Big picture

- •All languages
- Decidable

Turing machines

- •NP
- •P
- Context-free

Context-free grammars, push-down automata

•Regular

Automata, non-deterministic automata, regular expressions





- States), this DFA has 4 states
- Transitions

labelled with elements of the alphabet $\Sigma = \{0, 1\}$

Computation on input w:

- Begin in start state
- Read input string in a one-way fashion
- Follow the arrows matching input symbols
- When input ends: ACCEPT if in accept state

REJECT if not



Example: Input string



Example: Input string



Example: Input string



Example: Input string



Example: Input string



Example: Input string



Example: Input string

w = 0011 ACCEPT because end in accept state



Example: Input string



Example: Input string



Example: Input string



Example: Input string



Example: Input string



Example: Input string

w = 010 **REJECT**

because does not

end in accept state



Example: Input string w = 01 ACCEPT

- w = 010 REJECT
- w = 0011 ACCEPT
- w = 00110 REJECT



M recognizes language

L(M) = { w : w starts with 0 and ends with 1 }

L(M) is the language of strings causing M to accept

Example: 0101 is an element of L(M), 0101 $\in L(M)$



- 00 causes M to accept, so 00 is in $L(M) = 00 \in L(M)$
- 01 does not cause M to accept, so 01 not in L(M),

01 ∉ L(M)

- 0101 $\in L(M)$
- 01101100 $\in L(M)$
- 011010 ∉ L(M)



L(M) = {w : w has an even number of 1 }

Note: If there is no 1, then there are zero 1, zero is an even number, so M should accept.

Indeed 0000000 $\in L(M)$



L(M) = every possible string over {0,1}





• L(M) = ?



• L(M) = all strings over {0,1} except empty string ε = {0,1}* - { ε }



• L(M) = ?



- L(M) = { w : w starts and ends with same symbol }
- Memory is encoded in ... what ?



- L(M) = { w : w starts and ends with same symbol }
- Memory is encoded in states.

DFA have finite states, so finite memory

Convention:



L(M) = { w : w starts with 0 and ends with 1 }



Convention:



Don't need to write such arrows:

If, from some state, read symbol with no

corresponding arrow, imagine M goes into "sink state" that is not shown, and REJECT.

This makes pictures more compact.

Another convention:

List multiple transition on same arrow:





This makes pictures more compact.

Example $\sum = \{0,1\}$





$$L(M) = ?$$

Example
$$\sum = \{0,1\}$$

$$M =$$

$$\rightarrow O \xrightarrow{0,1} O \xrightarrow{0,1} O$$

$$L(M) = \sum^{2} = \{00, 01, 10, 11\}$$

Example from programming languages:

Recognize strings representing numbers:



Note: 0,...,9 means 0,1,2,3,4,5,6,7,8,9: 10 transitions

Example from programming languages:

Recognize strings representing numbers:





- Follow with arbitrarily many digits, but at least one
- Possibly put decimal point
- Follow with arbitrarily many digits, possibly none

Example from programming languages:

Recognize strings representing numbers:





- Input w = + REJECT
- Input w = -3.25 ACCEPT
- Input w = +2.35-. REJECT
Example $\Sigma = \{0, 1\}$

What about { w : w has same number of 0 and 1 }

• Can you design a DFA that recognizes that?

• It seems you need infinite memory

• We will prove later that there is no DFA that recognizes that language !

Next: formal definition of DFA

• Useful to prove various properties of DFA

 Especially important to prove that things CANNOT be recognized by DFA.

