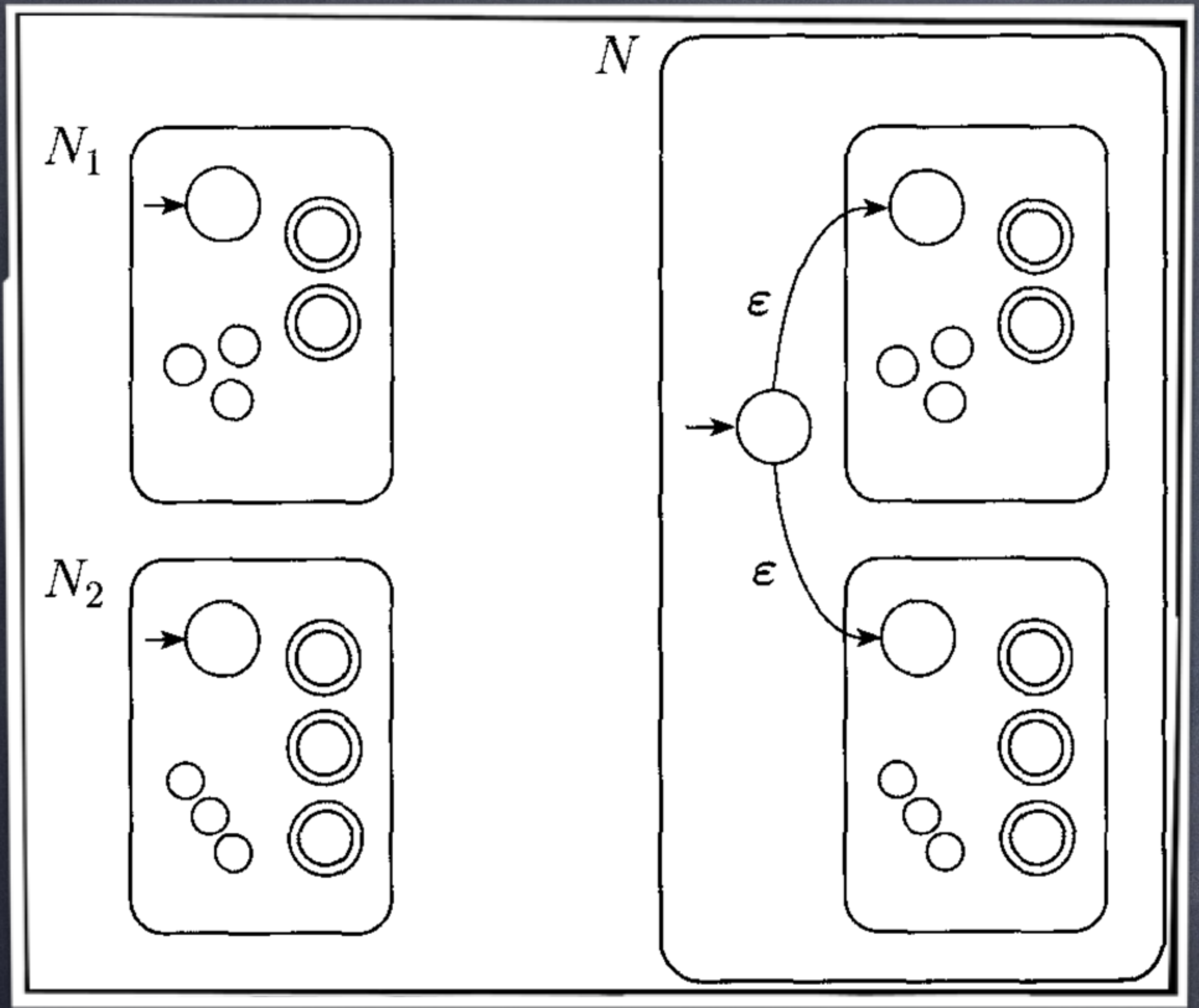
COMP-330
Theory of Computation

Fall 2019 -- Prof. Claude Crépeau

Lec. 6 : Myhill-Nerode Theorem and applications

Regular Operations

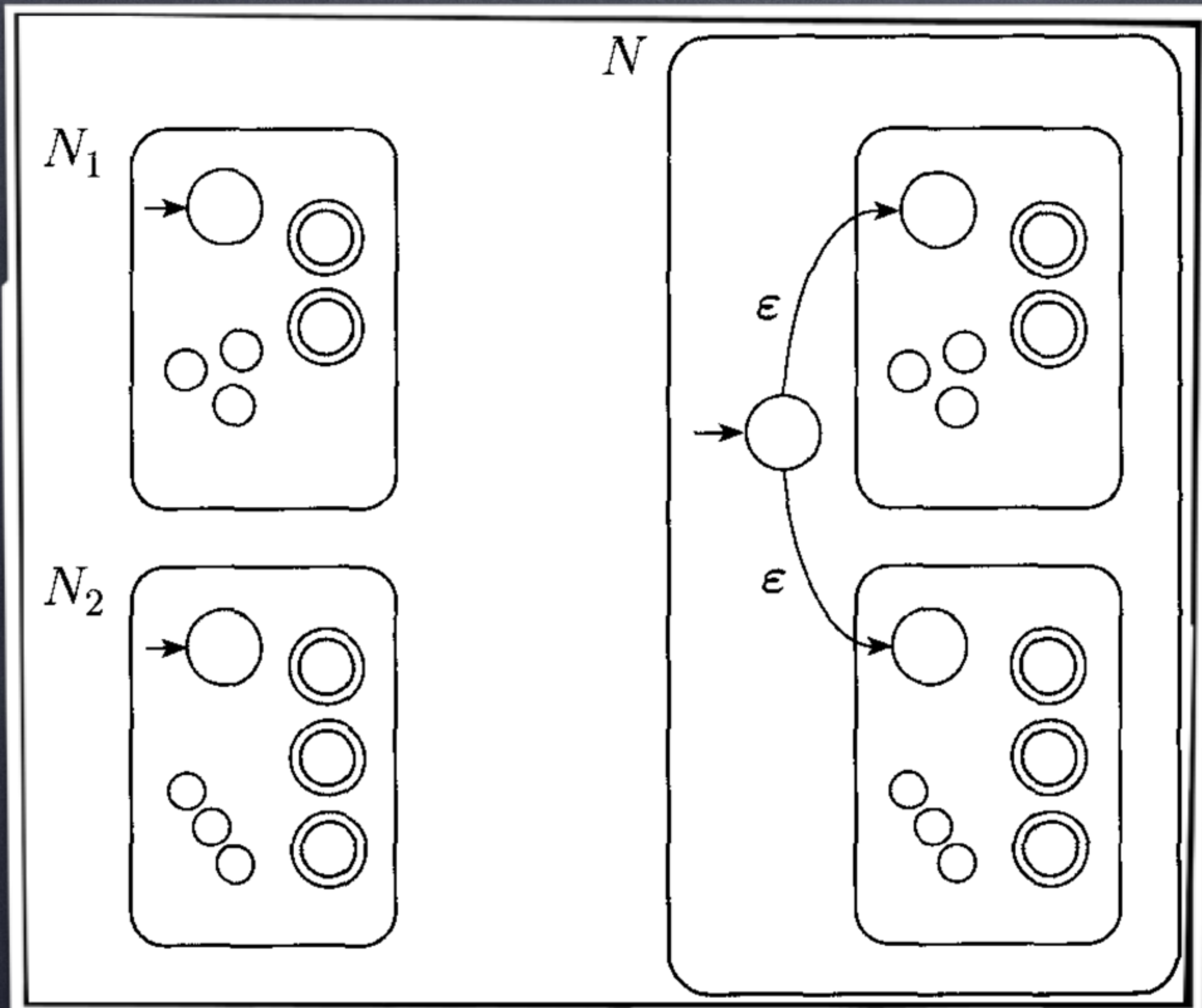
Regular Operations

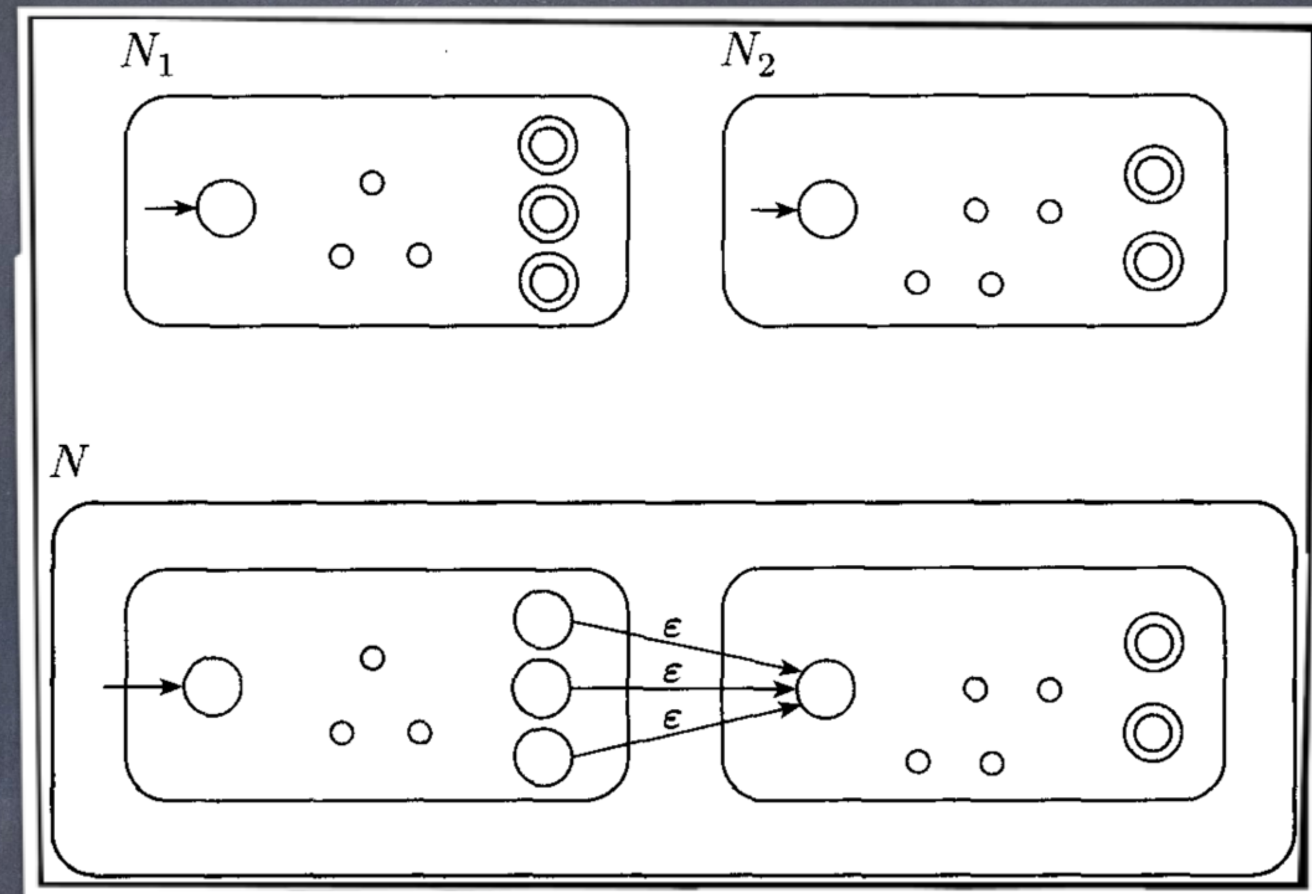


THEOREM 1.45

The class of regular languages is closed under the union operation.

