Properties of regular languages - The pumping lemma

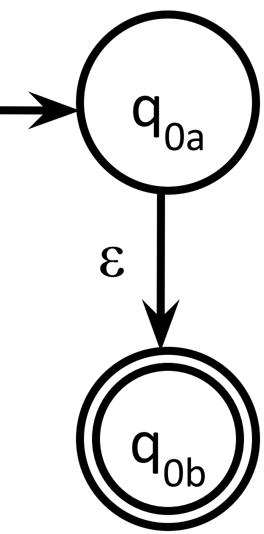
CHAPTER-4 (PART 1)

Regular Languages

- Regular languages are the languages which are accepted by a Finite Automaton.
- Not all languages are regular

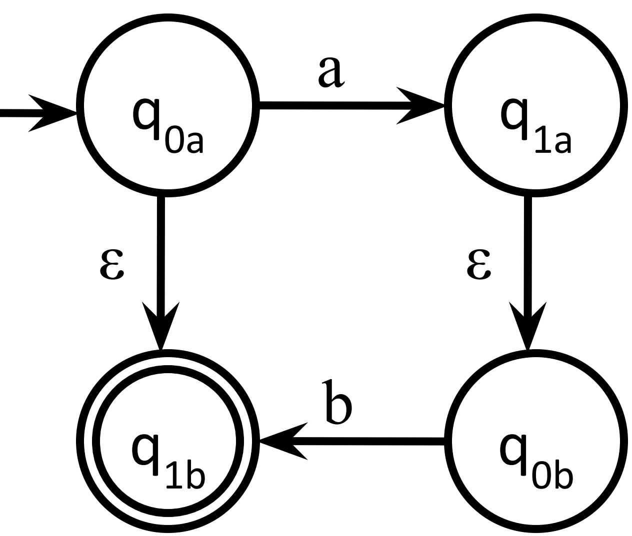
Non-Regular Languages

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



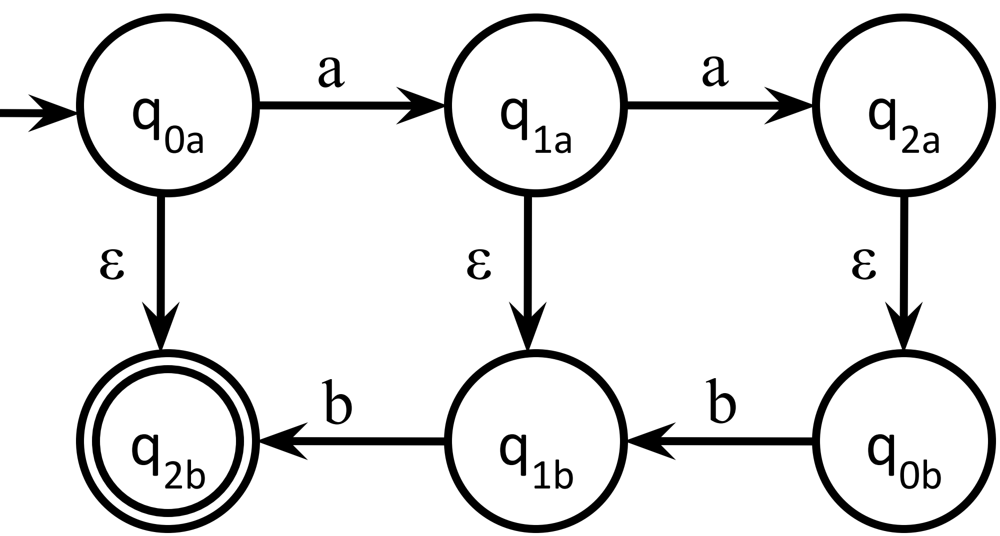
Non-Regular Languages

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



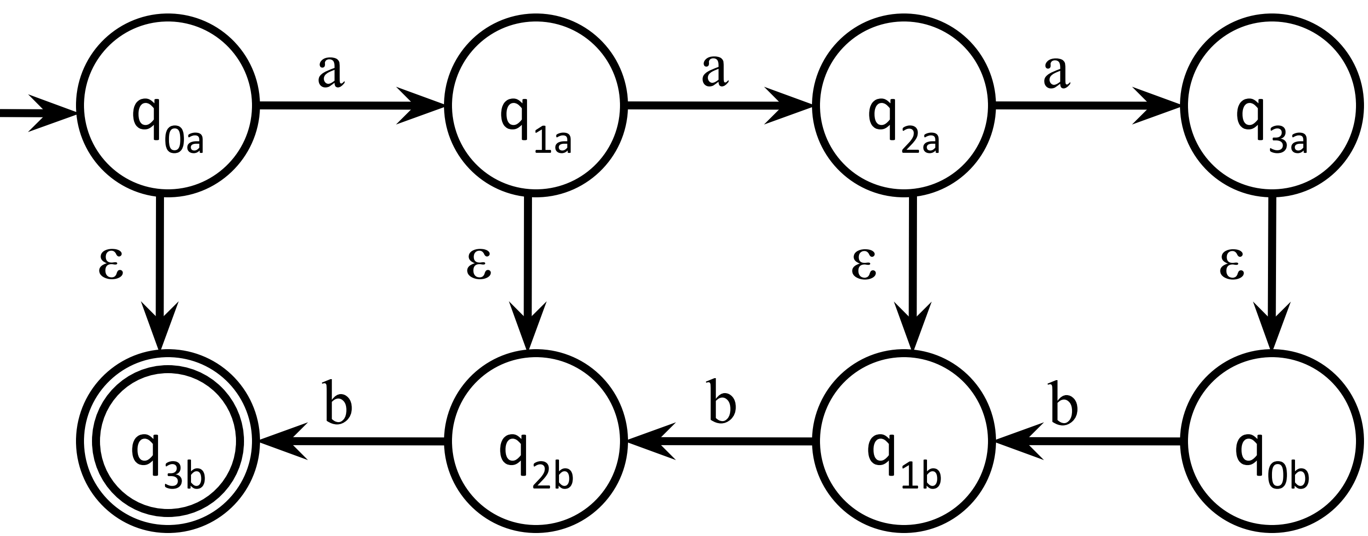
Non-Regular Languages

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



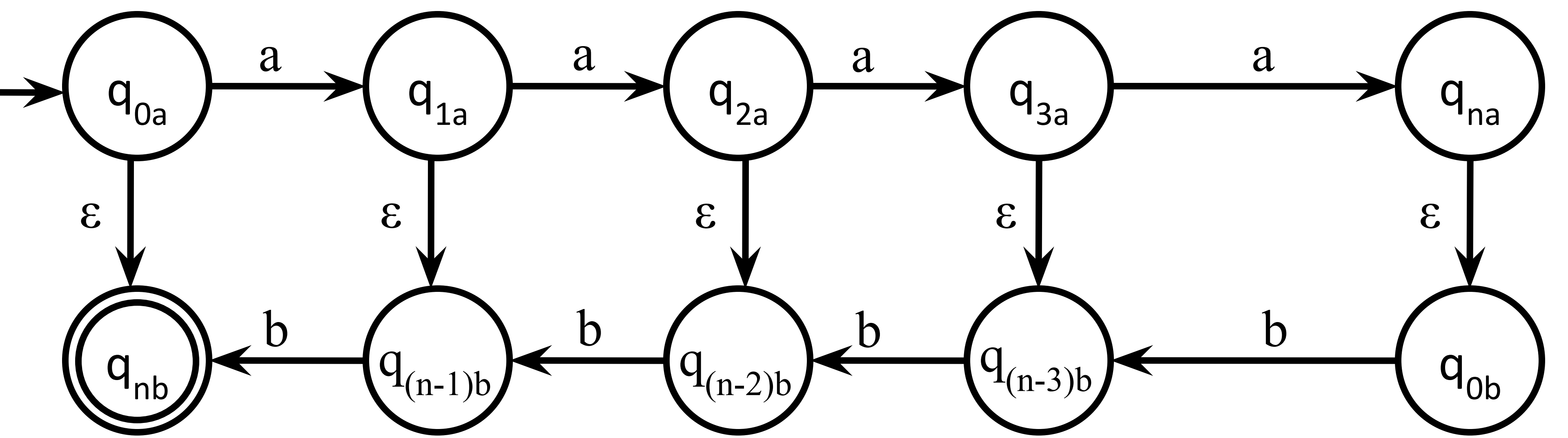
Non-Regular Languages

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



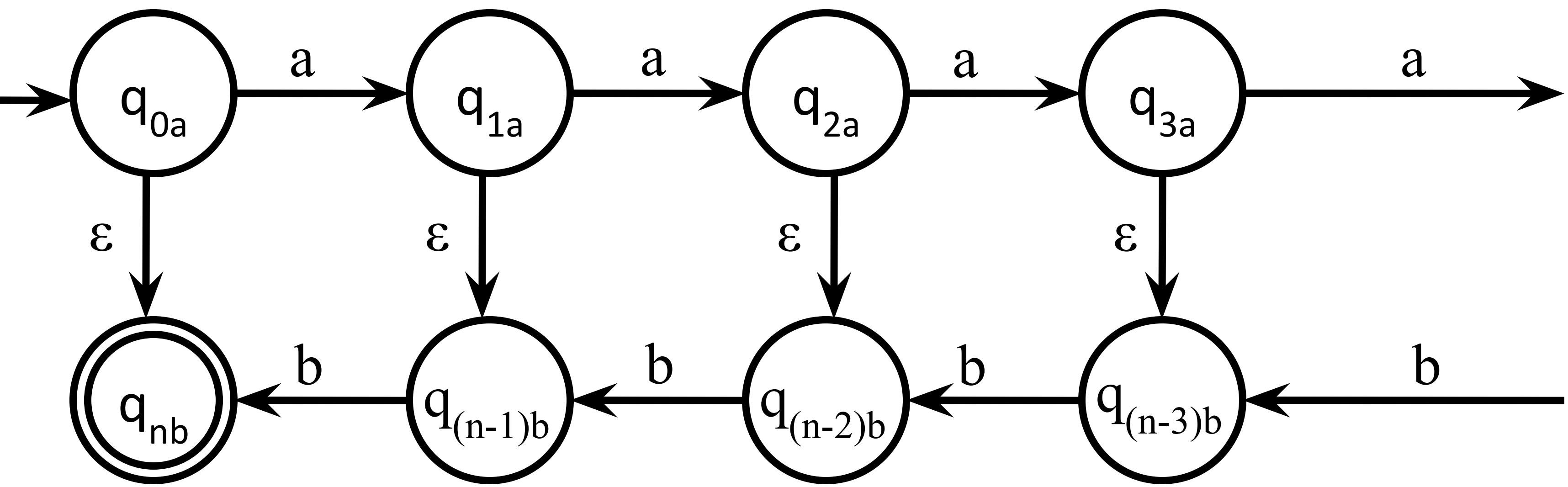
Non-Regular Languages