Useful to practice mathematical notation

State diagram of a DFA:





- •Some number of accept states ()
- •Labelled transitions exiting each state, _____ for every symbol in Σ

•Definition: A finite automaton (DFA) is a 5-tuple (Q, Σ , δ , q₀, F) where

- •Q is a finite set of states
- • Σ is the input alphabet
- • δ : Q X $\Sigma \rightarrow$ Q is the transition function
- $\bullet q_0$ in Q is the start state
- ${}^{\bullet}F \subseteq Q$ is the set of accept states

Q X Σ is the set of ordered pairs (a,b) : a \in Q, b \in Σ Example {q,r,s}X{0,1}={(q,0),(q,1),(r,0),(r,1),(s,0),(s,1)}



- •Q = { q_0, q_1 }
- $\bullet \Sigma = \{0,1\}$
- • $\delta(q_0, 0) = ?$



- •Q = { q_0, q_1 }
- • $\Sigma = \{0, 1\}$
- • $\delta(q_0, 0) = q_0 \quad \delta(q_0, 1) = ?$



- •Q = { q_0, q_1 }
- •∑ = {0,1}
- • $\delta(q_0, 0) = q_0 \quad \delta(q_0, 1) = q_1$
 - $\delta(q_1, 0) = q_1 \quad \delta(q_1, 1) = q_0$
- $\bullet q_0$ in Q is the start state

•F = ?



- •Q = { q_0, q_1 }
- $\bullet \Sigma = \{0,1\}$
- • $\delta(q_0, 0) = q_0 \quad \delta(q_0, 1) = q_1$
 - $\delta(q_1, 0) = q_1 \quad \delta(q_1, 1) = q_0$
- $\bullet q_0$ in Q is the start state
- •F = { q_0 } \subseteq Q is the set of accept states

•Definition: A DFA (Q, Σ , δ , q₀, F) accepts a string w if

•w = w₁ w₂ ... w_k where, $\forall 1 \le i \le k$, w_i is in Σ (the k symbols of w)

The sequence of k+1 states r₀, r₁, ..., r_k such that:
(1) r₀ = q₀, and
(2) r_{i+1} = δ(r_i, w_{i+1}) ∀ 0 ≤ i < k
has r_k in F
(r_i = state DFA is in after reading i-th symbol in w)



•Above DFA (Q, Σ , δ , q₀, F) accepts w = 011





We must show that

•The sequence of 3+1=4 states r_0 , r_1 , r_2 , r_3 such that:

(1)
$$r_0 = q_0$$

(2) $r_{i+1} = \delta(r_i, w_{i+1}) \forall 0 \le i < 3$
has r_3 in F



• $w = 011 = w_1 w_2 w_3$ $w_1 = 0 w_2 = 1 w_3 = 1$

- $r_0 = q_0$
- r₁ := ?



• $w = 011 = w_1 w_2 w_3$ $w_1 = 0 w_2 = 1 w_3 = 1$

- $r_0 = q_0$
- $r_1 = \delta(r_0, w_1) = \delta(q_0, 0) = q_0$ • $r_2 := ?$



• $w = 011 = w_1 w_2 w_3$ $w_1 = 0 w_2 = 1 w_3 = 1$

- $r_0 = q_0$
- $r_1 = \delta(r_0, w_1) = \delta(q_0, 0) = q_0$
- $r_2 = \delta(r_1, w_2) = \delta(q_0, 1) = q_1$
- r₃ := ?



()K

()NFi

- $r_0 = q_0$
- $r_1 = \delta(r_0, w_1) = \delta(q_0, 0) = q_0$
- $r_2 = \delta(r_1, w_2) = \delta(q_0, 1) = q_1$
- $r_3 = \delta(r_2, w_3) = \delta(q_1, 1) = q_0$
- $r_3 = q_0$ in F

• **Definition**: For a DFA M, we denote by L(M) the set of strings accepted by M:

L(M) := { w : M accepts w}

We say M accepts or recognizes the language L(M)

• Definition: A language L is regular if ∃ DFA M : L(M) = L In the next lectures we want to:

• Understand power of regular languages

• Develop alternate, compact notation to specify regular languages

Example: Unix command *grep '*\<*c.*h*\>' file selects all words starting with c and ending with h in *file*

• Understand power of regular languages:

- Suppose A, B are regular languages, what about
- not A := { w : w is not in A }
- A U B := { w : w in A or w in B }
- A o B := { $w_1 w_2$: w_1 in A and w_2 in B }
- $\bullet \ A^* \ := \{ \ w_1 \ w_2 \ \ldots \ w_k \ : k \geq 0 \ , \ w_i \ in \ A \ \ for \ every \ i \ \}$

• Are these languages regular?

