Lecture 9

Nonregular Languages

COT 4420

Theory of Computation

Section 4.3

Regular and Nonregular languages

- Regular languages

a^*b+a

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

$(a+b+c)aa(a+b+c)$

- Nonregular languages

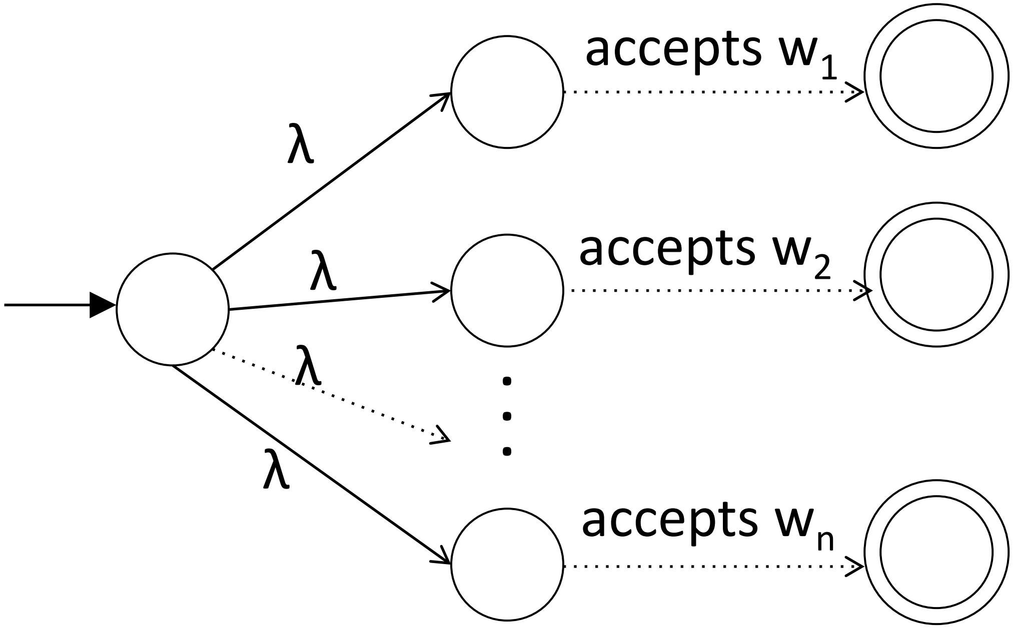
$\{a^n b^n : n \geq 0\}$

$\{waw^R : w \in \{a,b\}^*\}$

Regular and Nonregular languages

- All finite languages are regular

Suppose that $L = \{ w_1 , w_{2,} \dots w_n\}$



Nonregular Languages

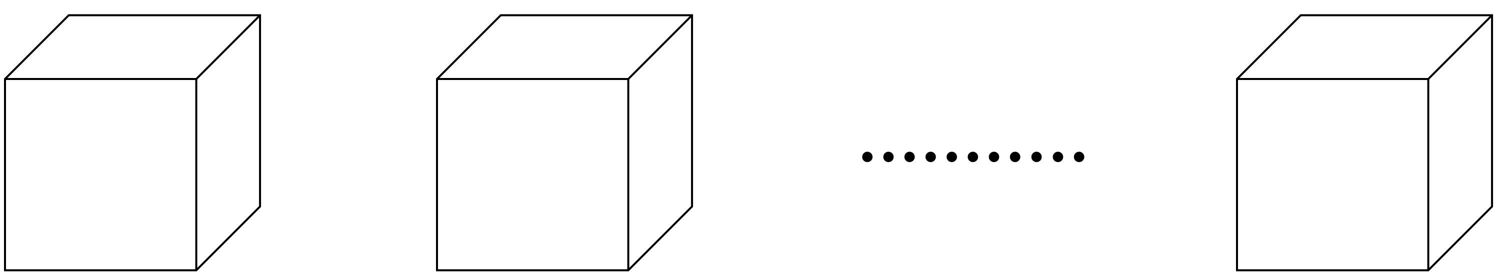
What about infinite regular languages or more generally non regular languages?

How can we prove in general that a languages L is not regular?

