GREIBACH NORMAL FORM

Dr. A. Beulah

AP/CSE

LEARNING OBJECTIVE

- To Understand the need of formal languages, and grammars (K3)
  - To Understand context free grammars

GREIBACH NORMAL FORM

- A context free grammar G is in GNF if every production is of the form $A \to a\alpha$ where $\alpha \in N^*$ and $a \in T$ (α may be ε) and $S \to \varepsilon$ is in G if $\lambda \in L(G)$, where S does not appear on the RHS of any production.
- Ex:

$S \to aAB \mid \varepsilon$

$A \to bC$

$B \to b$

GNF

$A \to a\alpha$

NT T NT

$\alpha \in N^*$

$S \to \varepsilon$

S ✗ RHS

$\varepsilon \in L(G)$

CNF

$A \to a$

$A \to BC$

$S \to \varepsilon$ S ✗ RHS

$L - \varepsilon$ $\varepsilon \in L(G)$

GREIBACH NORMAL FORM

- Examples: $G = (N, T, P, S)$
  - $G_1 = (\{S, A\}, \{a, b\}, S, \{S \rightarrow aSA \mid a, A \rightarrow aA \mid b\})$ GNF
  - $G_2 = (\{S, A\}, \{a, b\}, S, \{S \rightarrow AS \mid AAS, A \rightarrow SA \mid aa\})$ not GNF
- This grammar $S \rightarrow AB \quad A \rightarrow aA \mid bB \mid b \quad B \rightarrow b$

  is not in GNF
- This grammar $S \rightarrow aAB \mid bBB \mid bB$

  $A \rightarrow aA \mid bB \mid b$

  $B \rightarrow b$

  is in GNF

THEOREM

- Every context free language L can be generated by a context free grammar G in GNF.

CFG → CFL 'L'

GNF → L

- **Step 1 :**

Eliminate null productions and then construct a grammar in G in CNF generating L.

ε-, U, Useless → CNF

A → a

A → BC

Rename the variables as A_1, A_2,...... A_n with $S = A_1$.

Write G as $(\{A_1, A_2, \ldots\ldots A_n\}, T, P, A_1)$.

THEOREM

$A \to \underline{a}\alpha$
$\quad\; t \quad \underline{NT}$

- **Step 2: Derive the productions of the form**

  $$A_i \to a\gamma \text{ or}$$

  $$A_i \to A_j\gamma, \text{ where } j > i.$$

To obtain this convert the A_i productions (i=1,2, ... n-1) to the form $A_i \to A_j\gamma$ such that $j > i$.

THEOREM

- **Step 3: Derive the productions of the form**

$A_i \to a\gamma$ from $A_i \to A_j\gamma$

$A \to \cancel{A}$ $\quad$ $B \to B\,a\,A$

To obtain this eliminate the left recursion

$A \to A\,\alpha_1 \mid \ldots \mid A\,\alpha_m \mid \beta_1 \mid \ldots \mid \beta_n$ where $\beta_1 \ldots \beta_n$ do not start with A

$\Downarrow$ eliminate immediate left recursion

$A \to \beta_1 \mid \ldots \mid \beta_n$

$A \to \beta_1\,B \mid \ldots \mid \beta_n\,B$

$B \to \alpha_1 \mid \ldots \mid \alpha_m$

$B \to \alpha_1\,B \mid \ldots \mid \alpha_m\,B$

an equivalent grammar

THEOREM

- **Step 4: Modify A_i -productions to the form**

  $A_i \rightarrow a\gamma$ for i=1,2, ... n-1

- **Step 5: Modify B_i - productions to the form**

  $B_i \rightarrow a\gamma$

EXAMPLE

First Step

$S \to XA \mid BB$
$B \to b \mid SB$
$X \to b$
$A \to a$

$S = A_1$
$X = A_2$
$A = A_3$
$B = A_4$

$A_1 \to A_2A_3 \mid A_4A_4$
$A_4 \to b \mid A_1A_4$
$A_2 \to b$
$A_3 \to a$

CNF

New Labels

Updated CNF

EXAMPLE

$A_1 \rightarrow A_2A_3 \mid A_4A_4$
$A_4 \rightarrow b \mid A_1A_4$
$A_2 \rightarrow b$
$A_3 \rightarrow a$

