Today's learning goals

- 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

 HW2 solutions on Piazza
 Group HW2 due Saturday
 Exam 1 next week
 - Practice Questions on Website
 - Review session on Monday
 - No Discussion sections on Day of Exam
All roads lead to … regular sets?

Are there any languages over \{0,1\} that are not regular?

A. Yes: a language that is recognized by an NFA but not any DFA.

B. Yes: there is some finite language over \{0,1\} that is not described by any regular expression.

C. No: all languages over \{0,1\} are regular because that's what it means to be a language.

D. No: for each set of strings over \{0,1\}, some DFA recognizes that set.

E. None of the above.
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.
But where to start?
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.

\{100001001\} \{\varepsilon\} \{0\} \{1\} \{\varepsilon,0,1,00,01,10,11,000,\ldots\}

\{\text{Languages over } \{0,1\}\} = \{\text{set of strings over } \{0,1\}\} = \{\text{subset of } \{0,1\}\}
Counting regular languages over \{0,1\}

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

Each regular expression is a finite string over the alphabet \{0, 1, \varepsilon, \emptyset, (, ), 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.
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^n1^n \mid n \geq 0 \}

What are some strings in this set? 
\{ 0, 01, 001 \}

What are some strings not in this set? 
\{ 10, 001 \}

Is this set finite or infinite? 
Only infinite.

Compare to \( L(0^*1^*) = \{ 0^m1^n \mid m, n \geq 0 \} \supseteq \{ 0^n1^n \mid n \geq 0 \} \)

Design a DFA? NFA?
Example!

\{ 0^n1^n \mid n \geq 0 \}

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

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, $s = x y z$ such that

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

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, $s = x y z$ such that

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

Sipser p. 78 Theorem 1.70

# 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 $\geq p = |Q|$ 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 $z$. 
If A is recognized by DFA M with state diagram below, the computation of M on any string s of length \( \geq p = |Q| \) 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. \(|xy| \leq p\)
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  
  Ex 1.49
• Can be used to prove that a set is not regular … how?
Negation

There is p, where p is a pumping length for L

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

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

So p not being a pumping length of L means

\[ \exists s \left( |s| \geq p \land s \in L \land \forall x \forall y \forall z \left( (s = xyz \land |y| > 0 \land |xy| \leq p) \rightarrow \exists i (xy^iz \notin L) \right) \right) \]
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.
For next time

• Work on Group Homework 2 due Saturday

Pre class-reading for Friday: Examples 1.75, 1.77.
Using the Pumping Lemma

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

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

L = \{0^n1^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 \( \geq p \) that is not in L.
B. Find a string with length <p that is in L.
C. None of the above.

\[ \exists s \ (|s| \geq p \land s \in L \land \forall x \forall y \forall z \ ((s = xyz \land |y| > 0 \land |xy| \leq p) \rightarrow \exists i (xy^iz \notin L)) \]
Using the Pumping Lemma

\[ L = \{0^n1^n \mid n \geq 0\} \]

**CLAIM:** \( p \) is not a pumping length for \( L \).

**WTS**

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

Find a string \( s \) such that
1. \( |s| \geq p \)
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 \( L \).
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. \(|s| \geq p\) ?
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 \( L \)?
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 \( |s| = 2p \geq p \). 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.