CSC 3130: Automata theory and formal languages

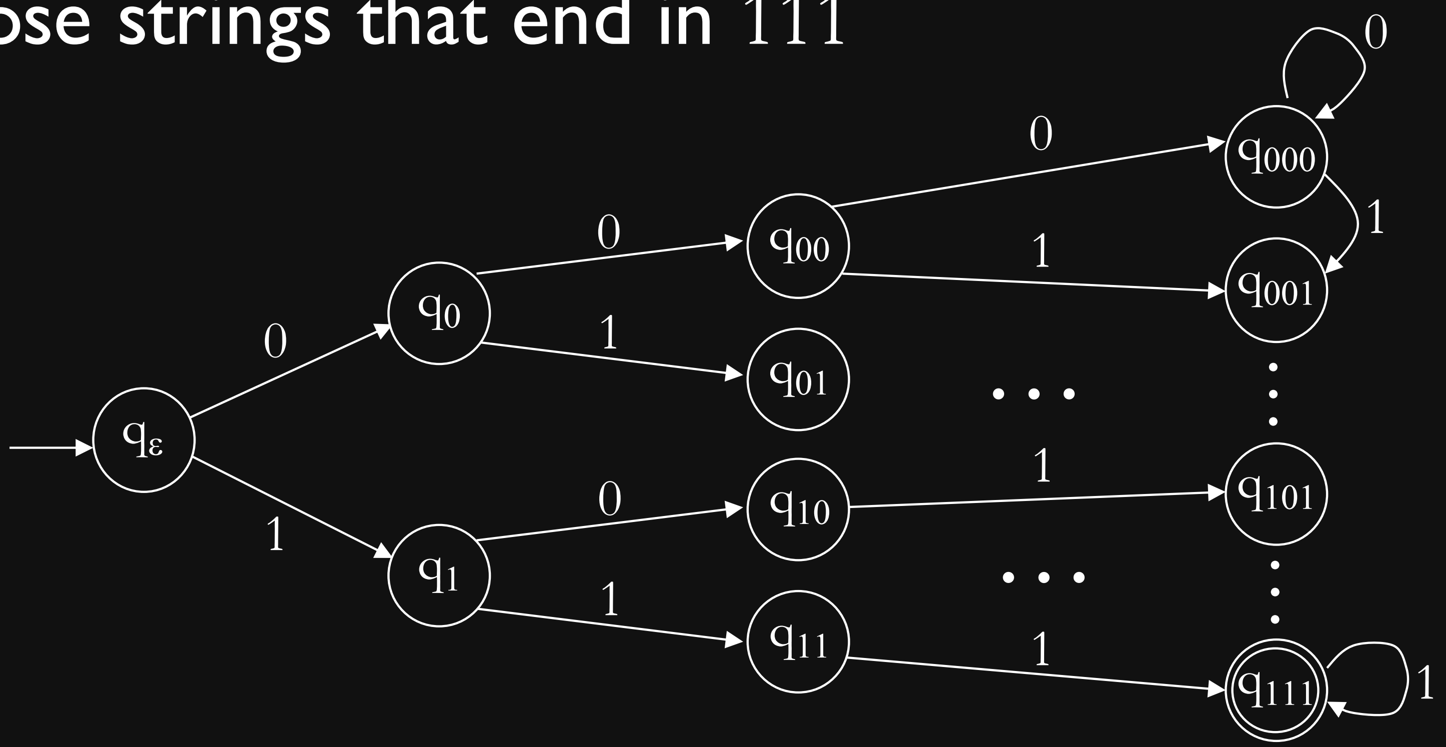
DFA minimization

Andrej Bogdanov

http://www.cse.cuhk.edu.hk/~andrejb/csc3130

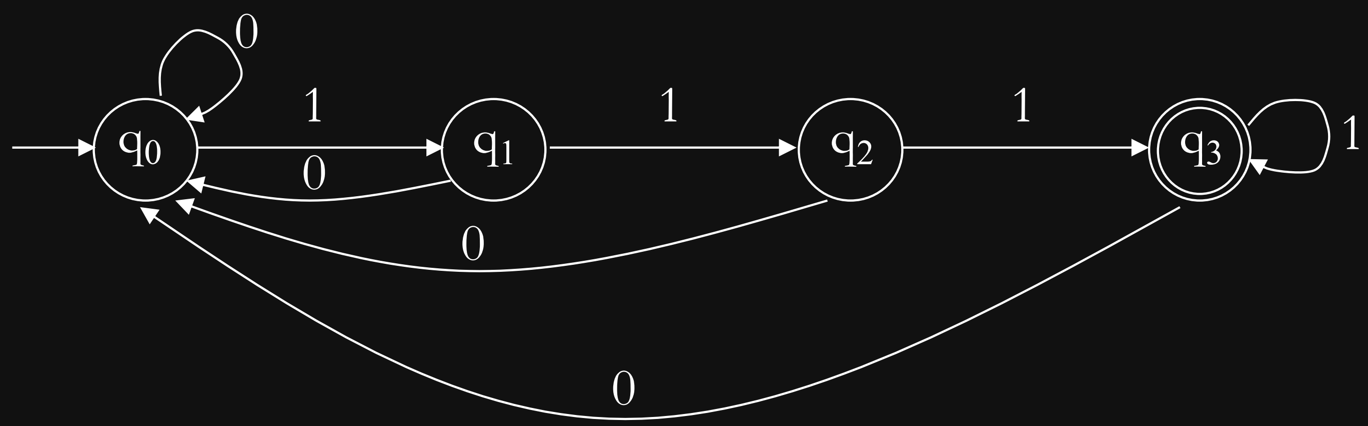
Example

- Construct a DFA over alphabet $\{0, 1\}$ that accepts those strings that end in 111



- This is big, isn't there a smaller DFA for this?

Smaller DFA

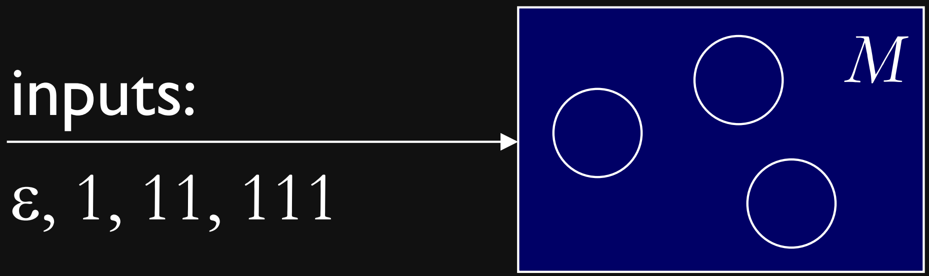


Can we do it with 3 states?

Even smaller DFA?

