Pumping Lemma for Regular Sets

Beulah A.
AP/CSE

Introduction

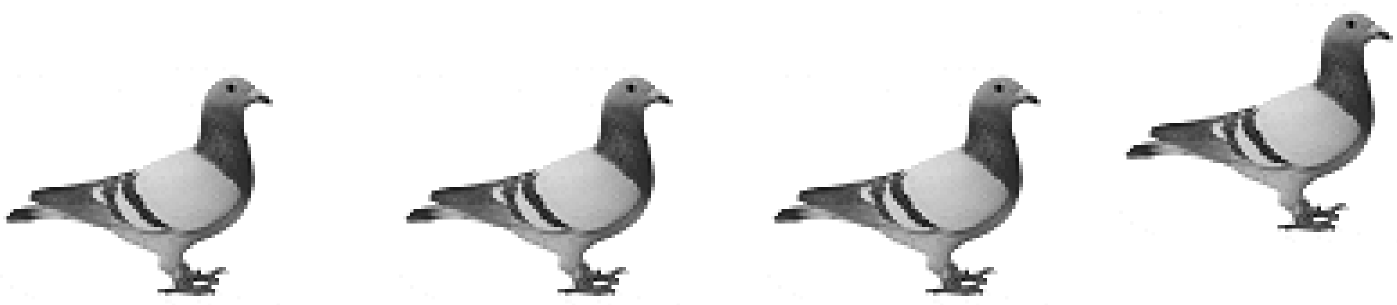
- A Regular language is a formal language that can be expressed using a regular expression
- A regular language satisfies the following equivalent properties:
  - it is the language accepted by a nondeterministic finite automaton
  - it is the language accepted by a deterministic finite automaton
  - it can be generated by a regular grammar
  - it can be generated by a prefix grammar
  - it can be accepted by a read-only Turing machine
- Regular set is a set of strings of a Regular Language
- For every regular language there is a FA that accepts the language

The Pigeonhole Principle

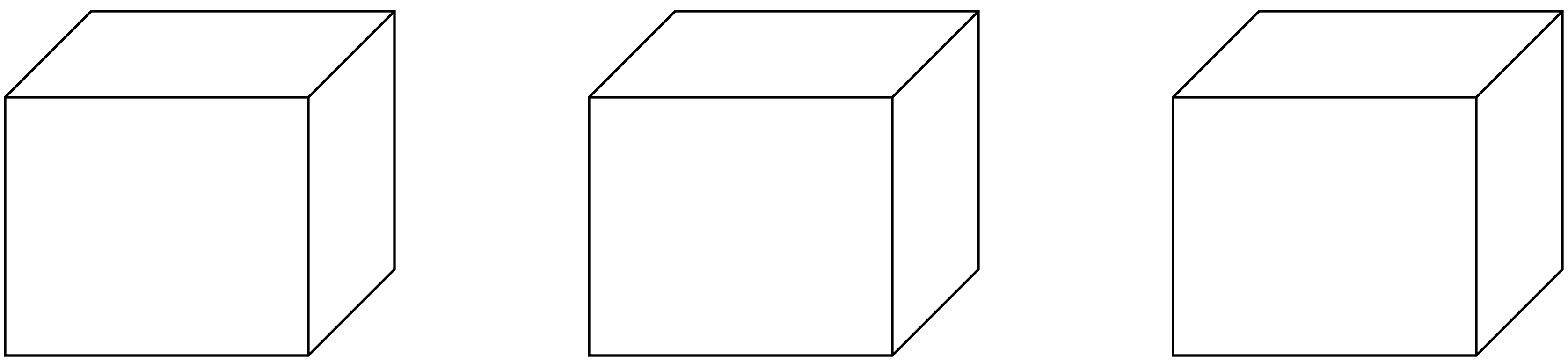
- If you put n pigeons into m pigeonholes, and $n > m > 0$, then at least at least two pigeons are in the same pigeonhole.

