Formal Language & Automata Theory

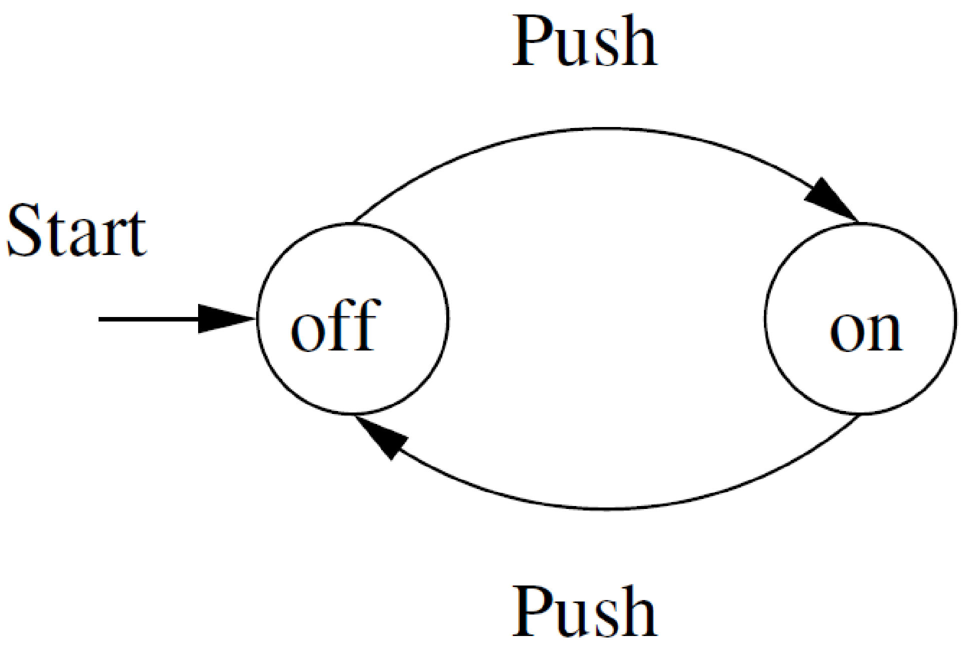
Prof. Sankhadeep Chatterjee

Topics

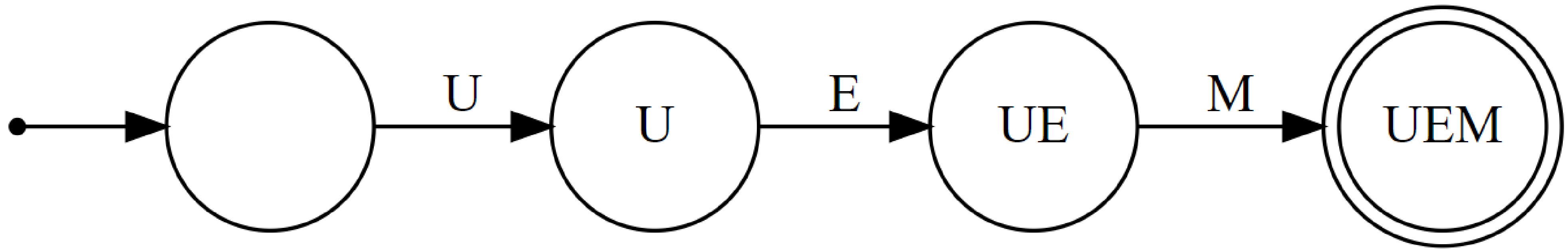
- Finite Automata
  - Deterministic Finite Automata (DFA)
  - Non-Deterministic Finite Automata (NFA)
- Extended Transition Functions
- Regular Languages
- Conversion of NFA to DFA
- Proof of Equivalence of DFA and NFA
- ϵ-Transition NFA
- ϵ-Closure
- ϵ-Transition Removal
- DFA Minimization

Finite Automata

- Revisiting Finite Automata for a “Switch”
- What are the different components?
  - States (e.g. ‘off’, ‘on’)
  - Actions (e.g. ‘Push’)
  - Effects
    - When in state ‘off’, taking action ‘Push’ changes the state to ‘on’
  - Start state

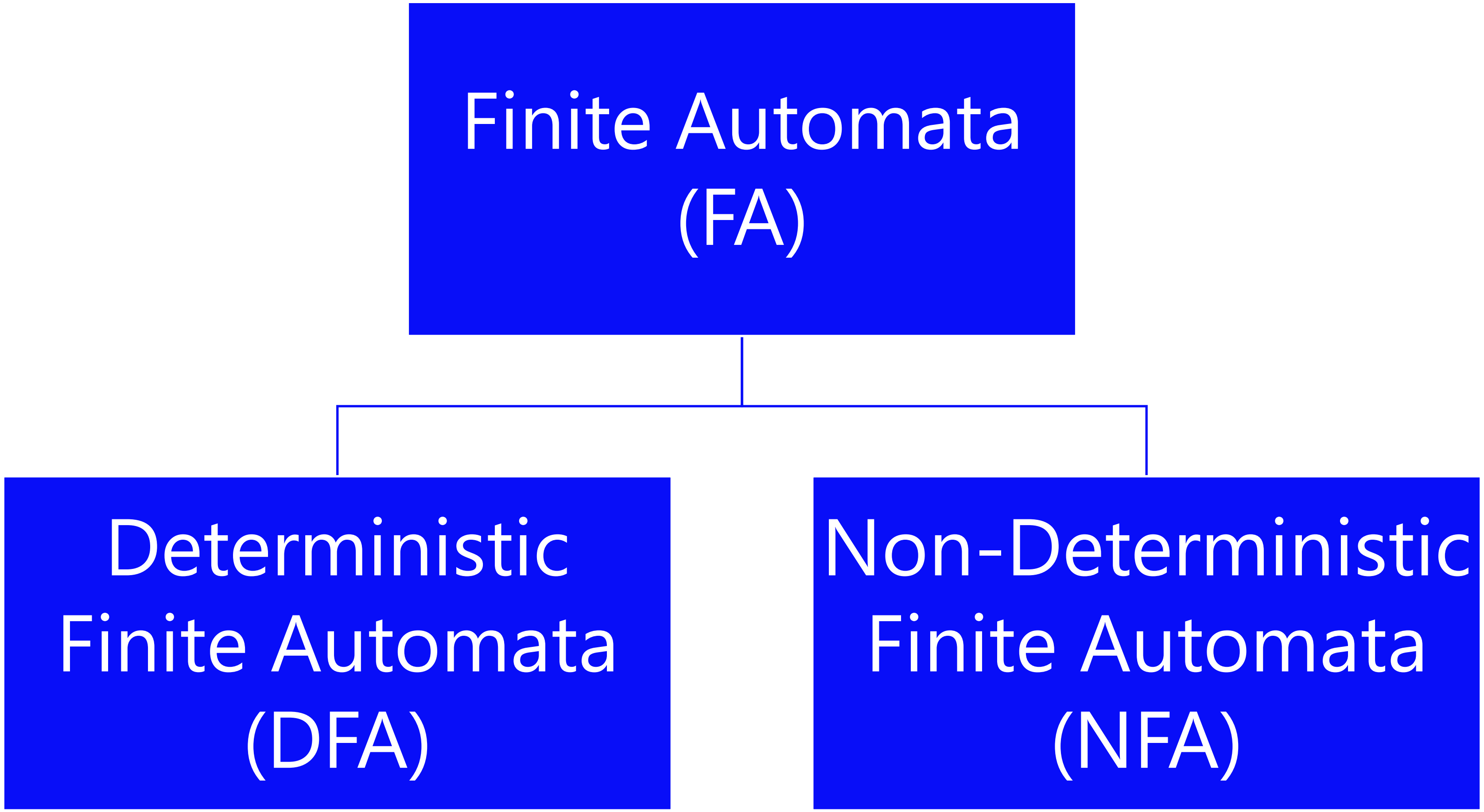


Finite Automata



- States
- Actions
- Effects
- Start state
- Final / Accepting State

Finite Automata



Deterministic Finite Automata

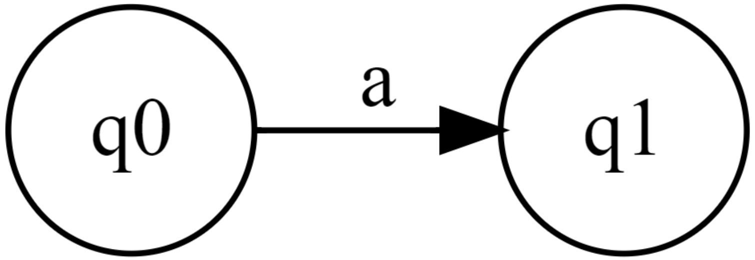
- Deterministic Finite Automata consists of the following:
  - A finite set of states, often denoted Q
  - A finite set of input symbols, often denoted Σ
  - A transition function, denoted by δ
  - A start state, one of the states in Q
  - A set of final or accepting states F $(F \subseteq Q)$
- Deterministic Finite Automata is represented as follows:

$$A = (Q, \Sigma, \delta, q_0, F)$$

Deterministic Finite Automata

- **Transition Function**:

