The Pumping Lemma for CFLs

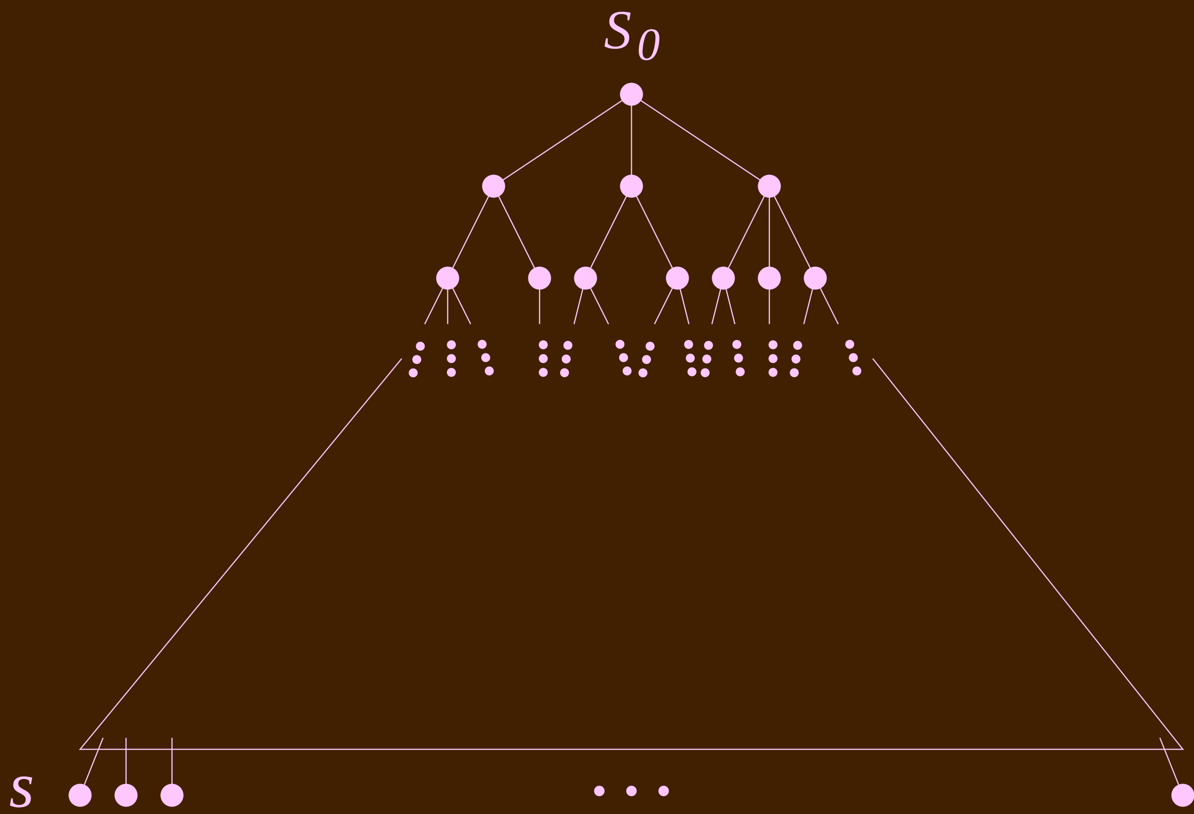
Mark Greenstreet, CpSc 421, Term 1, 2006/07

