COURSE: THEORY OF AUTOMATA COMPUTATION

TOPICS TO BE COVERED

- Non Regular Languages
- Pumping Lemma

NON REGULAR LANGUAGE
HOW CAN WE PROVE THAT A LANGUAGE IS NOT REGULAR?

Non-regular languages

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

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

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 or NFA or RE that accepts L

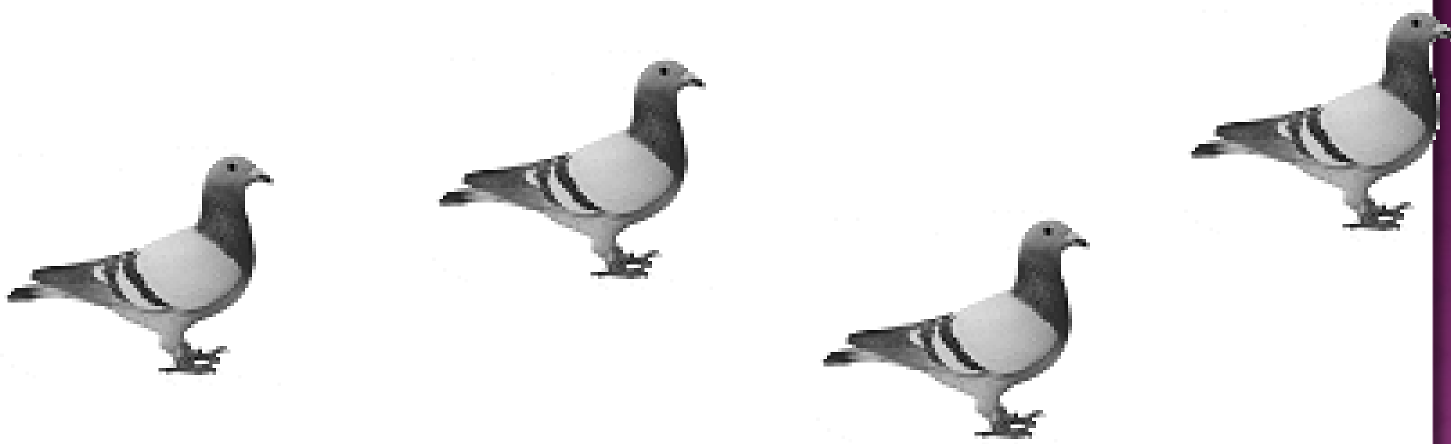
Difficulty: this is not easy to prove

(since there is an infinite number of them)

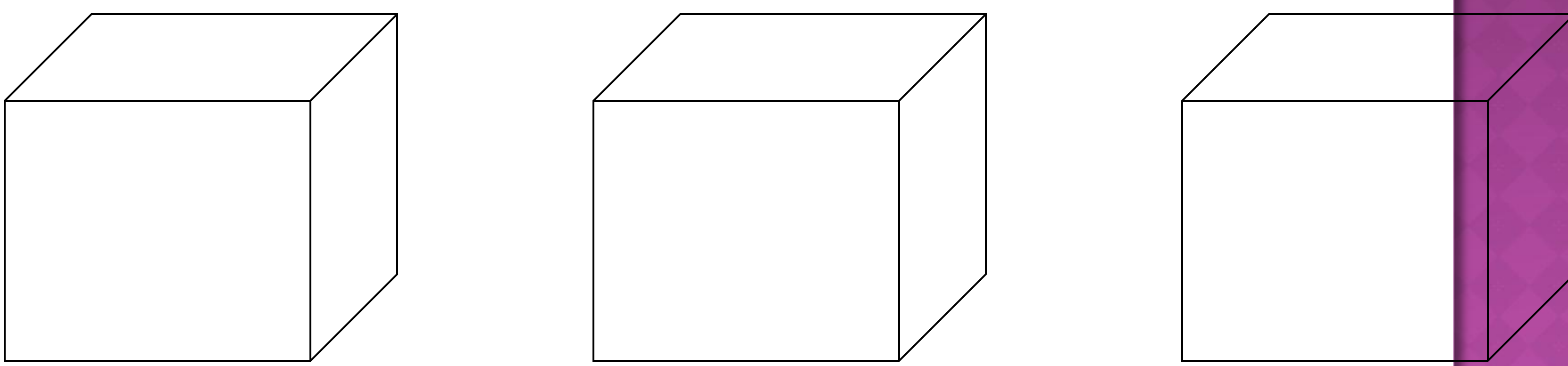
Solution: use the Pumping Lemma !!!