  $\delta(q_0, a) = q_1$ indicates that DFA enters state q_1 from state q_0 after processing input symbol a



- **Acceptance of Strings**:

  For any sequence of input symbols $a_1 a_2 \ldots a_i \ldots a_n$ if $\delta(q_{i-1}, a_i) = q_i$ and $q_n \in F$ then the string is said to be "ACCEPTED" by the DFA else it is called "REJECTED"

Deterministic Finite Automata

- $\boldsymbol{A = (Q, \Sigma, \delta, q_0, F)}$
  - $Q = \{q_0, q_1, q_2\}$
  - $\Sigma = \{0,1\}$
  - $\delta(q_0, 0) = q_1, \delta(q_0, 1) = q_0, \delta(q_1, 1) = q_1, \delta(q_1, 0) = q_2$
  - Start state = q_0
  - $F = \{q_2\}$

Can we represent it in a better way?

Deterministic Finite Automata

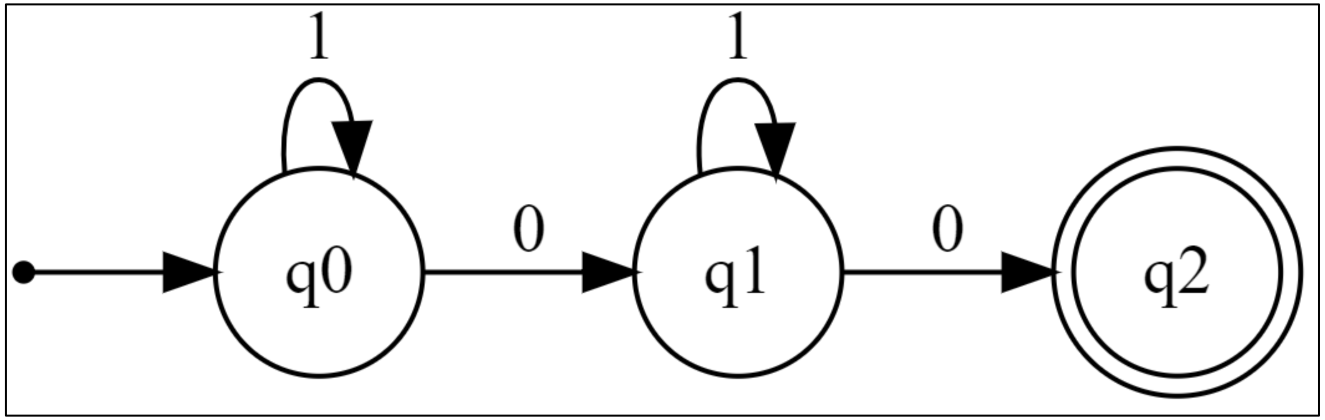
- **DFA representations:**
  - Transition Diagram
  - Transition Table

$$\boldsymbol{A = (\{q_0, q_1, q_2\}, \{0, 1\}, \delta, q_0, \{q_2\})}$$

- **Construction of Transition Diagram**:
  - For each state in Q there is a node
  - For each state p in Q and each symbol a in Σ if $\delta(p, a) = q$
    - There will be an arc from p to q with label a
  - There is an arrow into the start state
  - Final / Accepting states are marked with double circle

Deterministic Finite Automata

- $\boldsymbol{A = (Q, \Sigma, \delta, q_0, F)}$
  - $Q = \{q_0, q_1, q_2\}$
  - $\Sigma = \{0,1\}$
  - $\delta(q_0, 0) = q_1, \delta(q_0, 1) = q_0, \delta(q_1, 1) = q_1, \delta(q_1, 0) = q_2$
  - Start state = q_0
  - $F = \{q_2\}$

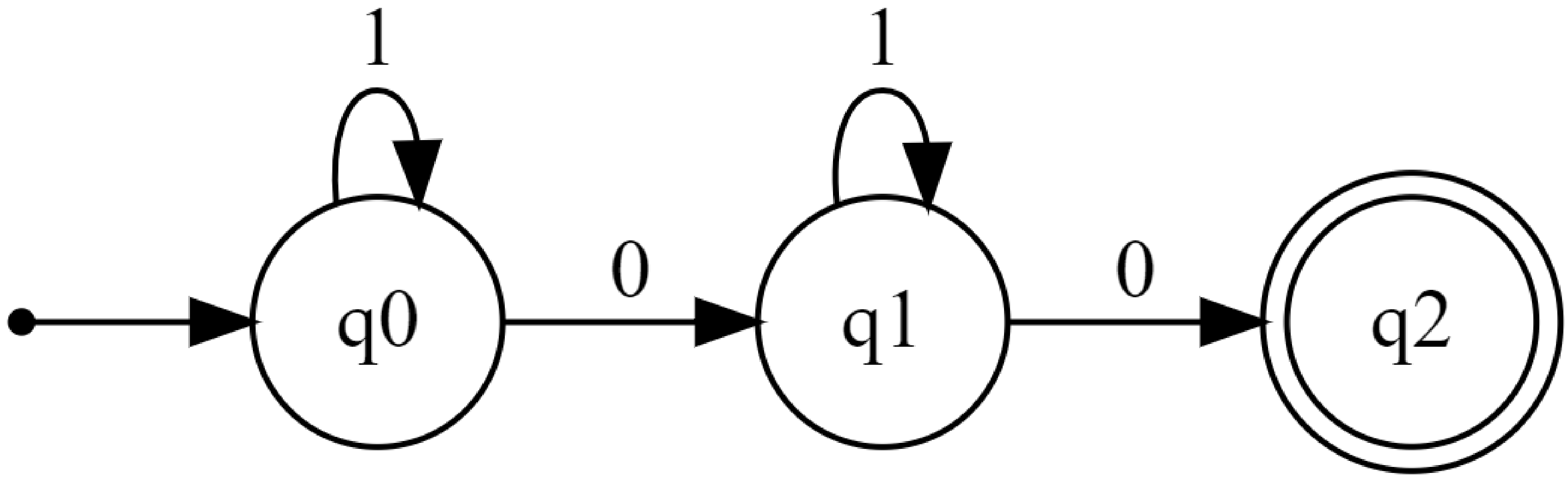


Deterministic Finite Automata

- **Transition Table:**

The entry for the row corresponding to state q and the column correspond input a is the state $\delta(q, a)$

State	0	1
$\rightarrow q_0$	q_1	q_0
q_1	q_2	q_1
$* \; q_2$	$\emptyset$	$\emptyset$



Deterministic Finite Automata

Extended transition Function:

$$\hat{\delta}(q, w) = p$$

- p is the state that the automaton reaches when starting in state q and processing the sequence of inputs w
- Consider $w = xa$ where a is the last symbol and x is a string then

$$\hat{\delta}(q, w) = \hat{\delta}(q, xa) = \delta\big(\hat{\delta}(q, x), a\big)$$

Extended Transition Function

$\hat{\delta}(q_0, 010)$

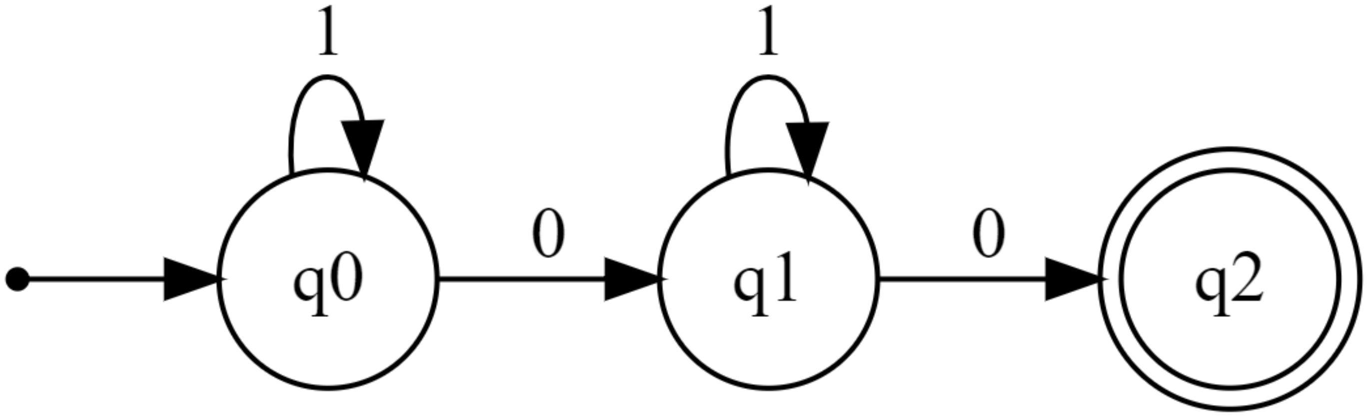
$= \delta\big(\hat{\delta}(q_0, 01), 0\big)$