The pigeonhole principle

n pigeons

m pigeonholes $n > m$



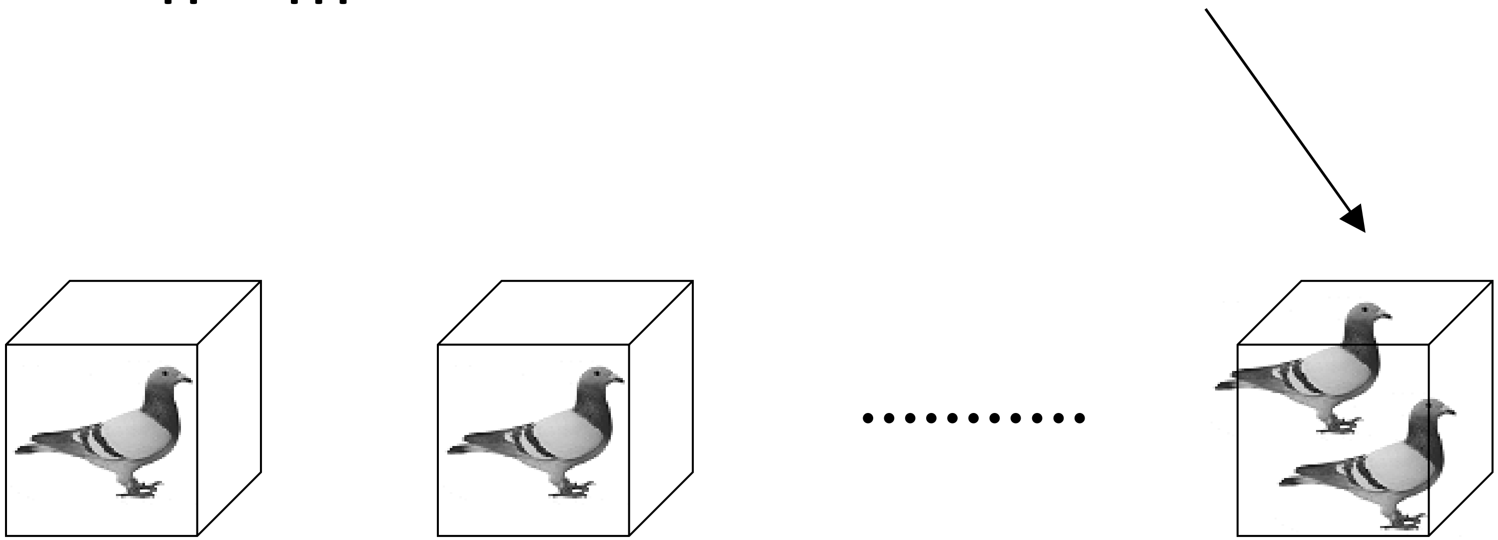
The pigeonhole principle

n pigeons

m pigeonholes

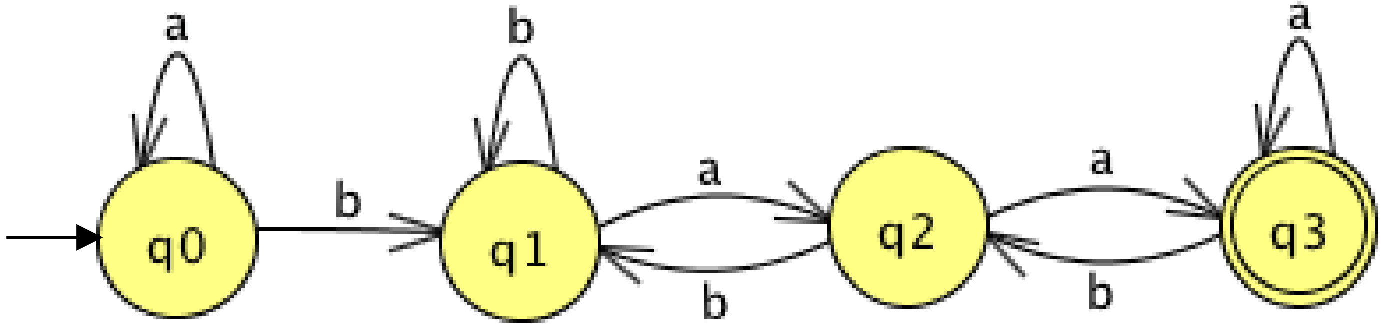
n > m

There is a pigeonhole with at least 2 pigeons



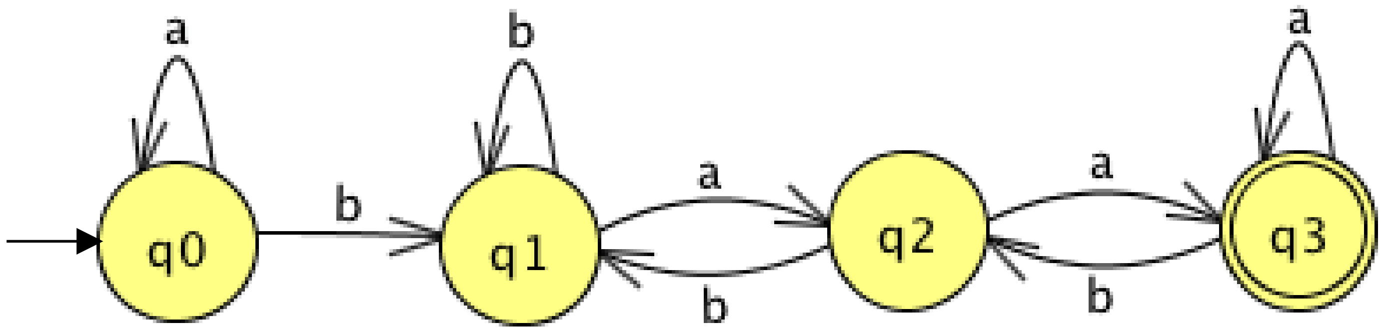
The pigeonhole principle and DFAs

- Suppose we have a DFA with 4 states



The pigeonhole principle and DFAs

- Suppose we have a DFA with 4 states



In a walk for baabb length > 4

$q_0 \rightarrow q_1 \rightarrow q_2 \rightarrow q_3 \rightarrow q_2 \rightarrow q_1$

State q_2 is repeated

The pigeonhole principle and DFAs

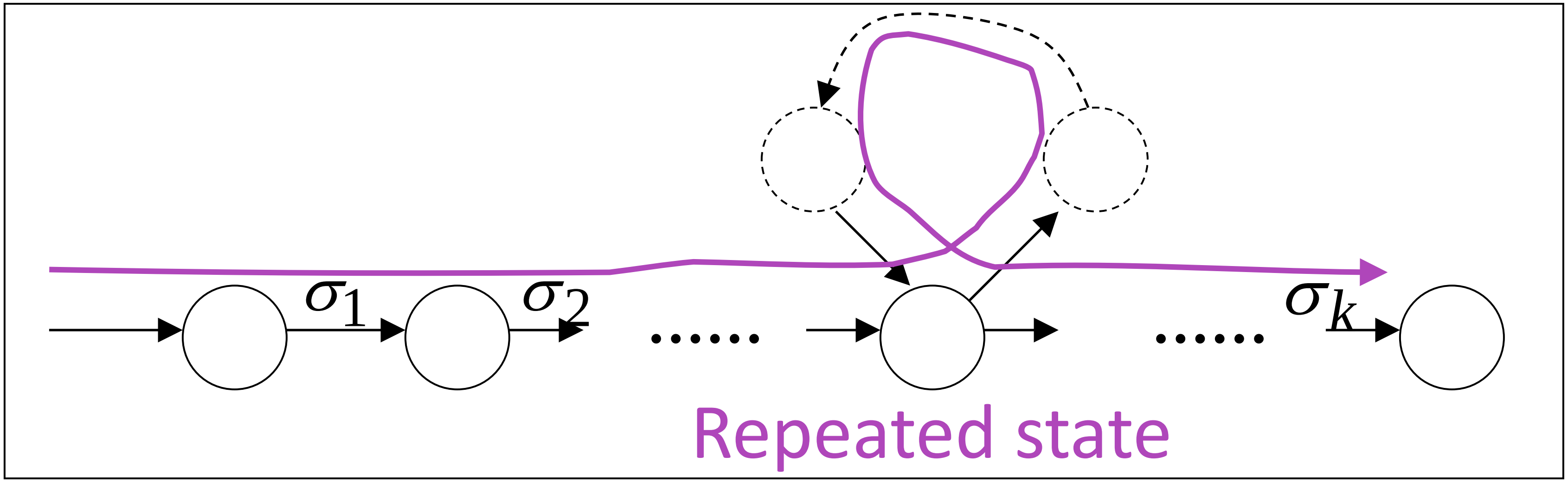
- The state is repeated as the result of pigeonhole principle.

