Properties of Context-Free Languages

A Pumping Lemma for CFLs and CFL Closure Properties

Monday, November 12, 2007
Reading: Stoughton 4.7, 4.10; Sipser 2.3

CS235 Languages and Automata

Department of Computer Science
Wellesley College

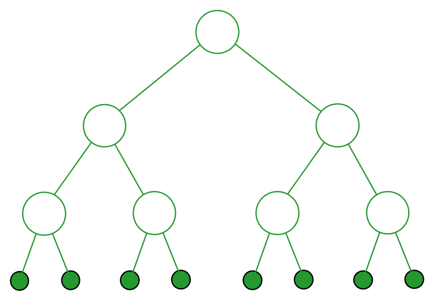
Goals

- Develop a pumping lemma for CFLs.
- Use the pumping lemma for CFLs to show that certain languages are not CFLs.
- Review closure properties for regular languages and discuss closure properties for context-free languages.

Height and # Leaves of Binary Trees

What is the maximum number of leaves in a binary tree of height h?

height	# leaves
0	
1	
2	
3	
h	



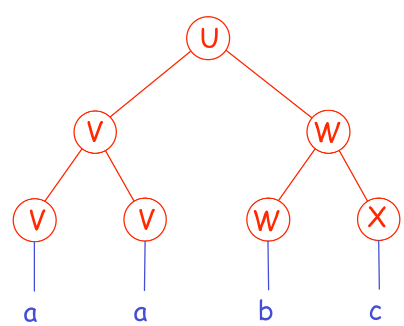
CFL Properties 29-3

Chomsky Normal Form Parse Trees

What is the maximum length of a string yielded by a Chomsky Normal Form parse tree of height h?

U → VW　　V → a
V → VV　　W → b
W → WX　　X → c

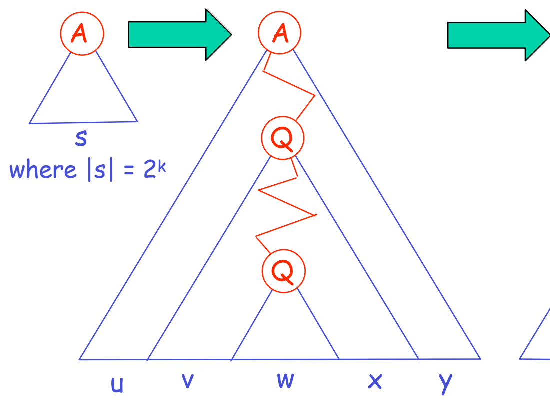
height	length
1	
2	
3	
h	

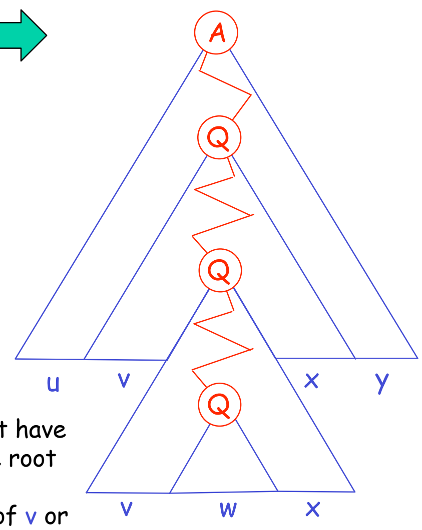


CFL Properties 29-4

Idea Behind Pumping Lemma For CFLs

Assume a Chomsky Normal Form grammar with k variables and start variable A





Any sufficiently long string *s* in a CFL *L* must have a CNF parse tree containing a path from the root A repeating some variable Q and so can be decomposed into *uvwxy*, where at least one of v or *x* is nonempty and $uv^iwx^iy \in L$ for any $i \in Nat$.

The Pumping Lemma for CFLs

The Pumping Lemma for Context-Free Languages

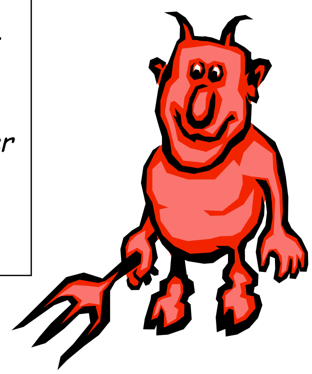
If *L* is a context-free language, there is a number *p* (the pumping length) such that any string *s* with length ≥ *p* can be expressed as *uvwxy*, where:

1. $|vx| > 0$ (at least one of v or x is nonempty)
2. $|vwx| \leq p$
3. $uv^iwx^iy \in L$ for each $i \in Nat$.

