01204211 Discrete Mathematics
Lecture 10a: Nondeterministic automata[1]

Jittat Fakcharoenphol

September 17, 2024

[1] Based on lecture notes of *Models of Computation* course by Jeff Erickson.

Review: DFA (Formal definitions)

A **finite-state machine** or a **deterministic finite-state automaton** (DFA) has five components:

- the input alphabet Σ,
- a finite set of states Q,
- a transition function $\delta : Q \times \Sigma \longrightarrow Q$
- a start state $s \in Q$, and
- a subset $A \subseteq Q$ of accepting states.

Review: Acceptance

One step move: from state q with input symbol a, the machine changes its state to $\delta(q, a)$.

Extension: from state q with input string w, the machine changes its state to $\delta^*(q, w)$ defined as

$$\delta^*(q, w) = \begin{cases} q & \text{if } w = \varepsilon, \\ \delta^*(\delta(q, a), x) & \text{if } w = ax. \end{cases}$$

The signature of δ^* is $Q \times \Sigma^* \longrightarrow Q$.

accepting w

For a finite-state machine with starting state s and accepting states A, it accepts string w iff

$$\delta^*(s, w) \in A.$$

Language of a DFA

$L(M)$

For a DFA M, let $L(M)$ be the set of all strings that M accepts. More formally, for $M = (\Sigma, Q, \delta, s, A)$,

$$L(M) = \{w \in \Sigma^* \mid \delta^*(s, w) \in A\}.$$

We refer to $L(M)$ as the language of M.

Acceptance

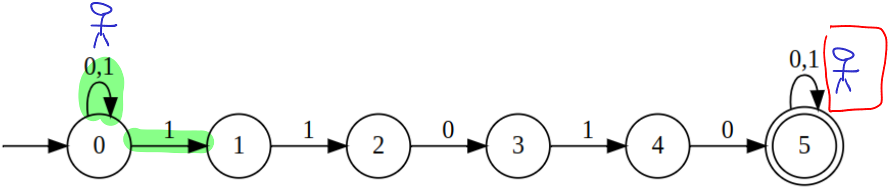
We also says M **accepts** $L(M)$.

New example 1

11010

111010

110111010011101111101 01...

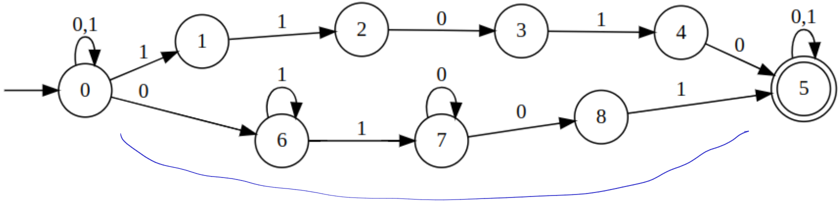


$(0+1)^* 11010 (0+1)^*$

New example 2

nondeterministic

$$(0+1)^*(11010 + 01^*10^*01)(0+1)^*$$



What's going on here?

More relaxed transitions

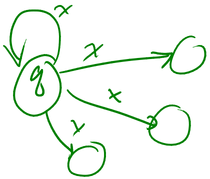
From state $q \in Q$, for input a, the machine can "possibly" change its state to many states.

More relaxed transitions

From state $q \in Q$, for input a, the machine can "possibly" change its state to many states.
New transition function δ :

More relaxed transitions

$\delta : Q \times \Sigma \rightarrow \boxed{Q}$ dfa



From state $q \in Q$, for input a, the machine can "possibly" change its state to many states.

New transition function $\delta : Q \times \Sigma \longrightarrow \boxed{2^Q}$.

We refer to this new kind of automaton as a **nondeterministic finite-state automaton** or **NFA**.

powerset set ของทุก subset ของ Q.