The Pigeonhole Principle

4 pigeons

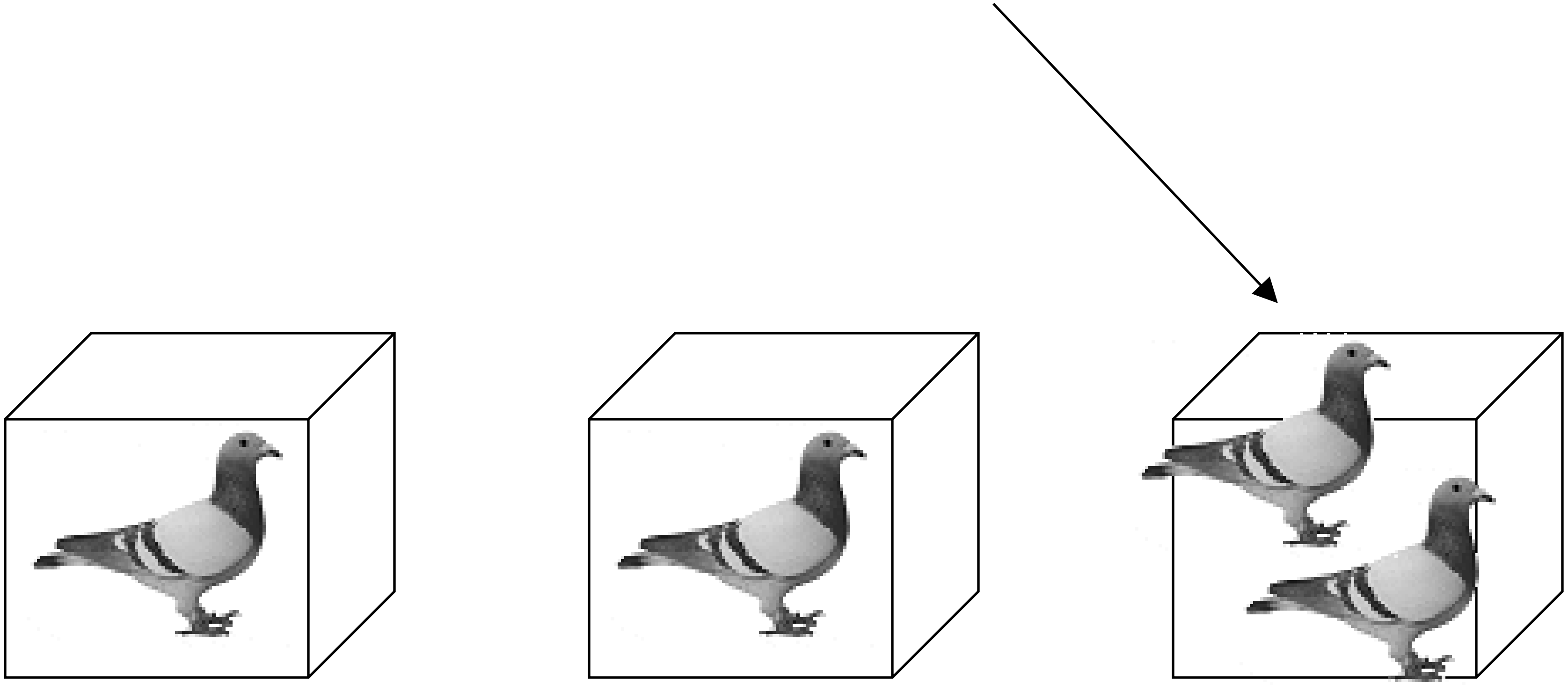


3 pigeonholes



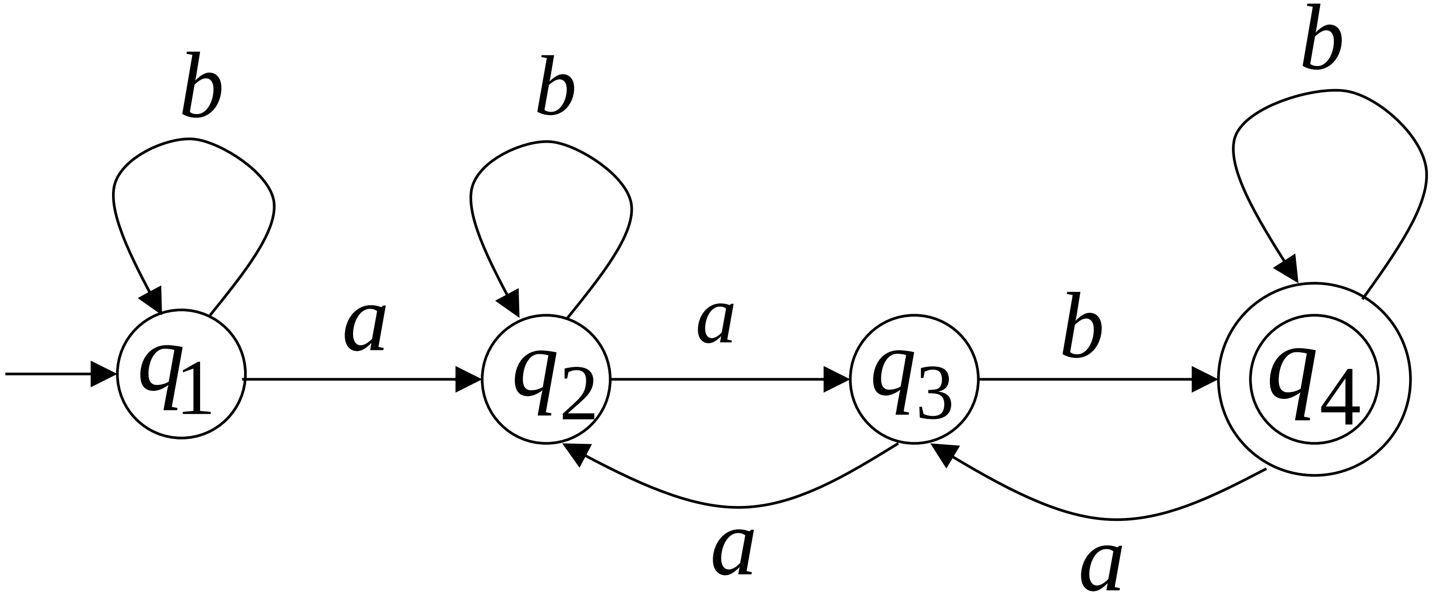
The Pigeonhole Principle

A pigeonhole must
contain at least two pigeons