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



Non-Regular Languages

- However $L = \{a^n b^n : n \geq 0\} = \cup_{n \geq 0} L_n$ doesn't seem to be a regular language at all!
- We need an infinite number of states to build this automaton!



Non-Regular Languages

- However $L = \{a^n b^n : n \geq 0\} = \bigcup_{n \geq 0} L_n$ doesn't seem to be a regular language at all!
- We need an infinite number of states to build this automaton!
- (Observe that you cannot use the fact that regular languages are closed under union because we have an infinite union)

Is this a proof?

NO! In fact consider:

L’ = {s : s contains equal number of a and b}

L’’ = {s : s contains equal number of ab and ba}

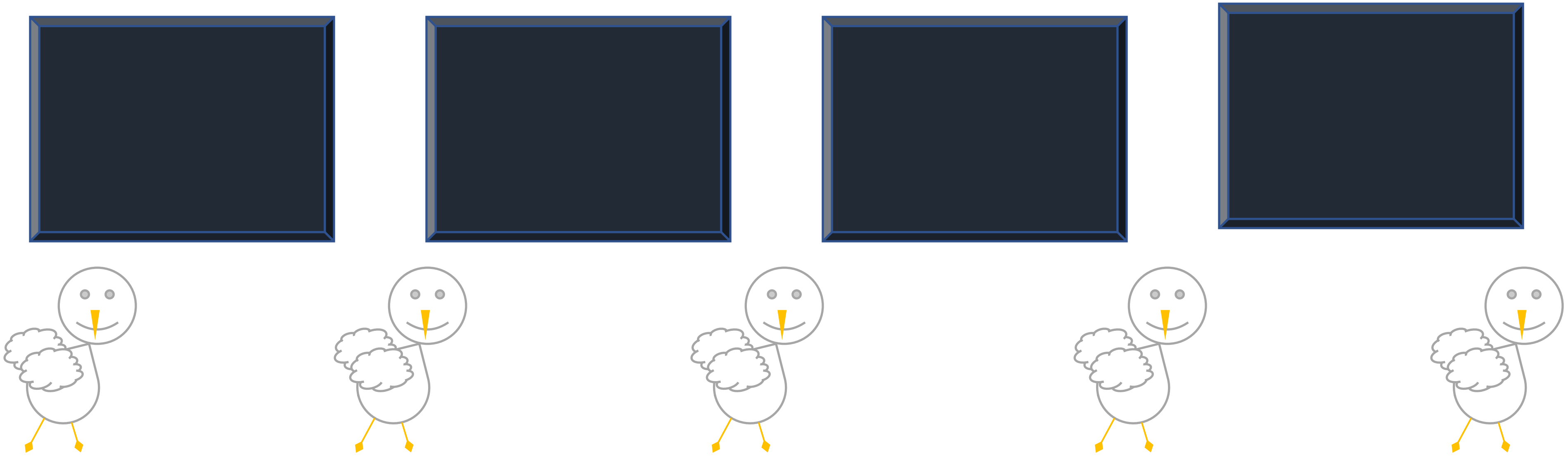
L’ is indeed not regular but L’’ is regular!