Regular Operations





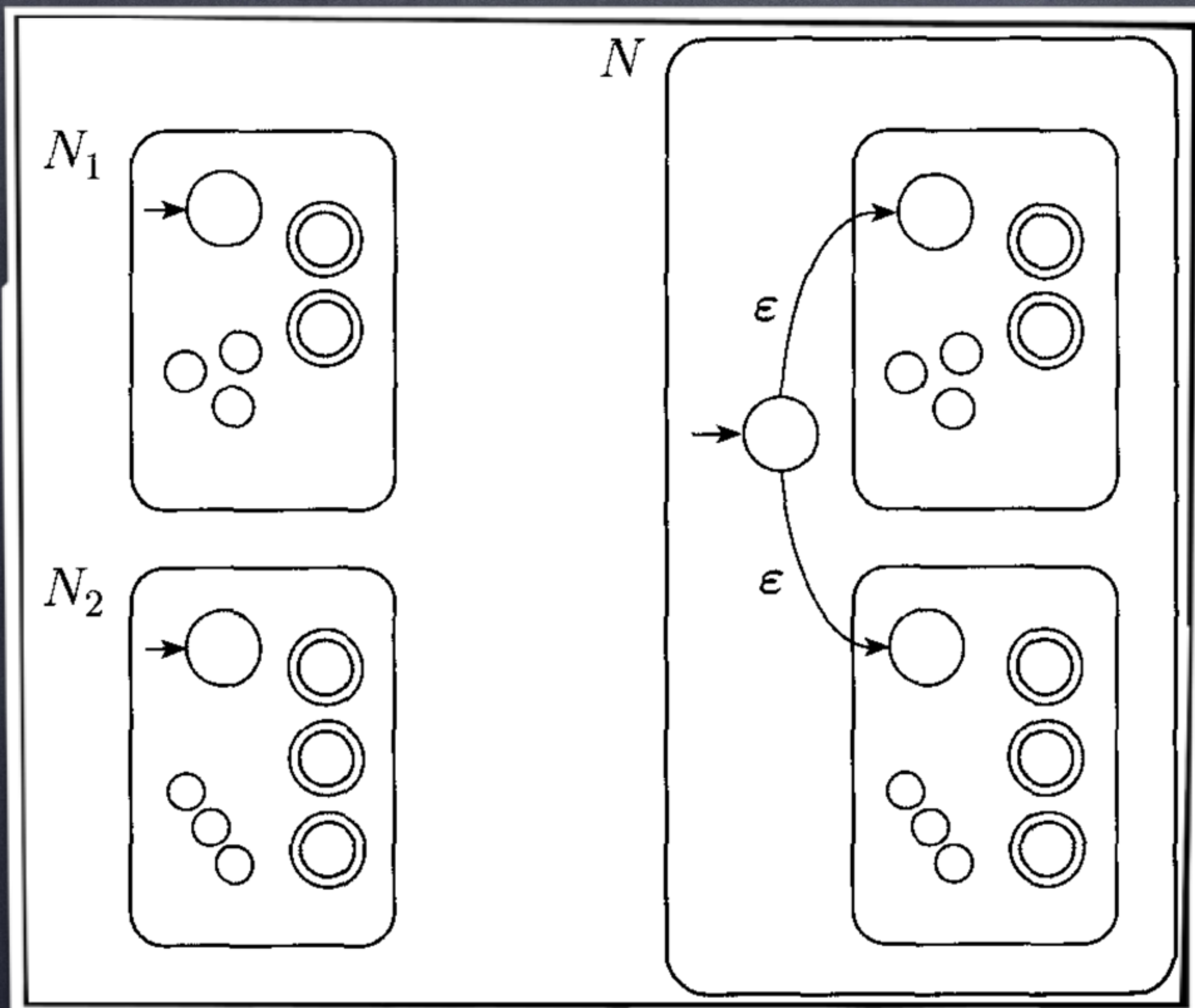
THEOREM 1.45

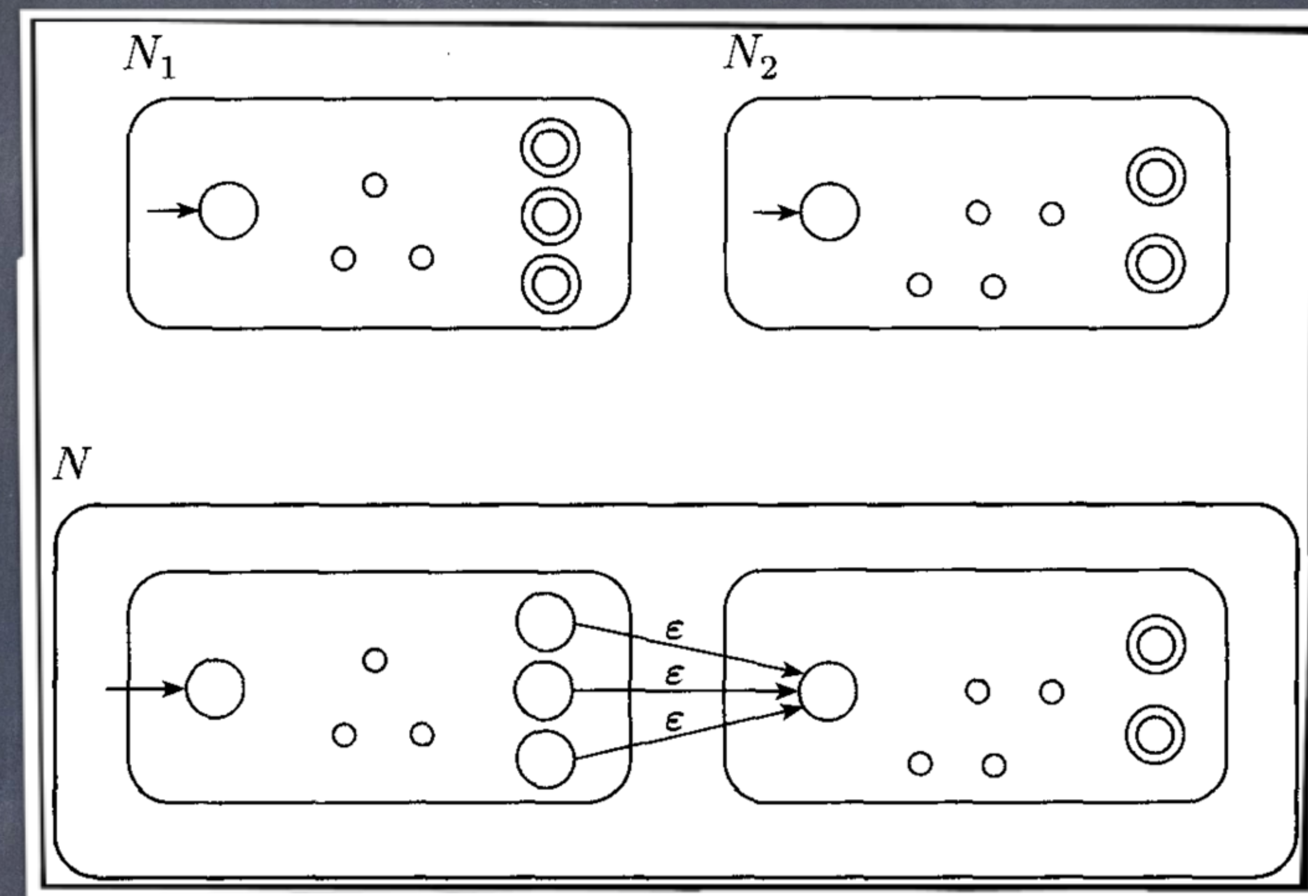
The class of regular languages is closed under the union operation.

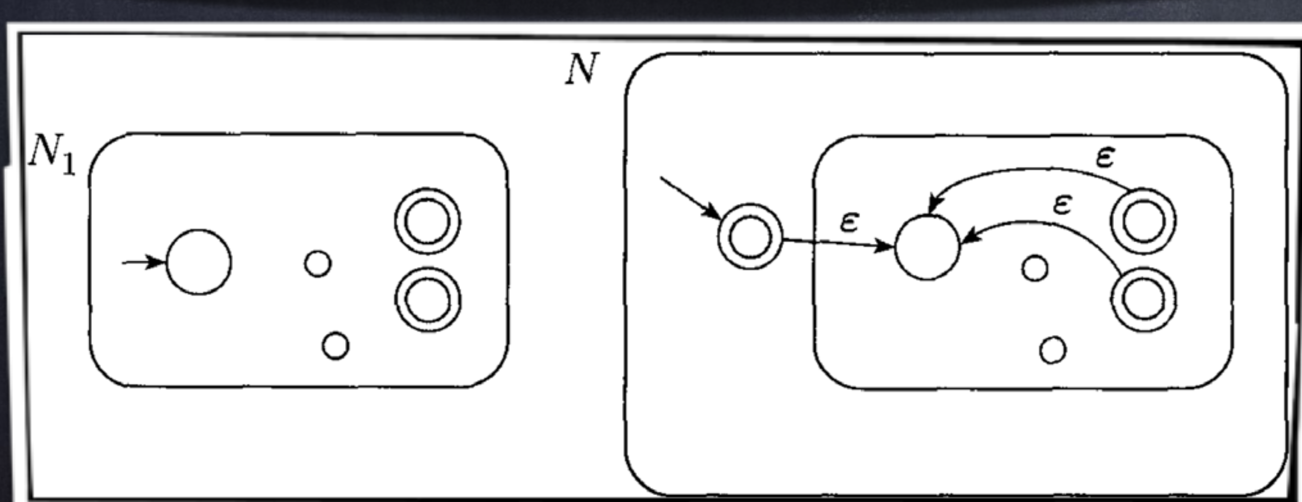
THEOREM 1.47

The class of regular languages is closed under the concatenation operation.

Regular Operations







THEOREM 1.45 ....

The class of regular languages is closed under the union operation.

THEOREM 1.47 ....

The class of regular languages is closed under the concatenation operation.