THE PIGEONHOLE PRINCIPLE

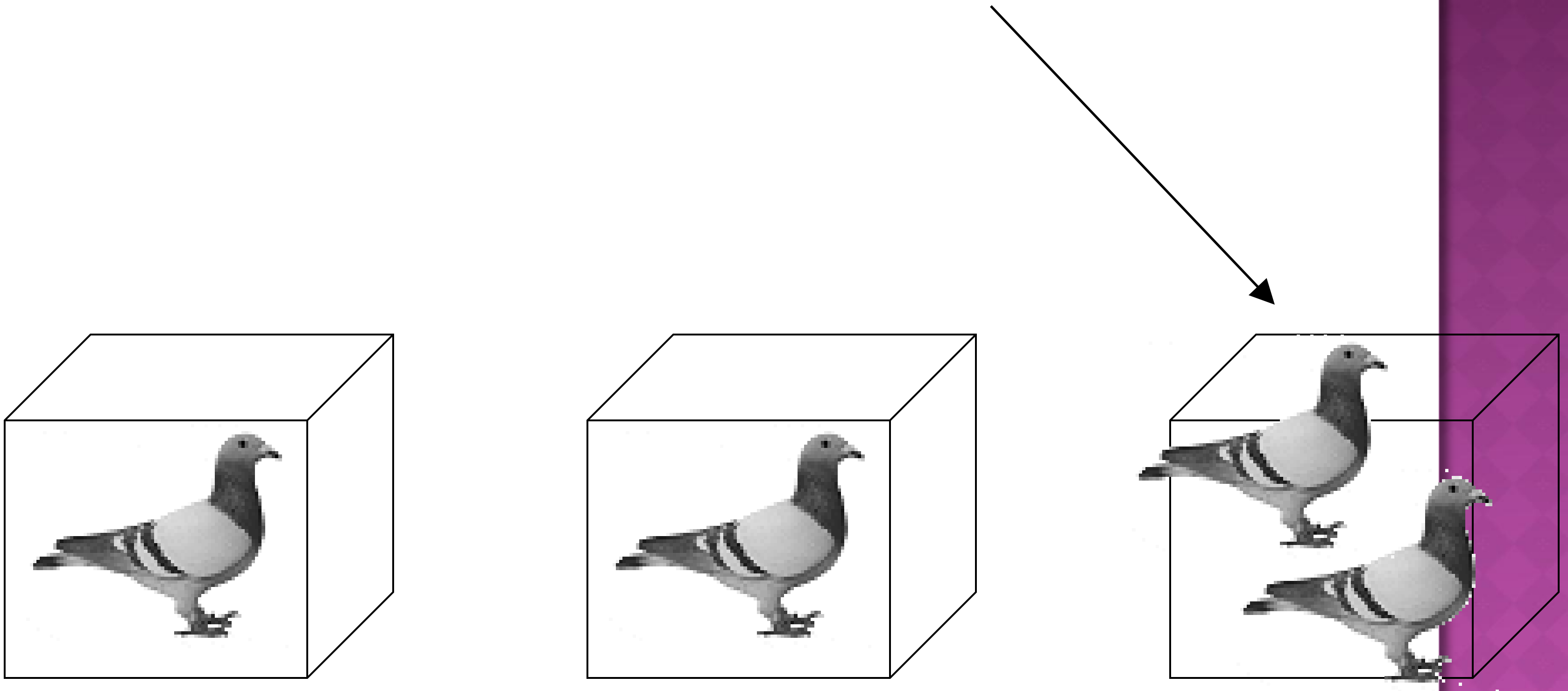
4 pigeons



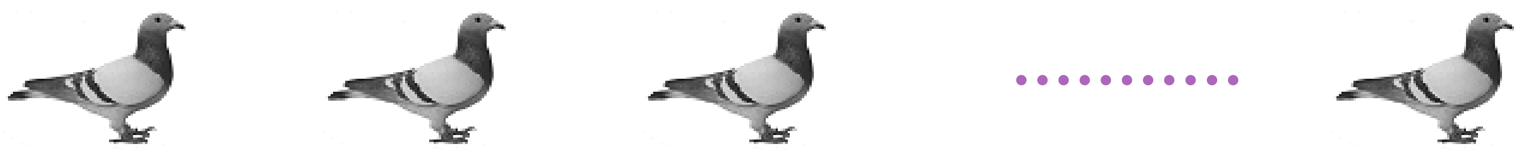
3 pigeonholes



A pigeonhole must contain at least two pigeons

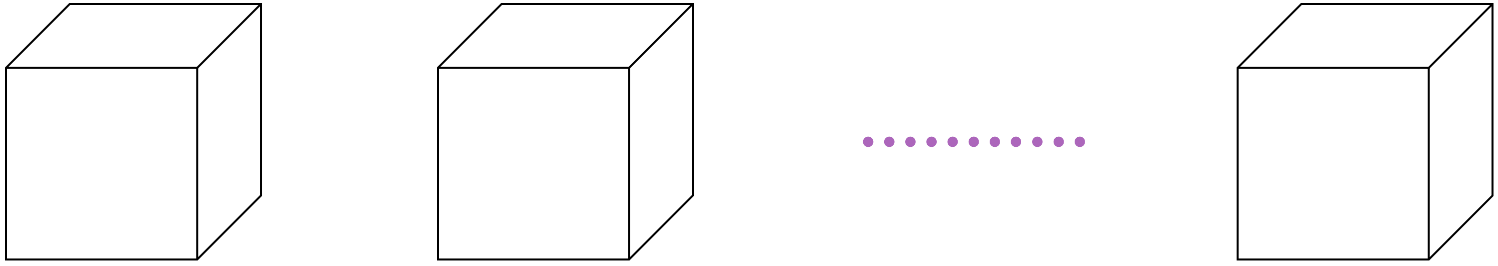


n pigeons



m pigeonholes

$n > m$



THE PIGEONHOLE PRINCIPLE

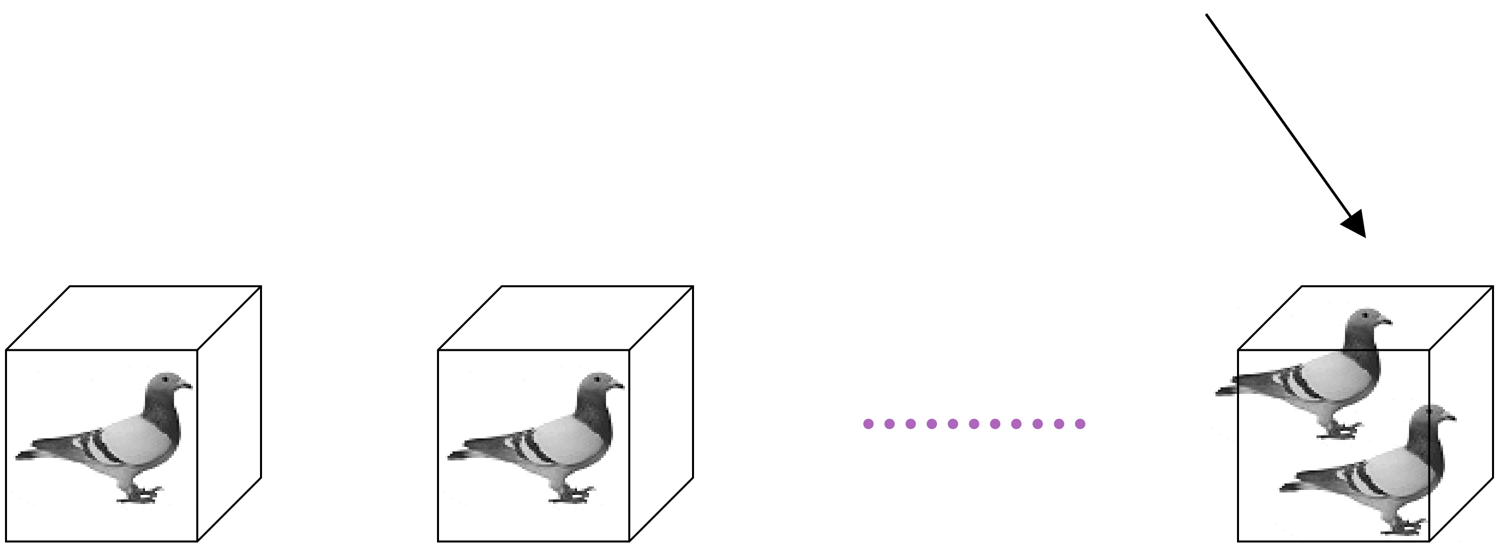
n pigeons

-

m pigeonholes

$$n > m$$

There is a pigeonhole with at least 2 pigeons

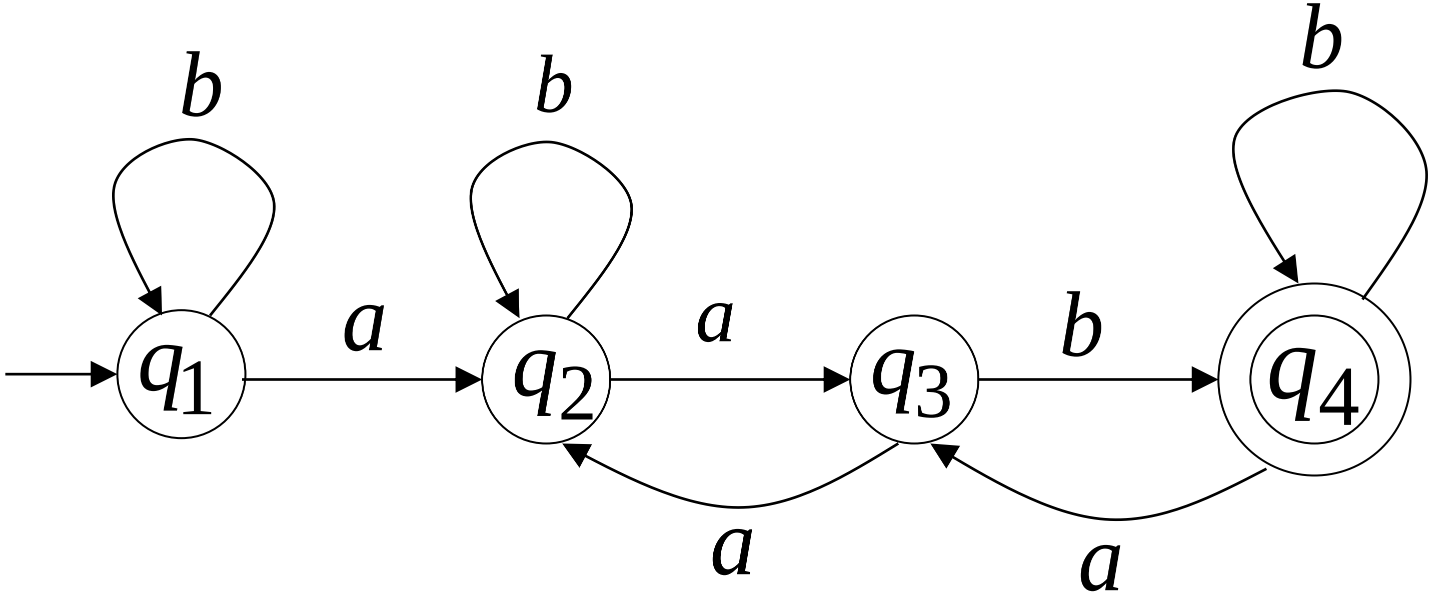


THE PIGEONHOLE PRINCIPLE

AND

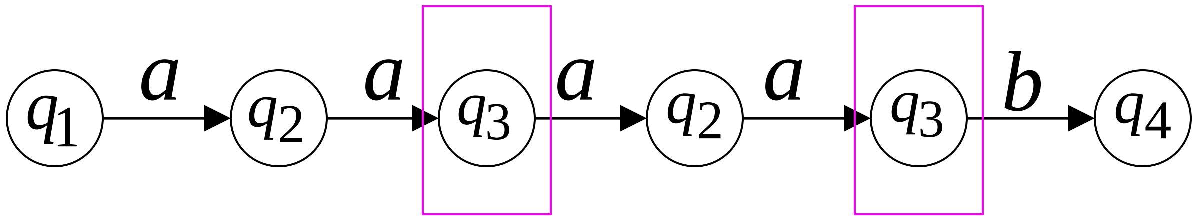
