CSC 339 – Theory of Computation
Fall 2023

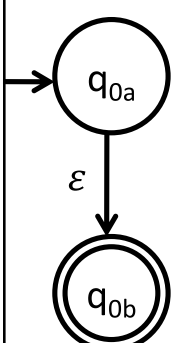
6. The Pumping Lemma

Outline

- Non-regular languages
- The pigeonhole principle
- The pumping lemma

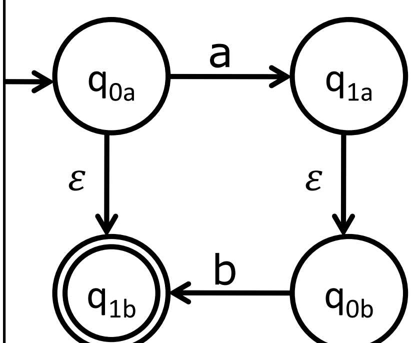
Non-Regular Languages

- $L_0 = \{a^k b^k : k \leq 0\} = \{\varepsilon\}$ is a regular language



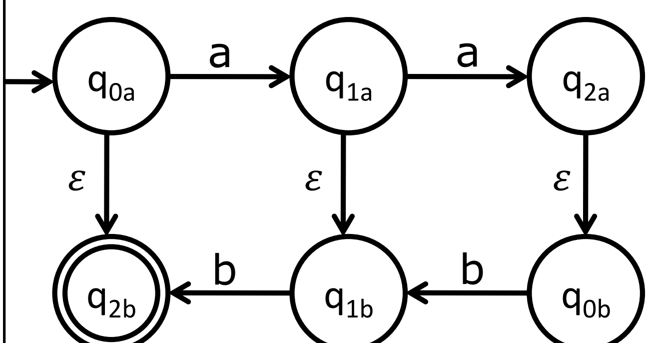
Non-Regular Languages

- $L_1 = \{a^k b^k : k \leq 1\} = \{\varepsilon, ab\}$ is regular



Non-Regular Languages

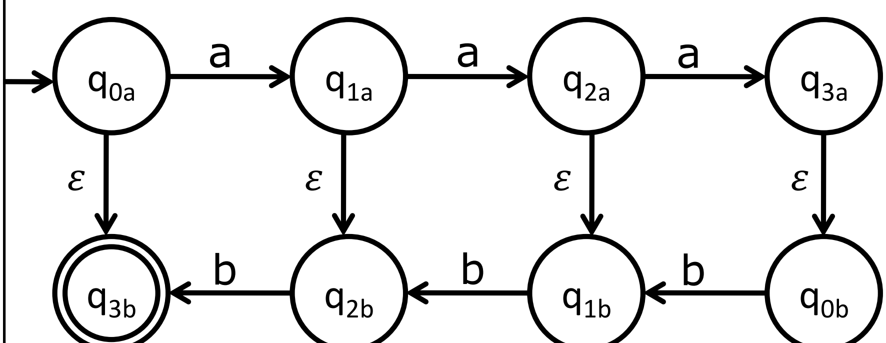
- $L_2 = \{a^k b^k : k \leq 2\} = \{\varepsilon, ab, aabb\}$ is regular



5

Non-Regular Languages

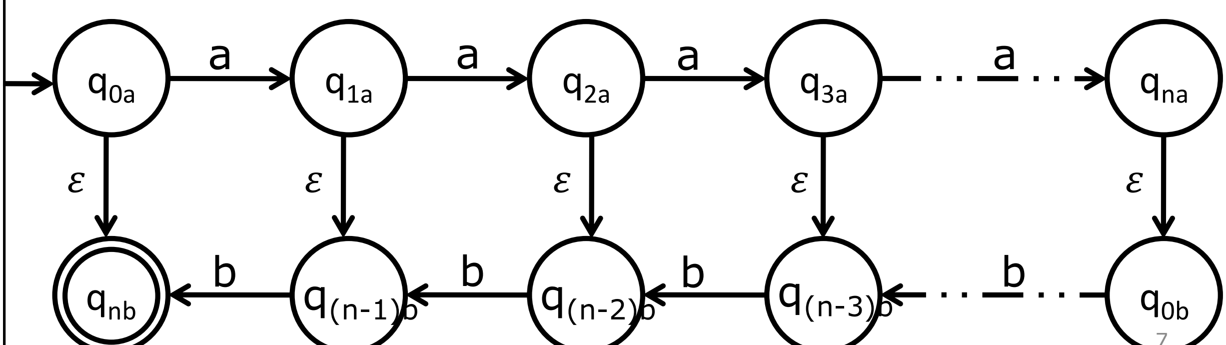
- $L_3 = \{a^k b^k : k \leq 3\}$ is regular