- Suppose we had a 3 state DFA M for L

  … let's imagine what happens when:



- By the pigeonhole principle, on two of these inputs M ends in the same state

Pigeonhole principle

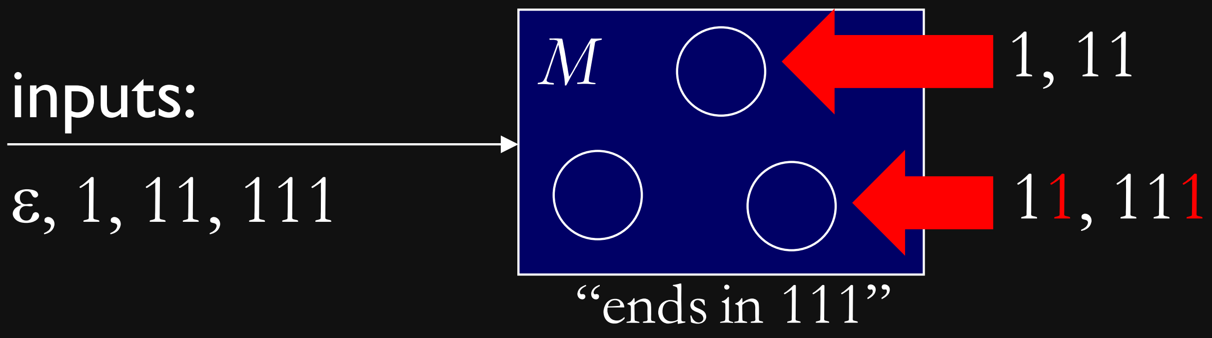
Suppose you are tossing m balls into n bins, and $m > n$. Then two balls end up in the same bin.

- Here, balls are inputs, bins are states:

If you have a DFA with n states and you run it on m inputs, and $m > n$, then two inputs end up in same state.

A smaller DFA

- Suppose M ends up in the same state after reading inputs $x = 1$ and $y = 11$

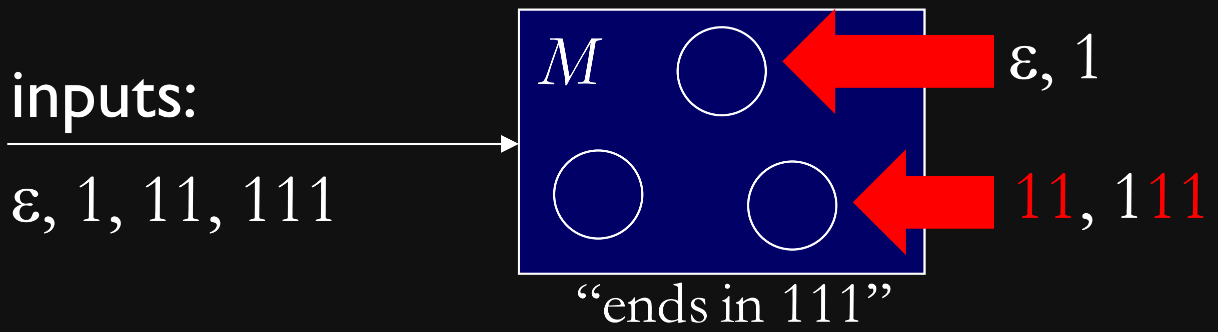


- Then after reading one more 1
  - The state of $x1 = 11$ should be rejecting
  - The state of $y1 = 111$ should be accepting

… but they are both the same state!

A smaller DFA

- Suppose M ends up in the same state after reading inputs $x = \varepsilon$ and $y = 1$



- Then after reading 11
  - The state of $x1 = 11$ should be rejecting
  - The state of $y1 = 111$ should be accepting

… but they are both the same state!

No smaller DFA!

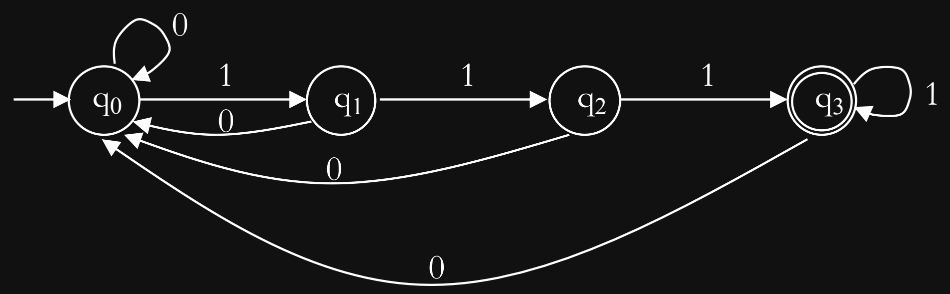
- After looking at all possible pairs for $x, y, x \neq y$

  (ε, 1) (ε, 11) (ε, 111) (1, 11) (1, 111) (11, 111)