Second Step $\boxed{A_i \rightarrow A_j\gamma \quad j > i}$

γ is a string of zero or more NTs

✗ $A_4 \rightarrow A_1A_4$

EXAMPLE

Second Step $\boxed{A_i \rightarrow A_j\gamma \quad j > i}$

$A_4 \longrightarrow A_1A_4$

$A_4 \longrightarrow A_2A_3A_4 \mid A_4A_4A_4 \mid b$

$A_4 \longrightarrow bA_3A_4 \quad \mid A_4A_4A_4 \quad \mid b$

$A_1 \longrightarrow A_2A_3 \mid A_4A_4$

$A_4 \longrightarrow b \mid A_1A_4$

$A_2 \longrightarrow b$

$A_3 \longrightarrow a$

EXAMPLE

$A_1 \rightarrow A_2A_3 \mid A_4A_4$
$A_4 \rightarrow bA_3A_4 \mid A_4A_4A_4 \mid b$
$A_2 \rightarrow b$
$A_3 \rightarrow a$

Third Step

Eliminate Left Recursions

✗ $A_4 \rightarrow A_4A_4A_4$

EXAMPLE

$A \rightarrow A\,\alpha_1 \mid \ldots \mid A\,\alpha_m \mid \beta_1 \mid \ldots \mid \beta_n$ where $\beta_1 \ldots \beta_n$ do not start with A

$\Downarrow$ eliminate immediate left recursion

$A \rightarrow \beta_1 \mid \ldots \mid \beta_n$

$A \rightarrow \beta_1\, B \mid \ldots \mid \beta_n\, B$

$B \rightarrow \alpha_1 \mid \ldots \mid \alpha_m$

$B \rightarrow \alpha_1\, B \mid \ldots \mid \alpha_m\, B$ an equivalent grammar

EXAMPLE

Third Step

Eliminate Left Recursions

$$A_4 \rightarrow bA_3A_4 \mid b \mid bA_3A_4B_4 \mid bB_4$$

$$B_4 \rightarrow A_4A_4 \mid A_4A_4B_4$$

$$A_1 \rightarrow A_2A_3 \mid A_4A_4$$

$$A_4 \rightarrow bA_3A_4 \mid A_4A_4A_4 \mid b$$

$$A_2 \rightarrow b$$

$$A_3 \rightarrow a$$

EXAMPLE

$A_1 \rightarrow A_2A_3 \mid A_4A_4$

$A_2 \rightarrow b$

$A_3 \rightarrow a$

$A_4 \rightarrow bA_3A_4 \mid b \mid bA_3A_4B_4 \mid bB_4$

$B_4 \rightarrow A_4A_4 \mid A_4A_4B_4$

EXAMPLE

Fourth Step

$A_1 \rightarrow A_2A_3 \mid A_4A_4$

$A_2 \rightarrow b$

$A_3 \rightarrow a$

$A_4 \rightarrow bA_3A_4 \mid b \mid bA_3A_4B_4 \mid bB_4$

$B_4 \rightarrow A_4A_4 \mid A_4A_4B_4$

Modify A_i -productions

$A_1 \rightarrow bA_3 \mid bA_3A_4A_4 \mid bA_4 \mid bA_3A_4B_4A_4 \mid bB_4A_4$

EXAMPLE

Fifth Step

Modify B_i -productions

$A_2 \rightarrow b$

$A_3 \rightarrow a$

$A_4 \rightarrow bA_3A_4 \mid b \mid bA_3A_4B_4 \mid bB_4$

$B_4 \rightarrow A_4A_4 \mid A_4A_4B_4$

$A_1 \rightarrow bA_3 \mid bA_3A_4A_4 \mid bA_4 \mid bA_3A_4B_4A_4 \mid bB_4A_4$

$B_4 \rightarrow bA_3A_4A_4 \mid bA_4 \mid bA_3A_4B_4A_4 \mid bB_4A_4 \mid bA_3A_4A_4 \mid bA_4 \mid bA_3A_4B_4A_4 \mid bB_4A_4$