<u>WE NEED A MATHEMATICAL PROOF!!!</u>

A proof that there is no FA that accepts L or L’.

Pigeonhole Principle

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

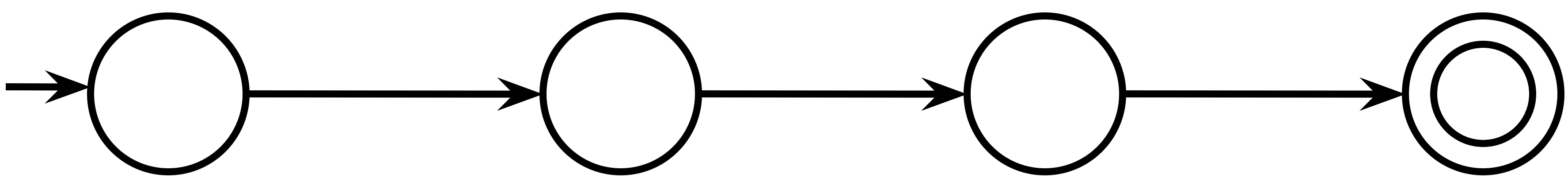


Pigeonhole Principle

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

PP and Finite Automata

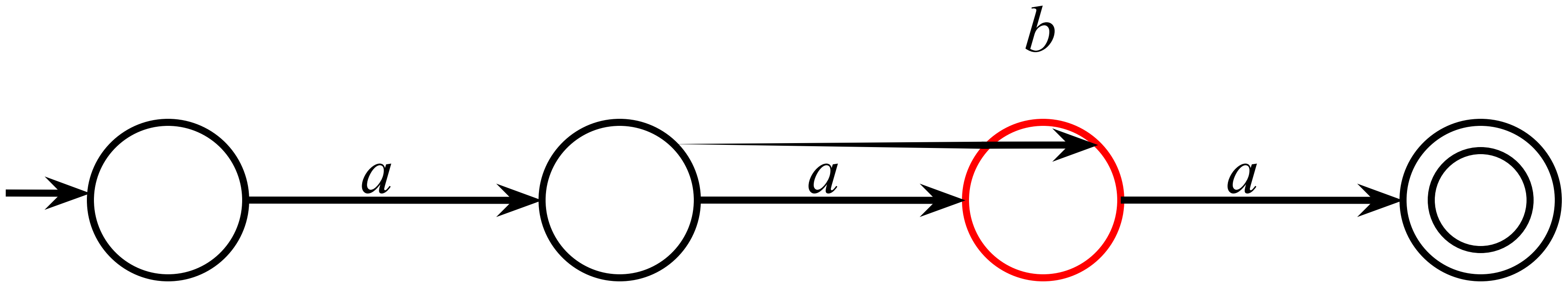
- If an automaton with n states accepts a string with length m (m≥n) then for every accepting path there should be at least one repeating state.



$$s = aaba$$

PP and Finite Automata

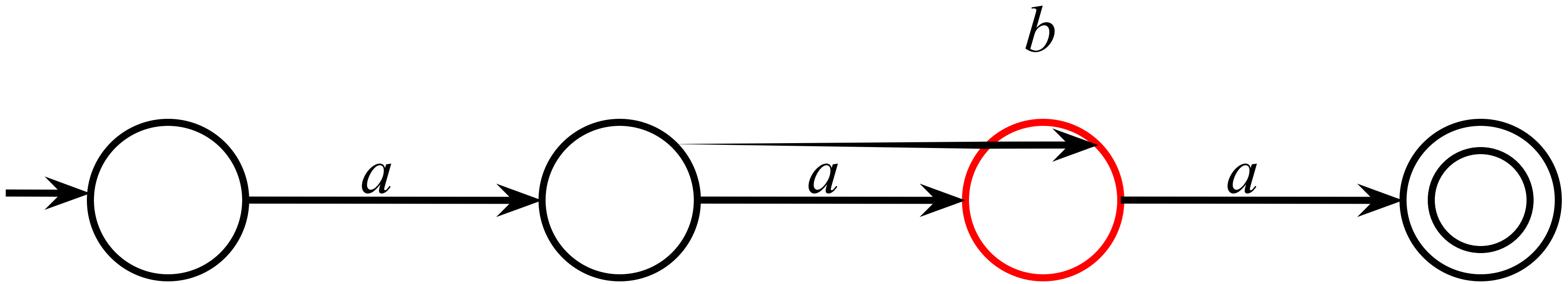
- If an automaton with n states accepts a string with length m (m≥n) then for every accepting path there should be at least one repeating state.