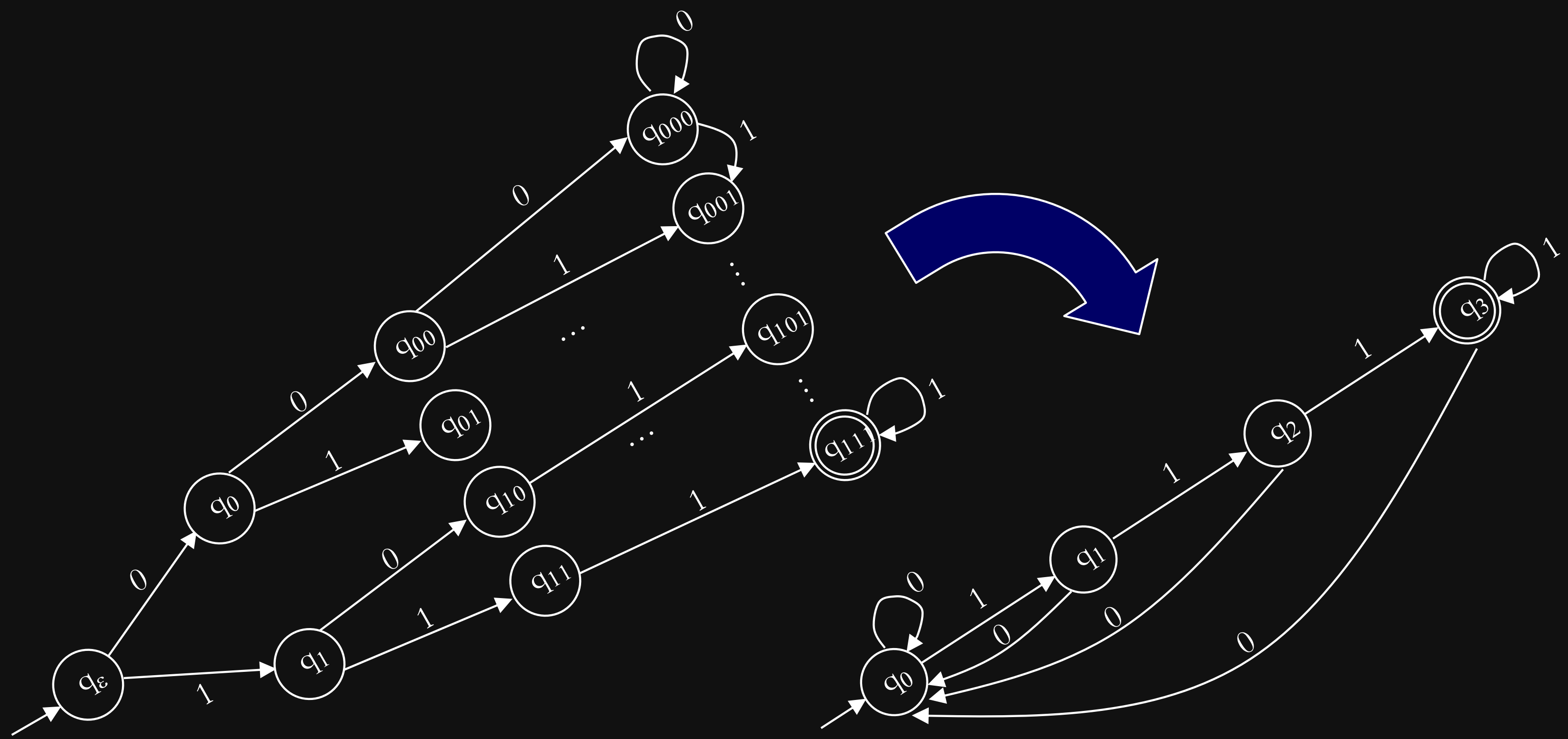
  we conclude that

  There is no DFA with 3 states for L

- So, this DFA is minimal



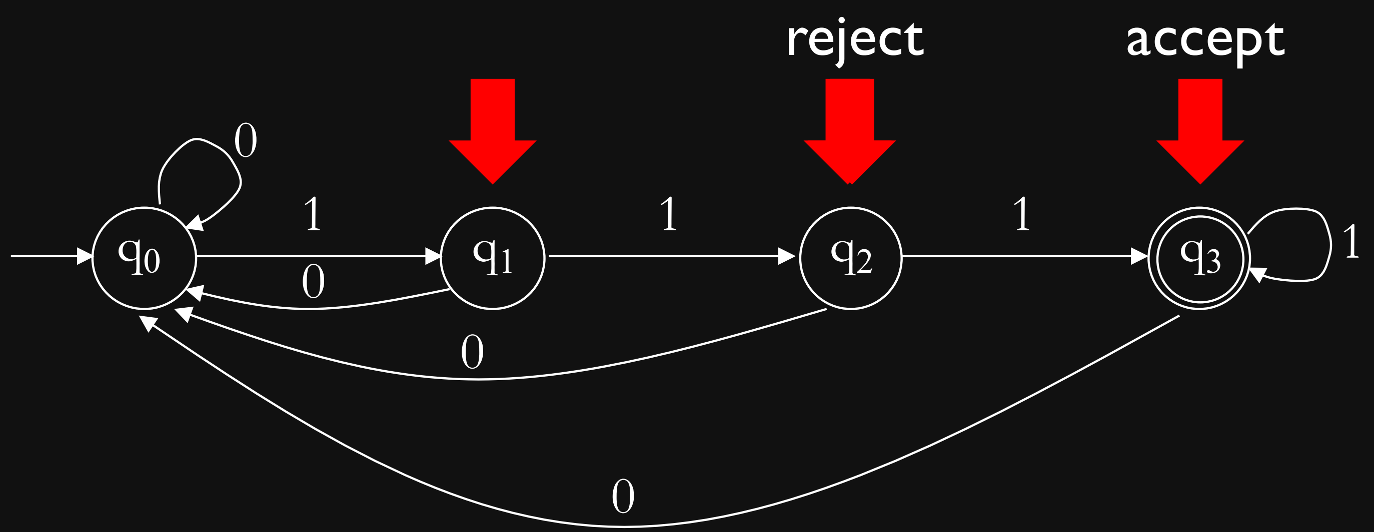
DFA minimization



We will show how to turn any DFA for L into the minimal DFA for L

Minimal DFAs and distinguishable states

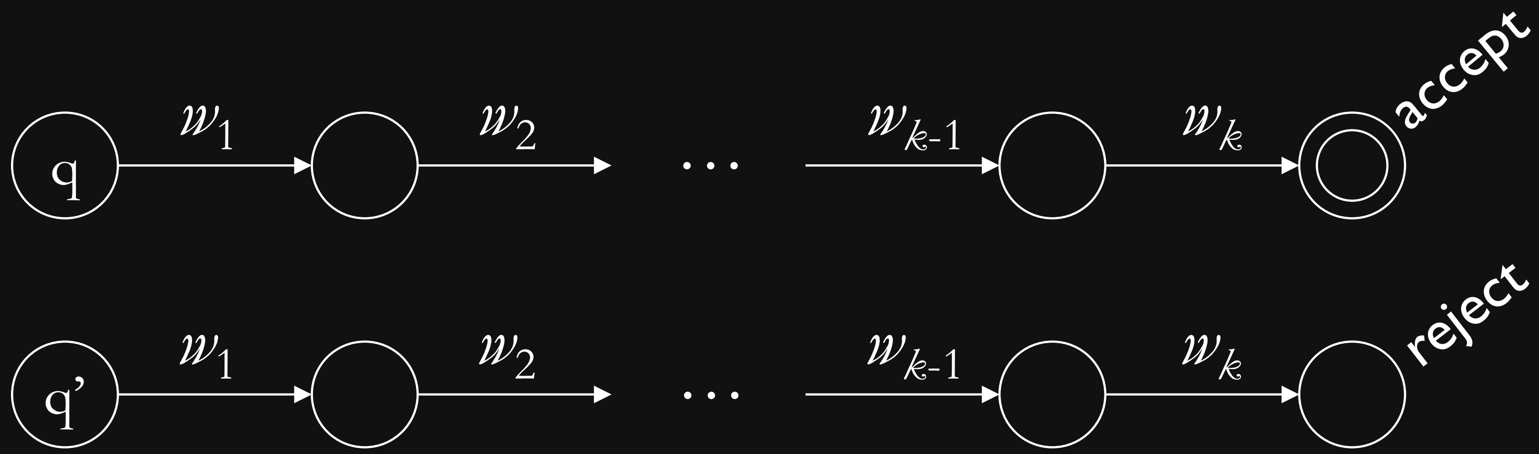
- First, we have to understand minimal DFAs:



