

Non-regular languages

Non-regular languages

$$\{a^n b^n : n \geq 0\}$$

$$\{vv^R : v \in \{a,b\}^*\}$$

Regular languages

$$a^*b$$

$$b^*c + a$$

$$b + c(a + b)^*$$

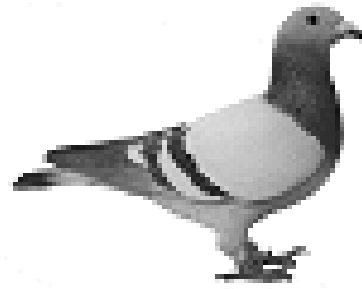
etc...

How can we prove that a language L is not regular?

Prove that there is no DFA that accepts L

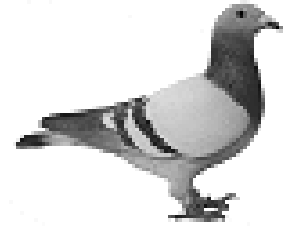
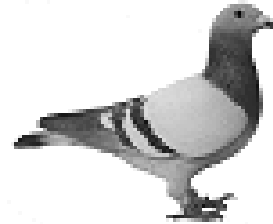
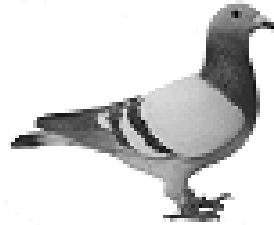
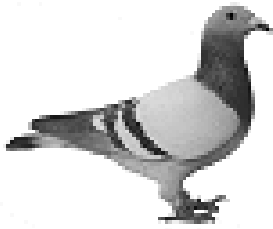
Problem: this is not easy to prove

Solution: the Pumping Lemma !!!

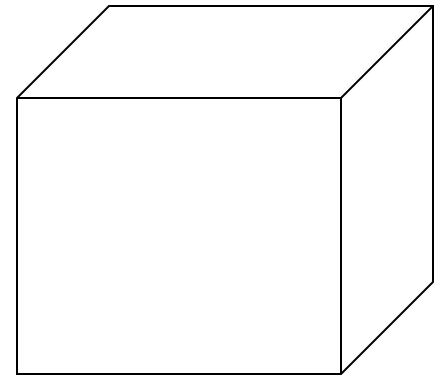
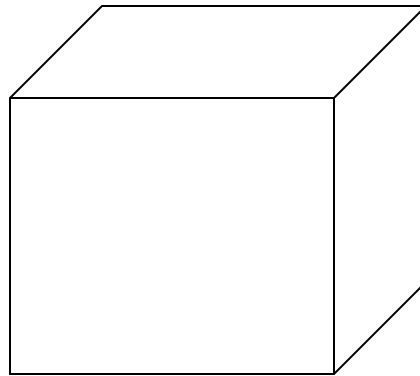
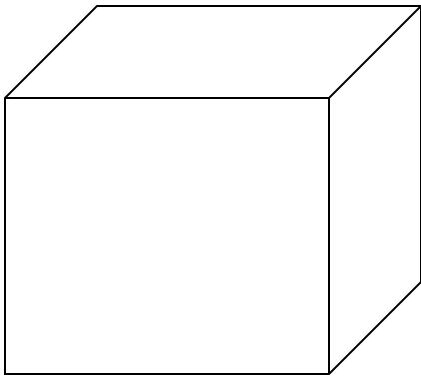


The Pigeonhole Principle

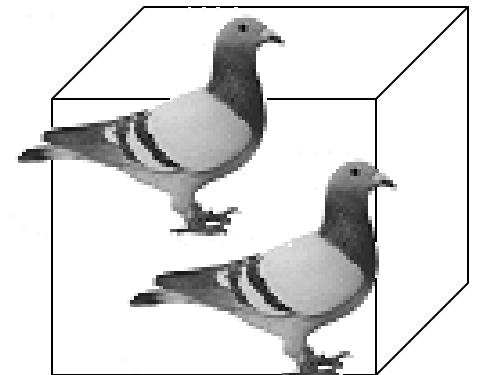
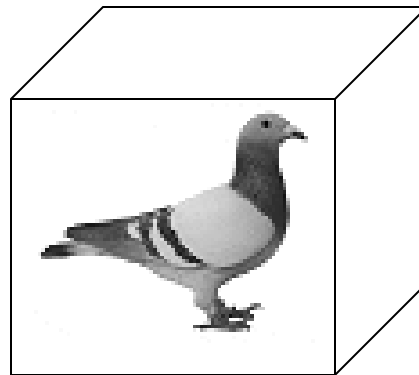
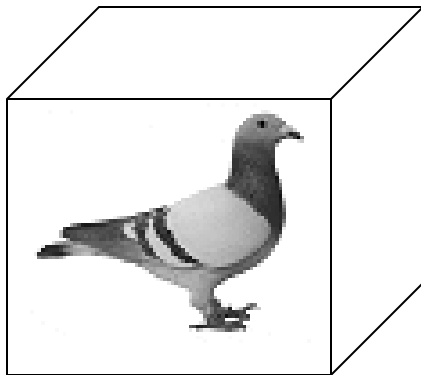
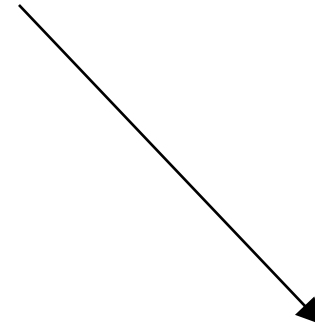
4 pigeons



3 pigeonholes



A pigeonhole must
contain at least two pigeons



n pigeons

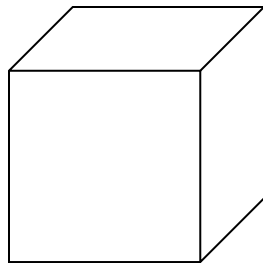
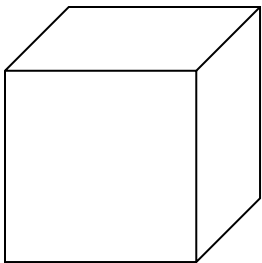


.....

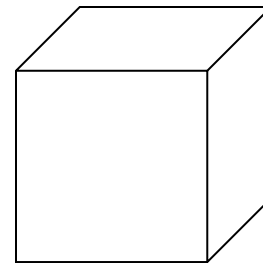


m pigeonholes

$n > m$



.....



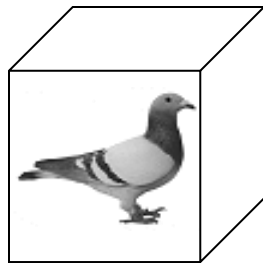
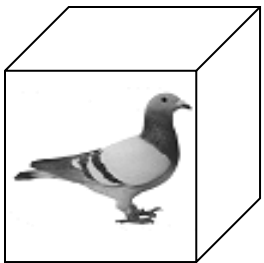
The Pigeonhole Principle

n pigeons

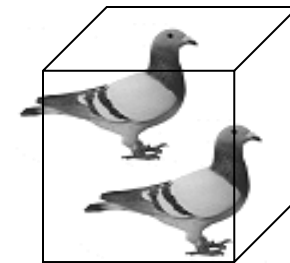
m pigeonholes

$$n > m$$

There is a pigeonhole
with at least 2 pigeons



.....

