Theory of Computation

Lecture 9: Pumping Lemma for CFLs

Outline

- Pumping Lemma for CFL
- Proof Idea
- Examples

From Sipser Chapter 2.3

Non Context Free Languages

- Just like there are non regular languages, there are languages that are not context free
  - They cannot be generated by a CFG
  - They cannot be generated by a PDA
- How do we show that a language is not context-free?
  - We use a pumping lemma for CFLs!

Pumping Lemma for CFGs

If A is a context-free language then there is a pumping length p such that, if s is any string in A with length ≥ p, then s may be divided into five pieces x = uvxyz satisfying 3 conditions:

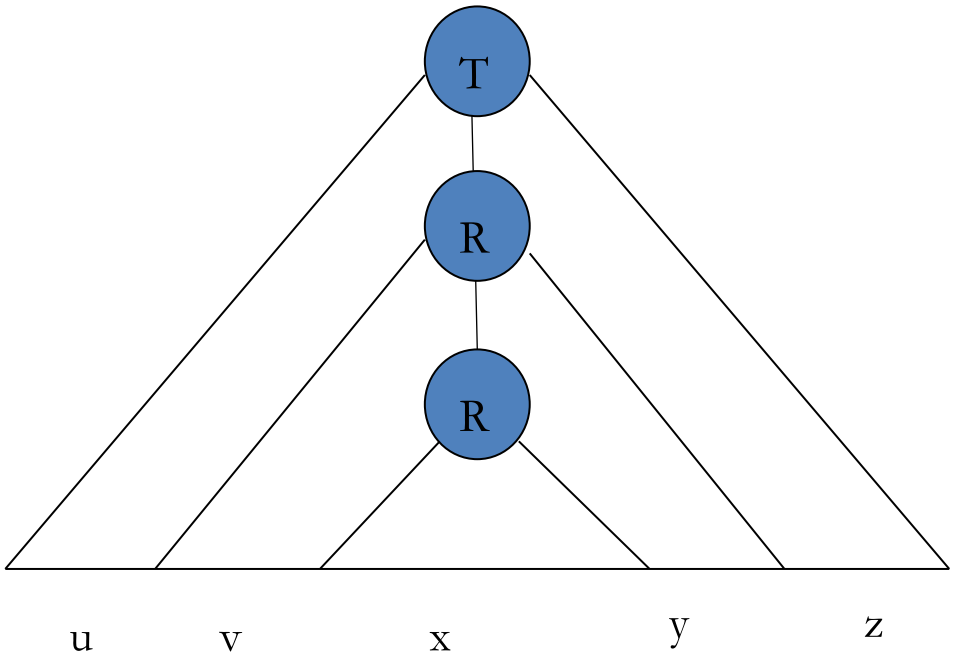
1. For each $i \geq 0$, $uv^ixy^iz \in A$
2. $|vy| > 0$, and
3. $|vxy| \leq p$

Proof Idea

For regular languages we applied the pigeonhole principle to the number of states to show that a state had to be repeated.