DFAS

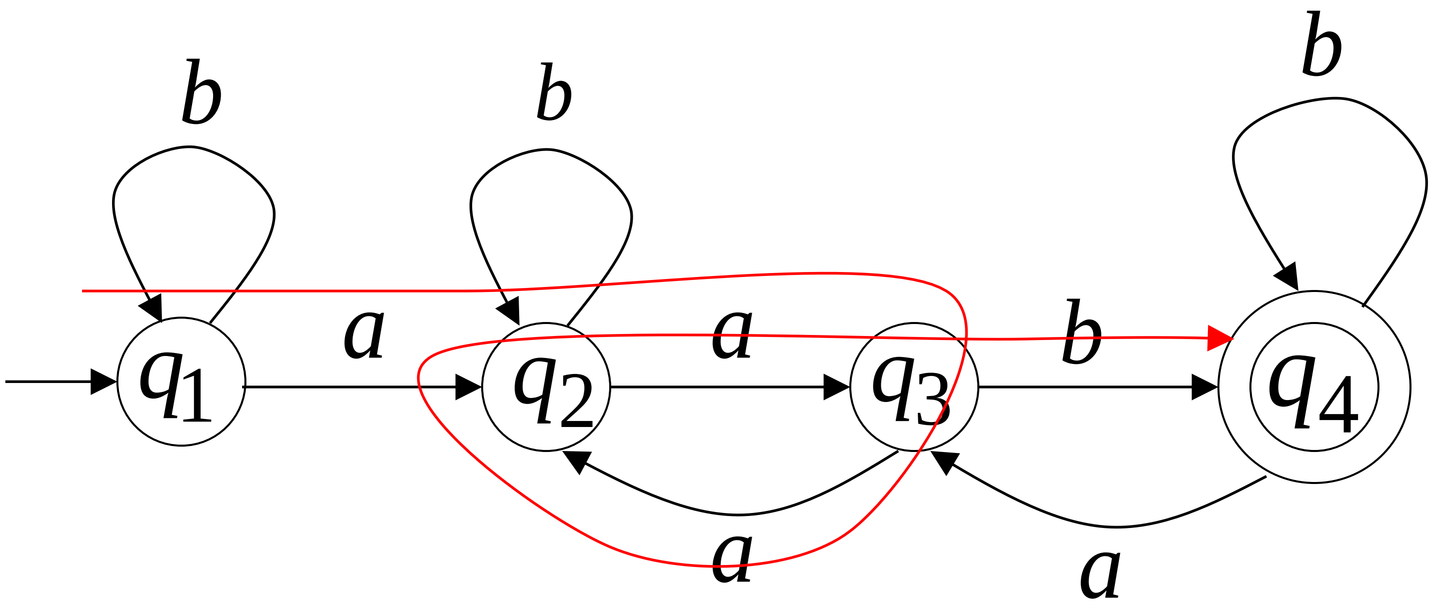
- Consider a DFA with 4 states



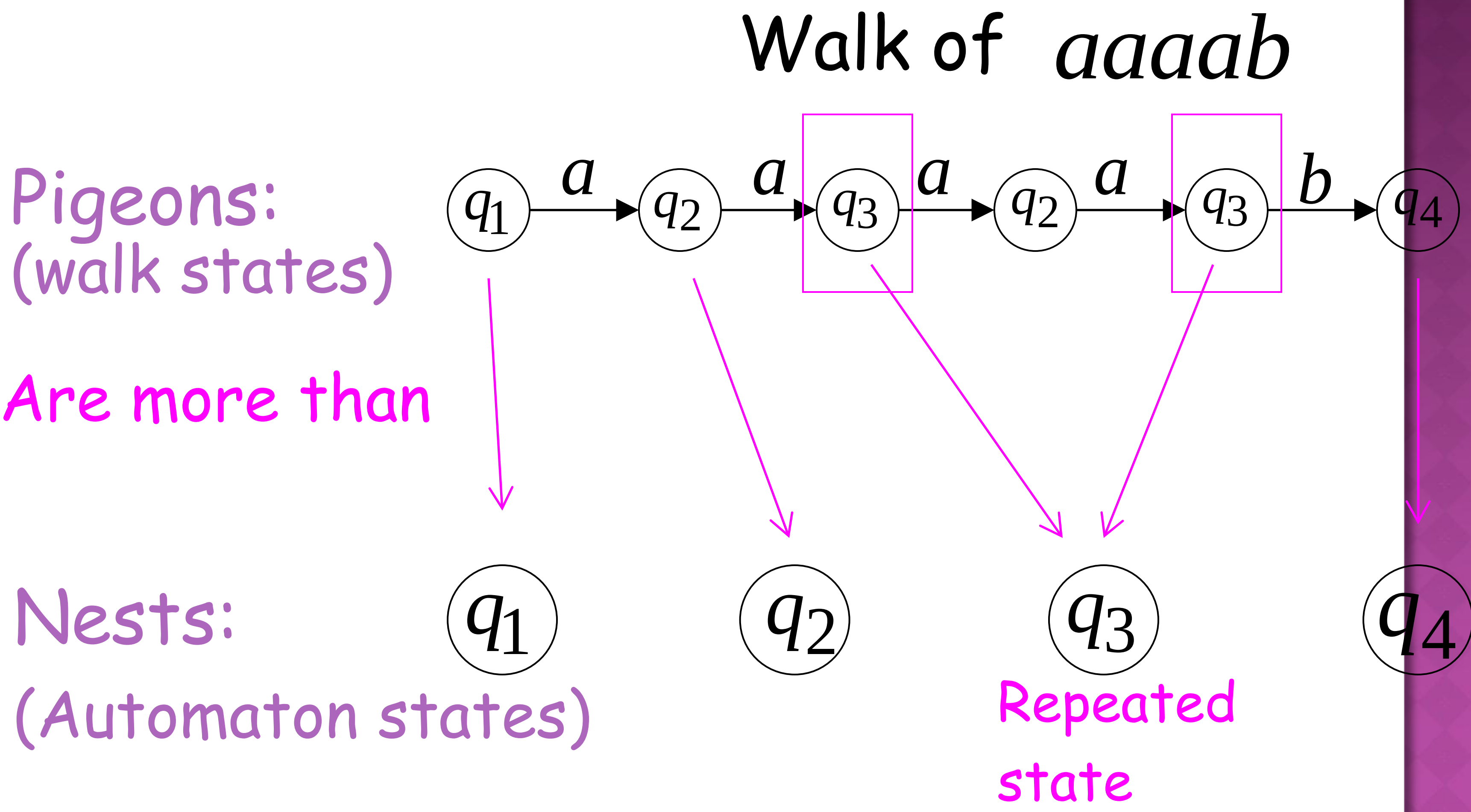
Consider the walk of a "long" string: *aaaab*
(length at least 4)

A state is repeated in the walk of *aaaab*





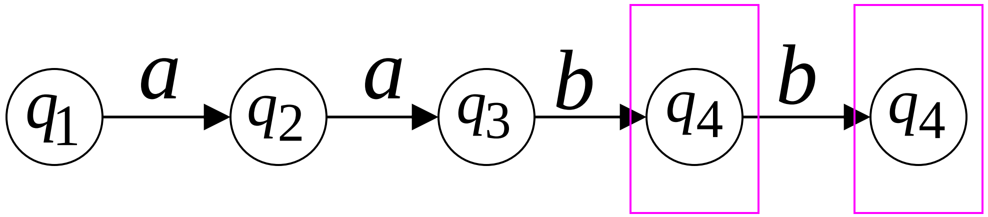
The state is repeated as a result of the pigeonhole principle

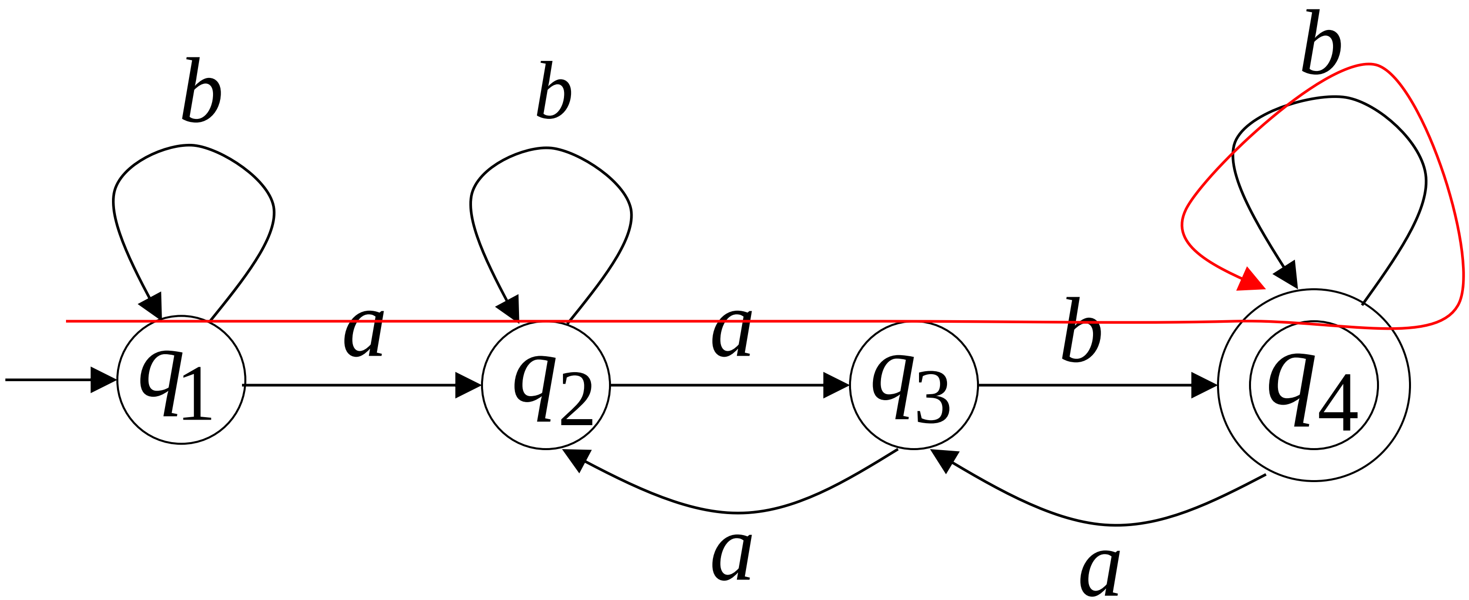


Consider the walk of a "long" string: $aabb$
(length at least 4)

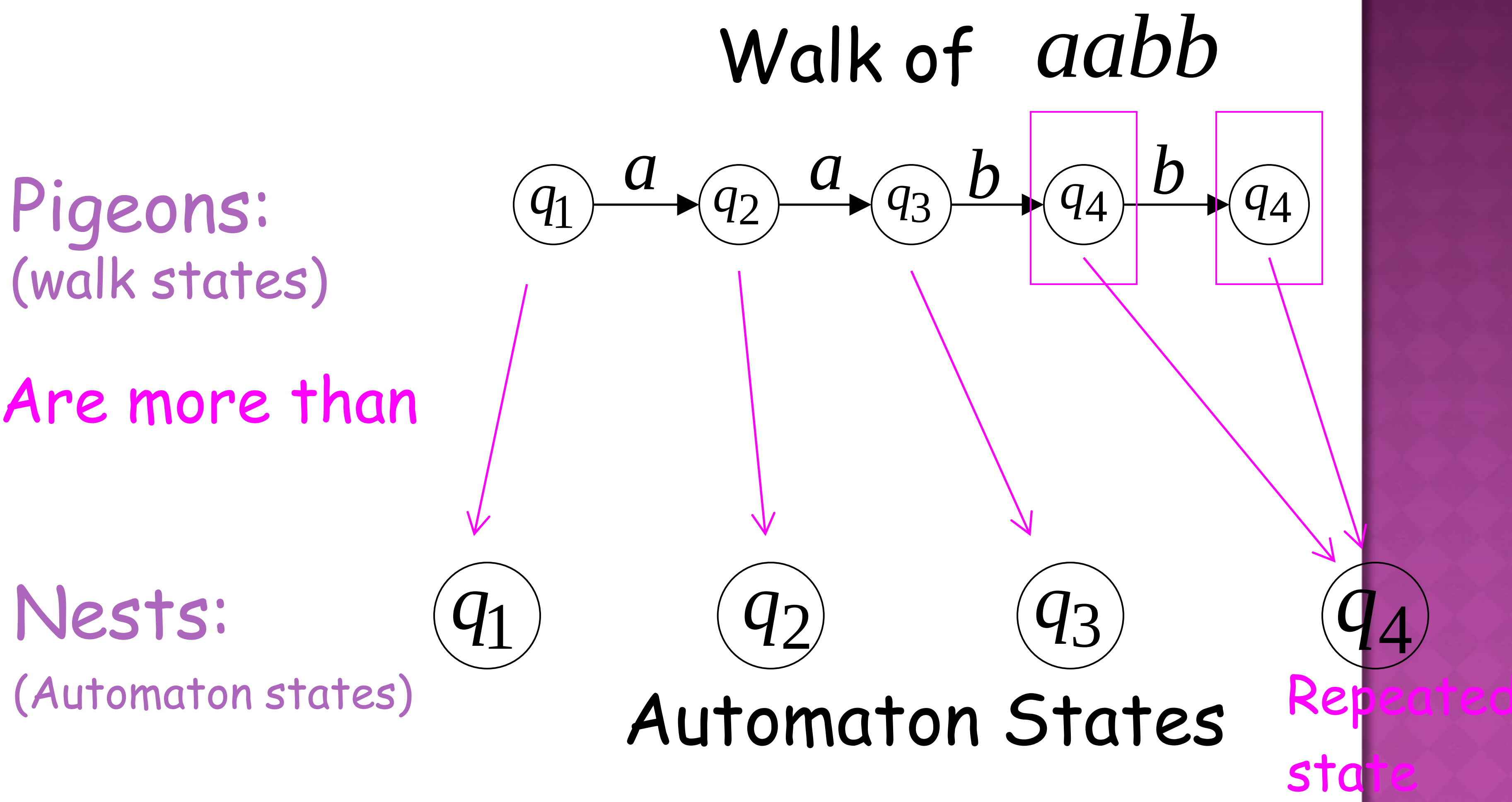
Due to the pigeonhole principle:

A state is repeated in the walk of $aabb$



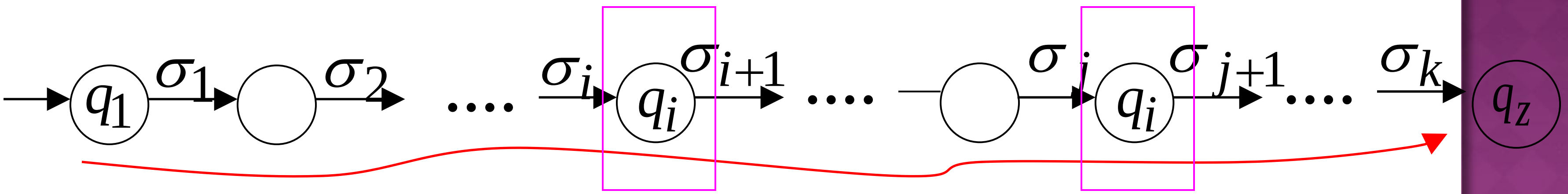


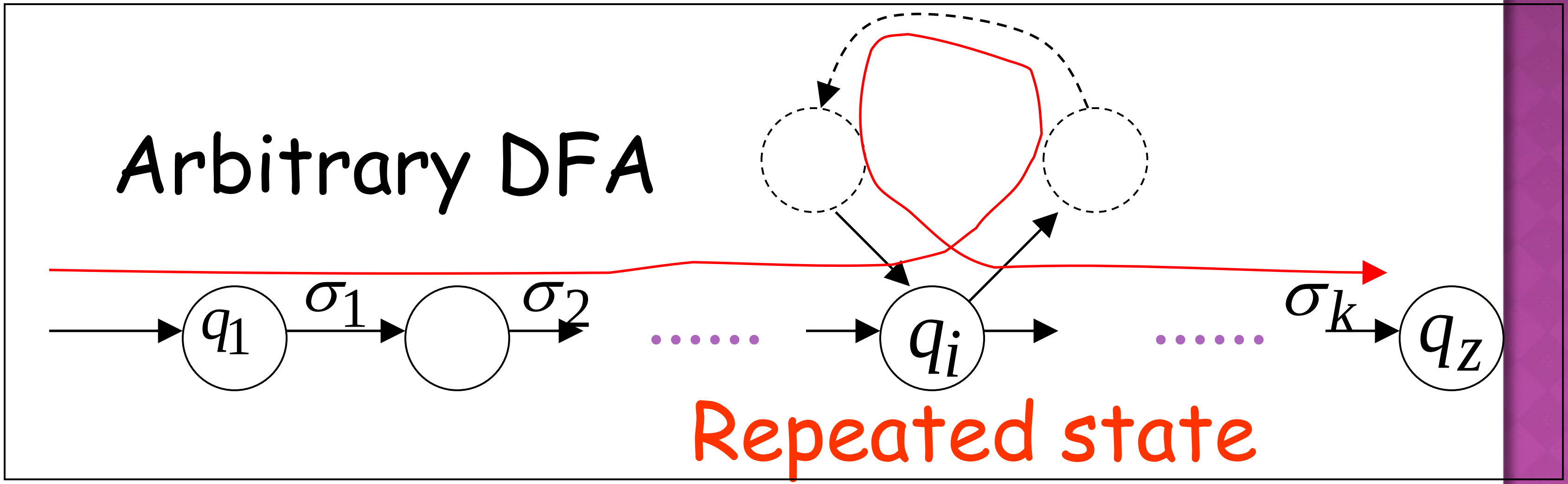
The state is repeated as a result of the pigeonhole principle



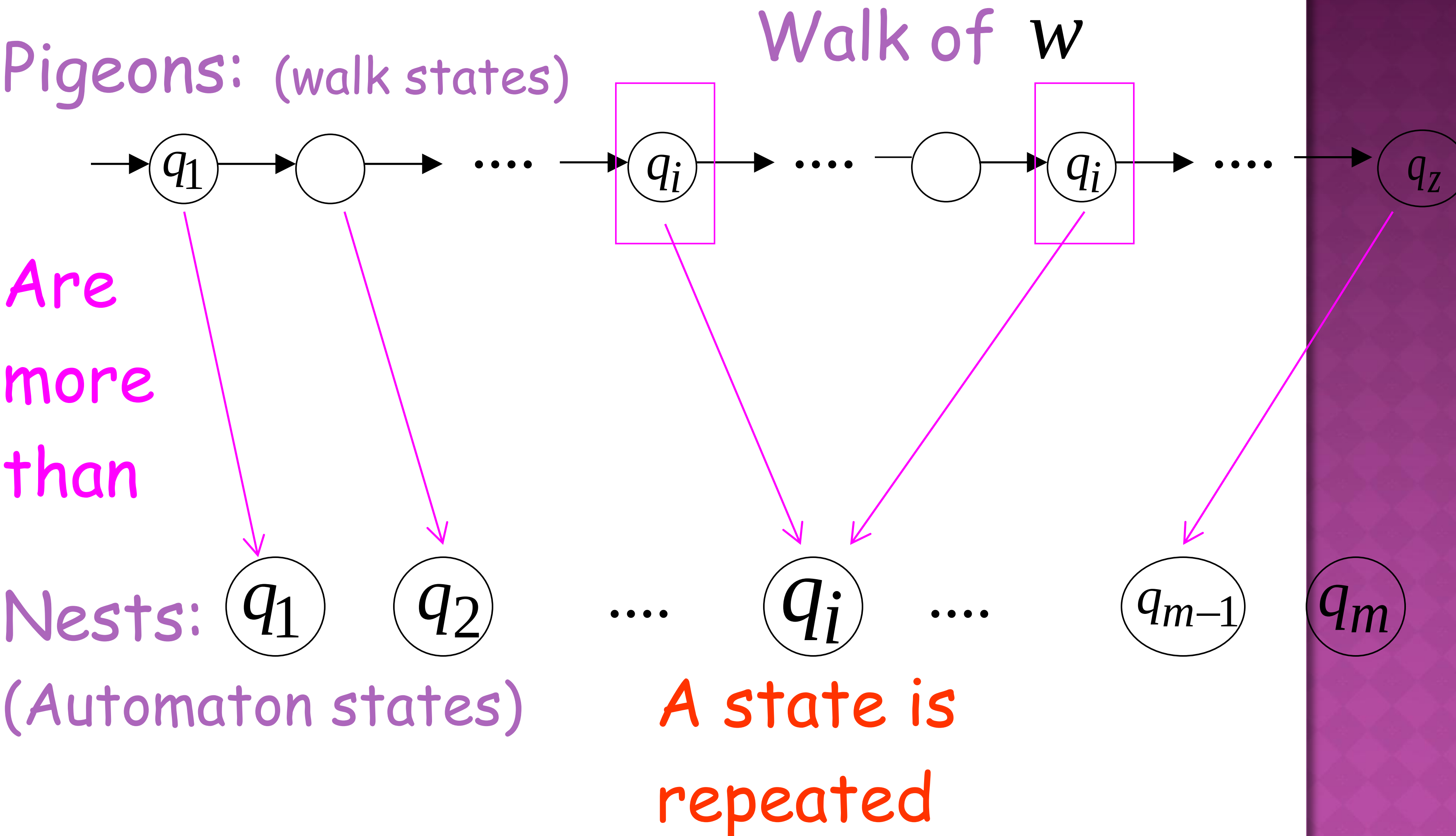
In General: If $|w| \geq$ #states of DFA,
by the pigeonhole principle,
a state is repeated in the walk w

Walk of $w = \sigma_1 \sigma_2 \cdots \sigma_k$





$|w| \geq \#\text{states of DFA} = m$

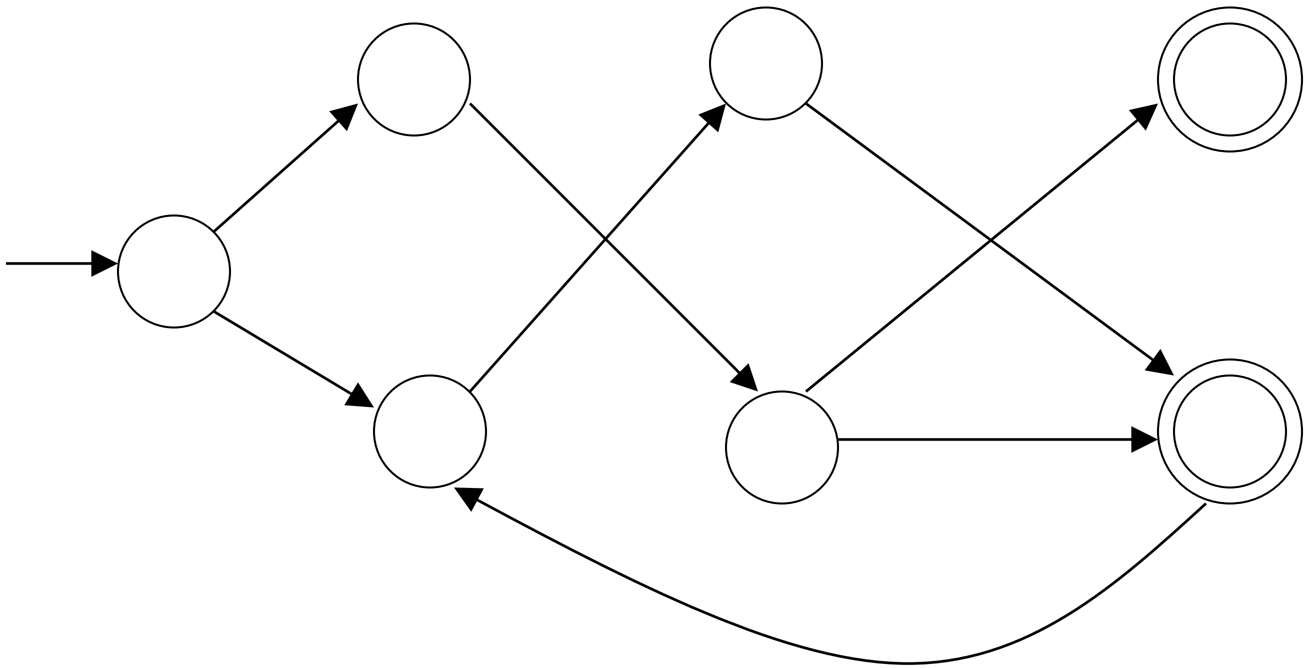


THE PUMPING LEMMA

Take an infinite regular language L

(contains an infinite number of strings)