$P(Q)$

NFA (Formal definitions)

A **nondeterministic finite-state automaton** (NFA) has five components:

- the input alphabet Σ,

NFA (Formal definitions)

A **nondeterministic finite-state automaton** (NFA) has five components:

- the input alphabet Σ,
- a finite set of states Q,

NFA (Formal definitions)

A **nondeterministic finite-state automaton** (NFA) has five components:

- the input alphabet Σ,
- a finite set of states Q,
- a transition function δ

NFA (Formal definitions)

A **nondeterministic finite-state automaton** (NFA) has five components:

- the input alphabet Σ,
- a finite set of states Q,
- a transition function $\delta : Q \times \Sigma \longrightarrow 2^Q$

NFA (Formal definitions)

A **nondeterministic finite-state automaton** (NFA) has five components:

- the input alphabet Σ,
- a finite set of states Q,
- a transition function $\delta : Q \times \Sigma \longrightarrow 2^Q$
- a start state $s \in Q$, and
- a subset $A \subseteq Q$ of accepting states.

NFA (Formal definitions)

A **nondeterministic finite-state automaton** (NFA) has five components:

- the input alphabet Σ,
- a finite set of states Q,
- a transition function $\delta : Q \times \Sigma \longrightarrow 2^Q$
- a start state $s \in Q$, and
- a subset $A \subseteq Q$ of accepting states.

Remark: δ can return the empty set $\emptyset$.

NFA (Formal definitions)

A **nondeterministic finite-state automaton** (NFA) has five components:

- the input alphabet Σ,
- a finite set of states Q,
- a transition function $\delta : Q \times \Sigma \longrightarrow 2^Q$
- a start state $s \in Q$, and
- a subset $A \subseteq Q$ of accepting states.

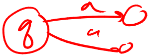
Remark: δ can return the empty set $\emptyset$.
What else do we need to define to "properly" talk about NFAs?

Transition

One step move: from state q with input symbol a, the machine changes its state to one of $\delta(q, a)$. $\longrightarrow$ จ. accepts w ถ้า

มีวิถีทางที่ไปจาก start state $\rightarrow$ accepting states

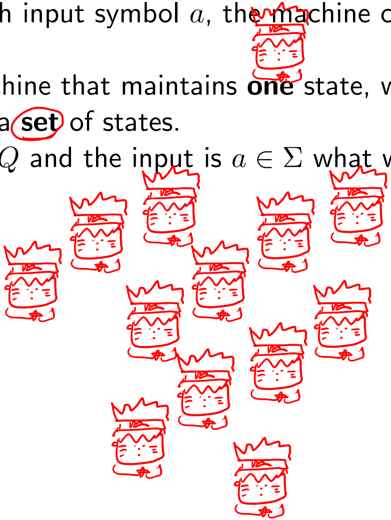
Transition



One step move: from state q with input symbol a, the machine changes its state to one of $\delta(q, a)$. $\subseteq Q$

Thus, instead of thinking of a machine that maintains **one** state, we can think of an NFA as a machine that maintains a **set** of states.

If the current set of states is $C \subseteq Q$ and the input is $a \in \Sigma$ what would the new set of states be?



Transition

One step move: from state q with input symbol a, the machine changes its state to one of $\delta(q, a)$.

Thus, instead of thinking of a machine that maintains **one** state, we can think of an NFA as a machine that maintains a **set** of states.

If the current set of states is $\boxed{C} \subseteq Q$ and the input is $a \in \Sigma$ what would the new set of states be?

Extension: from state q with input string w, the machine changes its set of states $\delta^*(q, w)$ defined as

$$\delta^*(q, w) = \begin{cases} \{q\} & \text{if } w = \varepsilon, \\ & \\ & w = \underbrace{a}_{} \cdot x \end{cases}$$

Transition

One step move: from state q with input symbol a, the machine changes its state to one of $\delta(q,a)$.

Thus, instead of thinking of a machine that maintains **one** state, we can think of an NFA as a machine that maintains a **set** of states.

If the current set of states is $C \subseteq Q$ and the input is $a \in \Sigma$ what would the new set of states be?

