CSE 396 Lecture Thu Mar 29 ~~Sept.~~ Apr. 2018 (1)

Top Hat 8137

Defn: A language L is a **Context-Free Language (CFL)** if there is a context-free grammar (CFG) G s.t. $L = L(G)$.

Context-Free Pumping Lemma: If L is a CFL then:

- there exists a number $N > 0$ such that
  - (for all $x \in L$ with $|x| \geq N$)
    - there exists a breakdown $x =: y u v w z$ such that $|uvw| \leq N$, u and w are not both ε, and: i.e. $uw \neq \varepsilon$ i.e. $|uw| > 0$ i.e. $u \neq \varepsilon$ or $w \neq \varepsilon$
  - for all $i \geq 0$ $x^{(i)} =_{def} y u^i v w^i z$ belongs to L.

L CFL ⇒ Blah
(¬Blah) ⇒ L is not a CFL

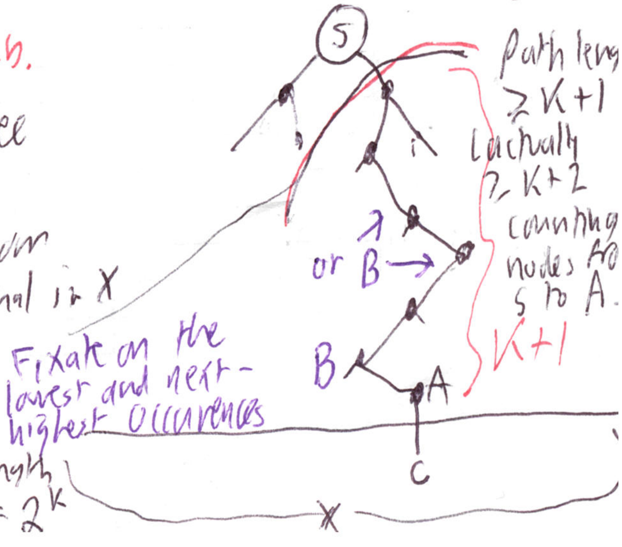
text $w =: u v x y z$ (w)

