BU CS 332 – Theory of Computation

Lecture 4:

- Non-regular languages
- Pumping Lemma

Reading:

Sipser Ch 1.4

Mark Bun

February 3, 2020

The Philosophical Question

- We've seen techniques for showing that languages are regular

Constructing DFAs
NFAs

Next time:
Regular expressions

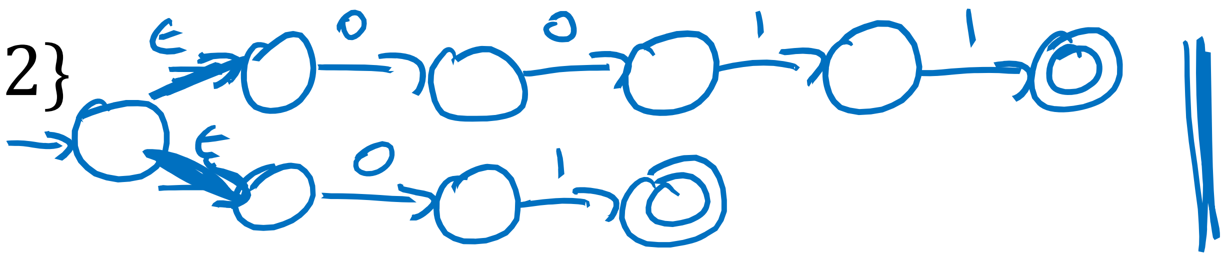
Closure properties

- Could it be the case that every language is regular?

Regular?

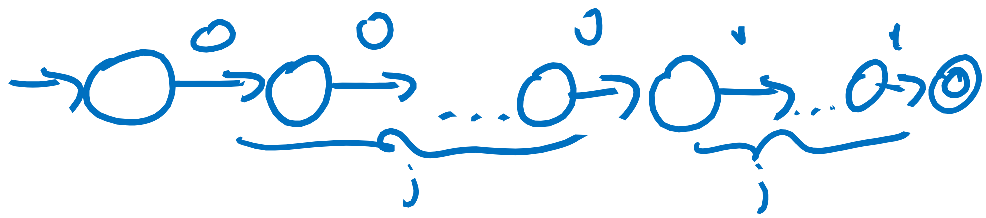
Construct an NFA for the following languages