minimal DFA ⟺ every pair of states is distinguishable

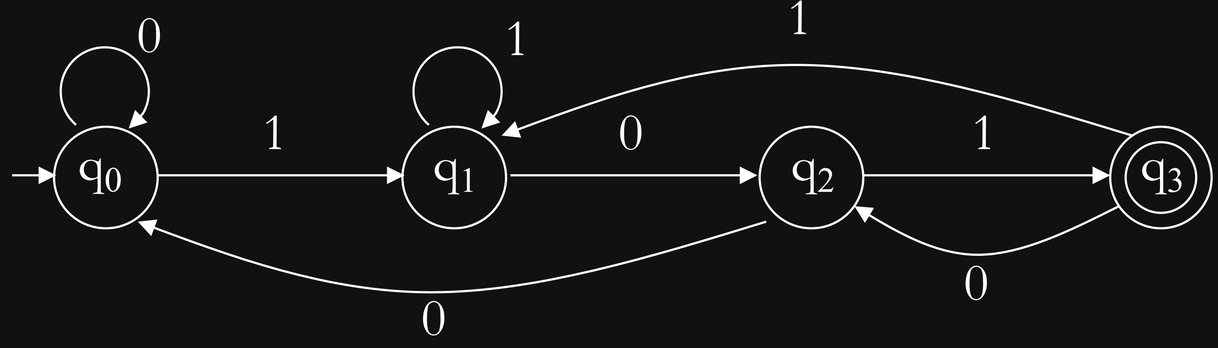
Distinguishable states

- Two states q and q' are distinguishable if



on the same continuation string $w_1w_2...w_k$, one accepts, but the other rejects

Examples of distinguishable states



(q_0, q_1) distinguishable by 01

(q_0, q_2) distinguishable by 1

(q_0, q_3) distinguishable by ε

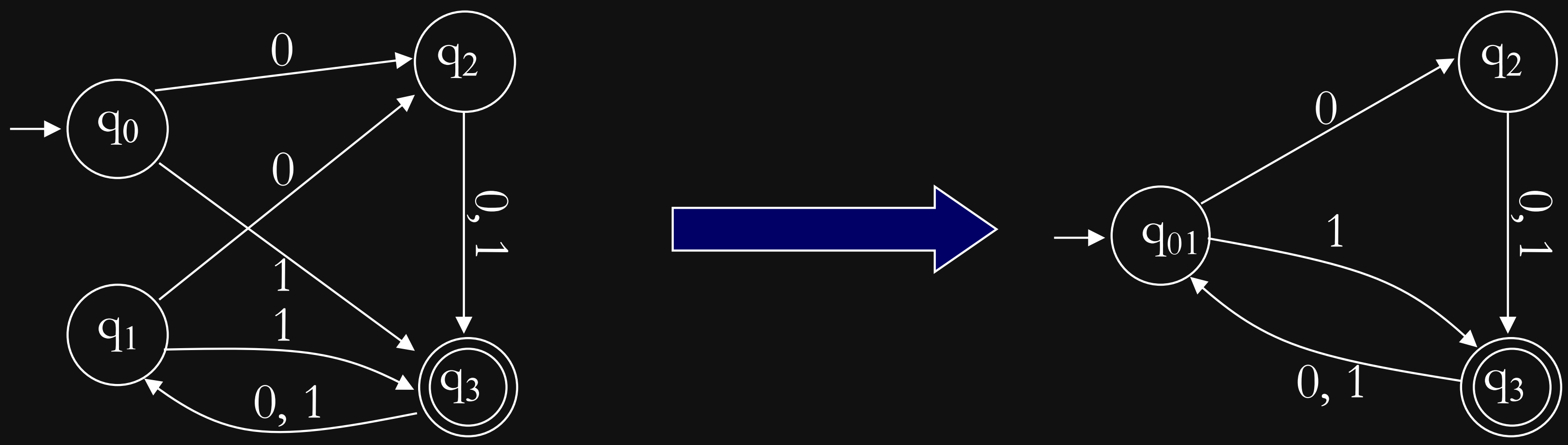
(q_1, q_2) distinguishable by 1

(q_1, q_3) distinguishable by ε

(q_2, q_3) distinguishable by ε

DFA is minimal

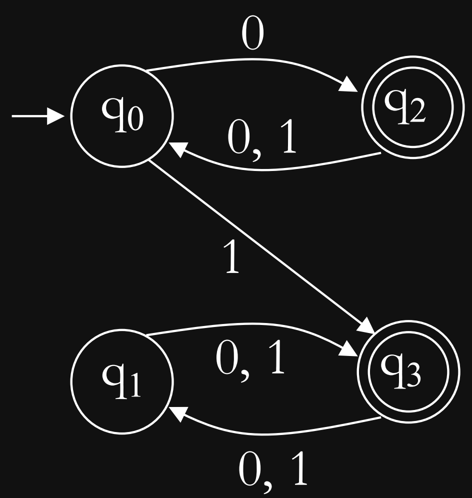
Examples of distinguishable states

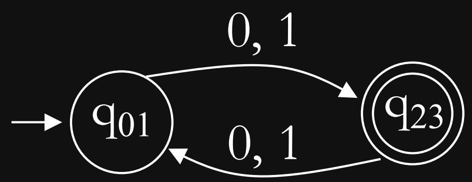


(q_0, q_3)	distinguishable by ε
(q_1, q_3)	distinguishable by ε
(q_2, q_3)	distinguishable by ε
(q_1, q_2)	distinguishable by 0
(q_0, q_2)	distinguishable by 0
(q_0, q_1)	indistinguishable

indistinguishable pairs can be merged

Examples of distinguishable states





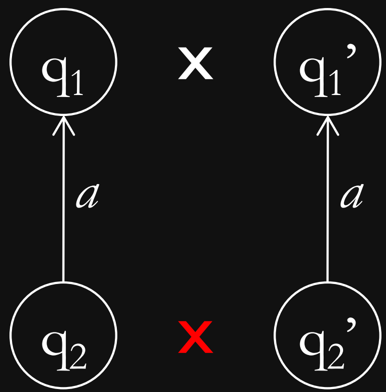
(q_0, q_2)	distinguishable by ε
(q_1, q_2)	distinguishable by ε
(q_0, q_3)	distinguishable by ε
(q_1, q_3)	distinguishable by ε
(q_0, q_1)	indistinguishable
(q_2, q_3)	indistinguishable

Finding (in)distinguishable states

Rule 1:

If q is accepting and q' is rejecting
Mark (q, q') as distinguishable (x)

Rule 2:

