

CHOMSKY NORMAL FORM

IT IS A THREE STEP PROCESS

ELIMINATE NULL PRODUCTION

ELIMINATE UNIT PRODUCTION

CONVERT CFG INTO CHOMSKY- NORMAL FORM

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

$$S \rightarrow ABA$$

$$A \rightarrow aA / \Lambda$$

$$B \rightarrow bB / \Lambda$$

NTs 'A' and B are directly nullable
Variables. ($A \rightarrow \Lambda$. and $B \rightarrow \Lambda$)

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

$$A \rightarrow aA \mid \lambda$$

\Downarrow

$$A \rightarrow aA \mid a$$

$$B \rightarrow bB \mid \lambda$$

\Downarrow

$$B \rightarrow bB \mid b$$

Definition 4.26 A Recursive Definition of the Set of Nullable Variables of G

1. Every variable A for which there is a production $A \rightarrow \Lambda$ is nullable.
2. If A_1, A_2, \dots, A_k are nullable variables (not necessarily distinct), and

$$B \rightarrow A_1 A_2 \dots A_k$$

is a production, then B is nullable.

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

'S' is indirectly nullable Variable

$$S \rightarrow ABA$$

$$S \rightarrow BA$$

$$S \rightarrow A$$

$$S \rightarrow \Lambda$$

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

$$S \rightarrow A B A$$

$$S \rightarrow A_1 B A_2$$

$$S \rightarrow A_1 B A_2 \mid B A_2 \mid A_1 A_2 \mid A_1 B$$

(Substituting Null for each nullable
Variable one at a time)

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

$$S \rightarrow A_1 B A_2 \mid B A_2 \mid A_1 B \mid A_1 A_2 \mid A_1 \mid A_2 \mid B$$

Substituting 'n' for 2 nullable variables at once.

Removing suffixes.

$$S \rightarrow ABA \mid BA \mid AA \mid AB \mid B \mid A$$

CHOMKSY NORMAL FORM : ELIMINATE UNIT PRODUCTION

Production of the form $NT \rightarrow NT$ is a Unit Production.

$$S \rightarrow ABA \mid AB \mid BA \mid AA \mid A \mid B$$

$$A \rightarrow aA \mid a$$

$$B \rightarrow bB \mid b$$

Remove $S \rightarrow A$: $S \rightarrow ABA \mid AB \mid BA \mid AA \mid aA \mid a \mid B$

Remove $S \rightarrow B$: $S \rightarrow ABA \mid AB \mid BA \mid AA \mid aA \mid a \mid b \mid bB$

CHOMKSY NORMAL FORM : CONVERSION INTO CNF

$$\begin{aligned} NT &\rightarrow NT \cdot NT \\ NT &\rightarrow T \end{aligned}$$

$$A \rightarrow aA \mid a \Rightarrow$$

$$X_a \rightarrow a \quad \checkmark$$

$$A \rightarrow X_a A \quad \checkmark$$

$$A \rightarrow a \quad \checkmark$$

$$B \rightarrow bB \mid b \Rightarrow$$

$$X_b \rightarrow b \quad \checkmark$$

$$B \rightarrow X_b B \quad \checkmark$$

$$B \rightarrow b \quad \checkmark$$

CHOMKSY NORMAL FORM : CONVERSION INTO CNF

$S \rightarrow ABA / AB / BA / AA / aA / a / bB / b$

$S \rightarrow AX_1 / AB / BA / AA / X_a A / a / X_b B / b$

$X_1 \rightarrow BA$

$A \rightarrow X_a A / a$

$B \rightarrow X_b B / b$

$X_a \rightarrow a$

$X_b \rightarrow b$

Answer

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

$$S \rightarrow TU \mid V$$

$$\checkmark T \rightarrow aTb \mid \Lambda$$

$$\checkmark U \rightarrow cU \mid \Lambda$$

$$V \rightarrow aVc \mid W$$

$$\checkmark W \rightarrow bW \mid \Lambda$$

$$W \rightarrow bW \mid b$$

$$U \rightarrow cU \mid c$$

$$T \rightarrow aTb \mid ab$$

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

$$S \rightarrow TU \mid V$$

$$\checkmark T \rightarrow aTb \mid \Lambda$$

$$\checkmark U \rightarrow cU \mid \Lambda$$

$$V \rightarrow aVc \mid W$$

$$\checkmark W \rightarrow bW \mid \Lambda$$

$$S \rightarrow TU \mid V \mid U \mid T$$

$$W \rightarrow bW \mid b$$

$$U \rightarrow cU \mid c$$

$$T \rightarrow Tb \mid ab$$

$$V \rightarrow aVc \mid ac \mid W$$

CHOMKSY NORMAL FORM : ELIMINATE UNIT PRODUCTION

$$\begin{aligned} S &\rightarrow TU \mid T \mid U \mid V & T &\rightarrow aTb \mid ab & U &\rightarrow cU \mid c \\ V &\rightarrow aVc \mid ac \mid W & W &\rightarrow bW \mid b \end{aligned}$$

$$W \rightarrow bW \mid b$$

$$V \rightarrow aVc \mid ac \mid bW \mid b$$

$$U \rightarrow cU \mid c$$

$$T \rightarrow aTb \mid ab$$

$$S \rightarrow TU \mid aTb \mid ab \mid cU \mid c \mid aVc \mid ac \mid bW \mid b$$