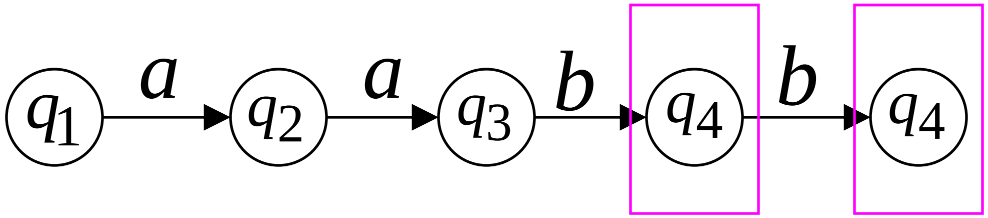
Pigeonhole Principle & DFAs

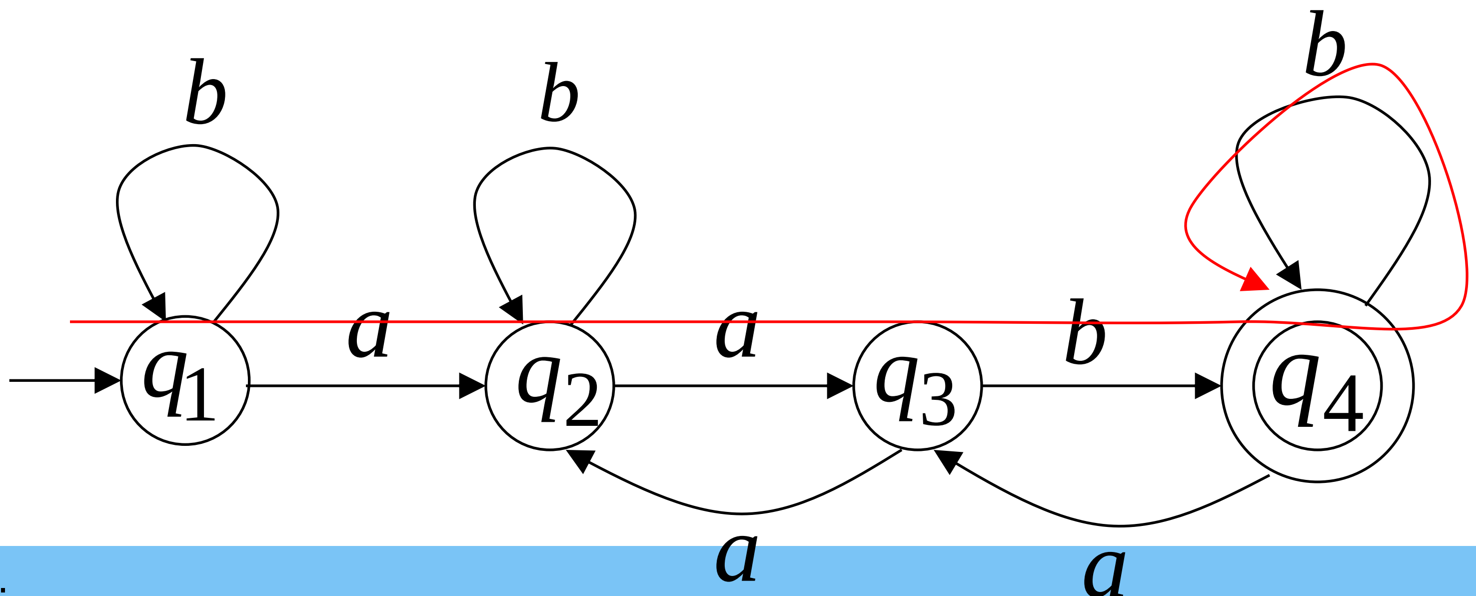
Consider a DFA with 4 states



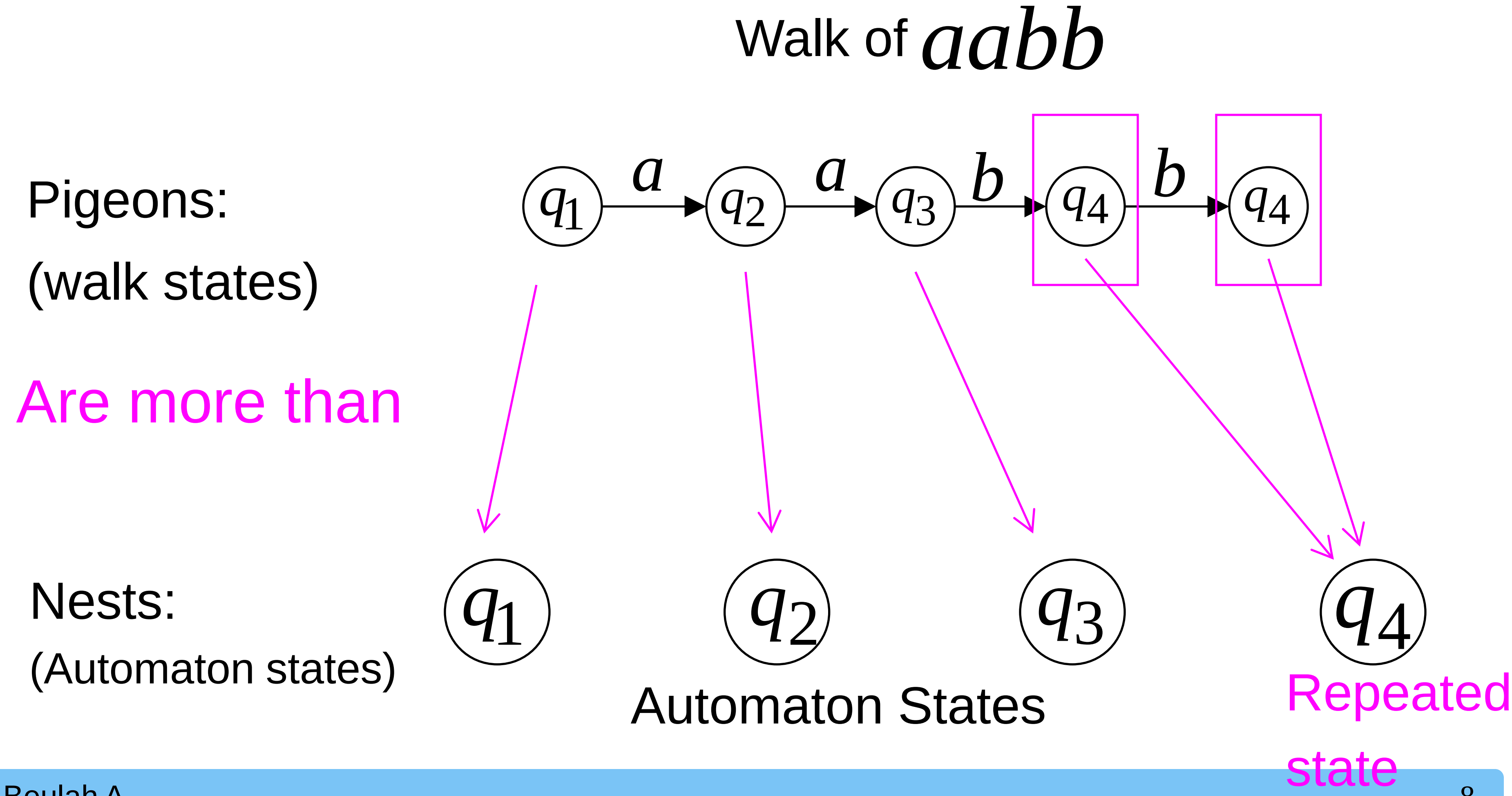
- Consider the walk of a “long” string: aabb (>=4)
-