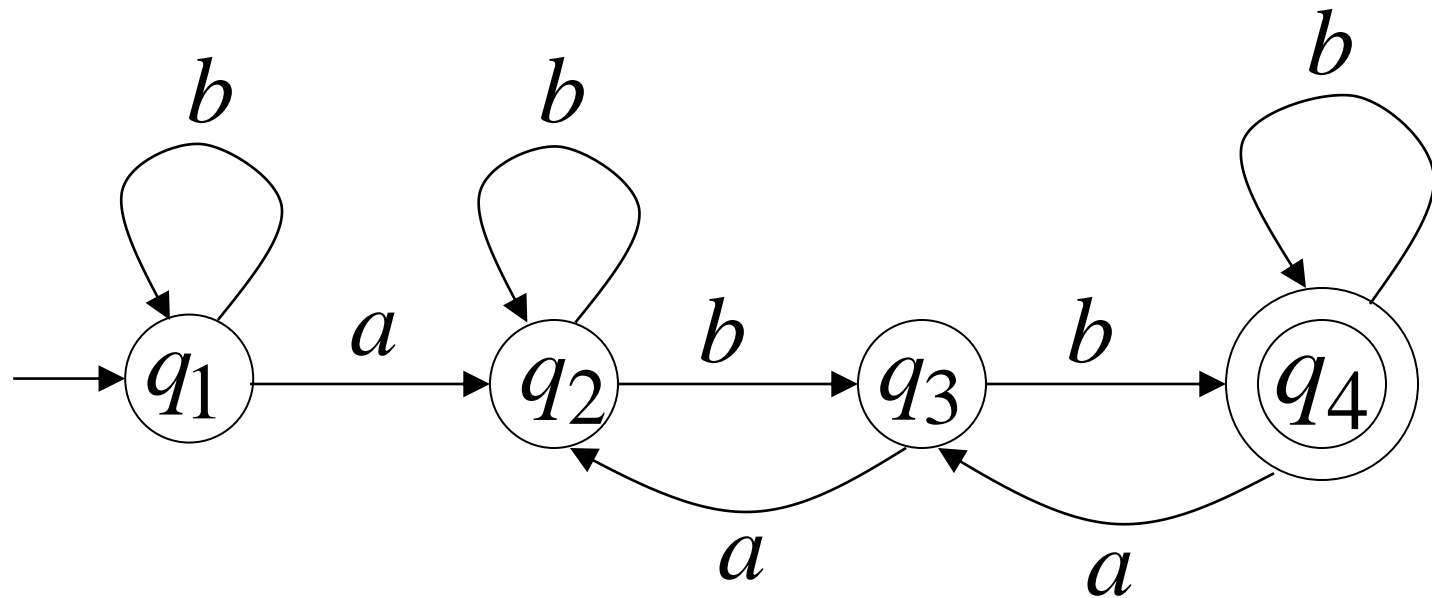


The Pigeonhole Principle

and

DFAs

DFA with 4 states



In walks of strings:

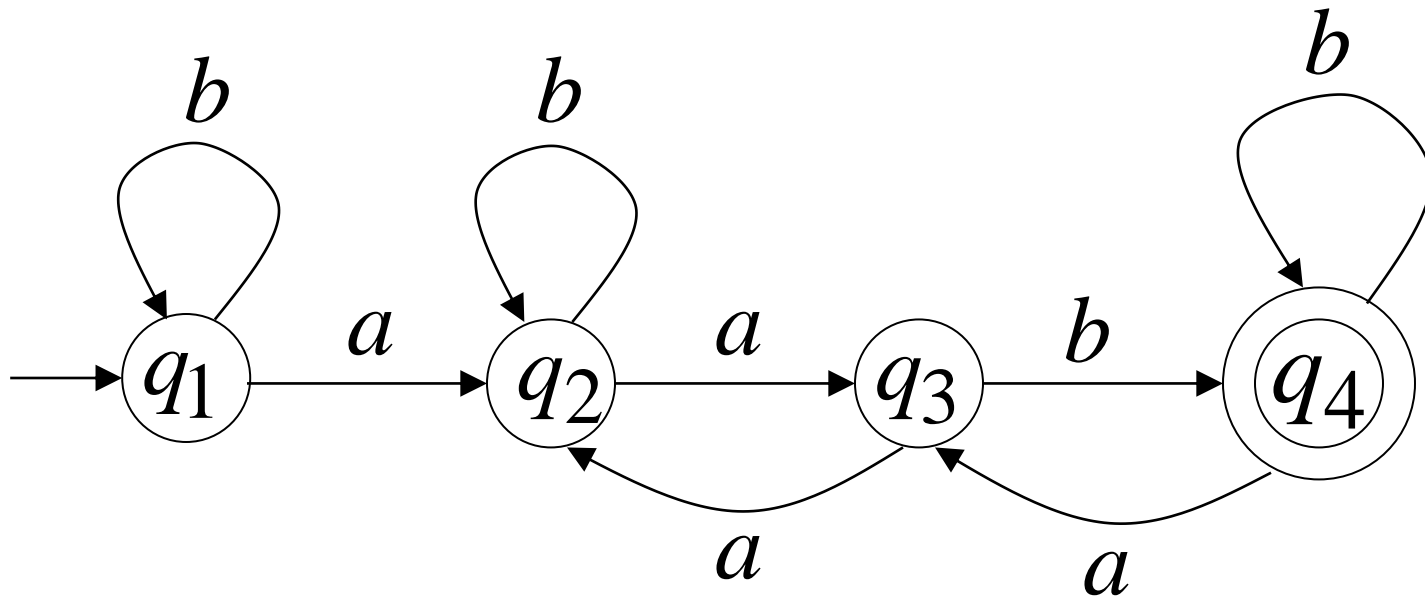
a

aa

aab

no state

is repeated



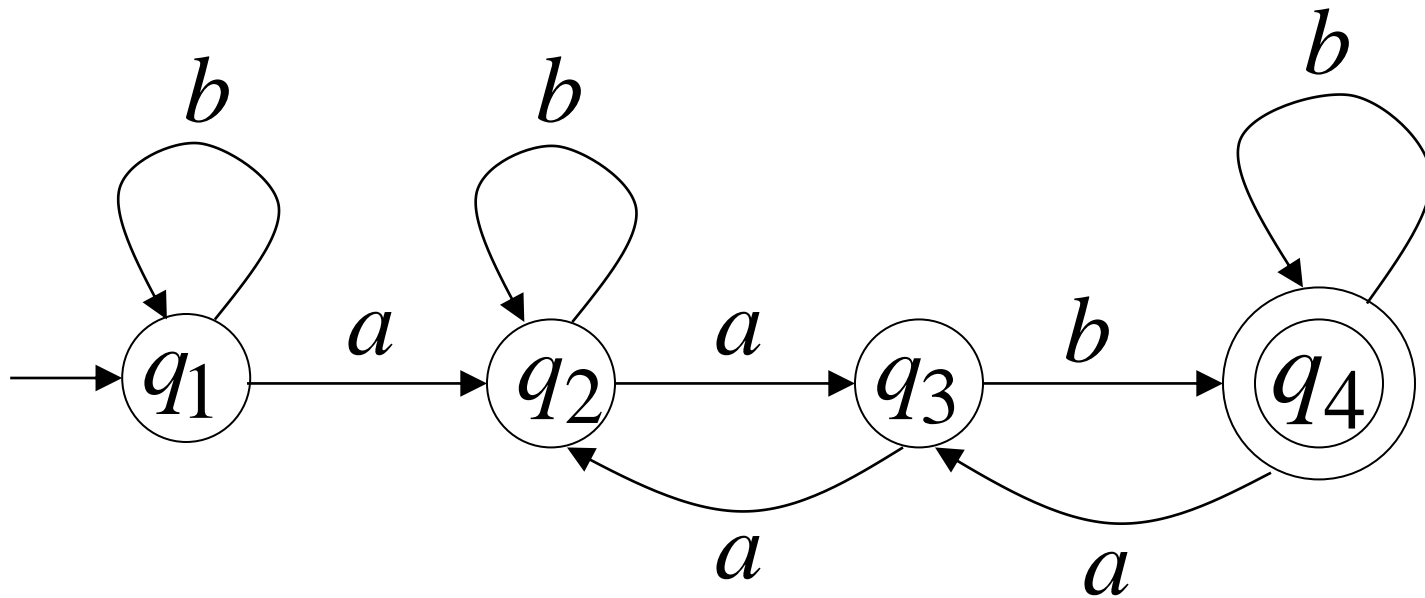
In walks of strings: $aabb$

$bbaa$

$abbabb$

$abbbabbabb...$

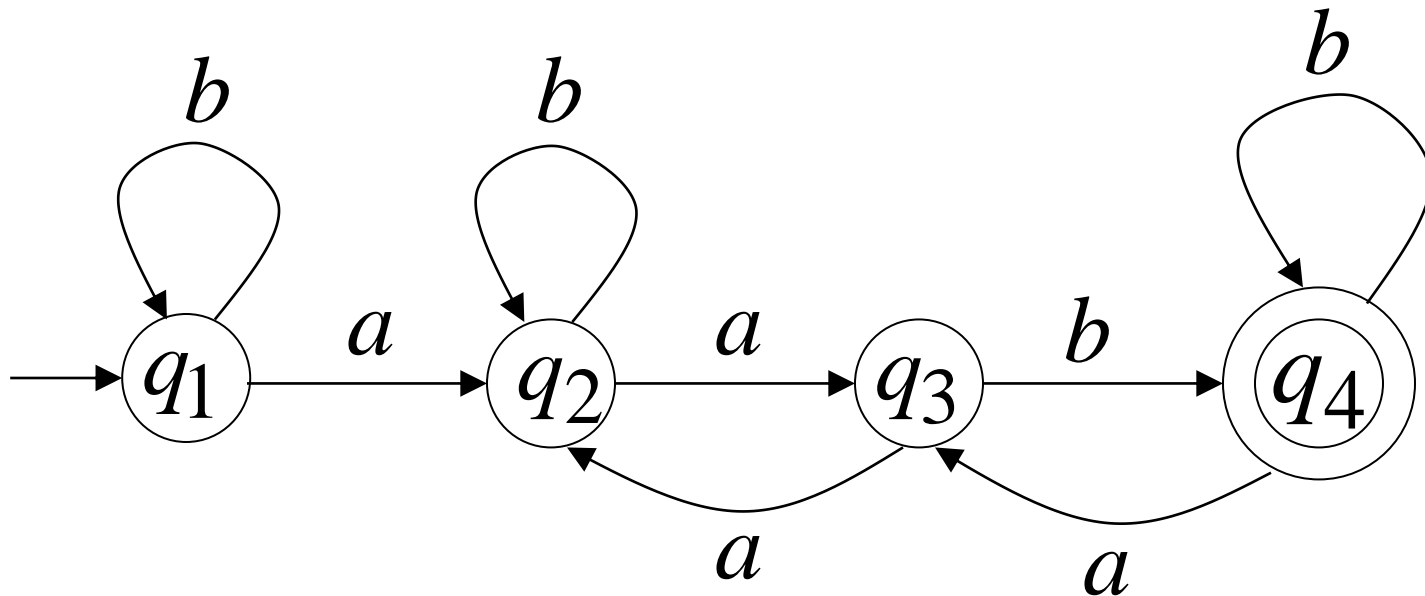
a state
is repeated



If string w has length $|w| \geq 4$:

Then the transitions of string w
are more than the states of the DFA

Thus, a state must be repeated

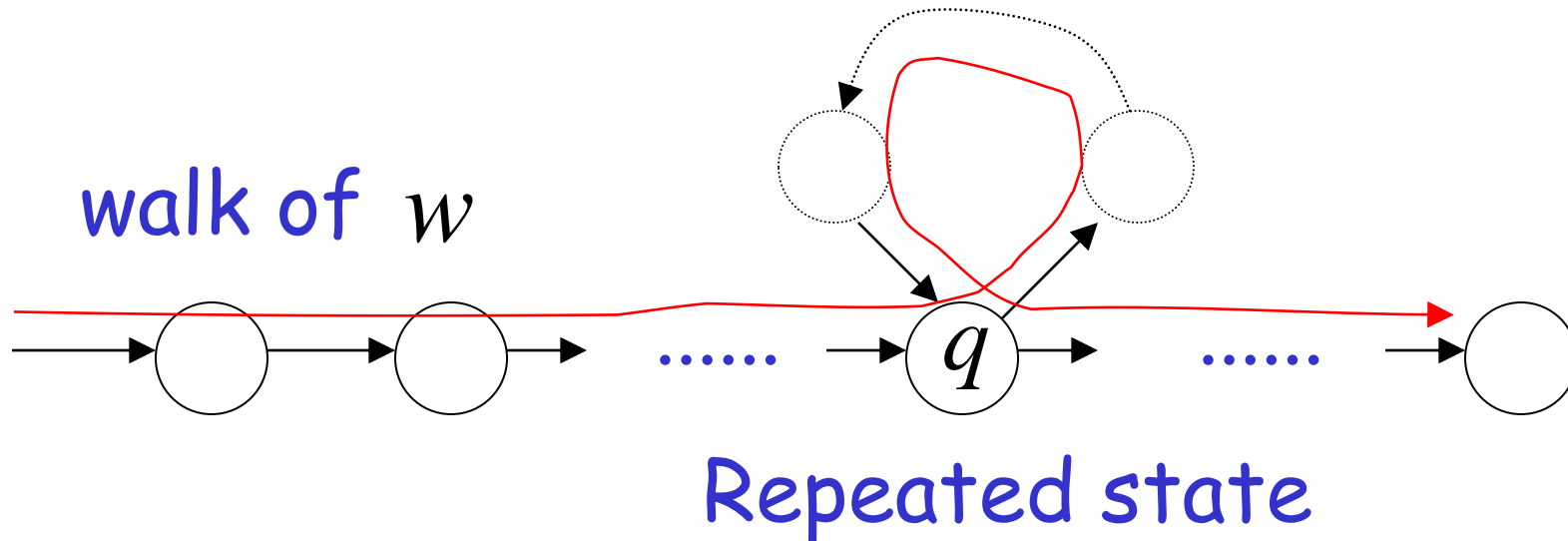


In general, for any DFA:

String w has length \geq number of states



A state q must be repeated in the walk of w

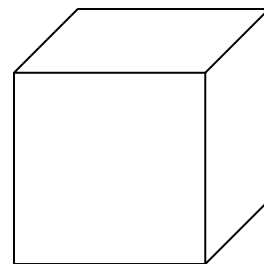


In other words for a string w :

