CSCI 361 Lecture 7:
Pumping Lemma

Shikha Singh

Announcements & Logistics

- **HW 2** was due last night
- **HW 3** will be released today and due next Wed (Oct 2)
- Hand in **reading questions # 4** and pick up **reading questions #2**
  - Low stakes incentive to follow along with textbook readings
  - If you have questions about the reading, bring them to class!

Last Time

- Discussed tools for identifying and proving languages are **not regular**
- Finite languages are always regular
- Infinite languages that have finite # of equivalence classes are regular
- Intuition: contain "repetitive patterns" that are not ***distinguishable***

Today

- Prove several languages are not regular
    - Using infinite fooling sets and MN Theorem
- Discuss alternate tools to show a language is not regular
    - Pumping lemma from the textbook
    - Properties of regular languages and known non-regular languages

Myhill-Nerode Theorem

Let L be a language over Σ^*, then L is regular **if and only if** the relation $\equiv_L$ over Σ^* has a finite number of equivalence classes.

Fooling Sets

Definition. A set of strings $S \subseteq \Sigma^*$ is a **fooling set** with respect to a language $L \subseteq \Sigma^*$ if every pair of strings in S is distinguishable with respect to each other.

- What is the difference between fooling set and a set consisting of one string from each equivalence class of $\equiv_L$?
  - Contains one string each from different equivalence classes but may not include have a representative from each class
  - Size of any fooling set for $L \leq$ # of equivalence class of $\equiv_L$

How To Come Up These Strings?

Some strings that may be in the fooling set (are mutually distinguishable):

- Prefixes of strings in L (correspond to DFA states from which an accept state is reachable)
  - The empty string ε
  - Whole strings that are in L
  - One element of each equivalence class
- Exactly one string that is not a prefix of any string in L
  - Why only one of these?
  - Which DFA states does this correspond to?

Myhill-Nerode Theorem

Maximum fooling set size of L

$=$ # equivalence classes of $\equiv_L$

$=$ minimum states of DFA for L

HW 3 Question. Show that if a regular language L has k equivalence classes then there is a DFA with k states for L.

Proving Non-Regularity

Problem. Prove that the language $L = \{a^i b^i \mid i \in \mathbb{N}\}$ is not regular.

Proof. Consider the infinite set $S = \{a^i | i \in \mathbb{N}\}$.

For any pair of strings a^i and a^j in S such that $i \neq j$, there exists a distinguishing suffix b^i such that. $a^i b^i \in L$ but $a^j b^i \notin L$.

Thus, L has an infinite fooling set, and thus an infinite # of equivalence classes and is not regular by Myhill-Nerode Theorem.

Using Closure and Known Regular Languages

Problem. Prove that the language

$L = \{w \in \{0,1\}^*$ has an equal number of 0s and 1s$\}$ is not regular.

Hint. Use the fact that $L = \{0^i 1^i \mid i \in \mathbb{N}\}$ *is not regular and closure properties of regular languages.*