There exists a DFA that accepts L



m states

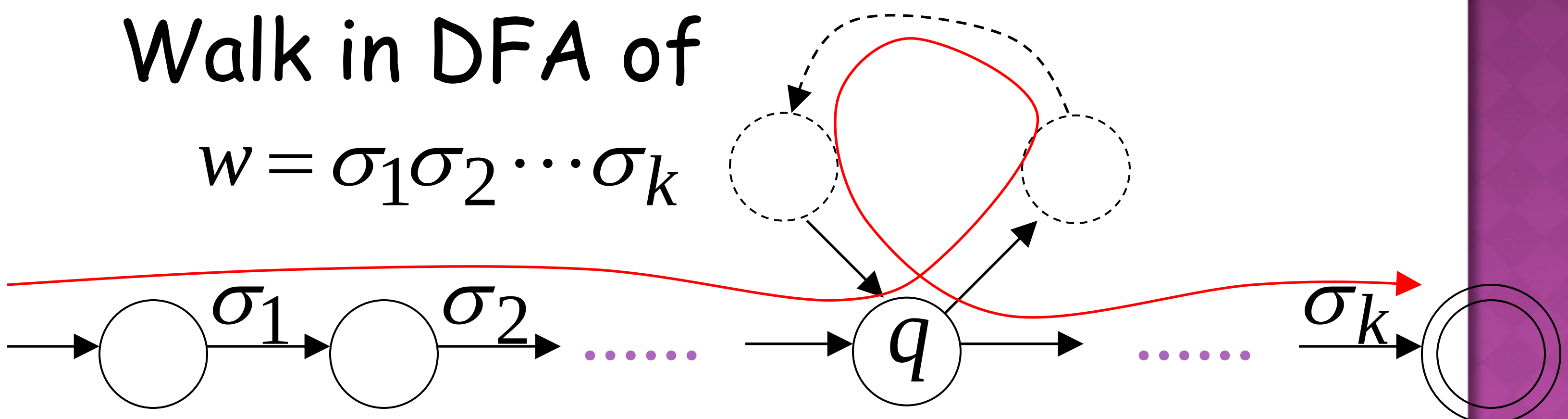
Take string $w \in L$ with $|w| \geq m$

(number of states of DFA)

then, at least one state is repeated in the walk of w

Walk in DFA of

$w = \sigma_1 \sigma_2 \cdots \sigma_k$

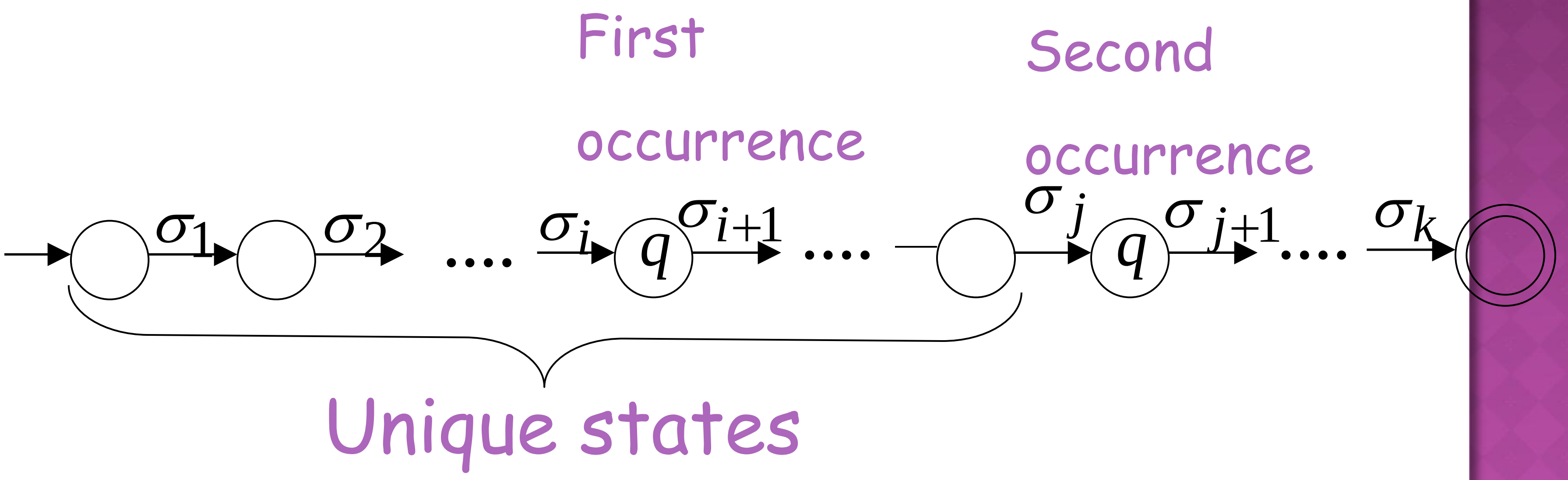


Repeated state in DFA

There could be many states repeated

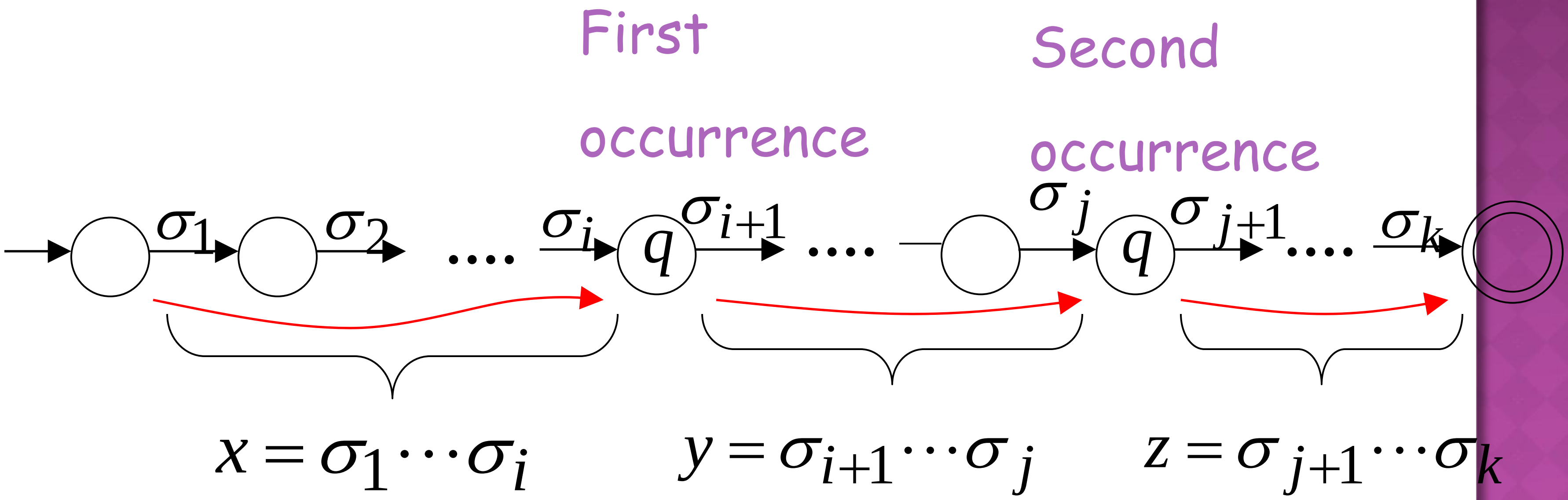
Take q to be the first state repeated

One dimensional projection of walk w:



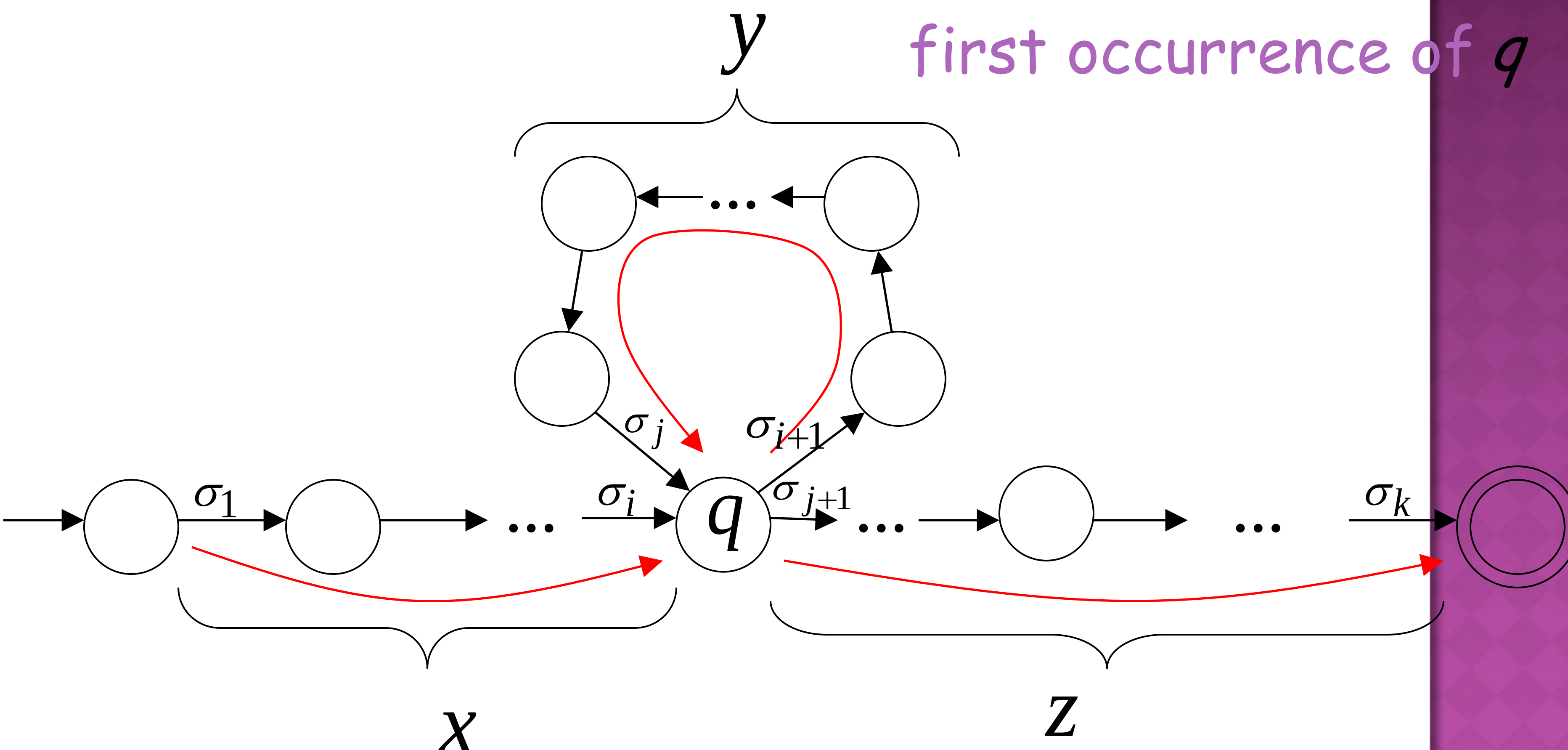
We can write $w = xyz$

One dimensional projection of walk w:

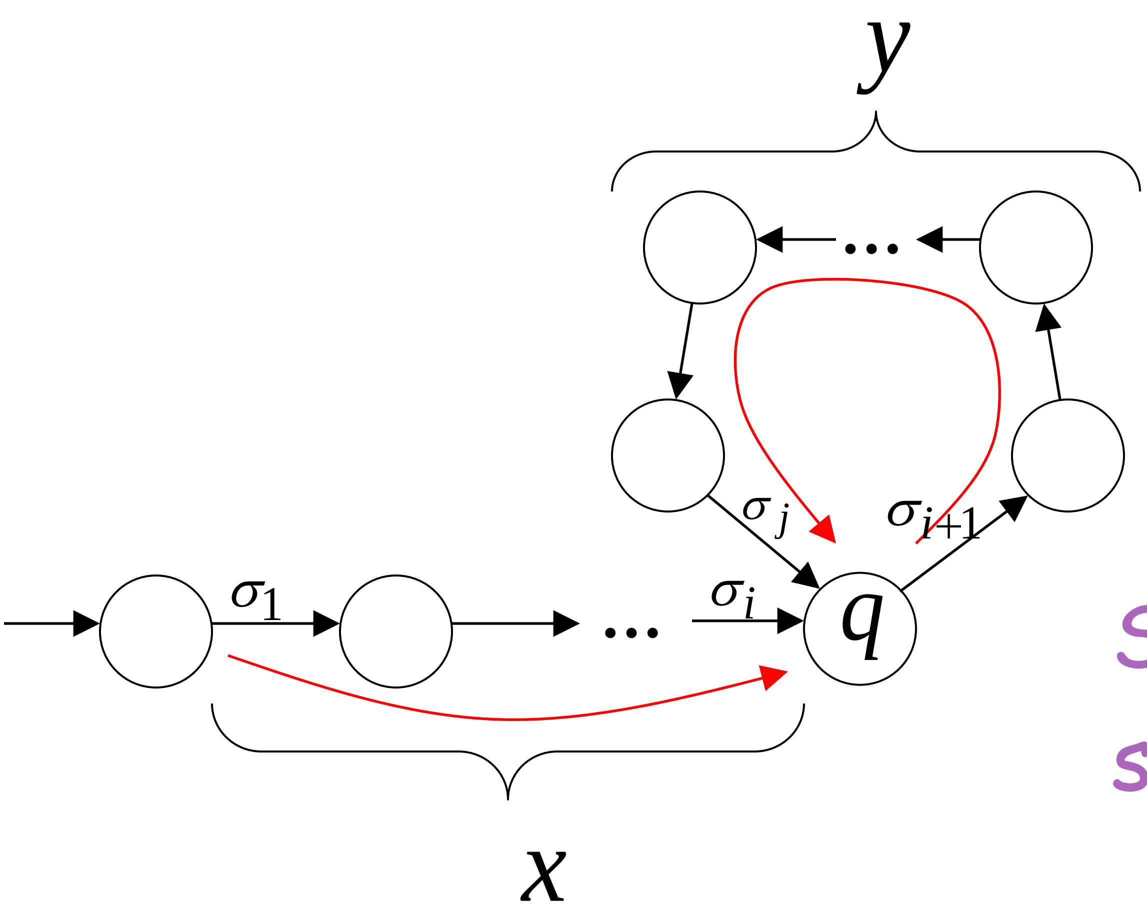


In DFA: $w = x\ y\ z$

contains only first occurrence of q



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



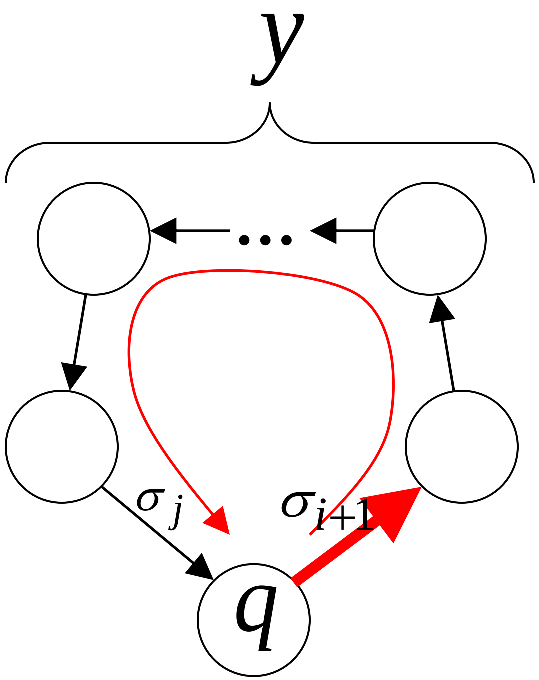
Unique States

Since, in xy no state is repeated

(except q)

