Chomsky normal form (CNF) I

- Purpose: a simplified form of grammars
- Every rule must be either

$$A \to BC$$

or

$$A \to a$$

- B, C are not start variables
- $a \in \Sigma$ so

$$A \to \epsilon$$

is not allowed.

Chomsky normal form (CNF) II

- However,

$$S \to \epsilon$$

  is allowed, where S is the start variable
- This form is useful later (but not in this chapter)
- To convert a CFG to a CNF, let's show an example first

Example to convert CFG to CNF I

- The original CFG

$$S \to ASA \mid aB$$
$$A \to B \mid S$$
$$B \to b \mid \epsilon$$

Example to convert CFG to CNF II

- Add

$$S_0 \to S$$

because the start variable cannot be on the right

$$\begin{aligned} S_0 &\to S \\ S &\to ASA \mid aB \\ A &\to B \mid S \\ B &\to b \mid \epsilon \end{aligned}$$

Example to convert CFG to CNF III

- Remove

$$B \to \epsilon$$

because ϵ cannot be on the right

$$S_0 \to S$$
$$S \to ASA \mid aB \mid a$$
$$A \to B \mid \epsilon \mid S$$
$$B \to b$$

Example to convert CFG to CNF IV

- Remove $A \to \epsilon$

$$S_0 \to S$$
$$S \to ASA \mid aB \mid a \mid AS \mid SA \mid S$$
$$A \to B \mid S$$
$$B \to b$$

- What if

$$B \to \epsilon$$

appears again? An infinite loop? We will discuss this issue later

Example to convert CFG to CNF V

- Remove $S \to S$ because the right-hand side cannot be a single variable

$$
\begin{aligned}
S_0 &\to S \\
S &\to ASA \mid aB \mid a \mid AS \mid SA \\
A &\to B \mid S \\
B &\to b
\end{aligned}
$$

Example to convert CFG to CNF VI

- Remove $S_0 \to S$

$$S_0 \to ASA \mid aB \mid a \mid AS \mid SA$$
$$S \to ASA \mid aB \mid a \mid AS \mid SA$$
$$A \to B \mid S$$
$$B \to b$$

Example to convert CFG to CNF VII

- Remove $A \to B, A \to S$

$$S_0 \to ASA \mid aB \mid a \mid AS \mid SA$$
$$S \to ASA \mid aB \mid a \mid AS \mid SA$$
$$A \to b \mid ASA \mid aB \mid a \mid AS \mid SA$$
$$B \to b$$

Example to convert CFG to CNF VIII

- Finally

$$
\begin{aligned}
S_0 &\to AA_1 \mid UB \mid a \mid AS \mid SA \\
S &\to AA_1 \mid UB \mid a \mid AS \mid SA \\
A &\to b \mid AA_1 \mid UB \mid a \mid AS \mid SA \\
A_1 &\to SA \\
U &\to a \\
B &\to b
\end{aligned}
$$