Proving Non-Regularity

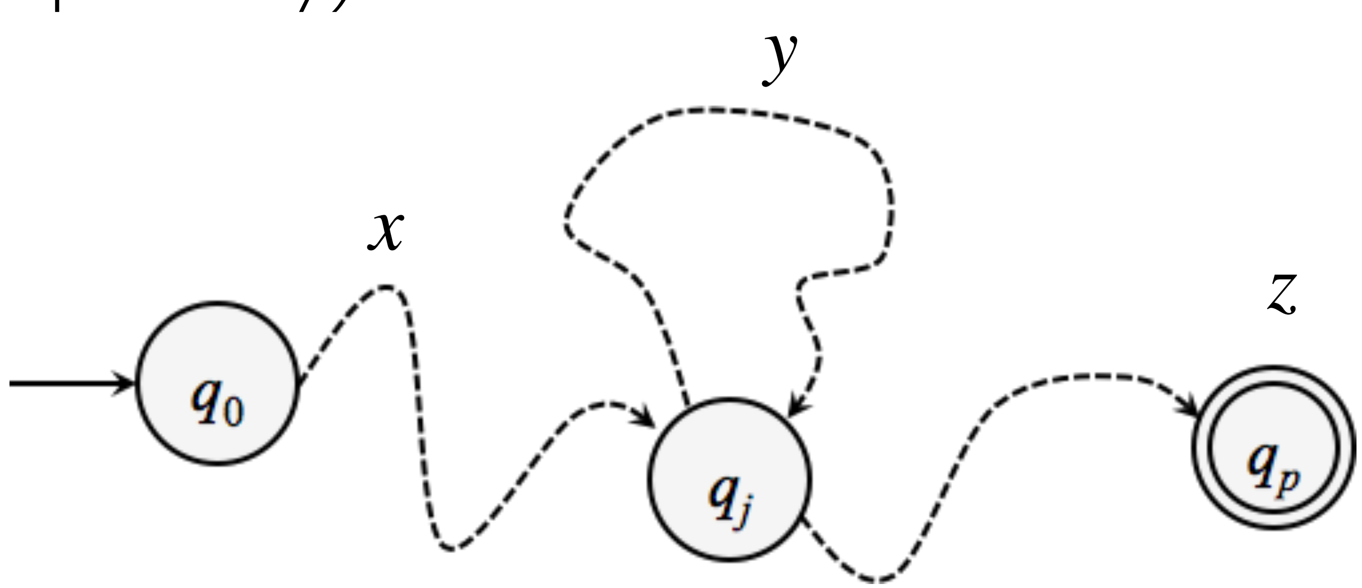
Problem. Prove that the language

$L = \{a^n \mid n \in \mathbb{N}$ and n is a power of $2\}$ is not regular.

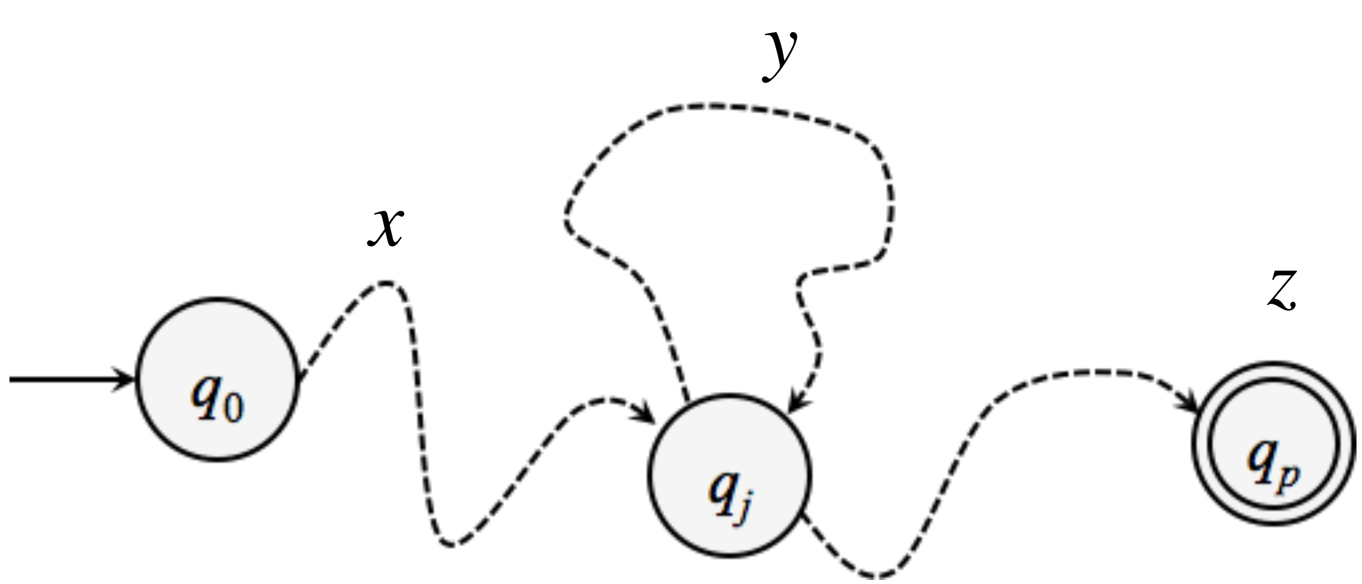
Hint. Identify and prove that L has an infinite fooling set.