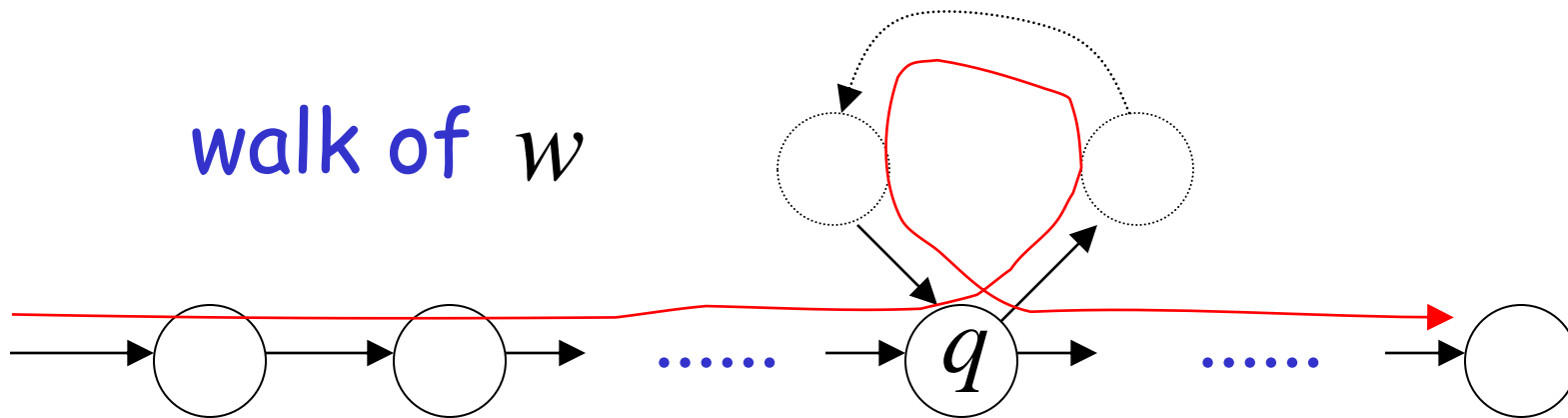
\xrightarrow{a} transitions are pigeons



(q) states are pigeonholes



walk of w

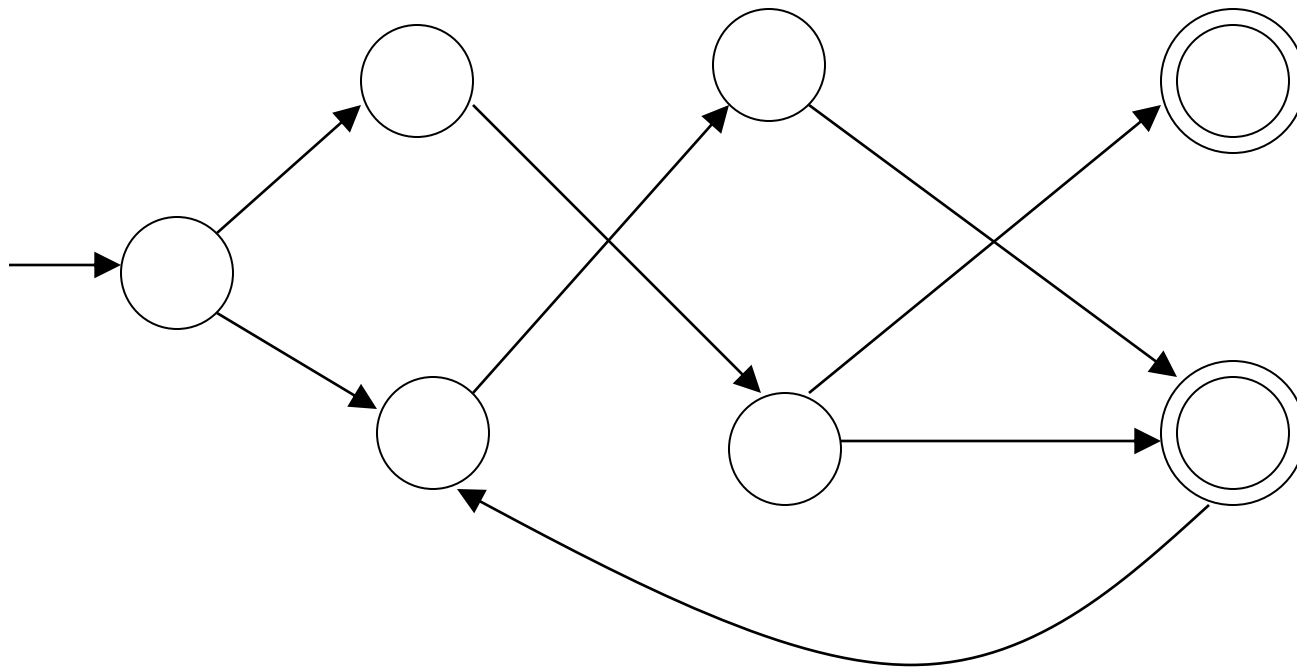


Repeated state

The Pumping Lemma

Take an **infinite** regular language L

There exists a DFA that accepts L



m
states

Take string w with $w \in L$

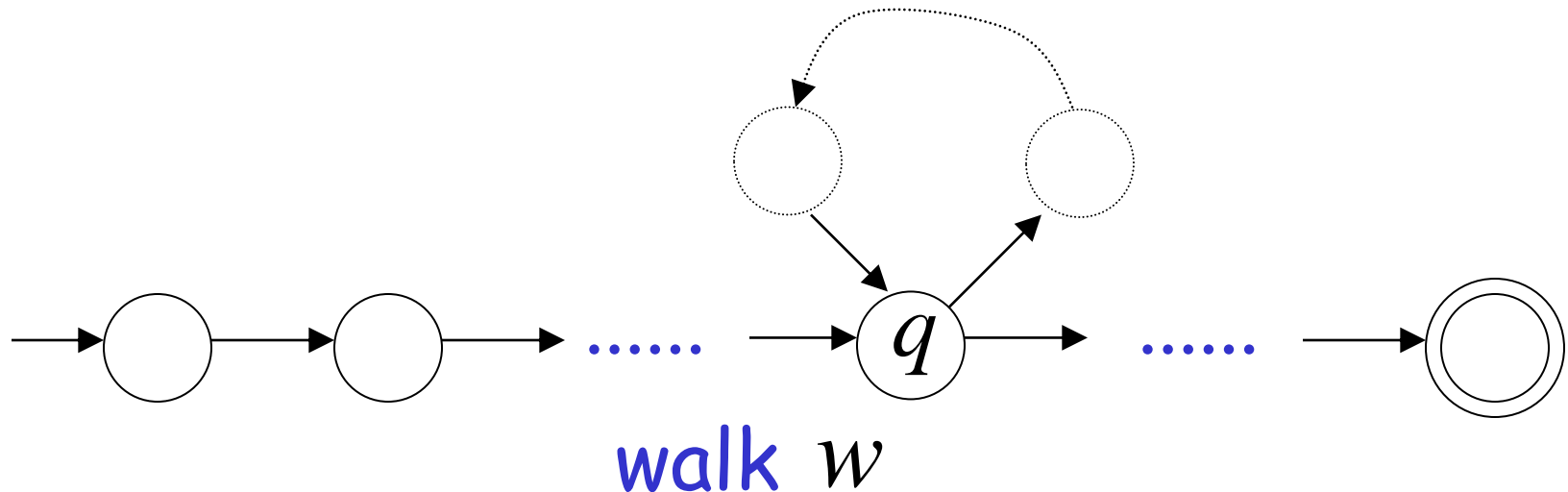
There is a walk with label w :



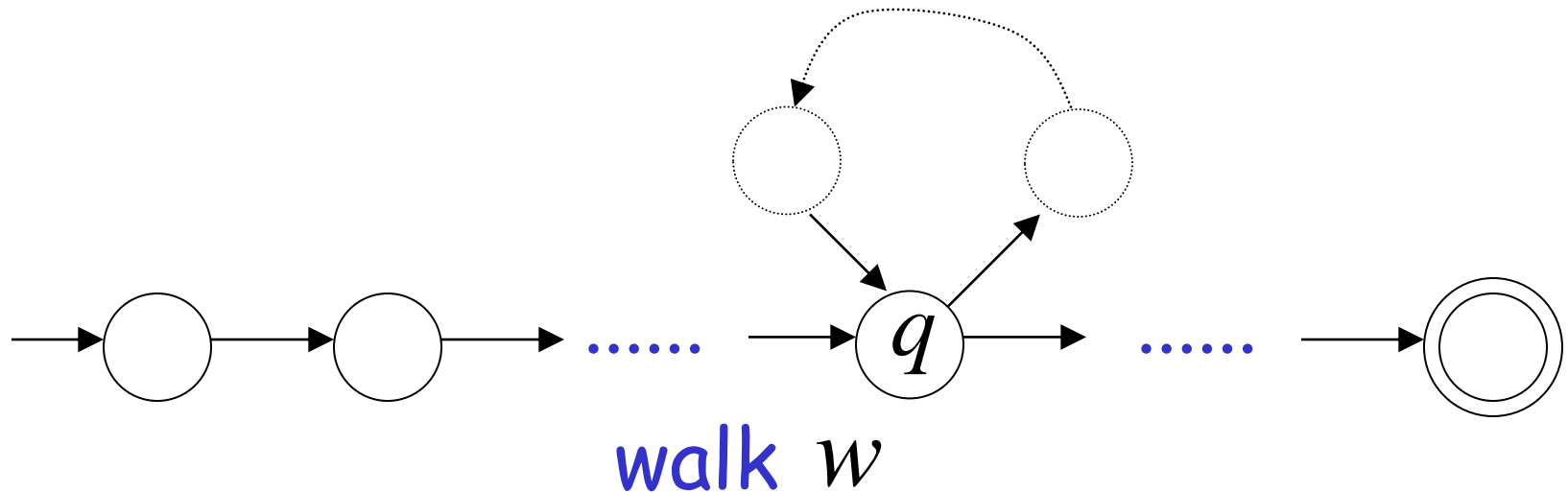
If string w has length $|w| \geq m$ (number of states of DFA)

then, from the pigeonhole principle:

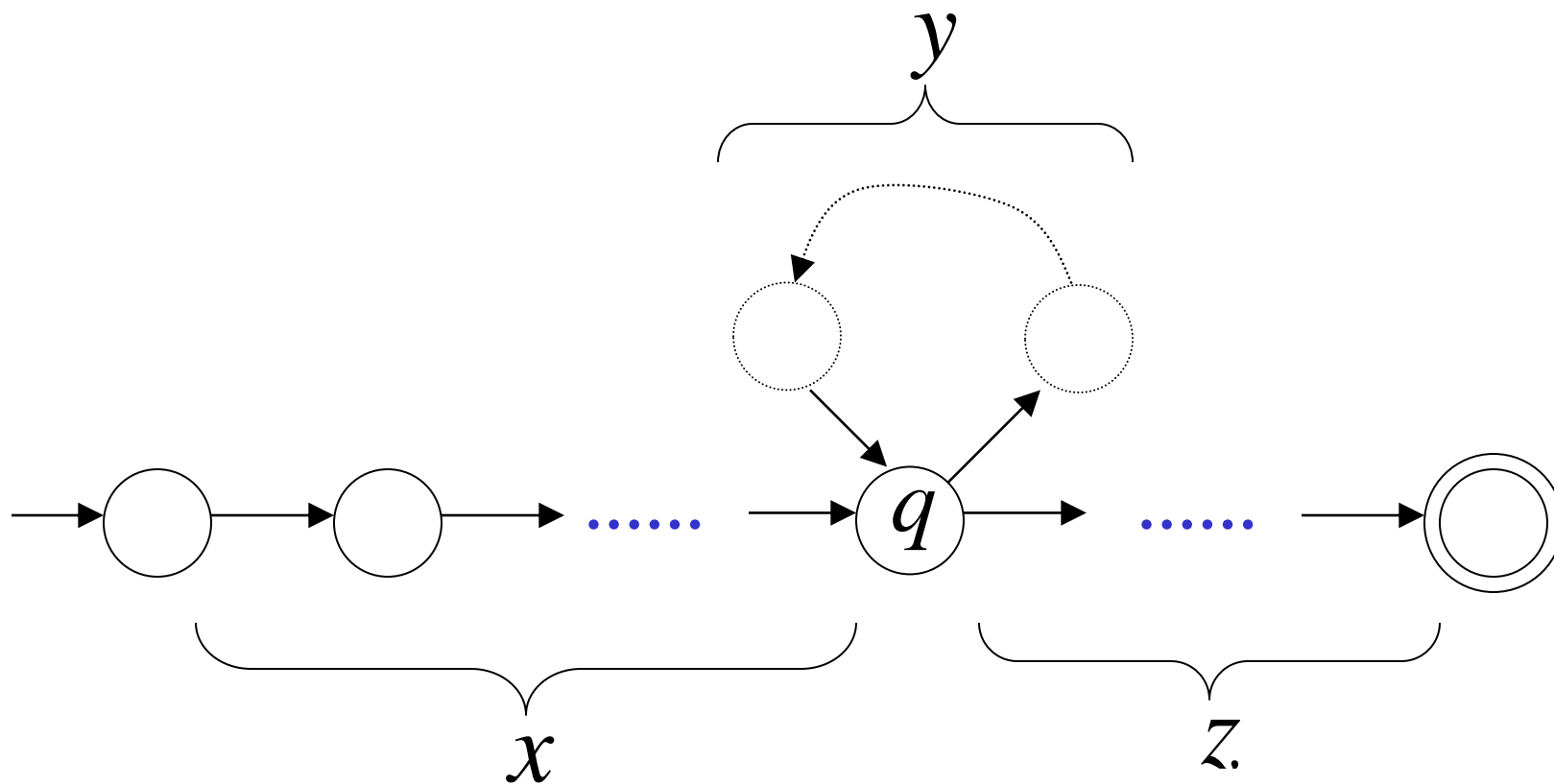
a state is repeated in the walk w



Let q be the first state repeated in the walk of w

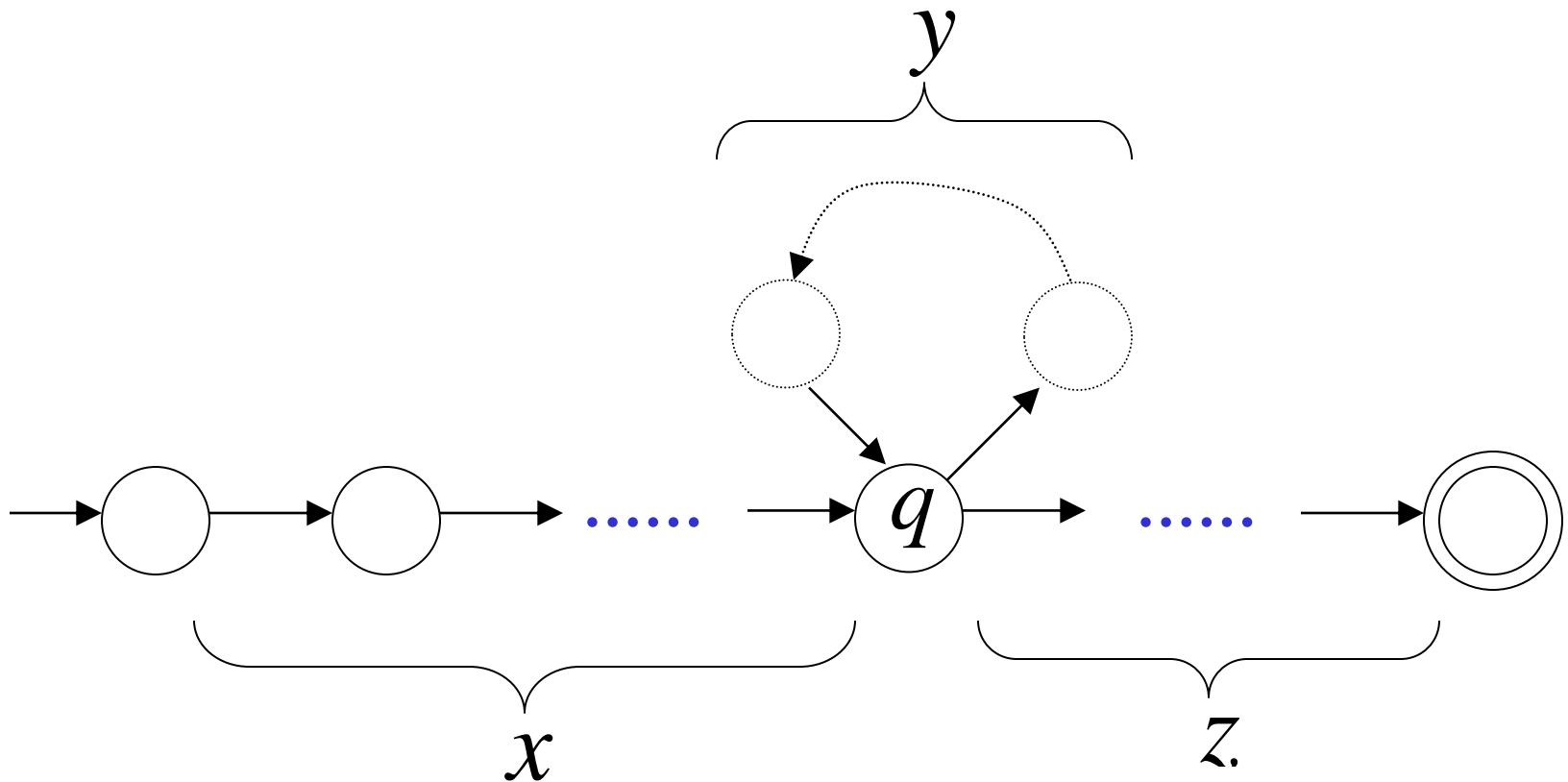


Write $w = x y z$

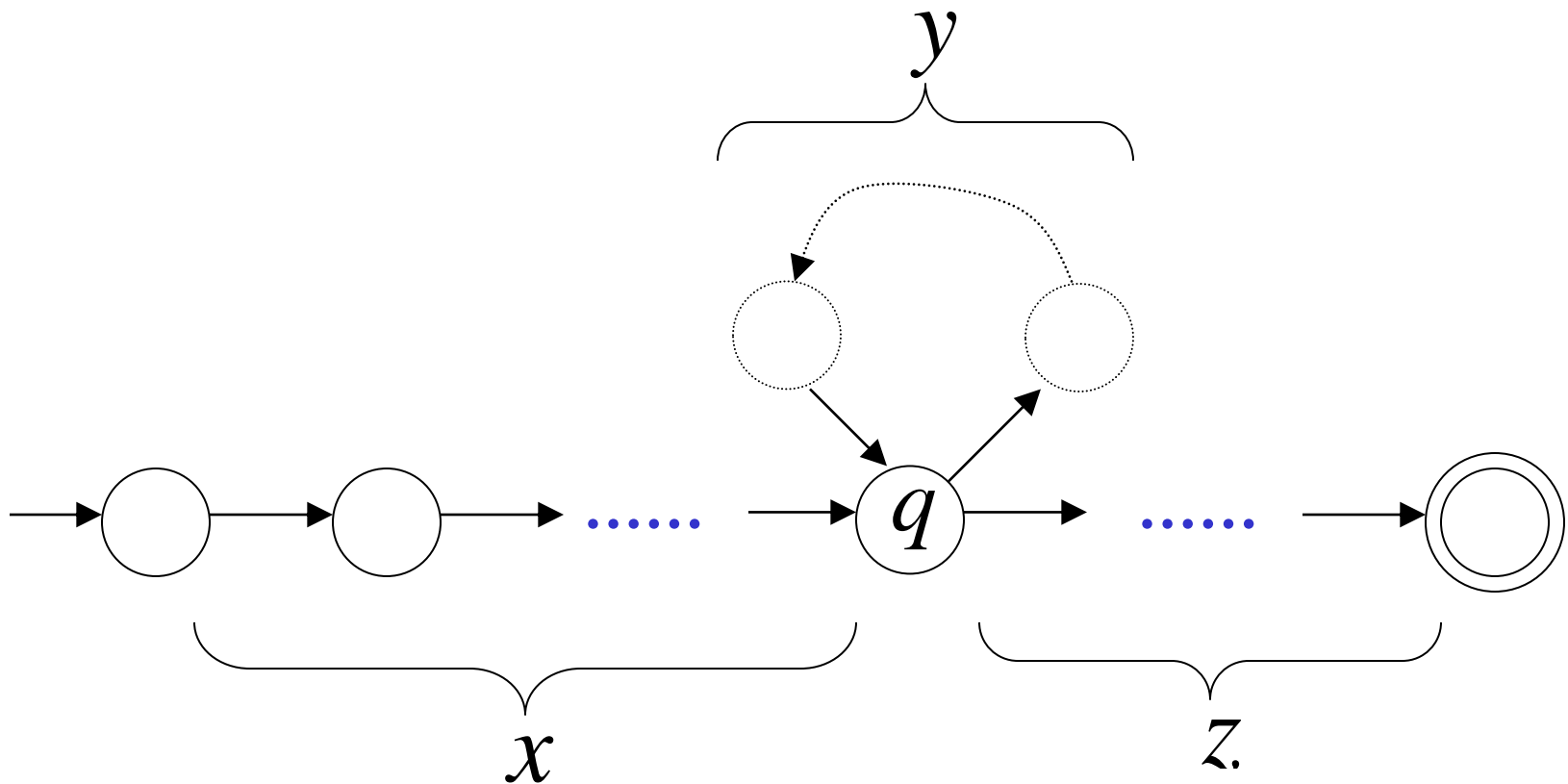


Observations: length $|x y| \leq m$ number
