NORMAL FORMS OF CONTEXT FREE GRAMMAR

Dr. A. Beulah AP/CSE





LEARNING OBJECTIVE

- To Understand the need of formal languages, and grammars (K3)
 - To Understand normal forms of CFG



NORMAL FORMS OF CONTEXT FREE GRAMMAR

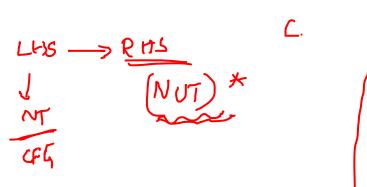
- Chomsky normal form
- Greibach normal form



 The Chomsky normal form places restrictions on the length and the composition of the right-hand side of a rule

• A context free grammar G is in CNF if every production is of the form $A \rightarrow a$ or $A \rightarrow BC$ and $S \rightarrow \epsilon$ is in G if $\epsilon \in L(G)$. When ϵ is in L(G) we assume that S does not appear on the R.H.S. of any other production.

– where B, C ∈ NT





Ex:

$$S \rightarrow AB$$

$$S \rightarrow \epsilon'$$

$$S \times RAIS$$

$$A \rightarrow \underline{a}$$

$$B \rightarrow \underline{b}$$



Algorithm Step 1

- Make sure that the following are satisfied:
 - No ε-productions (other than $S \rightarrow ε$)
 - No unit productions
 - No useless symbols



- Algorithm Step 2
- Eliminate terminals from RHS of productions
 - For each production $A \rightarrow X_1X_2...X_m$
 - where $X_i \in NT \cup T$

- − If m > 1, replace each **terminal** $a \in RHS$ of A
 - Add (if needed) $\underline{C_a \rightarrow a}$ for each $a \in T$, where each C_a is new non-terminal.
 - In production A, replace terminal a with corresponding C_a

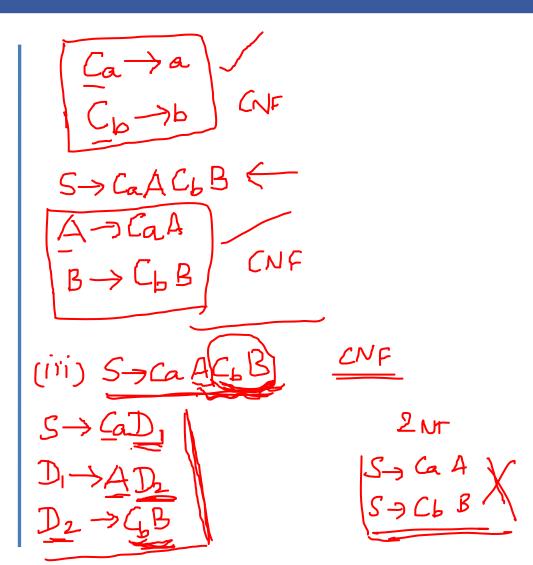


Algorithm Step 3

- Eliminate productions with long RHS:
 - For each production:
 - $A \rightarrow B_1 B_2 ... B_m$, m > 2, where $B_i \in NT$
 - replace with productions
 - $A \rightarrow B_1D_1$
 - $D_1 \rightarrow B_2 D_2$
 - D2 -> B3 D3
 - $D_{m-2} \rightarrow B_{m-1}B_m$
 - where $D_1...D_{m-2}$ are new non-terminals.



- S→aAbB
- A→aA | a
- B→bB | b





$$5 \rightarrow GaD_1$$
 $D_1 \rightarrow AD_2$
 $D_2 \rightarrow C_b R$
 $A \rightarrow GA / a$
 $B \rightarrow Cb B / b$
 $Ca \rightarrow a$
 $Cb \rightarrow b$
 $A \rightarrow a$
 $B \rightarrow a$