• Understand power of regular languages:

- Suppose A, B are regular languages, what about
- not A := { w : w is not in A }
- A U B := { w : w in A or w in B }
- A o B := { $w_1 w_2$: w_1 in A and w_2 in B }
- $\bullet \ A^* \ := \{ \ w_1 \ w_2 \ \ldots \ w_k \ : k \geq 0 \ , \ w_i \ in \ A \ \ for \ every \ i \ \}$

 Terminology: Are regular languages closed under not, U, o, * ?

If A is a regular language, then so is (not A)

If A is a regular language, then so is (not A)

If A is a regular language, then so is (not A)

Proof idea: Complement the set of accept states
Example

If A is a regular language, then so is (not A)

Proof idea: Complement the set of accept states
Example:



L(M) =

{ w : w has even number of 1}

If A is a regular language, then so is (not A)

Proof idea: Complement the set of accept statesExample:



- Theorem: If A is a regular language, then so is (not A)
 Proof:
 - Given DFA M = (Q, Σ , δ , q₀, F) such that L(M) = A.

This definition is the creative step of this proof, the rest is (perhaps complicated but) mechanical "unwrapping definitions"

- Theorem: If A is a regular language, then so is (not A)
 Proof:
 - Given DFA M = (Q, Σ , δ , q₀, F) such that L(M) = A.
 - **Define** DFA M' = (Q, Σ , δ , q₀, F'), where F' := not F.
- •We need to show L(M') = not L(M), that is:

- Theorem: If A is a regular language, then so is (not A)
 Proof:
 - Given DFA M = (Q, Σ , δ , q₀, F) such that L(M) = A.

Define DFA M' = (Q, Σ , δ , q₀, F'), where F' := not F.

- •We need to show L(M') = not L(M), that is:
- for any w, M' accepts w $\leftarrow \rightarrow$ M does not accept w.
- •So let w be any string of length k, and consider the k+1 states r_0 , r_1 , ..., r_k from the definition of accept:

(1)
$$r_0 = q_0$$
, and

(2)
$$r_{i+1} = \delta(r_i, w_{i+1}) \forall 0 \le i < k$$
.

How do we conclude?

- Theorem: If A is a regular language, then so is (not A)
 Proof:
 - Given DFA M = (Q, Σ , δ , q₀, F) such that L(M) = A.

Define DFA M' = (Q, Σ , δ , q₀, F'), where F' := not F.

- •We need to show L(M') = not L(M), that is:
- for any w, M' accepts w $\leftarrow \rightarrow$ M does not accept w
- •So let w be any string of length k, and consider the k+1 states r_0 , r_1 , ..., r_k from the definition of accept:

(1)
$$r_0 = q_0$$
, and

(2)
$$r_{i+1} = \delta(r_i, w_{i+1}) \forall 0 \le i < k$$
.

Note that r_k in F' $\leftarrow \rightarrow r_k$ not in F, since F' = not F.

What is a proof?

•A proof is an explanation, written in English, of why something is true.

•Every sentence must be logically connected to the previous ones, often by "so", "hence", "since", etc.

•Your audience is a human being, NOT a machine.

What is a proof?

Complement the set of accept states

To know a proof means to know all the pyramid

$$L(M) = \sum^2 = \{00, 01, 10, 11\}$$

What is a DFA M' : L(M') = not \sum^2 = all strings except those of length 2 ? Example $\sum = \{0,1\}$

$$\longrightarrow \bigcirc 0,1 \bigcirc$$

$$L(M') = not \sum^2 = \{0,1\}^* - \{00,01,10,11\}$$

Do not forget the convention about the sink state!