 of states
 of DFA

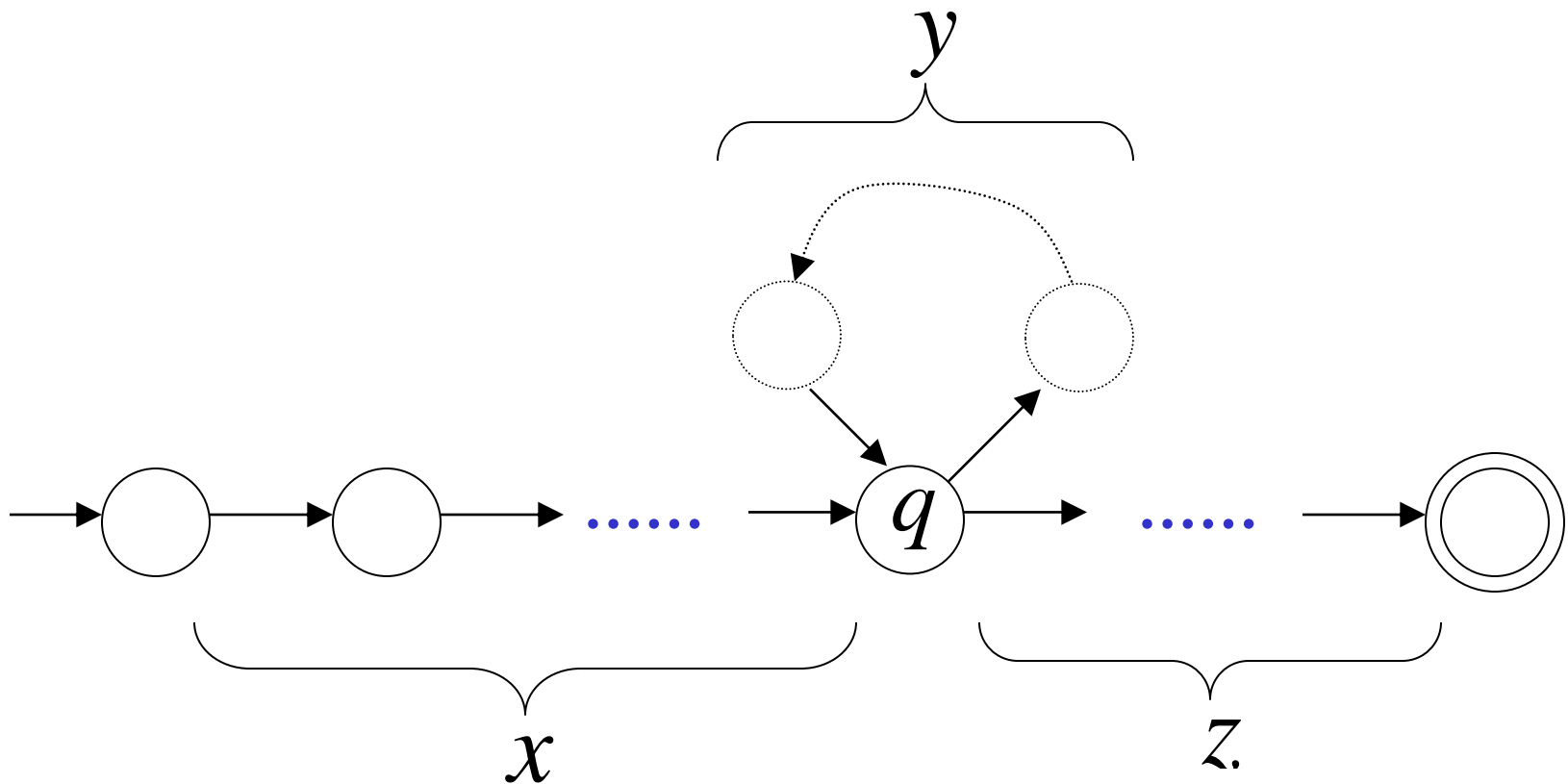
length $|y| \geq 1$



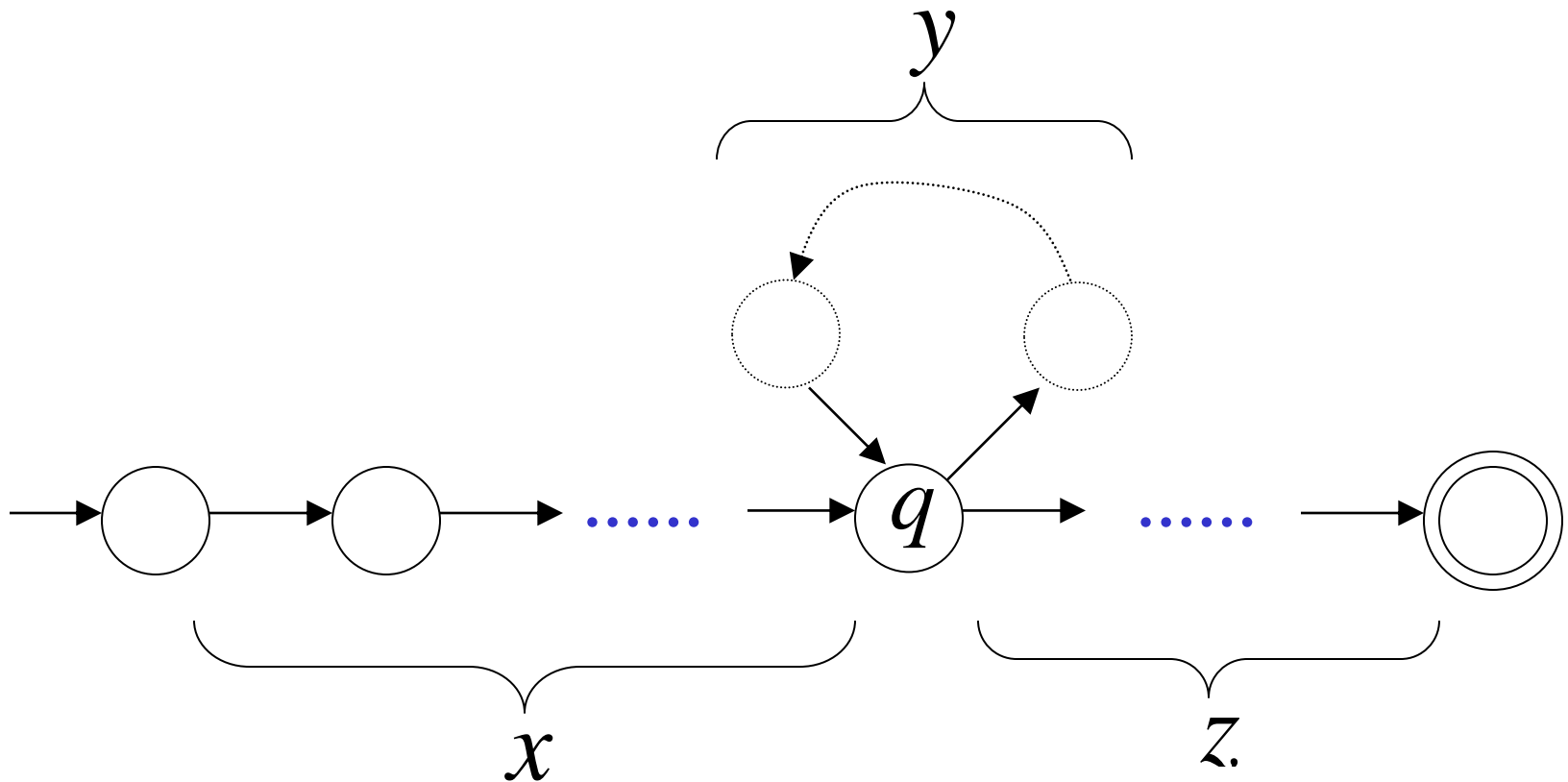
Observation: The string xzy is accepted



Observation: The string $x y y z$ is accepted



Observation: The string $x y y y z$ is accepted