A state is repeated in the walk of $aabb$





- The state is repeated as a result of the pigeonhole principle



Pumping Lemma

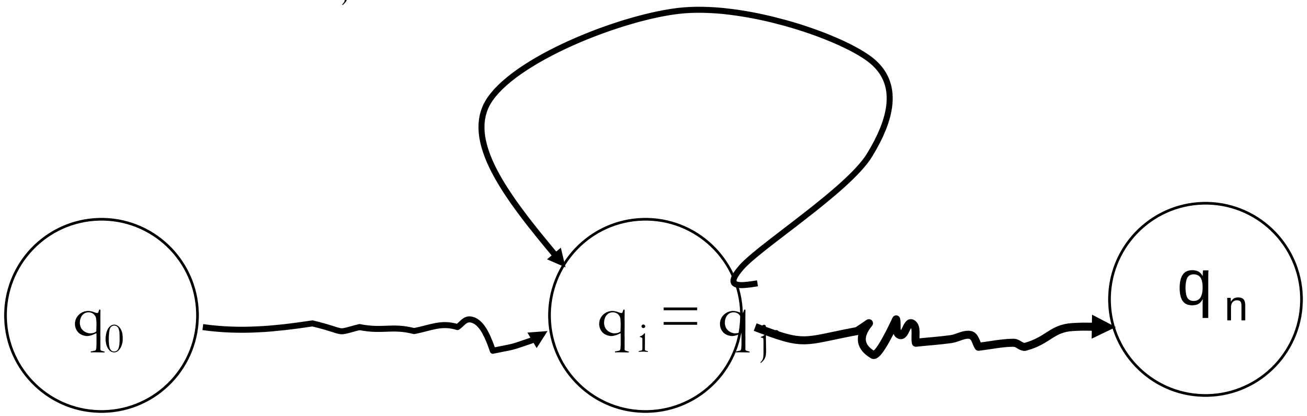
- Describes an essential property of all regular languages
- For a particular language, any sufficiently long string in the language contains a section, or sections, that can be removed, or repeated any number of times(pumping), with the resulting string remaining in that language
- The pumping lemma is often used to prove that a particular language is non-regular

Pumping Lemma

- Let L be a regular language. Then there is a constant n (which depends on L/ number of states in FA) such that for every string w in L such that $|w| \geq n$, we can break w into three strings, $w = xyz$, such that $y \neq \varepsilon$ ie $|y| > 0$, $|xy| \leq n$, and for all $i \geq 0$, xy^iz is also in L.
- Proof
  - Let n be $|Q|$.
  - If $w \in L$ and $|w| \geq n$. Let $w = a_1 a_2 \ldots a_m$, where $m \geq n$.
  - $\delta(q_0, a_1 a_2 \ldots a_i) = q_i$, $i = 1, 2, \ldots, m$.