$$s = aaba$$

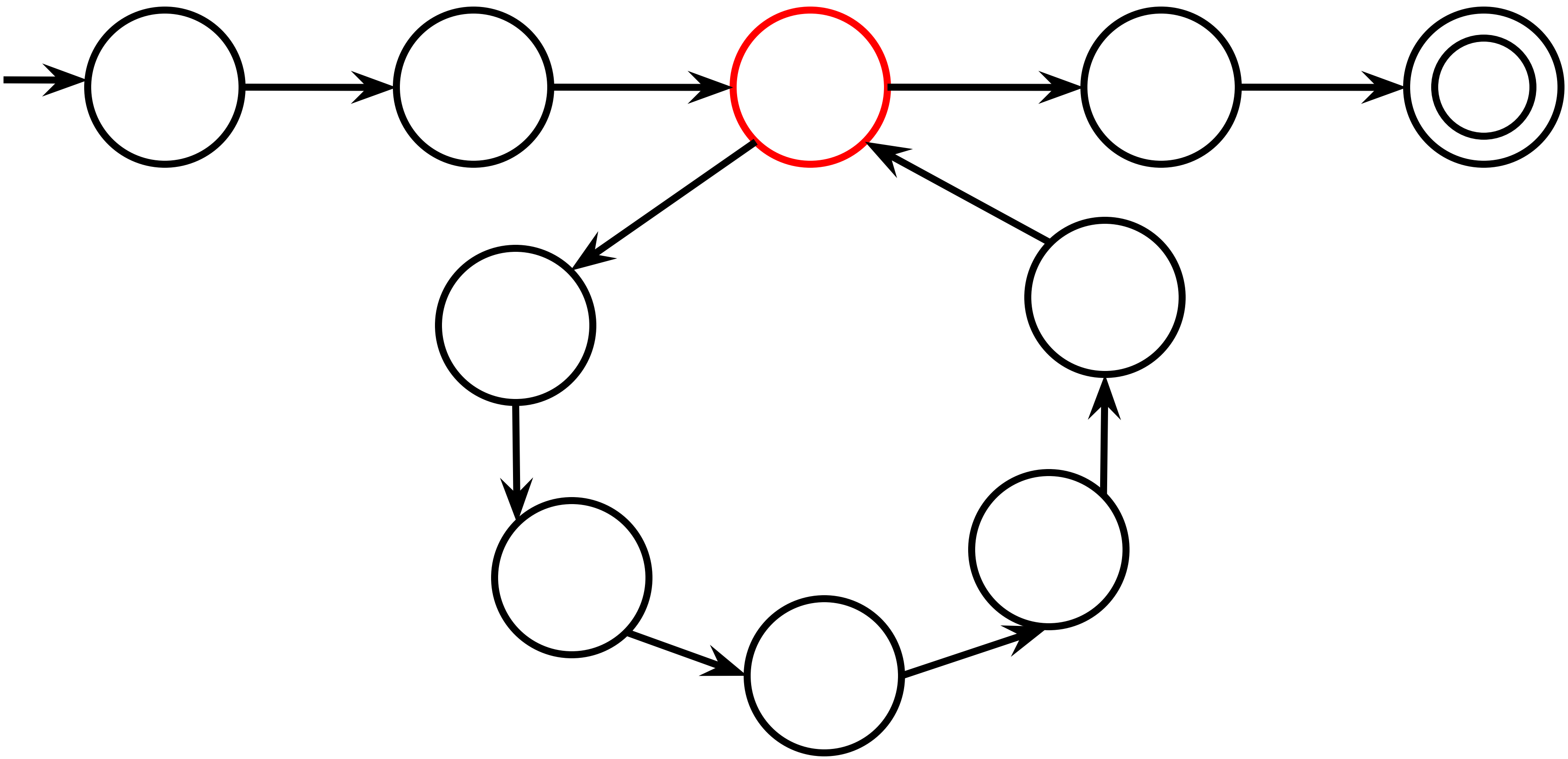
PP and Finite Automata

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

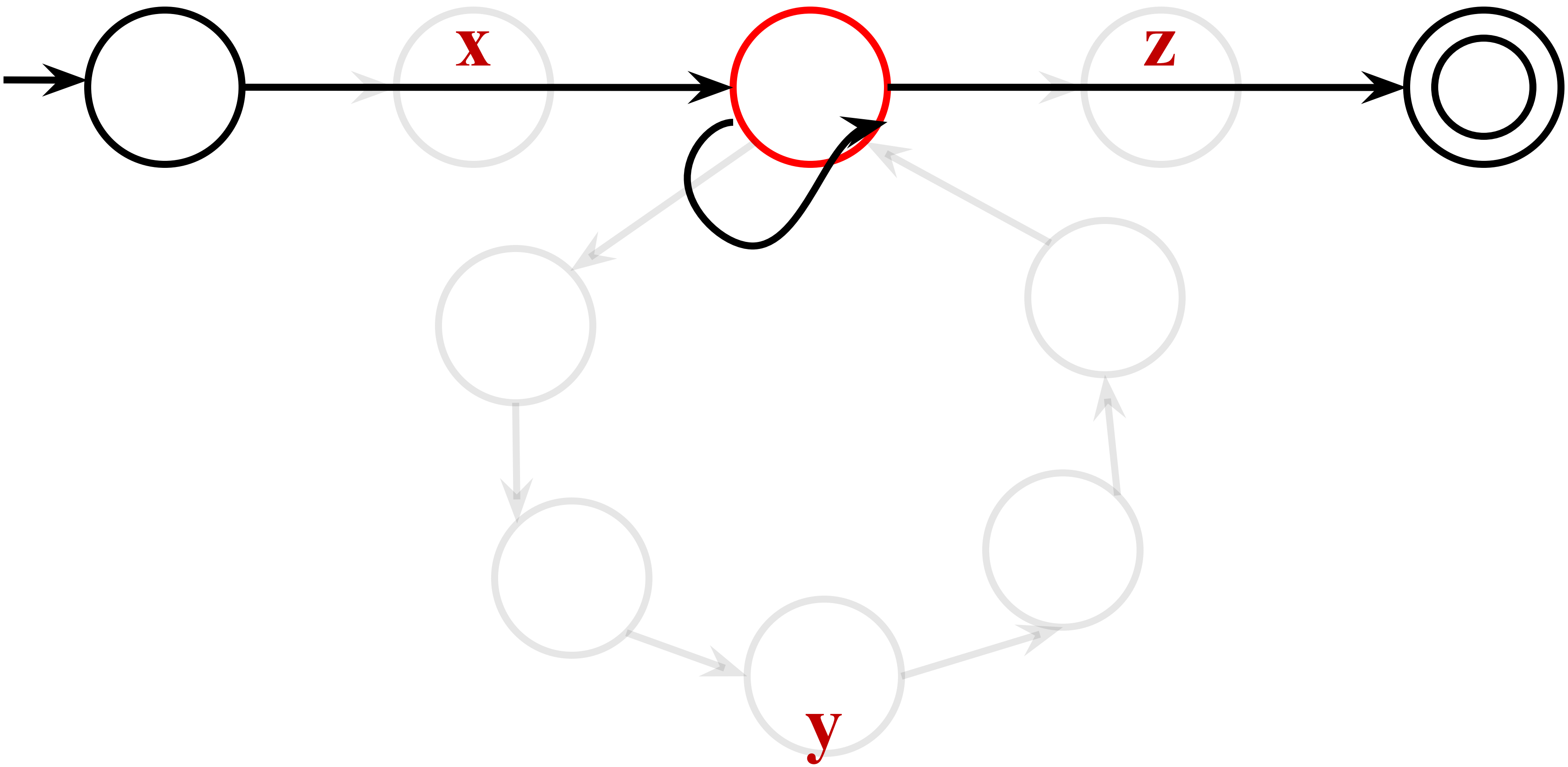


*Any string of the form aab*a should be accepted!*

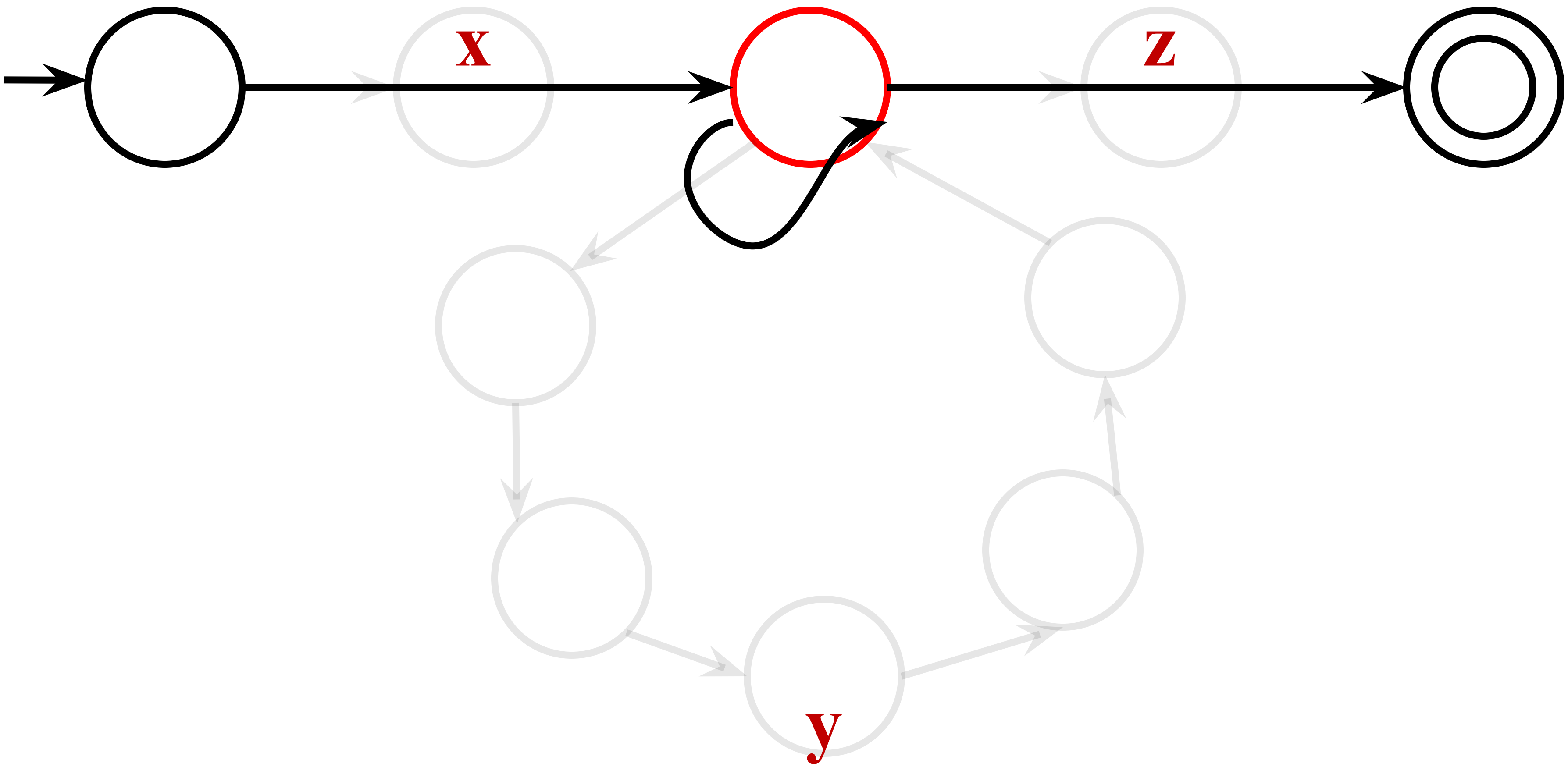
PP and FA continued



PP and FA continued

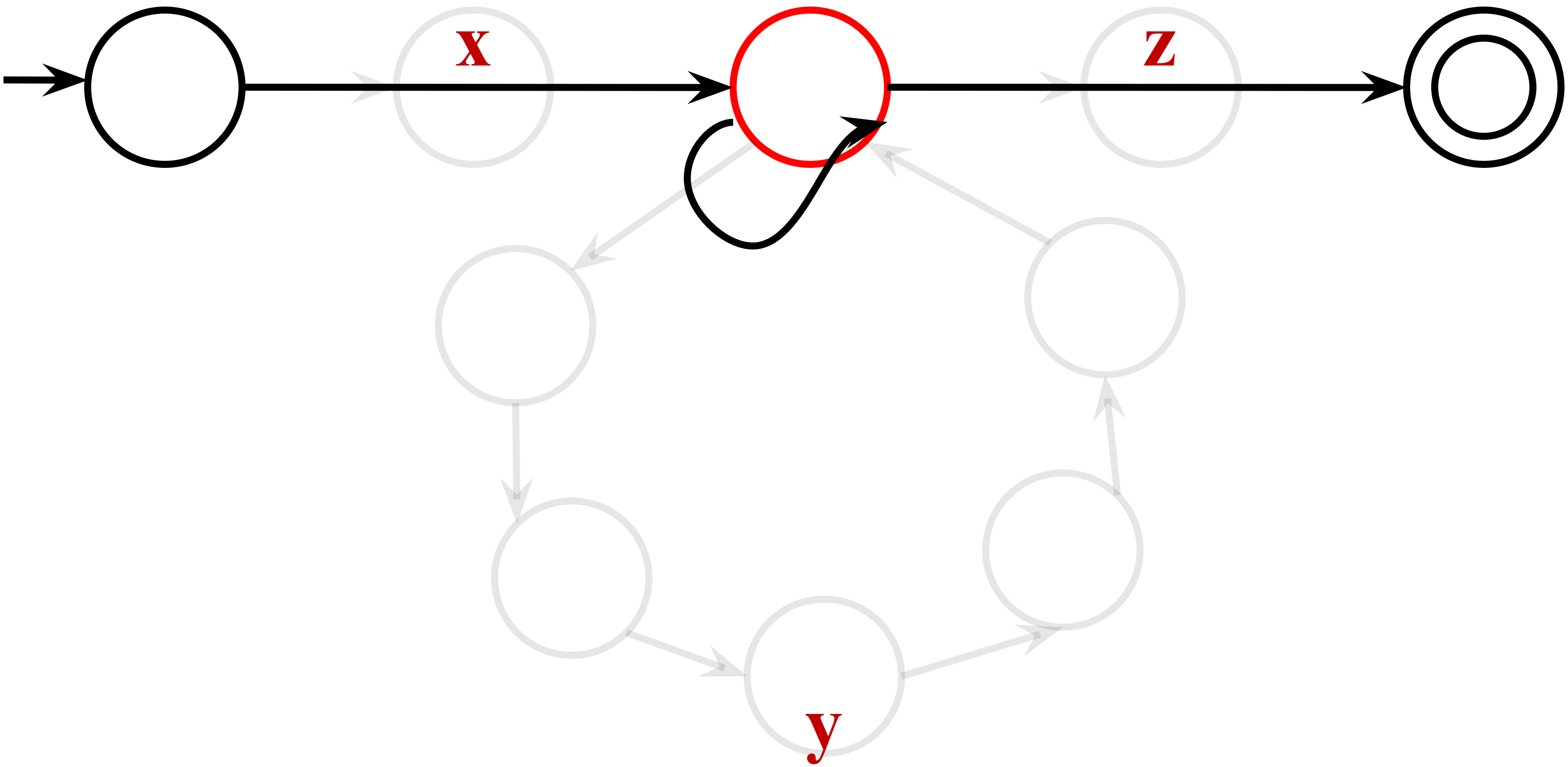


PP and FA continued



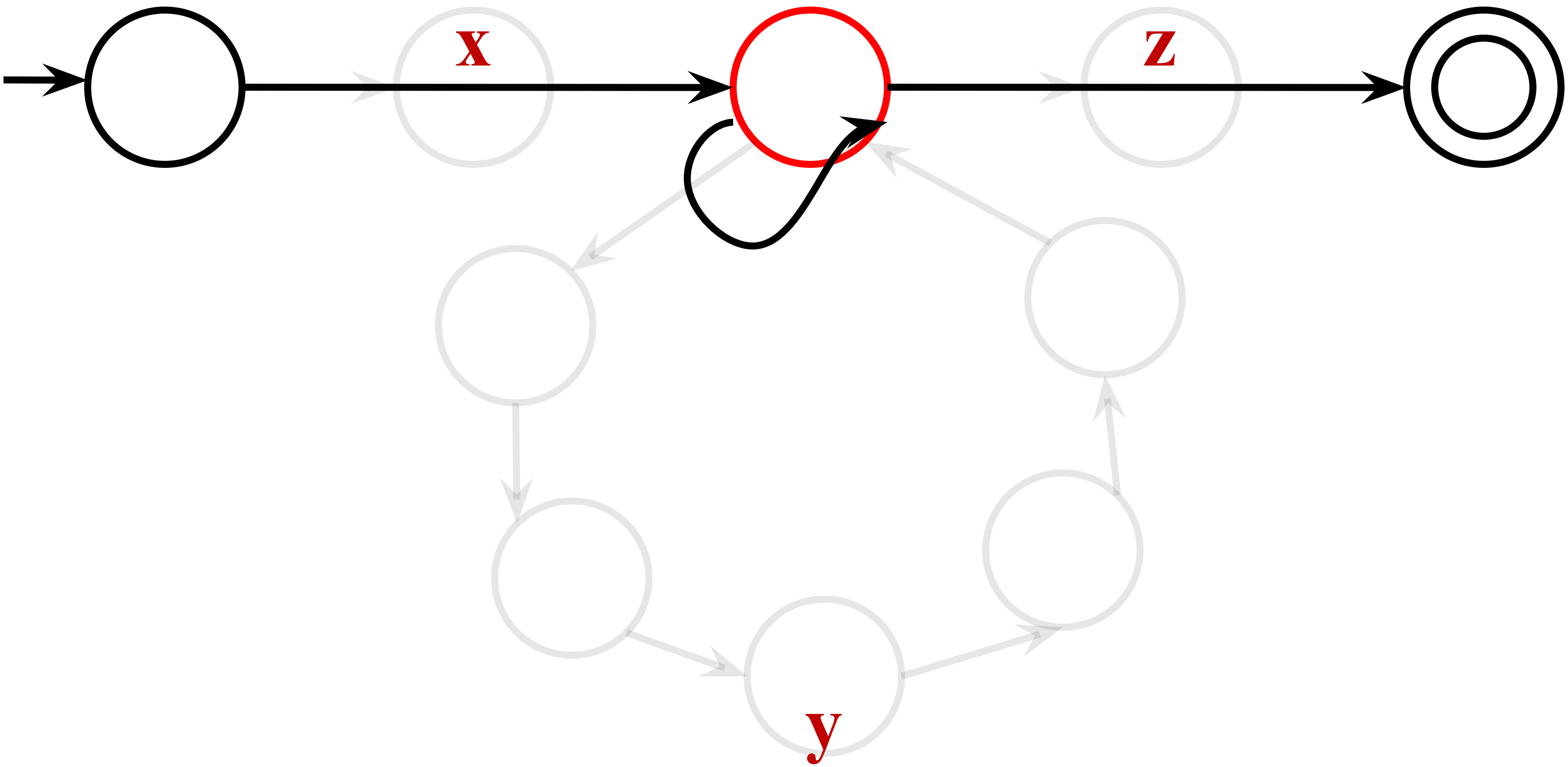
xyz is accepted

PP and FA continued



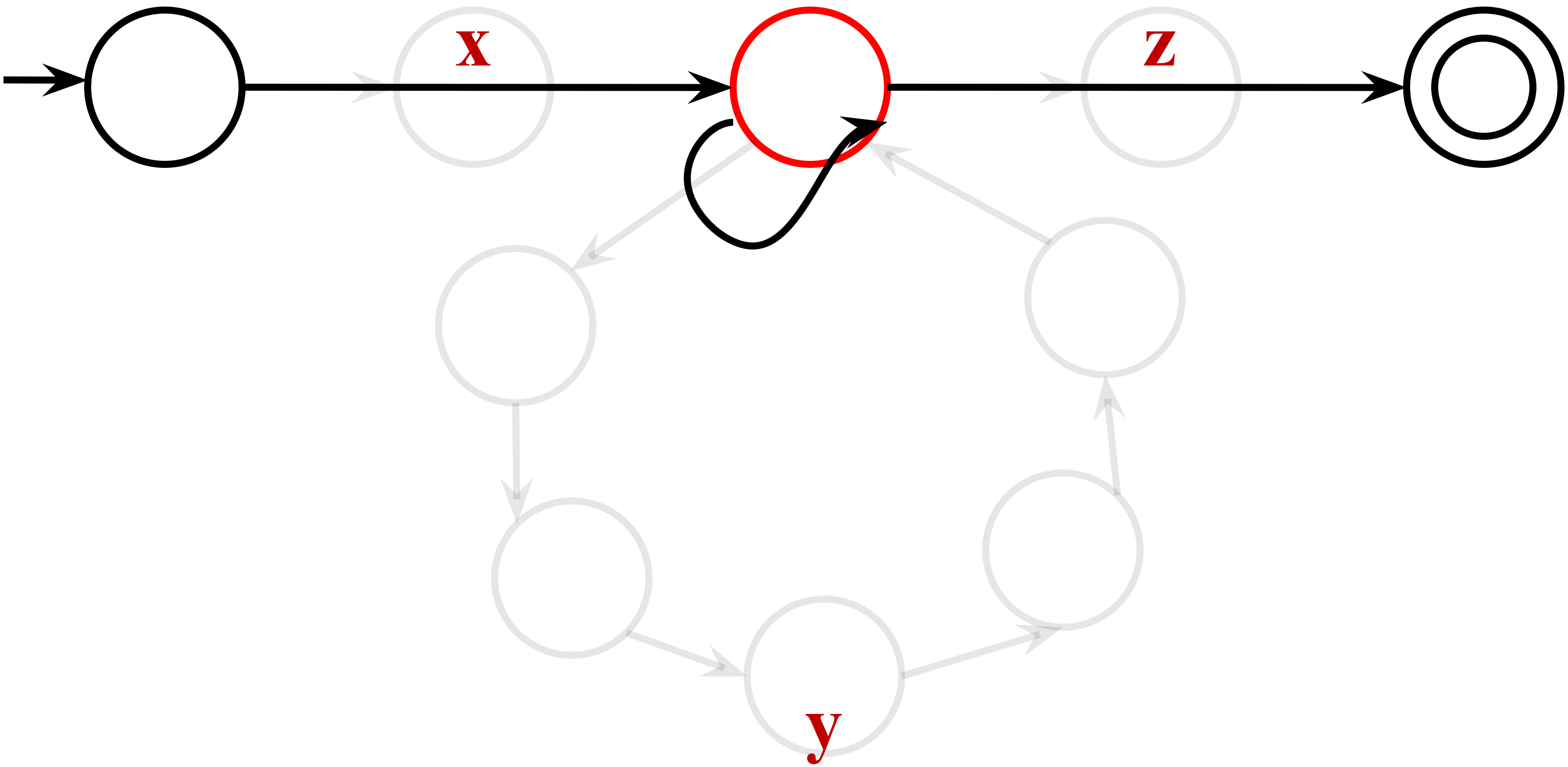
xyyz is accepted

PP and FA continued



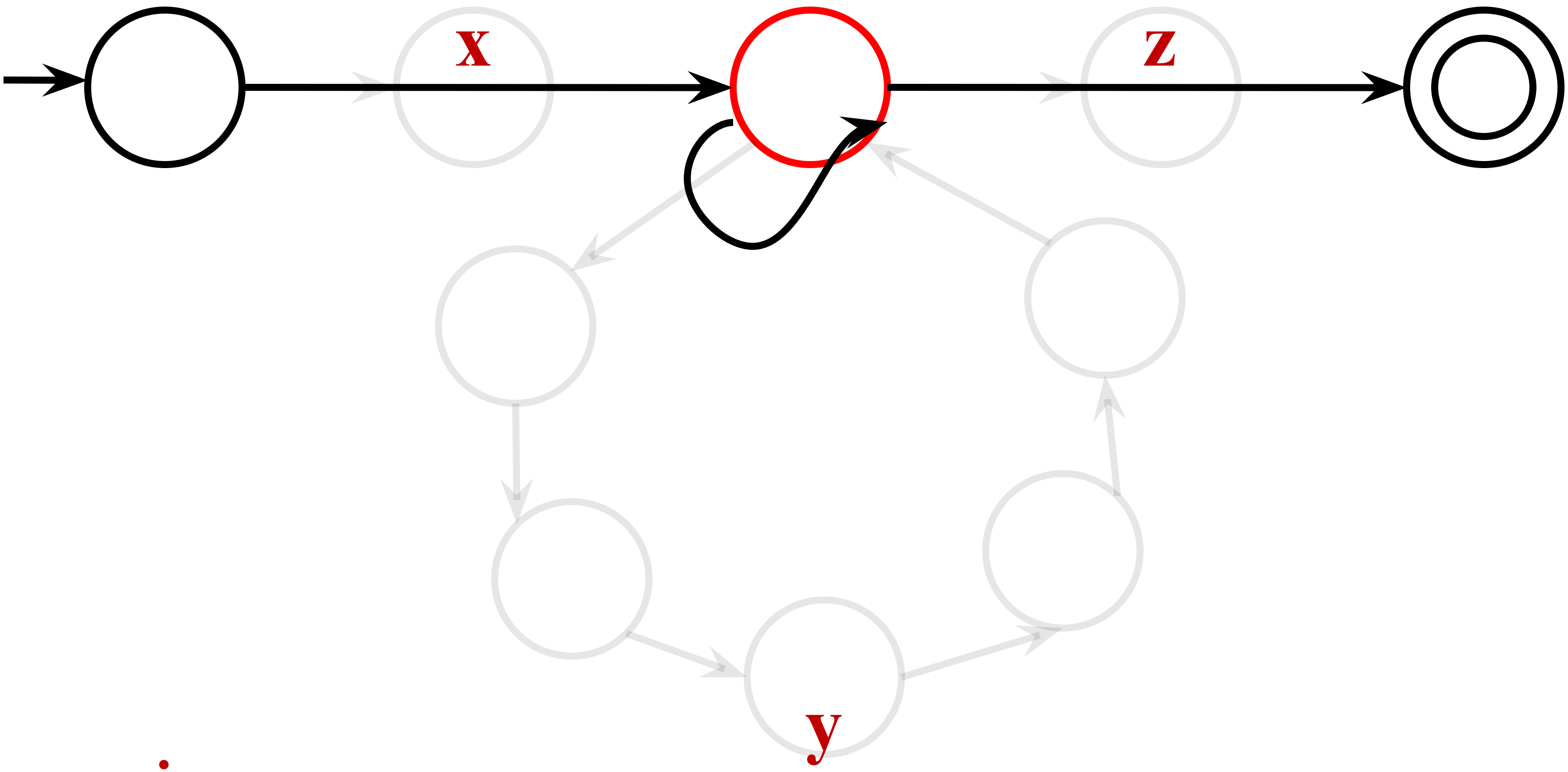
xyyyz is accepted

PP and FA continued



xz is accepted

PP and FA continued



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

The pumping lemma

For every <u>infinite</u> *regular* language L

