Feb 26

Lem 2.27. If a pushdown automaton recognizes some language, then it is CF.

<u>Proof</u> (hand-waving)

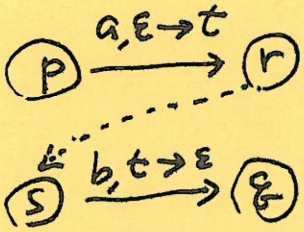
$$\text{PDA } P = (Q, \Sigma, \Gamma, \delta, q_0, \{q_{accept}\})$$

$$\Downarrow$$

$$\text{CFG } G$$

variables of G: A_{pq} for all states p, q.

Start variable: $A_{q_0, q_{accept}}$



- $A_{pq} \to a A_{rs} b$
- for $p, q, r \in Q$, add the rule
  $A_{pq} \to A_{pr} A_{rq}$
- for $p \in Q$, add $A_{p,p} \to \varepsilon$. $\square$

2.3 Non-Context-Free languages.

$$L = \{ a^n b^n c^n \mid n \geq 0 \}.$$

This is not CFL.

pumping lemma:

If A is context-free, then there is a number p (pumping length), s.t. if $s \in A$ with $|s| \geq p$, then s can be decomposed into $s = uvxyz$ satisfying

① $uv^ixy^iz \in A$, for $i \geq 0$

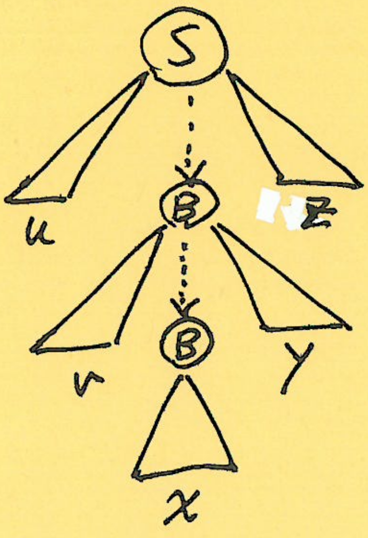
② $|vy| > 0$

③ $|vxy| \leq p$.

---

proof: We use a simpler proof compared with the textbook. We assume a grammar G generates $A - \{\varepsilon\}$ without using <u>ε-rules</u> or <u>unit rules</u>.

- Since the # of items on the right side of any production is bounded, say by K, the length of the derivation of any $w \in A$ must be at least $|w|/K$. As A is infinite, there exist arbitrarily long derivations.
- On the other hand, the # of variables in G is finite, there must be some variable that repeats on a path.



$$S \xRightarrow{*} uBz \xRightarrow{*} uvByz \xRightarrow{*} uvxyz = w$$

As $B \xRightarrow{*} vBy$ and $B \xRightarrow{*} x$, all strings uv^ixy^iz can be generated by the grammar, hence are in A.

- Furthermore, in $B \overset{*}{\Rightarrow} vBy$ and $B \overset{*}{\Rightarrow} x$ we can assume that no variable repeats (otherwise, use the repeated one as B). So, $|vxy| \leq |w| = p$.
- Finally, as there is no unit productions and ε-productions, v and y can't be both empty, or, $|vy| > 0$. $\square$

What is p? p is the shortest length of w, which makes the pumping lemma true.

Applications of pumping lemma.

Ex 1. $B = \{a^n b^n c^n \mid n \geq 0\}$ is not context-free.

proof: Assume that B is CF.

Pick $s = a^p b^p c^p$, p being the pumping length.

By the pumping lemma, s can be decomposed into $s = uvxyz$, s.t.
① $uv^ixy^iz \in B$, for $i \geq 0$,
② $|vy| > 0$,
③ $|vxy| \leq p$.

a) If v and y contain one type of symbols, uv^2xy^2z would contain more letters in this type, hence $\notin B$.

b) If v or y contain more than one type of symbols, uv^2xy^2z would be out of order (i.e. with subsequences a…b…a…b or b…c…b…c), hence $\notin B$

This is a contradiction to the pumping lemma.

$\therefore$ B is not CF. $\square$