6

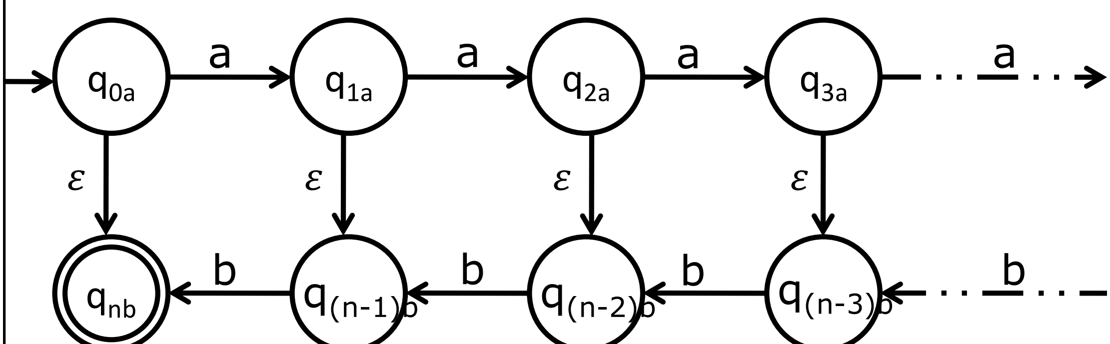
Non-Regular Languages

- $\forall\ n \geq 0$, $L_n = \{a^k b^k : k \leq n\}$ is regular



Non-Regular Languages

- However for any $n \geq 0$, $L_n = \{a^n b^n : n \geq 0\}$
  L_n is not a regular language.
- We need an infinite number of states to build this automaton!

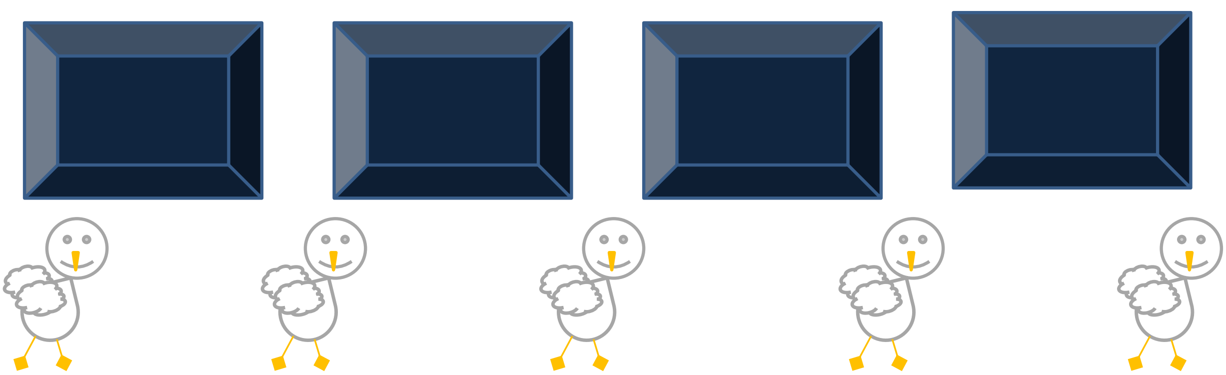


Proof

- We need mathematical proof that there is no FA that accepts L = $\{a^n b^n : n \geq 0\}$

The Pigeonhole Principle

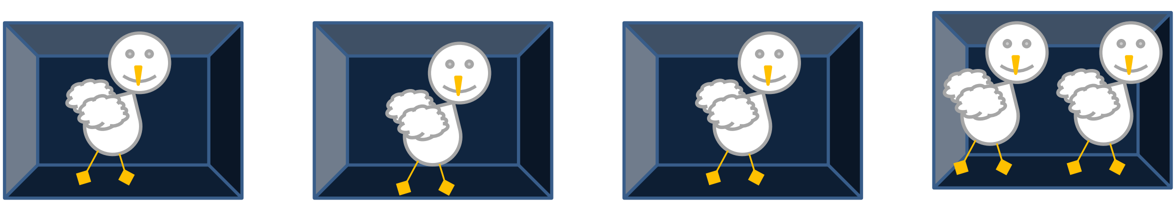
- If we have n holes and m pigeons ($m>n$) then there is a hole with at least two pigeons.



10

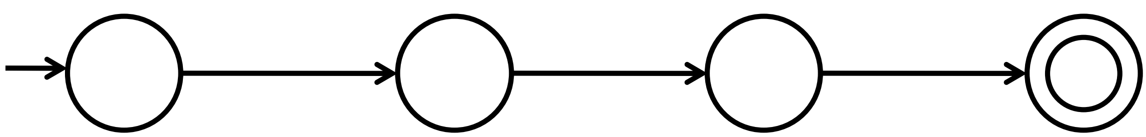
The Pigeonhole Principle (PP)

- If we have n holes and m pigeons ($m>n$) then there is a hole with at least two pigeons.