there exists a pumping length $n > 0$ such that

for any string s in L with length $|s| \geq n$

we can write $s = xyz$

with $|xy| \leq n$ 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 n to be M's number of states.
- L is infinite, there exists a string s with length greater than n.
- The number of states is n, the accepting path for s is of length at most n.
- The string is of length at least n, 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 n states the first repetition should take place ($|y| \geq 1$) in at most n transitions ($|xy| \leq n$).
- 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.

  (*What if L is finite?*)

Prove that L is not regular

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

Prove that L is not regular

- Given is an infinite language L
- If L is regular
- Pumping lemma holds:
  - *There exists* a pumping length n such that
  - *for all* <u>proper</u> strings s in L
  - *there is* a splitting of s in x,y,z (with the <u>desired properties</u>) such that
  - *for all* i xy^iz is in L.

Prove that L is not regular

- Given is an infinite language L
- If L is regular
- Pumping lemma holds:
  - *There exists* a pumping length n such that
  - *for all* proper strings s in L
  - *there is* a splitting of s in x,y,z (with the desired properties) such that
  - *for all* i xy^iz is in L.

$$\boxed{|s| \geq n}$$

$$\boxed{|y| \geq 1 \text{ and } |xy| \leq n}$$

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 n
  - *there exists* a proper string s in L such that
  - *for every possible* splitting of s in x,y,z (with the desired properties)
  - *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 n.
  - *Specify* a proper string s in L.
  - Show that *for every possible* splitting of s in x,y,z (with the desired properties),
  - *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 n.
  - *Specify* a proper string s in L.
  - Show that *for every possible* splitting of s in x,y,z (with the desired properties),
  - *there is* an i for which xy^iz is *not* in L.
- Contradiction!!!

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 n.
  - *Specify* a <u>proper</u> string s in L.
  - Show that *for every possible* splitting of s in x,y,z (with the <u>desired properties</u>),
  - *there is* an i for which xy^iz is *not* in L.
- Contradiction!!! **L is 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.

