



CSC-257

Theory Of Computation

(BSc CSIT, TU)

Ganesh Khatri
kh6ganesh@gmail.com

Chomsky Normal Form (CNF)

- A context free grammar $G = (V, T, P, S)$ is said to be in Chomsky Normal Form (CNF) if every production in G are in one of the two forms :

$A \rightarrow BC$ and

$A \rightarrow a$

Where $A, B, C \in V$ and $a \in T$

- Thus, a grammar in CNF is one which should not have :
 - ϵ -production
 - Unit productions
 - Useless symbols

Chomsky Normal Form

- **Theorem :** Every context free language (CFL) without ϵ -production can be generated by grammar in CNF.
- **Proof :**
 1. **Simplifications :** Eliminate ϵ -productions if any, Eliminate unit productions if any, and Eliminate useless symbols if any
 2. If all the productions are of the form $A \rightarrow a$ and $A \rightarrow BC$ with $A, B, C \in V$ and $a \in T$, we have done. Otherwise,
 3. If the right side is in the form $A \rightarrow aB$ or $A \rightarrow Ba$ where $a \in T$ and $A, B \in V$, then replace the production by $A \rightarrow XB$ and $X \rightarrow a$. Repeat this process for every production of the form $A \rightarrow aB$ or $A \rightarrow Ba$
 4. Replace each production of the form $A \rightarrow B_1B_2 \dots B_n$ where $n > 2$ with $A \rightarrow B_1C$ where $C \rightarrow B_2 \dots B_n$. And repeat this step for all the productions having two or more symbols on the right side. [$A, B_1, B_2, \dots B_n \in V$]
 5. Finally, all of the productions will be in the form as : $A \rightarrow BC$ or $A \rightarrow a$. This is certainly a grammar in CNF.

Chomsky Normal Form

- **Example :** Convert the following grammar into CNF.

$S \rightarrow AAC$

$A \rightarrow aAb \mid \epsilon$

$C \rightarrow aC \mid a$

- **Solution :**

- **1. Simplification :** grammar after removal of ϵ -productions is :

$S \rightarrow AAC \mid AC \mid C$

$A \rightarrow aAb \mid ab$

$C \rightarrow aC \mid a$

Chomsky Normal Form

- grammar after removal of unit productions is :
 $S \rightarrow AAC \mid AC \mid aC \mid a$
 $A \rightarrow aAb \mid ab$
 $C \rightarrow aC \mid a$
- There are no any useless symbols
- **2.** Resulting grammar is not in CNF
- **3.** Here, suppose $A' \rightarrow a$ and $B' \rightarrow b$, then our grammar will look like
 $S \rightarrow AAC \mid AC \mid A'C \mid a$
 $A \rightarrow A'AB' \mid A'B'$
 $C \rightarrow A'C \mid a$
 $A' \rightarrow a$
 $B' \rightarrow b$

Chomsky Normal Form

- **4.** Now suppose, $X \rightarrow AC$ and $Y \rightarrow AB'$, then grammar will be :

$S \rightarrow AX \mid AC \mid A'C \mid a$

$A \rightarrow A'Y \mid A'B'$

$C \rightarrow A'C \mid a$

$X \rightarrow AC$

$Y \rightarrow AB'$

$A' \rightarrow a$

$B' \rightarrow b$

- Now, above grammar is in CNF.

Chomsky Normal Form

- **Exercises** : Convert the following grammars in CNF.

1.

$$S \rightarrow ASA \mid aB$$

$$A \rightarrow B \mid S$$

$$B \rightarrow b \mid \epsilon$$

[**Solution** : <https://www.youtube.com/watch?v=FNPSInj3Vt0>]

2.

$$S \rightarrow a \mid aA \mid B$$

$$A \rightarrow aBB \mid \epsilon$$

$$B \rightarrow Aa \mid b$$

[**Solution** : <https://www.youtube.com/watch?v=7G0PwGrdIH8>]