PP and Finite Automata

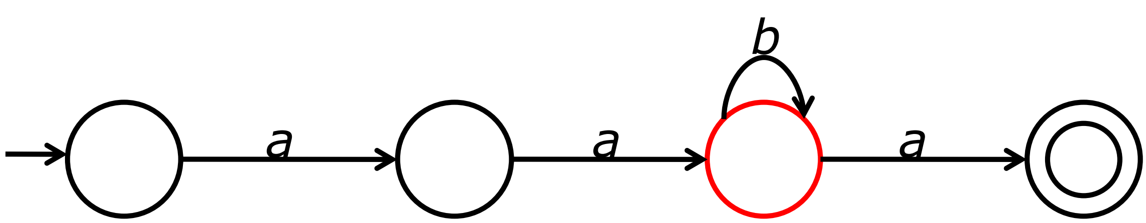
- If an automaton with n states accepts a string with length m ($m \geq n$), there should be at least one repeating state for every accepting path.



$s = aaba$

PP and Finite Automata

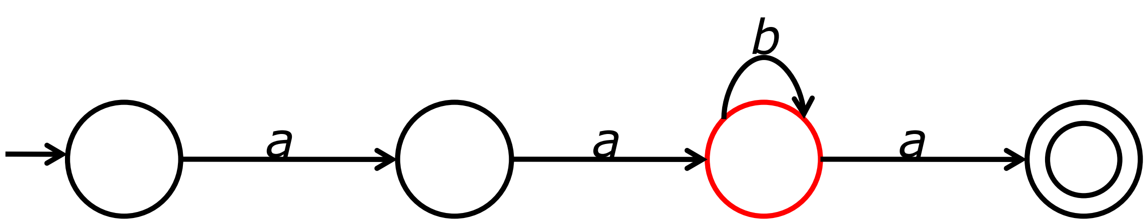
- If an automaton with n states accepts a string with length m ($m \geq n$), there should be at least one repeating state for every accepting path.



$s = aaba$

PP and Finite Automata

- If an automaton with n states accepts a string with length m ($m \geq n$), there should be at least one repeating state for every accepting path.

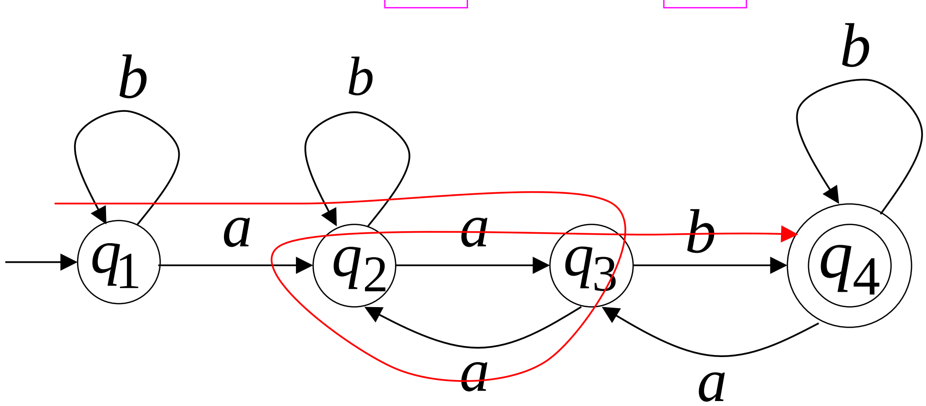


Any string of the form *aab*a* should be accepted!

Consider the walk of a "long" string: *aaaab*
(length at least 4)

A state is repeated in the walk of *aaaab*

$$q_1 \xrightarrow{a} q_2 \xrightarrow{a} \boxed{q_3} \xrightarrow{a} q_2 \xrightarrow{a} \boxed{q_3} \xrightarrow{b} q_4$$



15

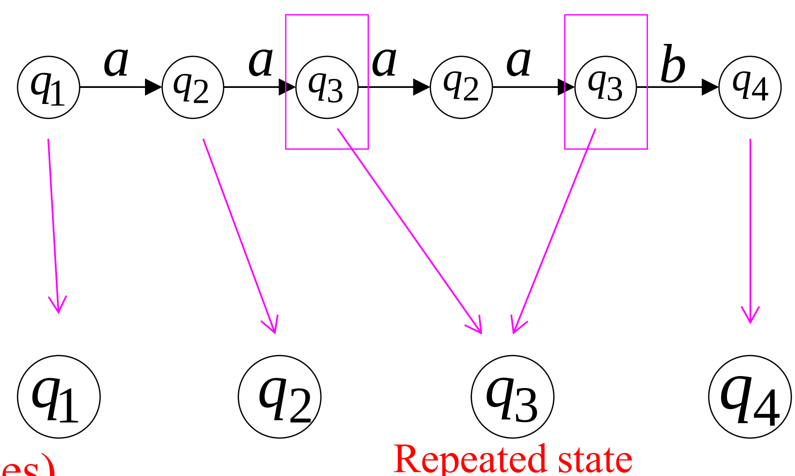
The state is repeated as a result of the pigeonhole principle

Walk of *aaaab*

Pigeons:
(walk states)