- Non-Context-Free Languages
- The Pumping Lemma for CFLs
- Examples

$A = \mathbf{a}^n\mathbf{b}^n\mathbf{c}^n$

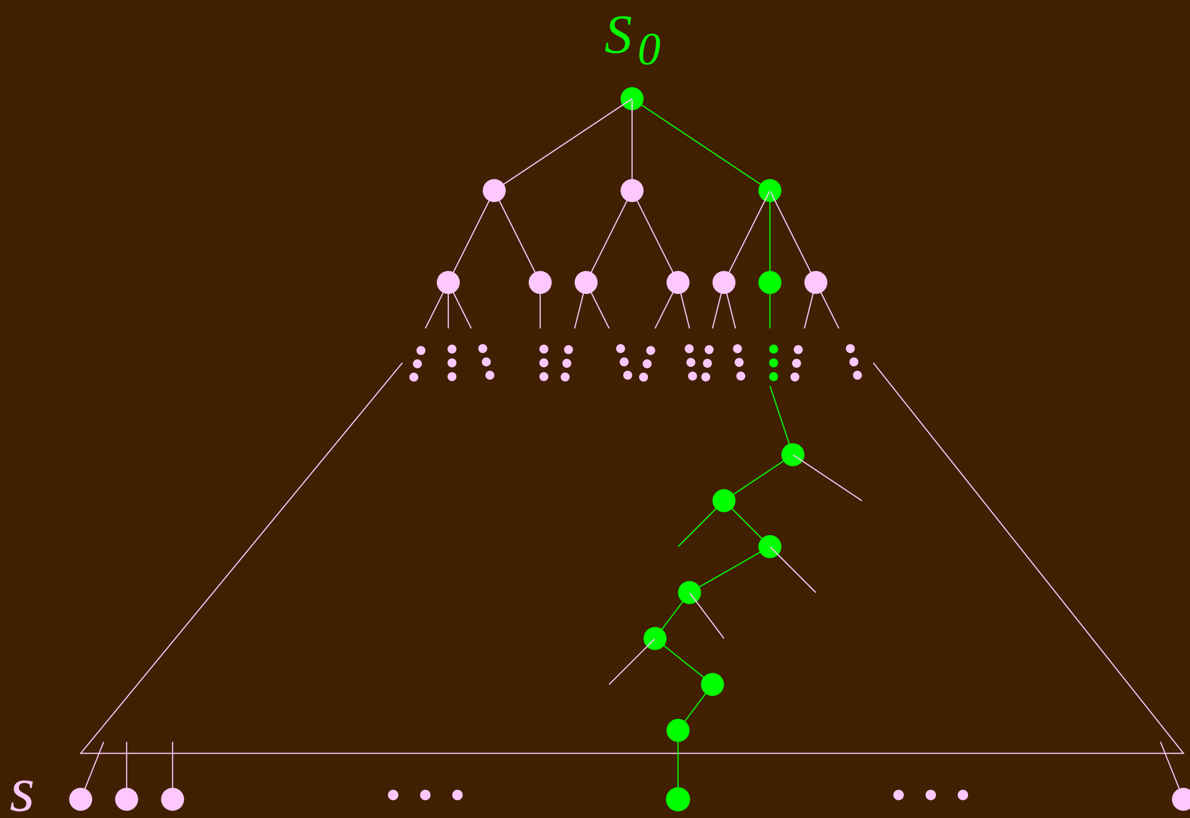
- Is A regular?
- Is A context free?
  - Can we construct a CFG that generates A?
  - Can we construct a PDA that recognizes A?

Pumping Lemma: The Main Idea



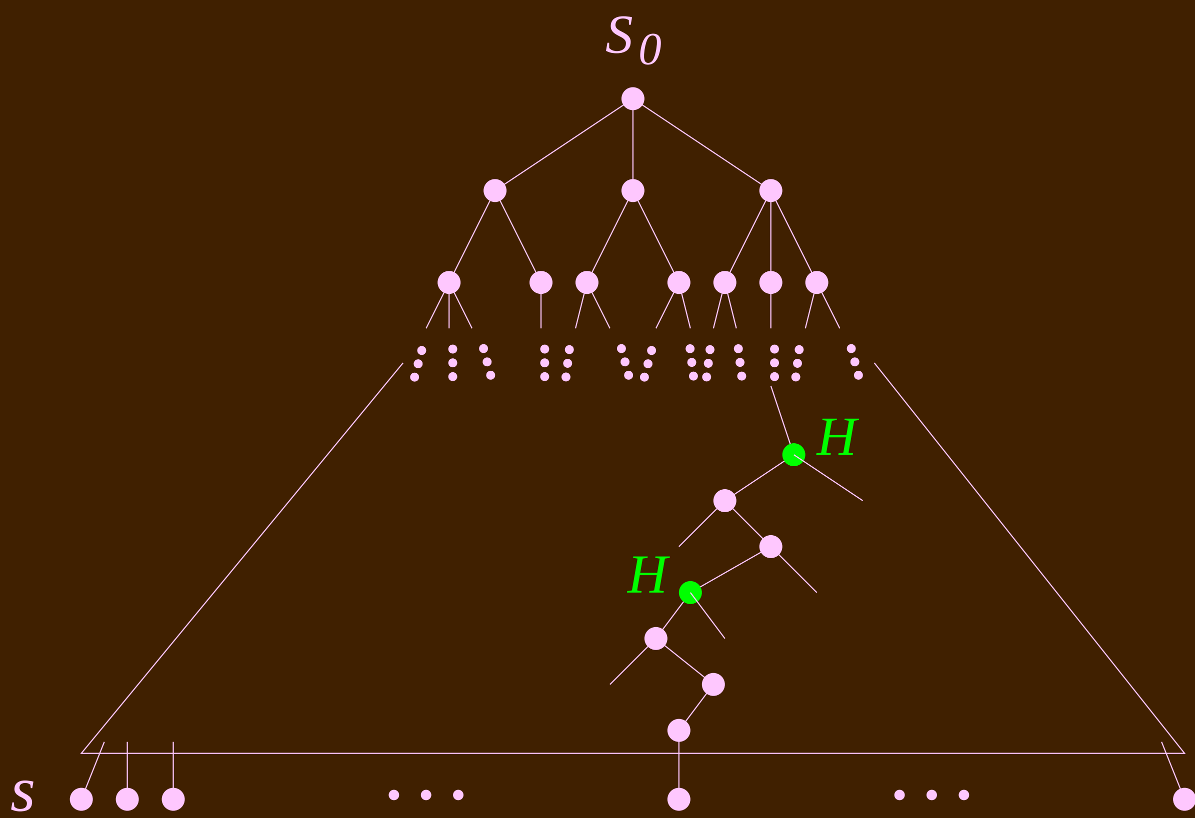
- Let G be a CFG. Let w be a very long string in $L(G)$. Consider the derivation tree for w.

