CSE 105
THEORY OF COMPUTATION

"Winter" 2018

http://cseweb.ucsd.edu/classes/wi18/cse105-ab/

Today's learning goals

Sipser Section 1.4

- Explain the limits of the class of regular languages
- Justify why the Pumping Lemma is true
- Apply the Pumping Lemma in proofs of nonregularity
- Identify some nonregular sets

Counting languages

How many languages over {0,1} are there?

A. Finitely many because {0,1} is finite.
B. Finitely many because strings are finite.
C. Countably infinitely many because {0,1}* is countably infinite.
D. Uncountably many because languages are in the power set of {0,1}*.
E. None of the above.

set of all languages = $P(\{0,1\}^*)$

Counting regular languages over {0,1}

| { regular languages} | ≤ | { regular expressions } |

Each regular expression is a finite string over the alphabet
{ 0 , 1 , ε , ø , (,) , U , * }

The set of strings over an alphabet is countably infinite.

Conclude: countably infinitely many regular languages.

Where we stand

Fact 1: There exist nonregular languages.

Fact 2: If we know some languages are nonregular, we can conclude others must be too. (cf. Discussion)

But, we don't have any specific examples of nonregular languages.

Yet.

Proving nonregularity

How can we prove that a set is non-regular?

A. Try to design a DFA that recognizes it and, if the first few attempts don't work, conclude there is none that does.
B. Prove that it's a strict subset of some regular set.
C. Prove that it's the union of two regular sets.
D. Prove that its complement is not regular.
E. I don't know.

Bounds on DFA

- in DFA, memory = states

- Automata can only "remember"…
  - …finitely far in the past
  - …finitely much information

- If a computation path visits the same state more than once, the machine can't tell the difference between the first time and future times it visited that state.

Example!

$\{ 0^n 1^n \mid n \geq 0 \}$ $= \{\varepsilon, 01, 0011, 000111, \ldots$

What are some strings in this set? $\{0^n 1^n \mid n \geq 0\} \subseteq L(0^* 1^*)$

What are some strings not in this set?

Compare to $L(0^*1^*)$ $= \{\varepsilon, 0, 00, 000, \ldots$ $01, 001, 0001, \ldots$ $\vdots$ $111, 1111, \ldots \}$

Design a DFA? NFA?

Example!

$\{0^n 1^n \mid n \geq 0\}$

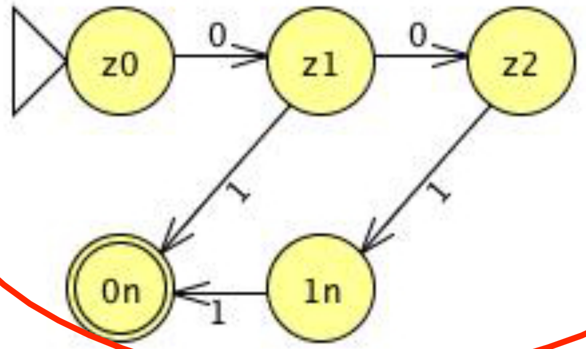
$\{0^n 1^k \mid n \geq 0, k \geq 0\} = 0^* 1^*$

$\{0^n 1^n \mid 0 \leq n \leq 75\}$

What are some strings in this set?

What are some strings not in this set?

Is this set finite or infinite?

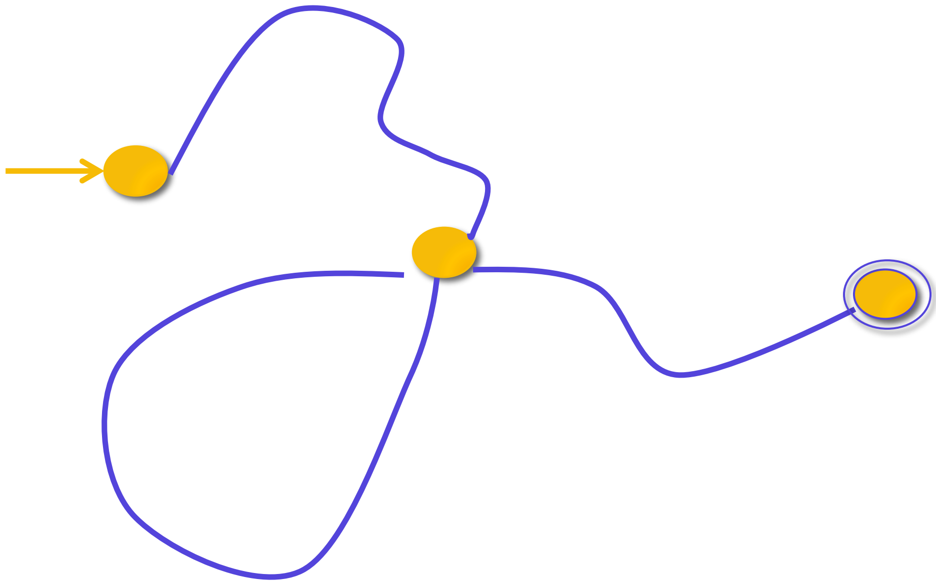


Compare to L(0^*1^*)

Design a DFA? NFA?