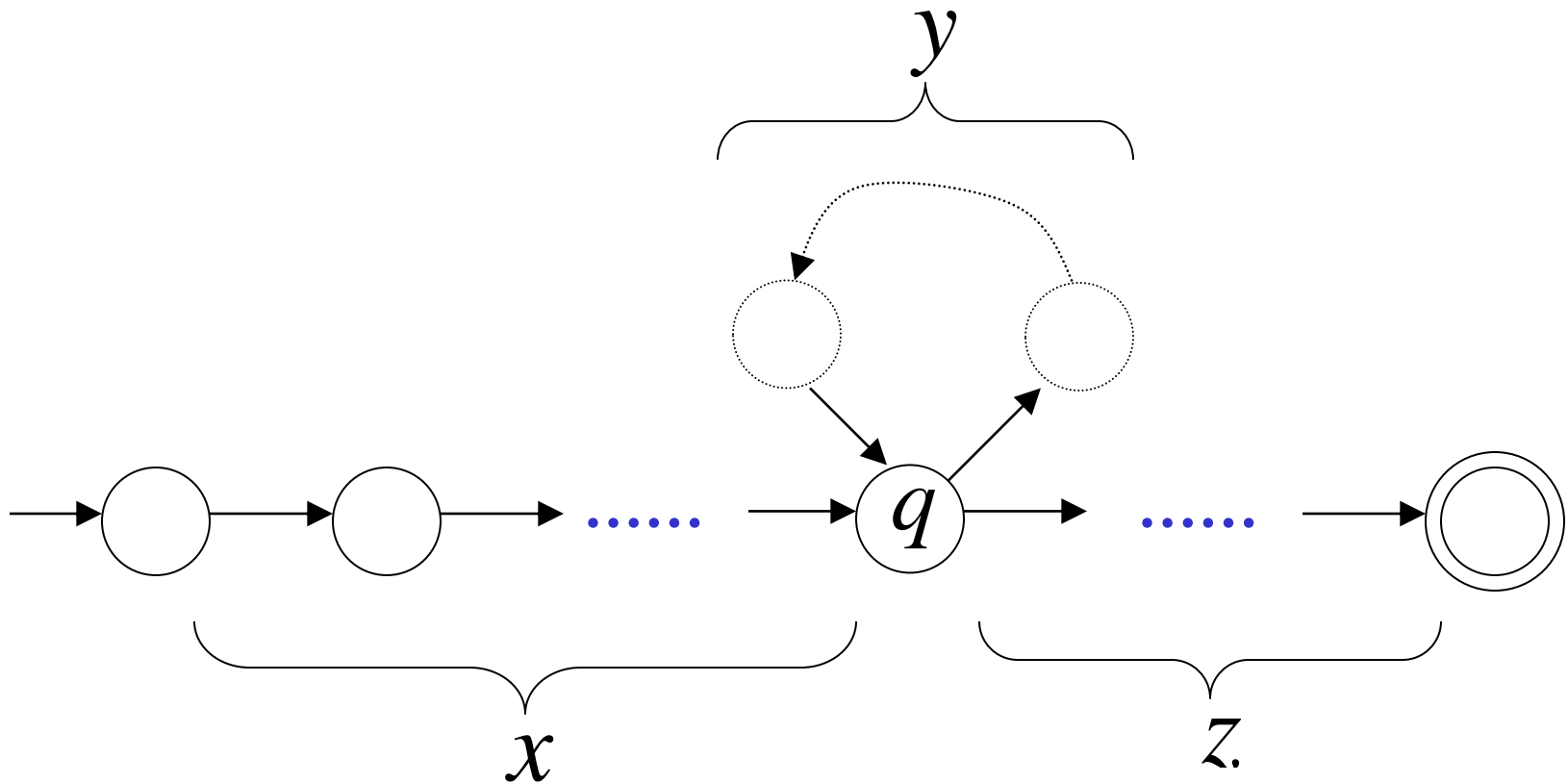
In General:

The string

$x y^i z$

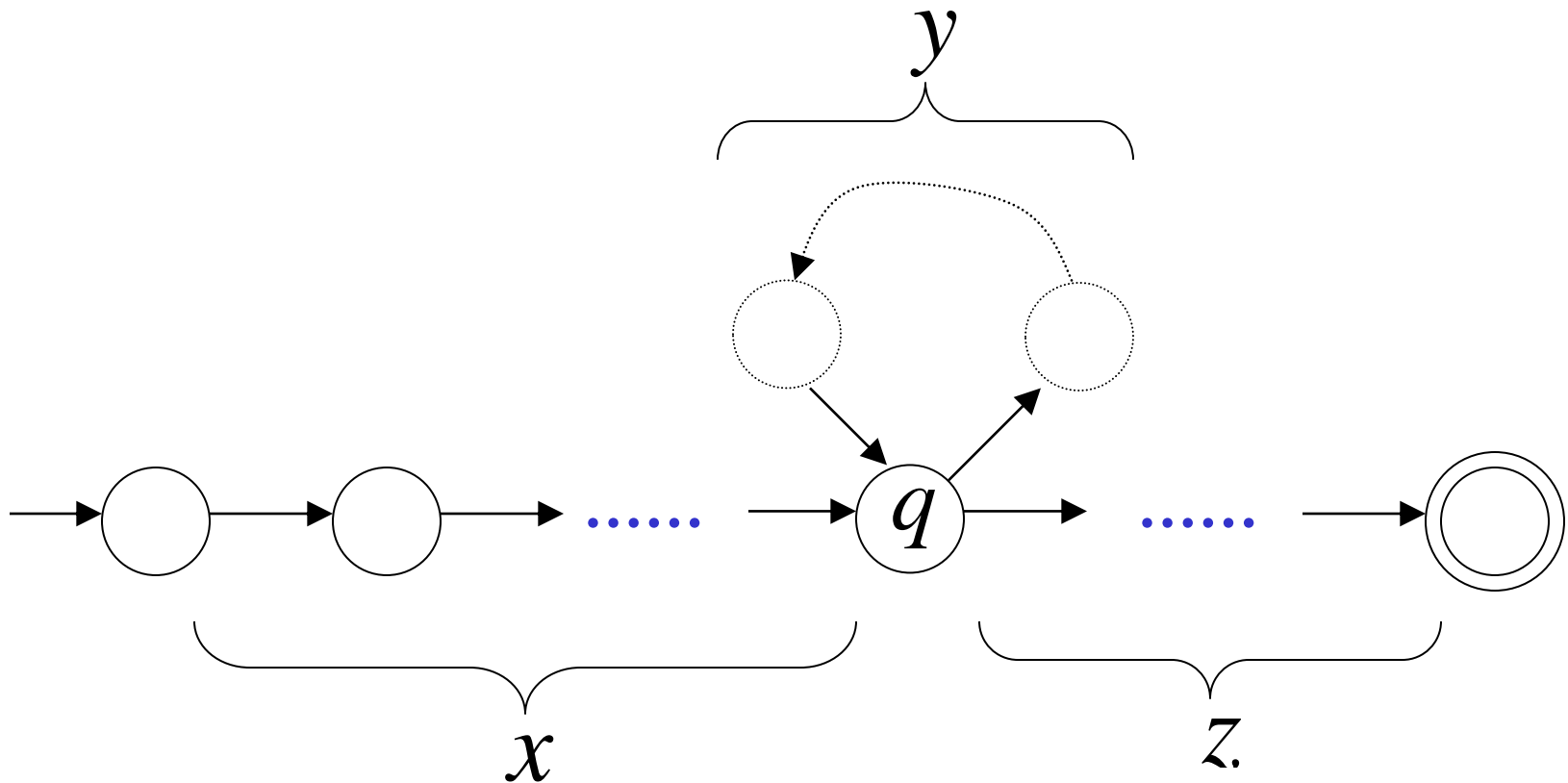
is accepted

$i = 0, 1, 2, \dots$



In General: $x y^i z \in L \quad i = 0, 1, 2, \dots$

Language accepted by the DFA



In other words, we described:

The Pumping Lemma !!!