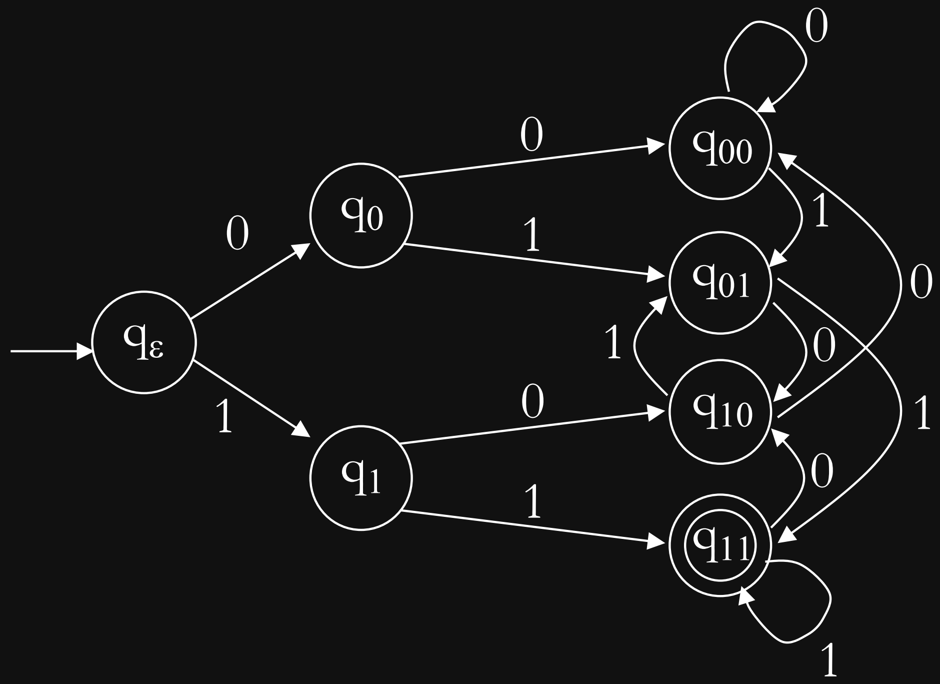


If (q_1, q_1') are marked,
Mark (q_2, q_2') as distinguishable (x)

Rule 3:

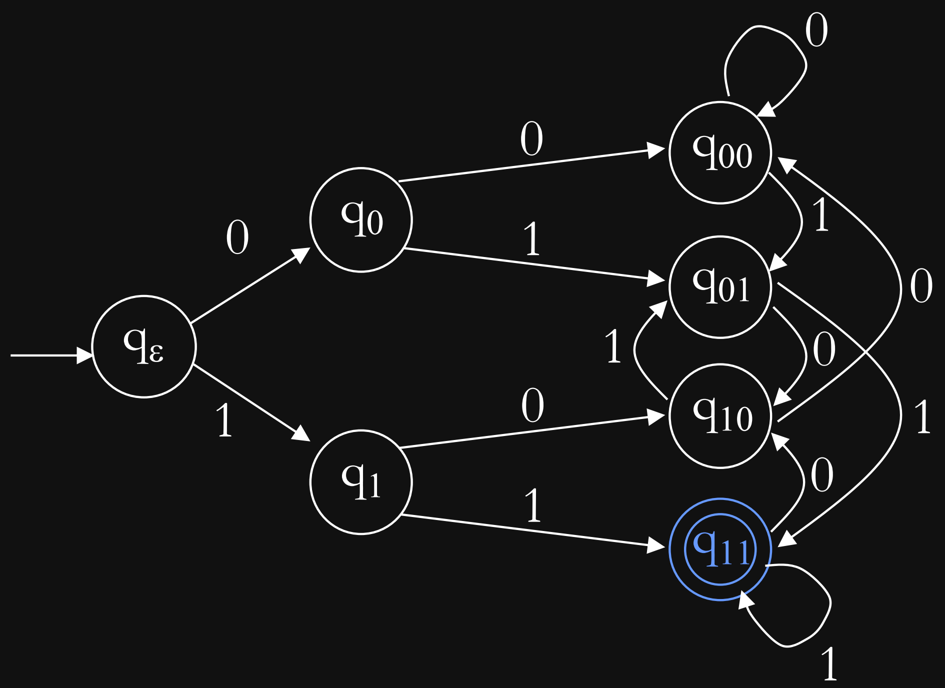
Unmarked pairs are indistinguishable
Merge them together

Example of DFA minimization



	q_ε	q_0	q_1	q_{00}	q_{01}	q_{10}
q_0						
q_1						
q_{00}						
q_{01}						
q_{10}						
q_{11}						

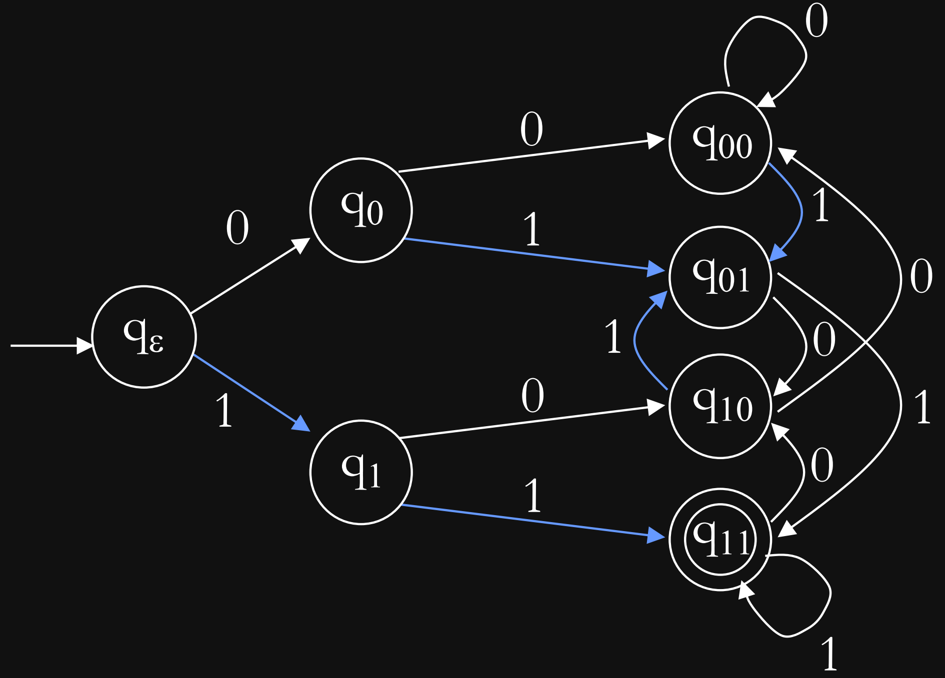
Example of DFA minimization



q_0						
q_1						
q_{00}						
q_{01}						
q_{10}						
q_{11}	x	x	x	x	x	x
	q_ε	q_0	q_1	q_{00}	q_{01}	q_{10}