Nests
(Automaton states)

Pigeons
(walk states)

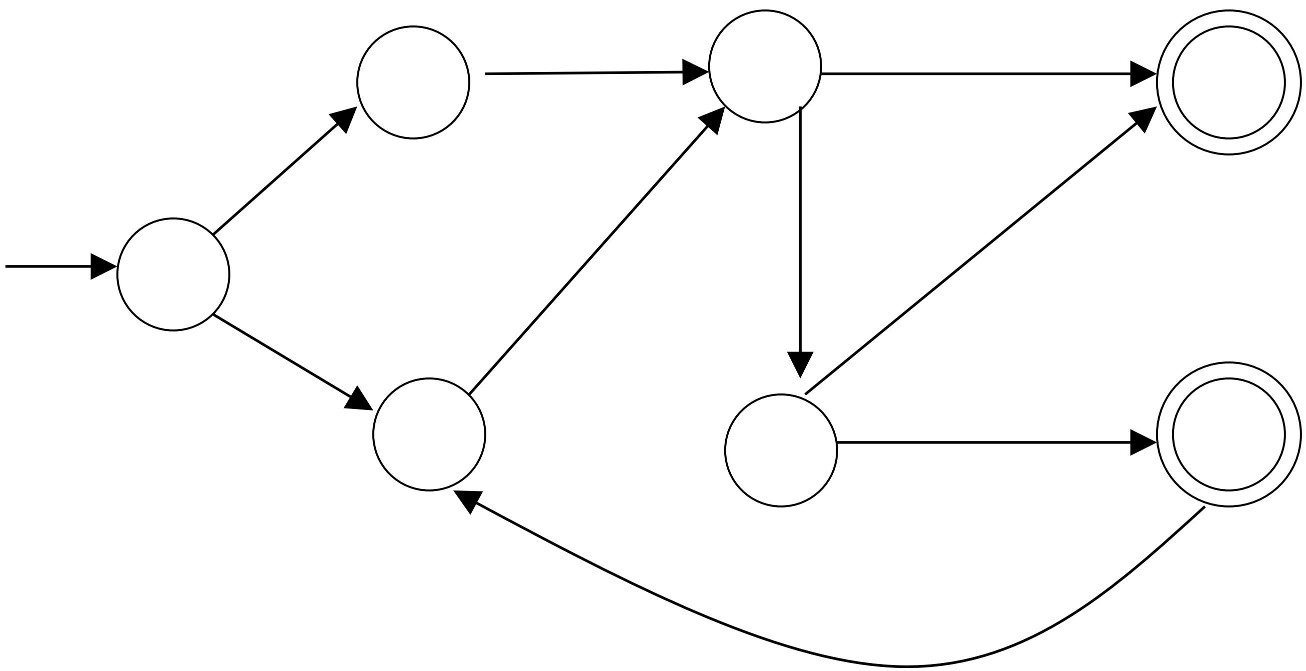
$q_0 \rightarrow q_1 \rightarrow q_2 \rightarrow q_3 \rightarrow q_2 \rightarrow q_1$

In general if |w| ≥ number of states of DFA by the pigeonhole principle, a state is repeated in the walk



Pumping Lemma

Suppose we have an infinite regular language L
We know that there exists a DFA for L (with m states) and a string |w| ≥ m accepted by the DFA.

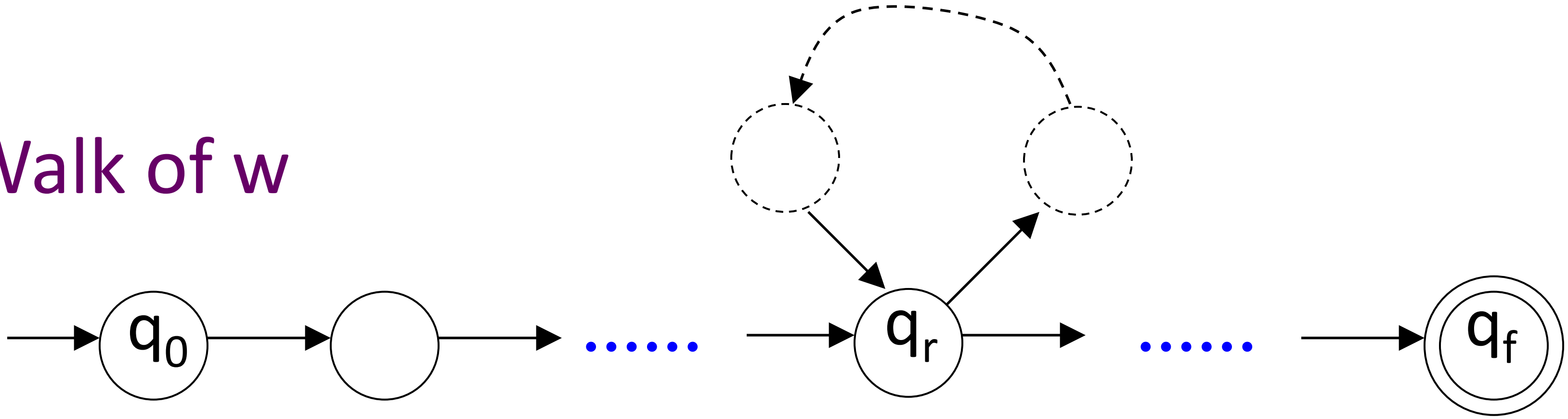


Pumping Lemma

Choose a string $w \in L$ such that $|w| \geq m$

Then from pigeonhole principle a state is repeated in the walk for w.