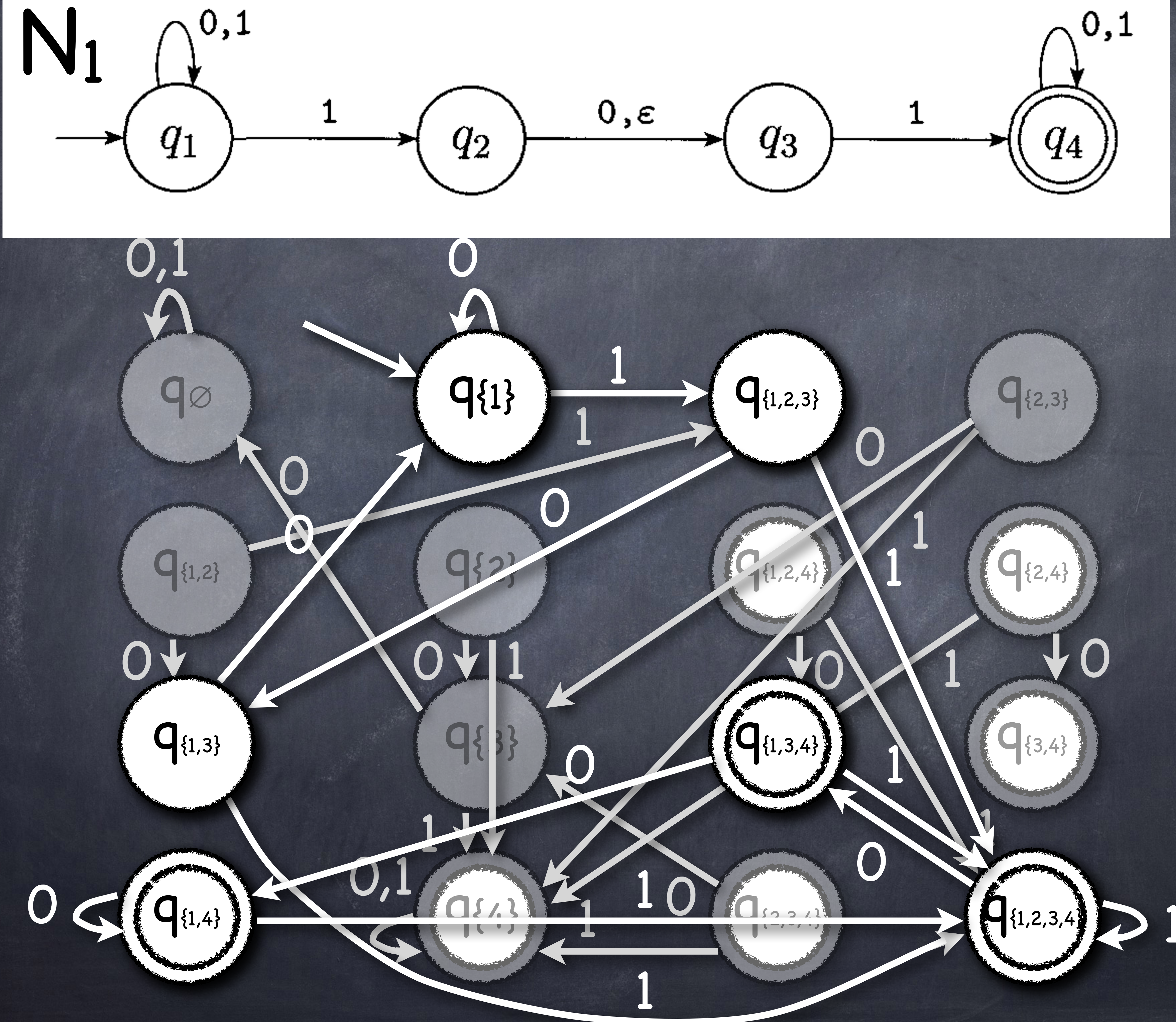
THEOREM 1.49 ....

The class of regular languages is closed under the star operation.

COMP 330 Fall 2012: Lectures Schedule

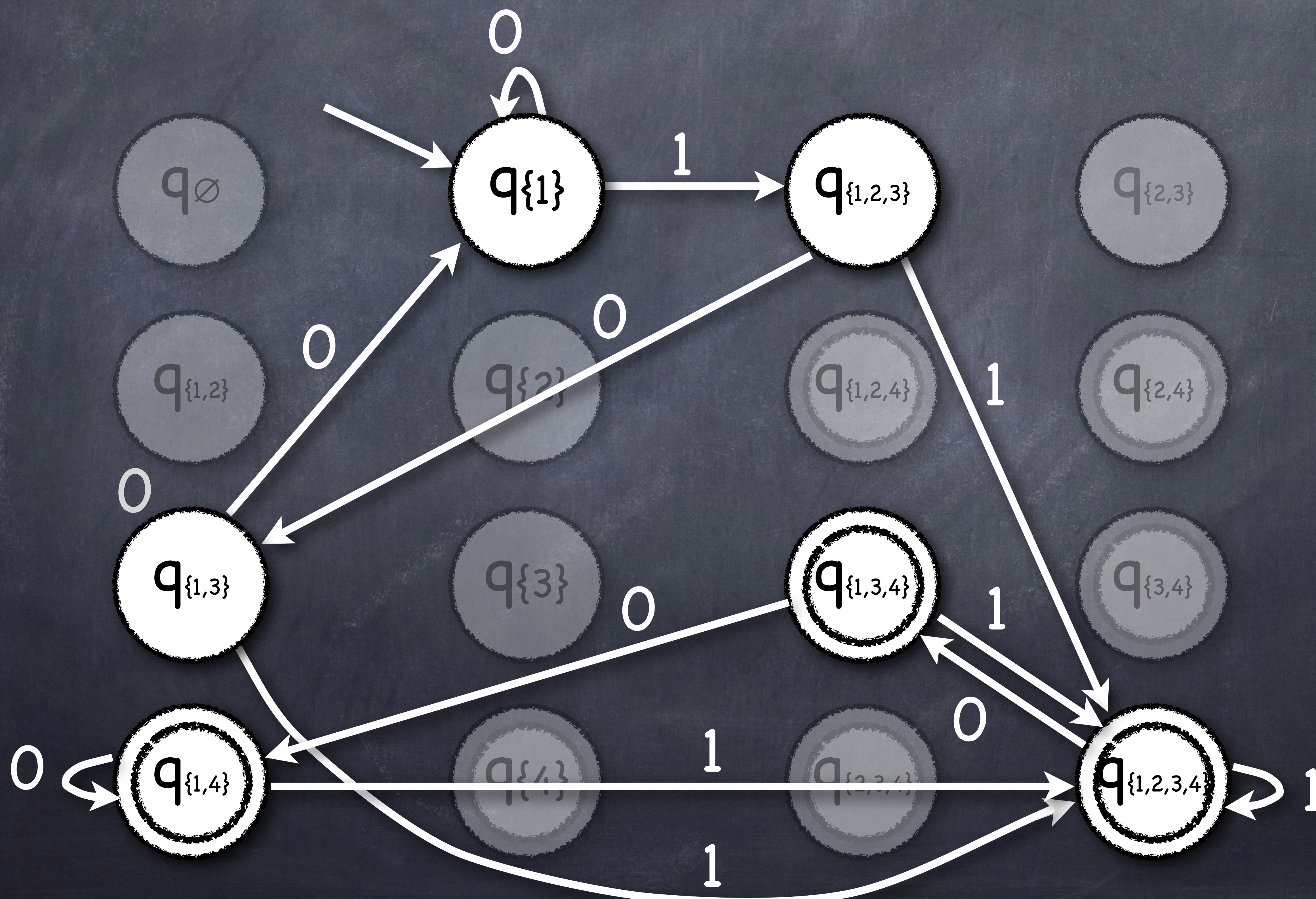
1-2. Introduction
 1.5. Some basic mathematics
2-3. Deterministic finite automata +Closure properties,
4. Nondeterministic finite automata
5. Determinization+Kleene's theorem
6. Myhill-Nerode theorem
7-8. Regular Expressions+GNFA
9. The pumping lemma
10. Duality
11. Labelled transition systems
13. MIDTERM

14. Context-free languages
15. Pushdown automata
16. Parsing
17. The pumping lemma for CFLs
18. Introduction to computability
19. Models of computation
 Basic computability theory
20. Reducibility, undecidability and Rice's theorem
21. Undecidable problems about CFGs
22. Post Correspondence Problem
23. Validity of FOL is RE / Gödel's and Tarski's thms
24. Universality / The recursion theorem
25. Degrees of undecidability
26. Introduction to complexity

