

CSE-233 : Section A  
Summer 2020

# Simplification of CFG

Reference:

- i) Book2 Theorem 2.9
- ii) [Neso Academy](#)

Md. Saidul Hoque Anik  
anik@cse.uiu.ac.bd

# Common Steps for Simplification

- ~~1. Reduction of CFG~~
2. Removal of Unit Production Rules ( $A \rightarrow B$ )
3. Removal of Null Production Rules ( $A \rightarrow \epsilon$ )

# Removal of Unit Production

For the following rules:

$$A \rightarrow B$$

$$B \rightarrow m \quad [m \text{ is a string containing both terminal and non-terminal. It's not a single terminal}]$$

We replace all occurrence of B in unit rules with the value of B (don't remove the rule  $B \rightarrow m$ )

After removal, the given rules will become:  $A \rightarrow m, B \rightarrow m$

# Practice

Identify and remove unit productions from the following grammar.

$$S \rightarrow XY, Y \rightarrow a, Y \rightarrow Z \mid b, Z \rightarrow M, M \rightarrow N, N \rightarrow a$$

# Removal of Null Productions

For the following rules:

$$A \rightarrow \varepsilon$$

For each null production, we create new rules replacing A with empty string in all possible combinations in a rule. For example,

$$S \rightarrow ABA$$

will becomes,  $S \rightarrow ABA \mid BA \mid AB \mid B$



First A is removed

# Removal of Null Productions

For the following rules:

$$A \rightarrow \varepsilon$$

For each null production, we create new rules replacing A with empty string in all possible combinations in a rule. For example,

$$S \rightarrow ABA$$

will becomes,  $S \rightarrow ABA \mid BA \mid AB \mid B$



Last A is removed

# Removal of Null Productions

For the following rules:

$$A \rightarrow \varepsilon$$

For each null production, we create new rules replacing A with empty string in all possible combinations in a rule. For example,

$$S \rightarrow ABA$$

will becomes,  $S \rightarrow ABA \mid BA \mid AB \mid B$



Both As are removed

# Practice

Identify and remove null productions from the following grammar.

$$S \rightarrow ABAC, A \rightarrow aA \mid \varepsilon, B \rightarrow bB \mid \varepsilon, C \rightarrow c$$



# Chomsky Normal Form

In Chomsky Normal Form (CNF) we have a restriction on the length of RHS in the production rules. Elements in RHS should either be two variables or a terminal.

A CFG is in Chomsky Normal Form if the all the productions are in any of the two following forms:

$$A \rightarrow a$$

$$A \rightarrow BC$$

(or  $S \rightarrow \epsilon$  if the language includes empty strings)

Where A, B and C are non-terminals and a is a terminal.

# Steps of Converting CFG to CNF

1. If  $S$  occurs on right side, create a new start symbol  $S'$  and add this production rule  $S' \rightarrow S$
2. Remove null productions
3. Remove unit productions
4. Replace rules with length  $> 2$  in RHS to  $A \rightarrow BC$  form. For example:  
 $A \rightarrow aBCDE$   
 $\Rightarrow A \rightarrow aX_1, X_1 \rightarrow BCDE$   
 $\Rightarrow A \rightarrow aX_1, X_1 \rightarrow BX_2, X_2 \rightarrow CDE$   
 $\Rightarrow A \rightarrow aX_1, X_1 \rightarrow BX_2, X_2 \rightarrow CX_3, X_3 \rightarrow DE$
5. If the right side of any production is in the form  $A \rightarrow aB$  where  $a$  is a terminal and  $A$  and  $B$  are non-terminals, then the production is replaced by  $A \rightarrow XB$  and  $X \rightarrow a$ . Repeat this step for every production which is of the form  $A \rightarrow aB$

# Example

Let's convert the following grammar in CNF. What will be the first step?

$$S \rightarrow ASA \mid aB$$

$$A \rightarrow B \mid S$$

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

# Example

New Start symbol added. 2<sup>nd</sup> step is to remove all null productions.

$$\begin{aligned} S_0 &\rightarrow S \\ S &\rightarrow ASA \mid aB \\ A &\rightarrow B \mid S \\ B &\rightarrow b \mid \epsilon \end{aligned}$$

# Example

Null productions removed. 3<sup>rd</sup> step is to remove all unit productions.

