CSE211-Formal Languages and Automata Theory

U2L13 – Pumping Lemma for CFL and Properties of CFL

Dr. P. Saravanan
School of Computing
SASTRA Deemed University

Agenda

- Recap of previous class
  - Normal Forms
- Pumping Lemma for CFL
  - Definition
  - Examples
- Closure Properties
- Decision Properties

Pumping Lemma for CFL

- **Theorem** (pumping lemma for CFL's)
  - Let L be a CFL. There exists an integer constant n such that if $z \in L$ with $|z| \geq n$, then we can write $z = uvwxy$, subject to the following conditions:
    1. $|vwx| \leq n$;
    2. $v, x \neq \varepsilon$ (that is, v, x are not both ε);
    3. for all $i \geq 0$, $uv^iwx^iy \in L$.
- Used to prove that the given language is not a context-free language

Example 1

- Prove by contradiction the language $L = \{0^m 1^m 2^m \mid m \geq 1\}$ is not a CFL by the pumping lemma.
- *Proof*.
  - Step 1: an integer constant $n=9$
  - Step 2: such that if $z \in L$ with $|z| \geq 9$, then we can write $z = 000111222$,
  - Step 3: Divide the string z into z= *uvwxy such that*

    $|vwx| \leq n$; and $v, x \neq \varepsilon$ (that is, v, x are not both ε);

$z =$	00	01	11	2	22
	U	V	W	X	Y

  - Step 4: for all $i \geq 0$, $uv^iwx^iy \in L$

For $i=0 \Rightarrow uv^0wx^0y \Rightarrow 001122 \in L$

For $i=1 \Rightarrow uv^1wx^1y$

$\Rightarrow 000111222 \in L$

For $i=2 \Rightarrow uv^2wx^2y$

$\Rightarrow 00\ 0101\ 11\ 22\ 22 \notin L$

So the given language is not a CFL

Example 2:

- Prove that $L=\{ww \mid w \in \{0, 1\}^*\}$ is not a CFL.

Closure Properties of CFL's

- **Some differences between CFL's and RL's**
  - CFL's are *not* closed under *intersection*, *difference*, or *complementation*
  - But the intersection or difference of a CFL and an RL is still a CFL.
  - We will introduce a new operation --- substitution.

Substitution

- **Definitions:**
  - A *substitution s* on an alphabet S is a function such that for each $a \in$ S, $s(a)$ is a language L_a over any alphabet (not necessarily S)
  - For a string $w = a_1a_2 \ldots a_n \in$ S*, $s(w) = s(a_1)s(a_2) \ldots s(a_n) = L_{a1}L_{a2} \ldots L_{an}$, i.e., $s(w)$ is a language which is the concatenation of all L_{ai}'s
  - Given a language L, $s(L) = \cup_{w \in L}\ s(w)$

Substitution Example

- A *substitution s* on an alphabet S = {0, 1} is defined as $S(0) = \{a^n b^n \mid n \geq 1\}$, $s(1) = \{aa, bb\}$.
- Let w = 01, then $s(w) = s(0)s(1) = \{a^n b^n \mid n \geq 1\}\{aa, bb\}$ $= \{a^n b^n aa \mid n \geq 1\} \cup \{a^n b^{n+2} \mid n \geq 1\}$.

Closure properties of CFL

The CFL's are closed under the following operations:

- 1. Union
- 2. Concatenation
- 3. Closure (*), and positive closure (+)
- 4. Homomorphism
- 5. Inverse Homomorphism
- 6. Reversal

Not Closed

- 1. Intersection
- 2. Difference and 3. Complementation

Intersection with an RL

- **Theorem 7.27**
  - If L is a CFL and R is an RL, then $L \cap R$ is a CFL.
- The following are true about CFL's L, L_1, and L_2, and an RL R:
- 1. $L - R$ is a CFL;
- 2. $\overline{L}$ is *not* necessarily a CFL;
- 3. $L_1 - L_2$ is *not* necessarily a CFL.

Decision Properties of CFL’s

- **Facts:**
  - Unlike RLs’ decision problems which are all solvable, *very little* can be said about CFL’s.
  - Only two problems *can* be *decided* for CFL’s:
    - whether the language is empty;
    - whether a given string is in the language.
  - Computational complexity for conversions between CFG’s and PDA’s will be investigated.

Decision Properties of CFL's

- **Testing Emptiness of CFL's**
- The problem of testing emptiness of a CFL *L* is *decidable*.
- **Testing Membership in a CFL**
- A way for solving the membership problem for a CFL *L* is to use the CNF of the CFG *G* for *L* in the following way:
  - The parse tree of an input string *w* of length *n* using the CNF grammar *G* has $2n - 1$ nodes.
  - We can generate all possible parse trees and check if a yield of them is *w*.
- The number of such trees is *exponential* in *n*.

Preview of Un-decidable CFL Problems

- **The following are undecidable CFL problems ---**
  - Is a given CFL inherently ambiguous?
  - Is the intersection of two CFL's empty?
  - Are two CFL's the same?
  - Is a given CFL equal to S*, where S is the alphabet of this language?

Summary

- Recap of previous class
  - Normal Forms
- Pumping Lemma for CFL
  - Definition
  - Examples
- Closure Properties
- Decision Properties

References

- John E. Hopcroft, Rajeev Motwani and Jeffrey D. Ullman, *Introduction to Automata Theory*, Languages, and Computation, Pearson, 3rd Edition, 2011.
- Peter Linz, An Introduction to Formal Languages and Automata, Jones and Bartle Learning International, United Kingdom, 6th Edition, 2016.

Next Class: **Unit III**

Context-Sensitive Language

Thank you.