Pumping Lemma

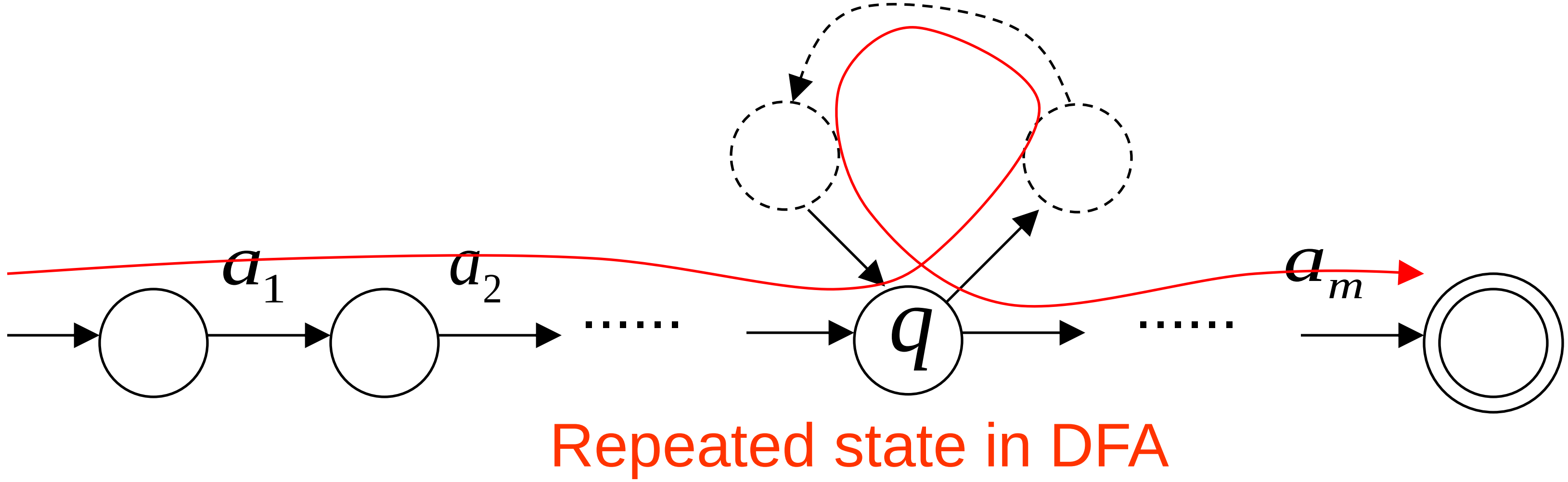
- Since there are only n states in Q and $m \geq n$, by the pigeon hole theorem there are two states of $q_0, q_1, q_2, \ldots,$ and q_n are same, say $0 \leq i < j \leq n$ and $q_i = q_j$.



Pumping Lemma

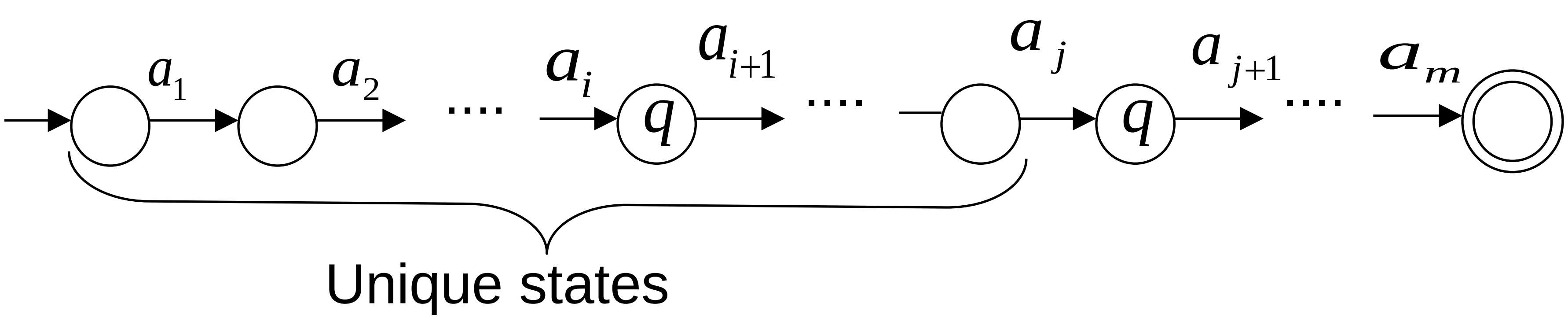
- If w $\in$L and $|w| \geq$ n, then, at least one state is repeated in the walk of w