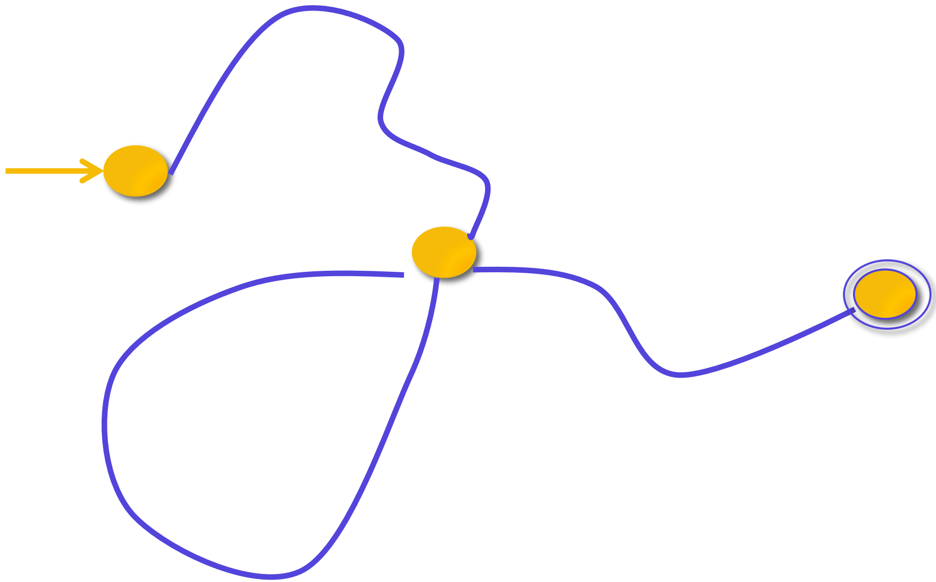
Pumping

- Focus on computation path through DFA



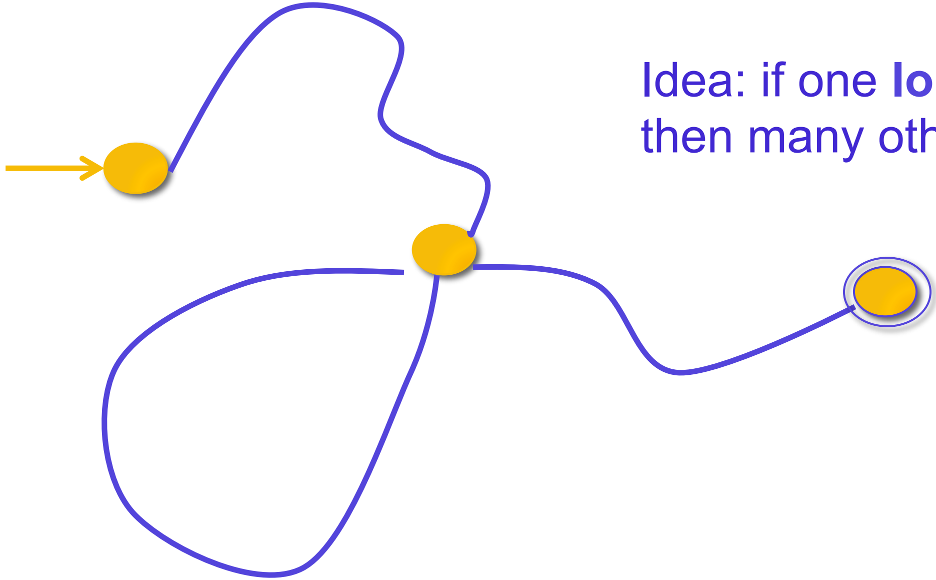
Pumping

- Focus on computation path through DFA



Pumping

- Focus on computation path through DFA



Idea: if one **long** string is accepted,
then many other similar strings have to be accepted too

Pumping Lemma

Sipser p. 78 Theorem 1.70

If A is a regular language, then there is a number p (*the pumping length*) where, if s is any string in A of length at least p, then s may be divided into three pieces, $\boxed{s = x\ y\ z}$ such that

- $|y| > 0$, and
- for each $i \geq 0$, $xy^iz \in A$,
- $|xy| \leq p$.