$= \delta(\delta(\delta(q_0, 0), 1), 0)$

$= \delta(\delta(q_1, 1), 0)$

$= \delta(q_1, 0)$

$= q_2$



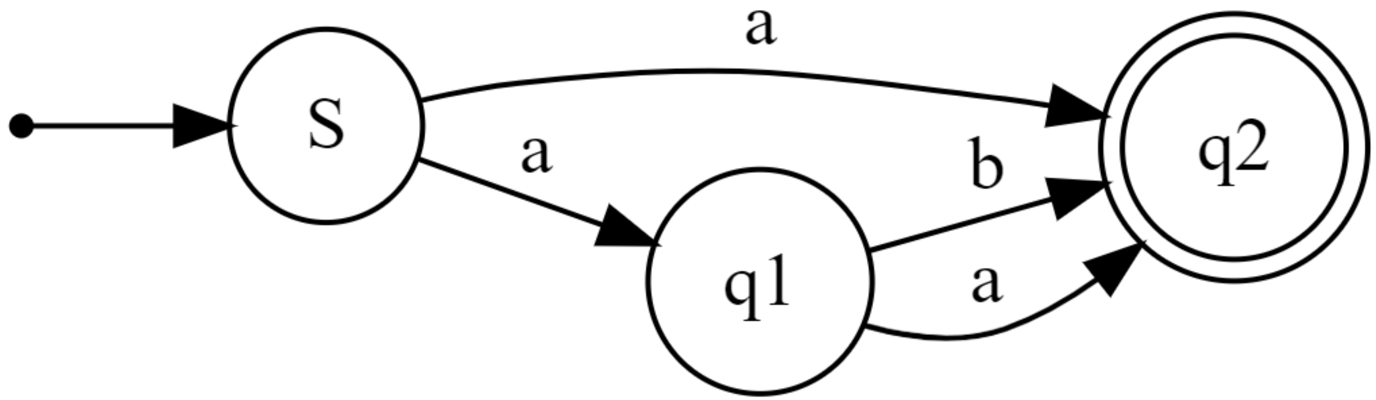
Regular Language

- Given a DFA $A = (Q, \Sigma, \delta, q_0, F)$ the **language of A** is defined as:

$$L(A) = \{w | \hat{\delta}(q_0, w) \text{ is in } F\}$$

- Language of the following DFA?

$$L = \{a, aa, ab\}$$

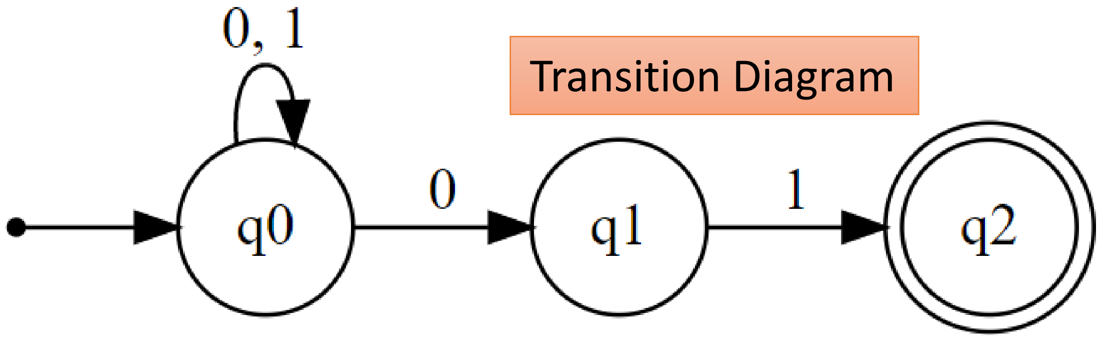


Non-Deterministic Finite Automata (NFA)

- Non-Deterministic Finite Automata consists of the following:
  - A finite set of states, often denoted Q
  - A finite set of input symbols, often denoted Σ
  - A transition function, denoted by δ (**Returns a subset of $\boldsymbol{Q}$**)
  - A start state, one of the states in Q
  - A set of final or accepting states F $(F \subseteq Q)$

Non-Deterministic Finite Automata (NFA)

- **$\boldsymbol{A = (Q, \Sigma, \delta, q_0, F)}$**
  - $Q = \{q_0, q_1, q_2\}$
  - $\Sigma = \{0,1\}$
  - $\delta(q_0, 0) = \{q_0, q_1\}, \delta(q_0, 1) = q_0, \delta(q_1, 1) = q_2$
  - Start state $= q_0$
  - $F = \{q_2\}$

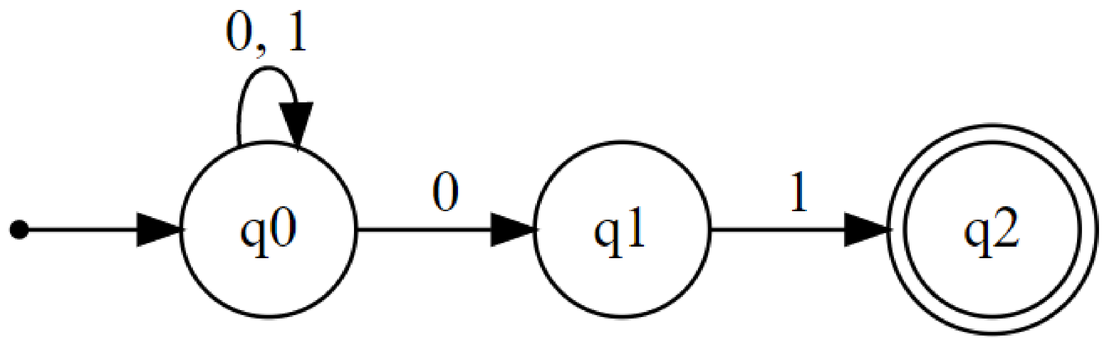


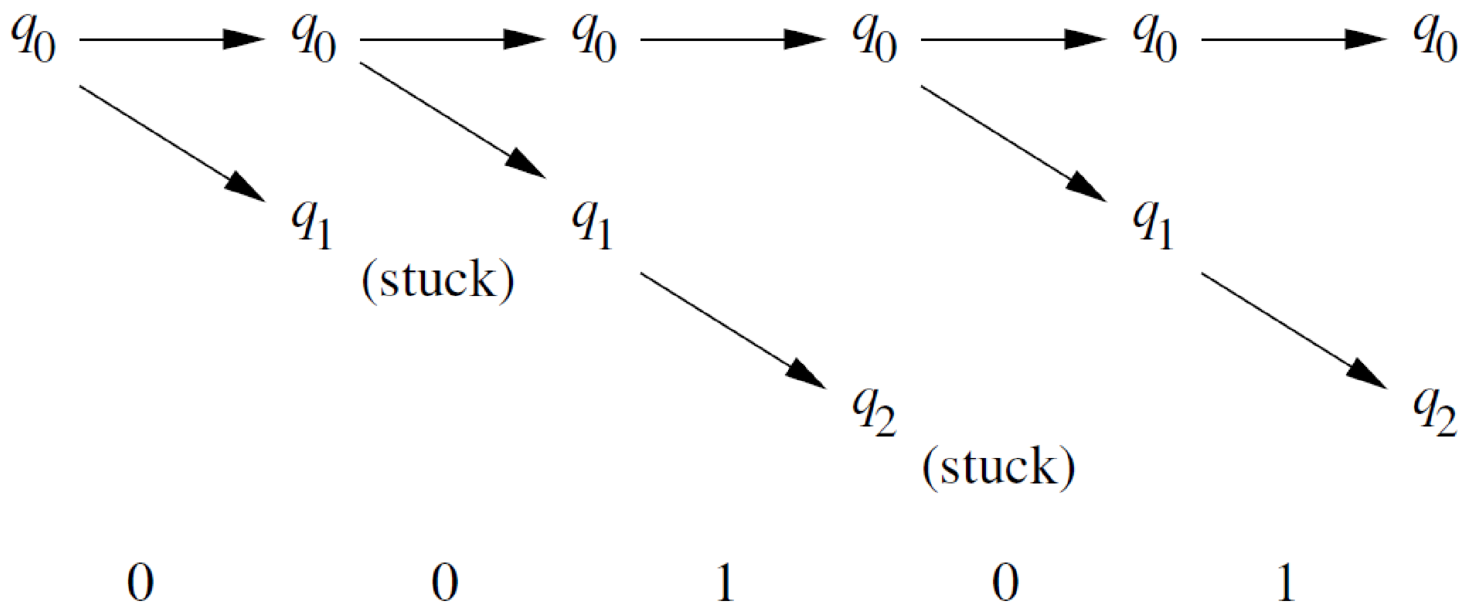
Transition Diagram

Transition Table

	0	1
$\rightarrow q_0$	$\{q_0, q_1\}$	$\{q_0\}$
q_1	$\emptyset$	$\{q_2\}$
$*q_2$	$\emptyset$	$\emptyset$

- How to process string 00101?