Extension: from state q with input string w, the machine changes its set of states $\delta^*(q,w)$ defined as

$$\delta^*(q,w) = \begin{cases} \{q\} & \text{if } w = \varepsilon, \\ \bigcup\limits_{r \in \delta(q,a)} \delta^*(r,x) & \text{if } w = ax. \end{cases}$$

The signature of δ^* is $Q \times \Sigma^* \longrightarrow 2^Q$.

Acceptance

accepting w

For a nondeterministic finite-state machine with starting state s and accepting states A, it accepts string w iff

$$\delta^*(s, w) \cap A \neq \emptyset.$$

Interpretation

Interpretation

- Clairvoyance. (สาพมู)

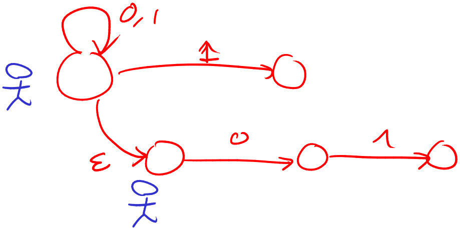
Interpretation

- Clairvoyance.
- Parallel threads. (ทาฎโท;)

Interpretation

- Clairvoyance.
- Parallel threads.
- Proofs/oracles.

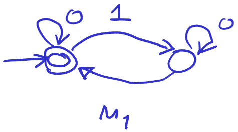
ε-transition

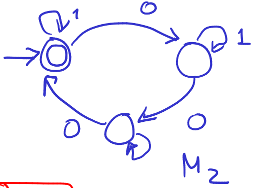


ε-transition

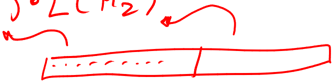
DFA M_1 DFA M_2

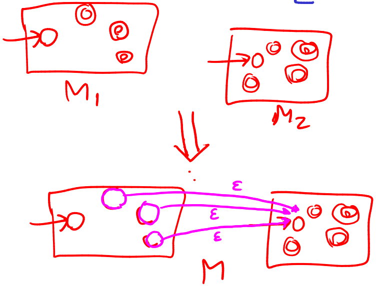
$L(M_1)$ $L(M_2)$

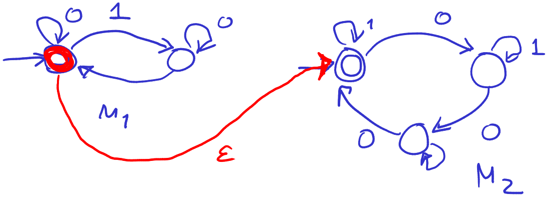




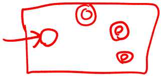
$L(M_1) \circ L(M_2)$







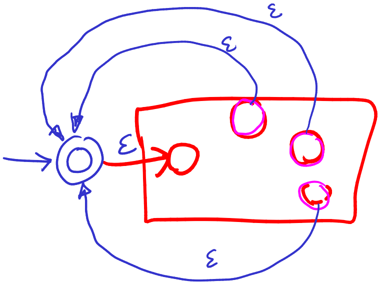
ε-transition



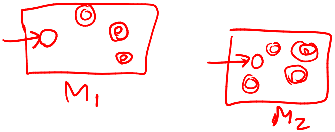
M_1

$L(M_1)$

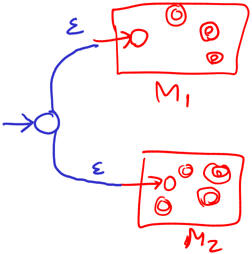
$L(M_1)^*$



ε-transition



$L(M_1) \cup L(M_2)$



ε-transition

1011

$1 0 \varepsilon 1 \varepsilon \varepsilon \varepsilon 1 \varepsilon \varepsilon$

ε

An NFA accepts string w iff there is a sequence of transitions

$$s \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \xrightarrow{a_3} q_3 \xrightarrow{a_4} \cdots \xrightarrow{a_{k-1}} q_{k-1} \xrightarrow{a_k} q_k,$$

where $q_k \in A$ and $w = a_1 a_2 \cdots a_k$ where $a_i \in \Sigma \cup \{\varepsilon\}$ for $1 \leq i \leq k$.