① q_{11} is distinguishable from all other states

Example of DFA minimization

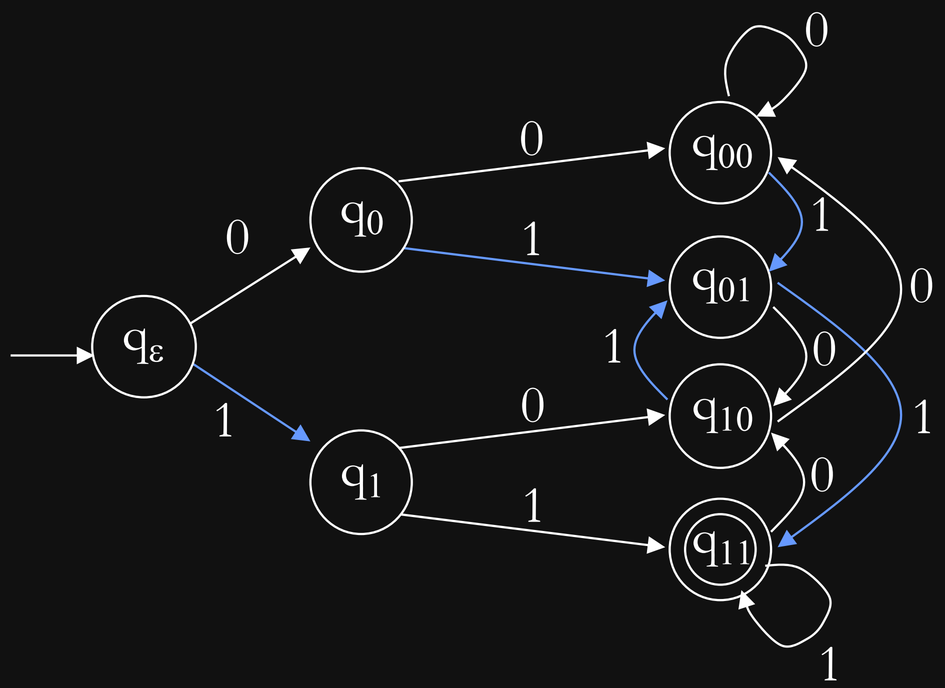


	q_ε	q_0	q_1	q_{00}	q_{01}	q_{10}
q_0						
q_1	x	x				
q_{00}			x			
q_{01}						
q_{10}			x			
q_{11}	x	x	x	x	x	x

② q_1 is distinguishable from q_ε, q_0, q_{00}, q_{10}

On transition 1, they go to distinguishable states

Example of DFA minimization

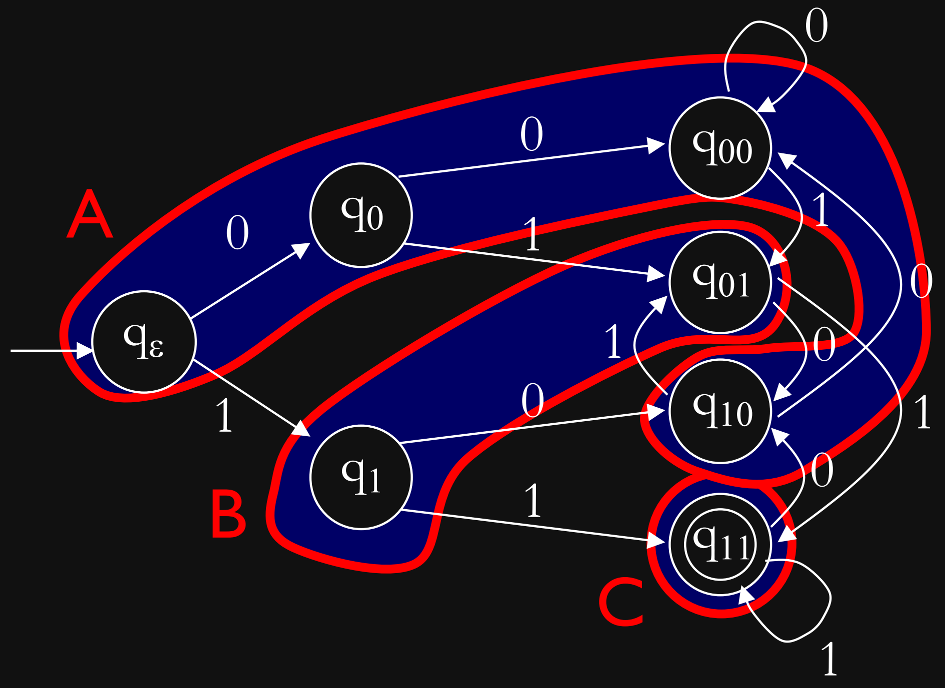


q_0						
q_1	x	x				
q_{00}			x			
q_{01}	x	x		x		
q_{10}			x		x	
q_{11}	x	x	x	x	x	x
	q_ε	q_0	q_1	q_{00}	q_{01}	q_{10}

② q_{01} is distinguishable from q_ε, q_0, q_{00}, q_{10}

On transition 1, they go to distinguishable states

Example of DFA minimization



	q_ε	q_0	q_1	q_{00}	q_{01}	q_{10}
q_0	A					
q_1	x	x				
q_{00}	A	A	x			
q_{01}	x	x	B	x		
q_{10}	A	A	x	A	x	
q_{11}	x	x	x	x	x	x

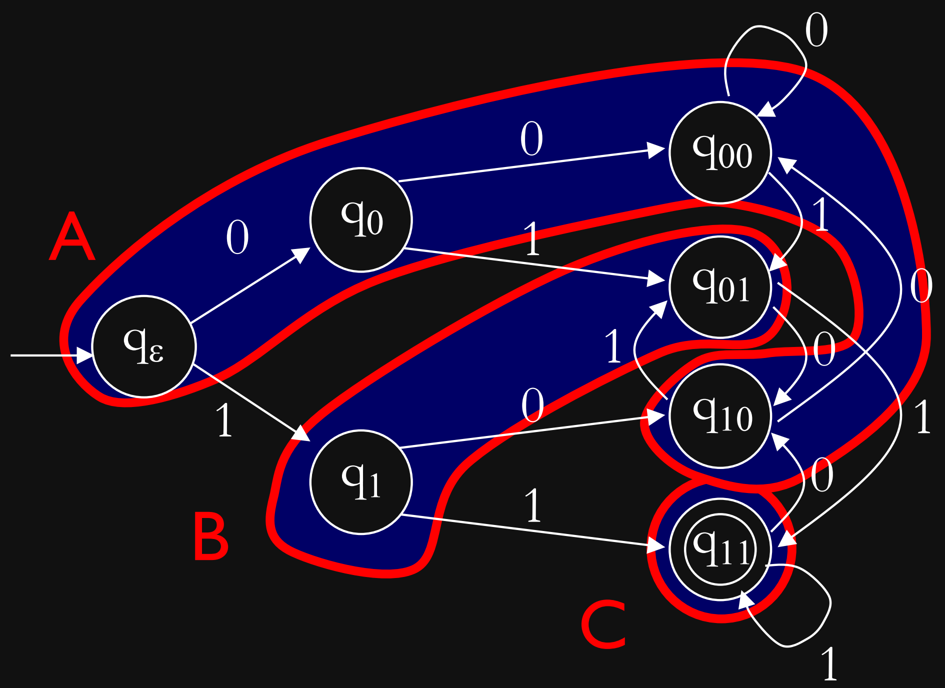
③ Merge states not marked distinguishable

