

Chomsky Normal Form (CNF)

Chomsky Normal Form (CNF)

- A Grammar G is in CNF if all of its productions are of the form:
 - $A \rightarrow BC$ or
 - $A \rightarrow a$ or
 - $S \rightarrow \varepsilon$,where $A, B, C \in V$, $a \in T$ and S is the starting symbol
- It has been proven that all CFG could be convert into CNF

Convert to CNF(1)

- Step 1: get rid of useless symbols and remove all ε -productions and unit productions

– Example:

$S \rightarrow ASB \mid AB$

$A \rightarrow aAS \mid aA \mid a$

$B \rightarrow SbS \mid bS \mid Sb \mid b \mid aAS \mid aA \mid a \mid bb$

Convert to CNF(2)

- Step 2: get rid of production whose bodies are mixes of terminals and variables, or consist of more than one terminal

$S \rightarrow ASB \mid AB$

$A \rightarrow CAS \mid CA \mid a$

$B \rightarrow SDS \mid DS \mid SD \mid b \mid CAS \mid CA \mid a \mid DD$

$C \rightarrow a$

$D \rightarrow b$

$S \rightarrow ASB \mid AB$

$A \rightarrow aAS \mid aA \mid a$

$B \rightarrow SbS \mid bS \mid Sb \mid b \mid aAS \mid aA \mid a \mid bb$

Convert to CNF(3)

- Step 3: break up production bodies longer than 2

$S \rightarrow AE \mid AB$

$A \rightarrow CF \mid CA \mid a$

$B \rightarrow SG \mid DS \mid SD \mid b \mid CF \mid CA \mid a \mid DD$

$C \rightarrow a$

$D \rightarrow b$

$E \rightarrow SB$

$F \rightarrow AS$

$G \rightarrow DS$

$S \rightarrow ASB \mid AB$

$A \rightarrow CAS \mid CA \mid a$

$B \rightarrow SDS \mid DS \mid SD \mid b \mid CAS \mid CA \mid a \mid DD$

$C \rightarrow a$

$D \rightarrow b$

Chomsky Normal Form (CNF)

A CFG is in Chomsky Normal Form if all its productions are of the form:

$$A \rightarrow BC \text{ or}$$

$$A \rightarrow a$$

where $A, B, C \in V$ and $a \in T$. Also, $S \rightarrow \varepsilon$ may be one of the productions.

Examples of CNF

Example 1:

$$S \rightarrow AB$$
$$A \rightarrow BC \mid CC \mid a$$
$$B \rightarrow CB \mid b$$
$$C \rightarrow c$$

Example 2:

$$S \rightarrow AB \mid BC \mid AC \mid \varepsilon$$
$$A \rightarrow BC \mid a$$
$$B \rightarrow AC \mid b$$
$$C \rightarrow AB \mid c$$

CNF

Is that all Context Free Grammars can be expressed in Chomsky Normal Form?

Consider the following simple grammar:

$$A \rightarrow cA \mid a$$

$$B \rightarrow ABC \mid b$$

$$C \rightarrow c$$

How to convert this grammar to CNF?

Conversion into CNF

Step 1: Convert every production into either:

$$A \rightarrow B_1 B_2 \dots B_n \text{ or}$$

$$A \rightarrow a$$

e.g. $A \rightarrow bCDeF$ becomes:

$$A \rightarrow BCDEF$$

$$B \rightarrow b$$

$$E \rightarrow e$$

Conversion into CNF

Step 2: Convert production of the form $A \rightarrow B_1 B_2 \dots B_n$ into $A \rightarrow C_1 C_2$:

e.g. $A \rightarrow BCDEF$ becomes:

$$A \rightarrow BX$$

$$X \rightarrow CY$$

$$Y \rightarrow DZ$$

$$Z \rightarrow EF$$

Class Discussion

Convert the following CFG into Chomsky Normal Form:

$$S \rightarrow \varepsilon$$

$$S \rightarrow ABBA$$

$$B \rightarrow bCb$$

$$A \rightarrow a$$

$$C \rightarrow c$$