- Suppose A, B are regular languages, what about
- not A := { w : w is not in A } REGULAR
- A U B := { w : w in A or w in B }
- A o B := { $w_1 w_2 : w_1 \text{ in } A \text{ and } w_2 \text{ in } B$ }
- $\bullet \ A^* \ := \{ \ w_1 \ w_2 \ \ldots \ w_k \ : k \geq 0 \ , \ w_i \ in \ A \ \ for \ every \ i \ \}$

•Theorem: If A, B are regular, then so is A U B

Proof idea: Take Cartesian product of states
 In a pair (q,q'),
 q tracks DFA for A,
 q' tracks DFA for B.

•Next we see an example. In it we abbreviate 1with 1






- •Theorem: If A, B are regular, then so is A U B •Proof:
 - Given DFA M_A = (Q_A, Σ , δ_A ,q_A, F_A) such that L(M) = A, DFA M_B = (Q_B, Σ , δ_B ,q_B, F_B) such that L(M) = B. Define DFA M = (Q, Σ , δ , q₀, F), where

Q := ?

- •Theorem: If A, B are regular, then so is A U B •Proof:
 - Given DFA M_A = (Q_A, Σ , δ_A , q_A, F_A) such that L(M) = A, DFA M_B = (Q_B, Σ , δ_B , q_B, F_B) such that L(M) = B. Define DFA M = (Q, Σ , δ , q₀, F), where Q := Q_A X Q_B q₀ := ?

- •Theorem: If A, B are regular, then so is A U B •Proof:
 - Given DFA M_A = $(Q_A, \Sigma, \delta_A, q_A, F_A)$ such that L(M) = A, DFA M_B = $(Q_B, \Sigma, \delta_B, q_B, F_B)$ such that L(M) = B. Define DFA M = $(Q, \Sigma, \delta, q_0, F)$, where $Q := Q_A X Q_B$ $q_0 := (q_A, q_B)$ F := ?

- •Theorem: If A, B are regular, then so is A U B •Proof:
 - Given DFA $M_A = (Q_A, \Sigma, \delta_A, q_A, F_A)$ such that L(M) = A, DFA $M_B = (Q_B, \Sigma, \delta_B, q_B, F_B)$ such that L(M) = B. Define DFA M = (Q, Σ, δ, q_0, F), where $Q := Q_A X Q_B$ $q_0 := (q_A, q_B)$
 - $\mathsf{F} := \{(\mathsf{q},\mathsf{q}') \in \mathsf{Q} : \mathsf{q} \in \mathsf{F}_{\mathsf{A}} \text{ or } \mathsf{q}' \in \mathsf{F}_{\mathsf{B}} \}$

δ((q,q'), v) := (?, ?)

- •Theorem: If A, B are regular, then so is A U B •Proof:
 - Given DFA M_A = (Q_A, Σ , δ_A ,q_A, F_A) such that L(M) = A, DFA M_B = (Q_B, Σ , δ_B ,q_B, F_B) such that L(M) = B. Define DFA M = (Q, Σ , δ , q₀, F), where
 - $$\begin{split} & \text{Q} := \text{Q}_{\text{A}} \times \text{Q}_{\text{B}} \\ & \text{q}_{0} := (\text{q}_{\text{A}}, \text{q}_{\text{B}}) \\ & \text{F} := \{(\text{q}, \text{q}') \in \text{Q} : \text{q} \in \text{F}_{\text{A}} \text{ or } \text{q}' \in \text{F}_{\text{B}} \} \\ & \delta(\ (\text{q}, \text{q}'), \ \text{v}) := (\delta_{\text{A}} \ (\text{q}, \text{v}), \ \delta_{\text{B}} \ (\text{q}', \text{v}) \) \end{split}$$
 - We need to show L(M) = A U B that is, for any w:
 M accepts w ←→ M_A accepts w or M_B accepts w

