

CHOMSKY NORMAL FORM (CNF)

- Defn:

spi form of context free grammar

3 rules:

$$A \rightarrow BC$$
$$A \rightarrow a$$
$$S \rightarrow E$$

A|B|C - nontermining

$a \Rightarrow \text{terminal}$

ε - empty string

$S \rightarrow$ start symbol.

Why use CAFE?

For efficient determining whether a given string can be generated by a given grammar.

Properties of CNF:-

① standardization : $\begin{cases} \rightarrow \text{simplifies parsing process} \\ \rightarrow \text{theoretical analysis of grammars} \end{cases}$

⑧ Binary structure : $A \rightarrow Bc$

③ No mixed terminals : 42 non terminal

y terminal, (on

g) 1 non terminal

Q: Consider the following context free grammar
into CNF:

$$S \rightarrow AB$$

$$A \rightarrow aA|a$$

$$B \rightarrow bB|b$$

Ans:-

Step 1: Introduce new nonterminal

$$\text{Let } x \rightarrow a$$

$$y \rightarrow b$$

Step 2: Rewrite rules using these new non-terminals

$$S \rightarrow AB$$

$$A \rightarrow xA|x$$

$$B \rightarrow yB|y$$

Now the grammar is in CNF

$$S \rightarrow AB \quad \checkmark$$

$$A \rightarrow xA|x \quad \checkmark$$

$$B \rightarrow yB|y \quad \checkmark$$

$$x \rightarrow a$$

$$y \rightarrow b$$

Any context free language is generated by a context-free grammar in CNF.

Proof:

Let $G = (N, T, R, S)$ → original CFG.

G_1 (Grammar)

$S \rightarrow ASA | aB$

$A \rightarrow B | S$

$A \rightarrow B$

$B \rightarrow b | \epsilon$

Conditions

Step 1: Introduce new start symbol
(not in right side)

$S_0 \rightarrow S$

Step 2: Remove ϵ production \rightarrow as ϵ is
the only production for start symbol

Step 3:- Remove unit production.

$A \rightarrow B$ where both are non
terminals

Step 4: Ensure binary production

at most 2 non terminals / terminals

Step 1: Start symbol

$$S_0 \rightarrow S$$

$$S \rightarrow ASA | aB$$

$$A \rightarrow B | \epsilon$$

$$B \rightarrow b | \epsilon$$

Step 2: Remove ϵ

$$B \rightarrow \epsilon \quad \downarrow \text{to remove this.}$$

1. For $S \rightarrow ASA$

$$A \rightarrow B \quad B \rightarrow \epsilon \quad \rightarrow \quad A \rightarrow \epsilon$$

Left side

Right side

Both

$$S \rightarrow SA$$

$$S \rightarrow AS$$

$$S \rightarrow S$$

2. For $S \rightarrow aB$ $B \rightarrow \epsilon \Rightarrow S \rightarrow a$.

Grammar.

$$S_0 \rightarrow S$$

$$S \rightarrow ASA | AA | aB | a$$

$$A \rightarrow B | S$$

$$B \rightarrow b$$

Step 3:

3. Remove unit production, ($A \rightarrow B$)

(2)

$$A \rightarrow B$$

$$A \rightarrow S$$

Replace $A \rightarrow B$ by productions of B
 $A \rightarrow S$ by productions of S

① $A \rightarrow B$ gives $A \rightarrow b$

② $A \rightarrow S$ gives $A \rightarrow ASA | AA | AB | a$

Grammar:

$S \rightarrow S$

$S \rightarrow ASA | AA | AB | a$

$A \rightarrow ASA | AA | AB | a | b$

$B \rightarrow b$

Step 1: Binary & terminal production.
(Grammar CNF)

1. $S \rightarrow ASA$ - 3 non terminals
new non terminal

$X \rightarrow SA$

Replace $S \rightarrow ASA$ with $S \rightarrow AX$.

2. $S \rightarrow aB$ - a terminal + non terminal
new non terminal

$Y \rightarrow a$

$S \rightarrow aB$ with $Y \rightarrow YB$

3. $111'8 \ A \rightarrow AB$
Replace with $A \rightarrow YB$ where $Y \rightarrow a$

After these adjustments

Grammar.

1. $S_0 \rightarrow S$

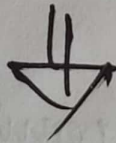
2. $S \rightarrow AX | AYB | a$

3. $A \rightarrow AX | AYB | a | b$

4. $B \rightarrow b$

5. $X \rightarrow SA$

6. $Y \rightarrow a$



Final Grammar.