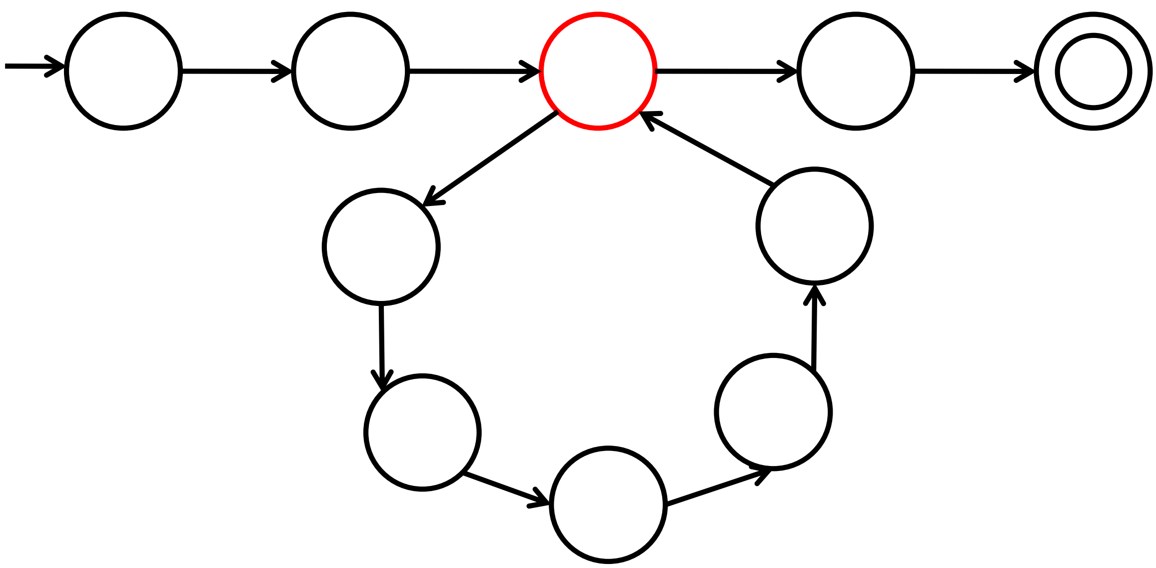
Are more than

Nests:
(Automaton states)