Pumping Lemma: The Main Idea



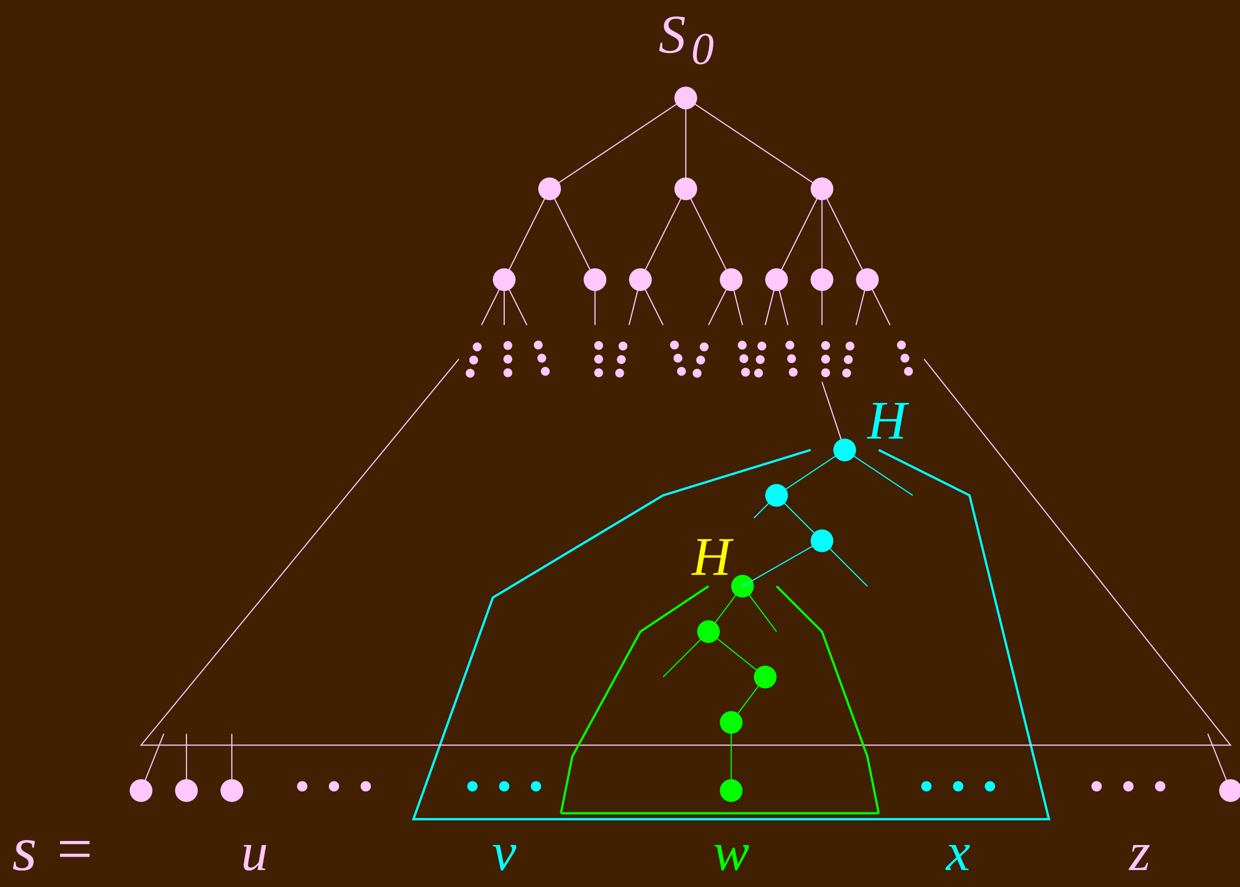
- Let G be a CFG. Let w be a very long string in $L(G)$. Consider the derivation tree for w.
- Choose some symbol in w that has a long path from S_0.

