

Teaching goals: The student is able to

- give the formal definition of Chomsky Normal Form and related notions
- convert a given context-free grammar to ChNF
- explain the CYK algorithm, apply to a given word and context-free grammar

IN-CLASS PROBLEMS

Problem 1 (About the conversion to ChNF). Recall the process of converting a context-free grammar to Chomsky Normal Form. Then answer the following questions. Justify.

- Find an example of a grammar in which there is a generating variable only reachable via nongenerating variables.
- When reducing a grammar, which variables do we need to remove first: nongenerating or unreachable?
- Is it possible for a reachable generating variable to become nongenerating after the removal of unreachable variables?
- When we want to break up a production rule with long body, what is the minimal number of Chomsky Normal Form rules we need to create?

Problem 2 (Convert to ChNF). Convert the following context-free grammars to Chomsky normal form:

- $G_1 = (\{S, A, B\}, \{0, 1\}, S, \mathcal{P})$, where

$$\mathcal{P} = \{S \rightarrow 0AB, \\ A \rightarrow 0A0 \mid 11, \\ B \rightarrow 0\}$$
- $G_2 = (\{S, A, B\}, \{0, 1\}, S, \mathcal{P})$, where

$$\mathcal{P} = \{S \rightarrow 0A10B10, \\ A \rightarrow 1A0 \mid \epsilon, \\ B \rightarrow 1B00 \mid \epsilon\}$$

Problem 3 (The CYK algorithm). Using the CYK algorithm determine if $w \in L(G)$.

- $w = 0110$, $G = (\{S, A, B\}, \{0, 1\}, S, \mathcal{P})$, where

$$\mathcal{P} = \{S \rightarrow 0 \mid AB, \\ A \rightarrow 1 \mid SA \mid SB, \\ B \rightarrow AS \mid BA \mid 0\}$$

- $w = 001100$, $G = G_1$ is the grammar from Problem ??(a)
- $w = 110011$, $G = G_1$ is the grammar from Problem ??(a)

EXTRA PRACTICE AND THINKING

Problem 4 (Convert to ChNF). Convert the following to Chomsky normal form:

- (a) $G = (\{S, A, B\}, \{0, 1\}, S, \mathcal{P})$ (b) $G = (\{S, E, F\}, \{(\cdot), *, +, \cdot, 1\}, S, \mathcal{P})$
- $$\begin{aligned} \mathcal{P} = \{ & S \rightarrow A \mid 0SA \mid \epsilon, & \mathcal{P} = \{ & S \rightarrow (E), \\ & A \rightarrow 1A \mid 1 \mid B1, & & E \rightarrow F + F \mid F * F, \\ & B \rightarrow 0B \mid 0 \mid \epsilon \} & & F \rightarrow S \mid 1 \} \end{aligned}$$

Problem 5 (The CYK algorithm). Using the CYK algorithm determine if $w \in L(G)$.

- (a) $w = abcb$, $G = (\{S, A, B, C\}, \{a, b, c\}, S, \mathcal{P})$, where

$$\begin{aligned} \mathcal{P} = \{ & S \rightarrow CA \mid CB, \\ & B \rightarrow CBA \mid CB \mid BA \mid BB, \\ & C \rightarrow ABC \mid BC, \\ & A \rightarrow a, B \rightarrow b, C \rightarrow c \} \end{aligned}$$

- (b) $w = 01010010$, $G = G_2$ is the grammar from Problem ??(b)
 (c) $w = 01010011$, $G = G_2$ is the grammar from Problem ??(b)