<Lecture 20>

2.3 Non-CFL

Thm 2.34. [Pumping lemma for CFL]. For any CFL A, there is a non-negative integer p ("pumping length") s.t. every string s where $|s| \geq p$ can be divided into $uvxyz$ satisfying

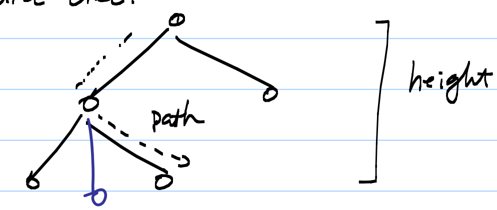
(1). for each $i \geq 0$, $uv^i x y^i z \in A$

(2) $|vy| > 0$,

(3) $|vxy| \leq p$

Preparation for proving the lemma

parse tree:



branching factor : largest number of children of any node, i.e. the maximum length of string in rhs of any rule

Lemma 1: The maximum length of a string that can be generated from a parse tree with height h and branching factor b is b^h

Proof of the pumping lemma.

- If L is context-free, there is a CFG G. Let b be the branching factor. Let $p = b^{|V|+1}$

- Let s be any string s.t. $|s| \geq p$. Since $b^{|V|+1} > b^{|V|}$ Any parse tree for s has height $\geq |V| + 1$

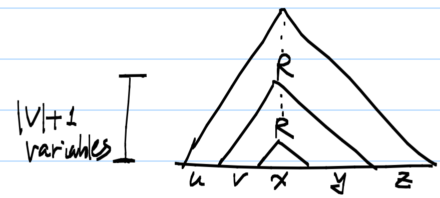
- Consider a parse tree T with the smallest number of nodes. T contains a path that contains $|V|+1$ variables and 1 terminal.

- Since there are $|V|$ variables, by pigeon hole principle there must be a variable repeated in the path.

- Let R be such a variable that repeats among the lowest $|V|+1$ variables on the path.

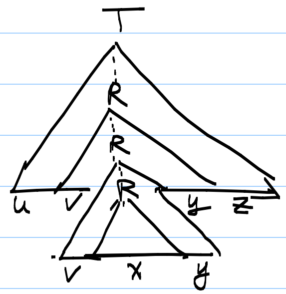
1) Why $uv^ixy^iz \in A$?

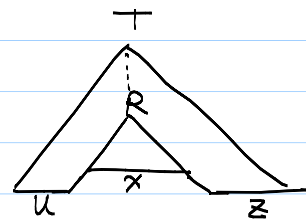
$$T \overset{*}{\Rightarrow} uRz \overset{*}{\Rightarrow} uvRyz$$



or $R \overset{*}{\Rightarrow} vRy$

$R \overset{*}{\Rightarrow} x$





$$T \overset{*}{\Rightarrow} uRz \overset{*}{\Rightarrow} uxz$$

$$T \overset{*}{\Rightarrow} uRz \overset{*}{\Rightarrow} uvRyz \overset{*}{\Rightarrow} uvvRyyz$$
$$= uv^2Ry^2z$$

- Why $|vy| > 0$?

- Why $|vxy| \le p$?

Example $B = \{0^n1^n2^n \mid n \ge 0\}$ is not CF

- Assume B is CF. There is a pumping length p. Choose $s = a^pb^pc^p$. Consider any division of s into uvxyz satisfying Cond 2,3.
- By condition 2, v or y is nonempty.
  By condition 3, ($|vxy| \le p$), vxy contains at most 2 different symbols.
  Therefore uv^2xy^2z contains at least one symbol that is repeated more than p times,
  and at least one symbol that is repeated p times
  Hence it doesn't belong to B, contradicting Condition 1