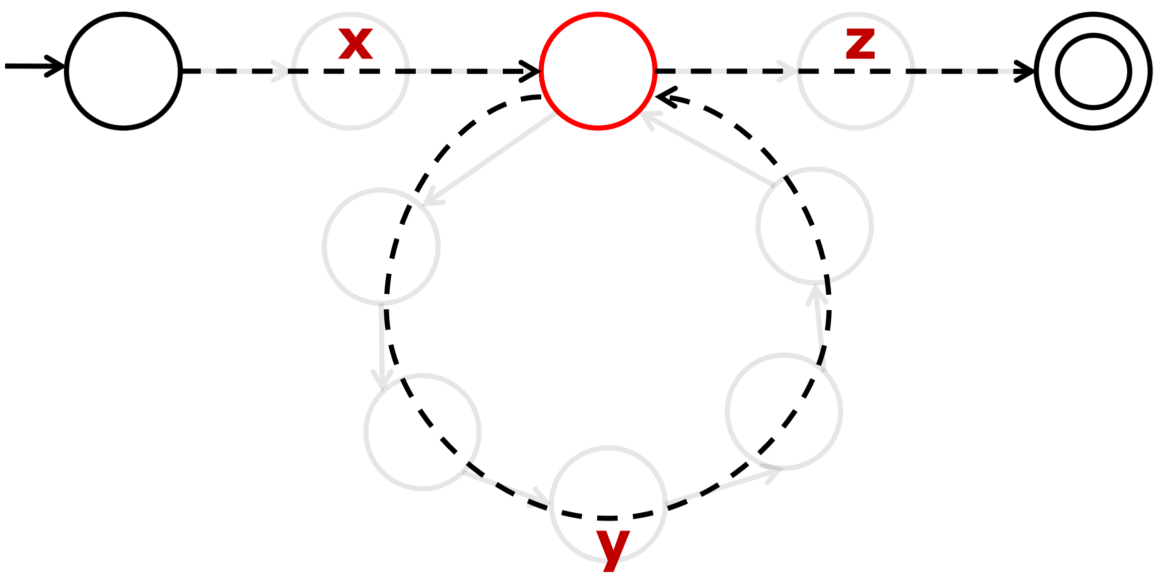
16

PP and Finite Automata



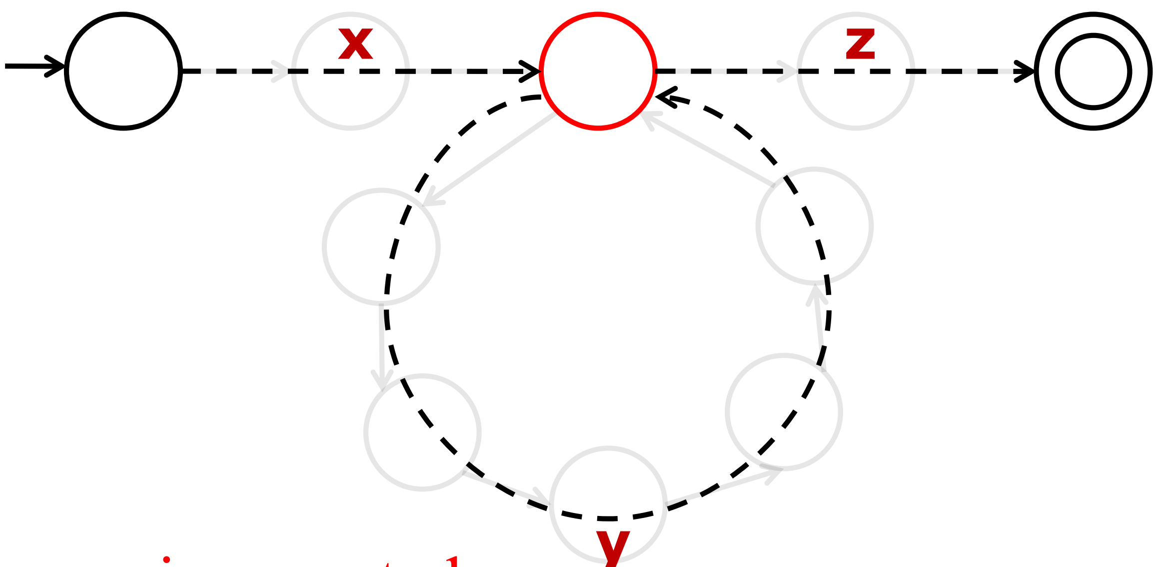
17

PP and Finite Automata



18

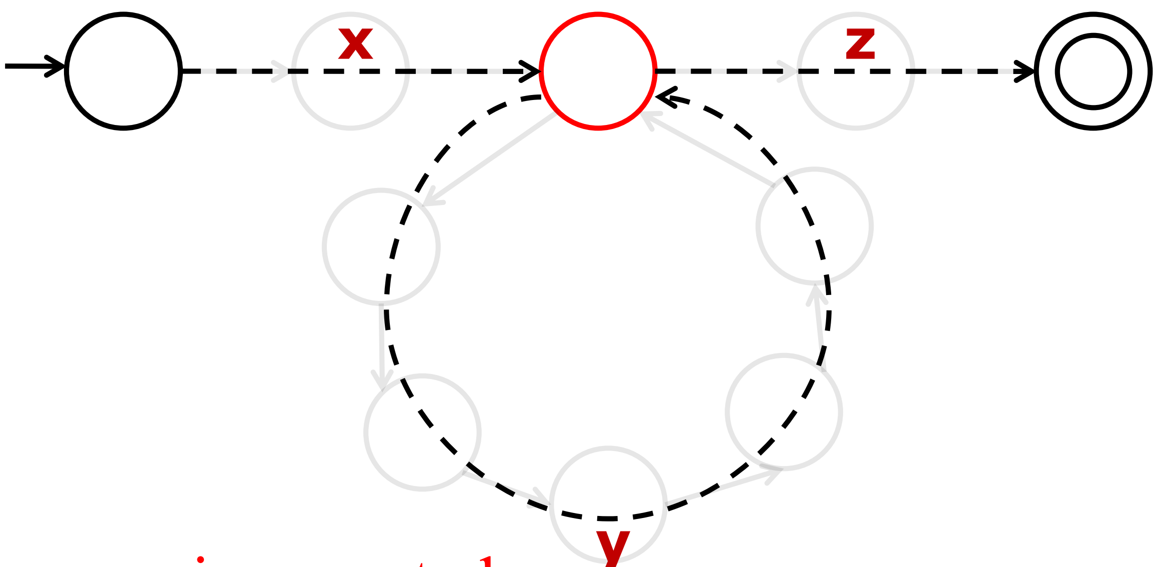
PP and Finite Automata



xyz is accepted

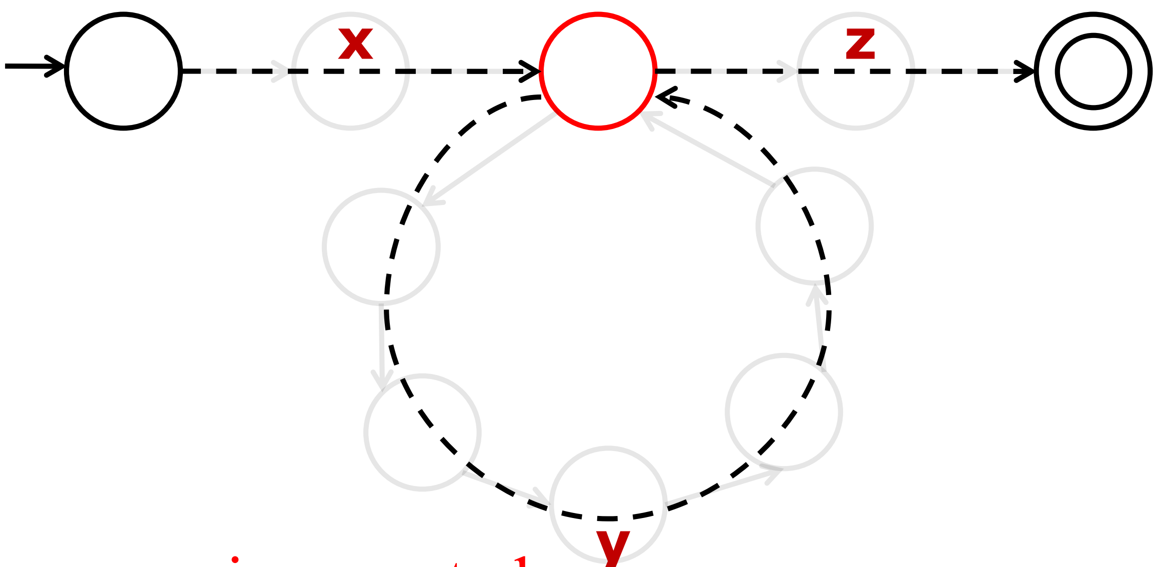
19

PP and Finite Automata



xyyz is accepted

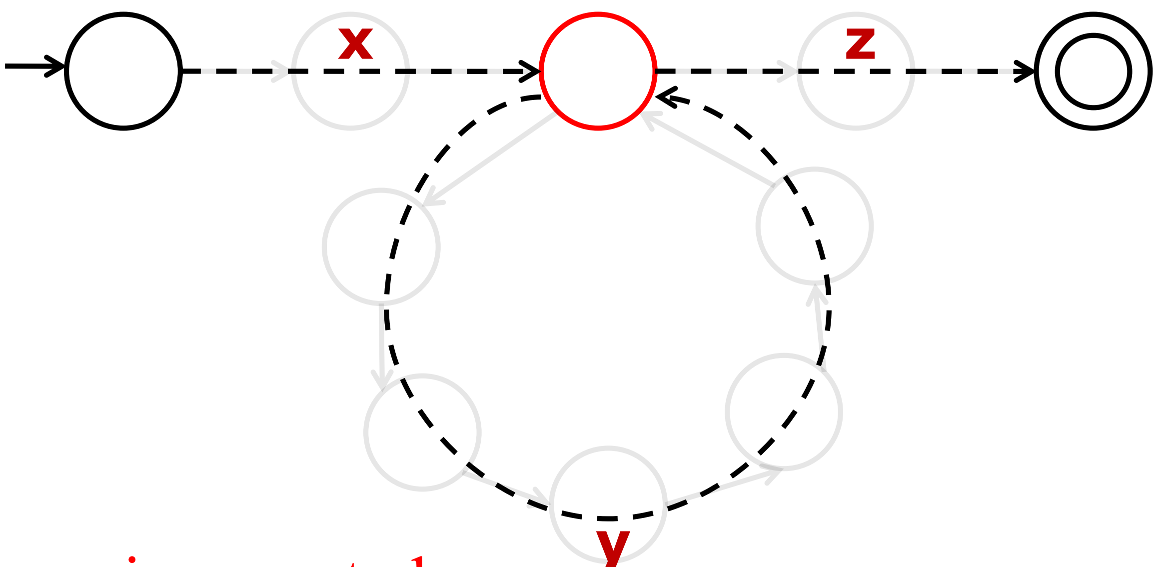
PP and Finite Automata



xyyyz is accepted

21

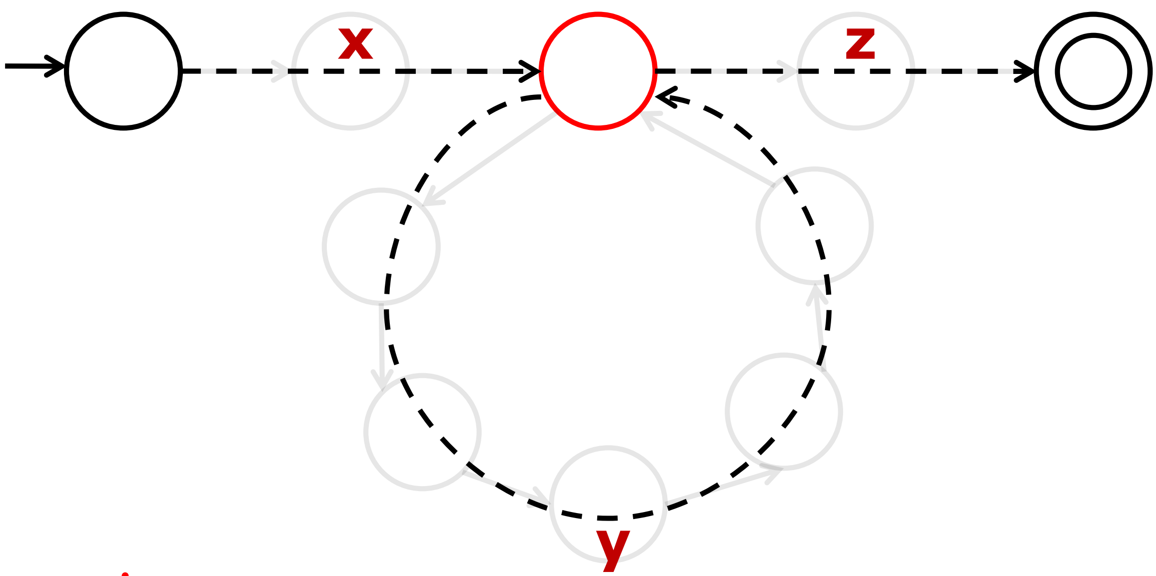
PP and Finite Automata



xz is accepted

22

PP and Finite Automata



xy^iz, for $i \geq 0$ is accepted

The pumping lemma

For every infinite regular language L
there exists a pumping length $k > 0$ such that
for any string s in L with length $|s| \geq k$
we can write $s = xyz$
with $|xy| \leq k$ and $|y| \geq 1$
such that xy^iz *in L* for every $i \geq 0$.

Proof

- If L is regular then there exists a DFA M which accepts L. Set k to be the number of states of M.
- L is infinite, there exists a string s with a length greater than k.
- The number of states of M is k.
- The string is of length at least k, there is a part in the path that is repeated.

Proof

- Split s into 3 parts x, y, z with y being the first repeated part.
- Since we have k states the first repetition should take place ($|y| \geq 1$) in at most k transitions ($|xy| \leq k$).