Walk of w

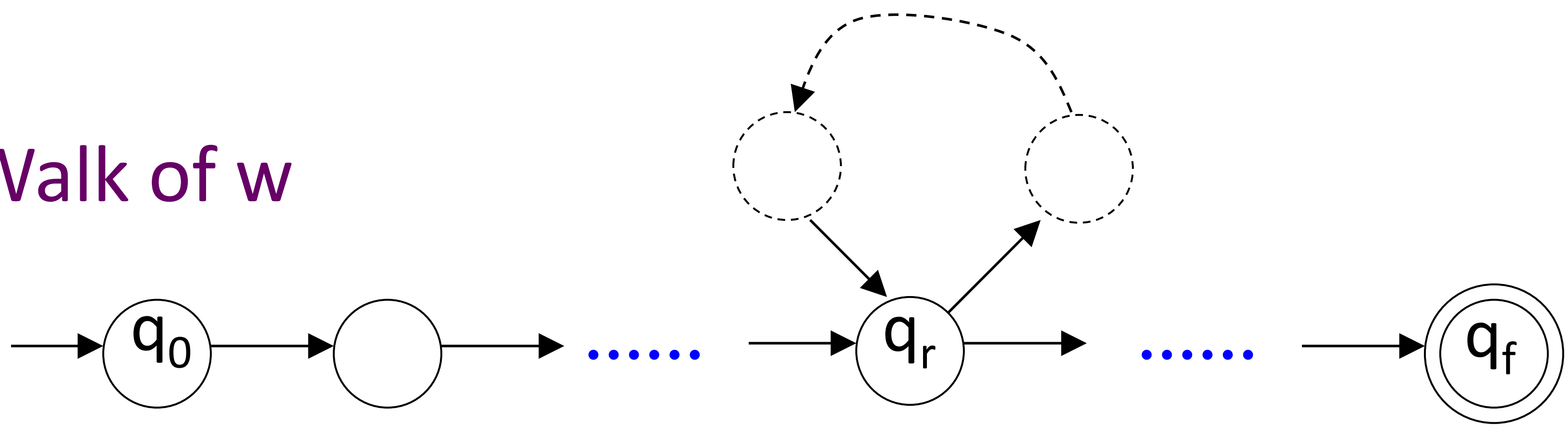


Pumping Lemma

For this string $w \in L$ with $|w| \geq m$,

let q_r be the first state to repeat:

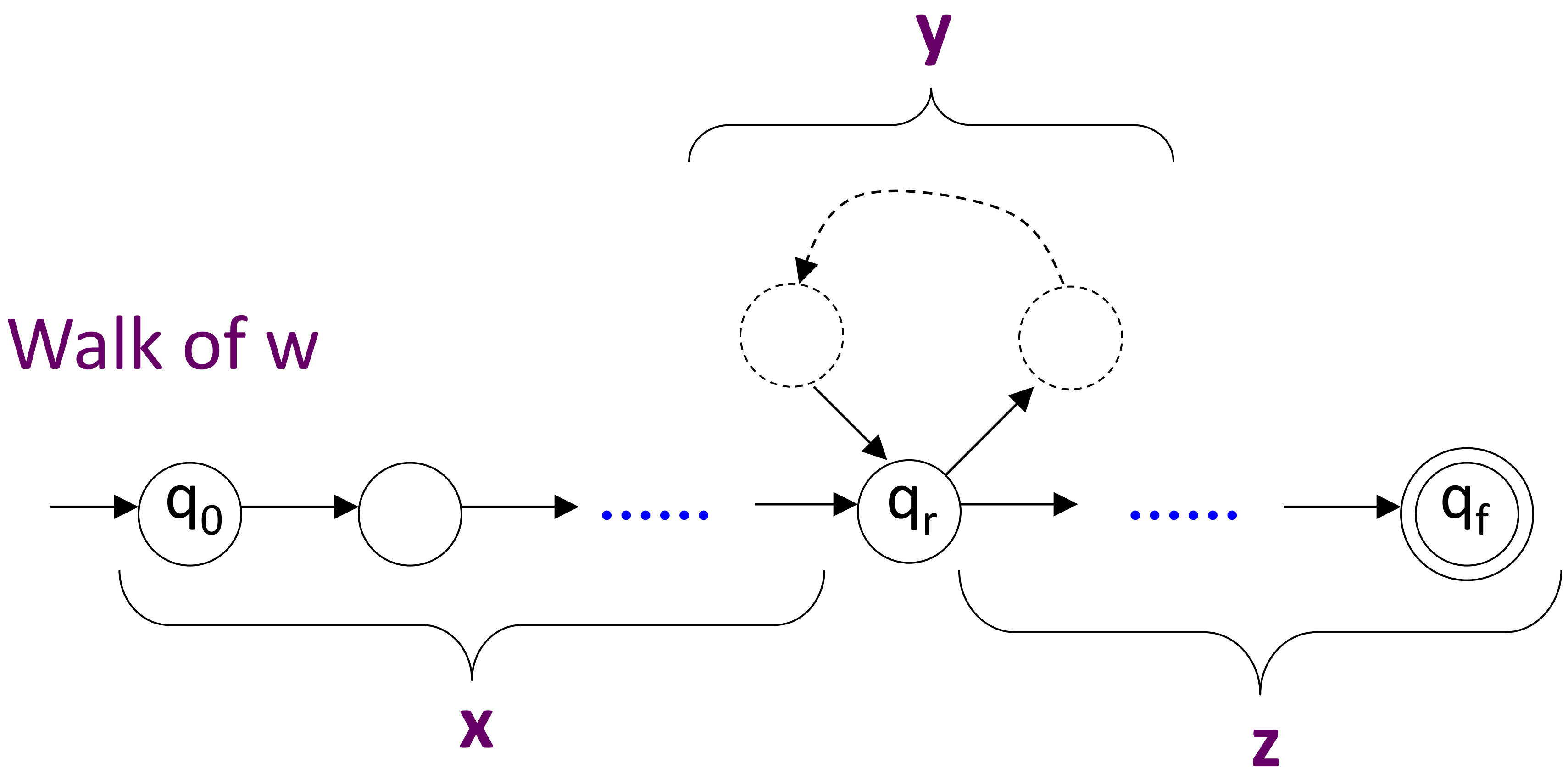
Walk of w



Pumping Lemma

We can write $w = x\ y\ z$

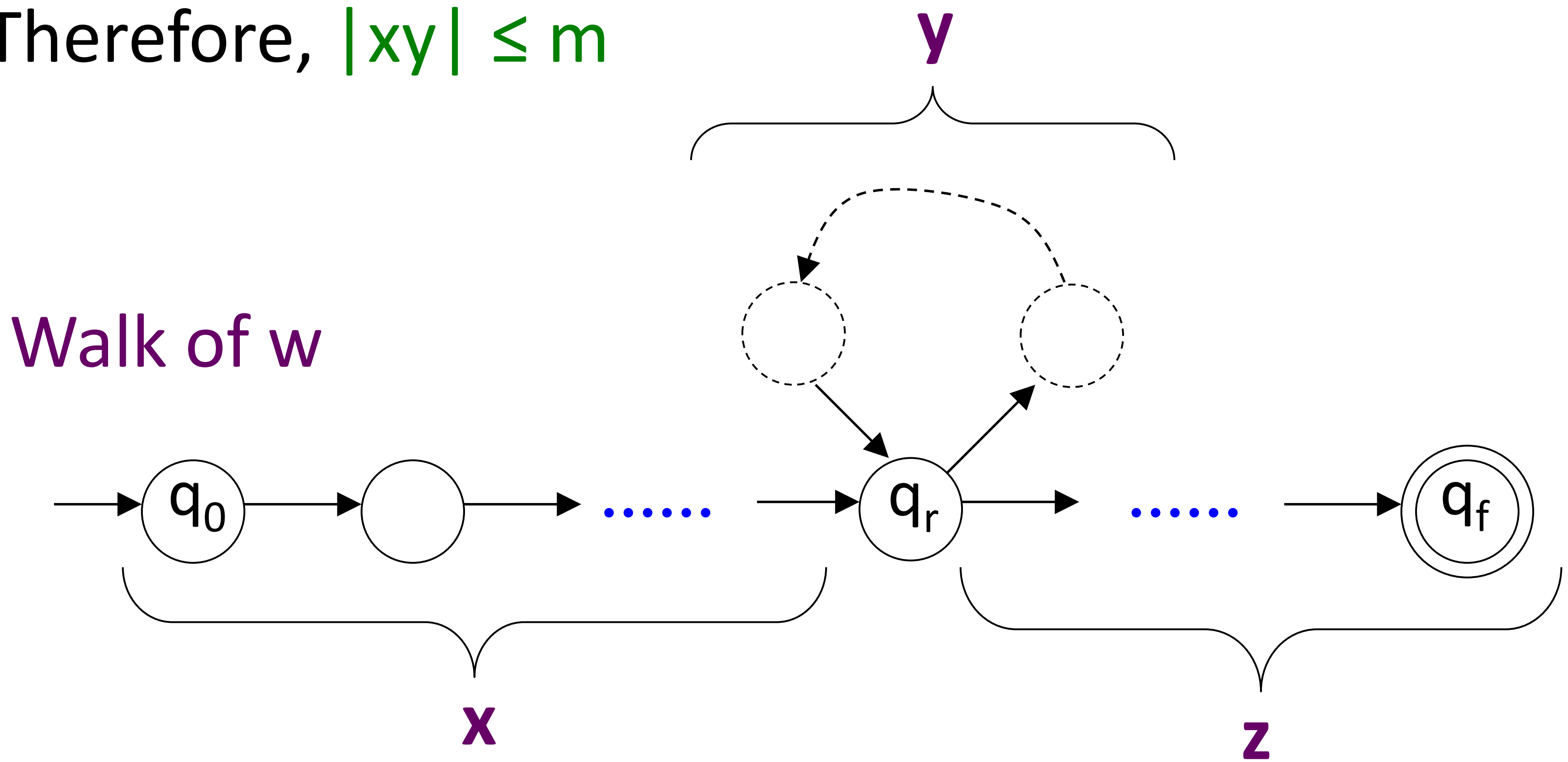
y is the substring between first and second occurrence of q_r



