A context-free grammar is in **Chomsky Normal Form (CNF)** if all the rules are of the form:

- $A \longrightarrow BC$
- $A \longrightarrow a$

A, B, C are variables (not necessarily distinct), a a terminal.

Every context-free grammar can be transformed to an equivalent grammar (i.e., generating the same set of strings of terminals) in CNF.

Given a grammar G in CNF, and a string $w = a_1 a_2 \ldots a_n$, we can test whether $w \in L(G)$ using the CYK algorithm.

The CYK algorithm, on input $\langle G, w = a_1 a_2 \ldots a_n \rangle$, builds an $n \times n$ table T, where each entry contains a subset of V.

At the end, $w \in L(G)$ if the start variable is in position $(1, n)$.

Let $V = \{X_1, X_2, \ldots, X_m\}$. Initialize T as follows:

for $(i = 1; i \leq n; i++)$
  for $(j = 1; j \leq m; j++)$
    Place variable X_j in (i, i) iff $X_j \longrightarrow a_i$ is a rule of G

Then, for $i < j$, the contents of (i, j) are computed as follows:

for $(k = i; k < j; k++)$
  if $(\exists X_p \in (i, k) \wedge \exists X_q \in (k+1, j)$ & $\exists$ rule $X_r \longrightarrow X_p X_q)$
   Put X_r in (i, j)

Eg.

x	(2,2)	(2,3)	(2,4)	**(2,5)**
x	x			(3,5)
x	x	x		(4,5)
x	x	x	x	(5,5)