Chapter 4
Syntax Analysis

Linear Grammar

A grammar is linear if it is context-free and all of its productions' right hand sides have at most one nonterminal.

A linear language is a language generated by some linear grammar.

Example

A simple linear grammar is G with N = {S}, Σ = {a, b}, P with start symbol S and rules

$S \rightarrow aSb \mid \varepsilon$

It generates the language $\{a^i b^i \mid i \geq 0\}$

Resolving Difficulties : Non-Reachable Variables

S -> abS | abA | abB
A -> cd
B -> aB
C -> dc

$V_{old} \leftarrow \emptyset$
$V_{new} \leftarrow \{S\}$
while $V_{old} \neq V_{new}$ **do**
  $V_{old} \leftarrow V_{new}$.
  for $X \in V_{old}$ **do**
    for $(X \rightarrow w) \in P$ **do**
      Add all variables appearing in w to V_{new}.
return V_{new}.

The C is a ***useless*** (Non-Reachable)variable

Resolving Difficulties : Non-Reachable Variables

$S_0 \rightarrow S \mid X \mid Z$
$S \rightarrow A$
$A \rightarrow B$
$B \rightarrow C$
$C \rightarrow Aa$
$X \rightarrow C$
$Y \rightarrow aY \mid a$
$Z \rightarrow \epsilon$.

$V_{old} \leftarrow \emptyset$
$V_{new} \leftarrow \{S_o\}$
while $V_{old} \neq V_{new}$ **do**
  $V_{old} \leftarrow V_{new}$.
  for $X \in V_{old}$ **do**
    for $(X \rightarrow w) \in P$ **do**
      Add all variables appearing in w to V_{new}.
return V_{new}.

The Y is a ***useless*** (Non-Reachable)variable

Resolving Difficulties

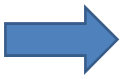
$S_0 \rightarrow S \mid X \mid \varepsilon$
$S \rightarrow A$
$A \rightarrow B$
$B \rightarrow C$
$C \rightarrow Aa$
$X \rightarrow C$

The rule $S \rightarrow A$ is redundant. We can replace any appearance of S by A, and reducing the number of variables by one. Rule of the form $S \rightarrow A$ is called a ***unit production***

The rule X→ C is also redundant. We can replace any appearance of X by C, and reducing the number of variables by one. Rule of the form X→ C is called ***a unit production***

Resolving Difficulties

$S_0 \rightarrow A \mid C \mid \varepsilon$
$A \rightarrow B$
$B \rightarrow C$
$C \rightarrow Aa$



$S_0 \rightarrow A \mid Aa \mid \varepsilon$
$A \rightarrow B$
$B \rightarrow Aa$

$S_0 \rightarrow A \mid Aa \mid \varepsilon$
$A \rightarrow Aa$

Resolving Difficulties: Non-Generating Variables

$S_0 \rightarrow A \mid Aa \mid \varepsilon$

$A \rightarrow Aa$

$V_{old} \leftarrow \emptyset$
$V_{new} \leftarrow V_{old}$
do
  $V_{old} \leftarrow V_{new}$
  for $X \in \mathcal{V}$ **do**
    for $(X \rightarrow w) \in P$ **do**
      if $w \in (\Sigma \cup V_{old})^*$ **then**
        $V_{new} \leftarrow V_{new} \cup \{X\}$
while $(V_{old} \neq V_{new})$
return V_{new}.

$V_{new} = \{S_0\}$ because $S_0 \rightarrow \varepsilon$

The variable A is also ***useless*** since we can note derive any word in Σ^* from A (because once we starting deriving from A we get into an infinite loop).

step-I ↓ step-II ↓

Remove Non-Generating and Non-Reachable Variables

$S \rightarrow AB \mid a$

$A \rightarrow b$

Non-generating {B}

Non-reachable {A}

$\Rightarrow$

$S \rightarrow a$

$S \rightarrow a$

$A \rightarrow b$

Remove Non-Generating and Non-Reachable Variables

S→AB|CA

B→ BC | AB

A→a

C→AB|b

Non-generating

{B}

S → CA

A → a

C → b

⟹

All are reachable

Variables

Remove Non-generating and Non-Reachable Variables

S→ aS | A | C

A→ a

B→aa

C→aCb

Non-generating

{C}

S → aS | A

A → a

B → aa

⟹

Non-reachable

{B}

S → aS | A

A → a

Try Yourself

Remove ***useless*** *variables*(Non-Generating and Non-Reachable Variables)

1. $P = \{S \to aAa, A \to Sb \mid bCC, C \to abb, E \to aC\}$
2. $P = \{S \to aBa \mid BC, A \to aC \mid BCC, C \to a, B \to bcc, D \to E, E \to d\}$
3. $P = \{S \to aAa, A \to bBB, B \to ab, C \to aB\}$
4. $P = \{S \to aS \mid AB, A \to bA, B \to AA\}$

Resolving Difficulties: Elimination of ε-production

S→ ABA
A→ aA| ε
B→ bB| ε

Here, A and B are directly nullable Variables.
A→ aA|a
B→ bB|b

Resolving Difficulties: Elimination of ε-production

S→ ABA
A→ aA| ε
B→ bB| ε

A→ aA | a
B→ bB | b

Here, A and B are **directly Nullable** Variables. S is **indirectly Nullable** variable

$S \rightarrow A_1BA_2 \mid A_1B \mid BA_2 \mid A_1A_2 \mid A_1 \mid A_2 \mid B$

S→ABA| AB | BA | AA |A | B

Resolving Difficulties: Elimination of ε-production

S→ aS |AB|a
A→ ε
B→ ε
D→ b

S→ aS|AB|A|B |a
D→ b

Here, A and B are **directly Nullable** Variables. S is **indirectly Nullable** variable.

Here, A and B are **Non-generating** variables.

Here, D is **Non-Reachable** variable.

S→ aS | a

Try Yourself

Remove ***Nullable*** *Variables* if any and generate grammar without ε-production

$S \rightarrow [E] \mid E$
$E \rightarrow T \mid E+T \mid E-T$
$T \rightarrow F \mid T*F \mid T/F$
$F \rightarrow a \mid b \mid c \mid \varepsilon$

Try Yourself

Remove ***useless*** *Variables and* ***Nullable*** *Variables*

$S \rightarrow a \mid aA \mid B \mid C$
$A \rightarrow aB \mid \varepsilon$
$B \rightarrow aA$
$D \rightarrow ddd$