ε-transition

$\delta(q, \varepsilon) = \{ — \}$

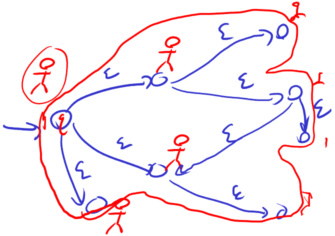
An NFA accepts string w iff there is a sequence of transitions

$$s \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \xrightarrow{a_3} q_3 \xrightarrow{a_4} \cdots \xrightarrow{a_{k-1}} q_{k-1} \xrightarrow{a_k} q_k,$$

where $q_k \in A$ and $w = a_1 a_2 \cdots a_k$ where $a_i \in \Sigma \cup \{\varepsilon\}$ for $1 \leq i \leq k$.
The transition function also changes its domain to $Q \times (\Sigma \cup \{\varepsilon\})$.

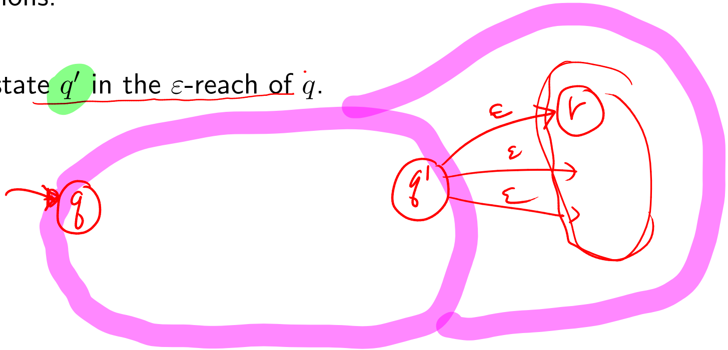
ε-transition: examples

ε-reach

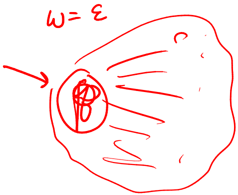


The ε-reach of state $q \in Q$ (denoted by ε-reach(q)) consists of all states r that satisfy one of the following conditions:

- $r = q$, or *
- $r \in \delta(q', \varepsilon)$ for some state q' in the ε-reach of q.

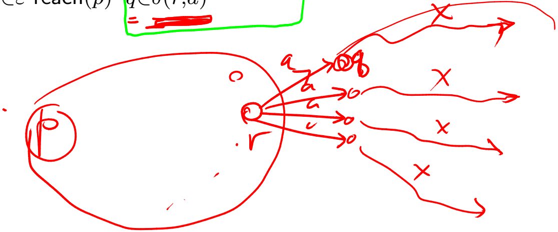


Extended transition function: δ^*

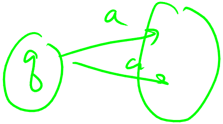


We define $\delta^* : Q \times \Sigma^* \longrightarrow 2^Q$ as follows:

$$\delta^*(p, w) = \begin{cases} \varepsilon\text{-reach}(p) & \boxed{\text{if } w = \varepsilon} \\ \displaystyle\bigcup_{r \in \varepsilon\text{-reach}(p)} \boxed{\bigcup_{q \in \delta(r,a)} \delta^*(q, x)} & \boxed{\text{if } w = ax.} \end{cases}$$