Pumping Lemma: The Main Idea



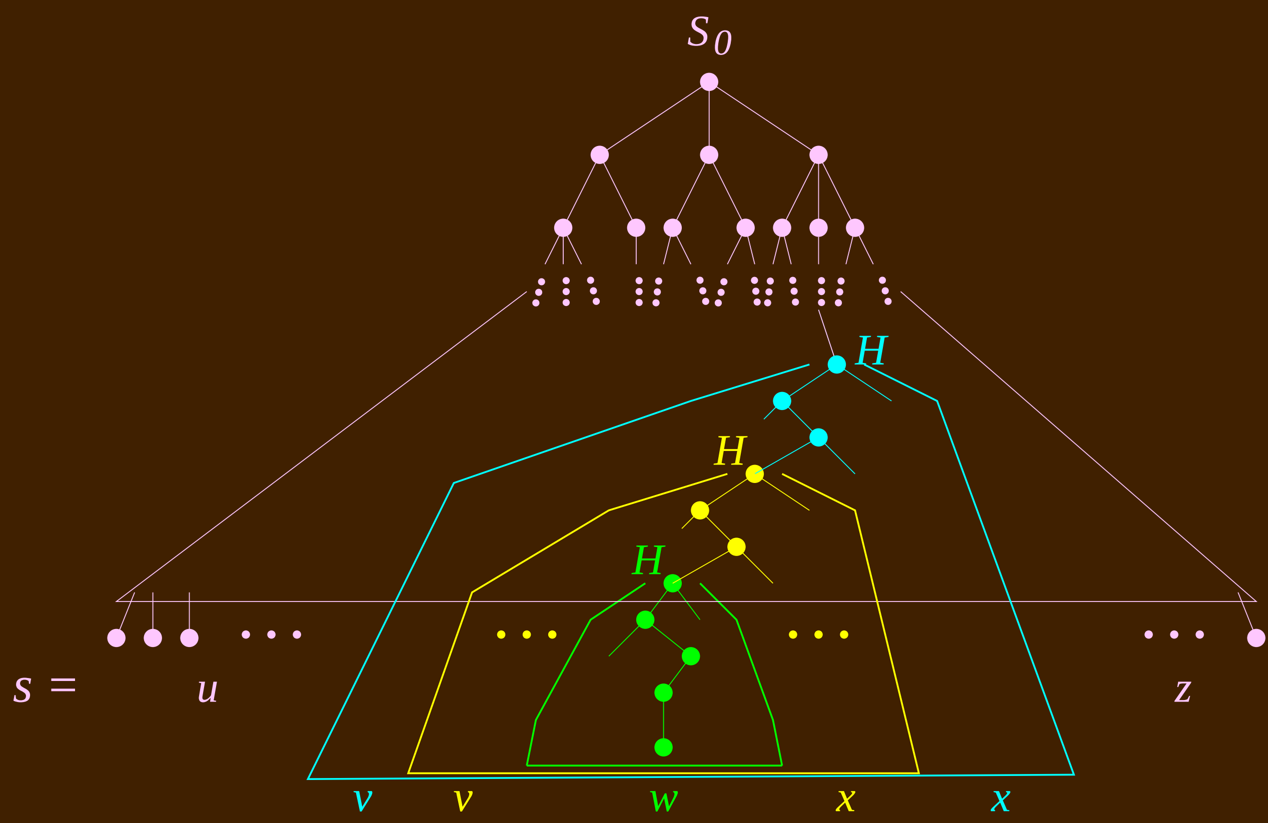
- Choose some symbol in w that has a long path from S_0.
- Some variable, H, must occur two or more times along this path (pigeon hole).