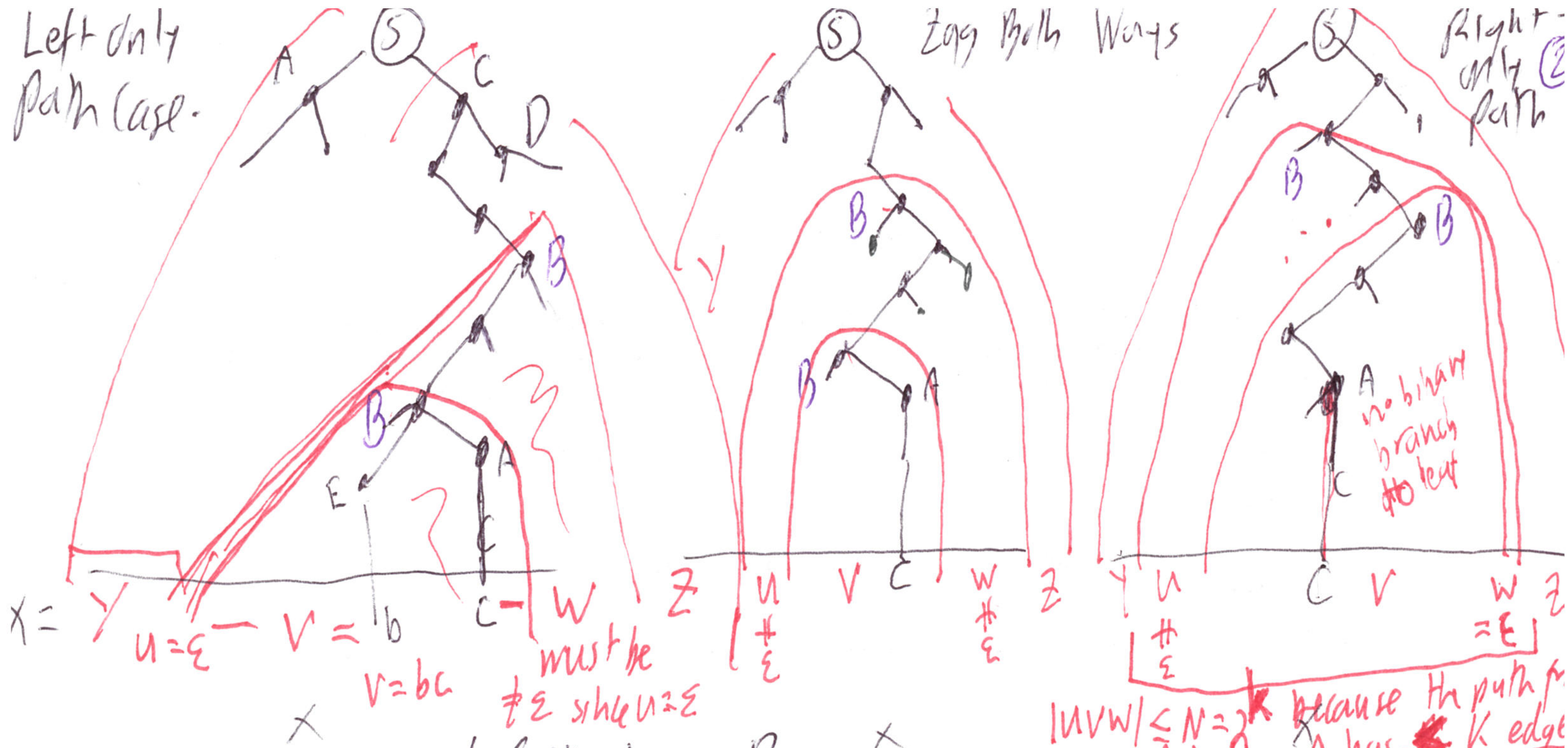
Note: $x^{(0)} = y v z$

Proof: Since L is a CFL, there is a grammar G in ChNF (gives binary trees, nicer to think a[bout]) such th[at] $L(G) = L \setminus \{\varepsilon\}$. We have $G = (V, \Sigma, R, S)$. Take $k = |V|$, $N = 2^k$.

Let any $x \in L = L(G)$, $|x| \geq N$ be given.

By $x \in L(G)$, we can take a parse tree T for x in the CFG G. By binary counting, T has at least $k+1$ edges in some path from the root S to a "leaf variable" that derived a terminal in x. Consider the bottom $k+1$ nodes of that path. By the Pigeonhole Principle, some $B \in V$ occurs at least twice in those bottom $k+1$ nodes.





Take v to be the yield of the lower B.
Take u to be everything to the left of v that is derived by the upper B.
Take w to be anything to the right of v that is derived by the upper B.
Take y to be the rest of x to the left of u, z the rest of x to right of w.

Let any $i \geq 0$ be given. Then $x^{(i)} = yu^ivw^iz \in L(G)$ because we can splice or extend the parse tree between the two B's. This results in a legal parse tree for $x^{(i)}$, so $x^{(i)} \in L(G) = L$. ☒

Contrapositive: Suppose L is any language such that

- for all $N > 0$
  - there exists an $x \in L$ with $|x| > N$ such that
    - for all breakdowns $x =: yuvwz$ s.t. $uw \neq \varepsilon$ and $|uvw| \leq N$
      - there exists $i \geq 0$ s.t. $x^{(i)} = yu^ivw^iz$ is not in L.

(∃∀∃∀... "blah")

Then L is not a CFL. ☒

Examples of non-CFLs:

$L = \{a^Nb^Nc^N : N \geq 1\}$

$L = \{a^nb^mc^nd^m : m, n \geq 1\}$

$L = \{ww : w \in \{0,1\}^*\}$

END OF LECTURE