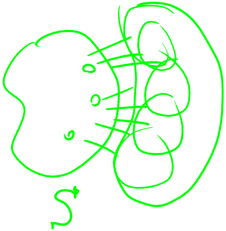


Notation abuse



We sometimes also write, for subset $S \subseteq Q$,

$$\delta(S, a) = \bigcup_{q \in S} \delta(q, a),$$



Notation abuse

We sometimes also write, for subset $S \subseteq Q$,

$$\delta(S, a) = \bigcup_{q \in S} \delta(q, a),$$

$$\delta^*(S, a) = \bigcup_{q \in S} \delta^*(q, a),$$

Notation abuse

We sometimes also write, for subset $S \subseteq Q$,

$$\delta(S, a) = \bigcup_{q \in S} \delta(q, a),$$

$$\delta^*(S, a) = \bigcup_{q \in S} \delta^*(q, a),$$

and

$$\varepsilon\text{-reach}(S) = \bigcup_{q \in S} \varepsilon\text{-reach}(q).$$

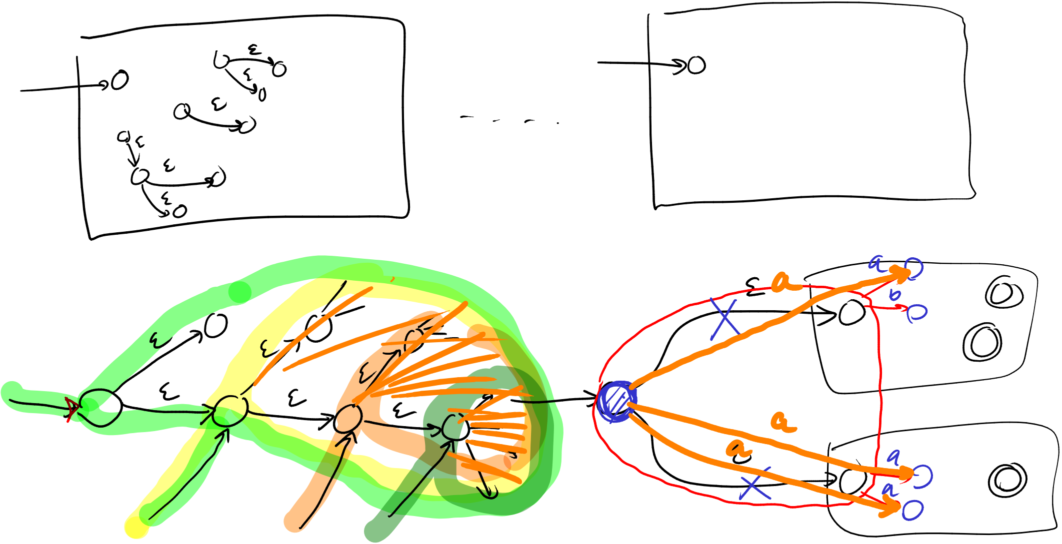
Extended transition function: δ^* (with shorter notation)

$\delta^*(\delta(q,a),x)$

We define $\delta^* : Q \times \Sigma^* \longrightarrow 2^Q$ as follows:

$$\delta^*(p,w) = \begin{cases} \varepsilon\text{-reach}(p) & \text{if } w = \varepsilon \\ \underbrace{\delta^*(\underbrace{\delta(\underbrace{\varepsilon\text{-reach}(p)}),a)},x)} & \text{if } w = ax. \end{cases}$$

Removing ε-transitions: idea



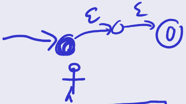
Lemma 1

For any NFA $M = (\Sigma, Q, \delta, s, A)$ *with* ε*-transitions, there is an NFA* $M' = (\Sigma, Q', \delta', s', A')$ *without* ε*-transitions such that* $L(M) = L(M')$.

Proof.

ให้ $Q' = Q$, $s' = s$,

และนิยาม δ' ดังนี้ , $\forall q \in Q'$, $\forall a \in \Sigma$



$$\delta'(q,a) = \delta(\varepsilon\text{-reach}(q), a) , \quad \boxed{\delta'(q,\varepsilon) = \phi}$$

และให้ $A' = \{ q \in Q' \mid \varepsilon\text{-reach}(q) \cap A \neq \emptyset \}$

• แสดงว่า $L(M) = L(M')$, สังเกตว่า $\delta(q,w) = \delta'(q,w)$

$\equiv$ (ละไว้)

$\square$

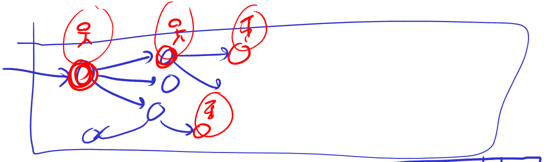
Main question

- We see that ε-transitions does not add any "power" to the machine.
- Does nondeterminism add any power to NFA (over typical DFA)?