Extended Transition Function of NFA

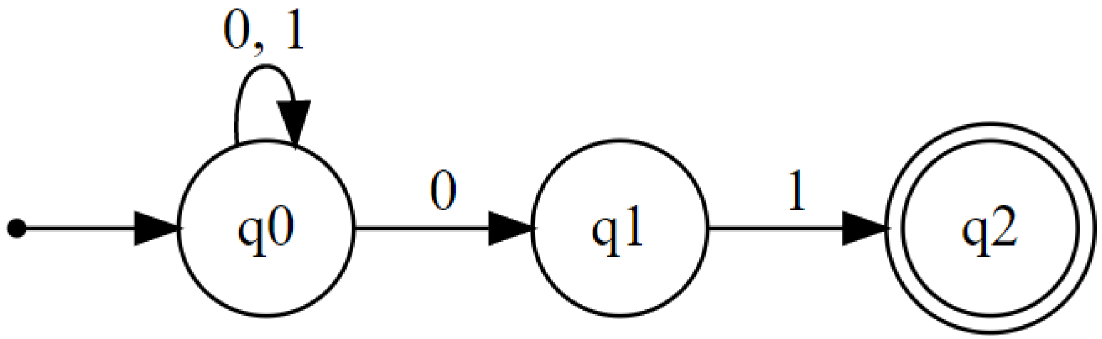
- Consider $w = xa$ where a is the last symbol and x is a rest of w.
- Assume $\hat{\delta}(q, x) = \{p_1, p_2, \ldots p_k\}$

If

$$\bigcup_{i=1}^{k} \delta(p_i, a) = \{r_1, r_2, \ldots r_m\}$$

then

$$\hat{\delta}(q, w) = \{r_1, r_2, \ldots r_m\}$$



	0	1
$\rightarrow q_0$	$\{q_0, q_1\}$	$\{q_0\}$
q_1	$\emptyset$	$\{q_2\}$
$*q_2$	$\emptyset$	$\emptyset$

➢ $\hat{\delta}(q_0, 00)$

$$= \delta\left(\hat{\delta}(q_0, 0), 0\right) = \delta(\{q_0, q_1\}, 0)$$

$$= \bigcup_{i=0}^{1} \delta(q_i, 0) = \delta(q_0, 0) \cup \delta(q_1, 0)$$

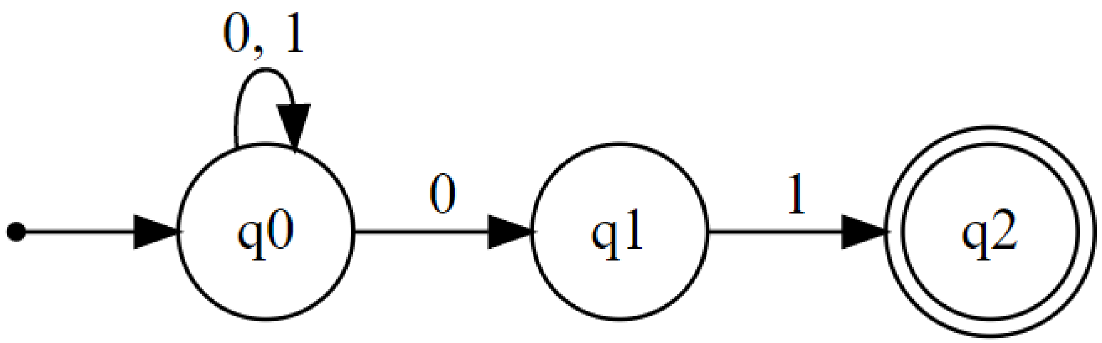
$$= \{q_0, q_1\} \cup \emptyset = \{q_0, q_1\}$$

NFA to DFA Conversion

- Consider DFA $\boldsymbol{D = (Q_D, \Sigma, \delta_D, \{q_0\}, F_D)}$
- Consider NFA $\boldsymbol{N = (Q_N, \Sigma, \delta_N, q_0, F_N)}$
- Conversion:
  - Σ remains same
  - $Q_D \subseteq P(Q_N)$ (Note: $P(Q_N)$ is power set of Q_N)
  - For each set $S \subseteq Q_N$ and for each a in Σ
$$\delta_D(S, a) = \bigcup_{p\ in\ S} \delta_N(p, a)$$
  - Start state of D is a set that contains only the start state of N
  - $F_D \subseteq P(Q_N)$ and for each $S \in F_D$, $S \cap F_N \neq \emptyset$

NFA to DFA Conversion

Subset Construction:

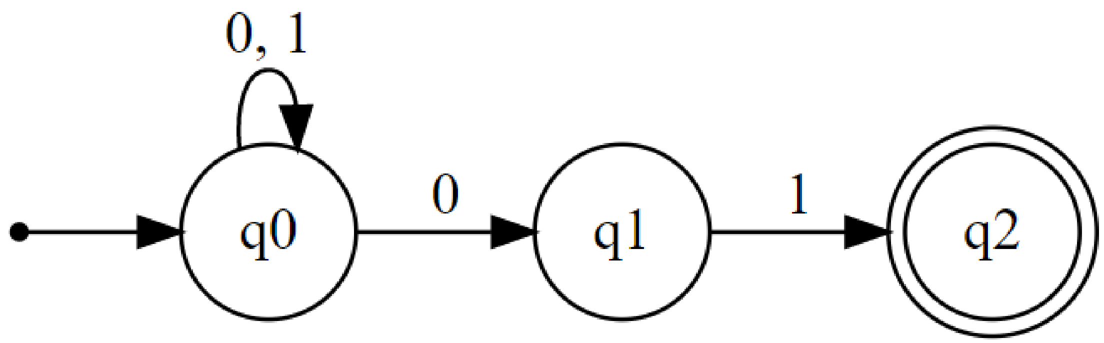


	0	1
$\rightarrow q_0$	$\{q_0, q_1\}$	$\{q_0\}$
q_1	$\emptyset$	$\{q_2\}$
$*q_2$	$\emptyset$	$\emptyset$

	0	1
$\emptyset$	$\emptyset$	$\emptyset$
$\rightarrow \{q_0\}$	$\{q_0, q_1\}$	$\{q_0\}$
$\{q_1\}$	$\emptyset$	$\{q_2\}$
$*\{q_2\}$	$\emptyset$	$\emptyset$
$\{q_0, q_1\}$	$\{q_0, q_1\}$	$\{q_0, q_2\}$
$*\{q_0, q_2\}$	$\{q_0, q_1\}$	$\{q_0\}$
$*\{q_1, q_2\}$	$\emptyset$	$\{q_2\}$
$*\{q_0, q_1, q_2\}$	$\{q_0, q_1\}$	$\{q_0, q_2\}$

Can we improve NFA to DFA conversion?

