

First Sets

Dr. Mattox Beckman

Illinois Institute of Technology
Department of Computer Science

Objectives

- Be able to explain the purpose of a first set.
- Be able to compute the first set.

The Problem

- Given a grammar for a language L , how can we recognize a sentence in L ?
- Solution: Divide and Conquer: Given a symbol E ...
 - What symbols indicate that the symbol E is just starting? (First Set)
 - What symbols should we expect to see after we have finished parsing an E ?

Misleadingly simple example: $S \rightarrow xEy$ $\text{First}(E) = \{z, q\}$
 $E \rightarrow zE$ $\text{Follow}(E) = \{y\}$
 $E \rightarrow q$

- Important because a parser can see only a few tokens at once.

Algorithm

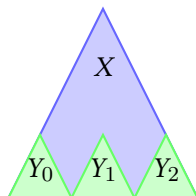
We can compute the FIRST set by a simple iterative algorithm.

For each symbol X .

- ① if X is a terminal, then $First(X) = \{X\}$
- ② if there is a production $X \rightarrow \epsilon$, then add ϵ to $First(X)$.
- ③ if there is a production $X \rightarrow Y_1 Y_2 \cdots Y_n$, then add $First(Y_1 Y_2 \cdots Y_n)$ to $First(X)$:
 - If $First(Y_1)$ does not contain ϵ , then $First(Y_1 Y_2 \cdots Y_n) = First(Y_1)$.
 - Otherwise, $First(Y_1 Y_2 \cdots Y_n) = First(Y_1) / \epsilon \cup First(Y_2 \cdots Y_n)$
 - If all of Y_1, Y_2, \dots, Y_n have ϵ then add ϵ to $First(X)$.

Diagram

$$X \rightarrow Y_0 Y_1 Y_2$$



- if there is a production $X \rightarrow Y_1 Y_2 \cdots Y_n$, then add $First(Y_1 Y_2 \cdots Y_n)$ to $First(X)$:
 - If $First(Y_1)$ does not contain ϵ , then $First(Y_1 Y_2 \cdots Y_n) = First(Y_1)$.
 - Otherwise, $First(Y_1 Y_2 \cdots Y_n) = First(Y_1)/\epsilon \cup First(Y_2 \cdots Y_n)$
 - If all of Y_1, Y_2, \dots, Y_n have ϵ then add ϵ to $First(X)$.

Small Examples

Example 1

$S \rightarrow x A B$

First set of S is $\{x\}$

Example 2

$A \rightarrow \epsilon$

$A \rightarrow y$

$A \rightarrow z q$

First set of A is $\{y, z, \epsilon\}$

Example 3

$B \rightarrow A q$

$B \rightarrow r$

First set of B is $\{y, z, q, r\}$

Example 4

$C \rightarrow A A$

$C \rightarrow B$

First set of C is $\{y, z, q, r, \epsilon\}$

First Set Example

Grammar

$$S \rightarrow \text{if } E \text{ then } S ;$$
$$S \rightarrow \text{print } E ;$$
$$E \rightarrow E + E$$
$$E \rightarrow P \text{ id}$$
$$P \rightarrow * P$$
$$P \rightarrow \epsilon$$

Result

$$S = \{ \}$$
$$E = \{ \}$$
$$P = \{ \}$$

Action

Step 1: Create a list of symbols.

First Set Example

Grammar

$$S \rightarrow \text{if } E \text{ then } S; \Leftarrow$$
$$S \rightarrow \text{print } E; \Leftarrow$$
$$E \rightarrow E + E$$
$$E \rightarrow P \text{ id}$$
$$P \rightarrow * P \Leftarrow$$
$$P \rightarrow \epsilon \Leftarrow$$

Result

$$S = \{\text{if, print}\}$$
$$E = \{\}$$
$$P = \{\epsilon, *\}$$

Action

Step 2: Add terminals starting productions, and all ϵ .

First Set Example

Grammar

$$S \rightarrow \text{if } E \text{ then } S ;$$
$$S \rightarrow \text{print } E ;$$
$$E \rightarrow E + E$$
$$E \rightarrow P \text{ id} \Leftarrow$$
$$P \rightarrow * P$$
$$P \rightarrow \epsilon$$

Result

$$S = \{\text{if, print}\}$$
$$E = \{*, \text{id}\}$$
$$P = \{\epsilon, *\}$$

Action

Step 3: Check productions. Add $\text{First}(Pid)$ to $\text{First}(E)$.

First Set Example

Grammar

$$S \rightarrow \text{if } E \text{ then } S ;$$
$$S \rightarrow \text{print } E ;$$
$$E \rightarrow E + E \leftarrow$$
$$E \rightarrow P \text{ id}$$
$$P \rightarrow * P$$
$$P \rightarrow \epsilon$$

Result

$$S = \{\text{if, print}\}$$
$$E = \{*, \text{id}\}$$
$$P = \{\epsilon, *\}$$

Action

Step 4: Check productions: $E \rightarrow E + E$ adds nothing. We're done.