The Pumping Lemma:

- Given a infinite regular language L
- there exists an integer m
- for any string $w \in L$ with length $|w| \geq m$
- we can write $w = x y z$
- with $|x y| \leq m$ and $|y| \geq 1$
- such that: $x y^i z \in L \quad i = 0, 1, 2, \dots$

Applications of the Pumping Lemma

Theorem: The language $L = \{a^n b^n : n \geq 0\}$
is not regular

Proof: Use the Pumping Lemma

$$L = \{a^n b^n : n \geq 0\}$$

Assume for contradiction
that L is a regular language

Since L is infinite
we can apply the Pumping Lemma

$$L = \{a^n b^n : n \geq 0\}$$

Let m be the integer in the Pumping Lemma

Pick a string w such that: $w \in L$

$$\text{length } |w| \geq m$$

We pick $w = a^m b^m$

Write: $a^m b^m = x y z$

From the Pumping Lemma

it must be that length $|x y| \leq m, |y| \geq 1$

$$xyz = a^m b^m = \overbrace{a \dots a}^m \overbrace{a \dots a \dots a b \dots b}^m$$

$x \quad y \quad z$

Thus: $y = a^k, k \geq 1$

$$x y z = a^m b^m$$

$$y = a^k, \quad k \geq 1$$

From the Pumping Lemma: $x y^i z \in L$

$$i = 0, 1, 2, \dots$$

Thus: $x y^2 z \in L$

$$x y z = a^m b^m$$

$$y = a^k, \quad k \geq 1$$

From the Pumping Lemma: $x y^2 z \in L$

$$xy^2z = \overbrace{a \dots a a \dots a a \dots a a \dots a}^{m+k} \overbrace{b \dots b}^m \in L$$

$\underbrace{\hspace{1.5cm}}_x \quad \underbrace{\hspace{1.5cm}}_y \quad \underbrace{\hspace{1.5cm}}_y \quad \underbrace{\hspace{2.5cm}}_z$

Thus: $a^{m+k} b^m \in L$

$$a^{m+k}b^m \in L \qquad k \geq 1$$

BUT: $L = \{a^n b^n : n \geq 0\}$



$$a^{m+k}b^m \notin L$$

CONTRADICTION!!!

Therefore: Our assumption that L
is a regular language is not true

Conclusion: L is not a regular language

Non-regular languages $\{a^n b^n : n \geq 0\}$



Regular languages

Non-regular languages

$$L = \{vv^R : v \in \Sigma^*\}$$



Regular languages

Theorem: The language

$$L = \{vv^R : v \in \Sigma^*\} \quad \Sigma = \{a,b\}$$

is not regular

Proof: Use the Pumping Lemma

$$L = \{vv^R : v \in \Sigma^*\}$$

Assume for contradiction
that L is a regular language

Since L is infinite
we can apply the Pumping Lemma

$$L = \{vv^R : v \in \Sigma^*\}$$

Let m be the integer in the Pumping Lemma

Pick a string w such that: $w \in L$ and

$$\text{length } |w| \geq m$$

We pick $w = a^m b^m b^m a^m$

Write $a^m b^m b^m a^m = x y z$

From the Pumping Lemma

it must be that length $|x y| \leq m, |y| \geq 1$

$$xyz = \overbrace{a \dots a}^m \overbrace{a \dots a}^m \overbrace{a \dots a}^m \overbrace{a \dots a}^m$$
$$\underbrace{a \dots a}_{x} \underbrace{a \dots a}_{y} \underbrace{a \dots a \dots a \dots a}_{z}$$

Thus: $y = a^k, k \geq 1$

$$x y z = a^m b^m b^m a^m$$

$$y = a^k, \quad k \geq 1$$

From the Pumping Lemma: $x y^i z \in L$
 $i = 0, 1, 2, \dots$

Thus: $x y^2 z \in L$

$$x y z = a^m b^m b^m a^m \qquad y = a^k, \quad k \geq 1$$

From the Pumping Lemma: $x y^2 z \in L$

$$xy^2z = \overbrace{a \dots a}^{m+k} \overbrace{a \dots a}^m \overbrace{a \dots a}^m \overbrace{a \dots a}^m \in L$$

x
 y
 y
 z

Thus: $a^{m+k} b^m b^m a^m \in L$

$$a^{m+k}b^mb^ma^m \in L \quad k \geq 1$$

BUT: $L = \{vv^R : v \in \Sigma^*\}$



$$a^{m+k}b^mb^ma^m \notin L$$

CONTRADICTION!!!

Therefore: Our assumption that L
is a regular language is not true

Conclusion: L is not a regular language

Non-regular languages

$$L = \{a^n b^l c^{n+l} : n, l \geq 0\}$$



Regular languages

Theorem: The language

$$L = \{a^n b^l c^{n+l} : n, l \geq 0\}$$

is not regular

Proof: Use the Pumping Lemma

$$L = \{a^n b^l c^{n+l} : n, l \geq 0\}$$

Assume for contradiction
that L is a regular language

Since L is infinite
we can apply the Pumping Lemma

$$L = \{a^n b^l c^{n+l} : n, l \geq 0\}$$

Let m be the integer in the Pumping Lemma

Pick a string w such that: $w \in L$ and

$$\text{length } |w| \geq m$$

We pick $w = a^m b^m c^{2m}$

Write $a^m b^m c^{2m} = x y z$

From the Pumping Lemma

it must be that length $|x y| \leq m, |y| \geq 1$

$$xyz = \overbrace{a \dots a}^m \overbrace{a \dots a}^m \overbrace{ab \dots bc \dots cc \dots c}^{2m}$$
$$\underbrace{\hspace{1.5cm}}_x \underbrace{\hspace{1.5cm}}_y \underbrace{\hspace{4.5cm}}_z$$

Thus: $y = a^k, k \geq 1$

$$x y z = a^m b^m c^{2m}$$

$$y = a^k, \quad k \geq 1$$

From the Pumping Lemma: $x y^i z \in L$
 $i = 0, 1, 2, \dots$

Thus: $x y^0 z = xz \in L$

$$x y z = a^m b^m c^{2m} \qquad y = a^k, \quad k \geq 1$$

From the Pumping Lemma: $xz \in L$

$$xz = \overbrace{a \dots a}^{m-k} \overbrace{a \dots a}^m \overbrace{b \dots b}^m \overbrace{c \dots c}^{2m} \in L$$

$$\underbrace{\hspace{1.5cm}}_x \underbrace{\hspace{4.5cm}}_z$$

Thus: $a^{m-k} b^m c^{2m} \in L$

$$a^{m-k} b^m c^{2m} \in L \quad k \geq 1$$

BUT: $L = \{a^n b^l c^{n+l} : n, l \geq 0\}$



$$a^{m-k} b^m c^{2m} \notin L$$

CONTRADICTION!!!

Therefore: Our assumption that L
is a regular language is not true

Conclusion: L is not a regular language

Non-regular languages

$$L = \{a^{n!} : n \geq 0\}$$



Regular languages

Theorem: The language $L = \{a^{n!} : n \geq 0\}$
is not regular

$$n! = 1 \cdot 2 \cdots (n-1) \cdot n$$

Proof: Use the Pumping Lemma

$$L = \{a^{n!} : n \geq 0\}$$

Assume for contradiction
that L is a regular language

Since L is infinite
we can apply the Pumping Lemma

$$L = \{a^{n!} : n \geq 0\}$$

Let m be the integer in the Pumping Lemma

Pick a string w such that: $w \in L$

$$\text{length } |w| \geq m$$

We pick $w = a^{m!}$

Write $a^{m!} = x y z$

From the Pumping Lemma

it must be that length $|x y| \leq m, |y| \geq 1$

$$xyz = a^{m!} = \overbrace{a \dots a}^m \overbrace{a \dots a}^{m!-m}$$
$$\underbrace{\hspace{1.5cm}}_x \underbrace{\hspace{1.5cm}}_y \underbrace{\hspace{4cm}}_z$$

Thus: $y = a^k, 1 \leq k \leq m$

$$x y z = a^{m!}$$

$$y = a^k, \quad 1 \leq k \leq m$$

From the Pumping Lemma: $x y^i z \in L$
 $i = 0, 1, 2, \dots$

Thus: $x y^2 z \in L$

$$x y z = a^{m!}$$

$$y = a^k, \quad 1 \leq k \leq m$$

From the Pumping Lemma: $x y^2 z \in L$

$$xy^2z = \overbrace{a \dots a a \dots a a \dots a a \dots a}^{m+k} \overbrace{a \dots a}^{m!-m} \in L$$

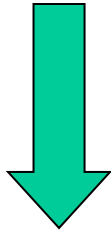
$\underbrace{a \dots a}_x$
 $\underbrace{a \dots a}_y$
 $\underbrace{a \dots a}_y$

$\underbrace{a \dots a}_z$

Thus: $a^{m!+k} \in L$

$$a^{m!+k} \in L \qquad 1 \leq k \leq m$$

Since: $L = \{a^{n!} : n \geq 0\}$



There must exist p such that:

$$m!+k = p!$$

However: $m!+k \leq m!+m$ for $m > 1$

$$\leq m!+m!$$

$$< m!m + m!$$

$$= m!(m+1)$$

$$= (m+1)!$$



$$m!+k < (m+1)!$$



$$m!+k \neq p! \quad \text{for any } p$$

$$a^{m!+k} \in L \qquad 1 \leq k \leq m$$

BUT: $L = \{a^{n!} : n \geq 0\}$



$$a^{m!+k} \notin L$$

CONTRADICTION!!!

Therefore: Our assumption that L
is a regular language is not true

Conclusion: L is not a regular language