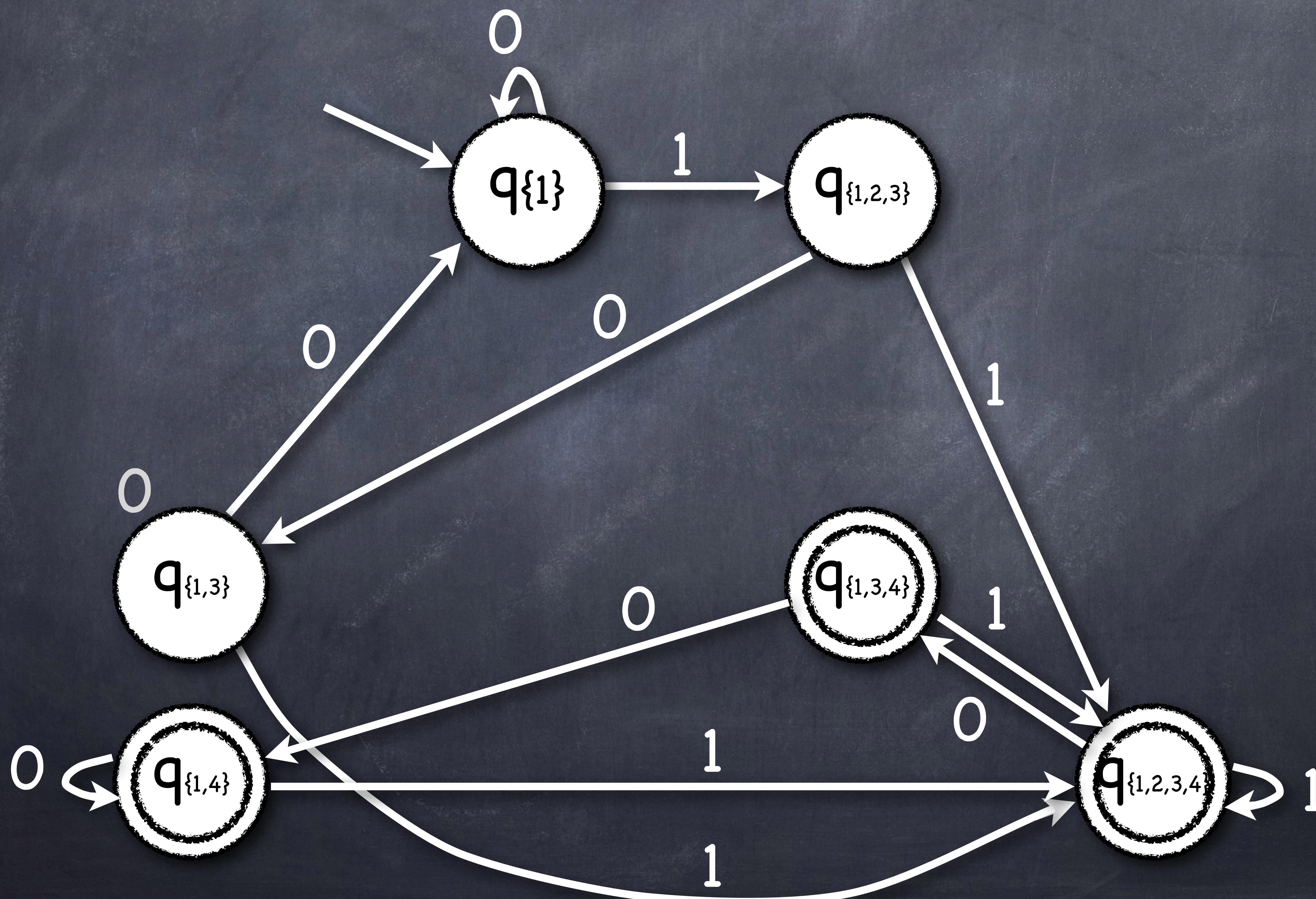


N1
q1
q2
q3
q4
0,1
1
0,ε
1
0,1
q∅
q{1}
q{1,2,3}
q{2,3}
q{1,2}
q{2}
q{1,2,4}
q{2,4}
q{1,3}
q{3}
q{1,3,4}
q{3,4}
q{1,4}
q{4}
q{2,3,4}
q{1,2,3,4}

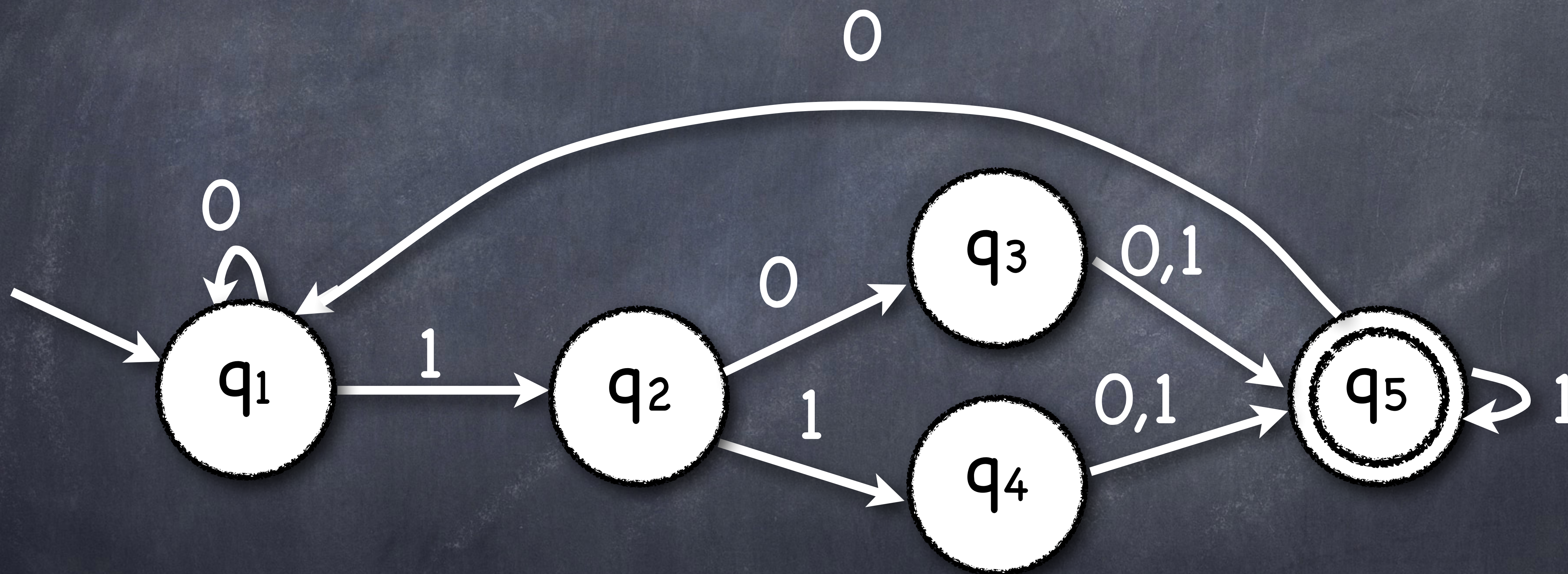
Unreachable States



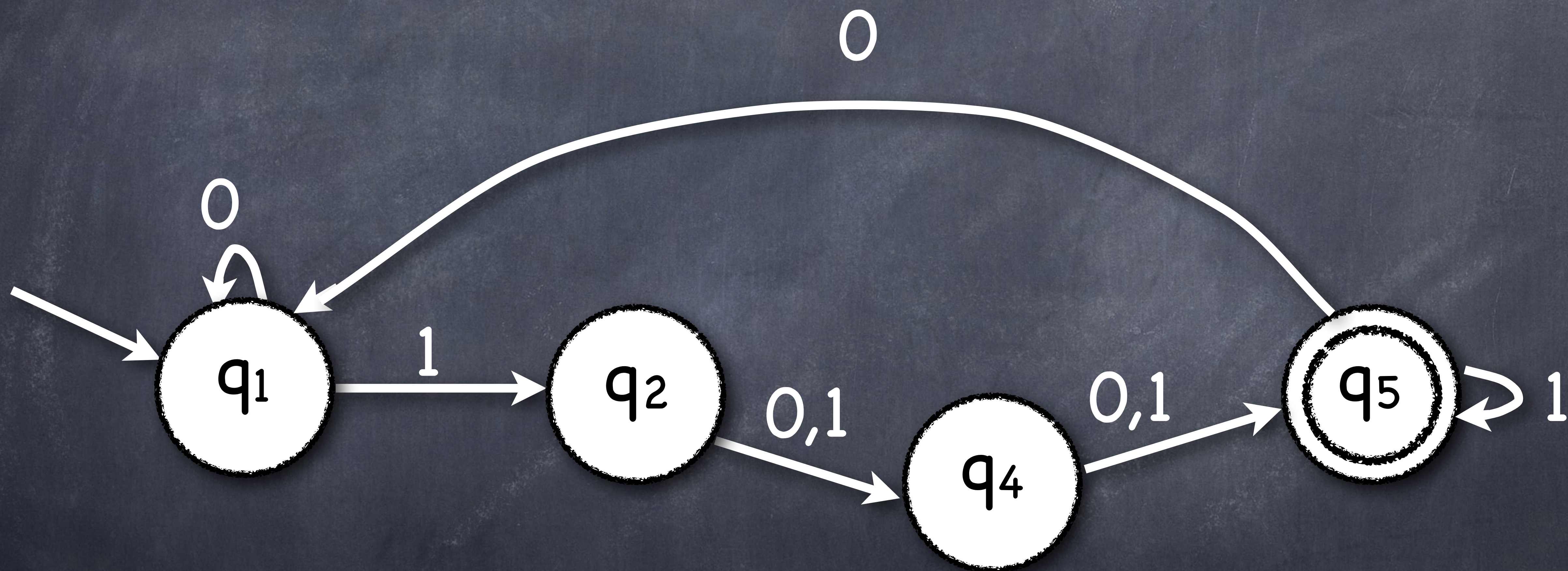
Reachable States



Redondant States



Redondant States



Myhill-Nerode Theorem

John R. Myhill

Anil Nerode

Myhill-Nerode Theorem

Myhill-Nerode Theorem

- Let x and y be strings and L be a language.

Myhill-Nerode Theorem

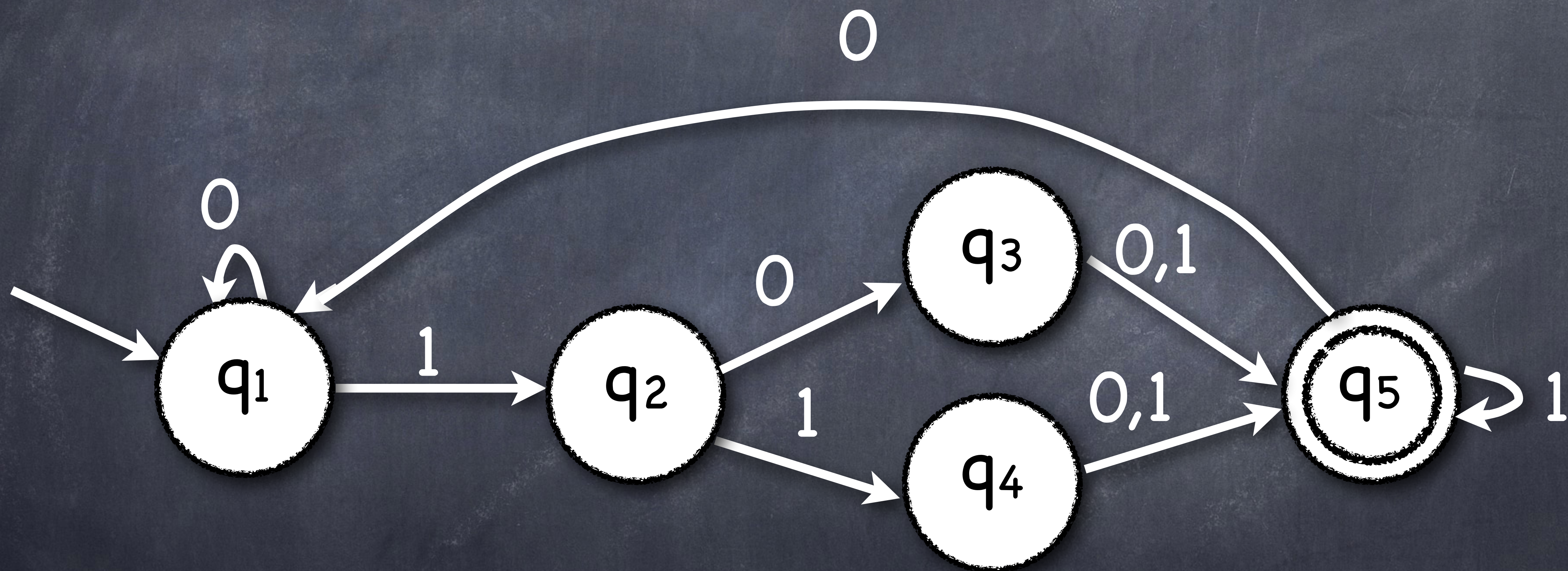
- Let x and y be strings and L be a language.
- We say that x and y are distinguishable by L if there exists a z such that $xz \in L$ and $yz \notin L$ or $yz \in L$ and $xz \notin L$.

