Chapter 4
Syntax Analysis

Topics

- FIRST and FOLLOW Sets
- Conditions of Predictive Parser

Predictive Parser

- Aided by two functions associated with a grammar.
- FIRST and FOLLOW allow to fill in the entries of a predictive parsing table.
- FIRST(α) set of a string of grammar symbols α is the set of terminals that begin the strings derived from α.
- FOLLOW(A) set of a nonterminal A is the set of terminals that can appear immediately to the right of A in some sentential form

FIRST Set

FIRST(X), all grammar symbols X, apply 1-3 until no more terminals or ε can be added to any FIRST set.

1. If X is terminal, FIRST(X) is {X}
2. If X $\rightarrow$ ε is a production, add ε to FIRST(X)
3. If X is nonterminal and X $\rightarrow$ $Y_1\ Y_2\ \ldots\ Y_k$ is a production, place *a* in FIRST(X) if for some *i*, *a* is in FIRST(Y_i), and ε is in all of FIRST(Y_1), …, FIRST(Y_{i-1}). If ε is in FIRST(Y_j) for all j=1,2,…k, add ε to FIRST(X).

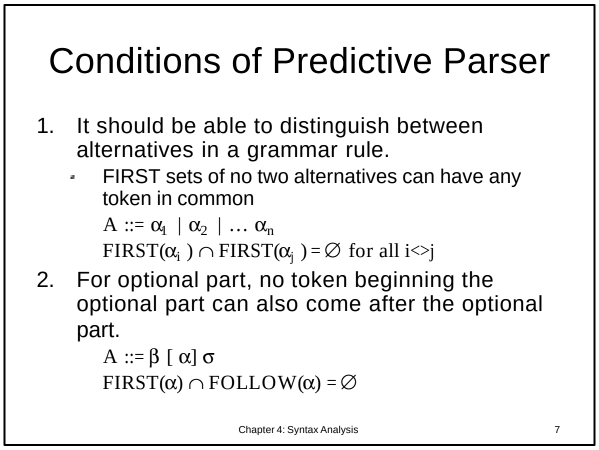
FOLLOW Set

FOLLOW(A), all nonterminals A, apply 1-3 until nothing can be added to any FOLLOW set.

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

Constructing a Predictive Parsing Table

- For each production $A \rightarrow \alpha$
  1. For each terminal a in FIRST(α), add $A \rightarrow \alpha$ to M[A,a].
  2. If $\in$ is in FIRST(α), add $A \rightarrow \alpha$ to M[A,b] for each terminal b in FOLLOW(A).
  3. If $\in$ is in FIRST(α) and \$ in FOLLOW(A), add $A \rightarrow \alpha$ to M[A,\$].
  4. Make each undefined entry of M error



Conditions of Predictive Parser

1. It should be able to distinguish between alternatives in a grammar rule.
   - FIRST sets of no two alternatives can have any token in common

     $A ::= \alpha_1 \mid \alpha_2 \mid \ldots \alpha_n$

     $\mathrm{FIRST}(\alpha_i) \cap \mathrm{FIRST}(\alpha_j) = \emptyset$ for all i<>j
2. For optional part, no token beginning the optional part can also come after the optional part.

   $A ::= \beta\ [\ \alpha]\ \sigma$

   $\mathrm{FIRST}(\alpha) \cap \mathrm{FOLLOW}(\alpha) = \emptyset$