Example

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

Proof:

s is proper: $|s| = 2k$ is greater than 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 the desired properties:

- $|xy| \leq k$ 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 the desired properties:

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

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

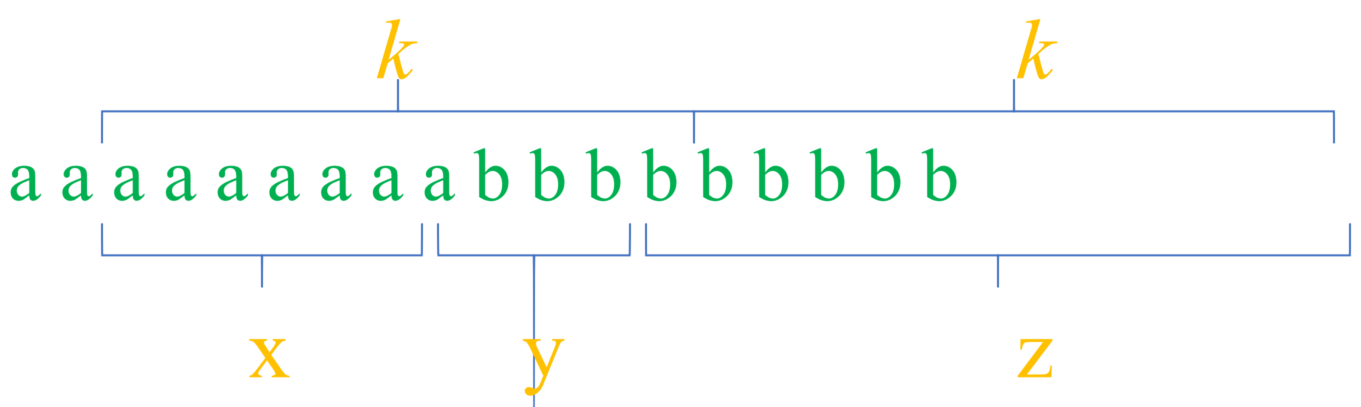
Example

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

Proof:

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

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



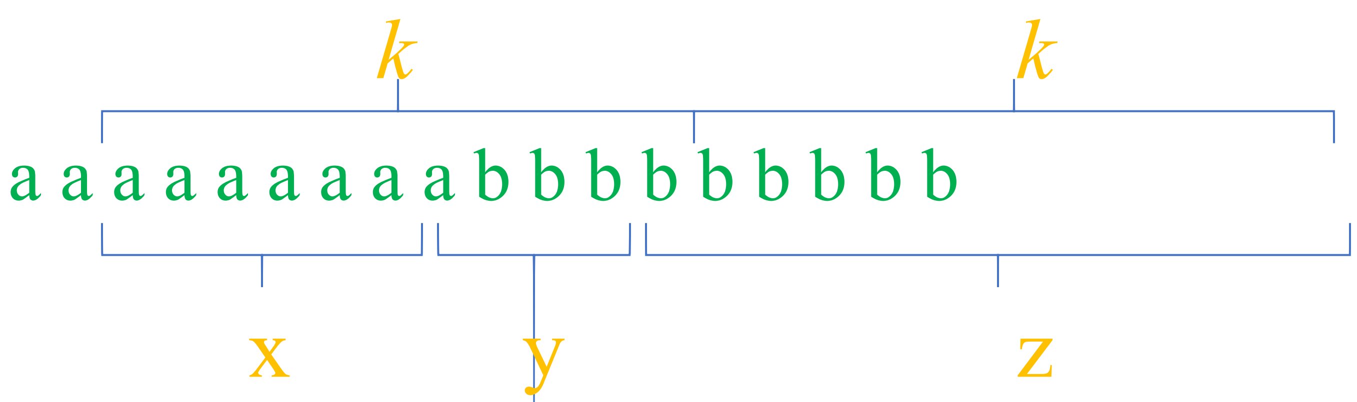
Example

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

Proof:

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

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



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 the desired properties

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

a a a a a a a a a a a a b b b b b b b b b

x y y 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 the desired properties

- xy^2z is not in L!

CONTRADICTION!

a a a a a a a a a a a a b b b b b b b b b

x y y z

Try it yourself

- Show that the following language is not a regular language:

$$L_p' = \{ab^jc^j \mid j \geq 0\}$$

- Members: a, abc, abbcc, abbbccc, …

Try it yourself

- Show that the following language is not a regular language:

$$L_p' = \{ab^jc^j \mid j \geq 0\}$$

- Negation of the pumping lemma (reminder):
    - *Fix* an *arbitrary* pumping length n.
    - *Specify* a proper string s in L.
    - Show that *for every possible* splitting of s in x,y,z (with the desired properties),
    - *there is* an i for which xy^iz is *not* in L.

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!

Show that L cannot be pumped to prove that L is not 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!

Do not try to show that L can be pumped in order to prove that L is regular.

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

1. L_p is not regular.

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

1. L_p is not regular.

Proof:

- $L_p' = \{ab^j c^j : j \geq 0\}$ is not regular.

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

1. L_p is not regular.

Proof:

- $L_p' = \{ab^j c^j : j \geq 0\}$ is not regular.
- $L_p' = L_p \cap L(ab^*c^*)$.

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

1. L_p is not regular.

Proof:

- $L_p' = \{ab^j c^j : j \geq 0\}$ is not regular.
- $L_p' = L_p \cap L(ab^*c^*)$.
- If L_p was regular then L_p' would be regular too.

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Proof:

Prove that: *There is* a pumping length n such that *for any* <u>proper</u> s in L_p *there exists* a splitting xyz of s with the <u>desired properties</u> such that *for all* i xy^iz is in L_p.

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Proof:

Choose pumping length n = 2

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Proof:

Choose pumping length n = 2

- For i=0, j=0, k ≥2 : $s = c^k$

  Set $x=\varepsilon$, $y = c$, $z = c^{k-1}$

  For every i≥0, xy^iz is in L_p

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Proof:

Choose pumping length $n = 2$

- For $i=0$, $j\geq 1$, $k \geq 0$: $s = b^j c^k$

    Set $x=\varepsilon$, $y = b$, $z = b^{j-1}c^k$

        For every $i\geq 0$, xy^iz is in L_p

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Proof:

Choose pumping length n = 2

- For i=1 and any j ≥ 1: $s = ab^jc^j$

Set $x=\varepsilon$, $y = a$, $z = b^jc^j$

For every $i\geq 0$, xy^iz is in L_p

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Proof:

Choose pumping length n = 2

- For i=2 and any j, k ≥ 0: $s = aab^jc^k$

Set $x=\varepsilon$, $y = aa$, $z = b^jc^k$

For every $i \geq 0$, xy^iz is in L_p

Pumping Lemma - Converse

- For example consider the language

$$L_p = \{a^i b^j c^k : i,j,k \geq 0, \text{ if } i=1 \text{ then } j=k\}.$$

2. L_p can be pumped

Proof:

Choose pumping length n = 2

- For $i \geq 3$ and any $j, k \geq 0$: $s = aaa^{i-2}b^jc^k$

Set $x=\varepsilon$, $y = a$, $z\ aa^{i-2}b^jc^k$

For every $i \geq 0$, xy^iz is in L_p

The Universe