CHOMSKY NORMAL FORM (CNF)

- Defn:

spl form of context free grammar

3 rules:

$A \to BC$

$A \to a$

$S \to \epsilon$

A, B, C - nonterminal

a → terminal

ε - empty string

S → start symbol.

Why use CNF?

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

Properties of CNF:-

① Standardization : → simplifies parsing process
→ theoretical analysis of grammars

② Binary structure : $A \to BC$

③ No mixed terminals : → 2 non terminals (or)
→ 1 terminal.
→ 1 non terminals

Q: consider the following context free gra[mmar]
into CNF :

$S \rightarrow AB$
$A \rightarrow aA | a$
$B \rightarrow bB | b$

Ans :-
Step 1: introduce new nonterminals

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

$S \rightarrow AB$ ✓
$A \rightarrow XA | X$ ✓
$B \rightarrow YB | Y$ ✓
$X \rightarrow a$
$Y \rightarrow b$.

...eorem 2.9.

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

proof:

Let $G = (N, T, \vec{R}, S)$ → original CFG.

G_{16} (Grammar)

$S \to ASA \mid aB$

$A \to B \mid S$

$B \to b \mid \epsilon$

$A \to B$

Conditions

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

$S_0 \to S$

Step 2: Remoe ϵ production → as ϵ is the only production for start symbol

Step 3:- Remove unit production.

$A \to B$ where both are non terminal's

Step 4: Ensure Binary Production

more than 2 non terminals / terminals

Step1: Start symbol

$S_0 \to S$

$S \to ASA | aB$

$A \to B | S$

$B \to b | \epsilon$

step 2: Remove ϵ

$B \to \epsilon$ To remove this.

1. For $S \to ASA$

$A \to B$ $B \to \epsilon$ $\to$ $A \to \epsilon$

Left side	Right side	Both
$S \to SA$	$S \to AS$	$S \to S$

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

Grammar.

$S_0 \to S$

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

$A \to B | S$

$B \to b$

Step3:

3. Remove unit production. $(A \to B)$

(2) $A \to B$

$A \to S$

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

① $A \to B$ gives $A \to b$
② $A \to S$ gives $A \to ASA | AA | aB | a$

Grammar:

$S_0 \to S$
$S \to ASA | AA | aB | a$
$A \to ASA | AA | aB | a | b$
$B \to b$.

Step 4:- Binary & terminal Production.
(Grammar CNF)

1. $S \to ASA$ - 3 non terminals

new non terminal
$X \to SA$

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

2. $S \to aB$ → a terminal + non terminal

new non terminal
$Y \to a$

$S \to aB$ with $S \to YB$

3. III'ly A → aB

Replace with A → YB where Y → a

After these adjustments

Grammar.

1. S0 → S
2. S → AX | AA | YB | a
3. A → AX | AA | YB | a | b
4. B → b
5. X → SA
6. Y → a

⇓

Final Grammar.