Lecture 6
Simplification of Context-Free Grammars and Normal Forms

COSE215: Theory of Computation

Seunghoon Woo

Fall 2023

Contents

Context-Free Grammars

- **CFG can contain some useless rules**
  - E.g., Given a context-free grammar G with the production rules
    - $S \rightarrow aSb \mid \lambda \mid A$
    - $A \rightarrow aA$
  - Here, $S \rightarrow A$ plays no role
    - A cannot be transformed into a terminal string (never lead to a sentence)
  - Therefore, we can remove $S \rightarrow A$ and $A \rightarrow aA$
    - The language is unaffected!

Context-Free Grammars

- **Three simplification techniques**
    1. Eliminating λ $-$productions
    2. Eliminating unit productions
    3. Eliminating useless variables

Simplification of Context-Free Grammars

- **Eliminating λ−productions**
  - Any production of a CFG of the form $A \to \lambda$ is called a **λ−production**
  - Any variable A for which the derivation $A \overset{*}{\Rightarrow} \lambda$ is possible is called **nullable**
- **A grammar may generate a language not containing λ yet have some λ-productions or nullable variables**
  - In this case, the λ-productions can be removed

Simplification of Context-Free Grammars

- **Eliminating $\lambda-$productions: example**
  - Consider the grammar
    - $S \rightarrow aS_1b$
    - $S_1 \rightarrow aS_1b \mid \lambda$

  which generates the language $L(G) = \{a^nb^n : n \geq 1\}$ (where $\lambda \notin L(G)$)
  - Therefore, we can remove $\lambda-$productions
    - $S \rightarrow aS_1b \mid ab$
    - $S_1 \rightarrow aS_1b \mid ab$

Simplification of Context-Free Grammars

- **How to eliminate λ−productions?**
  1. Find all nullable variables
  2. Construct a new CFG with production rules produced by replacing nullable variables with λ in all combinations, except for the λ-production

Simplification of Context-Free Grammars

- **Eliminating $\lambda-$productions: example**
  - Consider the grammar
    - $S \to ABaC$
    - $A \to BC$
    - $B \to b \mid \lambda$
    - $C \to D \mid \lambda$
    - $D \to d$

Simplification of Context-Free Grammars

- **Eliminating $\lambda-$productions: example**
  - Consider the grammar
    - $S \rightarrow ABaC$
    - $A \rightarrow BC$
    - $B \rightarrow b \mid \lambda$
    - $C \rightarrow D \mid \lambda$
    - $D \rightarrow d$
  - Nullable variables: (A, B, C)

Simplification of Context-Free Grammars

- **Eliminating λ−productions: example**
  - Consider the grammar
    - $S \rightarrow ABaC$
    - $A \rightarrow BC$
    - $B \rightarrow b \mid \lambda$
    - $C \rightarrow D \mid \lambda$
    - $D \rightarrow d$
  - Nullable variables: (A, B, C)
  - Replace nullable variables with λ (e.g., replace A, B, C to λ in S)

Simplification of Context-Free Grammars

- **Eliminating $\lambda-$productions: example**
  - Result grammar with no $\lambda-$productions
    - $S \rightarrow ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Aa \mid Ba \mid a$
    - $A \rightarrow BC \mid B \mid C$
    - $B \rightarrow b$
    - $C \rightarrow D$
    - $D \rightarrow d$

Simplification of Context-Free Grammars

- **Eliminating unit productions**
  - Any production of a CFG of the form $A \rightarrow B$ is called a **unit production**
  - A pair of variables $(A, B) \in V \times V$ is a **unit pair** if $A \overset{*}{\Rightarrow} B$
    - ❖ (A, A) is a unit pair for all $A \in V$
- **How to eliminate unit productions?**
  1. Find all unit pairs
  2. Construct a new CFG by adding all possible non-unit productions for each unit pair (A, B)

Simplification of Context-Free Grammars

- **Eliminating unit productions: example**
  - Consider the grammar
    - $S \to Aa \mid B$
    - $A \to a \mid bc \mid B$
    - $B \to A \mid bb$

Simplification of Context-Free Grammars

- **Eliminating unit productions: example**
  - Consider the grammar
    - $S \rightarrow Aa \mid B$
    - $A \rightarrow a \mid bc \mid B$
    - $B \rightarrow A \mid bb$
  - All unit pairs: (S, S), (A, A), (B, B), (S, A), (S, B), (A, B), (B, A)