Myhill-Nerode Theorem

- Let x and y be strings and L be a language.
- We say that x and y are <u>distinguishable</u> by L if there exists a z such that $xz \in L$ and $yz \notin L$ or $yz \in L$ and $xz \notin L$.
- If x and y are indistinguishable by L we write $x \equiv_L y$, ($\equiv_L$ is an equivalence relation). If x, y are distinguishable by L we write $x \not\equiv_L y$.

Distinguishable Strings

$011 \not\equiv_L 1$

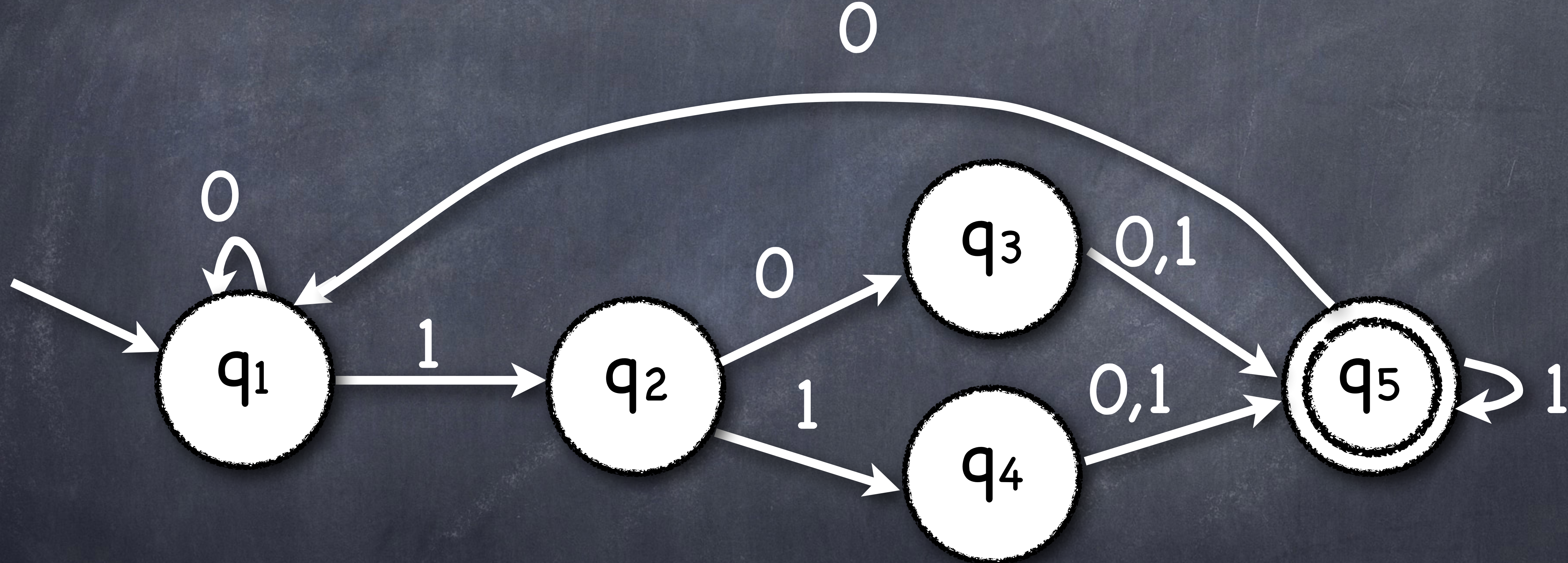


there exists a z such that xz∈L and yz∉L.

z=0 is such that 0110∈L while 10∉L.

Indistinguishable Strings

$011 \equiv_L 010$



There does not exist a z such that xz∈L and yz∉L nor yz∈L and xz∉L. For all z, xz and yz are both in L or neither in L.

Myhill-Nerode Theorem

Myhill-Nerode Theorem

- Let L be a language and X a set of strings.

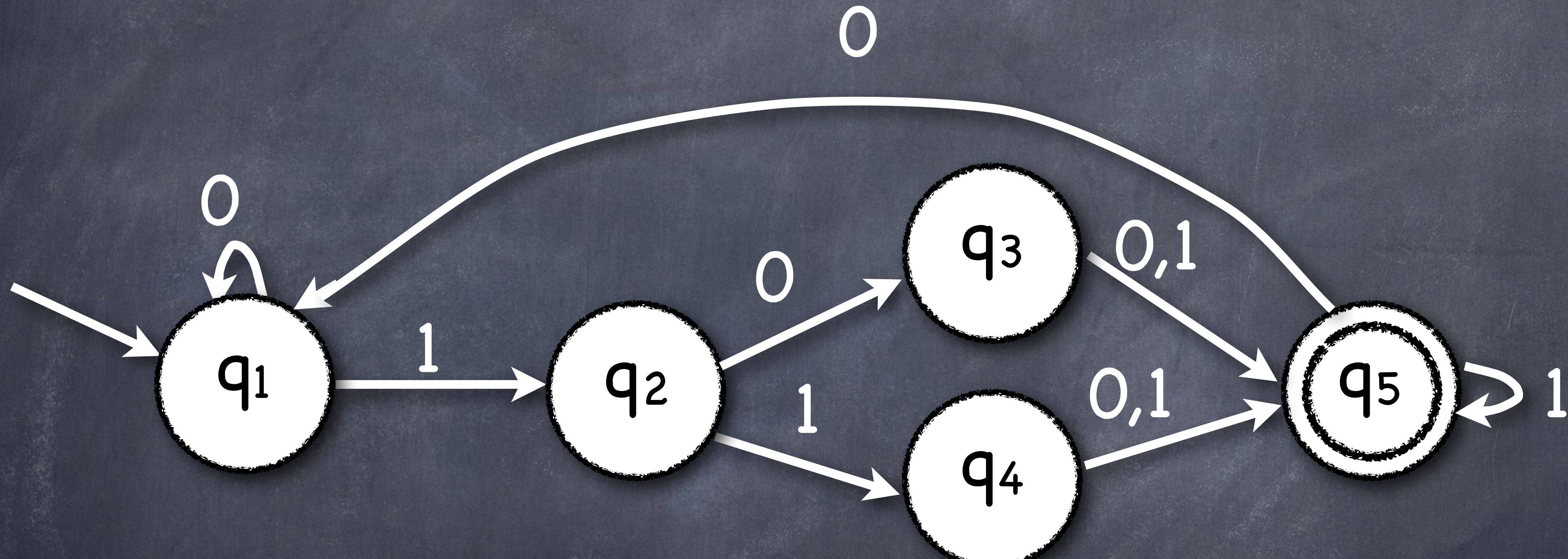
Myhill-Nerode Theorem

- Let L be a language and X a set of strings.
- We say that X is pairwise <u>distinguishable</u> by L if every two elements in X are distinguishable by L (For all x, x′ in X, $x \not\equiv_L x'$).

Myhill-Nerode Theorem

- Let L be a language and X a set of strings.
- We say that X is pairwise distinguishable by L if every two elements in X are distinguishable by L (For all x, x′ in X, $x \not\equiv_L x'$).
- Define the index of L to be the size of a maximum set X that is pairwise distinguishable by L. The index may be finite or infinite.

Distinguishable Strings



While this automaton has 5 states, the index of L is only 4:

ε, 1, 11 and 111

111∈L while 11∉L, 111∈L while 1∉L, 111∈L while ε∉L,

111∈L while 11∉L, 111∈L while 1∉L, 111∈L while 11∉L.

Myhill-Nerode Theorem

Myhill-Nerode Theorem

a. If L is recognized by a DFA with k states, then L has index at most k.

Myhill-Nerode Theorem

a. If L is recognized by a DFA with k states, then L has index at most k.

b. If the index of L is a finite number k, then it is recognized by a DFA with k states.

Myhill-Nerode Theorem

a. If L is recognized by a DFA with k states,
then L has index at most k.

b. If the index of L is a finite number k,
then it is recognized by a DFA with k states.

c. L is regular iff it has finite index.
This index is the size of the smallest DFA recognizing L.

Myhill-Nerode Theorem

a. If L is recognized by a DFA with k states, then L has index at most k.

Myhill-Nerode Theorem

a. If L is recognized by a DFA with k states, then L has index at most k.

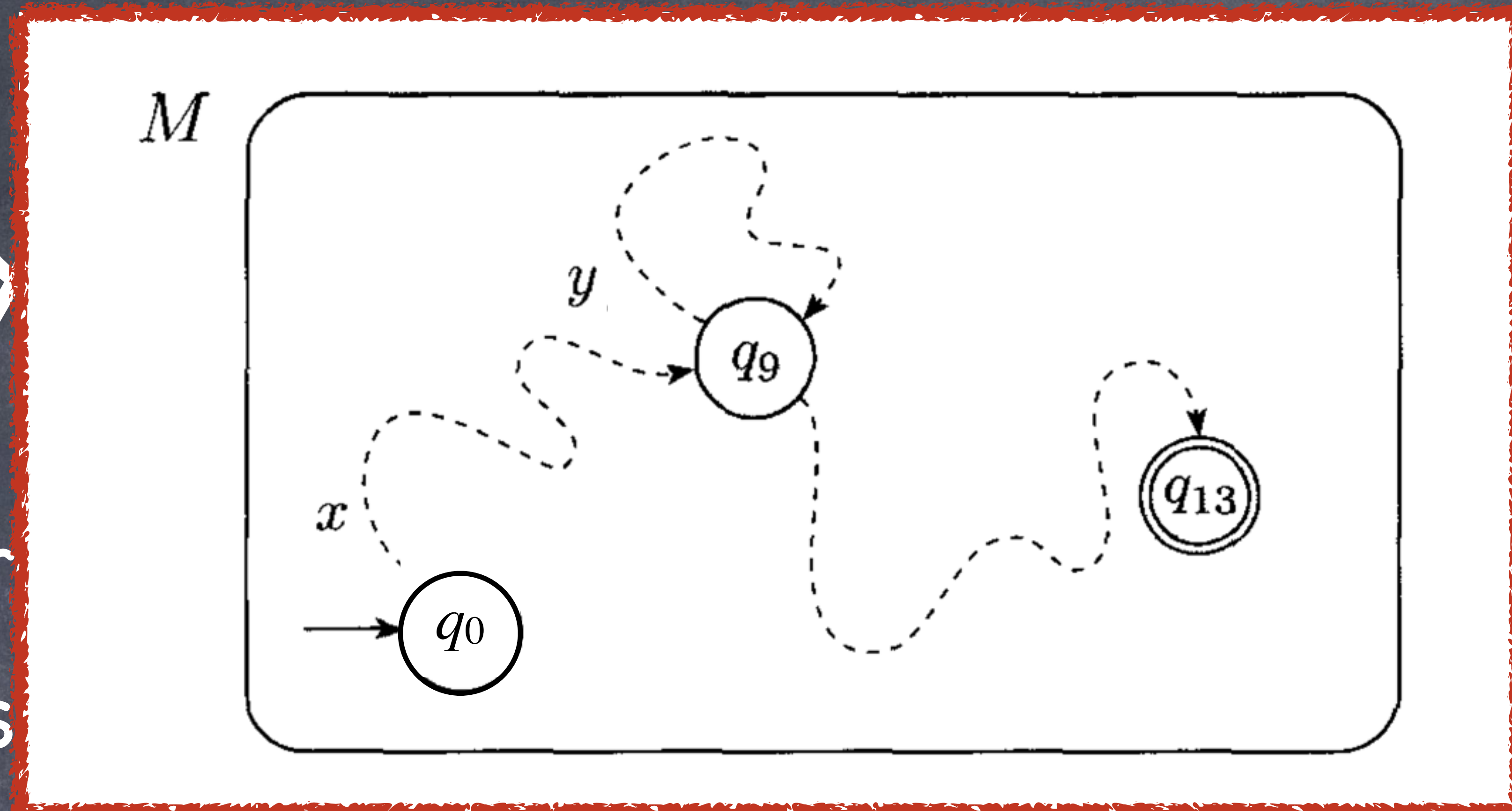
- Let M be a k state DFA recognizing L.