•Proof of M accepts w \rightarrow M_A accepts w or M_B accepts w

- •Suppose that M accepts w of length k.
- •From the definitions of accept and M, the sequence

$$\begin{aligned} &(s_0 , t_0) = q_0 = (q_A , q_B), \\ &(s_{i+1} , t_{i+1}) = \delta((s_i , t_i) , w_{i+1}) = (\ \delta_A(s_i , w_{i+1}) , \ \delta_B(t_i , w_{i+1}) \ \forall 0 \leq i < k \\ &has \ (s_k , t_k) \in \ \end{aligned}$$

- •Proof of M accepts $w \rightarrow M_A$ accepts w or M_B accepts w
- •Suppose that M accepts w of length k.
- •From the definitions of accept and M, the sequence

$$(s_0, t_0) = q_0 = (q_A, q_B),$$

 $(s_{i+1}, t_{i+1}) = \delta((s_i, t_i), w_{i+1}) = (\delta_A(s_i, w_{i+1}), \delta_B(t_i, w_{i+1}) \forall 0 \le i < k$
has $(s_k, t_k) \in F.$

•By our definition of F, what can we say about (s_k, t_k) ?

- •Proof of M accepts $w \rightarrow M_A$ accepts w or M_B accepts w
- •Suppose that M accepts w of length k.
- •From the definitions of accept and M, the sequence

$$(s_0, t_0) = q_0 = (q_A, q_B),$$

 $(s_{i+1}, t_{i+1}) = \delta((s_i, t_i), w_{i+1}) = (\delta_A(s_i, w_{i+1}), \delta_B(t_i, w_{i+1}) \forall 0 \le i < k$
has $(s_k, t_k) \in F.$

- •By our definition of F, $s_k \in F_A$ or $t_k \in F_B$.
- •Without loss of generality, assume $s_k \in F_A$.
- •Then M_A accepts w because the sequence

$$s_0 = q_A$$
, $s_{i+1} = \delta_A (s_i, w_{i+1})$ ∀0 ≤ i < k,
has $s_k \in F_A$.

•Proof of M accepts w $\bigstar M_A$ accepts w or M_B accepts w •W/out loss of generality, assume M_A accepts w, |w|=k. •From the definition of M_A accepts w, the sequence $r_0 := q_A, r_{i+1} := \delta_A (r_i, w_{i+1}) \forall 0 \le i < k$, has r_k in ? •Proof of M accepts w $\bigstar M_A$ accepts w or M_B accepts w •W/out loss of generality, assume M_A accepts w, |w|=k. •From the definition of M_A accepts w, the sequence $r_0 := q_A, r_{i+1} := \delta_A (r_i, w_{i+1}) \forall 0 \le i < k$, has r_k in F_A .

- •Define the sequence of k+1 states $t_0 := q_B$, $t_{i+1} := \delta_B (t_i, w_{i+1}) \forall 0 \le i < k$.
- •M accepts w because the sequence ???????????? (recall states in M are pairs)

•Proof of M accepts w $\bigstar M_A$ accepts w or M_B accepts w •W/out loss of generality, assume M_A accepts w, |w|=k. •From the definition of M_A accepts w, the sequence $r_0 := q_A, r_{i+1} := \delta_A (r_i, w_{i+1}) \forall 0 \le i < k$, has r_k in F_A .

•Define the sequence of k+1 states $t_0 := q_B$, $t_{i+1} := \delta_B (t_i, w_{i+1}) \forall 0 \le i < k$.

•M accepts w because the sequence $(r_0, t_0) = q = (q_A, q_B),$ $(r_{i+1}, t_{i+1}) = \delta((r_i, t_i), w_{i+1}) = (\delta_A(r_i, w_{i+1}), \delta_B(t_i, w_{i+1}) \forall 0 \le i < k$ has (r_k, t_k) in F, by our definition of F.