Simplification of Context-Free Grammars

- **Eliminating unit productions: example**
  - Consider the grammar
    - $S \to Aa \mid B$
    - $A \to a \mid bc \mid B$
    - $B \to A \mid bb$
  - All unit pairs: (S, S), (A, A), (B, B), (S, A), (S, B), (A, B), (B, A)
  - Consider non-unit productions in (S, S), (A, A), and (B, B)
    - $S \to Aa$
    - $A \to a \mid bc$
    - $B \to bb$

Simplification of Context-Free Grammars

- **Eliminating unit productions: example**
  - Consider non-unit productions in (S, S), (A, A), and (B, B)
    - ❖ $S \rightarrow Aa$
    - ❖ $A \rightarrow a \mid bc$
    - ❖ $B \rightarrow bb$
  - Consider possible non-unit productions in (S, A), (S, B), (A, B), (B, A)
    - ❖ $S \rightarrow a \mid bc \mid bb$
    - ❖ $A \rightarrow bb$
    - ❖ $B \rightarrow a \mid bc$

Original grammar

- ❖ $S \rightarrow Aa \mid B$
- ❖ $A \rightarrow a \mid bc \mid B$
- ❖ $B \rightarrow A \mid bb$

Simplification of Context-Free Grammars

- **Eliminating unit productions: example**
  - Result grammar with no unit productions
    - $S \to Aa \mid a \mid bc \mid bb$
    - $A \to a \mid bc \mid bb$
    - $B \to a \mid bc \mid bb$

Simplification of Context-Free Grammars

- **Eliminating unit productions: example**
  - Design an equivalent CFG with no unit productions
    - $E \to E + T \mid T$
    - $T \to T * F \mid F$
    - $F \to (E) \mid a$

Simplification of Context-Free Grammars

- **Eliminating unit productions: example**
  - Result grammar with no unit productions
    - ❖ $E \rightarrow E + T \mid T * F \mid (E) \mid a$
    - ❖ $T \rightarrow T * F \mid (E) \mid a$
    - ❖ $F \rightarrow (E) \mid a$

Simplification of Context-Free Grammars

- **Eliminating useless variables**
  - A variable $A \in V$ is said to be **useful** if and only if there is at least one $w \in L(G)$ such that $S \overset{*}{\Rightarrow} xAy \overset{*}{\Rightarrow} w$ (where $x, y \in (V \cup T)^*$)
  - In other words, a variable is useful if and only if it occurs in at least one derivation
  - A variable that is not useful is called **useless**

Simplification of Context-Free Grammars

- **Eliminating useless variables: example**
  - Consider CFG with production rules
    - $S \to Aa \mid a \mid bc \mid bb$
    - $A \to a \mid bc \mid bb$
    - $B \to a \mid bc \mid bb$
  - The variable B is **useless**!

Simplification of Context-Free Grammars

- **How to eliminate useless variables?**
  1. Find all useless variables
  2. Remove them!

Simplification of Context-Free Grammars

- **Example**
  - Consider CFG with production rules
    - $S \to aS \mid A \mid C$
    - $A \to a$
    - $B \to aa$
    - $C \to aCb$

Simplification of Context-Free Grammars

- **Example**
  - Consider CFG with production rules
    - $S \rightarrow aS \mid A \mid C$
    - $A \rightarrow a$
    - $B \rightarrow aa$
    - $C \rightarrow aCb$
  - B is unreachable from the start variable
  - C cannot derive any words

Simplification of Context-Free Grammars

- **Example**
  - Result grammar with no useless variables
    - $S \to aS \mid A$
    - $A \to a$

Simplification of Context-Free Grammars

- **Practice**
  - **Design an equivalent CFG without unit productions**
    - $S \to 0A \mid 1B \mid C$
    - $A \to 0S \mid 00$
    - $B \to 1 \mid A$
    - $C \to 01$

Simplification of Context-Free Grammars

- **Practice**
  - **Design an equivalent CFG without unit productions**
    - $S \rightarrow 0A \mid 1B \mid C$
    - $A \rightarrow 0S \mid 00$
    - $B \rightarrow 1 \mid A$
    - $C \rightarrow 01$

Simplification of Context-Free Grammars

- **Practice**
  - **Design an equivalent CFG without λ −productions**
    - $S \rightarrow ABC$
    - $A \rightarrow a \mid bbD$
    - $B \rightarrow a \mid \lambda$
    - $C \rightarrow b \mid \lambda$
    - $D \rightarrow c \mid d$

Simplification of Context-Free Grammars

- **Practice**
  - **Design an equivalent CFG without λ −productions**
    - $S \to ABC$
    - $A \to a \mid bbD$
    - $B \to a \mid \lambda$
    - $C \to b \mid \lambda$
    - $D \to c \mid d$

Next Lecture

- **Class review for the midterm exam**