Given a CFL A and a string x, how to determine if $x \in L$?

Consider CFG generating A.

$$\text{CFG} = (\Sigma, \Gamma, P)$$

Chomsky Normal Form

A CFG (Σ, I, P) is in Chomskey Normal Form if every production in P has one of following two forms:

$$B \to a$$

$$B \to CD$$

Converting a CFG to Chomsky Normal Form

For every $a \in \Sigma$, introduce a new production $T_a \rightarrow a$

For every production $B \rightarrow \gamma$ in P, replace terminals in γ by newly introduced non-terminals unless $\gamma \in \Sigma$

Productions now have the following forms:

$$B \to a$$

Replace $B \to C_1 \cdots C_k$ $\leftarrow$ $B \to C_1 C_2 \cdots C_k \,,\; k > 1$

by
$$B \to C_1 R_2$$
$$R_2 \to C_2 R_3$$
$$R_3 \to C_3 R_4$$
$$\vdots$$
$$R_{k-1} \to C_{k-1} C_k$$

new non-term

$$B \to \epsilon$$

$$B \to C\, T_\epsilon$$
$$T_\epsilon \to \epsilon$$

Replace $C \to \gamma_1 B \gamma_2 B \gamma_3$

by
$$C \to \gamma_1 \gamma_2 B \gamma_3$$
$$C \to \gamma_1 \gamma_2 \gamma_3$$
$$C \to \gamma_1 B \gamma_2 \gamma_3$$
$$C \to \gamma_1 B \gamma_2 B \gamma_3$$

Idea: Given string $x \in \Sigma^*$, $x = a_1 a_2 a_3 \cdots a_n$,
compute the set of non-terminals T_{ij} that generate
string $x_{ij} = a_i a_{i+1} \cdots a_j$ for every i & j.

Computing $T_{i,i}$: $T_{i,i} = \{ C \mid C \to a_i$ is a production $\}$

Computing $T_{i,i+1}$: For every production $B \to C_1 C_2$, add B to $T_{i,i+1}$ if $C_1 \in T_{i,i}$ & $C_2 \in T_{i+1,i+1}$

Computing $T_{i,j}$: For every k, $i \le k < j$, for every production $B \to C_1 C_2$, add B to $T_{i,j}$ if $C_1 \in T_{i,k}$ & $C_2 \in T_{k+1,j}$

Finally, check if $S \in T_{1,n}$.

<u>Time Complexity</u>: $O(n^3)$

<u>Time Complexity for regular sets</u>: $O(n)$