q_ε, q_0, q_{00}, q_{10} are equivalent → group A

q_1, q_{01} are equivalent → group B

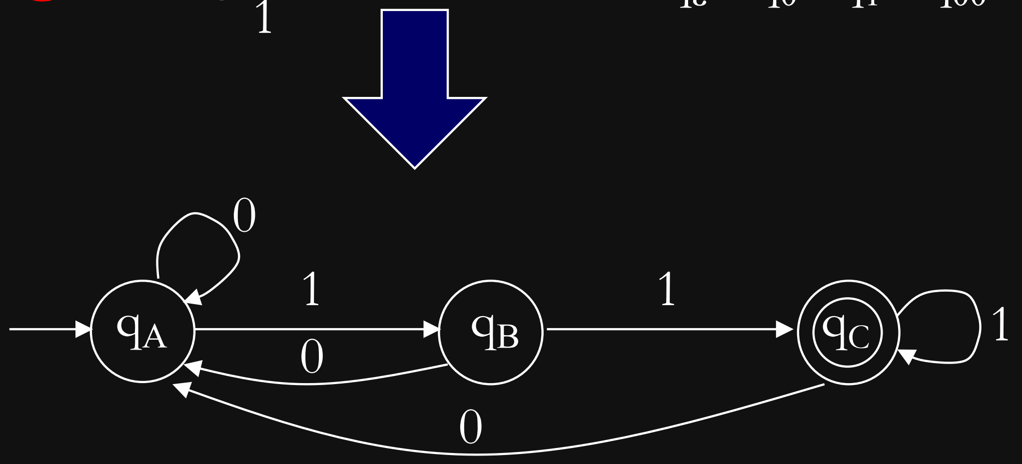
q_{11} cannot be merged → group C

Example of DFA minimization



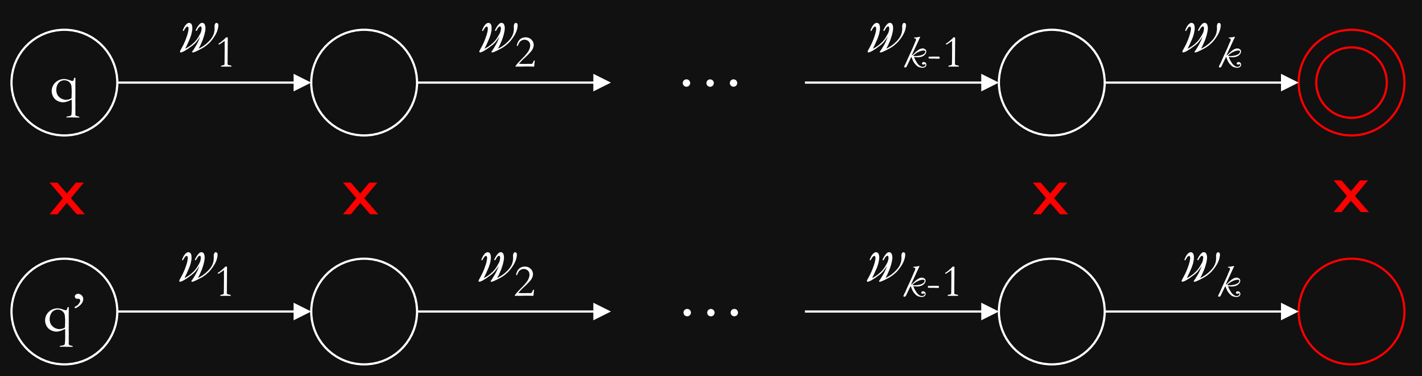
	q_ε	q_0	q_1	q_{00}	q_{01}	q_{10}
q_0	A					
q_1	x	x				
q_{00}	A	A	x			
q_{01}	x	x	B	x		
q_{10}	A	A	x	A	x	
q_{11}	x	x	x	x	x	x

minimized DFA:



Food for thought

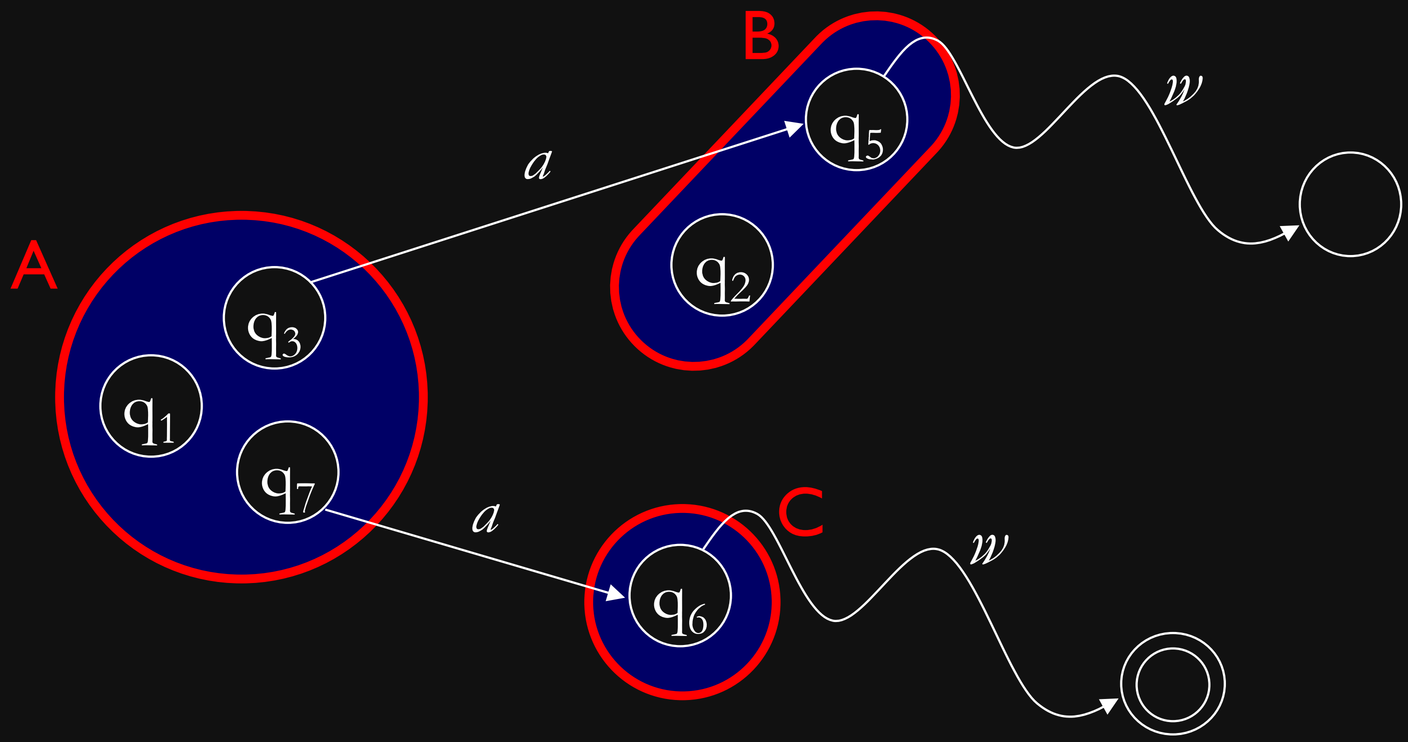
- Why does method find all distinguishable pairs?



