CSCI 340: Computational Models

Context-Free Languages

Chapter 17

Department of Computer Science

Closure Properties of Context-Free Languages

- Union, Product, Kleene closure, complement, and intersection of regular languages are *all regular*
- What operations of context-free languages are still context-free?

Closure Properties of Context-Free Languages

- Union, Product, Kleene closure, complement, and intersection of regular languages are *all regular*
- What operations of context-free languages are still context-free?
- Union $L_1 + L_2$
- Concatenation $L_1 L_2$
- Kleene closure ${L_1}^*$

Union

Theorem

If L_1 and L_2 are context-free languages, then their union, $L_1 + L_2$, is also a context-free language. In other words, the context-free languages are closed under union

Proof (by grammars).

- The CFG for L_1 has start state S – rename it S_1
- The CFG for L_2 has start state S – rename it S_2
- To avoid collisions with non-terminal states, append $_1$ if it belonged to the first CFG and $_2$ if it belonged to the second CFG
- Introduce a new start state, S and create the production:

$$S \longrightarrow S_1 \mid S_2$$

- All words with derivations starting with $S \longrightarrow S_1$ belong to L_1 and all words with derivations starting with $S \longrightarrow S_2$ belong to L_2

□

Union – Example

Example

Consider the two languages L_1 and L_2:

L_1 be PALINDROME with CFG:

$$S \longrightarrow aSa \mid bSb \mid a \mid b \mid \Lambda$$

L_2 be $\{ a^n b^n \}$ with CFG:

$$S \longrightarrow aSb \mid \Lambda$$

Example

1. L_1 = EVENPALINDROME
2. L_2 = ODDPALINDROME

Union – Alternative Proof

Proof (by Machines).

- PDA_1 has a START state
- PDA_2 has a START state
- "merge" these two START states together
- Any input string which reaches ACCEPT either went through a path along PDA_1 or PDA_2

□

Concatenation

Theorem

If L_1 and L_2 are context-free languages, then so is L_1L_2. In other words, the context-free languages are closed under product

Proof.

- Let CFG_1 and CFG_2 be context-free grammars that generate L_1 and L_2 respectively
- Relabel all nonterminals by appending $_1$ for every nonterminal in CFG_1 and appending $_2$ for every nonterminal in CFG_2
- Create a new production for S:

$$S \rightarrow S_1 S_2$$

- Any word generated by this CFG has a "front" part derived from S_1 and a "rear" part derived from S_2
- The two sets cannot cross over and interact with each other because the sets of non-terminals are disjoint

□

Concatenation – Example

Example

Let L_1 be PALINDROME and CFG_1 be

$$S \rightarrow aSa \mid bSb \mid a \mid b \mid \Lambda$$

Let L_2 be $\{\ a^n b^n\ \}$ and CFG_2 be

$$S \rightarrow aSb \mid \Lambda$$

Concatenation – Example

Example

Let L_1 be PALINDROME and CFG_1 be

$$S \to aSa \mid bSb \mid a \mid b \mid \Lambda$$

Let L_2 be $\{ a^n b^n \}$ and CFG_2 be

$$S \to aSb \mid \Lambda$$

But can we prove concatenation with machines?

- Should the TAPE be empty after processing L_1?
- Should the STACK be empty after processing L_1?

Kleene Closure

Theorem

If L is a context-free language, then L^* is one too. In other words, the context-free languages are closed under the Kleene star.

Proof (by construction).

- Let us start with a CFG for the language L – it has start symbol S
- Relabel S as S_1 (replacing all occurrences)
- Create a new production for non-terminal S:

$$S \longrightarrow S_1 S \mid \Lambda$$

- We are able to apply the S production exactly once (producing λ), twice (producing exactly what was accepted originally), or n times (producing the closure)

□

Example

PALINDROME = $S \longrightarrow aSa \mid bSb \mid a \mid b \mid \Lambda$

Intersection and Complement

Theorem (sort of)

The intersection of two context-free languages may or may not be context-free

Intersection and Complement

Theorem (sort of)

The intersection of two context-free languages may or may not be context-free

Consider two regular languages – then $\{a^n b^n a^m\}$ and $\{a^n b^m a^m\}$

Theorem (sort of)

The complement of two context-free languages may or may not be context-free

Intersection and Complement

Theorem (sort of)

The intersection of two context-free languages may or may not be context-free

Consider two regular languages – then $\{a^n b^n a^m\}$ and $\{a^n b^m a^m\}$

Theorem (sort of)

The complement of two context-free languages may or may not be context-free

Consider a regular language – then consider $(L_1{}' + L_2{}')' = L_1 \cap L_2$

Mixing Context-Free and Regular Languages

Claim

The union of a context-free language and a regular language must be context-free because the regular language itself is context-free.

Example

1. PALINDROME (nonregular context-free)
2. $(\mathbf{a} + \mathbf{b})^*$ (regular and contains PALINDROME)

Mixing Context-Free and Regular Languages

Claim

The union of a context-free language and a regular language must be context-free because the regular language itself is context-free.

Example

1. PALINDROME (nonregular context-free)
2. $(\mathbf{a} + \mathbf{b})^*$ (regular and contains PALINDROME)

The union is regular

Example

1. $\mathbf{a}^*$ (regular)
2. PALINDROME (nonregular context-free and contains $\mathbf{a}^*$)

Mixing Context-Free and Regular Languages

Claim

The union of a context-free language and a regular language must be context-free because the regular language itself is context-free.

Example

1. PALINDROME (nonregular context-free)
2. $(\mathbf{a} + \mathbf{b})^*$ (regular and contains PALINDROME)

The union is regular

Example

1. $\mathbf{a}^*$ (regular)
2. PALINDROME (nonregular context-free and contains $\mathbf{a}^*$)

The union is nonregular context-free

Intersection of CFLs and RLs

Theorem

The intersection of a context-free language and a regular language is always context-free

Homework 9b

5. (4pt) Using (1) the theorems on slides 2, 5, and 7; (2) a little ingenuity; and (3) the recursive definition of regular languages – provide a new proof that all regular languages are context-free
6. (2pt ea) Find CFGs for the following languages:
   - All words that start with a or are of the form $a^n b^n$
   - All words in EVEN-EVEN*
   - All words that start with ODD-PALINDROME and end with EVEN-PALINDROME
7. (4pt) Find a CFG for $a^x b^y a^z$ where $x + z = y$
8. (2pt ea) Which of the following are context-free?
   - EQUAL ∩ { $a^n b^n a^n$ }
   - EVEN-EVEN' ∩ PALINDROME
   - { $a^n b^n$ }' ∩ PALINDROME