Remove  $\epsilon$  rules  $B \rightarrow \epsilon$ , shown on the left, and  $A \rightarrow \epsilon$ , shown on the right.

$$S_0 \rightarrow S$$

$$S \rightarrow ASA \mid aB \mid \mathbf{a}$$

$$A \rightarrow B \mid S \mid \epsilon$$

$$B \rightarrow \mathbf{b}$$

$$S_0 \rightarrow S$$

$$S \rightarrow ASA \mid aB \mid \mathbf{a} \mid \mathbf{SA} \mid \mathbf{AS} \mid \mathbf{S}$$

$$A \rightarrow B \mid S$$

$$B \rightarrow \mathbf{b}$$

# Example

Removal of unit productions.

**3a.** Remove unit rules  $S \rightarrow S$ , shown on the left, and  $S_0 \rightarrow S$ , shown on the right.

$$\begin{aligned} S_0 &\rightarrow S \\ S &\rightarrow ASA \mid aB \mid a \mid SA \mid AS \\ A &\rightarrow B \mid S \\ B &\rightarrow b \end{aligned}$$

$$\begin{aligned} S_0 &\rightarrow \mathbf{ASA \mid aB \mid a \mid SA \mid AS} \\ S &\rightarrow ASA \mid aB \mid a \mid SA \mid AS \\ A &\rightarrow B \mid S \\ B &\rightarrow b \end{aligned}$$

# Example

Removal of unit productions.

3b. Remove unit rules  $A \rightarrow B$  and  $A \rightarrow S$ .

$$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

$$S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

$$A \rightarrow S \mid \mathbf{b}$$

$$B \rightarrow \mathbf{b}$$

$$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

$$S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$$

$$A \rightarrow \mathbf{b} \mid \mathbf{ASA} \mid \mathbf{aB} \mid \mathbf{a} \mid \mathbf{SA} \mid \mathbf{AS}$$

$$B \rightarrow \mathbf{b}$$

# Example

Converting to proper form

$$S_0 \rightarrow AA_1 \mid UB \mid \mathbf{a} \mid SA \mid AS$$

$$S \rightarrow AA_1 \mid UB \mid \mathbf{a} \mid SA \mid AS$$

$$A \rightarrow \mathbf{b} \mid AA_1 \mid UB \mid \mathbf{a} \mid SA \mid AS$$

$$A_1 \rightarrow SA$$

$$U \rightarrow \mathbf{a}$$

$$B \rightarrow \mathbf{b}$$



# Practice

## Converting CFG to CNF

1.

$S \rightarrow ASB$   
 $A \rightarrow aAS \mid a \mid \varepsilon$   
 $B \rightarrow SbS \mid A \mid bb$

3.

$S \rightarrow aXbY$   
 $X \rightarrow aX \mid \varepsilon$   
 $Y \rightarrow bY \mid \varepsilon$

5.

$S \rightarrow S_1 \mid S_2$   
 $S_1 \rightarrow S_1b \mid Ab \mid \varepsilon$   
 $A \rightarrow aAb \mid ab$   
 $S_2 \rightarrow S_2a \mid Ba \mid \varepsilon$   
 $B \rightarrow bBa \mid ba$

7.

$S \rightarrow aX \mid bY \mid b \mid ZZc$   
 $X \rightarrow Yaa \mid abZ \mid \varepsilon$   
 $Y \rightarrow bXXb \mid ab \mid cZ$   
 $Z \rightarrow a \mid b \mid XZ \mid \varepsilon$

2.

$S \rightarrow XY$   
 $X \rightarrow abb \mid aXb \mid \varepsilon$   
 $Y \rightarrow c \mid cY$

4.

$S \rightarrow aXbX$   
 $X \rightarrow aY \mid bY \mid \varepsilon$   
 $Y \rightarrow X \mid c$

6.

$S \rightarrow DBC \mid Ba$   
 $B \rightarrow 0B1 \mid 01 \mid \varepsilon$   
 $C \rightarrow aCb \mid aC \mid Bb$   
 $D \rightarrow bD \mid D$

8.

$A \rightarrow BAB \mid B \mid \varepsilon$   
 $B \rightarrow 00 \mid \varepsilon$