Pumping Lemma

Sipser p. 78 Theorem 1.70

If A is a regular language, then there is a number p (*the pumping length*) where, if s is any string in A of length at least p, then s may be divided into three pieces, $s = xyz$ such that

- $|y| > 0$, and
- for each $i \geq 0$, $xy^iz \in A$,
- $|xy| \leq p$.

states in DFA recognizing A

Transition labels along loop

Pumping Lemma

If A is recognized by DFA M with state diagram below,

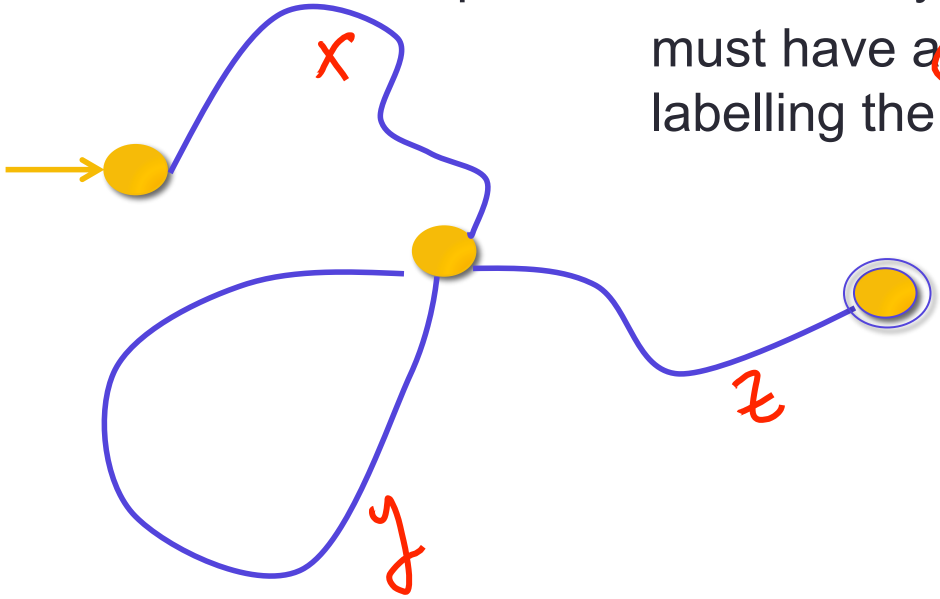
the computation of M on any string s of length >= p = |Q|

must have a **loop**. Divide s into the strings

labelling the path before the loop x ,

the loop itself y , and

from the loop to the accept state z



Pumping Lemma

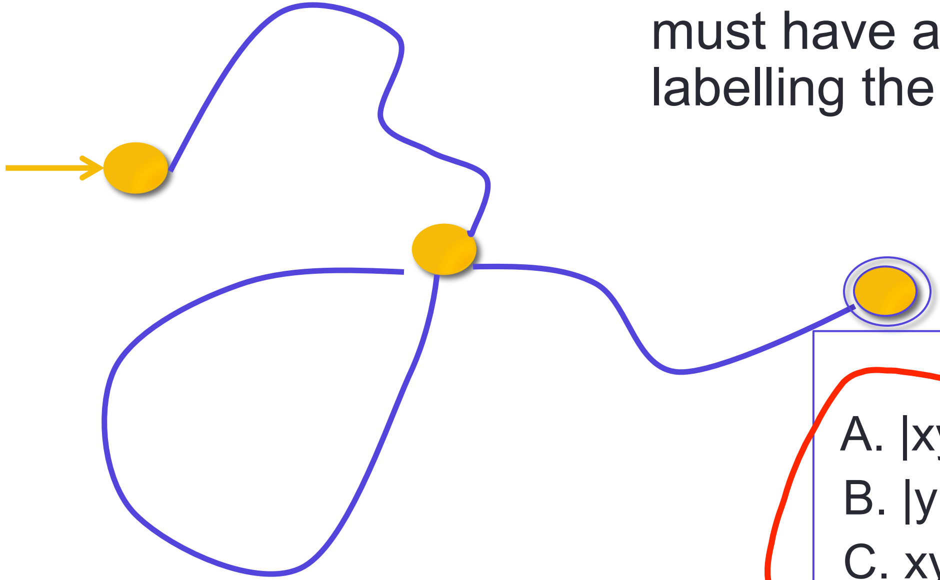
If A is recognized by DFA M with state diagram below,

the computation of M on any string s of length >= p = |Q|

must have a **loop**. Divide s into the strings labelling the path before the loop x ,

the loop itself y , and

from the loop to the accept state z



Which of the following is true?

A. |xy| <= p

B. |y| > 0

C. xy^iz is accepted by M for all i

D. All of A,B,C

E. None of them

Pumping Lemma

- True for **all** (but not only) regular sets.
  - Can't be used to prove that a set ***is*** regular
  - Can be used to a prove that a set ***is not*** regular … how?