- Suppose A, B are regular languages, what about
- not A := { w : w is not in A } REGULAR
- A U B := { w : w in A or w in B } REGULAR
- A o B := { $w_1 w_2 : w_1$ in A and w_2 in B }
- $\bullet \ A^* \ := \{ \ w_1 \ w_2 \ \ldots \ w_k \ : k \geq 0 \ , \ w_i \ in \ A \ \ for \ every \ i \ \}$

• Other two are more complicated!

 Plan: we introduce NFA prove that NFA are equivalent to DFA reprove A U B, prove A o B, A* regular, using NFA Non deterministic finite automata (NFA)

 DFA: given state and input symbol, unique choice for next state, deterministic:

•Next we allow multiple choices, non-deterministic

We also allow ε-transitions:
 can follow without reading anything







Intuition of how it computes:

- Accept string w if there is a way to follow transitions that ends in accept state
- •Transitions labelled with symbol in $\Sigma = \{a, b\}$ must be matched with input
- • ϵ transitions can be followed without matching



Example:

- Accept a (first follow ε -transition)
- Accept baaa

ANOTHER Example of NFA



Example:

- Accept bab (two accepting paths, one uses the ϵ -transition)
- Reject ba (two possible paths, but neither has final state = q₁)

 Definition: A non-deterministic finite automaton (NFA) is a 5-tuple (Q, Σ, δ, q₀, F) where

- •Q is a finite set of states
- • Σ is the input alphabet
- • δ : Q X (Σ U { ϵ }) \rightarrow Powerset(Q)
- $\bullet q_0$ in Q is the start state
- •F \subseteq Q is the set of accept states

•Recall: Powerset(Q) = set of all subsets of Q Example: Powerset({1,2}) = ? •Definition: A non-deterministic finite automaton (NFA) is a 5-tuple (Q, Σ , δ , q₀, F) where

- •Q is a finite set of states
- • Σ is the input alphabet
- • δ : Q X (Σ U { ϵ }) \rightarrow Powerset(Q)
- $\bullet q_0$ in Q is the start state
- ${}^{\bullet}\mathsf{F} \subseteq \mathsf{Q}$ is the set of accept states

•Recall: Powerset(Q) = set of all subsets of Q Example: Powerset($\{1,2\}$) = { \emptyset , {1}, {2}, {1,2} }



•Example: above NFA is 5-tuple (Q, Σ , δ , q_0 , F) where •Q = { q_0 , q_1 } • Σ = {0,1}

•δ(q₀,0) = ?



• $\Sigma = \{0, 1\}$ • $\delta(q_0, 0) = \{q_0\} \quad \delta(q_0, 1) = ?$



• $\delta(q_0, 0) = \{q_0\} \quad \delta(q_0, 1) = \{q_0, q_1\} \quad \delta(q_0, \varepsilon) = ?$



•Example: above NFA is 5-tuple (Q, Σ , δ , q_0 , F) where •Q = { q_0 , q_1 } • Σ = {0,1} • $\delta(q_0, 0) = \{q_0\} \quad \delta(q_0, 1) = \{q_0, q_1\} \quad \delta(q_0, \epsilon) = \emptyset$ $\delta(q_1, 0) = ?$



- • $\Sigma = \{0, 1\}$
- $\begin{aligned} \bullet \delta(q_0, 0) &= \{q_0\} \quad \delta(q_0, 1) = \{q_0, q_1\} \quad \delta(q_0, \varepsilon) = \emptyset \\ \delta(q_1, 0) &= \emptyset \quad \delta(q_1, 1) = ? \end{aligned}$



$$\begin{split} \bullet \Sigma &= \{0,1\} \\ \bullet \delta(q_0,0) &= \{q_0\} \quad \delta(q_0,1) = \{q_0,q_1\} \quad \delta(q_0,\epsilon) = \emptyset \\ \delta(q_1,0) &= \emptyset \quad \delta(q_1,1) = \emptyset \quad \delta(q_1,\epsilon) = ? \end{split}$$



