Chomsky Normal Form

- -> A standardized form for a CFG.
- One of the main goals is to be able to check if string w is generated by grammar G.
- -> What we want: Simple rules that are every to check.
- -> What we don't want: loops, Empty definations Vieles rules.

Consider $S \rightarrow 0S1 \mid 1SOS1 \mid T$ $T \rightarrow S \mid E$

10 (0000 10 \$ 10 TOO 10 \$ 10 DO

- cac

Deb 2.8: A content-free grammar is in Chamsley

Normal Form if every rule is of the form $A \rightarrow BC$ $A \rightarrow BC$ $A \rightarrow CDC$ $A \rightarrow a$

where a is a terminal, A, B, C are variables and B and C are not the start variable. Also, we allow $S \rightarrow E$.

One of the advantages of the Chamshy Normal Form is the "predictability" in the derivation of a string. Any string of lingth requires exactly 2n-1 steps in its defination.

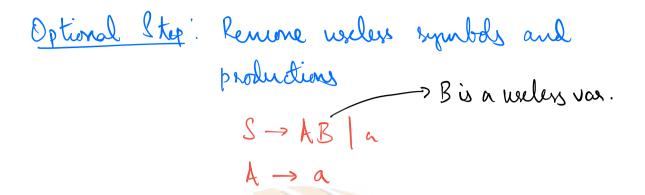
EXERCISE!

Also Chousley NF results in an efficient algorithm for checking if w is generated by Gr.

Natural Question at this stage. What if there is no equivalent CFh that is in the Chamsley Normal Form?

Theorem 2.9: Any CFL is generated by a CFG in Chameley Normal Form.

Proof: We provide a process that demonstrates have to correct any CFG into Chamshy NF.



1. Add new start nariable So.

This guarantees that So does not appear in the RHS of any rule.

2. Remone & rules.

Sony there is the rule $A \rightarrow E$. Then modify rules with A in RHS.

If R -> u Av was wrule, then replace it with

R -> u Av | uv

(here u, v are string of terminals and variables)

If R > A is a rule, then replace it with R > A 18

unles h > E was already removed.

3. Remone unit rules like A -> B.

If there is a rule B-su, add A-su unless we removed that rule.

4. Restaucture rules with long RHS.

Example: S -> TST/aB

T -> B/S

B -> b/E

1. Add lo as new start variable.

 $2 \leftarrow 62$

$$S \longrightarrow TST \mid aB$$

$$T \longrightarrow B \mid S$$

$$B \longrightarrow b \mid E$$

$$S_0 \rightarrow S$$

$$S \rightarrow TST \mid aB \mid a$$

$$T \rightarrow B \mid E \mid S$$

$$B \rightarrow b$$

$$\begin{array}{c} T \rightarrow B \mid S \\ R \rightarrow b \end{array}$$

 $T \rightarrow S$ $S_0 \rightarrow TST \mid TS \mid ST \mid aB \mid a$ $S \rightarrow TST \mid TS \mid ST \mid aB \mid a$ $T \rightarrow TST \mid TS \mid ST \mid aB \mid a \mid b$ $S \rightarrow b$

4. Restaucture rules with more than one symbol in the RHS.

Let TS = Z, a = A. $S_0 \Rightarrow ZT \mid TS \mid ST \mid AB \mid a$ $S \Rightarrow ZT \mid TS \mid ST \mid AB \mid a$ $T \Rightarrow ZT \mid TS \mid ST \mid AB \mid a \mid b$ $Z \Rightarrow TS$ $A \Rightarrow a$ $B \Rightarrow b$

Now the above CFG is in Chamdley Normal Form and by construction, we have evered that it is equivalent to the original gamenas.