EXAMPLE

$A_1 \rightarrow bA_3 \mid bA_3A_4A_4 \mid bA_4 \mid bA_3A_4B_4A_4 \mid bB_4A_4$

$A_2 \rightarrow b$

$A_3 \rightarrow a$

$A_4 \rightarrow bA_3A_4 \mid b \mid bA_3A_4B_4 \mid bB_4$

$B_4 \rightarrow bA_3A_4A_4 \mid bA_4 \mid bA_3A_4B_4A_4 \mid bB_4A_4 \mid bA_3A_4A_4 \mid bA_4 \mid bA_3A_4B_4A_4 \mid bB_4A_4$

EXAMPLE

- S->AA|a
- A->SS|b

Step1

ε, Unit, Useless ×

CNF ✓

$$\boxed{\begin{array}{l} A \to a \\ A \to BC \end{array}}$$

Rename $S = A1$

$A = A2$

$A_1 \to A_2 A_2 \mid a$

$A_2 \to A_1 A_1 \mid b$

Step2:

$A_i \to a?$

$A_i \to A_j ?$

$j > i$

$A_1 \to A_2 A_2 \mid a$ ✓

$2 > 1$

$A_1 \to a$ ✓

$A_2 \to A_1 A_1 \mid b$

$A_2 \to b$ ✓

$1 \not> 2$

$j \not> i$ $j > i$

$j < i$ ×

$j > i$

$A_2 \to A_2 A_2 A_1 \mid a A_1 \mid b$

$i = i$ LR

Step3

$A_1 \to A_2 A_2 \mid a$ ✓

$A_2 \to A_2 A_2 A_1 \mid a A_1 \mid b$

$A \to A \alpha$

EXAMPLE

Eliminate left Recursion (Immediate)

$A \to A\alpha_1 \mid A\alpha_2 \cdots A\alpha_m \mid \beta_1 \mid \beta_2 \cdots \mid \beta_n$

$A \to \beta_1 \mid \beta_2 \cdots \beta_n$

$A \to \beta_1 B \mid \beta_2 B \cdots \beta_n B$

$B \to \alpha_1 \mid \alpha_2 \cdots \alpha_m$

$B \to \alpha_1 B \mid \alpha_2 B \cdots \alpha_m B$

$\underbrace{A_2}_{A} \to \underbrace{A_2}_{A} \underbrace{A_2 A_1}_{\alpha} \mid \underbrace{aA_1}_{\beta_1} \mid \underbrace{b}_{\beta_2}$

$A_2 \to aA_1 \mid b$

$A_2 \to aA_1B_2 \mid bB_2$

$B_2 \to A_2A_1$

$B_2 \to A_2A_1B_2$

$A_1 \to A_2A_2 \mid a$

$A_2 \to aA_1 \mid b$

$A_2 \to aA_1B_2 \mid bB_2$

$B_2 \to A_2A_1$

$B_2 \to A_2A_1B_2$

Step 4

Modify A_i - productions

$A_1 \to aA_1A_2 \mid bA_2 \mid aA_1B_2A_2 \mid bB_2A_2 \mid a$

$A_2 \to aA_1 \mid b \mid aA_1B_2 \mid bB_2$

$B_2 \cdots\cdots$

Step 5

Modify B_i - productions

EXAMPLE

$B_2 \rightarrow aA_1A_1 \mid bA_1 \mid aA_1B_2A_1 \mid bB_2A_1$

$B_2 \rightarrow aA_1A_1B_2 \mid bA_1B_2 \mid aA_1B_2A_1B_2 \mid bB_2A_1B_2$

Final Grammar

$A_1 \rightarrow \ldots$

$A_2 \rightarrow \ldots$

$B_2 \rightarrow \ldots$

SUMMARY

- Definition of GNF, theorem

TEST YOUR KNOWLEDGE

- The entity which generate Language is termed as:
  a) Automata
  b) Tokens
  c) Grammar
  d) Data
- The minimum number of productions required to produce a language consisting of palindrome strings over $\sum=\{a,b\}$ is
  a) 3
  b) 7 c) 5 d) 6

REFERENCE

- Hopcroft J.E., Motwani R. and Ullman J.D, “Introduction to Automata Theory, Languages and Computations”, Second Edition, Pearson Education, 2008