Myhill-Nerode Theorem

a. If L is recognized by a DFA with k states, then L has index at most k.

- Let M be a k state DFA recognizing L.
- Suppose L has index larger than k.

My… …em

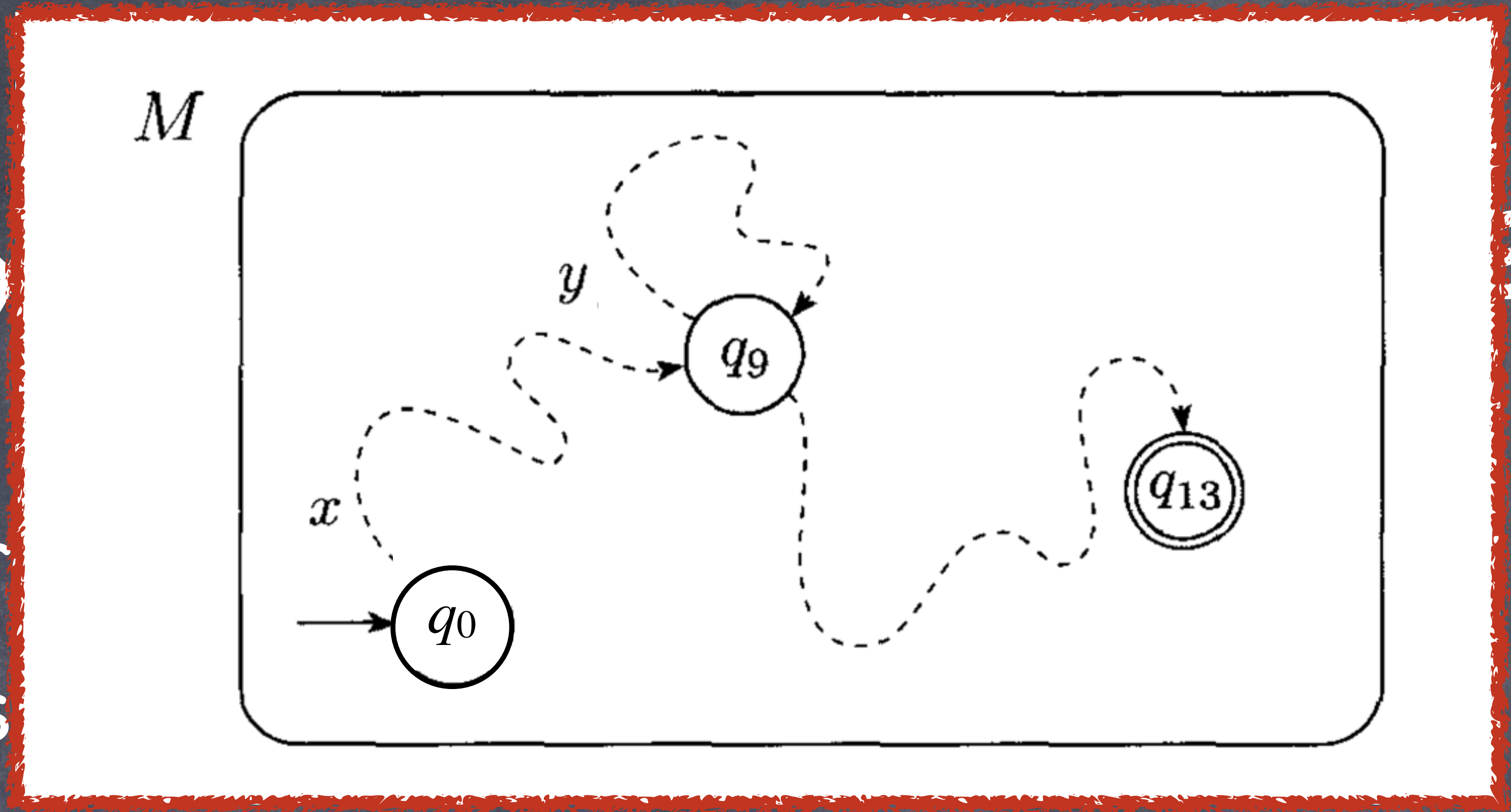


a. If L is r…

then L has…

- Let M be a k state DFA recognizing L.
- Suppose L has index larger than k.

My... ...em



a. If L is r...

then L has...

- Let M be a k state DFA recognizing L.
- Suppose L has index larger than k.
- Some X with k+1 elements is distinguishable by L. But since the number of states < k+1 there must exist x,y in X such that $\delta(q_0,x) = \delta(q_0,y)$. But then, x and y are not distinguishable.
  A contradiction.

Myhill-Nerode Theorem

b. If the index of L is a finite number k,

then it is recognized by a DFA with k states.

Myhill-Nerode Theorem

b. If the index of L is a finite number k,

then it is recognized by a DFA with k states.

- Let $X=\{s_1,\ldots,s_k\}$ be pairwise distinguishable by L.

Myhill-Nerode Theorem

b. If the index of L is a finite number k,

then it is recognized by a DFA with k states.

- Let $X=\{s_1,\ldots,s_k\}$ be pairwise distinguishable by L.
- Let $Q=\{q_1,\ldots,q_k\}$ be the states of a DFA recognizing L and define $\delta(q_i,a)=q_j$ s.t. $s_j \equiv_L s_i a$.

Myhill-Nerode Theorem

b. If the index of L is a finite number k,

then it is recognized by a DFA with k states.

- Let $X=\{s_1,\ldots,s_k\}$ be pairwise distinguishable by L.
- Let $Q=\{q_1,\ldots,q_k\}$ be the states of a DFA recognizing L and define $\delta(q_i,a)=q_j$ s.t. $s_j \equiv_L s_i a$.
- Let q_0 be the q_i s.t. $s_i \equiv_L \varepsilon$. Let $F=\{ q_i \mid s_i \in L \}$.

Myhill-Nerode Theorem

b. If the index of L is a finite number k,

then it is recognized by a DFA with k states.

- Let $X=\{s_1,\ldots,s_k\}$ be pairwise distinguishable by L.
- Let $Q=\{q_1,\ldots,q_k\}$ be the states of a DFA recognizing L and define $\delta(q_i,a)=q_j$ s.t. $s_j \equiv_L s_i a$.
- Let q_0 be the q_i s.t. $s_i \equiv_L \varepsilon$. Let $F=\{ q_i \mid s_i \in L \}$.
- M is s.t. $\{ s \mid \delta(q_0,s)=q_i \} = \{ s \mid s \equiv_L s_i \}$.

Myhill-Nerode Theorem

c. L is regular iff it has finite index.
This index is the size of the smallest DFA recognizing L.

($\Longrightarrow$) L is regular implies the existence of a DFA recognizing L. By (a), L has index at most k.
($\Longleftarrow$) If L has index k then by (b) there exists a DFA with k states (i.e. L is regular).

- As for the minimality, if the index of L is not the size of the minimal DFA then there exists a DFA with index-1 states recognizing L. But this is impossible by part (a).

Minimizing via Myhill-Nerode Theorem

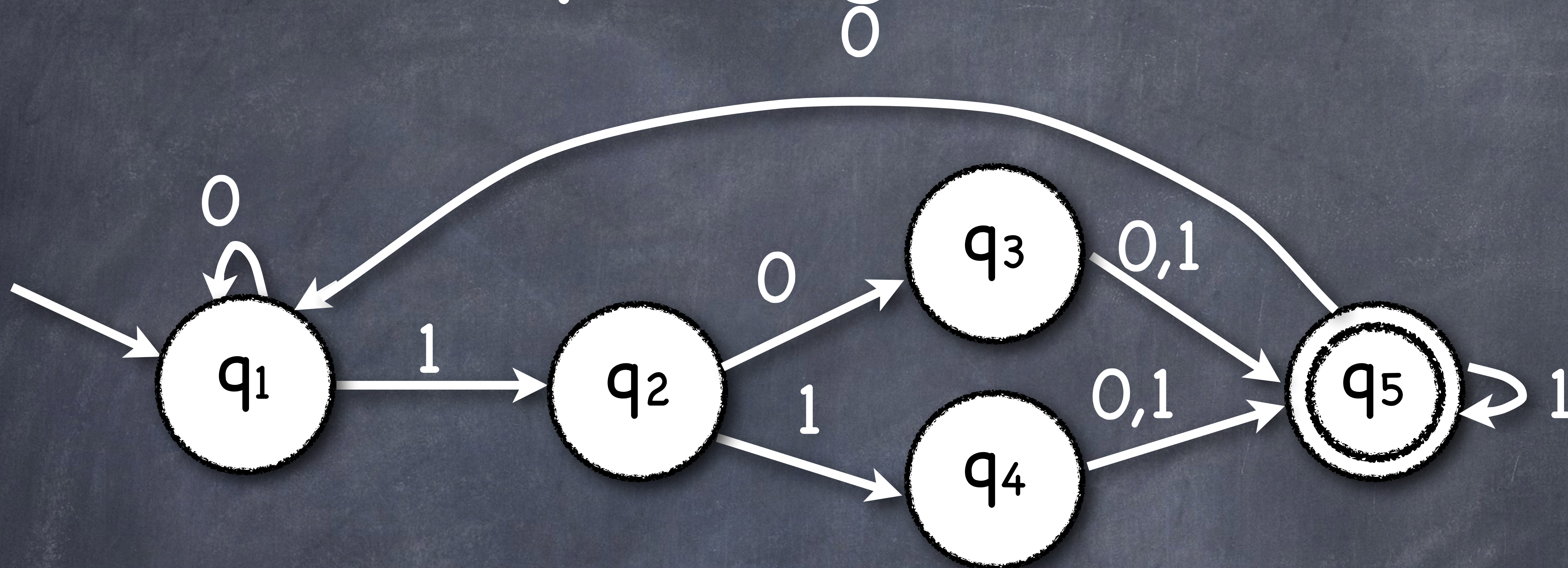
Minimizing via Myhill-Nerode Theorem

- Let L be a regular language. Compute the index of L by finding the set X of all the strings that are pairwise distinguishable by L.