Pumping Lemma: Intuition

- If DFA M has p states then M visits a state more than once on any string with length at least p
  - Number of states visited = length of string $+ 1$
- Let $w = xyz$ be the string that is accepted such that y is component in between the first repeated state (q_j)
  - Then xy^iz should also be accepted (can "pump" the middle piece repeatedly)



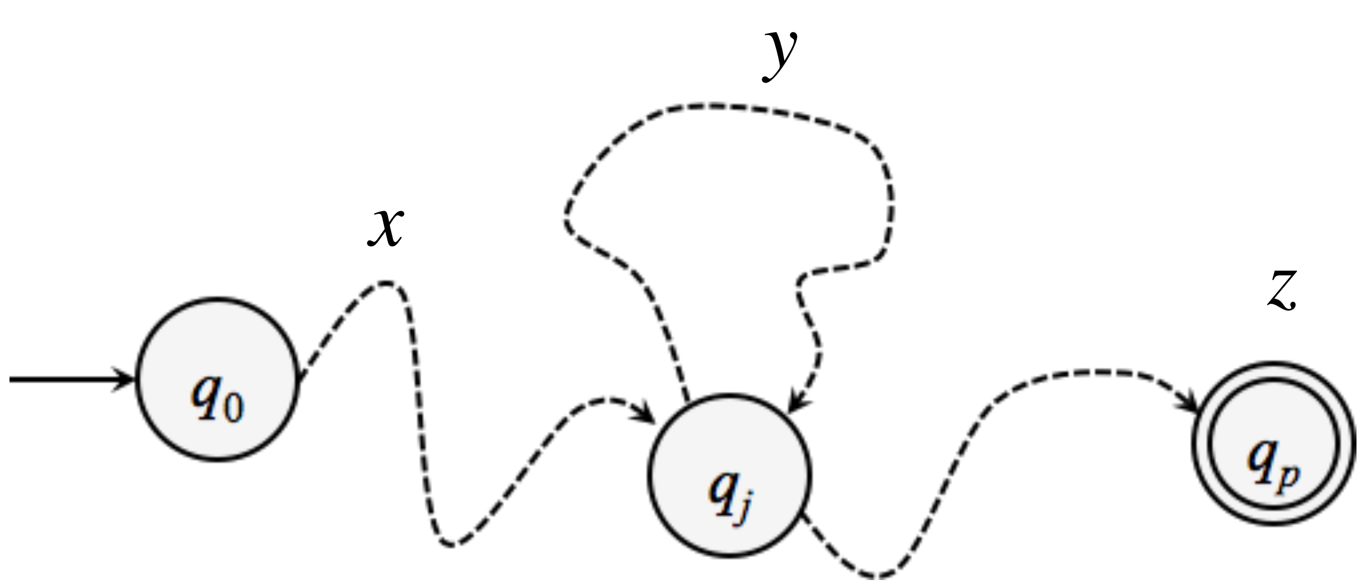

PUMP ALL THE STRINGS!
imgflip.com


y
x
z
q_0
q_j
q_p

Formal Statement

Pumping Lemma. If L is a regular language, then there exists a number p where if $w \in L$ is any string of length at least p, then w may be divided into three pieces $w = xyz$ such that:

1. $|y| > 0$
2. $|xy| \leq p$ (y must appear amongst the first p symbols)
3. for each $i \geq 0$, $xy^i z \in L$