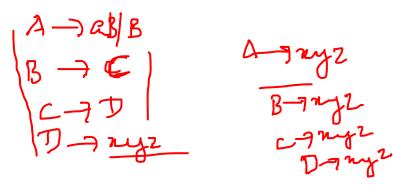


1. Original grammar (no chain rules, useless symbols, or ϵ -productions):

$$S \rightarrow X a Y \mid Y b$$

$$X \rightarrow YX a Y \mid a$$

$$Y \rightarrow SS \mid aX \mid b$$





1. Original grammar (no chain rules, useless symbols, or ε -productions):

$$S \rightarrow X a Y \mid Y b$$

 $X \rightarrow Y X a Y \mid a$
 $Y \rightarrow S S \mid a X \mid b$

2. Grammar after eliminating terminals from RHSs:

$$S \to X C_a Y \mid Y C_b$$

$$X \to Y X C_a Y \mid a$$

$$Y \to S S \mid C_a X \mid b$$

$$C_a \rightarrow a$$

$$C_b \rightarrow b$$

1. Original grammar (no chain rules, useless symbols, or ε -productions):

$$S \rightarrow X a Y \mid Y b$$

 $X \rightarrow Y X a Y \mid a$
 $Y \rightarrow S S \mid a X \mid b$

2. Grammar after eliminating terminals from RHSs:

$$S \rightarrow X C_a Y \mid Y C_b$$
 $C_a \rightarrow a$
 $X \rightarrow Y X C_a Y \mid a$ $C_b \rightarrow b$
 $Y \rightarrow S S \mid C_a X \mid b$

3. Grammar after eliminating long RHSs:

$$S \rightarrow X D_1 \mid Y C_b$$
 $D_1 \rightarrow C_a Y$ $C_a \rightarrow a$
 $X \rightarrow Y D_2 \mid a$ $D_2 \rightarrow X D_3$ $C_b \rightarrow b$
 $Y \rightarrow S S \mid C_a X \mid b$ $D_3 \rightarrow C_a Y$

- S→abAB
- A→bAa | ε
- B \rightarrow BAa | A | ϵ



• S→aSa | bSb | c



Original grammar (no chain rules, useless symbols, or ε -productions):

$$S \rightarrow aXYZ \mid a$$
 $X \rightarrow aX \mid a$
 $Y \rightarrow bcY \mid bc$ $Z \rightarrow cZ \mid c$

1. Grammar after eliminating terminals from RHSs:

$$S \rightarrow AXYZ \mid a$$
 $A \rightarrow a$
 $X \rightarrow AX \mid a$ $B \rightarrow b$
 $Y \rightarrow BCY \mid BC$ $C \rightarrow c$
 $Z \rightarrow CZ \mid c$

2. Grammar after eliminating long RHSs:

$$S \rightarrow AF \mid a$$
 $A \rightarrow a$ $F \rightarrow XG$
 $X \rightarrow AX \mid a$ $B \rightarrow b$ $G \rightarrow YZ$
 $Y \rightarrow BH \mid BC$ $C \rightarrow c$ $H \rightarrow CY$
 $Z \rightarrow CZ \mid c$



GREIBACH NORMAL FORM

- A context-free grammar is in **Greibach Normal Form** if every production is of the form $A \rightarrow aX$
 - where $A \in NT$, $X \in NT^*$, and $a \in \Sigma$
- Examples:
 - $-G_1 = (\{S, A\}, \{a, b\}, S, \{S \rightarrow aSA \mid a, A \rightarrow aA \mid b\})$
 - GNF
 - $-G_2 = (\{S, A\}, \{a, b\}, S, \{S \rightarrow AS \mid AAS, A \rightarrow SA \mid aa\})$
 - not GNF
- This grammar $S \rightarrow ABA \rightarrow aA \mid bB \mid b \quad B \rightarrow b$ is not in GNF
- This grammar $S \rightarrow aAB \mid bBB \mid bB$ $A \rightarrow aA \mid bB \mid bB \rightarrow b$

is in GNF



SUMMARY

- Different normal forms of Context free grammar
 - Chomsky normal form
 - Greibach normal form



TEST YOUR KNOWLEDGE

- The Grammar can be defined as: G=(V, ∑, p, S)
 In the given definition, what does S represents?
 - a) Accepting State
 - b) Starting Variable
 - c) Sensitive Grammar
 - d) None of these



LEARNING OUTCOME

On successful completion of this topic, the student will be able to:

Understand simplification of CFG (K3)



REFERENCE

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