Walk in DFA of $w = a_1 a_2 \cdots a_m$



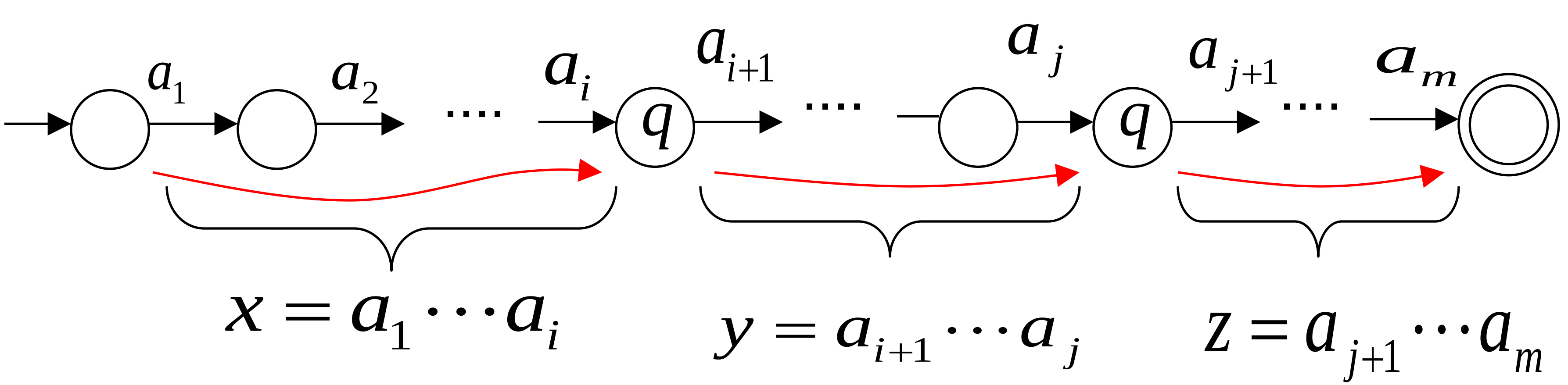
Pumping Lemma

- There could be many states repeated
- Take q to be the first state repeated



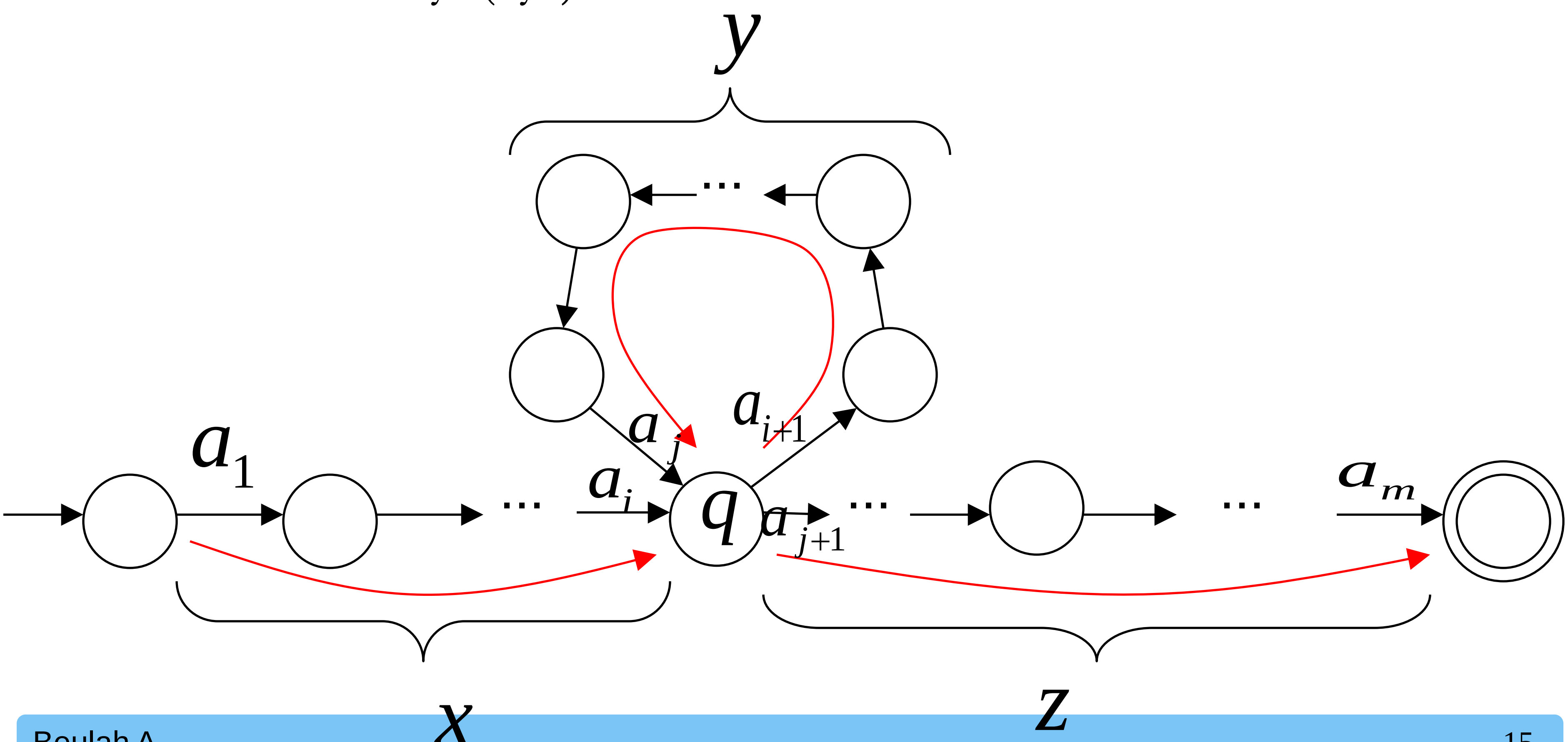
Pumping Lemma