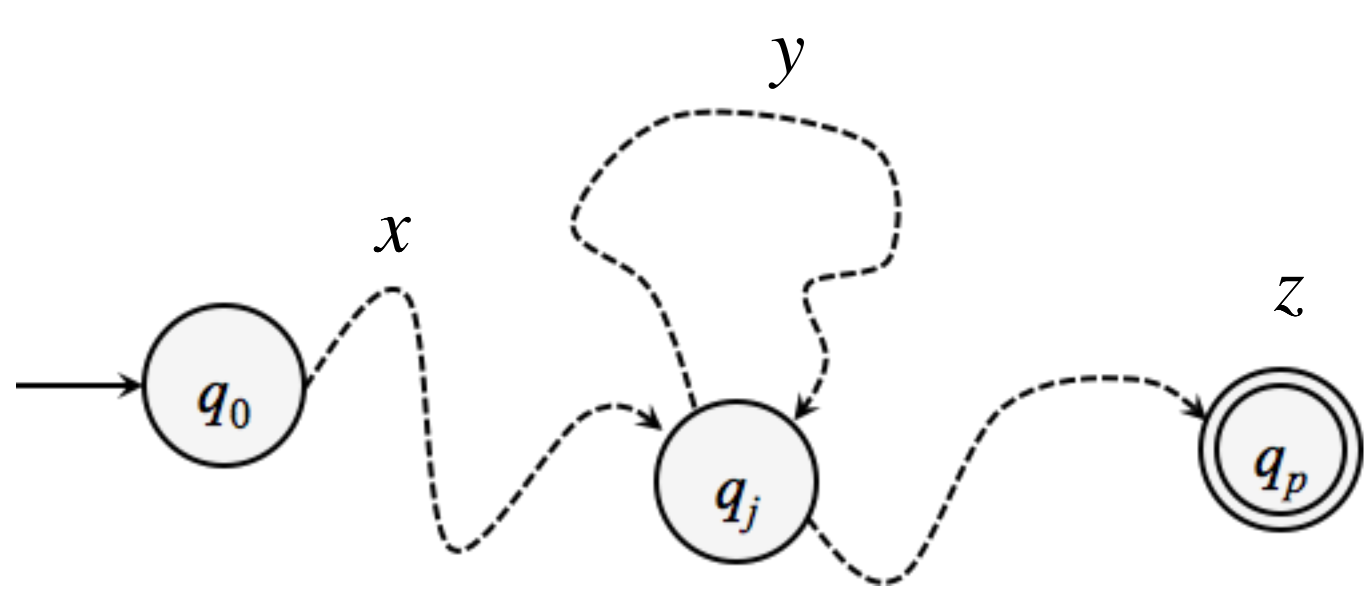
Questions

- Do all regular languages satisfy the pumping lemma?
- If a language satisfies the pumping lemma, does that mean it is regular?

Pumping Lemma Proof

Proof. Let DFA M for L have p states. Let $w = w_1 \cdots w_n$ such that $n \geq p$ and $q_0, q_1, \ldots, q_n$ be the states entered by M on w. M must revisit a state in the first p symbols. Let q_j and q_k be the first and second occurrence of this state.

Let $x = w_1 w_2 \cdots w_{j-1}$, $y = w_j w_{j+1} \cdots w_k$ and $z = w_{k+1} \cdots w_n$ which satisfies the conditions (1) and (2). Condition (3) follows from the fact that the strings xy^i are all ***indistinguishable*** wrt M.



Using Pumping Lemma

Problem. Prove that the language $L = \{a^i b^i \mid i \in \mathbb{N}\}$ is not regular using the pumping lemma.