NFA to DFA Conversion

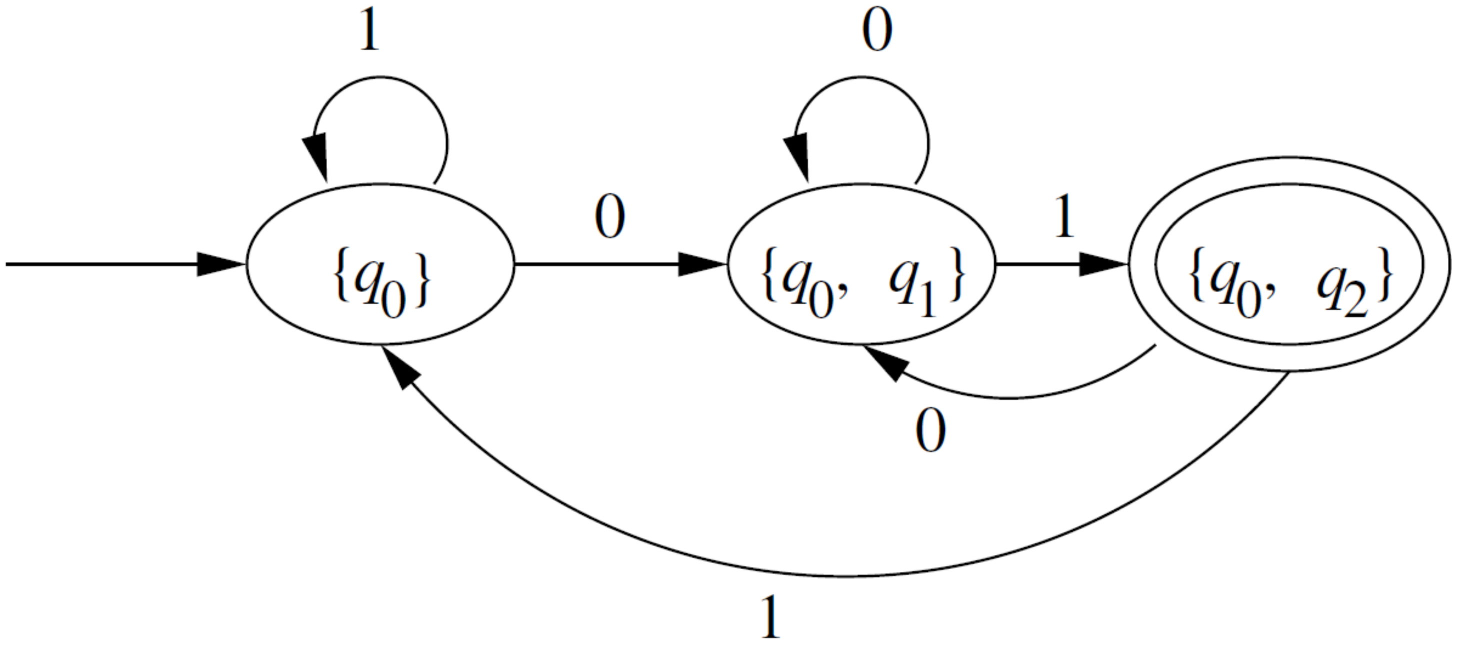


	0	1
$\rightarrow q_0$	$\{q_0, q_1\}$	$\{q_0\}$
q_1	$\emptyset$	$\{q_2\}$
$*q_2$	$\emptyset$	$\emptyset$

	0	1
$\rightarrow \{q_0\}$	$\{q_0, q_1\}$	$\{q_0\}$
$\{q_0, q_1\}$	$\{q_0, q_1\}$	$\{q_0, q_2\}$
$*\{q_0, q_2\}$	$\{q_0, q_1\}$	$\{q_0\}$

NFA to DFA Conversion

	0	1
$\rightarrow \{q_0\}$	$\{q_0, q_1\}$	$\{q_0\}$
$\{q_0, q_1\}$	$\{q_0, q_1\}$	$\{q_0, q_2\}$
$*\{q_0, q_2\}$	$\{q_0, q_1\}$	$\{q_0\}$



Equivalence of NFA and DFA

If DFA $D = (Q_D, \Sigma, \delta_D, \{q_0\}, F_D)$ is constructed from NFA $N = (Q_N, \Sigma, \delta_N, q_0, F_N)$ using subset construction then $L(D) = L(N)$

Proof Outline:

- Show the following by induction on $|w|$

$$\hat{\delta}_D(\{q_0\}, w) = \hat{\delta}_N(q_0, w)$$

Equivalence of NFA and DFA

- **Basis**

  Consider $|w| = 0$, then $w = \epsilon$

  Both $\hat{\delta}_D(\{q_0\}, \epsilon)$ and $\hat{\delta}_N(q_0, \epsilon)$ are $\{q_0\}$

  Hence, $\hat{\delta}_D(\{q_0\}, w) = \hat{\delta}_N(q_0, w)$

- **Induction**

  Assume that the $\hat{\delta}_D(\{q_0\}, w) = \hat{\delta}_N(q_0, w)$ holds for $|w| = n$

  Consider, $|\mathrm{w}| = \mathrm{n} + 1$ and $w = xa$ where a is the final symbol of w

  Hence, $|x| = n$ thus by induction hypothesis $\hat{\delta}_D(\{q_0\}, x) = \hat{\delta}_N(q_0, x)$

Equivalence of NFA and DFA

Consider that $\hat{\delta}_D(\{q_0\}, x) = \hat{\delta}_N(q_0, x) = \{p_1, p_2, \ldots p_k\}$

Further we get;

$$\hat{\delta}_N(q_0, w) = \hat{\delta}_N(q_0, xa) = \delta_N\big(\hat{\delta}_N(q_0, x), a\big) = \bigcup_{i=1}^{k} \delta_N(p_i, a)$$

Now, from subset construction we get;

$$\delta_D(\{p_1, p_2, \ldots p_k\}, a) = \bigcup_{i=1}^{k} \delta_N(p_i, a)$$

Equivalence of NFA and DFA

$$\hat{\delta}_D(q_0, w) = \hat{\delta}_D(\{q_0\}, xa) = \delta_D\big(\hat{\delta}_D(\{q_0\}, x), a\big)$$

$$= \delta_D(\{p_1, p_2, \dots p_k\}, a) = \bigcup_{i=1}^{k} \delta_N(p_i, a)$$

Both D and N accepts w iff $\hat{\delta}_D(\{q_0\}, w)$ and $\hat{\delta}_N(q_0, w)$ contains a state in F_N.

Hence, $L(D) = L(N)$

Operations on DFA

Complement of DFA:

If $D = (Q, \Sigma, \delta, q_0, F)$ then complement of D is denoted as follows:

$$\overline{D} = (Q, \Sigma, \delta, q_0, Q - F)$$

Non final states become final states and vice-versa.

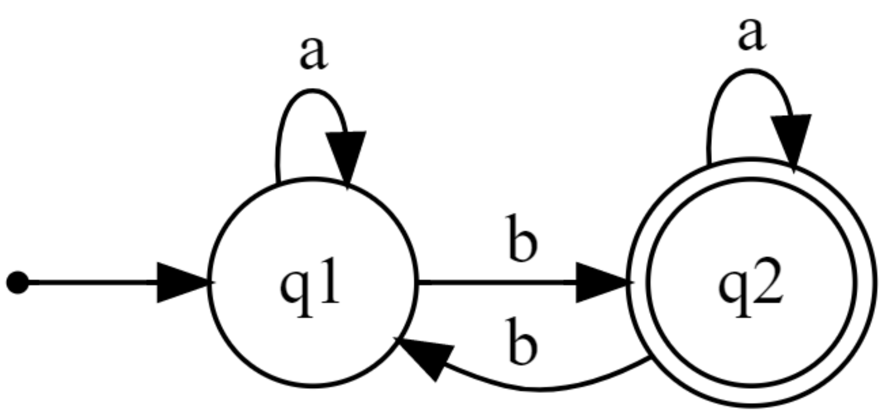
Every thing else remains same.

Operations on DFA

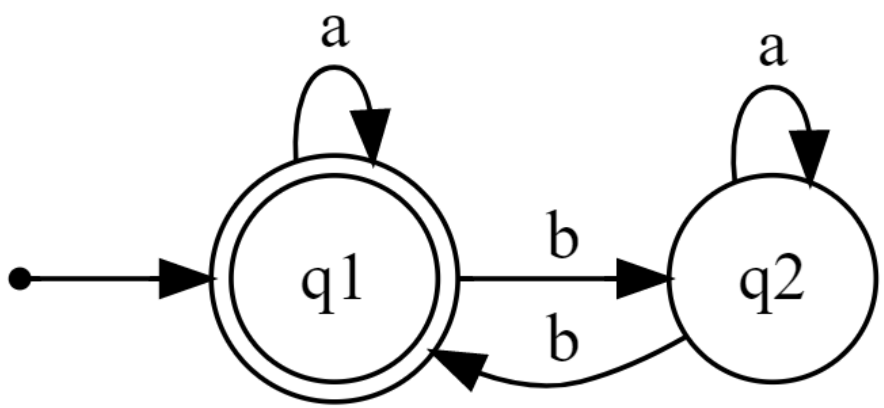
Complement of DFA (Example):

Design a DFA that accepts the following language and find its complement:

$L = \{w |\ w\ contains\ odd\ number\ of\ \boldsymbol{b}\}$ over $\Sigma = \{a, b\}$