- $\bullet \Sigma = \{0,1\}$
- $\begin{aligned} \bullet \delta(q_0, 0) &= \{q_0\} \quad \delta(q_0, 1) = \{q_0, q_1\} \quad \delta(q_0, \varepsilon) = \emptyset \\ \delta(q_1, 0) &= \emptyset \quad \delta(q_1, 1) = \emptyset \quad \delta(q_1, \varepsilon) = \{q_0\} \end{aligned}$

 $\bullet q_0$ in Q is the start state

•F = ?



- $\bullet \Sigma = \{0,1\}$
- • $\delta(q_0, 0) = \{q_0\} \quad \delta(q_0, 1) = \{q_0, q_1\} \quad \delta(q_0, \varepsilon) = \emptyset$
- $\delta(q_1, 0) = \emptyset \qquad \delta(q_1, 1) = \emptyset \qquad \delta(q_1, \varepsilon) = \{q_0\}$
- $\bullet q_0$ in Q is the start state
- $\bullet F = \{ \ q_1 \} \subseteq Q \ is \ the \ set \ of \ accept \ states$

•Definition: A NFA (Q, Σ , δ , q_0 , F) accepts a string w if \exists integer k, \exists k strings w_1 , w_2 , ..., w_k such that •w = $w_1 w_2 ... w_k$ where $\forall 1 \le i \le k$, $w_i \in \Sigma \cup \{\epsilon\}$ (the symbols of w, or ϵ)

•∃ sequence of k+1 states r_0 , r_1 , ..., r_k in Q such that:

• $r_0 = q_0$

•
$$r_{i+1} \in \delta(r_i, w_{i+1}) \ \forall \ 0 \le i < k$$

• r_k is in F

•Differences with DFA are in green



 $r_0 = ?$



$$w_1 = b$$
, $w_2 = a$, $w_3 = a$, $w_4 = \varepsilon$, $w_5 = a$

$$r_0 = q_0, r_1 = ?$$



$$w_1 = b$$
, $w_2 = a$, $w_3 = a$, $w_4 = \varepsilon$, $w_5 = a$

$$r_0 = q_0, r_1 = q_1, r_2 = ?$$

$$\mathsf{r}_1 \in \delta(\mathsf{r}_{_0},\mathsf{b}) = \{\mathsf{q}_{_1}\}$$



$$w_1 = b$$
, $w_2 = a$, $w_3 = a$, $w_4 = \varepsilon$, $w_5 = a$

$$r_0 = q_0, r_1 = q_1, r_2 = q_2, r_3 = ?$$

$$r_1 \in \delta(r_0, b) = \{q_1\} \quad r_2 \in \delta(r_1, a) = \{q_1, q_2\}$$



$$w_1 = b$$
, $w_2 = a$, $w_3 = a$, $w_4 = \varepsilon$, $w_5 = a$

$$r_0 = q_0, r_1 = q_1, r_2 = q_2, r_3 = q_0, r_4 = ?$$

$$\begin{split} \mathbf{r}_1 &\in \delta(\mathbf{r}_0, \mathbf{b}) = \{\mathbf{q}_1\} \quad \mathbf{r}_2 \in \delta(\mathbf{r}_1, \mathbf{a}) = \{\mathbf{q}_1, \mathbf{q}_2\} \\ \mathbf{r}_3 &\in \delta(\mathbf{r}_2, \mathbf{a}) = \{\mathbf{q}_0\} \end{split}$$