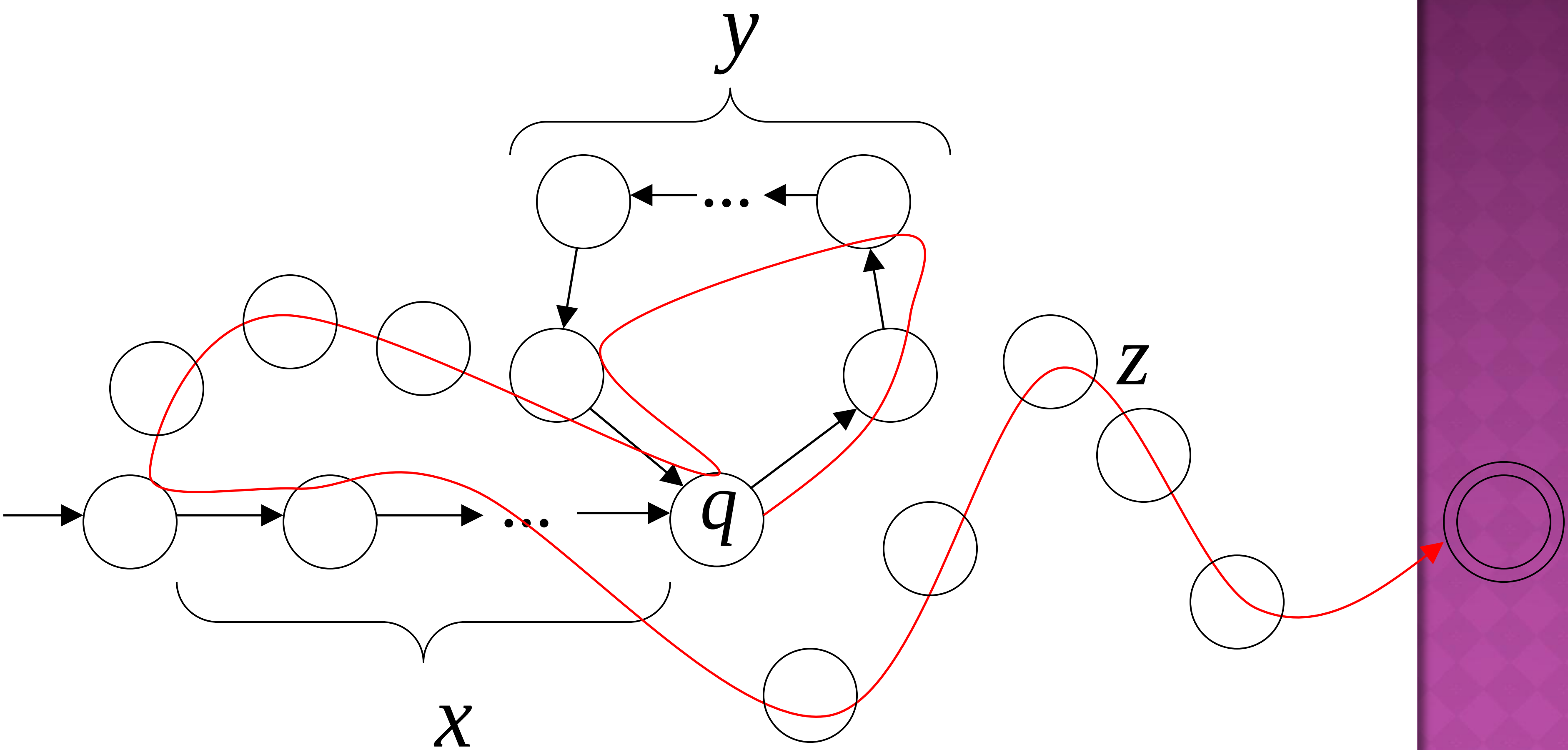
Observation: length $|y| \geq 1$

Since there is at least one transition in loop



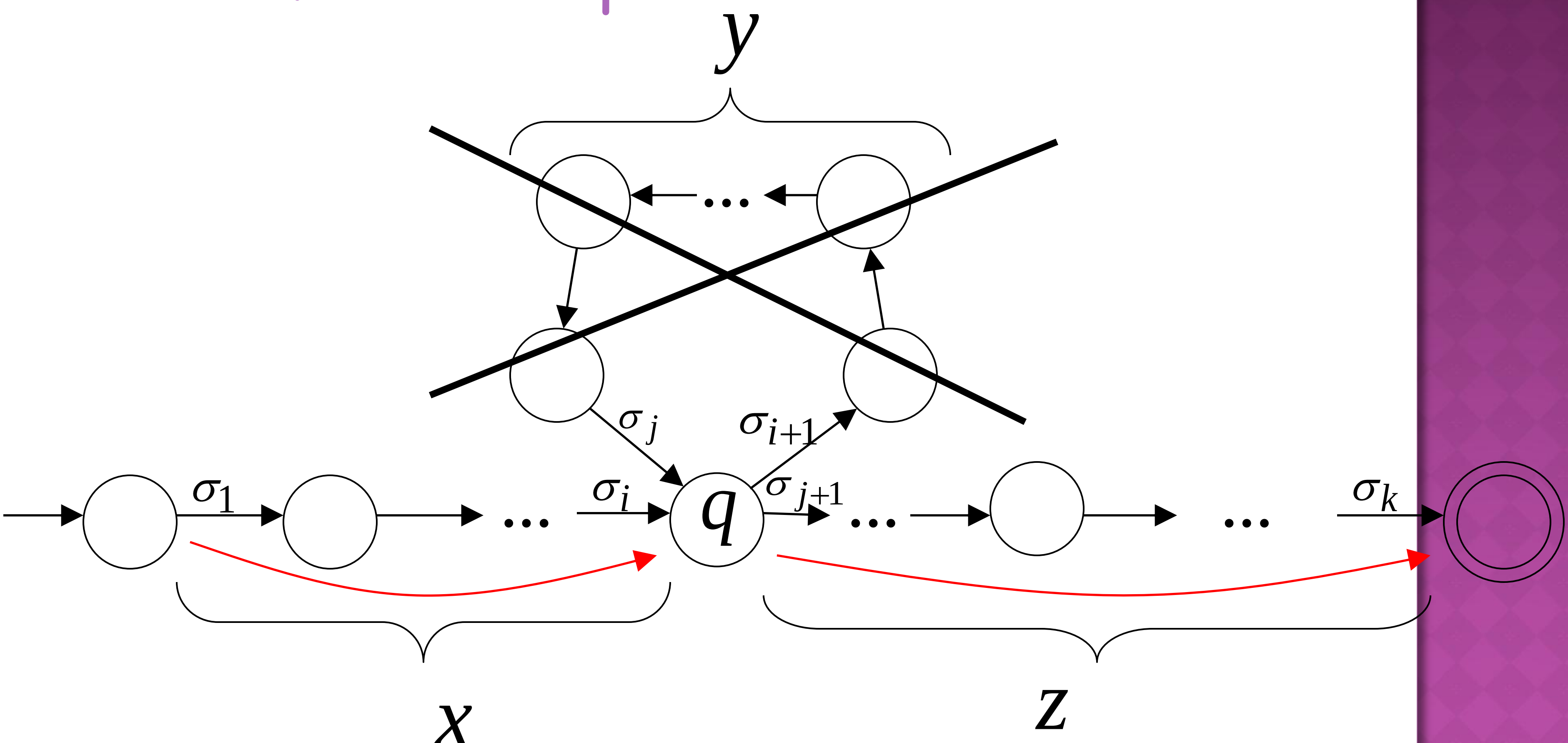
We do not care about the form of string z

z may actually overlap with the paths of x and y



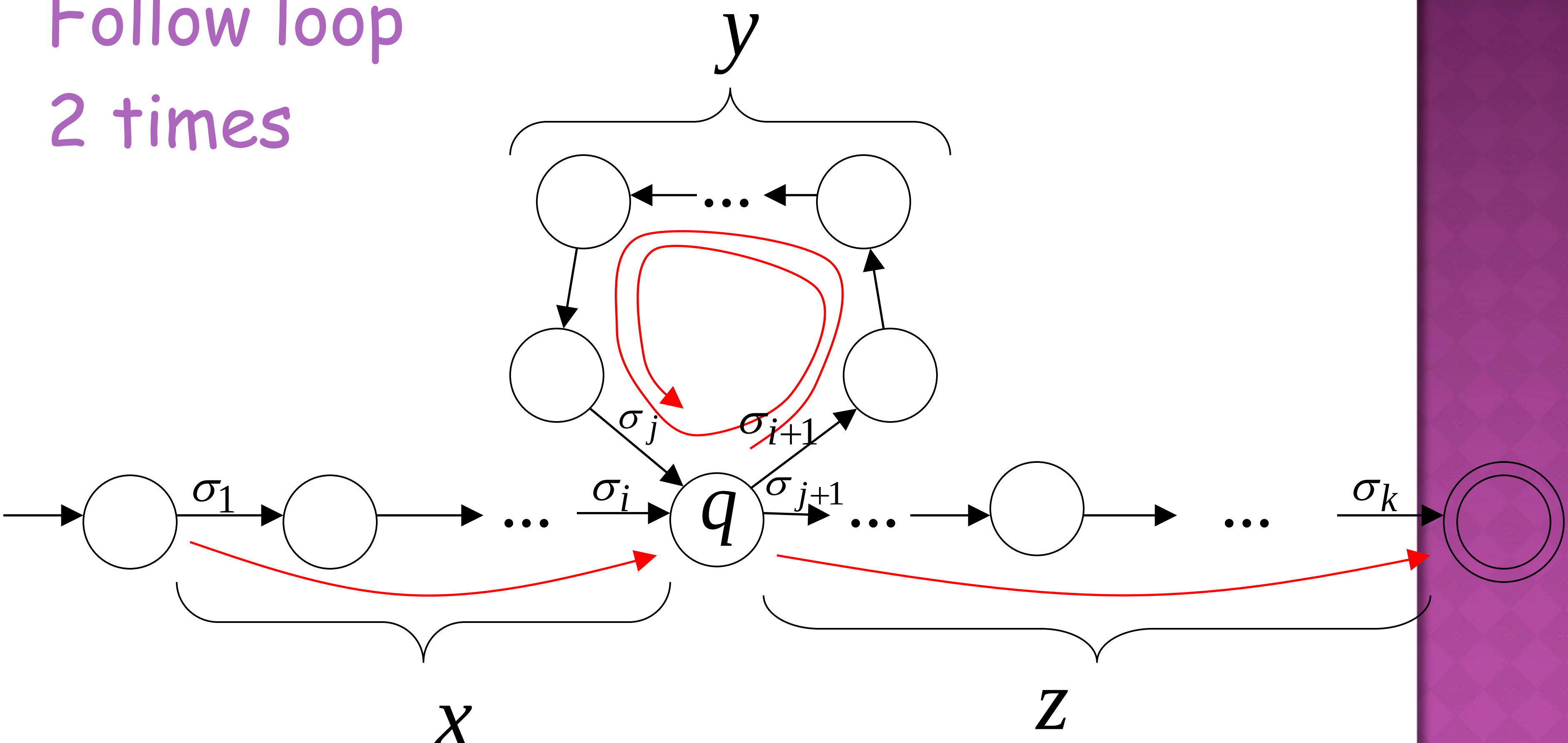
Additional string: The string $x\ z$ is accepted