Minimizing via Myhill-Nerode Theorem

- Let L be a regular language. Compute the index of L by finding the set X of all the strings that are pairwise distinguishable by L.
- All strings considered as x, y, xz and yz may be <u>shorter than the number of states</u> of a DFA accepting L. Every string which is longer is equivalent to a shorter one obtained by pumping down.

Computing Index



If we consider all $63 = 2^6 - 1$ strings of length up to 5, we get:

$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

Minimizing via Myhill-Nerode Theorem

- Let L be a regular language. Compute the index of L by finding the set X of all the strings that are pairwise distinguishable by L.
- Using part (b) of the Myhill-Nerode Theorem we construct a minimal DFA to accept L.

Minimal DFA

Minimal DFA

(b) Let $X=\{\varepsilon,1,10,100\}$ be pairwise distinguishable by L.

Minimal DFA

(b) Let $X=\{\varepsilon,1,10,100\}$ be pairwise distinguishable by L.

- Let $Q=\{q_\varepsilon,q_1,q_{10},q_{100}\}$ be the states of a DFA recognizing L

Minimal DFA

(b) Let X={ε,1,10,100} be pairwise distinguishable by L.

- Let $Q=\{q_\varepsilon,q_1,q_{10},q_{100}\}$ be the states of a DFA recognizing L

Minimal DFA

(b) Let X={ε,1,10,100} be pairwise distinguishable by L.

- Let $Q=\{q_\varepsilon,q_1,q_{10},q_{100}\}$ be the states of a DFA recognizing L
- Let q_ε be the initial state and $F=\{ q_{100} \}$.

Minimal DFA

(b) Let $X=\{\varepsilon,1,10,100\}$ be pairwise distinguishable by L.

- Let $Q=\{q_\varepsilon,q_1,q_{10},q_{100}\}$ be the states of a DFA recognizing L
- Let q_ε be the initial state and $F=\{ q_{100} \}$.

Minimal DFA

(b) Let $X=\{\varepsilon,1,10,100\}$ be pairwise distinguishable by L.

- Let $Q=\{q_\varepsilon,q_1,q_{10},q_{100}\}$ be the states of a DFA recognizing L
- Let q_ε be the initial state and $F=\{ q_{100} \}$.
- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.

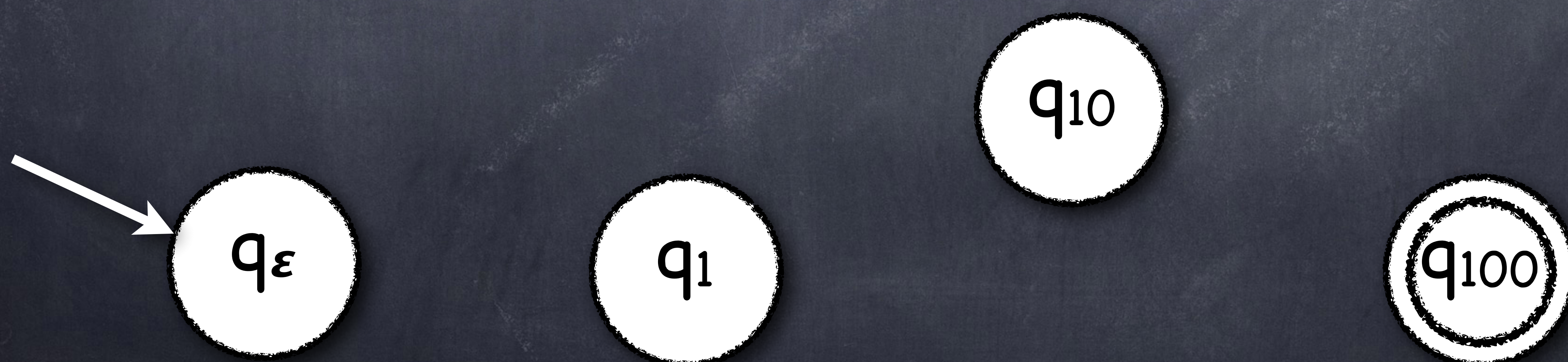
$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.



$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.

$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.

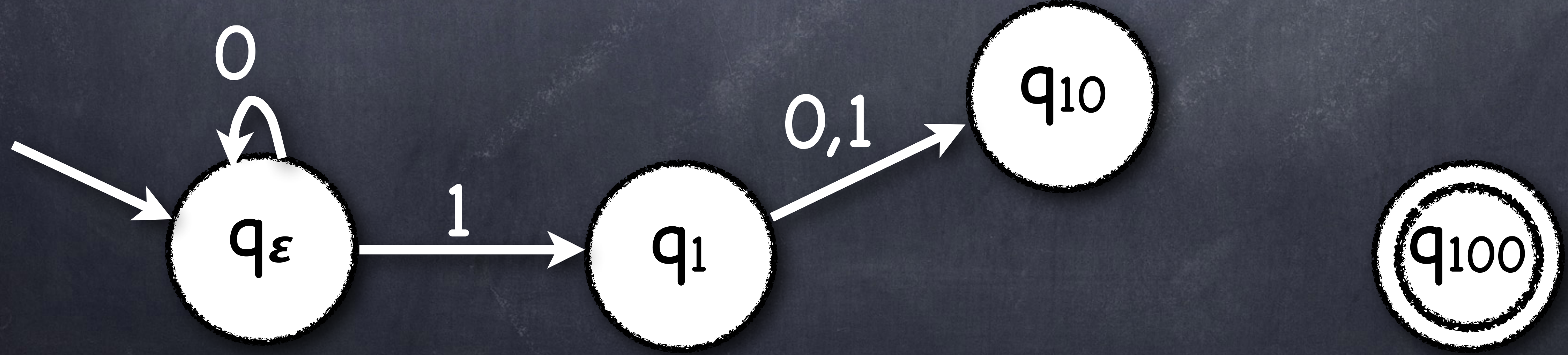
$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.



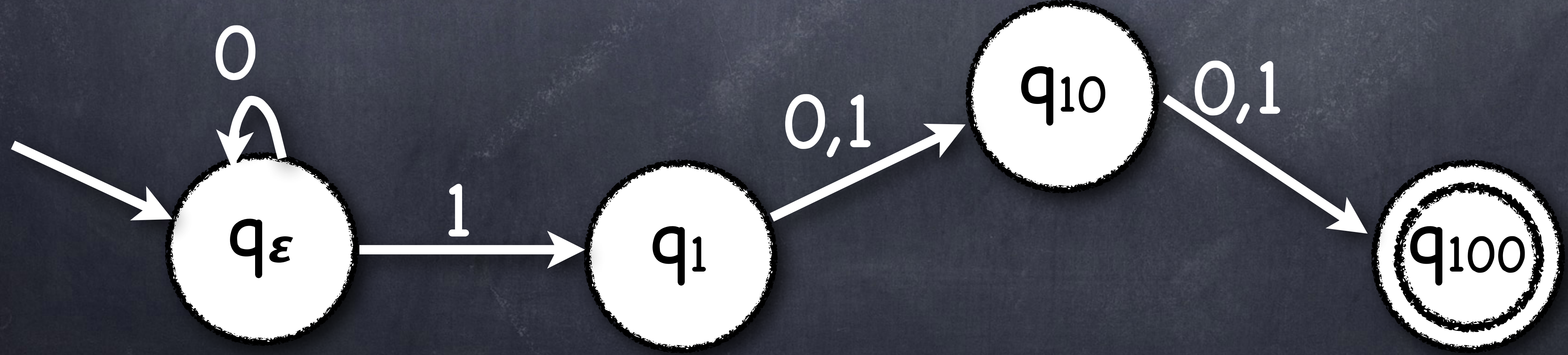
$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.



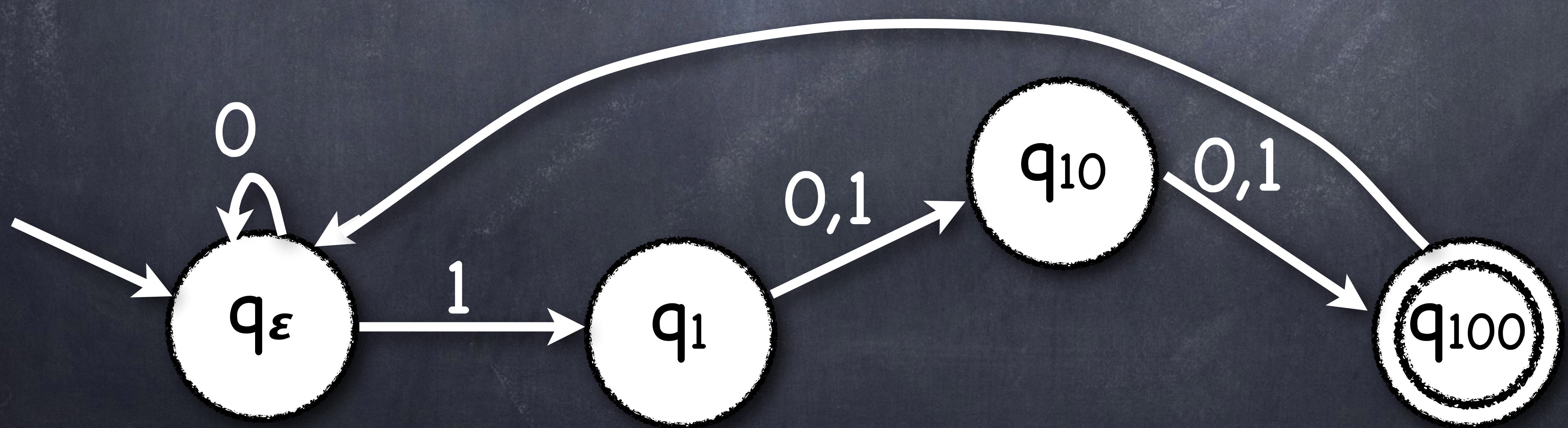
$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.