Because we work backwards

Food for thought

- Why are there no inconsistencies when we merge?



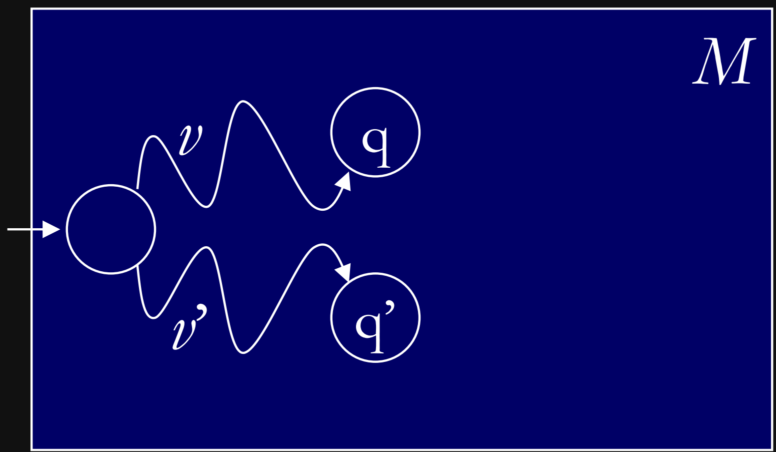
Because we only merge indistinguishable states

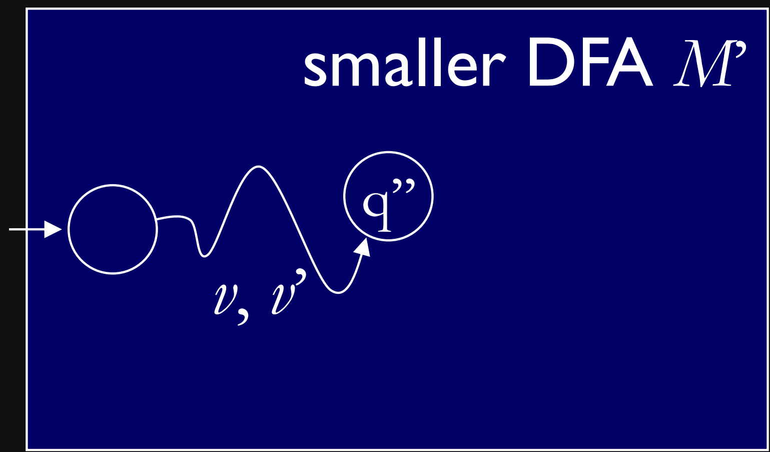
Food for thought

- Why is there no smaller DFA?

  Suppose there is

  By the pigeonhole principle this must happen:

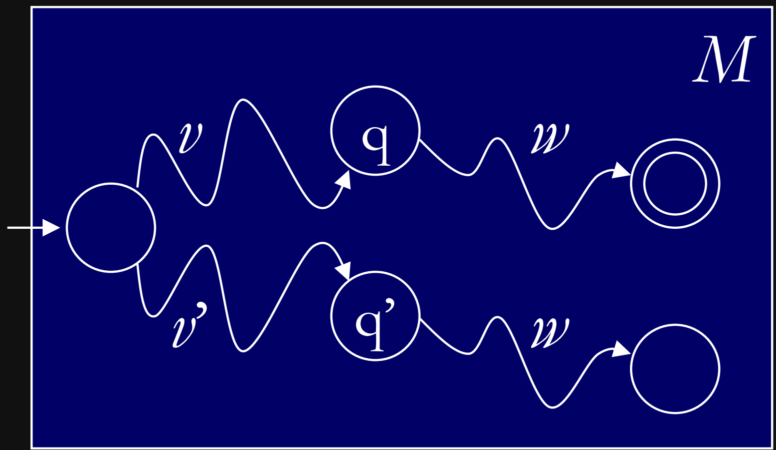


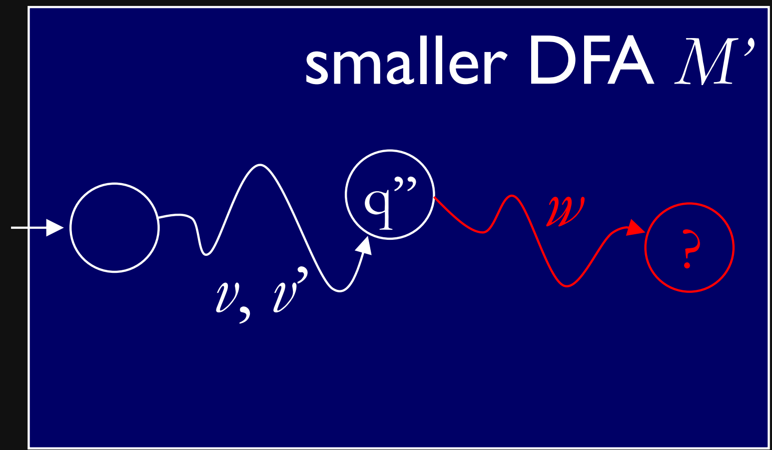


Food for thought

- Why is there no smaller DFA?

But then





Every pair of states
is distinguishable

q'' cannot exist!