$\{0^n1^n \mid 0 < n \leq 2\}$



$\{0^n1^n \mid 0 < n \leq \underbrace{k}_{=2}\}$ fixed constant k

For any $j \in \{1, \ldots, k\}$



$\{0^n1^n \mid n > 0\}$

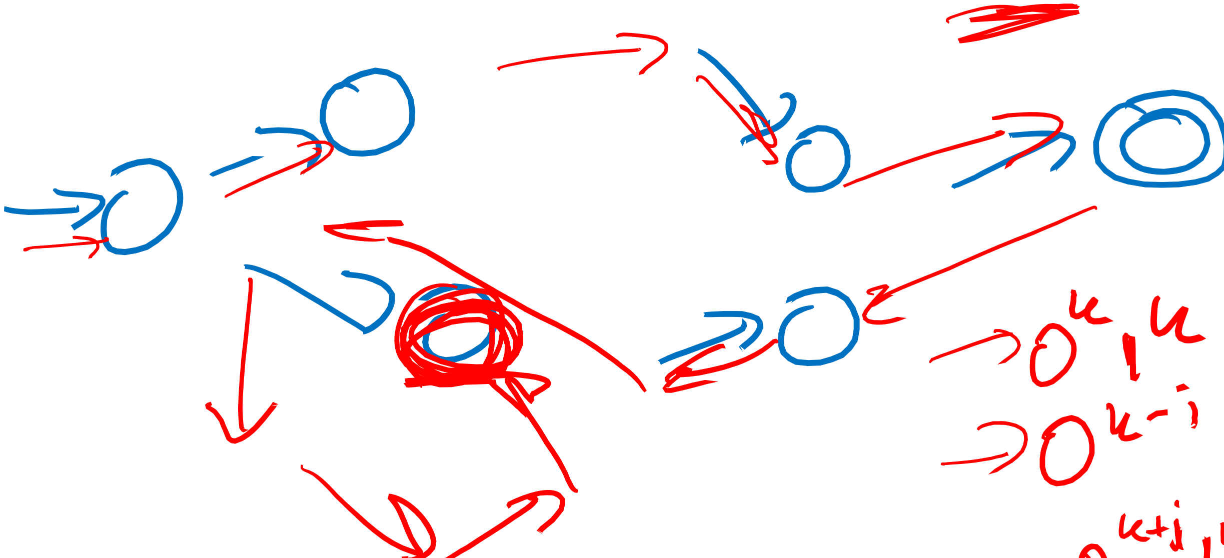
Proving a language is not regular

Theorem: $A = \{0^n 1^n \mid n > 0\}$ is not regular

Proof: (by contradiction)

Let M be a DFA with k states recognizing A

Consider running M on input $0^k 1^k$



$0^k 1^k \in A$

$0^{k-i}\, 0^i\, 0^i\, 1^k \in A$

$0^{k+i} 1^k$

$\times$

Regular or not?