- We can write w = xyz



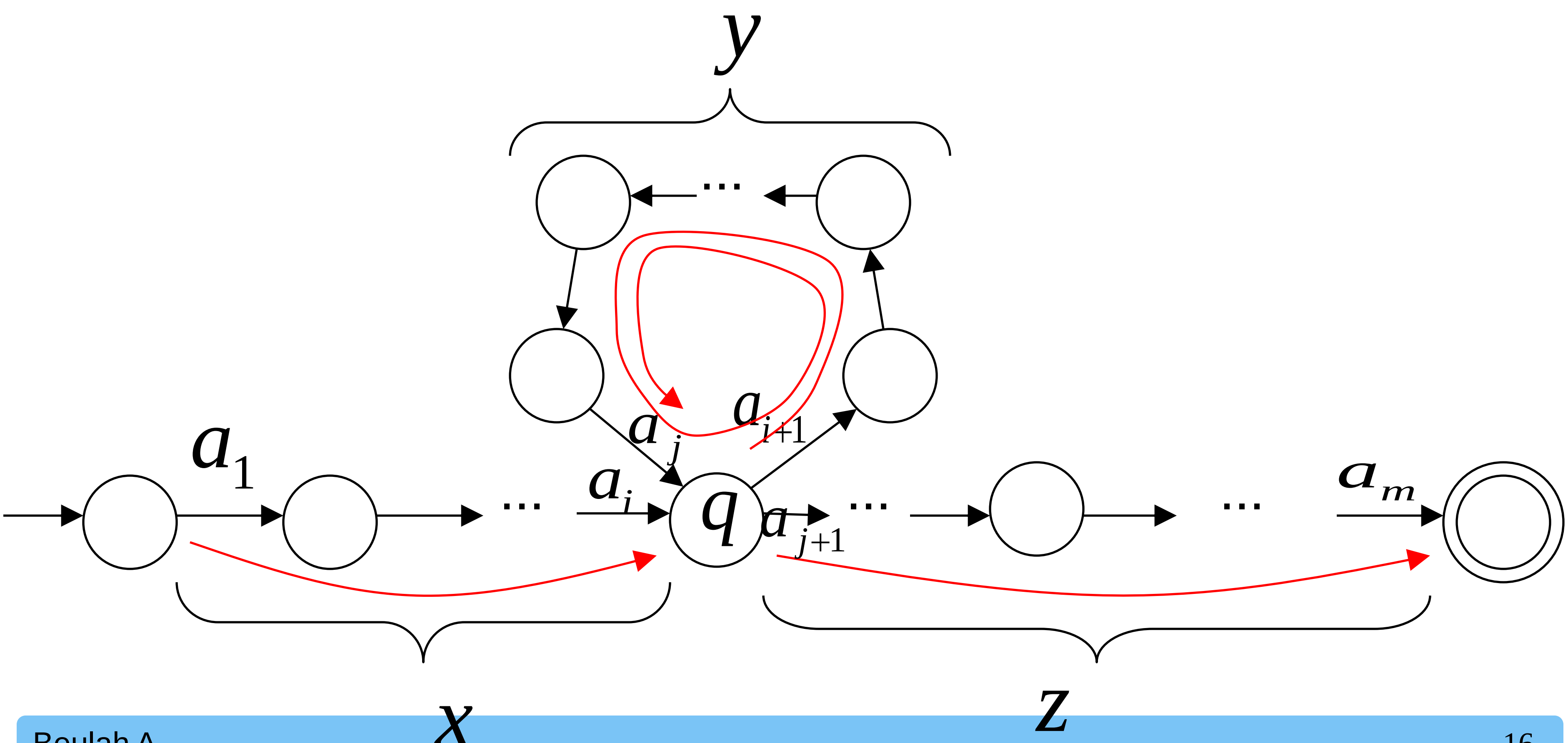
Pumping Lemma

- In DFA: write w = xyz
- 1 occurrence of y (xyz)



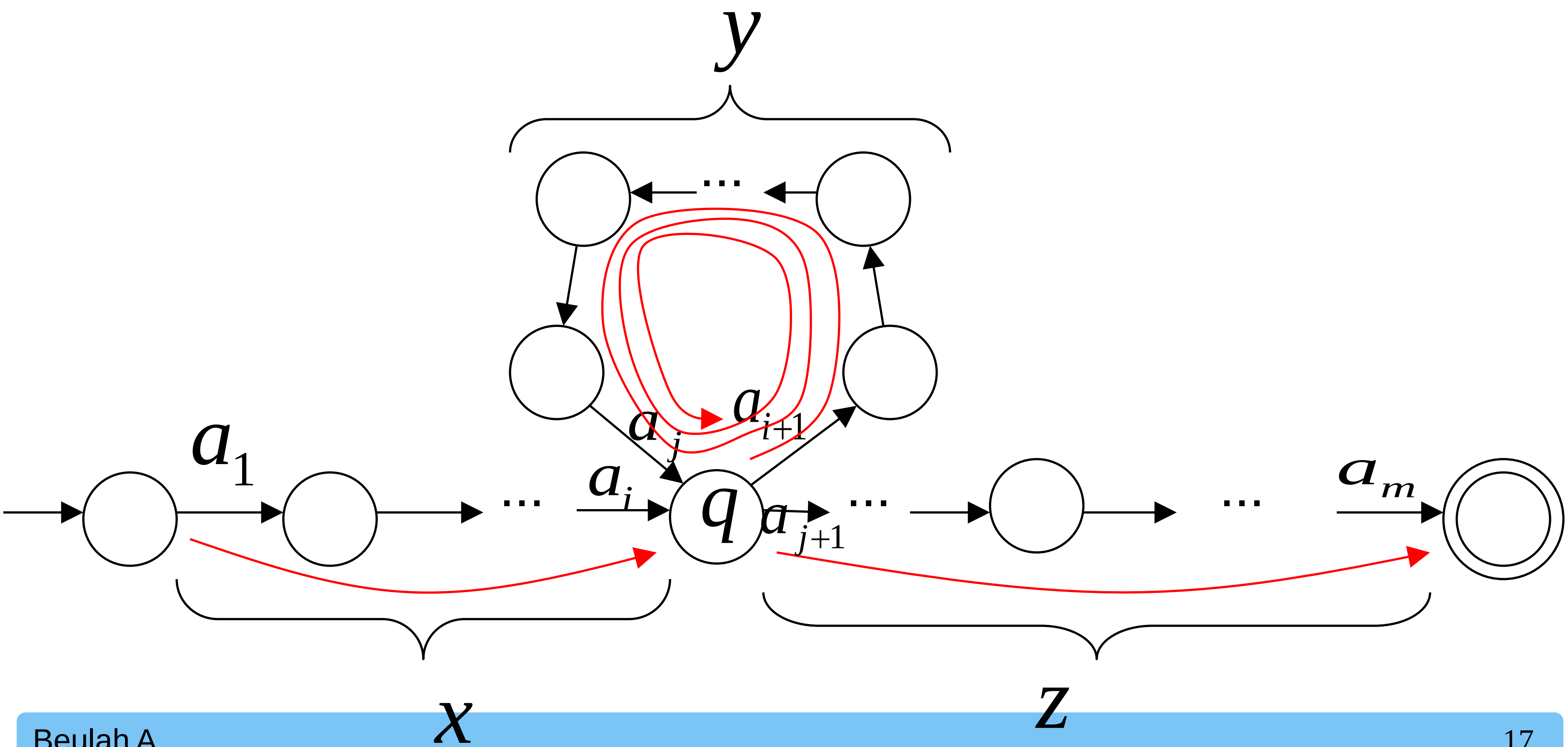
Pumping Lemma

- 2 occurrences of y (xyyz)