More Practice

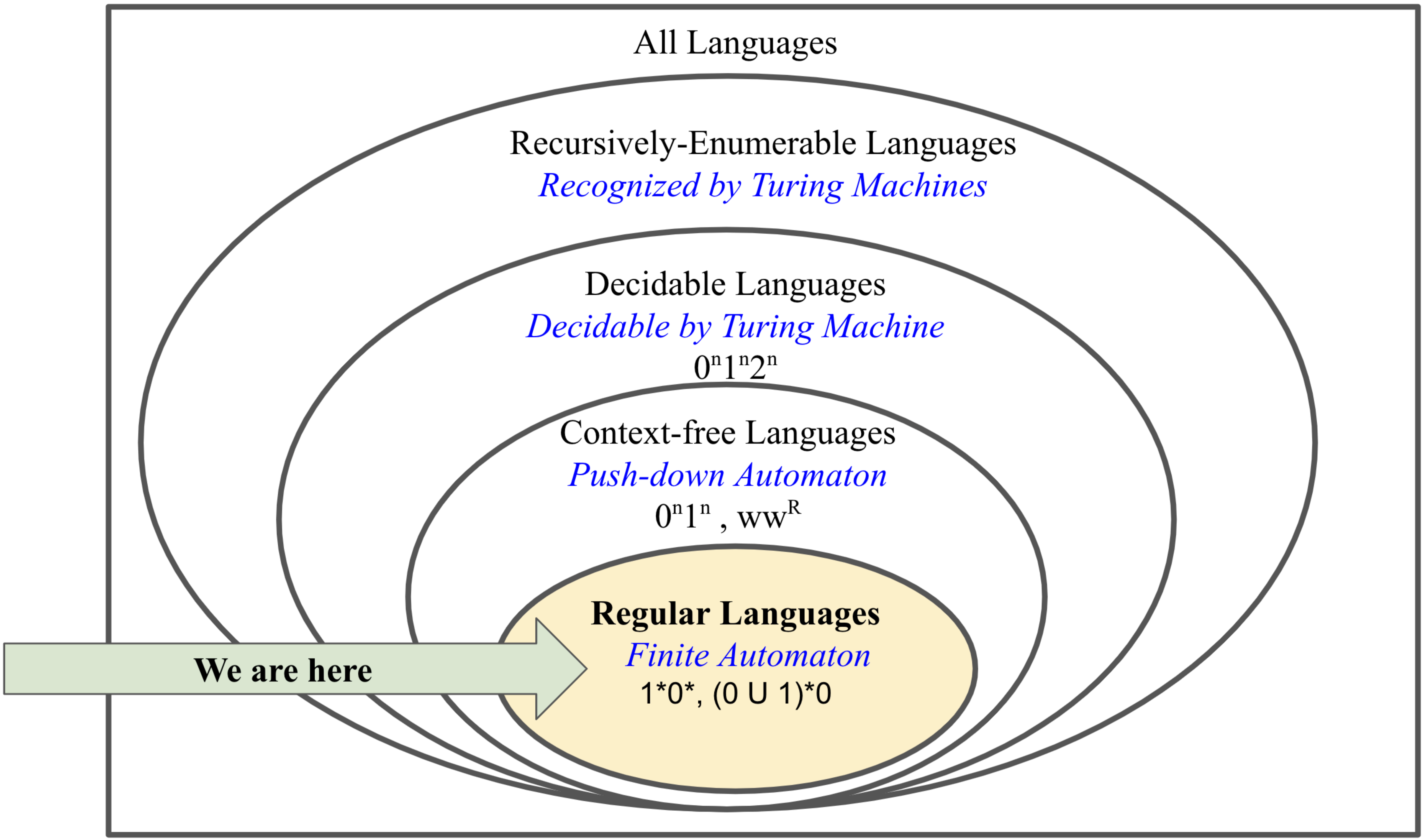
Problem. Prove that $L = \{ww \mid w \in \{0,1\}^*\}$ is not regular.

More Practice

Question. Is the language $L = \{(ab)^i \circ (ab)^i \mid i \geq 0\}$ regular?

More Practice

Question. Are all subsets of regular languages also regular?


All Languages
Recursively-Enumerable Languages
Recognized by Turing Machines
Decidable Languages
Decidable by Turing Machine
0n1n2n
Context-free Languages
Push-down Automaton
0n1n , wwR
Regular Languages
Finite Automaton
1*0*, (0 U 1)*0
We are here