Another First Set Example

Grammar

$$S \rightarrow Ax$$

$$S \rightarrow By$$

$$S \rightarrow z$$

$$A \rightarrow 1CB$$

$$A \rightarrow 2B$$

$$B \rightarrow 3B$$

$$B \rightarrow C$$

$$C \rightarrow 4$$

$$C \rightarrow \epsilon$$

Result

$$S = \{ \}$$

$$A = \{ \}$$

$$B = \{ \}$$

$$C = \{ \}$$

Action

Create a chart.

Another First Set Example

Grammar

$$S \rightarrow Ax$$

$$S \rightarrow By$$

$$S \rightarrow z \leftarrow$$

$$A \rightarrow 1CB \leftarrow$$

$$A \rightarrow 2B \leftarrow$$

$$B \rightarrow 3B \leftarrow$$

$$B \rightarrow C$$

$$C \rightarrow 4 \leftarrow$$

$$C \rightarrow \epsilon \leftarrow$$

Result

$$S = \{ z \}$$

$$A = \{ 1, 2 \}$$

$$B = \{ 3 \}$$

$$C = \{ \epsilon, 4 \}$$

Action

Add initial terminals and ϵ s.

Another First Set Example

Grammar

$$S \rightarrow Ax \Leftarrow$$

$$S \rightarrow By$$

$$S \rightarrow z$$

$$A \rightarrow 1CB$$

$$A \rightarrow 2B$$

$$B \rightarrow 3B$$

$$B \rightarrow C$$

$$C \rightarrow 4$$

$$C \rightarrow \epsilon$$

Result

$$S = \{z, 1, 2\}$$

$$A = \{1, 2\}$$

$$B = \{3\}$$

$$C = \{\epsilon, 4\}$$

Action

Add $First(Ax)$ to $First(S)$.

Another First Set Example

Grammar

$$S \rightarrow Ax$$
$$S \rightarrow By \leftarrow$$
$$S \rightarrow z$$
$$A \rightarrow 1CB$$
$$A \rightarrow 2B$$
$$B \rightarrow 3B$$
$$B \rightarrow C$$
$$C \rightarrow 4$$
$$C \rightarrow \epsilon$$

Result

$$S = \{z, 1, 2, 3\}$$
$$A = \{1, 2\}$$
$$B = \{3\}$$
$$C = \{\epsilon, 4\}$$

Action

Add $First(By)$ to $First(S)$. Note that there is still more to be added to $First(B)$! We will have to revisit this step later.

Another First Set Example

Grammar

$$S \rightarrow Ax$$
$$S \rightarrow By$$
$$S \rightarrow z$$
$$A \rightarrow 1CB$$
$$A \rightarrow 2B$$
$$B \rightarrow 3B$$
$$B \rightarrow C \leftarrow$$
$$C \rightarrow 4$$
$$C \rightarrow \epsilon$$

Result

$$S = \{z, 1, 2, 3\}$$
$$A = \{1, 2\}$$
$$B = \{3, 4, \epsilon\}$$
$$C = \{\epsilon, 4\}$$

Action

Add $First(C)$ to $First(B)$. At this point we should iterate again to see if anything changes.

Another First Set Example

Grammar

$$S \rightarrow Ax \Leftarrow$$

$$S \rightarrow By$$

$$S \rightarrow z$$

$$A \rightarrow 1CB$$

$$A \rightarrow 2B$$

$$B \rightarrow 3B$$

$$B \rightarrow C$$

$$C \rightarrow 4$$

$$C \rightarrow \epsilon$$

Result

$$S = \{z, 1, 2, 3\}$$

$$A = \{1, 2\}$$

$$B = \{3, 4, \epsilon\}$$

$$C = \{\epsilon, 4\}$$

Action

Add $First(Ax)$ to $First(S)$ again. Nothing happens...

Another First Set Example

Grammar

$$S \rightarrow Ax$$
$$S \rightarrow By \leftarrow$$
$$S \rightarrow z$$
$$A \rightarrow 1CB$$
$$A \rightarrow 2B$$
$$B \rightarrow 3B$$
$$B \rightarrow C$$
$$C \rightarrow 4$$
$$C \rightarrow \epsilon$$

Result

$$S = \{z, 1, 2, 3, 4, y\}$$
$$A = \{1, 2\}$$
$$B = \{3, 4, \epsilon\}$$
$$C = \{\epsilon, 4\}$$

Action

Add $First(By)$ to $First(S)$ again. The 4 gets propagated. Since B could be ϵ we need to add y .

Another First Set Example

Grammar

$$S \rightarrow Ax$$

$$S \rightarrow By$$

$$S \rightarrow z$$

$$A \rightarrow 1CB$$

$$A \rightarrow 2B$$

$$B \rightarrow 3B$$

$$B \rightarrow C \leftarrow$$

$$C \rightarrow 4$$

$$C \rightarrow \epsilon$$

Result

$$S = \{z, 1, 2, 3, 4, y\}$$

$$A = \{1, 2\}$$

$$B = \{3, 4, \epsilon\}$$

$$C = \{\epsilon, 4\}$$

Action

Add $First(C)$ to $First(B)$ again. We are done.