NTIN071 A&G: Tutorial 7 – Chomsky normal form, The CYK algorithm

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 w and context-free grammar G

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.

(a) Find an example of a grammar in which there is a generating variable only reachable via nongenerating variables.

(b) When reducing a grammar, which variables do we need to remove first: nongenerating or unreachable?

(c) Is it possible for a reachable generating variable to become nongenerating after the removal of unreachable variables?

(d) 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:

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

$$\begin{aligned}\mathcal{P} = \{S &\to 0AB,\\ A &\to 0A0 \mid 11,\\ B &\to 0\}\end{aligned}$$

(b) $G_2 = (\{S, A, B\}, \{0, 1\}, S, \mathcal{P})$, where

$$\begin{aligned}\mathcal{P} = \{S &\to 0A10B10,\\ A &\to 1A0 \mid \epsilon,\\ B &\to 1B00 \mid \epsilon\}\end{aligned}$$

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

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

$$\begin{aligned}\mathcal{P} = \{S &\to 0 \mid AB,\\ A &\to 1 \mid SA \mid SB,\\ B &\to AS \mid BA \mid 0\}\end{aligned}$$

(b) $w = 001100$, $G = G_1$ is the grammar from Problem 2(a)

(c) $w = 110011$, $G = G_1$ is the grammar from Problem 2(a)

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

(a) $G = (\{S, A, B\}, \{0, 1\}, S, \mathcal{P})$

$$\begin{aligned}\mathcal{P} = \{S &\to A \mid 0SA \mid \epsilon,\\ A &\to 1A \mid 1 \mid B1,\\ B &\to 0B \mid 0 \mid \epsilon\}\end{aligned}$$

(b) $G = (\{S, E, F\}, \{(,), *, +, , 1\}, S, \mathcal{P})$

$$\begin{aligned}\mathcal{P} = \{S &\to (E),\\ E &\to F + F \mid F * F,\\ F &\to S \mid 1\}\end{aligned}$$

Problem 5 (The CYK algorithm)**.** Using the CYK algorithm determine if $w \in L(G)$.
(a) $w = abcbb$, $G = (\{S, A, B, C\}, \{a, b, c\}, S, \mathcal{P})$, where

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

(b) $w = 01010010$, $G = G_2$ is the grammar from Problem 2(b)

(c) $w = 01010011$, $G = G_2$ is the grammar from Problem 2(b)