Context-Free Grammars

Lecture 16
Section 5.1

Robb T. Koether

Hampden-Sydney College

Wed, Sep 28, 2016

Outline

Outline

Context-Free Grammars

Definition (Context-free grammar)

A context-free grammar, abbreviated CFG, is a 4-tuple (V, T, S, P) with variables V, terminals T, start symbol S, and productions P, where the productions have the following form:

- The left-hand side is a single variable.
- The right-hand side is any string of variables and terminals.
- That is, every production is of the form

$$A \to x,$$

where $x \in (V \cup T)^*$.

Outline

Example

Example (Context-free grammar)

Let

- $V = \{E\}$
- $T = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, +, *, (,)\}$
- $S = E$
- P consists of the rules

$$E \to E + E \mid E * E \mid (E) \mid \mathbf{a} \mid \mathbf{b} \mid \mathbf{c}$$

Examples

Example (Context-free grammars)

Find CFGs for the following languages (none of which are regular).

- $\{\mathbf{a}^n\mathbf{b}^n \mid n \geq 0\}$
- $\{\mathbf{a}^n\mathbf{b}^m \mid n \geq m \geq 0\}$
- $\{\mathbf{a}^n\mathbf{b}^m \mid m \geq n \geq 0\}$

Outline

Example

Example (Equal number of a's and b's)

- Find a CFG for the language

$$\{w \mid w \text{ contains an equal number of } \mathbf{a}\text{'s and } \mathbf{b}\text{'s}\}.$$

Example

Example (Equal number of a's and b's)

- Let w be a string with an equal number of **a**'s and **b**'s.
- There are four cases:
  - $w = \mathbf{a}z\mathbf{a}$
  - $w = \mathbf{a}z\mathbf{b}$
  - $w = \mathbf{b}z\mathbf{a}$
  - $w = \mathbf{b}z\mathbf{b}$

  where z is a string in Σ^*.
- In each case, what do we know about z?

Example

Example (Equal number of a's and b's)

- In the cases $w = \mathbf{a}z\mathbf{b}$ and $w = \mathbf{b}z\mathbf{a}$, z is again a string with an equal number of **a**'s and **b**'s.
- So we may use the rules $S \to \mathbf{a}S\mathbf{b}$ and $S \to \mathbf{b}S\mathbf{a}$.

Example

Example (Equal number of **a**'s and **b**'s)

- In the cases $w = \mathbf{a}z\mathbf{a}$ and $w = \mathbf{b}z\mathbf{b}$, the string w must necessarily split into a concatenation of strings

$$w = w_1 w_2,$$

where w_1 and w_2 each have an equal number of **a**'s and **b**'s.

- Why?
- So we may add the rule $S \to SS$.
- These rules are sufficient.

Example

Example (Ratio of **a**'s to **b**'s)

- Find a CFG for the language

$$\{w \mid w \text{ contains twice as many } \mathbf{a}\text{'s as } \mathbf{b}\text{'s}\}.$$

- Find a CFG for the language

$$\{w \mid w \text{ contains } 3/4 \text{ as many } \mathbf{a}\text{'s as } \mathbf{b}\text{'s}\}.$$

- Find a CFG for the language

$$\{w \mid \text{ the ratio of } \mathbf{a}\text{'s to } \mathbf{b}\text{'s is } m : n\}.$$

Outline

Definition (Yields)

A string u yields a string v if we can apply *one* grammar rule to u and get v. We write $u \Rightarrow v$.

Derivations

Definition (Derives)

A string u derives a string v if there is a sequence

$$u_1, u_2, \ldots, u_k,$$

with $k \geq 1$, where $u_1 = u$, $u_k = v$, and

$$u_1 \Rightarrow u_2 \Rightarrow \cdots \Rightarrow u_k.$$

The sequence is called a derivation. We write $u \stackrel{*}{\Rightarrow} v$.

Outline

Leftmost and Rightmost Derivations

Definition (Leftmost derivation)

A leftmost derivation of a string is a derivation in which, at each step, the leftmost variable is replaced with a string.

Definition (Rightmost derivation)

A rightmost derivation of a string is a derivation in which, at each step, the rightmost variable is replaced with a string.

Example

Example (Leftmost and rightmost derivations)

Using the grammar

$$S \to SS \mid \mathbf{a}S\mathbf{b} \mid \mathbf{b}S\mathbf{a} \mid \lambda,$$

find leftmost and rightmost derivations of **abab.**

Outline

Ambiguity

- Some grammars provide two or more fundamentally different ways to derive a string.
- For example, **abab** can be derived in two different ways using the grammar rules

$$S \to SS \mid \mathbf{a}S\mathbf{b} \mid \mathbf{b}S\mathbf{a} \mid \lambda.$$

Ambiguity

- Some grammars provide two or more fundamentally different ways to derive a string.
- For example, **abab** can be derived in two different ways using the grammar rules

$$S \to SS \mid \mathbf{a}S\mathbf{b} \mid \mathbf{b}S\mathbf{a} \mid \lambda.$$

- Does the grammar indicate which **a**'s should be paired with which **b**'s?

Ambiguity

Definition (Ambiguous grammar)

A grammar is ambiguous if its language contains a string that has more than one leftmost derivation under that grammar.

Definition (Inherently ambiguous language)

A language is inherently ambiguous if every grammar for that language is ambiguous.

Ambiguity

- Consider again the grammar rules

$$S \rightarrow SS \mid \mathbf{a}S\mathbf{b} \mid \mathbf{b}S\mathbf{a} \mid \lambda.$$

- Find two different leftmost derivations of **abab.**

Example

Example (Ambiguous grammar)

- Consider the grammar

$$E \to E + E \mid E * E \mid (E) \mid \mathbf{a} \mid \mathbf{b} \mid \mathbf{c}.$$

- Derive the string $\mathbf{a} + \mathbf{b} * \mathbf{c}$ in two different ways.
- Is this grammar ambiguous?
- Is this language inherently ambiguous?

Example

Example (Unambiguous grammar)

The same language can be derived unambiguously from the following grammar.

$$E \to E + T \mid T$$
$$T \to T * F \mid F$$
$$F \to (E) \mid \mathbf{a} \mid \mathbf{b} \mid \mathbf{c}$$

Outline

Assignment

Assignment

- Section 5.1 Exercises 1, 2, 4, 9cde, 10, 12bf, 15, 23, 24, 25.