Games vs. Demons, Revisited

Suppose you want to prove that a language L is not context-free.
The proof (by contradiction) can be viewed as a game with a demon:

1. *You:* give the demon the language *L*.
2. *Demon:* gives you the pumping length *p*.
3. *You:* give the demon a string *s* with $|s| \geq p$.
4. *Demon:* divides *s* into *uvwxy* such that $|vx| > 0$ and $|vwx| \leq p$.
   (Warning: typically more cases to consider than in nonregular language proofs.)
5. *You:* win if you can give the demon an *i* such that $uv^iwx^iy \notin L$.



$0^n1^n2^n$ is Not a CFL

Viewed as game vs. a demon:

1. *You:* give the demon the language $L = \{0^n1^n2^n \mid n \text{ in Nat}\}$
2. *Demon:* gives you the pumping length *p*
3. *You:* give the demon what string *s* with $|s| \geq p$?
4. *Demon:* divides *s* into *uvwxy* such that $|vx| > 0$ and $|vwx| \leq p$.
   What are the possible divisions for your *s*?
5. *You:* win if you give the demon an *i* such that $uv^iwx^iy \notin L$.

$\{0^i 1^j 2^k \mid i \leq j \leq k\}$ is Not a CFL

Viewed as game vs. a demon:

1. *You:* give the demon the language $L = \{0^i 1^j 2^k \mid i \leq j \leq k\}$
2. *Demon:* gives you the pumping length p
3. *You:* give the demon what string s with $|s| \geq p$?
4. *Demon:* divides s into $uvwxy$ such that $|vx| > 0$ and $|vwx| \leq p$. What are the possible divisions for your s?
5. *You:* win if you give the demon an i such that $uv^iwx^iy \notin L$.

$\{ww \mid w \in \{0,1\}^*\}$ is Not a CFL

Viewed as game vs. a demon:

1. *You:* give the demon the language $L = \{ww \mid w \text{ in } \{0,1\}^*\}$
2. *Demon:* gives you the pumping length p
3. *You:* give the demon what string s with $|s| \geq p$?
   (Hint: 0^p10^p1 doesn't work!)
4. *Demon:* divides s into $uvwxy$ such that $|vx| > 0$ and $|vwx| \leq p$. What are the possible divisions for your s?
5. *You:* win if you give the demon an i such that $uv^iwx^iy \notin L$.

What is a Closure Property?

A set S is **closed** under an n-ary operation f iff $x_1, \ldots, x_n \in S$ implies $f(x_1, \ldots, x_n) \in S$

Examples:

- Bool is closed under negation, conjunction, disjunction.
- Nat is closed under + and * but not – and /.
- Int is closed under +, *, and -, but not /.
- Rat is closed under +, *, -, and / (except division by 0).

Closure Properties for Regular Languages

Suppose L, L_1, and L_2 are regular languages.
Then all the following are regular languages:

- L_1L_2
- L^*
- $\overline{L}$
- $L_1 \cup L_2$
- $L_1 \cap L_2$
- $L_1 - L_2$
- L^R

Closure Properties for CFLs

Suppose L, L_1, and L_2 are CFLs.
Then all the following are CFLs:

- L_1L_2
- L^*
- $L_1 \cup L_2$
- L^R

What about Complement & Intersection?

CFLs are not closed under intersection.

Counterexample:
Both $0^n1^n2^m$ and $0^m1^n2^n$ are CFLs,,
but $0^n1^n2^m \cap 0^n1^n2^m = 0^n1^n2^n$ is not.

CFLs are not closed under complement.

Proof by Contradiction (use deMorgan!)

Counterexample:
Let $L_{ww} = \{ww \mid w \text{ in } \{0,1\}^*\}$.

Surprisingly, $\overline{L_{ww}}$ is a CFL!