Chomsky Normal Form

CNF:

A CFG G is in CNF if every production is of the form $A \to a$ or $A \to BC$ and $S \to \Lambda$ is in G if $\Lambda \in L(G)$.

e.g. We consider a grammar G whose productions are $S \to BC \mid \Lambda$, $B \to b$, $C \to c$ then we can say G is in CNF.

86

Null production

Defn A Context-free grammar may have productions of the form $A \to \Lambda$. The production $A \to \Lambda$ is just used to erase A. So a production of the form $A \to \Lambda$, where $A \in V_N$ i.e. a variable or non-terminal of G, is called a null production.

- A variable A in a Context-free Grammar is nullable if $A \overset{*}{\Rightarrow} \Lambda$.

Theorem: If $G = (V_N, \Sigma, P, S)$ is a CFG, then we can find a CFG G_1 having no null productions such that

$L(G_1) = L(G) - \{\Lambda\}$

Conversion of G_1 from G

step-1 Construction of the set of nullable variables

i) $W_1 = \{A \in V_N \mid A \to \Lambda \text{ is in } P\}$

ii) $W_{i+1} = W_i \cup \{A \in V_N \mid$ there exists a production $A \to \alpha$ with $\alpha \in W_i^*\}$

(7)

step-2 Construction of P':

i) Any production whose R.H.S. does not have any nullable variable is included in P'

ii) If $A \to X_1 X_2 \cdots X_k$ is in P, the productions of the form $A \to \alpha_1 \alpha_2 \cdots \alpha_k$ are included in P', where $\alpha_i = X_i$ if $X_i \notin W$ or $\alpha_i = X_i$ if $X_i \in W$ and $\alpha_1 \alpha_2 \cdots \alpha_k \neq \Lambda$ and rule ii) will gives various productions in P'. The productions are obtained either by not erasing any nullable variable on the R.H.S of $A \to X_1 X_2 \cdots X_k$ or by erasing some or all nullable variables.

Prob:

Consider the grammar G whose productions are $S \to aS \mid AB$, $A \to \wedge$, $B \to \wedge$, $D \to b$, Construct a grammar G_1 without null productions generating $L(G) - \{\wedge\}$.

(10)

Elimination of Unit Productions

Defn A CFG may have productions of the form $A \to B$, $A, B \in V_N$, called unit production or chain rule.

Theorem
If G is the original grammar and G_1 is the grammar without unit productions then we can say $L(G) = L(G_1)$

Conversion of G_1 from G

Step-1 Construction of the set of variables derivable from A, defined as $W_i(A)$ recursively as follows:

$$W_0(A) = \{A\}$$

$$W_{i+1}(A) = W_i(A) \cup \{B \in V_N \mid C \to B \text{ is in } P \text{ with } C \in W_i(A)\}$$

Step-2 Constructions of A-productions in G_1

The A-productions in G_1 are either

a) the nonunit production in G, or

b) $A \rightarrow \alpha$ whenever $B \rightarrow \alpha$ is in G with $B \in W(A)$ and $\alpha \notin V_N$.

Prob 1: (11) Let G be $S \rightarrow AB$, $A \rightarrow a$, $B \rightarrow C/b$, $C \rightarrow D$, $D \rightarrow E$, $E \rightarrow a$. Eliminate unit productions and obtain an equivalent grammar

Conversion of CNF from the grammar which is not in CNF

Step 1 Elimination of null production and Unit productions:

Step 2 Elimination terminals on R.H.S.

Step 3 Restricting the number of variables on R.H.S.

Prob 1:

Reduce the following grammar G_2 to CNF, G_2 is $S \to aAD$, $A \to aB \mid bAB$, $B \to b$, $D \to d$

12

Prob 2:

Find ~~the~~ a grammar in CNF equivalent to $S \to aAbB$, $A \to aA \mid a$, $B \to bB \mid b$.

Prob 3:

Find a grammar in CNF equivalent to the grammar

$$S \to \sim S \mid [S \supset S] \mid p \mid q$$