Pumping Lemma

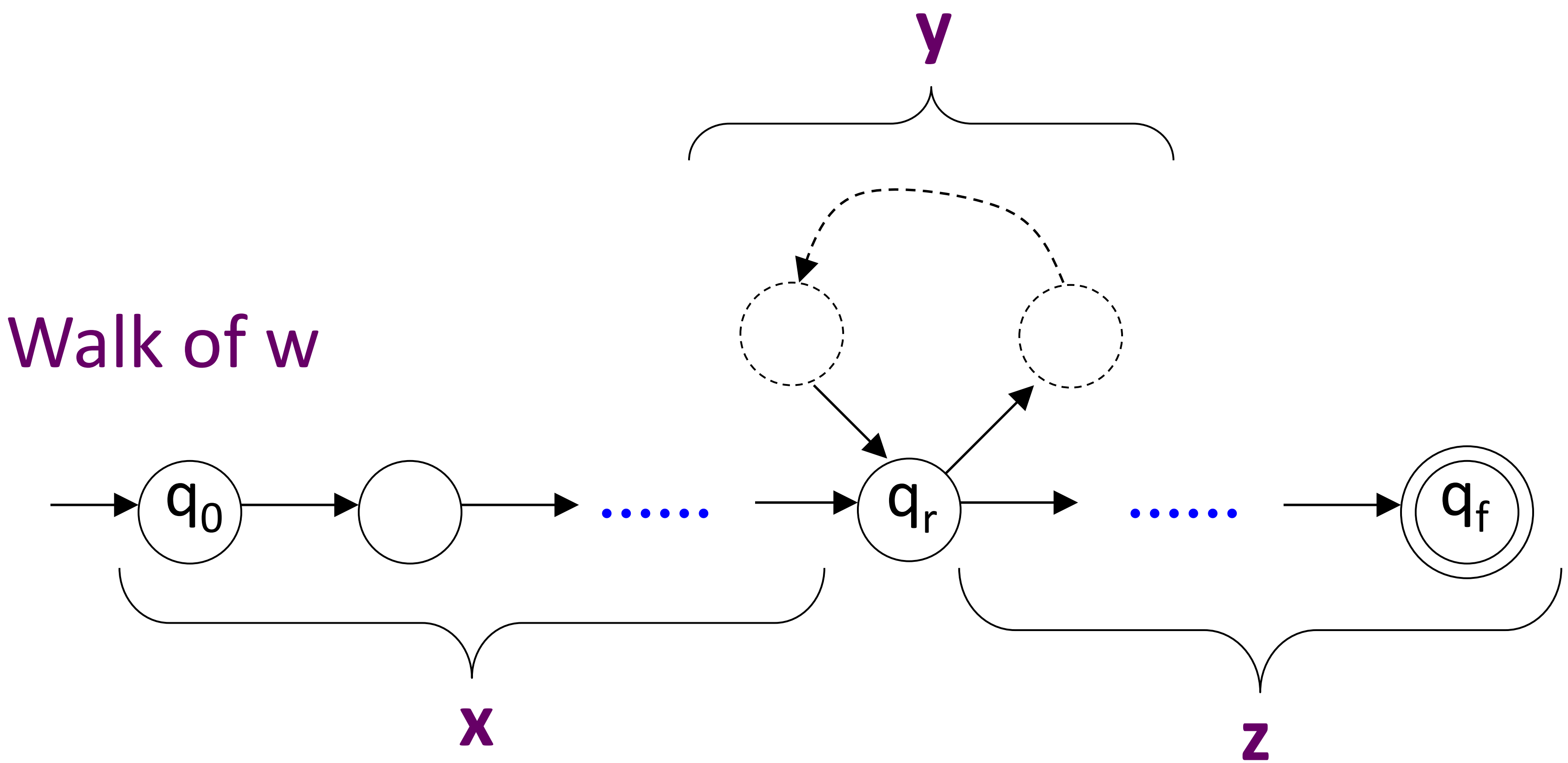
In xy there is no state repeated (except first reoccurrence of q_r).

