Lecture 6
Simplification of Context-Free Grammars and Normal Forms

COSE215: Theory of Computation

Seunghoon Woo

Fall 2023

Contents

- **Normal Forms**

Normal Forms

- **There are many normal forms we can establish for CFGs**
- **We consider two major normal forms**
  - Chomsky Normal Form (CNF)
  - Greibach Normal Form (GNF)

Normal Forms

- **Advantages**
  - Grammars in this form are far easier to analyze
  - This can remove ambiguity
  - All derivations can be represented by binary trees (for CNF)
  - Useful for providing the equivalence of CFG and PDA (for GNF)
  - …

Chomsky Normal Forms

- **Definition**
  - Strictly limiting the number of symbols on the right of a production
  - A CFG (V, T, S, P) is in Chomsky normal form if all Ps are of the form
    - ❖ $A \rightarrow BC$ ($A \in V$ and $B, C \in V \setminus \{S\}$)

      or
    - ❖ $A \rightarrow x$ ($x \in T$)

      or
    - ❖ $S \rightarrow \lambda$

Chomsky Normal Forms

- **Any CFG can be converted to CNF**
  - Example
    - $S \rightarrow ABa$
    - $A \rightarrow aab$
    - $B \rightarrow Ac$

Chomsky Normal Forms

- Example
  - $S \rightarrow ABa$
  - $A \rightarrow aab$
  - $B \rightarrow Ac$

- **Any CFG can be converted to CNF**

  <u>Step (0):</u> If S in the right-hand side of a rule, add a new start variable S' and a production $S' \rightarrow S$

Chomsky Normal Forms

- Example
  - $S \rightarrow ABa$
  - $A \rightarrow aab$
  - $B \rightarrow Ac$

- **Any CFG can be converted to CNF**

  Step (1): Eliminate $\lambda-$productions, unit productions, and useless variables

Chomsky Normal Forms

- Example
  - $S \rightarrow ABa$
  - $A \rightarrow aab$
  - $B \rightarrow Ac$

- **Any CFG can be converted to CNF**

  Step (2): If a terminal symbol a appears in a right-hand side of a rule, replace it with a new variable A and add a production $A \rightarrow a$

Chomsky Normal Forms

- Example
  - $S \rightarrow ABa$
  - $A \rightarrow aab$
  - $B \rightarrow Ac$

- **Any CFG can be converted to CNF**

  Step (2): If a terminal symbol a appears in a right-hand side of a rule, replace it with a new variable A and add a production $A \rightarrow a$

  - Consider the original grammar
    - Add $B_a \rightarrow a$, $B_b \rightarrow b$ and $B_c \rightarrow c$
    - Then the original production rules can be changed to
      - $S \rightarrow ABB_a$
      - $A \rightarrow B_aB_aB_b$
      - $B \rightarrow AB_c$
      - $B_a \rightarrow a$
      - $B_b \rightarrow b$
      - $B_c \rightarrow c$

Chomsky Normal Forms

- Example
  - ❖ $S \rightarrow ABa$
  - ❖ $A \rightarrow aab$
  - ❖ $B \rightarrow Ac$

- **Any CFG can be converted to CNF**

  Step (3): If a rule has more than two variables in the right-hand side, replace them with a chain of variables

  - Consider the original grammar
    - ❖ Split production rules for S and A
    - ❖ Then the original production rules can be changed to
      - $S \rightarrow AD_1$
      - $D_1 \rightarrow BB_a$
      - $A \rightarrow B_aD_2$
      - $D_2 \rightarrow B_aB_b$
      - $B \rightarrow AB_c$
      - $B_a \rightarrow a$
      - $B_b \rightarrow b$
      - $B_c \rightarrow c$

Chomsky Normal Forms

- Example
  - $S \rightarrow ABa$
  - $A \rightarrow aab$
  - $B \rightarrow Ac$

- **Any CFG can be converted to CNF**

  Step (4): If λ is contained in the original CFG, then add a production $S \rightarrow \lambda$ (or $S' \rightarrow \lambda$)

  - $S \rightarrow AD_1$
  - $D_1 \rightarrow BB_a$
  - $A \rightarrow B_aD_2$
  - $D_2 \rightarrow B_aB_b$
  - $B \rightarrow AB_c$
  - $B_a \rightarrow a$
  - $B_b \rightarrow b$
  - $B_c \rightarrow c$