A is regular → pumping lemma is true

pumping lemma fails → A is not regular

Negation

flash-back to CSE 20 ☺

- Pumping lemma ``**There is** p, where p is a pumping length for L"

- Given a specific number p, it being a pumping length for L means

$$\forall s \left((|s| \geq p \wedge s \in L) \rightarrow \exists x \exists y \exists z \left(s = xyz \wedge |y| > 0 \wedge |xy| \leq p \wedge \forall i (xy^i z \in L) \right) \right)$$

- So p **not** being a pumping length of L means

$$\exists s \left(|s| \geq p \wedge s \in L \wedge \forall x \forall y \forall z \left((s = xyz \wedge |y| > 0 \wedge |xy| \leq p) \rightarrow \exists i (xy^i z \notin L) \right) \right)$$

Proof strategy

To prove that a language L is **not** regular

- Consider arbitrary positive integer p.
- Prove that p isn't a pumping length for L.

- Conclude that L does not have any pumping length and is therefore not regular.

Using the Pumping Lemma

Claim: The set $L = \{0^n 1^n \mid n \geq 0\}$ is not regular.

let $p > 0$. let $s = 0^p 1^p \in L$

Using the Pumping Lemma

Claim: The set $L = \{0^n 1^n \mid n \geq 0\}$ is not regular.

Proof: Consider an arbitrary positive integer. WTS p is not a pumping length for L.

How? Want to show that there is some string that *should* be pump'able but isn't.

Using the Pumping Lemma

L = $\{0^n 1^n \mid n \geq 0\}$ **CLAIM:** p is not a pumping length for L.

How would you prove the claim?

A. Find a string with length >= p that is not in L.

B. Find a string with length <p that is in L.

C. None of the above.

$$\exists s\left(|s| \geq p \wedge s \in L \wedge \forall x \forall y \forall z\left((s=xyz \wedge |y|>0 \wedge |xy| \leq p) \rightarrow \exists i(xy^i z \notin L)\right)\right)$$

Using the Pumping Lemma

$L = \{0^n 1^n \mid n \geq 0\}$ **CLAIM:** p is not a pumping length for L.

WTS

$$\exists s\left(|s| \geq p \wedge s \in L \wedge \forall x \forall y \forall z\left((s = xyz \wedge |y| > 0 \wedge |xy| \leq p) \rightarrow \exists i(xy^iz \notin L)\right)\right)$$

Find a string s such that

$s = 0^p 1^p$

1. |s| >= p
2. s is in L
3. No matter how we cut s into three (viable) pieces, some related string obtained by repeating the middle part falls out of L.

Using the Pumping Lemma

Claim: The set L = **$\{0^n1^n \mid n \geq 0\}$** is not regular.

Proof: Consider an arbitrary positive integer. WTS p is not a pumping length for L.

Consider the string

$$s = 0^p1^p.$$

1. |s| >= p ? yes
2. s is in L ? yes
3. No matter how we cut s into three (viable) pieces, some related string obtained by repeating the middle part falls out of L ?

we shall see.

Using the Pumping Lemma

Claim: The set L = $\{0^n1^n \mid n \geq 0\}$ is not regular.

Proof: Consider an arbitrary positive integer p. WTS p is not a pumping length for L.

Consider the string $s = 0^p1^p$.

$|s| \geq p$ ✓

$s = xyz$

$|y| > 0$

$|xy| \leq p$

$x = 0^k \quad y = 0^m \quad z = 0^r 1^p$

$k + m + r = p$

$0^k 0^m 0^r 1^p \in L$

$0^k (0^m)^2 0^r 1^p \notin L$

$k + 2m + r \neq p$

Using the Pumping Lemma

Claim: The set L = **$\{0^n1^n \mid n \geq 0\}$** is not regular.

Proof: Consider an arbitrary positive integer. WTS p is not a pumping length for L. Consider the string $s = 0^p1^p$. Then, s is in L and $|s| = 2p >= p$. Consider any division of s into three parts

$$s = xyz \text{ with } |y|>0, |xy| \leq p.$$

Since $|xy| \leq p$, $x=0^k$, $y = 0^m$, $z = 0^r1^p$ with $k+m+r = p$, and since $|y| > 0$, $m>0$. Picking $i=0$: $xy^iz = xz = 0^k0^r1^p = 0^{k+r}1^p$, which is not in L because $k+r < p$. Thus, no p can be a pumping length for L and L is not regular.


WOW!

Another example

Claim: The set $\{a^n b^m a^n \mid m,n \geq 0\}$ is not regular.

Proof: Consider an arbitrary positive integer. WTS p is not a pumping length for L.

Consider the string

$$s = ???$$

$a^p b a^p$ $a^p a^p$

exercise:

1. |s| >= p ?
2. s is in L ?
3. No matter how we cut s into three (viable) pieces, some related string obtained by "repeating" the middle part falls out of L ?

For next time

- Work on Group Homework 2 **due Saturday**

Pre class-reading for Monday: Examples 1.75, 1.77.