Therefore, $|xy| \leq m$



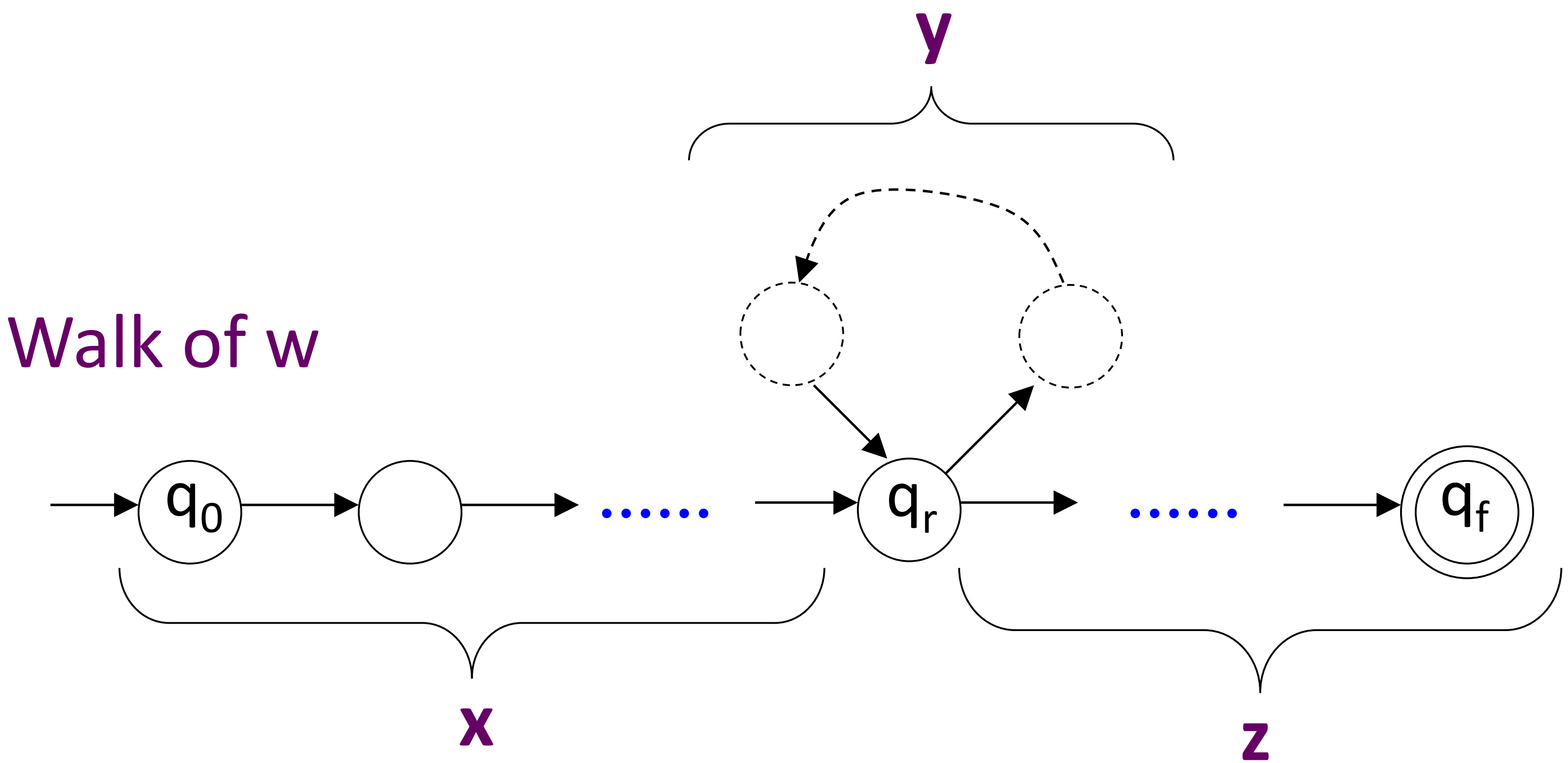
Pumping Lemma

In y there is at least one transition, and therefore, $|y| \geq 1$



Pumping Lemma

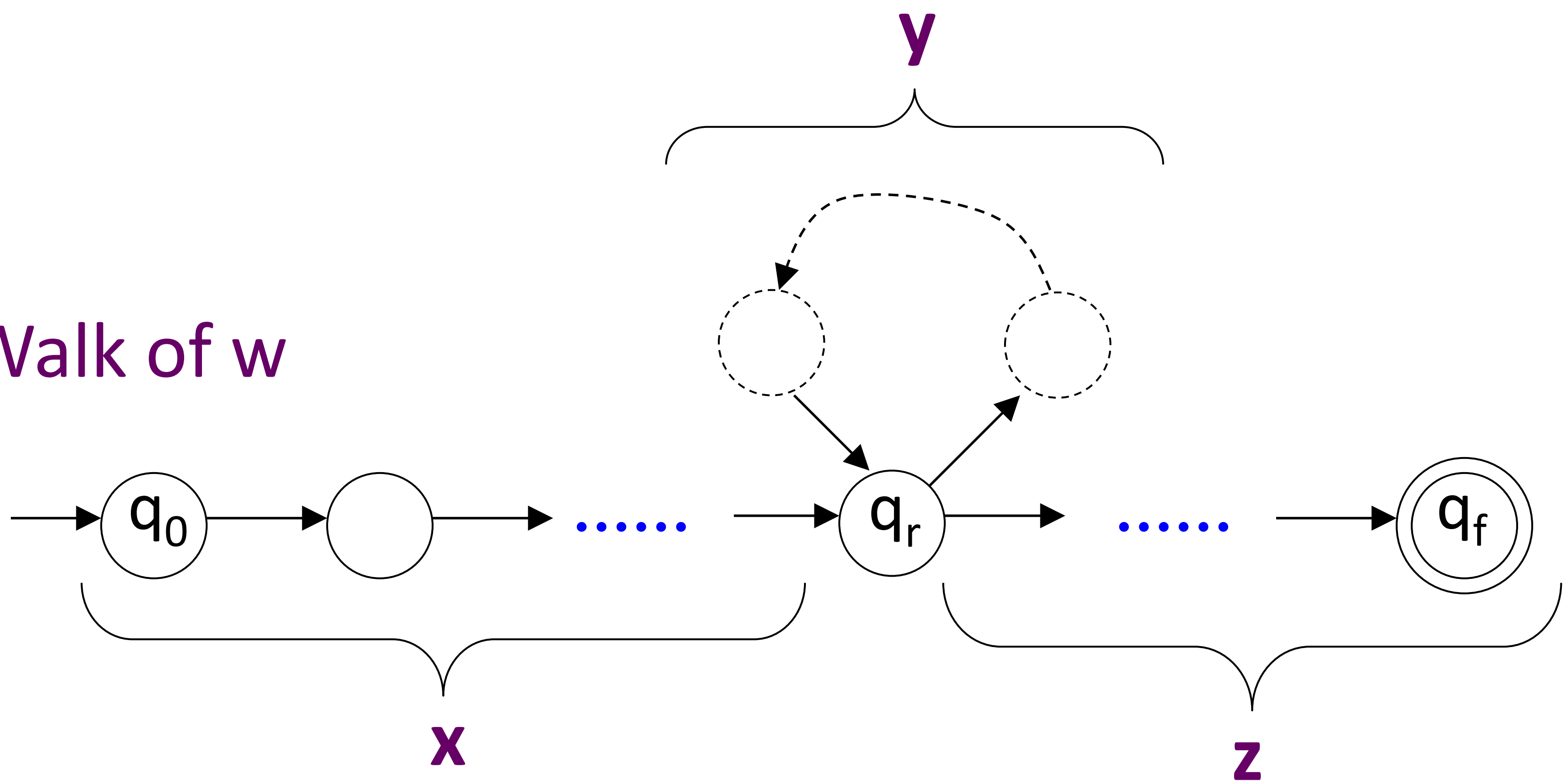
String xz is accepted if we do not follow the loop at all.