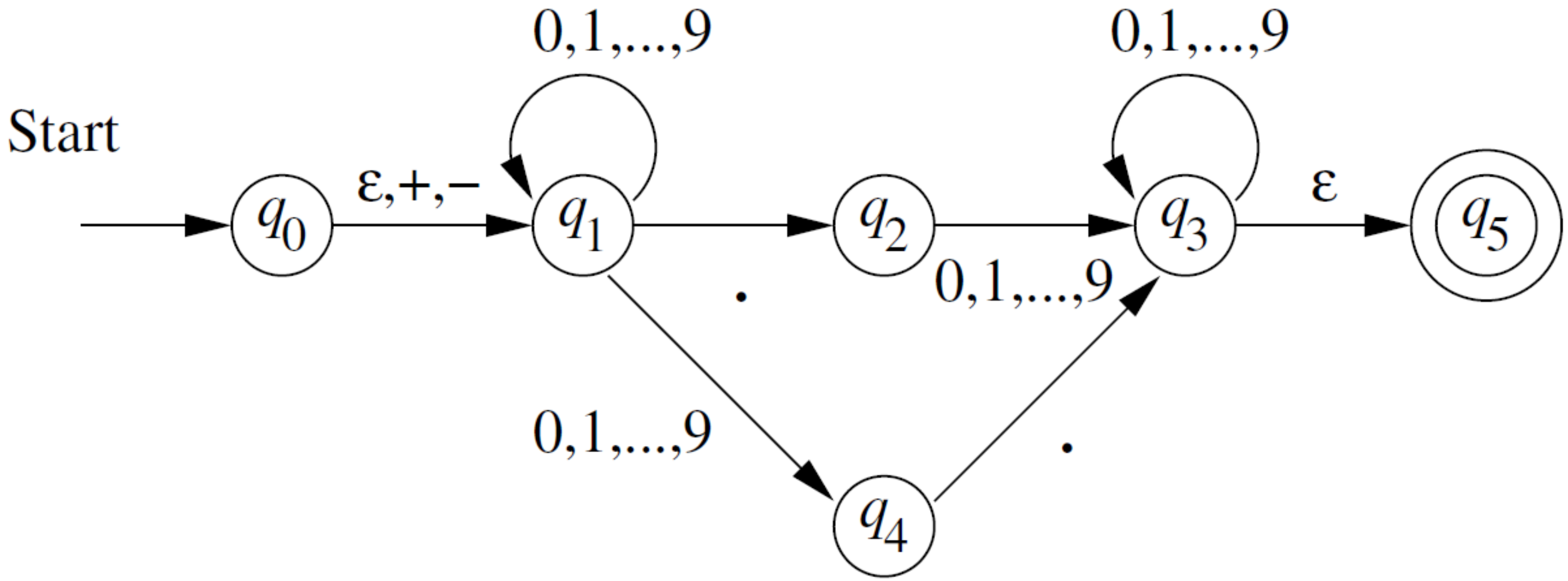
Original DFA accepting **odd** number of **b**



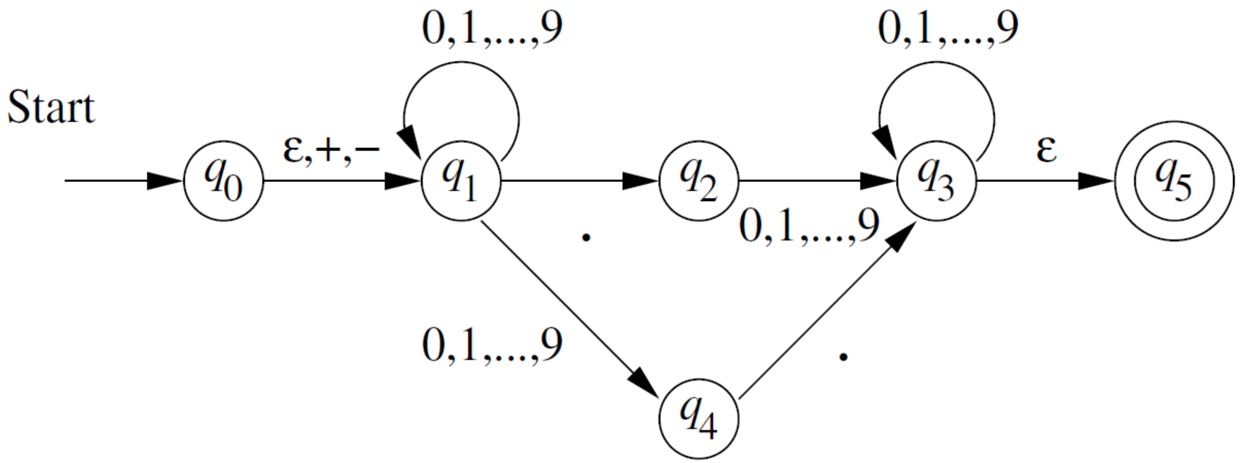
Complemented DFA accepting **even** number of **b**

NFA with ϵ Transition

- NFA where ϵ-transitions are allowed
- Only difference with NFA: $\delta: Q \times (\Sigma \cup \epsilon) \rightarrow P(Q)$
- Example: ϵ-NFA that accepts decimal numbers



NFA with ϵ Transition



Transition Diagram

	ϵ	$+,-$	$.$	$0,1,\ldots,9$
q_0	$\{q_1\}$	$\{q_1\}$	$\emptyset$	$\emptyset$
q_1	$\emptyset$	$\emptyset$	$\{q_2\}$	$\{q_1, q_4\}$
q_2	$\emptyset$	$\emptyset$	$\emptyset$	$\{q_3\}$
q_3	$\{q_5\}$	$\emptyset$	$\emptyset$	$\{q_3\}$
q_4	$\emptyset$	$\emptyset$	$\{q_3\}$	$\emptyset$
q_5	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$

Transition Table

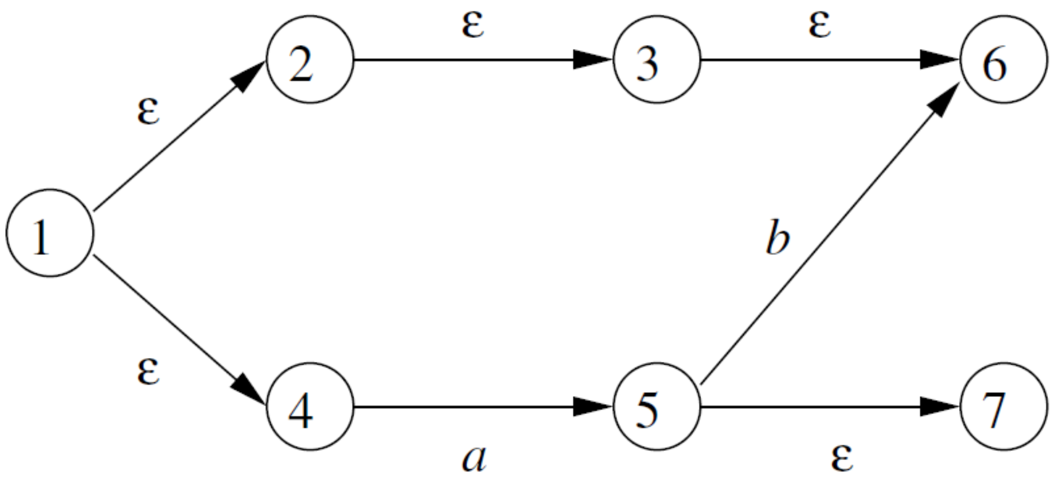
ϵ-closure

Find ϵ-close of state q by following all transitions out of q with label ϵ

$\underline{ECLOSE(q):}$
$E \ = \{q\}$

1. For all p in $\delta(q, \epsilon) - q$
   i. $E \leftarrow E \cup ECLOSE(p)$

2. Return E



ECLOSE(1) = {1,2,3,6,4}

Transition Function of ϵ-NFA

- Consider $w = xa$ where a is the last symbol and x is a rest of w.
- Assume $\hat{\delta}(q, x) = \{p_1, p_2, \ldots p_k\}$