$$w_1 = b$$
, $w_2 = a$, $w_3 = a$, $w_4 = \varepsilon$, $w_5 = a$

$$r_0 = q_0, r_1 = q_1, r_2 = q_2, r_3 = q_0, r_4 = q_2, r_5 = ?$$

$$\begin{split} \mathbf{r}_{1} &\in \delta(\mathbf{r}_{_{0}}, \mathbf{b}) = \{\mathbf{q}_{_{1}}\} \quad \mathbf{r}_{2} \in \delta(\mathbf{r}_{_{1}}, \mathbf{a}) = \{\mathbf{q}_{_{1}}, \mathbf{q}_{_{2}}\} \\ \mathbf{r}_{3} &\in \delta(\mathbf{r}_{_{2}}, \mathbf{a}) = \{\mathbf{q}_{_{0}}\} \quad \mathbf{r}_{4} \in \delta(\mathbf{r}_{_{3}}, \varepsilon) = \{\mathbf{q}_{_{2}}\} \end{split}$$


$$w_1 = b$$
, $w_2 = a$, $w_3 = a$, $w_4 = \varepsilon$, $w_5 = a$

Accepting sequence of 5+1 = 6 states:

$$\mathbf{r}_{0} = \mathbf{q}_{0}, \quad \mathbf{r}_{1} = \mathbf{q}_{1}, \quad \mathbf{r}_{2} = \mathbf{q}_{2}, \quad \mathbf{r}_{3} = \mathbf{q}_{0}, \quad \mathbf{r}_{4} = \mathbf{q}_{2}, \quad \mathbf{r}_{5} = \mathbf{q}_{0}$$

Transitions:

$$\begin{aligned} \mathbf{r}_{1} \in \delta(\mathbf{r}_{0}, \mathbf{b}) &= \{\mathbf{q}_{1}\} & \mathbf{r}_{2} \in \delta(\mathbf{r}_{1}, \mathbf{a}) &= \{\mathbf{q}_{1}, \mathbf{q}_{2}\} \\ \mathbf{r}_{3} \in \delta(\mathbf{r}_{2}, \mathbf{a}) &= \{\mathbf{q}_{0}\} & \mathbf{r}_{4} \in \delta(\mathbf{r}_{3}, \mathbf{\epsilon}) &= \{\mathbf{q}_{2}\} & \mathbf{r}_{5} \in \delta(\mathbf{r}_{4}, \mathbf{a}) &= \{\mathbf{q}_{0}\} \end{aligned}$$

•NFA are at least as powerful as DFA, because DFA are a special case of NFA

•Are NFA more powerful than DFA?

•Surprisingly, they are not:

•Theorem:

For every NFA N there is DFA M : L(M) = L(N)

- •Theorem:
- For every NFA N there is DFA M : L(M) = L(N)

- •Construction without ϵ transitions
- •Given NFA N (Q, Σ , δ , q, F)
- •Construct DFA M (Q', Σ , δ ', q', F') where:
- •Q' := Powerset(Q)
- •q' = {q}
- $\bullet F' = \{ S : S \in Q' \text{ and } S \text{ contains an element of } F \}$
- $\delta'(S, a) := U_{s \in S} \delta(s, a)$

= { t : t $\in \delta$ (s,a) for some s \in S }

•It remains to deal with ϵ transitions

 Definition: Let S be a set of states.
 E(S) := { q : q can be reached from some state s in S traveling along 0 or more ε transitions }

•We think of following ϵ transitions at beginning, or right after reading an input symbol in Σ

- •Theorem:
- For every NFA N there is DFA M : L(M) = L(N)

- •Construction including ϵ transitions
- •Given NFA N (Q, Σ , δ , q, F)
- •Construct DFA M (Q', Σ , δ ', q', F') where:
- •Q' := Powerset(Q)
- •q' = E({q})
- $\bullet F' = \{ \ S : S \in Q' \ and \ S \ contains \ an \ element \ of \ F \}$
- δ'(S, a) := E(U_{s ∈ S} δ(s,a))
 = { t : t ∈ E(δ (s,a)) for some s ∈ S }







































We can delete the unreachable states.






































NFA

We can delete the unreachable states.

