Chomsky Normal Form

Ref: prof chih-jen lin
Neso Academy

Chomsky Normal Form

A CFG is in Chomsky Normal Form if the Productions are in the following forms –

- $A \rightarrow a$
- $A \rightarrow BC$
- $S \rightarrow \varepsilon$

where A, B, and C are non-terminals and **a** is terminal.

Chomsky Normal Form

- 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

- 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

Chomsky Normal Form

- The original CFG

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

Chomsky Normal Form

- Add

$$S_0 \to S$$

because the start variable cannot be on the right

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

Chomsky Normal Form

- Remove

$$B \to \epsilon$$

because ϵ cannot be on the right

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

Chomsky Normal Form

- Remove $A \to \epsilon$

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

- What if

$$B \to \epsilon$$

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

Chomsky Normal Form

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

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

Chomsky Normal Form

- Remove $S_0 \to S$

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

Chomsky Normal Form

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

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

Chomsky Normal Form

- Finally

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

Convert Context-free to Chomsky Normal Form

A procedure to summarize what we have done in the example

- Add

$$S_0 \to S$$

  So start state not on the right
- Remove $A \to \epsilon$, where A is not the start state:

Convert Context-free to Chomsky Normal Form

For any rule of

$$\cdots \rightarrow uAv$$

add

$$\cdots \rightarrow uv$$

We discuss the issue of a possible infinite loop later

Convert Context-free to Chomsky Normal Form

- Remove

$$A \to B$$

because the right hand cannot have a single variable. For any

$B \to u$, where u is a string of variables and terminals

we

remove $A \to B$ and $B \to u$, and add $A \to u$

Convert Context-free to Chomsky Normal Form

unless $A \to u$ is a unit rule previously removed (this setting avoids the possible infinite loop)

- After this, we have either

$$A \to u_1 \cdots u_k, u_i \in V \text{ or } \Sigma;$$

and

$$\text{if } k = 1, \text{ then } u_i \in \Sigma$$

Convert Context-free to Chomsky Normal Form

- Replace the right side with

$$A \to u_1 A_1$$
$$A_1 \to u_2 A_2$$
$$\ldots$$

- Replace any u_i in the above rules with U_i
- Add

$$U_i \to u_i \text{ if } u_i \in \Sigma$$

Infinite Loop

- Original rules

$$S \to B \mid \epsilon$$
$$B \to S \mid \epsilon$$

- Add S_0

$$S_0 \to S$$
$$S \to B \mid \epsilon$$
$$B \to S \mid \epsilon$$

Infinite Loop

- Remove $S \to \epsilon$

$$S_0 \to S \mid \epsilon$$
$$S \to B$$
$$B \to S \mid \epsilon$$

- Remove $B \to \epsilon$

$$S_0 \to S \mid \epsilon$$
$$S \to B \mid \epsilon$$
$$B \to S$$

Infinite Loop

- No need to add $S \rightarrow \epsilon$
- Reason: $S \rightarrow \epsilon$ has been handled; see line -8 of p109 in the textbook.

CNF - Examples

Problem
Convert the following CFG into CNF

S -> a B c C

CNF - Examples

Problem
Convert the following CFG into CNF
S -> a B c C

CFG IN Chomsky normal form.
S -> H_1 H_0
H_0 -> C
H_1 -> H_3 H_2
H_2 -> c
H_3 -> H_4 H_5
H_4 -> a
H_5 -> B

CNF - Examples

Problem
Convert the following CFG into CNF

CFG
S -> a B B B | b A A A

CNF - Examples

Problem
Convert the following CFG into CNF

CFG
S -> a B B B | b A A A

CFG IN Chomsky normal form
S -> H_1 H_0 | H_3 H_2
H_0 -> B
H_1 -> H_4 H_0
H_2 -> A
H_3 -> H_5 H_2
H_4 -> H_6 H_0
H_5 -> H_7 H_2
H_6 -> a
H_7 -> b

CNF - Examples

Problem
Convert the following CFG into CNF

CFG
S -> b A A

CNF - Examples

Problem
Convert the following CFG into CNF

CFG
S -> b A A

CFG IN Chomsky normal form.
S -> H_1 H_0
H_0 -> A
H_1 -> H_2 H_0
H_2 -> b

CNF - Examples

Problem
Convert the following CFG into CNF

CFG
S -> C b A D A

CNF - Examples

Problem
Convert the following CFG into CNF

CFG
S -> C b A D A

CFG IN Chomsky normal form.
S -> H_1 H_0
H_0 -> A
H_1 -> H_3 H_2
H_2 -> D
H_3 -> H_4 H_0
H_4 -> H_5 H_6
H_5 -> C
H_6 -> b