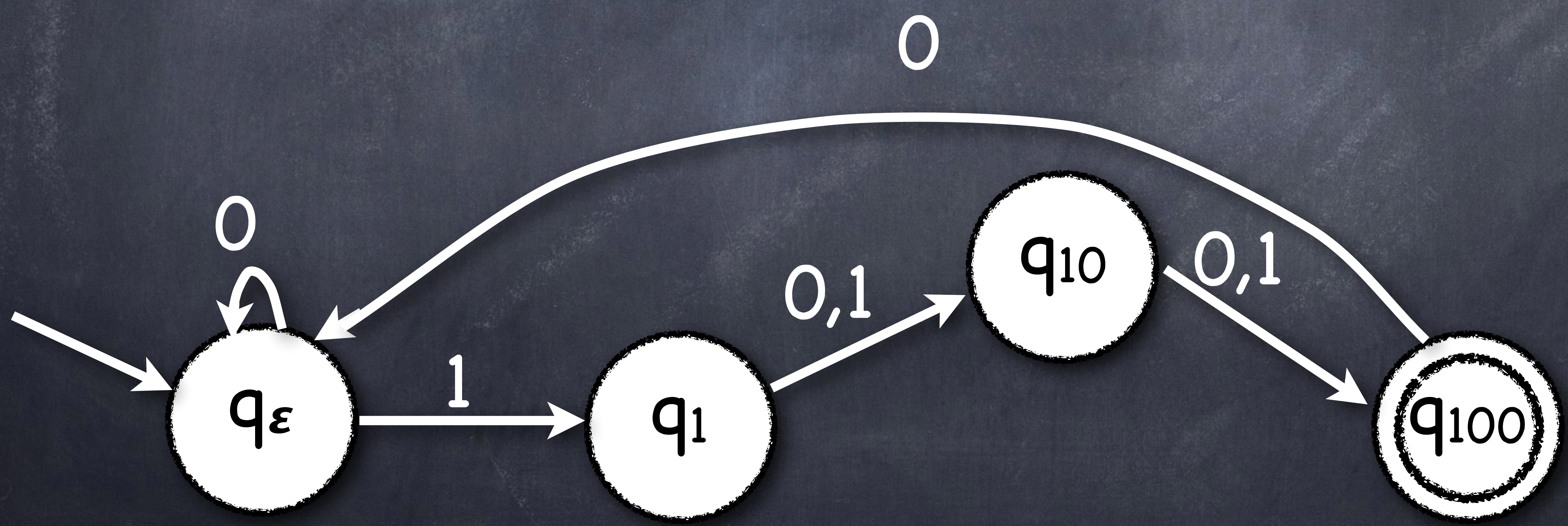
$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w, a) = q_{w'}$ s.t. $w' \equiv_L wa$.



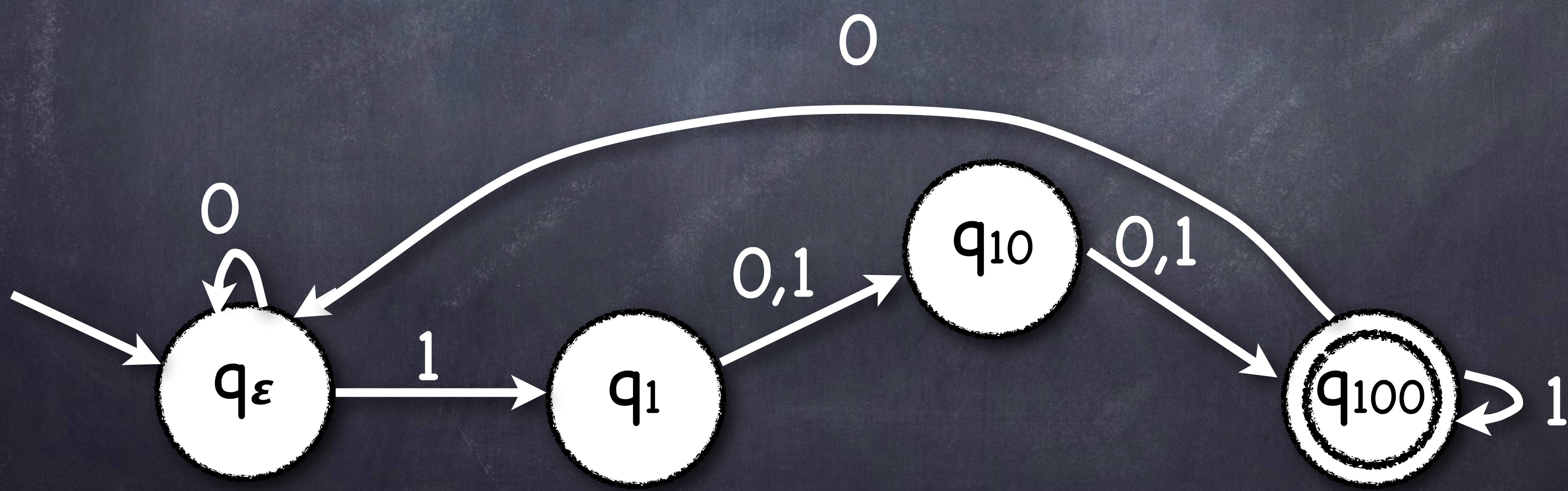
$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w, a) = q_{w'}$ s.t. $w' \equiv_L wa$.



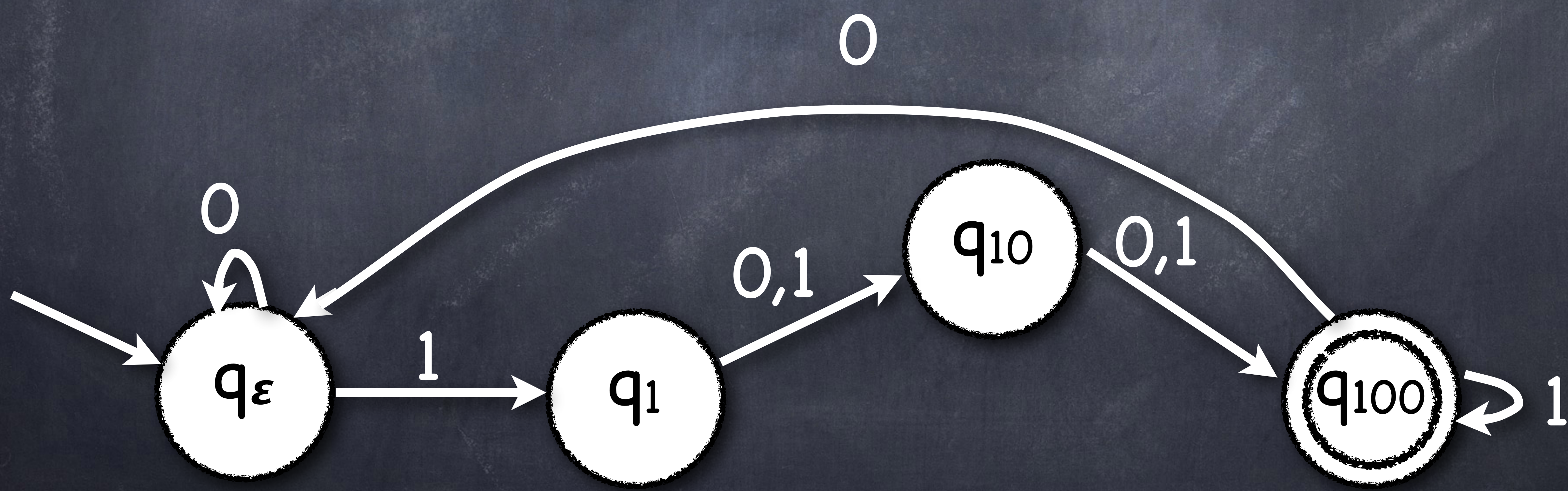
$\varepsilon \equiv_L 0 \equiv_L 00 \equiv_L 000 \equiv_L 0000 \equiv_L 1000 \equiv_L 1010 \equiv_L 1100 \equiv_L 1110 \equiv_L$
$00000 \equiv_L 01000 \equiv_L 01010 \equiv_L 01100 \equiv_L 01110 \equiv_L 10000 \equiv_L 10010 \equiv_L 10100 \equiv_L 10110 \equiv_L 11000 \equiv_L 11010 \equiv_L 11100 \equiv_L 11110$

$1 \equiv_L 01 \equiv_L 001 \equiv_L 0001 \equiv_L 00001 \equiv_L 10001 \equiv_L 10101 \equiv_L 11001 \equiv_L 11101$

$10 \equiv_L 11 \equiv_L 010 \equiv_L 011 \equiv_L 0010 \equiv_L 0011 \equiv_L 00010 \equiv_L 00011$

$100 \equiv_L 101 \equiv_L 110 \equiv_L 111 \equiv_L 0100 \equiv_L 0101 \equiv_L 0110 \equiv_L 0111 \equiv_L 1001 \equiv_L 1011 \equiv_L 1101 \equiv_L 1111 \equiv_L$
$00100 \equiv_L 00101 \equiv_L 00110 \equiv_L 00111 \equiv_L 01001 \equiv_L 01011 \equiv_L 01101 \equiv_L 01111 \equiv_L 10011 \equiv_L 10111 \equiv_L 11011 \equiv_L 11111$

- Define $\delta(q_w,a)=q_{w'}$ s.t. $w' \equiv_L wa$.
- M is s.t. $\{ s \mid \delta(q_\varepsilon,s)=q_w \} = \{ s \mid s \equiv_L w \}$.



Application of the Myhill-Nerode Theorem

$B = \{ 0^n1^n \mid n \geq 0 \}$ is non-regular because it has infinite index.

Consider the set $X = \{ 0^n \mid n \geq 0 \}$. It's an infinite set that is pairwise distinguishable by B.

Proof: For all n, 0^n is distinguishable from all previous 0^i, $0 \leq i \leq n-1$, because there exists a $z = 1^n$ such that $0^n z \in B$ while $0^i z \notin B$, $0 \leq i \leq n-1$.

QED

Application of the Myhill-Nerode Theorem

$F = \{ ww \mid w \in \Sigma^* \}$ is non-regular because it has infinite index.

Consider the set $X = \{ 0^n 1 \mid n \geq 0 \}$. It's an infinite set that is pairwise distinguishable by F.

Proof: For all n, 0^n1 is distinguishable from all previous 0^i1, $0 \leq i \leq n-1$, because there exists a $z = 0^n1$ such that $0^n1z \in B$ while $0^i1z \notin B$, $0 \leq i \leq n-1$.

QED

COMP-330
Theory of Computation

Fall 2019 -- Prof. Claude Crépeau

Lec. 6 : Myhill-Nerode Theorem and applications

Application of the Myhill-Nerode Theorem

Given two regular expressions R and R′ we can find out whether they generate the same regular language or not :

- From R and R′, compute NFAs N and N′ accepting L(R) and L(R′) (Lemma 1.55).
- Compute equivalent DFAs M and M′ (Thm 1.39).
- Using part (b) we construct minimal DFAs W and W′ for each of them.
- L(R)=L(R′) iff W≈W′
  (≈ means "identical up to renaming of states").