

The construction of both top-down and bottom-up parsers is aided by two functions  $FIRST()$  and  $FOLLOW()$ .

Define  $FIRST(\alpha)$ , where  $\alpha$  is any string of grammar symbols, to be the set of **terminals** that begin strings derived from  $\alpha$ . If  $\alpha \Rightarrow^* \epsilon$ , then  $\epsilon$  is also in  $FIRST(\alpha)$ .

To compute  $FIRST(X)$

1. If  $X$  is a terminal, then  $FIRST(X) = \{X\}$
2. If  $X$  is a nonterminal and  $X \rightarrow Y_1 Y_2 \dots Y_k$  is a production, then place  $a$  in  $FIRST(X)$  if for some  $i$ ,  $a$  is in  $FIRST(Y_i)$ , and  $\epsilon$  is all of  $FIRST(Y_1), \dots, FIRST(Y_{i-1})$ ; that is:  $Y_1 \dots Y_{i-1} \Rightarrow^* \epsilon$ . If  $\epsilon$  is in  $FIRST(Y_j)$  for all  $j = 1, 2, \dots, k$ , then add  $\epsilon$  to  $FIRST(X)$ . For example, everything in  $FIRST(Y_1)$  is surely in  $FIRST(X)$ . If  $Y_1$  does not derive  $\epsilon$ , then we add nothing more to  $FIRST(X)$  but if  $Y_1 \Rightarrow^* \epsilon$ , then we add  $FIRST(Y_2)$ , and so on.
3. If  $X \rightarrow \epsilon$  is a production, then add  $\epsilon$  to  $FIRST(X)$ .

To compute  $FOLLOW(A)$

1. Place  $\$$  in  $FOLLOW(S)$  where  $S$  is the start symbol, and  $\$$  is the input right endmarker.
2. If there is a production  $A \rightarrow \alpha B \beta$ , then everything in  $FIRST(\beta)$  except  $\epsilon$  is in  $FOLLOW(B)$ .
3. If there is a production  $A \rightarrow \alpha B$ , or a production  $A \rightarrow \alpha B \beta$ , where  $FIRST(\beta)$  contains  $\epsilon$ , then everything in  $FOLLOW(A)$  is in  $FOLLOW(B)$ .

Consider the grammar.

|   | <i>left<br/>side</i> | <i>right side</i>       |
|---|----------------------|-------------------------|
| 1 | $E$                  | $\rightarrow E + T$     |
| 2 | $E$                  | $\rightarrow T$         |
| 3 | $T$                  | $\rightarrow T^* F$     |
| 4 | $T$                  | $\rightarrow F$         |
| 5 | $F$                  | $\rightarrow (E)$       |
| 6 | $F$                  | $\rightarrow \text{id}$ |

Table 1. Set of productions expressions

| Step | Action                     | Justification                                |
|------|----------------------------|--|
| 1    | $FIRST(E) = FIRST(T)$      | FIRST, rule 2                                |
| 2    | $FIRST(T) = FIRST(F)$      | FIRST, rule 2                                |
| 3    | $FIRST(F) = \{\text{id}\}$ | FIRST, rule 1 applied to productions 5 and 6 |

| Step | Action   | Justification                       |
|------|--|-------------------------------------|
| 1    | $FOLLOW(E) = \$$   | FOLLOW, rule 1                      |
| 2    | $FOLLOW(E) = FOLLOW(E) \cup \{\text{id}\} = \{\$, \text{id}\}$ | FOLLOW, rule 2, productions 1 and 5 |
| 3    | $FOLLOW(T) = FOLLOW(E) = \{\$, +, \}\}$                        | FOLLOW, rule 2, production 2        |
| 4    | $FOLLOW(T) = FOLLOW(T) \cup \{*\} = \{\$, *, +, \}\}$          | FOLLOW, rule 2, production 3        |
| 5    | $FOLLOW(F) = FOLLOW(T) = \{\$, *, +, \}\}$                     | FOLLOW, rule 3, production 4        |

|     | $FIRST()$    | $FOLLOW()$          |
|-----|--------------|---------------------|
| $E$ | $\{(), id\}$ | $\{\$, +, )\}$      |
| $F$ | $\{(), id\}$ | $\{\$, +, *, )\}$   |
| $T$ | $\{(), id\}$ | $\{\$, +, *, , )\}$ |