$C = \{ w \mid w$ has equal number of 1s and 0s$\}$

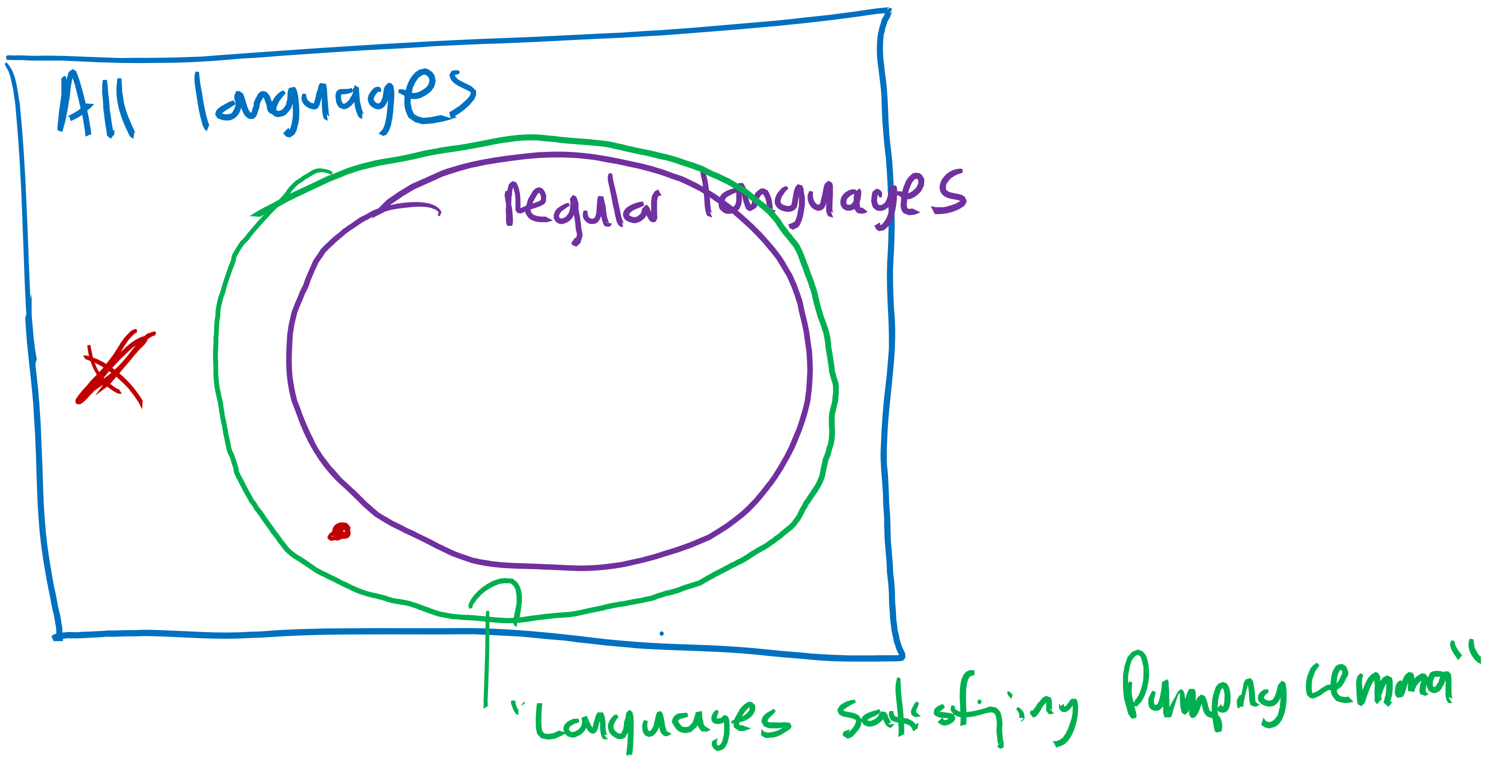
Not regular

$D = \{ w \mid w$ has equal number of 10s and 01s$\}$

Regular!

The Pumping Lemma

A **systematic** way to prove that a language is not regular



Why do we teach this?

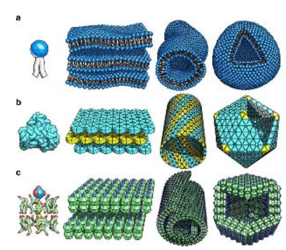
Cons:

- The statement is difficult (5 quantifiers!)
- Some non-regular languages can still be pumped

Myhill-Nerode Theorem

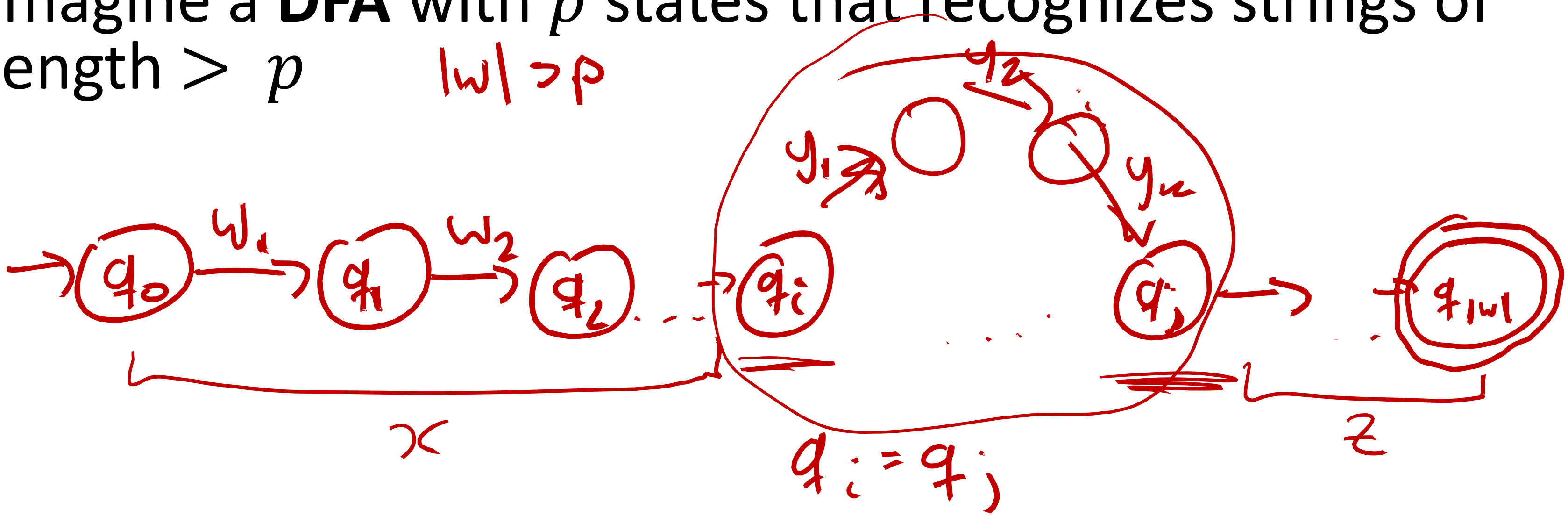
Pros:

- Proof illuminates essential structure of finite automata
- Generalizes to other models of computation / classes of languages (CFLs, self-assembly)
- Applying it can be fun!



Intuition for the Pumping Lemma

Imagine a **DFA** with p states that recognizes strings of length $> p$ $|w| > p$



$i = 0, 1, 2, \ldots$

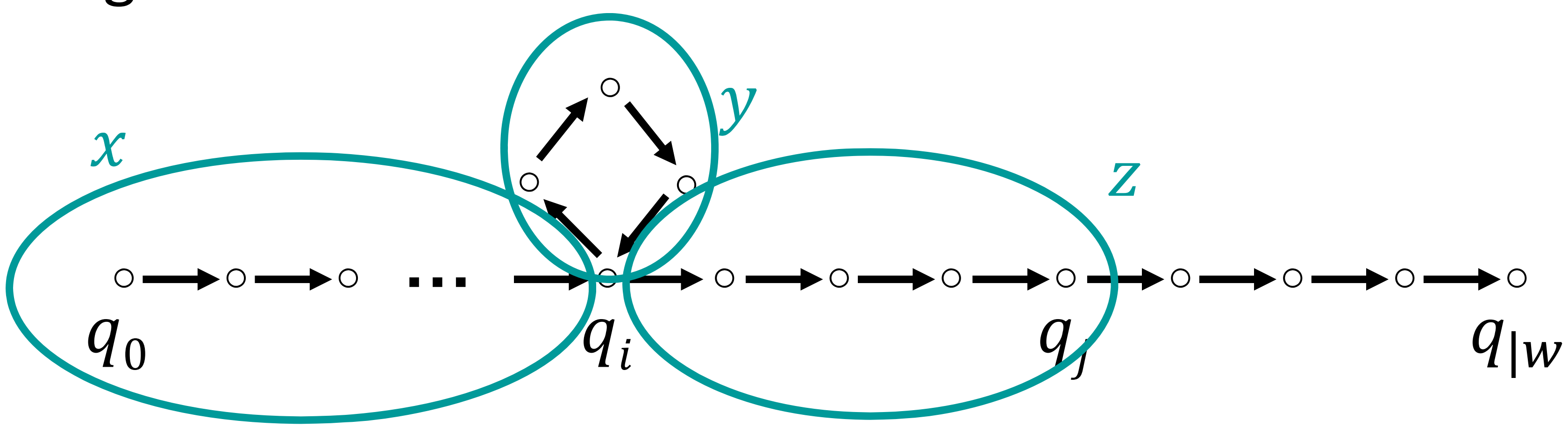
$w = xyz \in A$

$w = xyyz \in A$, $xyyyz \in A$, ... $xy^iz \in A$

Idea: If you can go around the cycle once, you can go around 0 or 2,3,4... times

Pumping Lemma (Informal)

Let L be a regular language. Let w be a “long enough” string in L.



Then we can write $w = xyz$ such that $xy^iz \in L$ for every $i \geq 0$.

$i = 0$: xz

$i = 1$: xyz

$i = 2$: xyyz

$i = 3$: xyyyz

Pumping Lemma (Formal)

Let L be a regular language.