Chomsky Normal Forms

- **Any CFG can be converted to CNF**
  - Example
    - $S \rightarrow aSb \mid \lambda$

Chomsky Normal Forms

- **Any CFG can be converted to CNF**
  - Example
    - $S \rightarrow aSb \mid \lambda$

$S' \rightarrow S$
$S \rightarrow aSb \mid \lambda$

Chomsky Normal Forms

- **Any CFG can be converted to CNF**
  - Example
    - $S \rightarrow aSb \mid \lambda$

$S' \rightarrow S$ $S \rightarrow aSb \mid \lambda$

➡

$S' \rightarrow aSb \mid ab$ $S \rightarrow aSb \mid ab$

Chomsky Normal Forms

- **Any CFG can be converted to CNF**
  - Example
    - $S \rightarrow aSb \mid \lambda$

$$\begin{array}{l} S' \rightarrow S \\ S \rightarrow aSb \mid \lambda \end{array} \Rightarrow \begin{array}{l} S' \rightarrow aSb \mid ab \\ S \rightarrow aSb \mid ab \end{array} \Rightarrow \begin{array}{l} S' \rightarrow ASB \mid AB \\ S \rightarrow ASB \mid AB \\ A \rightarrow a \\ B \rightarrow b \end{array}$$

Chomsky Normal Forms

- **Any CFG can be converted to CNF**
  - Example
    - $S \rightarrow aSb \mid \lambda$

$$\begin{aligned} S' &\rightarrow S \\ S &\rightarrow aSb \mid \lambda \end{aligned} \quad \Rightarrow \quad \begin{aligned} S' &\rightarrow aSb \mid ab \\ S &\rightarrow aSb \mid ab \end{aligned} \quad \Rightarrow \quad \begin{aligned} S' &\rightarrow ASB \mid AB \\ S &\rightarrow ASB \mid AB \\ A &\rightarrow a \\ B &\rightarrow b \end{aligned} \quad \Rightarrow \quad \begin{aligned} S' &\rightarrow AS_1 \mid AB \\ S &\rightarrow AS_1 \mid AB \\ S_1 &\rightarrow SB \\ A &\rightarrow a \\ B &\rightarrow b \end{aligned}$$

Chomsky Normal Forms

- **Any CFG can be converted to CNF**
  - Example
    - $S \rightarrow aSb \mid \lambda$

$$S' \rightarrow S \\ S \rightarrow aSb \mid \lambda$$

➡

$$S' \rightarrow aSb \mid ab \\ S \rightarrow aSb \mid ab$$

➡

$$S' \rightarrow ASB \mid AB \\ S \rightarrow ASB \mid AB \\ A \rightarrow a \\ B \rightarrow b$$

➡

$$S' \rightarrow AS_1 \mid AB \\ S \rightarrow AS_1 \mid AB \\ S_1 \rightarrow SB \\ A \rightarrow a \\ B \rightarrow b$$

➡

$$S' \rightarrow AS_1 \mid AB \mid \lambda \\ S \rightarrow AS_1 \mid AB \\ S_1 \rightarrow SB \\ A \rightarrow a \\ B \rightarrow b$$

Greibach Normal Forms

- **Definition**
  - Restrict not the length of a production, but the **positions**
  - A CFG (V, T, S, P) is in Greibach normal form if all Ps are of the form
    - $A \rightarrow ax$ ($a \in \Sigma$ and $x \in V^*$)
  - Conversion is not always a simple matter..

Greibach Normal Forms

- **Definition**
  - Example
    - $S \rightarrow AB$
    - $A \rightarrow aA \mid bB \mid b$
    - $B \rightarrow b$

Greibach Normal Forms

- **Definition**
  - Example
    - $S \rightarrow AB$
    - $A \rightarrow aA \mid bB \mid b$
    - $B \rightarrow b$
  - This CFG can be converted to GNF as follows
    - $S \rightarrow aAB \mid bBB \mid bB$
    - $A \rightarrow aA \mid bB \mid b$
    - $B \rightarrow b$

Next Lecture

- **A membership algorithm for CFG**