- Since the path under y is a loop we can follow it as many times as we want (maybe none).
- Thus xy^iz for any $i \geq 0$ should lead us to the same accepting state as xyz.

Prove that L is not regular

- Given is an infinite language L.

Prove that L is not regular

- Given is an infinite language L
- If L is regular

Prove that L is not regular

- Given L an infinite language
- If L is regular
- Then the Pumping lemma holds:
  - *There exists* a pumping length k such that
  - *for all* strings s ($|s| \geq k$) in L
  - *there is* a splitting of s in x, y, and z ($|xy| \leq k$ and $|y| \geq 1$) such that *for all* i, xy^iz is in L.

29

Prove that L is not regular

- Given is an infinite language L
- If L is regular the pumping lemma holds.
- The negation of the pumping lemma:
  - *For all* pumping lengths k
  - *there exist* a string s ($|s| \geq k$) in L such that
  - *for every possible* splitting of s in x, y, and z ($|xy| \leq k$ and $|y| \geq 1$) *there is* an i for which xy^iz is *not* in L.

Prove that L is not regular

- Given is an infinite language L
- Assume L is regular (pumping lemma holds)
- Prove the negation of the pumping lemma:
  - *Fix* an *arbitrary* pumping length k.
  - *Specify* a string s ($|s| \geq k$) in L.
  - Show that *for every possible splitting* of s in x, y, and z ($|xy| \leq k$ and $|y| \geq 1$)
  - *there is* an i for which xy^iz is *not* in L.
- Contradiction! L is then not regular.

Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Assume that L is regular. The pumping lemma holds!

Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Fix an arbitrary pumping length k for L.

Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

The string $s = a^k b^k$ should be in the language.

34

Example

- L = $\{a^n b^n: n \geq 0\}$ is not regular.

Proof:

$|s| \geq k$: $|s| = 2k$ is greater than k

35

Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Consider all possible splittings of $a^k b^k$ in the form xyz with:

- $|xy| \leq k$ and
- $|y| \geq 1$

Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Consider all possible splittings of $a^k b^k$ in the form xyz with:

- $|xy| \leq k$ and
- $|y| \geq 1$

$$\overbrace{a\ a\ a\ a\ a\ a\ a\ a\ a}^{k}\ \overbrace{b\ b\ b\ b\ b\ b\ b\ b\ b}^{k}$$

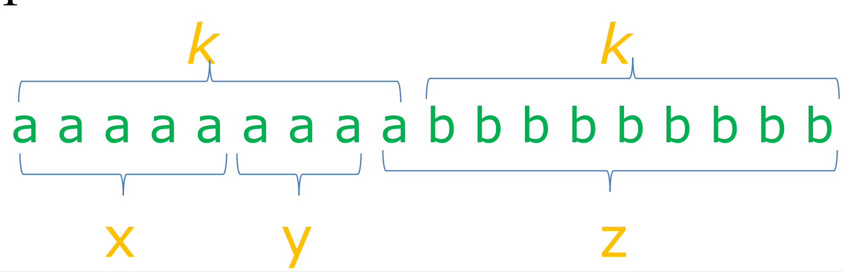
Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Consider all possible splittings of $a^k b^k$ in the form xyz with:

- $|xy| \leq k$ and
- $|y| \geq 1$



Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Consider all possible splittings of $a^k b^k$ in the form xyz with ($|xy| \leq k$ and $|y| \geq 1$)

- $y = a^m$, for $1 \leq m \leq k$

$$\overbrace{a\ a\ a\ a\ a\ a\ a\ a\ a}^{k}\ \overbrace{b\ b\ b\ b\ b\ b\ b\ b\ b}^{k}$$

$$\underbrace{a\ a\ a\ a\ a}_{x}\ \underbrace{a\ a\ a}_{y}\ \underbrace{a\ b\ b\ b\ b\ b\ b\ b\ b\ b}_{z}$$

39

Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Consider all possible splittings of $a^k b^k$ in the form xyz with ($|xy| \leq k$ and $|y| \geq 1$)

- $y = a^m$, for $1 \leq m \leq k$
- for i =2, $xy^2z = a^{k+m}b^k$ is not in L!

$$\underbrace{a\ a\ a\ a\ a}_{x}\ \underbrace{a\ a\ a}_{y}\ \underbrace{a\ a\ a}_{y}\ \underbrace{a\ b\ b\ b\ b\ b\ b\ b\ b\ b}_{z}$$

Example

- L = $\{a^n b^n : n \geq 0\}$ is not regular.

Proof:

Consider all possible splittings of $a^k b^k$ in the form xyz with ($|xy| \leq k$ and $|y| \geq 1$)

- xy^2z is not in L!

CONTRADICTION!

$$\underbrace{a\;a\;a\;a\;a}_{x}\;\underbrace{a\;a\;a}_{y}\;\underbrace{a\;a\;a}_{y}\;\underbrace{a\;b\;b\;b\;b\;b\;b\;b\;b\;b}_{z}$$

How to use the pumping lemma

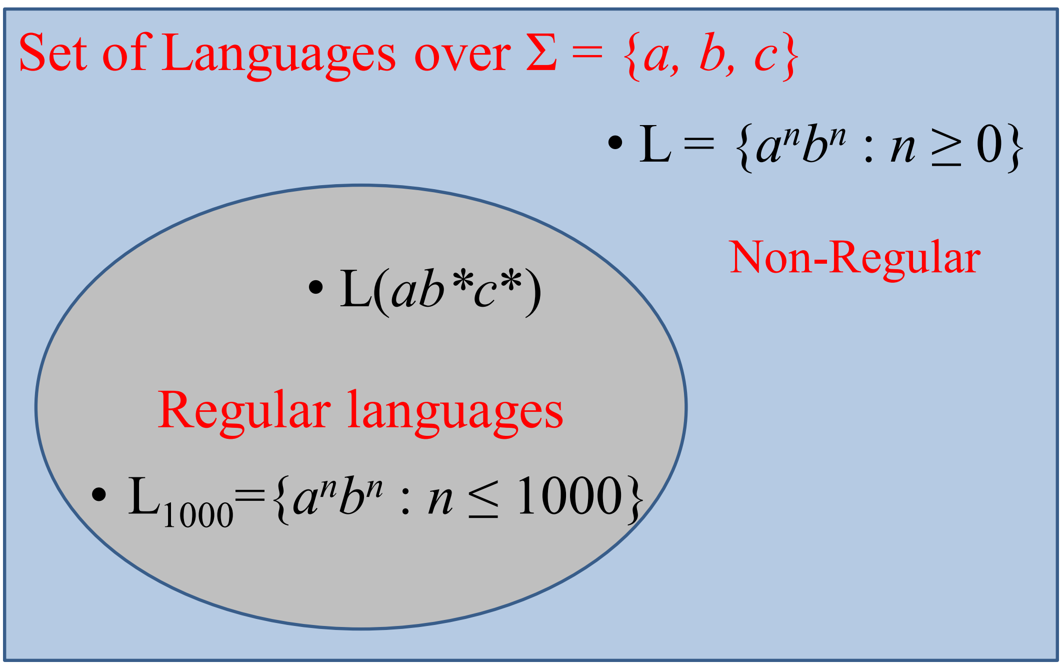
- The pumping lemma mentions that if L is a regular language then it can be pumped.
- The contrapositive is true: If L cannot be pumped then it shouldn't be regular!

How not to use the pumping lemma

- The pumping lemma mentions that if L is a regular language then it can be pumped.
- The converse is not true: If a language can be pumped this doesn't mean that it is regular!

43

Language universe



44

Summary

- Every language of finite size has to be regular
  - We can easily construct an NFA that accepts every string in the language.
- Therefore, every non-regular language has to be of infinite size.

Summary

To prove that an infinite language L is not regular:

1. Assume the opposite: L is regular.
2. The pumping lemma should hold for L.
3. Use the pumping lemma to obtain a contradiction.
4. Therefore, L is not regular.