The Pumping Lemma

A Technique for Proving that Languages are Nonregular

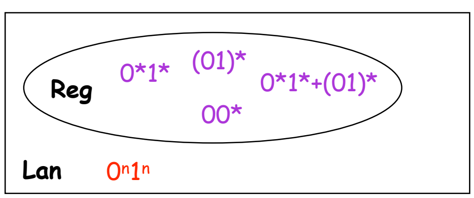
Thursday, October 24, 2007
Reading: Stoughton 3.12, Sipser 1.4

CS235 Languages and Automata

Department of Computer Science
Wellesley College

Nonregular Languages: Overview

1. Show the language $\{0^n1^n \mid n \text{ in Nat}\}$ is not regular.



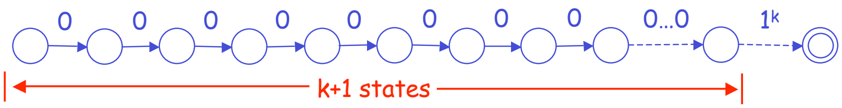
2. Generalize the technique for #1 by developing the **pumping lemma**.

3. Give examples of using the pumping lemma (sometimes in conjunction with closure properties of regular languages) to prove-by-contradiction that certain languages aren't regular.

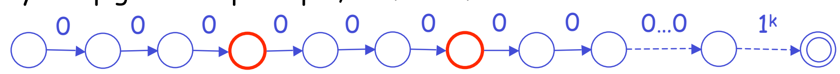
0^n1^n is Not a Regular Language

Proof by Contradiction: Suppose 0^n1^n ***is*** a regular language.
Then it is accepted by a DFA. Suppose the DFA has k states.

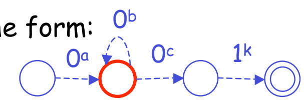
Now consider the labeled path for accepting the string 0^k1^k:



By the pigeonhole principle, 2 of the first k+1 states must be the same:



So the path has the form:



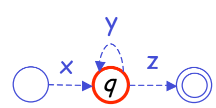
where a + b + c = k and b > 0

This means the DFA also accepts strings $0^a0^{ib}0^c1^k$ for any $i \in Nat$.
But for $i \neq 1$, these strings do not have the form 0^n1^n for some *n*.
This contradicts the assumption that there is a DFA for 0^n1^n. ✗

Generalizing the Technique: The Pumping Lemma

Any infinite regular language *L* has a weakly-simplified regular expression with a * ⇒ it is accepted by a DFA with a loop.

Any sufficiently long string *s* in *L* must traverse the loop, and so can be decomposed into *xyz*, where *y* is nonempty and $xy^iz \in L$ for any $i \in Nat$. We say that the substring *y* of *s* can be **pumped**.



The Pumping Lemma
If *L* is a regular language, there is a number *p* (the pumping length) such that any string *s* with length ≥ *p* can be expressed as *xyz*, where:

1. $|y| > 0$
2. $|xy| \leq p$
3. $xy^iz \in L$ for each $i \in Nat$.

Proof sketch: Let *p* be the number of states in a DFA for *L* and *q* be the first repeated state in the path for s (which must exist by the pigeonhole principle). Use *q* to divide *s* into *xyz*.

Using the Pumping Lemma to Prove L Nonregular

Towards a contradiction, assume L is regular.

By the pumping lemma, there is a p such that all strings $s \in L$ with length $\geq p$ can be pumped.

Find ***some*** string $s \in L$ with length $\geq p$ for which pumping is problematic. I.e., ***every*** decomposition of s into xyz with $|y| > 0$ and $|xy| \leq p$ leads to a string $xy^iz \notin L$ for ***some*** $i \in Nat$.

Therefore, the assumption that L is regular is false. X

This can be viewed as game vs. a demon:

1. *You:* give the demon the language L
2. *Demon:* gives you p
3. *You:* give the demon string s with $|s| \geq p$.
4. *Demon:* divides s into xyz such that $|y| > 0$ and $|xy| \leq p$
5. *You:* give the demon an i such that $xy^iz \notin L$.

0^n1^n revisited

Viewed as game vs. a demon:

1. *You:* give the demon the language 0^n1^n
2. *Demon:* gives you p
3. *You:* give the demon what string s with $|s| \geq p$?
4. *Demon:* divides s into xyz such that $|y| > 0$ and $|xy| \leq p$
5. *You:* give the demon an i such that $xy^iz \notin L$.

L_2 = {w | w has equal # of 0s and 1s}

Viewed as game vs. a demon:

1. *You:* give the demon the language L_2
2. *Demon:* gives you p
3. *You:* give the demon what string s with $|s| \geq p$?

4. *Demon:* divides s into xyz such that $|y| > 0$ and $|xy| \leq p$
5. *You:* give the demon an i such that $xy^iz \notin L$.

Another approach for L_2

Suppose L_2 is regular.

Then $L_2 \cap$ 0*1* is regular. Why?

So L_2 can't be regular. Why?

$L_3 = \{ww \mid w \in \{0,1\}^*\}$

Viewed as game vs. a demon:

1. *You:* give the demon the language L_3
2. *Demon:* gives you p
3. *You:* give the demon what string s with $|s| \geq p$?

4. *Demon:* divides s into xyz such that $|y| > 0$ and $|xy| \leq p$
5. *You:* give the demon an i such that $xy^iz \notin L$.

$L_4 = \{0^i1^j \mid i > j\}$

Viewed as game vs. a demon:

1. *You:* give the demon the language L_4
2. *Demon:* gives you p
3. *You:* give the demon what string s with $|s| \geq p$?

4. *Demon:* divides s into xyz such that $|y| > 0$ and $|xy| \leq p$
5. *You:* give the demon an i such that $xy^iz \notin L$.

$L_5 = \{1^{n^2} \mid n \geq 0\}$

Viewed as game vs. a demon:

1. *You:* give the demon the language L_5
2. *Demon:* gives you p
3. *You:* give the demon what string s with $|s| \geq p$?

4. *Demon:* divides s into xyz such that $|y| > 0$ and $|xy| \leq p$
5. *You:* give the demon an i such that $xy^iz \notin L$.