Pumping Lemma

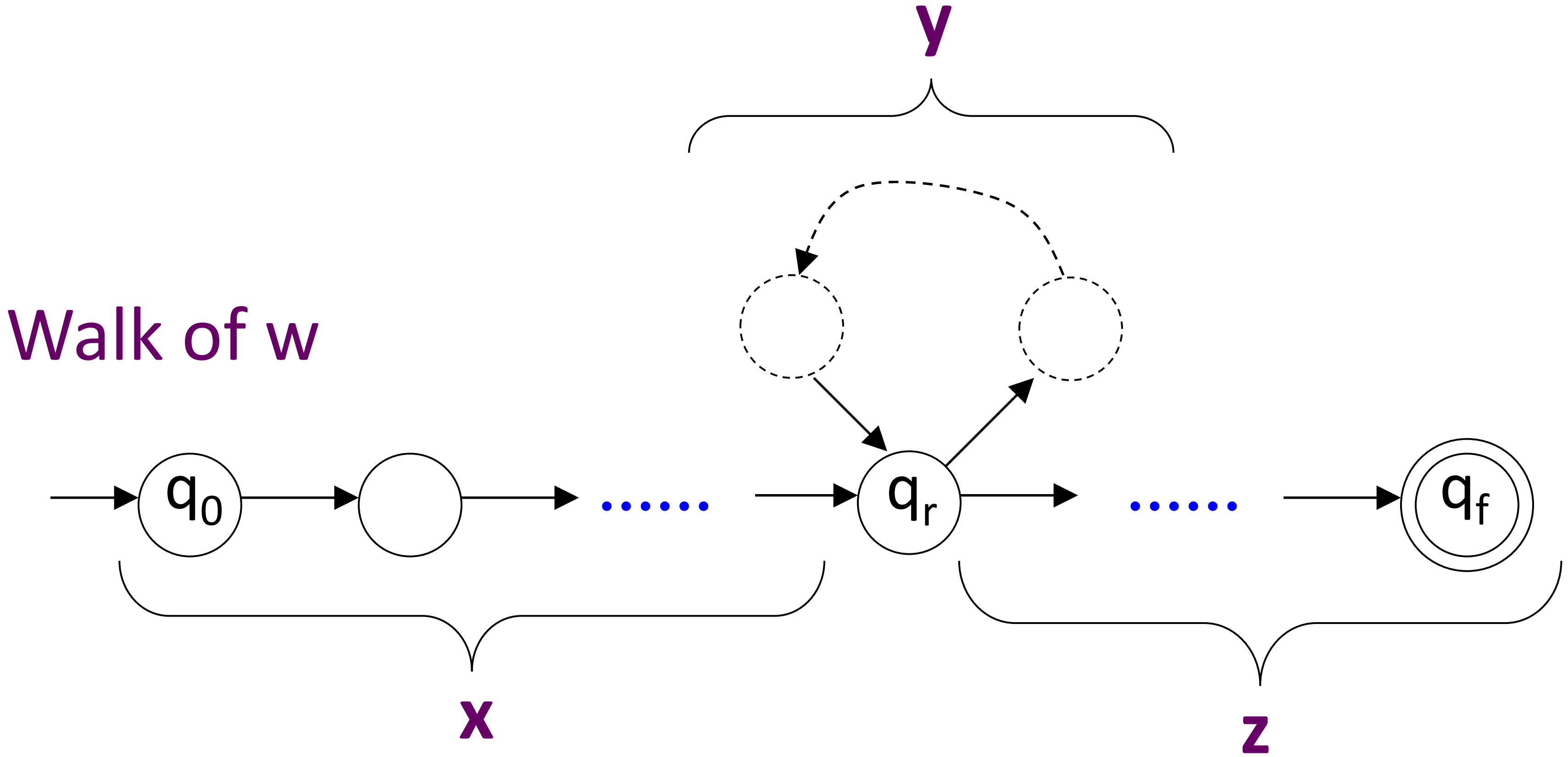
String xyz, xyyz, xyyyz, … are accepted if we follow the y loop multiple times …

Walk of w



Pumping Lemma

Therefore, string $xy^iz \in L, \quad i = 0,1,2,\ldots$



Pumping Lemma

Theorem: Let L be an infinite regular language. Then there exists a positive integer m such that for any $w \in L$ with $|w| \geq m$, w can be decomposed as $w = xyz$ with

$|xy| \leq m$ and

$|y| \geq 1$

And $w_i = xy^iz$ is also in L for all i = 0,1,2,...

Pumping Lemma

- We can only use the pumping lemma to show certain languages are **not regular**.
- You cannot use this theorem for proving that a language is regular.

Using pumping lemma to prove a language L is not regular

1. Assume by contradiction that L is regular
2. Let m be the integer for pumping lemma
3. Pick a string $w \in L$, $|w| \geq m$
4. w can be decomposed as w = xyz such that $|xy| \leq m$ & $|y| \geq 1$, and $xy^iz \in L$ for all $i \geq 0$.
5. Show that $w' = xy^iz$ is not in L for some i.
6. This results in a contradiction since pumping lemma says $xy^iz \in L$ for all i=0,1,2,3,…

Pumping Lemma Example

Question: Prove that the language L = { $a^n b^n$: n ≥ 0} is not regular.

Answer: Proof by contradiction: assume L is regular.

Choose m as in the pumping lemma.

Pick w ∈ L such that |w| ≥ m:

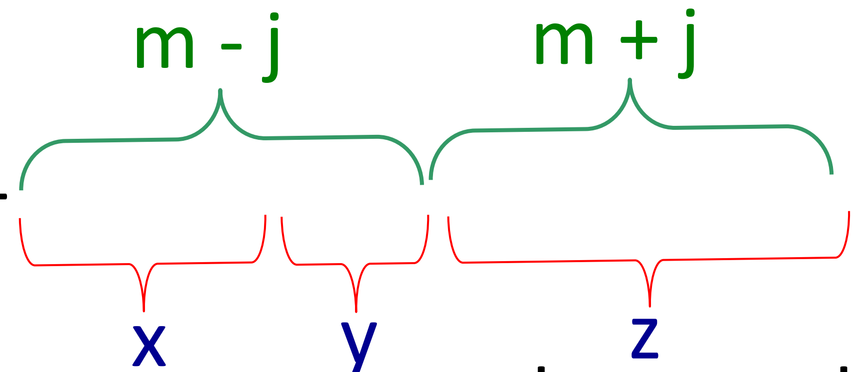
$$w = a^m b^m$$

Clearly w is in L.

Pumping Lemma Example

There exists a decomposition such that $w = xyz$ and $|xy| \leq m$ & $|y| \geq 1$. We do not know the exact decomposition but it must be that the bounds hold. As an example:

For some $0 \leq j \leq m-1$



$w = a^m b^m =$ a..aa...aa...aaab.............b

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

$x = a^{m-j-k}$ $y = a^k$ $z = a^j b^m$

Pumping Lemma Example

- Now we need to show that w' = xy^iz is not in L for some i.
- w = xyz
- $x = a^{m-j-k}$ $y = a^k$ $z = a^jb^m$
- Since y is all a's and at least 1 a, the number of
- a's in xy^i (and including in z) will become greater than the number of b's given a sufficiently high value of i.

Is w' ∈ L ?

No! CONTRADICTION!

Pumping Lemma
Example 1

The assumption that L is a regular language is not true.

Therefore, L is not regular.