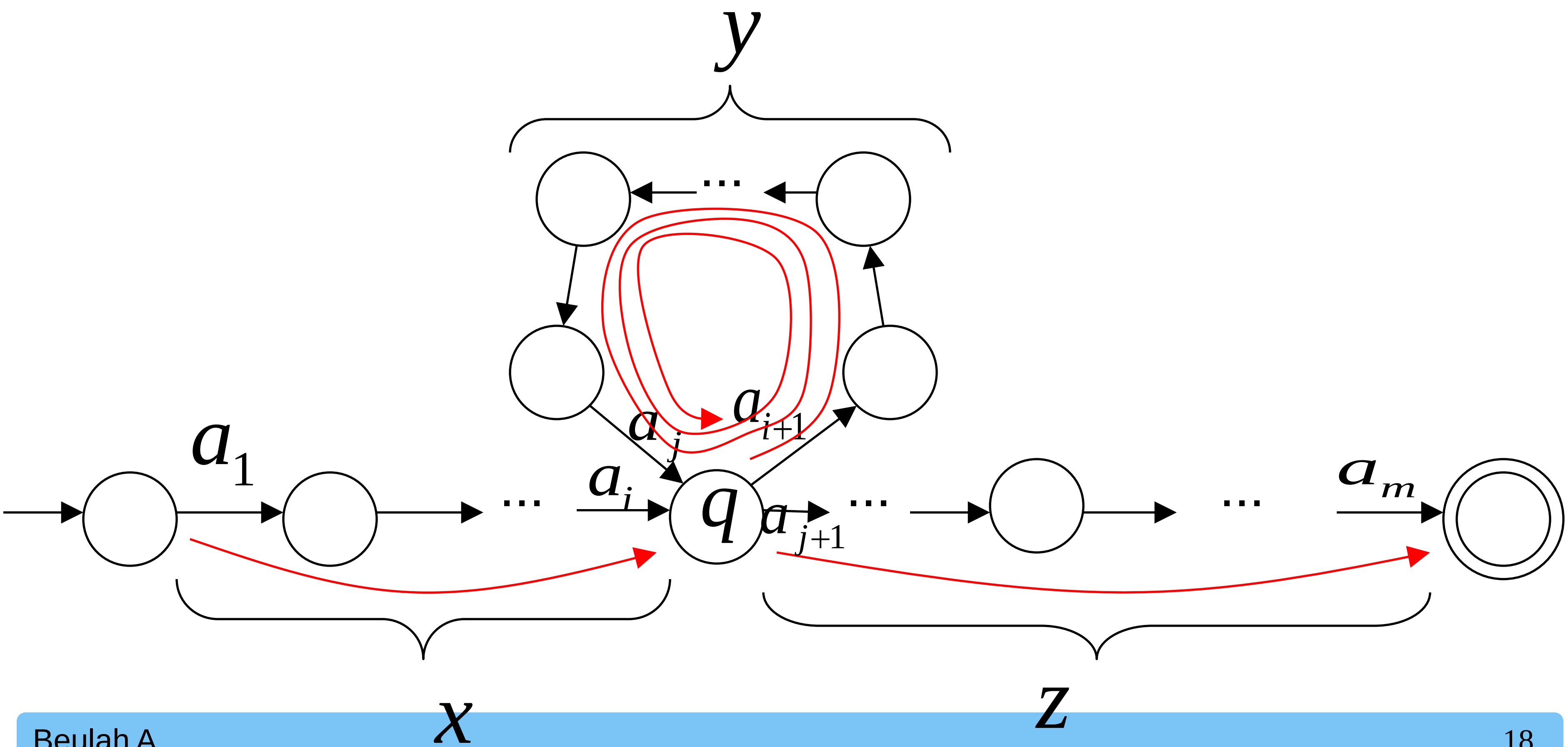
Pumping Lemma

- 3 occurrences of y (xyyyz)



Pumping Lemma

- Many occurrences of y (xy^iz)
- xy^iz is also in L



Pumping Lemma

- $\delta(q_0, a_1a_2\ldots a_i) = q_i = q_j$
- $\delta(q_i, a_{i+1}..a_j) = q_i$, and
- $\delta(q_j, a_{j+1}\ldots a_m) = q_n$

It is obvious that $\delta(q_i, y^i) = q_i$ for $i \geq 0$.

So, if the FA accepts $w = xyz$, it also accepts xy^iz.

Application

- Useful to prove a language L is not a regular set
- Method
  - Select an arbitrary ‘n’
  - Choose a string w in L where $|w| \geq n$
  - For any partition of $w = xyz$ such that
    - $|xy| \leq n$ and $|y| \geq 1$, show a contradiction;
    - i.e. show that there is a string xy^kz not in L;
    - k will depend on n, x, y, and z

Summary

- Definition of Pumping lemma – Regular Language
- Application of pumping lemma

Test Your Knowledge

- If we select a string w such that w∈L, and w=xyz. Which of the following portions cannot be an empty string?
  a) x
  b) y
  c) z
  d) all of the mentioned
- Which of the following one can relate to the given statement:
  Statement: If n items are put into m containers, with n>m, then atleast one container must contain more than one item.
  a) Pumping lemma
  b) Pigeon Hole principle
  c) Count principle
  d) None of the mentioned

Reference

- Hopcroft J.E., Motwani R. and Ullman J.D, “Introduction to Automata Theory, Languages and Computations”, Second Edition, Pearson Education, 2008