CHOMKSY NORMAL FORM : CONVERSION INTO CNF

$$S \rightarrow TU \mid aTb \mid ab \mid cU \mid c \mid aVc \mid ac \mid bW \mid b$$

$$T \rightarrow aTb \mid ab$$

$$U \rightarrow cU \mid c$$

$$V \rightarrow aVc \mid ac \mid bW \mid b$$

$$W \rightarrow bW \mid b$$



$$X_a \rightarrow a$$

$$X_b \rightarrow b$$

$$X_c \rightarrow c$$

$$S \rightarrow TU \mid X_a T X_b \mid X_a X_b \mid X_c U \mid c \mid X_a V X_c \mid X_a X_c \mid X_b W \mid b$$

$$T \rightarrow X_a T X_b \mid X_a X_b$$

$$U \rightarrow X_c U \mid c$$

$$V \rightarrow X_a V X_c \mid X_a X_c \mid X_b W \mid b$$

$$W \rightarrow X_b W \mid b$$

CHOMKSY NORMAL FORM : CONVERSION INTO CNF

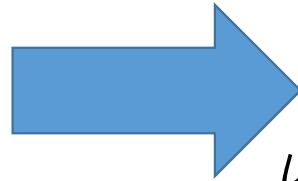
$$S \rightarrow TU \mid X_a T X_b \mid X_a X_b \mid X_c U \mid c \mid X_a V X_c \mid X_a X_c \mid X_b W \mid b$$

$$T \rightarrow X_a T X_b \mid X_a X_b$$

$$U \rightarrow X_c U \mid c$$

$$V \rightarrow X_a V X_c \mid X_a X_c \mid X_b W \mid b$$

$$W \rightarrow X_b W \mid b$$



$$S \rightarrow TU \mid X_a Y_1 \mid X_a X_b \mid X_c U \mid c \mid X_a Y_2 \mid X_a X_c \mid X_b W \mid b$$

$$Y_1 \rightarrow T X_b$$

$$Y_2 \rightarrow V X_c$$

$$T \rightarrow X_a Y_3 \mid X_a X_b$$

$$Y_3 \rightarrow T X_b$$

$$U \rightarrow X_c U \mid c$$

$$V \rightarrow X_a Y_4 \mid X_a X_c \mid X_b W \mid b$$

$$Y_4 \rightarrow V X_c$$

$$W \rightarrow X_b W \mid b$$

CHOMKSY NORMAL FORM : ELIMINATE NULL PRODUCTION

$$S \rightarrow SS \mid (S) \mid \Lambda$$

\Downarrow

$$S \rightarrow SS \mid (S) \mid () \mid S$$

CHOMKSY NORMAL FORM : ELIMINATE UNIT PRODUCTION

$$S \rightarrow SS \mid (s) \mid () \mid S$$

\Downarrow

$$S \rightarrow SS \mid (s) \mid ()$$

CHOMKSY NORMAL FORM : CONVERSION INTO CNF

$$S \rightarrow SS \mid (S) \mid ($$

✓ $X_c \rightarrow ($

✓ $X_c \rightarrow)$

$$S \rightarrow SS \mid X_c S X_c \mid X_c X_c$$

$$X_c \rightarrow ($$

$$X_c \rightarrow)$$

$$S \rightarrow SS \mid X_c P \mid X_c X_c$$

$$P \rightarrow S X_c$$

ADVANTAGES OF CHOMSKY NORMAL FORM

- There is no practical significance of Null when we want to generate/derive any string from grammar. Removing null production gives a language $L' = L - \{\epsilon\}$
- Keeping Unit production is an extra and unnecessary step when we want to generate a string (Parsing Process in compiler)
- Chomsky Normal form converts grammar representation into Binary Tree from M-Ary tree representation. It is possible to count the number of steps required to generate/derive the actual string from the grammar.

TRY YOURSELF

G has productions

$$\begin{aligned} S &\rightarrow AaA \mid CA \mid BaB & A &\rightarrow aaBa \mid CDA \mid aa \mid DC \\ B &\rightarrow bB \mid bAB \mid bb \mid aS & C &\rightarrow Ca \mid bC \mid D & D &\rightarrow bD \mid \Lambda \end{aligned}$$