Pumping Lemma: The Main Idea



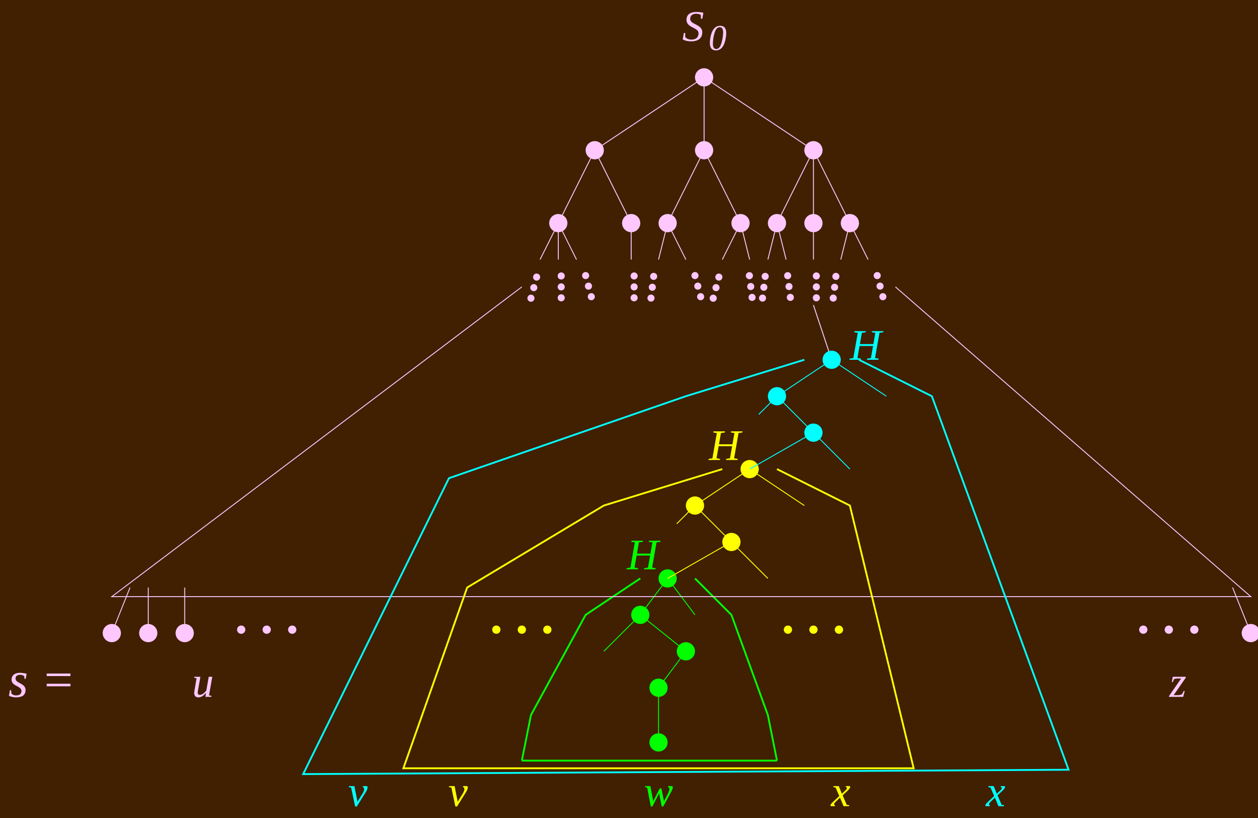
- Some variable, H, must occur two or more times along this path (pigeon hole).
- Consider the derivation trees rooted at each of these two instances of H.

Pumping Lemma: The Main Idea



- Consider the derivation trees rooted at each of these two instances of H.
- We can replace the tree rooted at the second with the tree rooted at the first.

Pumping Lemma: The Main Idea



- We can replace the tree rooted at the second with the tree rooted at the first.
- $\therefore\ \forall i > 0.\ uv^i wx^i z \in L(A)$.

Pumping Lemma: Formal Statement

- Let A be a CFL.
- There exists a constant, p such that for any $s \in A$ with $|s| \geq p$, there exists strings u, v, w, x, and z such that:
  - $s = uvxyz$;
  - $|vy| > 0$;
  - $|vxy| \leq p$; and
  - for all $i \geq 0$, $uv^ixy^iz \in A$.

Pumping Lemma: Formal Proof

- Let A be a CFL.
- Let $G = (V, \Sigma, R, S_0)$ be a CNF CFG for A.
- Let $p = 2^{|V|+1}$.
- Let $s \in A$ with $|s \geq p$.
- Let T be a derivation tree for s.
- CNF derivation trees are binary trees except that they have a layer of degree-1 nodes just before the leaves. Thus,

$$\begin{aligned} height(T) &\geq \lceil \log_2(|s|) \rceil + 1 \\ &\geq \lceil \log_2(p) \rceil + 1 \\ &\geq (|V| + 1) + 1 \\ &= |V| + 2 \end{aligned}$$

- There is a path from S_0 to a leaf that goes through at least $|V| + 1$ variables. Thus, it visits some variable twice.
- Now, apply the idea suggested on the "Key Idea" slide.

$\mathbf{a}^n\mathbf{b}^n\mathbf{c}^n$ Is Not Context Free

$\{w \mid \exists x.\ w = xx\}$ Is Not Context Free