Do not follow loop



Additional string: The string $x\ y\ y\ z$ is accepted

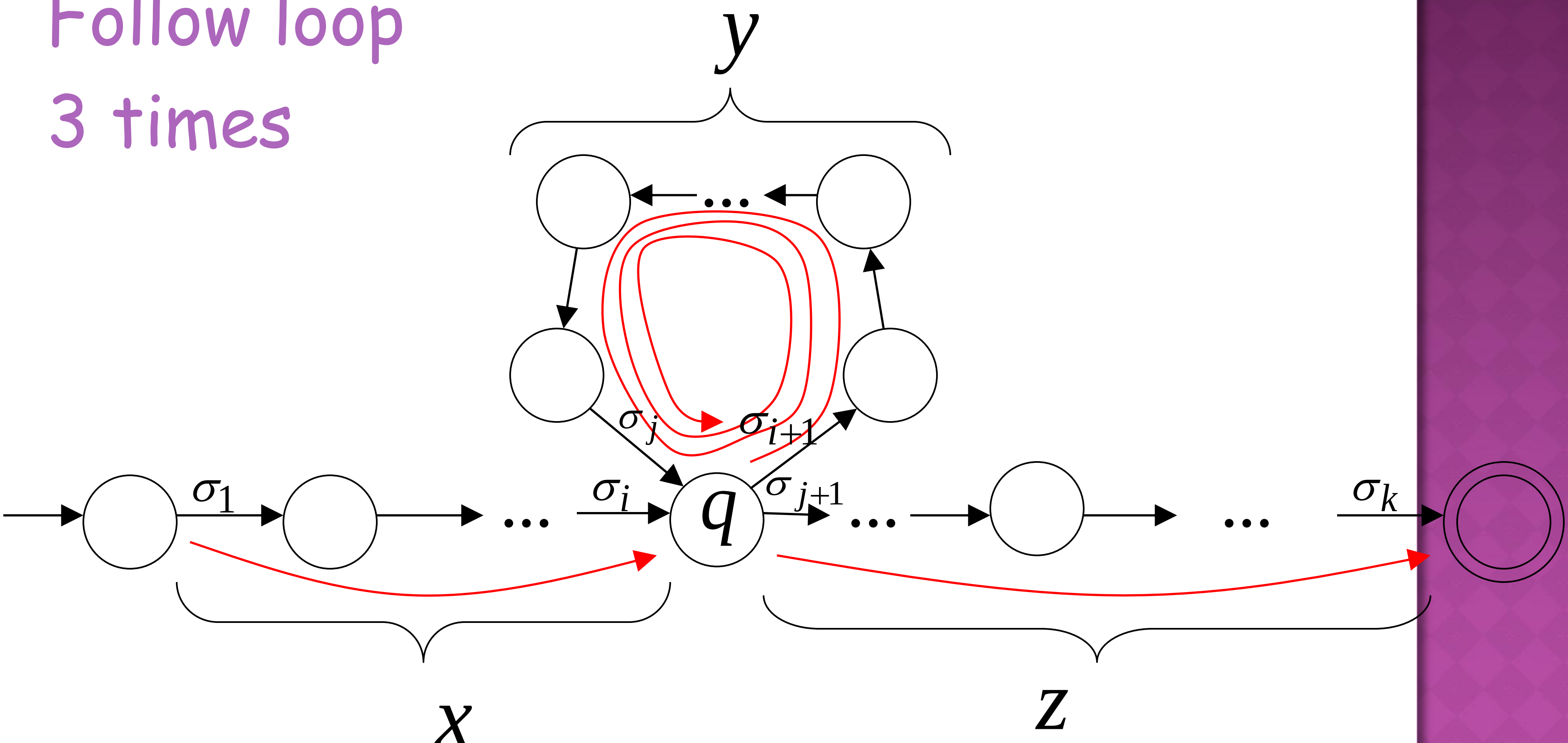
Follow loop 2 times



Additional string:

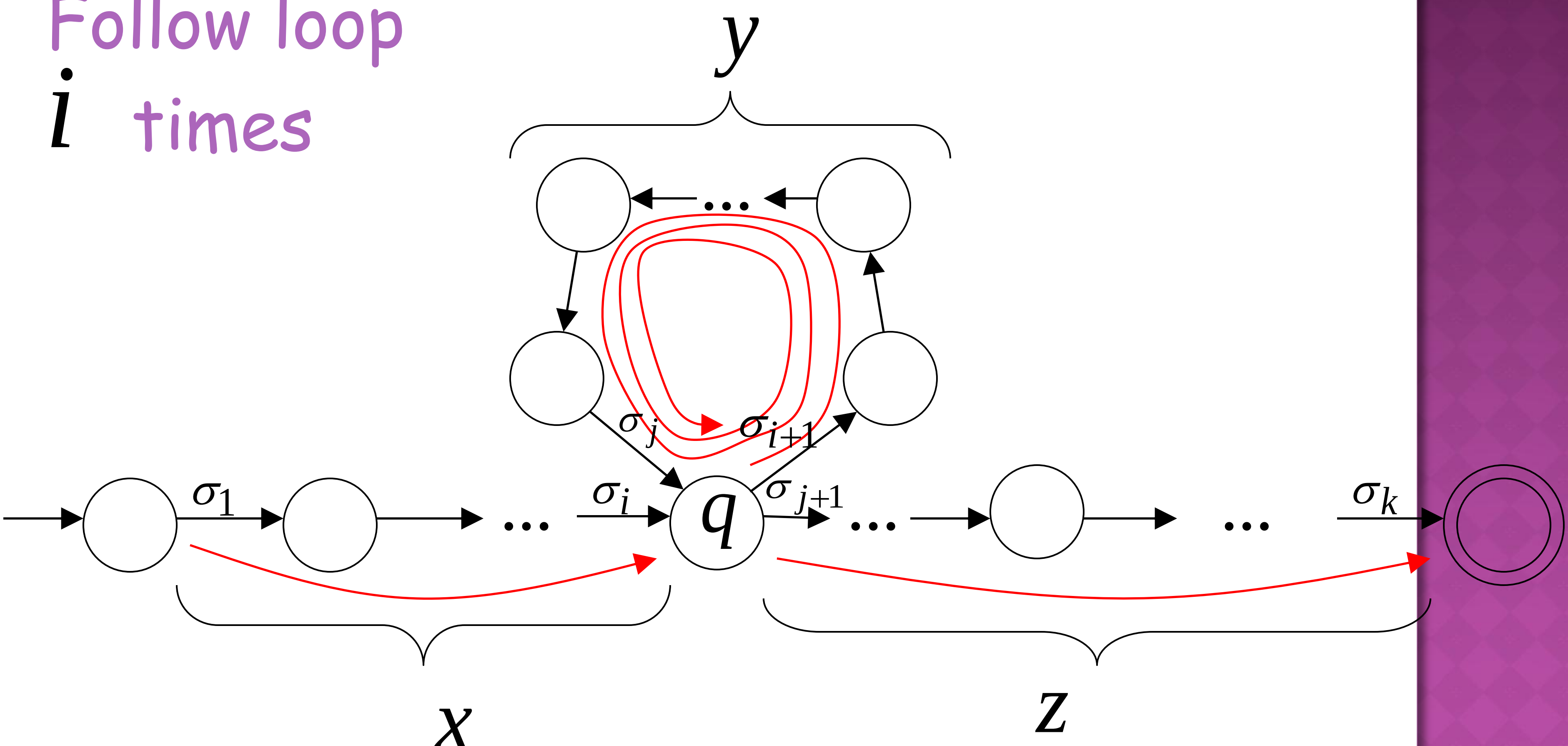
The string $x\,y\,y\,y\,z$ is accepted

Follow loop 3 times



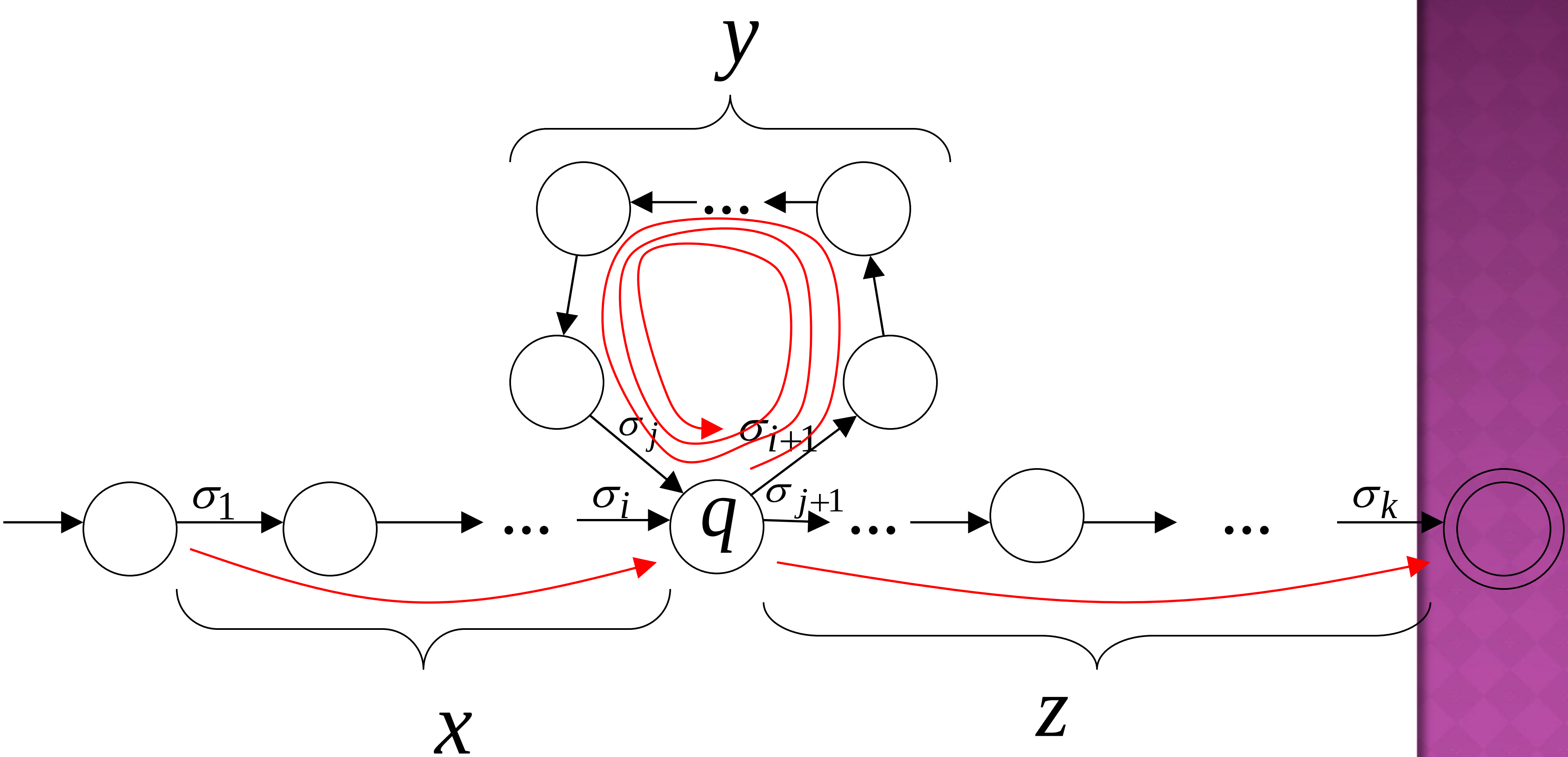
In General: The string $x\,y^i\,z$ is accepted $i = 0, 1, 2, \ldots$

Follow loop i times



Therefore: $x\,y^i\,z \in L \qquad i = 0, 1, 2, \ldots$

Language accepted by the DFA



- Given a infinite regular language L

THE PUMPING LEMMA:

- there exists an integer m (critical length)
- 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: $x\, y^i\, z \in L$ $\quad i = 0, 1, 2, \ldots$

In the book:

Critical length m = Pumping length p