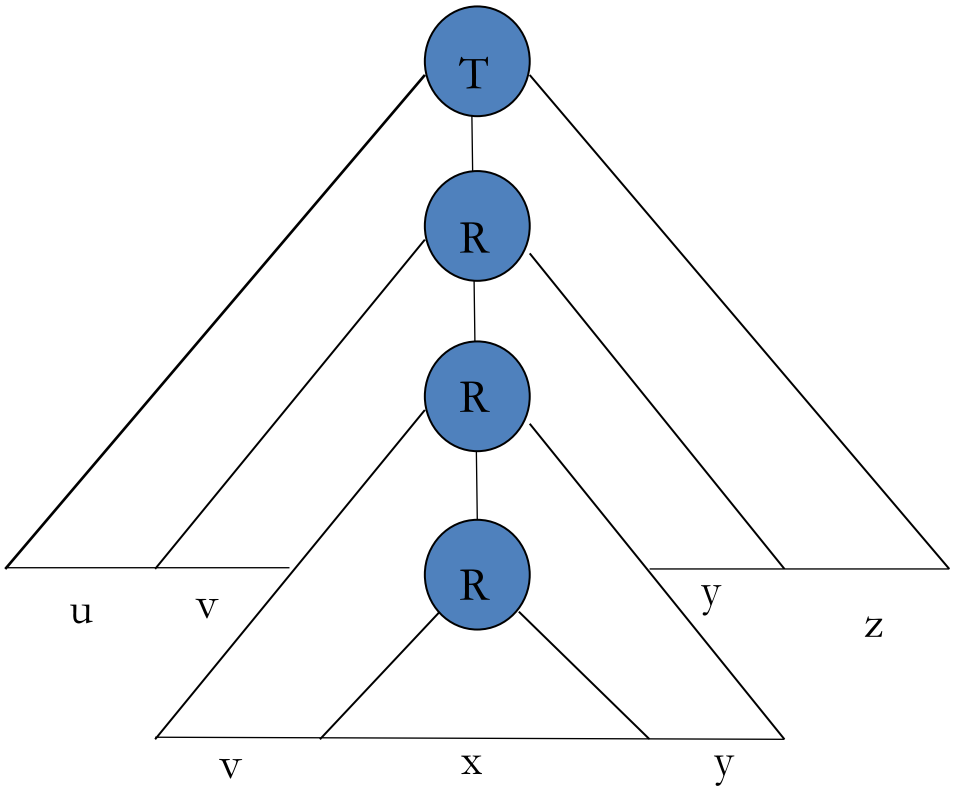
- Here we apply the same principle to the number of variables in the CFG to show that some variable will need to be repeated given a sufficiently long string
  - We will call this variable R and assume it can derive X

Proof Idea Sketch



Consider the derivation: T => uRz =>uvRyz =>uvRyz,
where R => x and R => vRy

Proof Idea Sketch



We can replace R => x with R => vRy and "pump" vRy to derive uvixyiz which therefore must also belong to the language

Example 2.36

Use the pumping lemma to show that the language B = $\{a^n b^n c^n \mid n \geq 0\}$ is not a CFL

- What string should we pick?
  - Select the string $s = a^p b^p c^p$. $s \in B$ and $|s| > p$
- How should we break up s?
  - We have to consider all possible ways of breaking down s and show that in no case the pumping lemma can be applied
  - Condition 2 says that v and y cannot both be empty
  - By condition 3 $|vxy| \leq p$ so u and z cannot both be empty
  - v and y each only contain one type of symbol (one symbol is left out)
    - uv^2xy^2z cannot contain equal number of a's, b's and c's
  - Either v or y contains more than one type of symbol
    - Pumping will violate the separation of a's, b's and c's

Example 2.38

Use the pumping lemma to show that the language $D = \{ww \mid w \in \{0,1\}^*\}$ is not a CFL

- What string should we pick?
  - Consider $s = \{0^p10^p1\}$
    - But this can be pumped– try generating uvxyz so it can be pumped
    - Hint: we need to straddle the middle for this to work
    - Solution: $u=0^{p-1}$, $v=0$, $x=1$, $y=0$, $z=0^{p-1}1$
    - Check it. Does $uv^2xy^2z \in D$? Does $uwz \in D$?

Example 2.38

Use the pumping lemma to show that the language $D = \{ww \mid w \in \{0,1\}^*\}$ is not a CFL

- What string should we pick?
  - Choose $s = \{0^p1^p0^p1^p\}$. How can we partition it?
    - First note that vxy must straddle midpoint. Otherwise pumping makes 1st half ≠ 2nd half
      - » since $|vxy| < p$, it is all 1's in left case and all 0's in right case
    - If does straddle midpoint, if pump down, then not of form ww since neither 1's or 0's match in each half

Conclusion on Pumping Lemma

- Important similarities in the version for Regular Languages and Contex Free Languages
- Used to show that a language is not RL or CFL
- Remember common structure of the proof using pidgeonhole principle
- When applying it pay attention to the quantifiers:
  - Is sufficient to show that one string from the language cannot be pumped
  - We need to show that it cannot be pumped for all possible ways of breaking it down
- Remember correct direction of the implication:
  - If a language violates the pumping lemma then it is guaranteed to not be Regular/Contex Free
  - Viceversa is not true!