Then there exists a "pumping length" p such that

For every $w \in L$ where $|w| \geq p$,

w can be split into three parts $w = xyz$ where:

1. $|y| > 0$ ⇐
2. $|xy| \leq p$
3. $xy^iz \in L$ for all $i \geq 0$

$w = a$
$x = \epsilon$
$y = a$
$z = \epsilon$

Example:
Let $L = \{w \mid$ all a's in w appear before all b's$\}$; $p = 1$

$w = aab$
$x = \epsilon$
$y = a$
$z = ab$

$w = aaabb$
$x = \epsilon$
$y = a$
$z = aabb$

Using the Pumping Lemma

Theorem: $A = \{0^n 1^n \mid n > 0\}$ is not regular

Proof: (by contradiction)

Assume instead that A is regular. Then A has a pumping length p.

What happens if we try to pump $0^p 1^p$?

$xy = 0^k$
$y = 0^j \quad 0 < j \le k$
$\Rightarrow x = 0^{k-j}$

If A is regular, w can be split into $w = xyz$, where

$\overbrace{00000 \ldots 00}^{p}\overbrace{11 \ldots 111}^{p}$
$x \quad y$

1. $|y| > 0 \Rightarrow y = 0^j$ for $0 < j \le k$
2. $|xy| \le p \Rightarrow xy = 0^k$ for some $0 < k \le p$
3. $xy^iz \in A$ for all $i \ge 0$

$$xyyz = \underbrace{0^{k-j}}_{x}\,\underbrace{0^j}_{y}\,\underbrace{0^j}_{y}\,\underbrace{0^{p-k}1^p}_{z} = 0^{p+j}1^p \notin A \quad \times$$

General Strategy for proving L is not regular

Proof by contradiction: assume L is regular.

Then there is a pumping length p.

1) Choose a string w, $|w| \geq p$

2) Show that for every decomposition $w = xyz$ (satisfying conditions of PL), w cannot be pumped

3) $\times \Rightarrow L$ nonregular

Pumping Lemma as a game

1. YOU pick the language L to be proved nonregular.
2. ADVERSARY picks a possible pumping length p.
3. YOU pick w of length at least p.
4. ADVERSARY divides w into x, y, z, obeying rules of the Pumping Lemma: $|y| > 0$ and $|xy| \leq p$.
5. YOU win by finding $i \geq 0$, for which $xy^i z$ is not in L.

If *regardless* of how the ADVERSARY plays this game, you can always win, then L is nonregular

Example: Palindromes

$0110 \in L$
$0011 \notin L$
$010010 \in L$
$101 \notin L$

Claim: $L = \{ww^R \mid w \in \{0,1\}^*\}$ is not regular

Proof: Assume L is regular with pumping length p

1. Find $w \in L$ with $|w| > p$

   $w = 0^p 1^p 1^p 0^p$

2. Show that w cannot be pumped

   $w = xyz$
   $|xy| \le p$

   Intuitively

$$\underbrace{\overset{x}{00}\,0\,\overset{y}{0}\ldots 0}_{p}\ \overbrace{\underbrace{1\ldots 1}_{2p}\ \underbrace{000\ldots 0}_{p}}^{z}$$

$x = 0^j \quad 0 \le j < p$
$y = 0^k \quad 0 < k \le p-j$
$z = 0^{p-(j+k)} 1^{2p} 0^p$

$xyyz = 0^j 0^{2k} 0^{p-(j+k)} 1^{2p} 0^p$
$= 0^{p+k} 1^{2p} 0^p \notin L$

Example: Palindromes

Claim: $L = \{ww^R \mid w \in \{0,1\}^*\}$ is not regular

Proof: Assume L is regular with pumping length p

1. Find $w \in L$ with $|w| > p$

2. Show that w cannot be pumped
   Formally If $w = xyz$ with $|xy| \leq p$, then…

Now you try!

Claim: $L = \{0^i 1^j \mid i > j \geq 0\}$ is not regular

Proof: Assume L is regular with pumping length p

1. Find $w \in L$ with $|w| > p$

2. Show that w cannot be pumped
   Intuitively