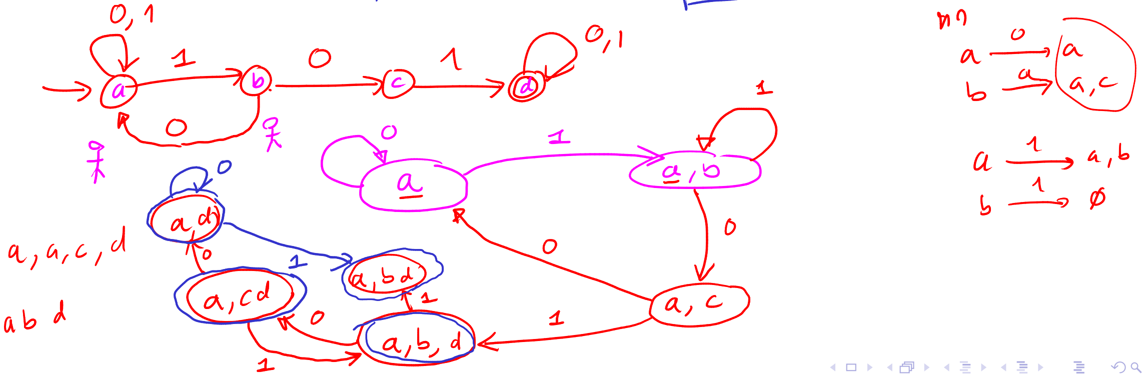
Simulating parallel machines

NFA N

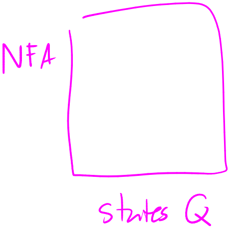
(ไม่มี ε-trans)

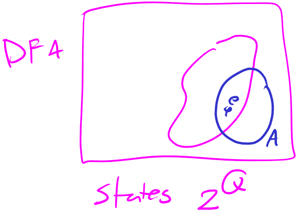


การจำลองการทำงาน : จำ $\boxed{\text{set ของ states}}$ ที่ไปถึงได้



Subset construction: idea





NFA to DFA: subset construction

$\delta' : Q' \times \Sigma \to Q'$ $\delta : Q \times \Sigma \to 2^Q$

Given an NFA $M = (\Sigma, Q, \delta, s, A)$, we can construct an equivalent DFA $M' = (\Sigma, Q', \delta', s', A')$ as follows:

- Let $Q' = 2^Q$,
- $s' = \{s\}$,

NFA to DFA: subset construction

ไม่มี ε-tran

Given an NFA $M = (\Sigma, Q, \delta, s, A)$, we can construct an equivalent DFA $M' = (\Sigma, Q', \delta', s', A')$ as follows:

- Let $Q' = 2^Q$,
- $s' = \{s\}$,
- $\boxed{A'} = \{S \subseteq Q \mid S \cap A \neq \emptyset\}$,
- and let $\delta' : Q' \times \Sigma \longrightarrow Q'$ be such that

NFA to DFA: subset construction

abd

Given an NFA $M = (\Sigma, Q, \delta, s, A)$, we can construct an equivalent DFA $M' = (\Sigma, Q', \delta', s', A')$ as follows:

- Let $Q' = 2^Q$,
- $s' = \{s\}$,
- $A' = \{S \subseteq Q \mid S \cap A \neq \emptyset\}$,
- and let $\delta' : Q' \times \Sigma \longrightarrow Q'$ be such that

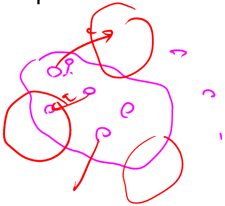
is a subset of Q

$$\delta'(q', a) = \boxed{\bigcup_{p \in q'} \delta(p, a)},$$

NFA

for all $q' \subseteq Q$ and $a \in \Sigma$.

