\}$$

And we calculate the intersections of all the  $SELECT(A \rightarrow \alpha)$ s from the same nonterminal:

$$SELECT(A \rightarrow BCc) \cap SELECT(A \rightarrow gDB) = \{a, b, c, d\} \cap \{g\} = \emptyset$$

$$SELECT(B \rightarrow bCDE) \cap SELECT(B \rightarrow \varepsilon) = \{b\} \cap \{a, c, d, g, f, Z_0\} = \emptyset$$

$$SELECT(C \rightarrow DaB) \cap SELECT(C \rightarrow ca) = \{a, d\} \cap \{c\} = \emptyset$$

$$SELECT(D \rightarrow dD) \cap SELECT(D \rightarrow \varepsilon) = \{d\} \cap \{a, b, c, g, f, Z_0\} = \emptyset$$

$$SELECT(E \rightarrow gAf) \cap SELECT(E \rightarrow c) = \{g\} \cap \{c\} = \emptyset$$

As can be seen from above, all the intersections are empty, so the grammar is a LL(1) grammar.