If

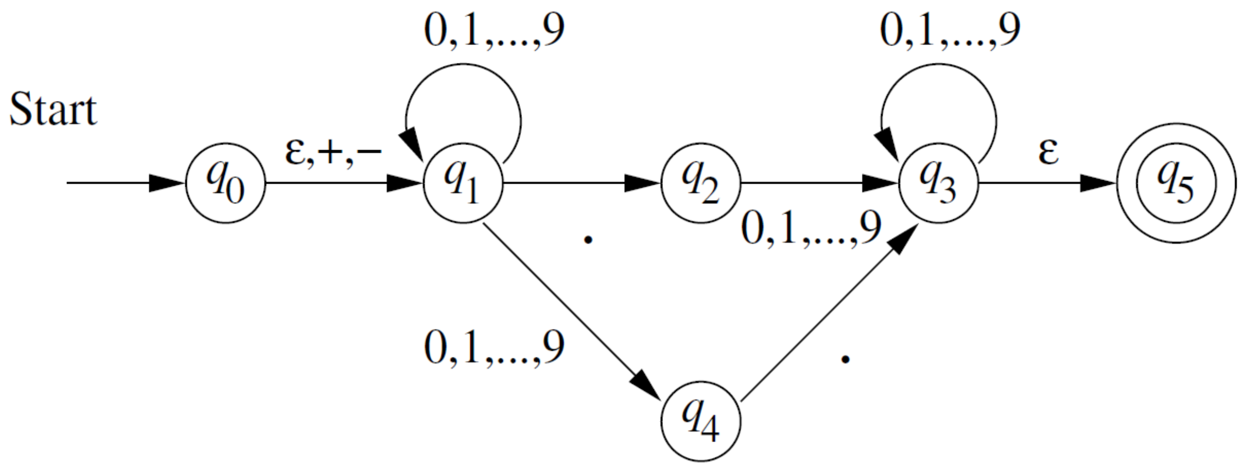
$$\bigcup_{i=1}^{k} \delta(p_i, a) = \{r_1, r_2, \ldots r_m\}$$

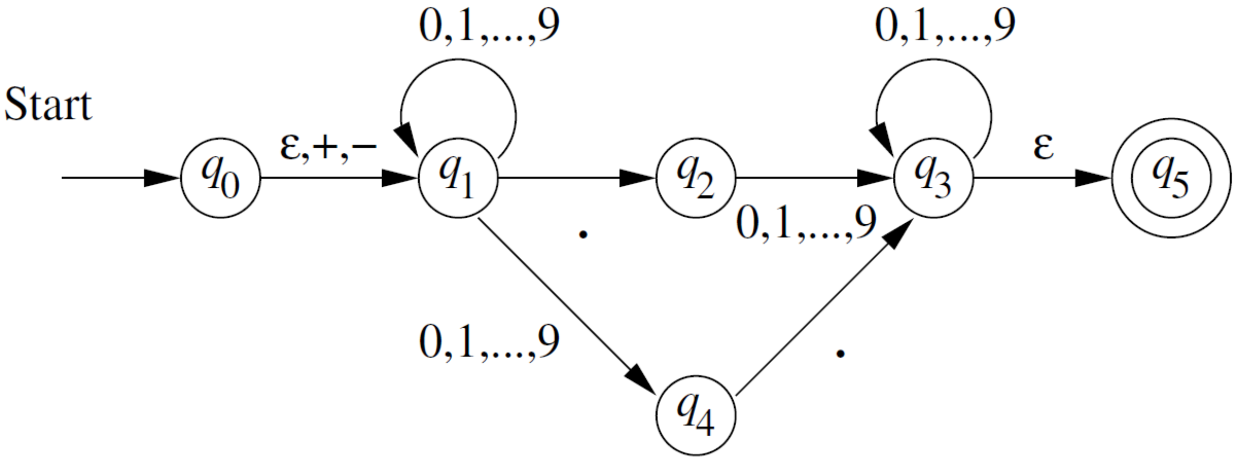
then

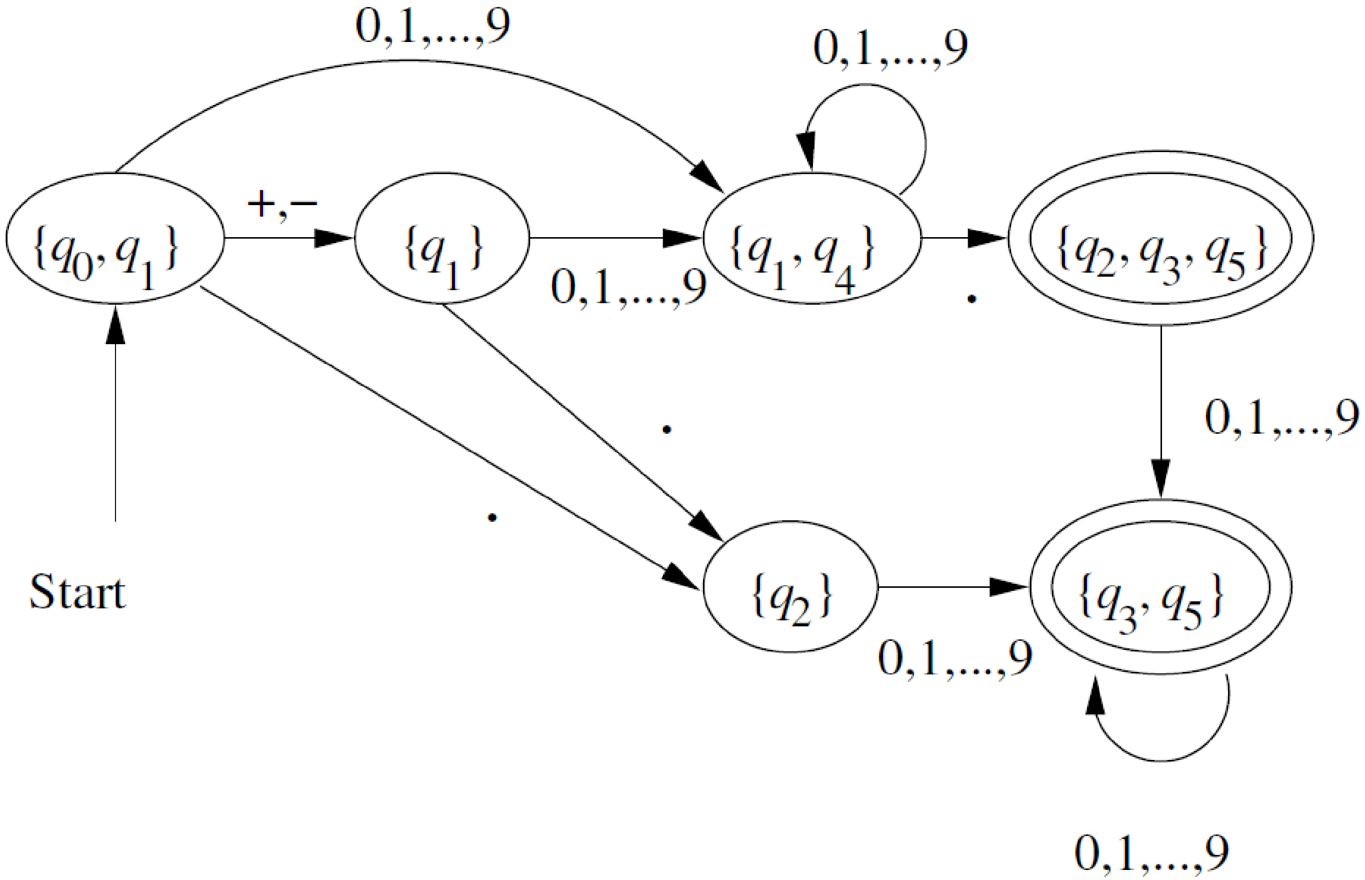
$$\hat{\delta}(q, w) = ECLOSE\{r_1, r_2, \ldots r_m\}$$

Eliminating ϵ-Transition

- Consider ϵ-NFA $\boldsymbol{E = (Q_E, \Sigma, \delta_E, q_0, F_E)}$
- Consider DFA $\boldsymbol{D = (Q_D, \Sigma, \delta_D, q_D, F_D)}$
- Conversion:
  - $Q_D = \{S | S \subseteq Q_E and\ \epsilon - closed\}$
  - Σ remains same
  - $\delta_D(S, a)$ is computed as follows:
    - Assume $\mathrm{S} = \{p_1, p_2, \ldots p_k\}$
    - $\bigcup_{i=1}^{k} \delta_E(p_i, a) = \{r_1, r_2, \ldots r_m\}$
    - $\delta_D(S, a) = ECLOSE(\{r_1, r_2, \ldots r_m\})$
  - $q_D = ECLOSE(q_0)$
  - $F_D = \{S | S \subseteq Q_D\ and\ S \cap F_E \neq \emptyset\}$




Start
ε,+,−
0,1,...,9
.
0,1,...,9
0,1,...,9
0,1,...,9
.
ε
q_0
q_1
q_2
q_3
q_4
q_5