$q' \in Q'$



NFA

Example

$\text{reg} \rightarrow \text{NFA} \rightleftarrows \text{DFA}$

??

Kleene's Theorem

Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Kleene's Theorem

Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Steps:

- Every DFA can be transformed into an equivalent NFA.

Kleene's Theorem

Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Steps:

- Every DFA can be transformed into an equivalent NFA. (trivial)

Kleene's Theorem

Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Steps:

- Every DFA can be transformed into an equivalent NFA. (trivial)
- Every NFA can be transformed into an equivalent DFA.

Kleene's Theorem

Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Steps:

- Every DFA can be transformed into an equivalent NFA. (trivial)
- Every NFA can be transformed into an equivalent DFA. (done)

Kleene's Theorem

Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Steps:

- Every DFA can be transformed into an equivalent NFA. (trivial)
- Every NFA can be transformed into an equivalent DFA. (done)
- Every regular expression can be transformed into an equivalent NFA.

Kleene's Theorem

Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Steps:

- Every DFA can be transformed into an equivalent NFA. (trivial)
- Every NFA can be transformed into an equivalent DFA. (done)
- Every regular expression can be transformed into an equivalent NFA. (TODO)

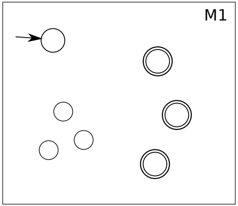
Kleene's Theorem

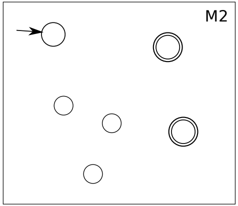
Every language L can be described by a regular expression if and only if L is the language accepted by a DFA.

Steps:

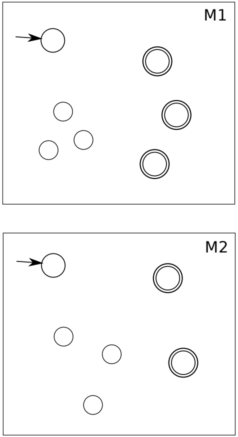
- Every DFA can be transformed into an equivalent NFA. (trivial)
- Every NFA can be transformed into an equivalent DFA. (done)
- Every regular expression can be transformed into an equivalent NFA. (TODO)
- Every NFA can be transformed into an equivalent regular expression. (only idea) *

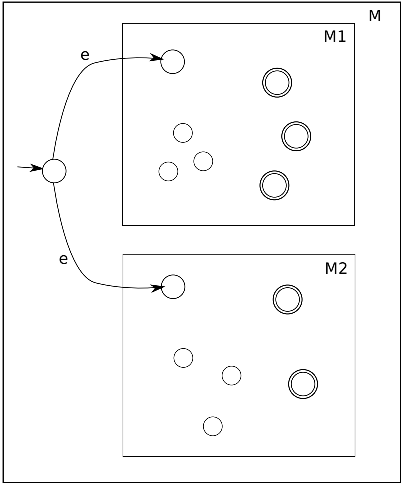
Warm-up: union of DFA $\Longrightarrow$ NFA



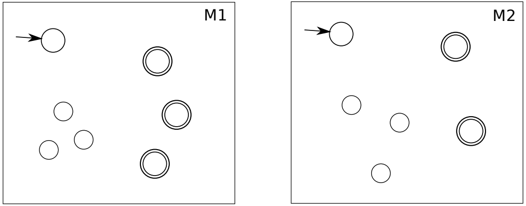


Warm-up: union of DFA $\Longrightarrow$ NFA





Concatenation: idea

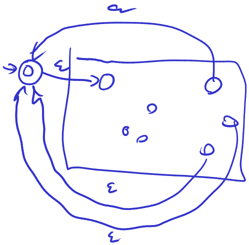


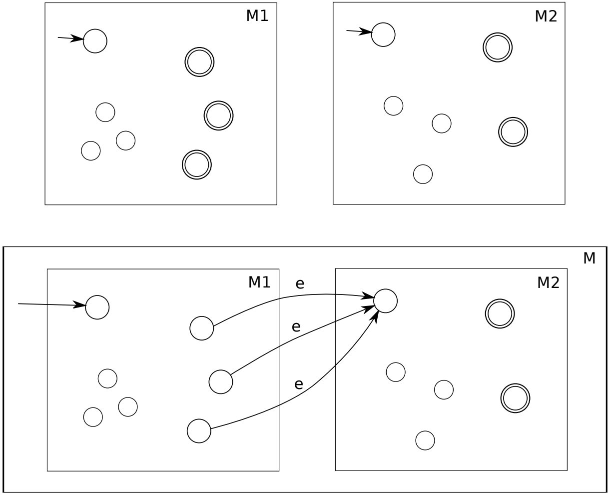
Concatenation: idea

$(0+1)(00)^*$

$1^* + (10)^* + (100)^*$

$L(M_1)^*$





Stronger claim

Our goal is to prove:

Lemma 2

Every regular language is accepted by a nondeterministic finite-state automaton.

Stronger claim

Our goal is to prove:

Lemma 2

Every regular language is accepted by a nondeterministic finite-state automaton.

But we will prove a "stronger" claim.

Lemma 3 (Thompson's algorithm)

Every regular language is accepted by a nondeterministic finite-state automaton with exactly one accepting state, which is different from its start state.

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

- $R = \emptyset$:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

- $R = \emptyset$:
- $R = \varepsilon$:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

- $R = \emptyset$:
- $R = \varepsilon$:
- $R = a$ for some $a \in \Sigma$:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

- $R = \emptyset$:
- $R = \varepsilon$:
- $R = a$ for some $a \in \Sigma$:
- $R = ST$ for some regular expression S and T:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

- $R = \emptyset$:
- $R = \varepsilon$:
- $R = a$ for some $a \in \Sigma$:
- $R = ST$ for some regular expression S and T:
- $R = S + T$ for some regular expression S and T:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

- $R = \emptyset$:
- $R = \varepsilon$:
- $R = a$ for some $a \in \Sigma$:
- $R = ST$ for some regular expression S and T:
- $R = S + T$ for some regular expression S and T:
- $R = S^*$ for some regular expression S:

Proof (Thompson's algorithm).

Consider any regular expression R over alphaget Σ. We prove that there is an NFA N that accepts the language described by R by induction.

Induction hypothesis: for any subexpression S of R, there is an NFA that accepts the language described by S.

We denote an NFA with this notation:

There are 6 cases:

- $R = \emptyset$:
- $R = \varepsilon$:
- $R = a$ for some $a \in \Sigma$:
- $R = ST$ for some regular expression S and T:
- $R = S + T$ for some regular expression S and T:
- $R = S^*$ for some regular expression S:

In all cases, the language $L(R)$ is accepted by an NFA with exactly one accepting state which is different from its start state, as required. $\square$

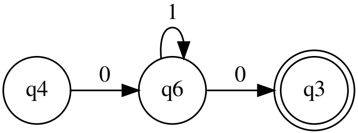
Example: $1 + 00$

Example: $(1 + 00)^*$

Example: $(1+00)^* + 1^*0$

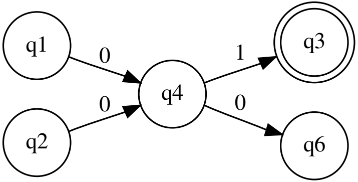
NFA to Regular expressions

State elimination: example 1



$0\,1^*0$

State elimination: example 2



State elimination: example 3