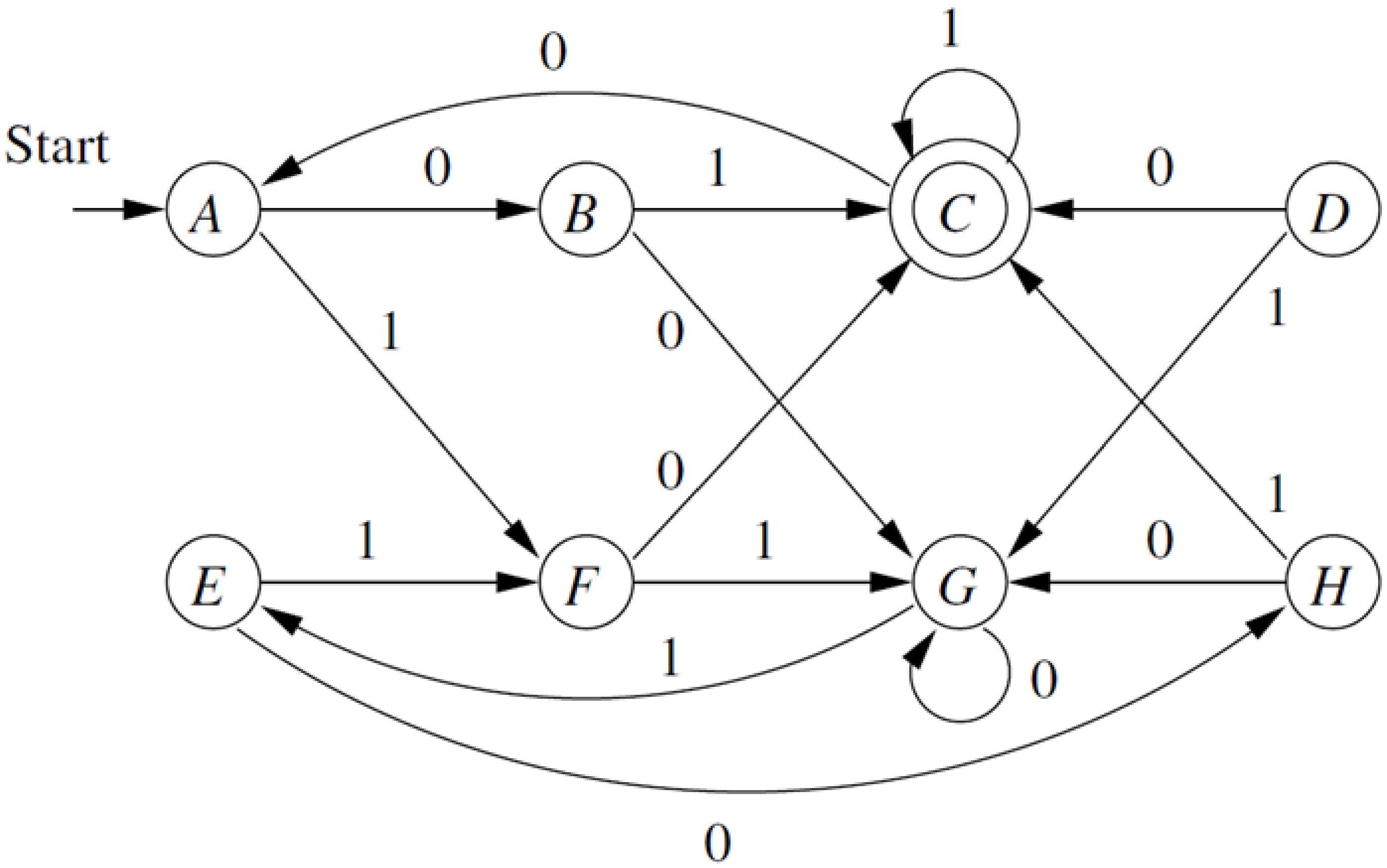

0,1,...,9
0,1,...,9
+,−
0,1,...,9
.
0,1,...,9
.
.
0,1,...,9
0,1,...,9
Start
$\{q_0, q_1\}$
$\{q_1\}$
$\{q_1, q_4\}$
$\{q_2, q_3, q_5\}$
$\{q_2\}$
$\{q_3, q_5\}$

Distinguishable States of DFA

- If p is an accepting state and q is nonaccepting, then the pair $\{p, q\}$ is distinguishable.

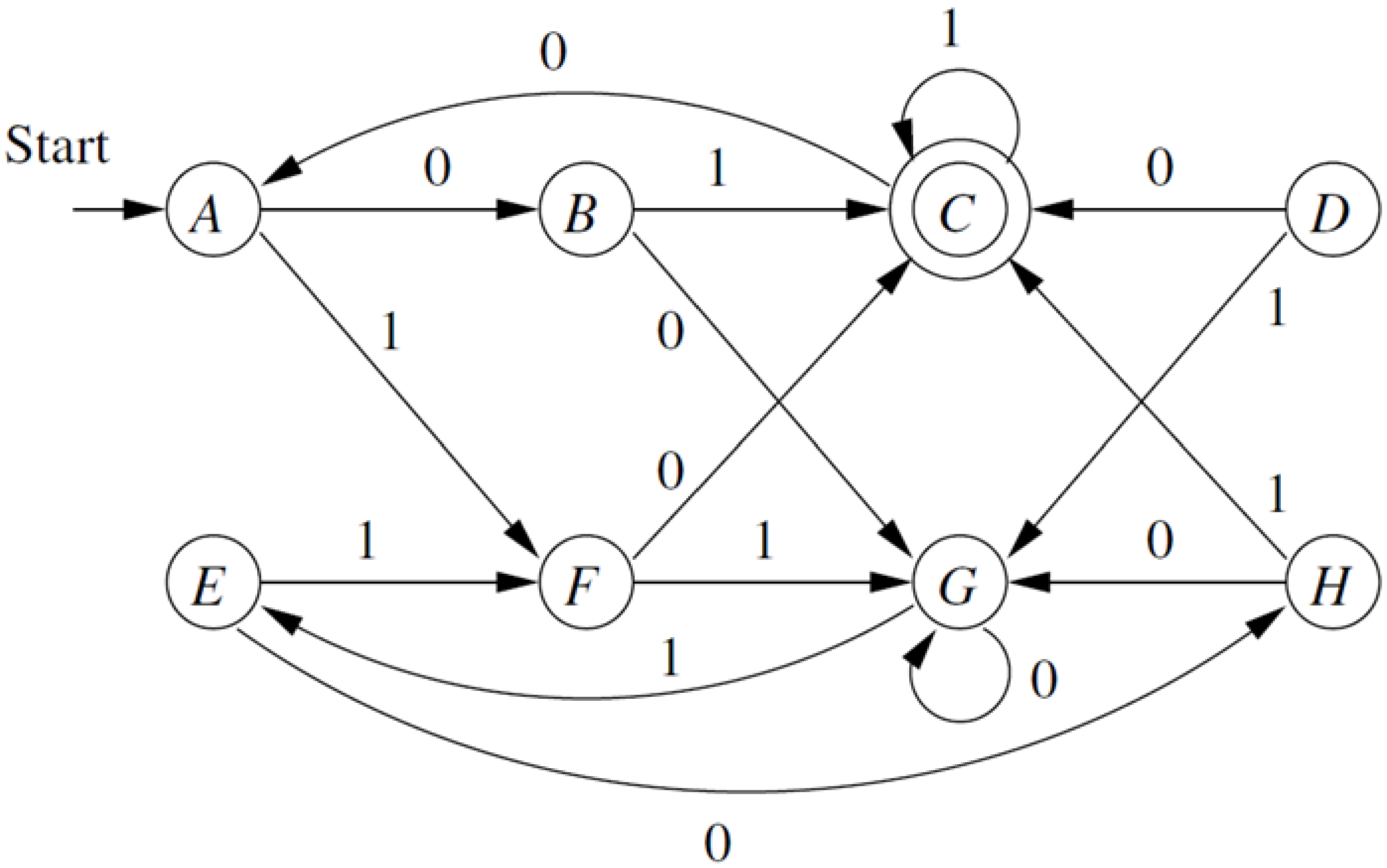
- Let p and q be states such that for some input symbol a, $r = \delta(p, a)$ and $s = \delta(q, a)$ are a pair of states known to be distinguishable. Then, $\{p, q\}$ is distinguishable.

Table Filling Algorithm



	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x	x	x	
H	x		x	x	x	x	x

DFA Minimization



- After partitioning the states into equivalent blocks we get:

  ({A, E}, {B, H}, {C}, {D, F}, {G})

DFA Minimization