eecs2001

Introduction to the Theory of Computation

Lecture 10

Ruth Urner

Febuary 8, 2023

Non-deterministic finite automata

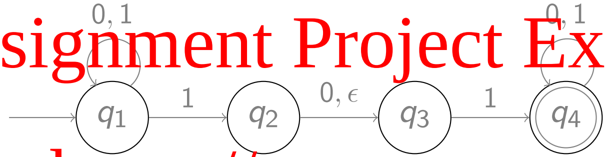
Reading: ITC Section 1.2

DFA versus NFAs

To show that regular languages are also closed under the concatenation and star operations, it is convenient to employ the concept of a Non-deterministic Finite Automaton (NFA for short).

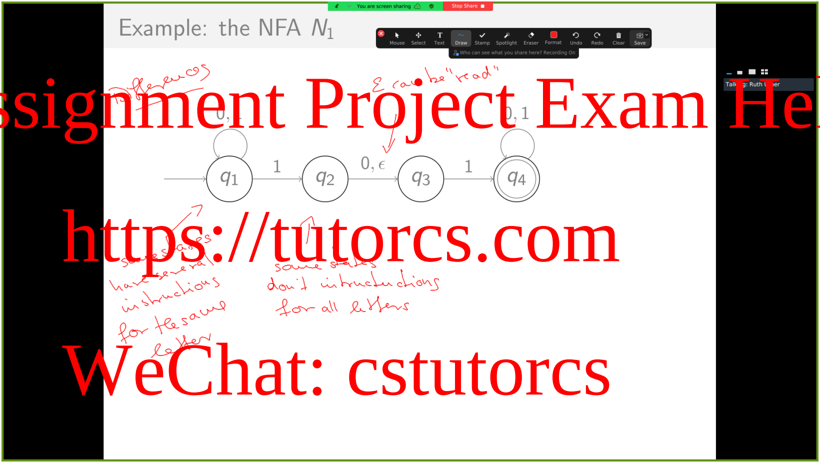
In the automata we have seen so far, every state has exactly one outgoing edge for every symbol. That is, for every state, the "reaction" to any input symbol is uniquely determined. Such an automaton is also called a Deterministic Finite Automaton (DFA for short).

We will generalize this to also allow non-unique (think randomized or parallelized) computations. NFAs are a model for such computations.

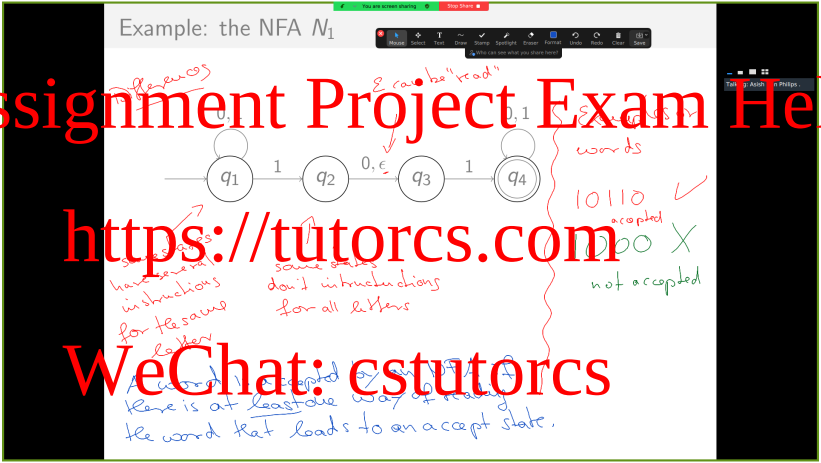


For the machine N_1 we get

$$L(N_1) = \{\mathbf{w} \mid \mathbf{w} \text{ contains 11 or 101 as a substring}\}$$



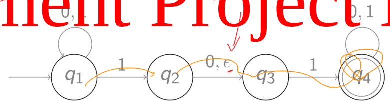
Example: the NFA N_1
differences
ε can be "read"
0, 1
0, 1
q_1
1
q_2
0, ε
q_3
1
q_4
some states have several instructions for the same letter
some states don't instructions for all letters


Example: the NFA N_1
differences
ε can be "read"
0, 1
0, 1
1
0, ε
1
q_1
q_2
q_3
q_4
words
10110
accepted
not accepted
some states have several instructions for the same letter
some states don't instructions for all letters
there is at least one way of reading the word that leads to an accept state.

Example: the NFA N_1

differences

ε can be "read"



Examples of words

10110 ✓ accepted

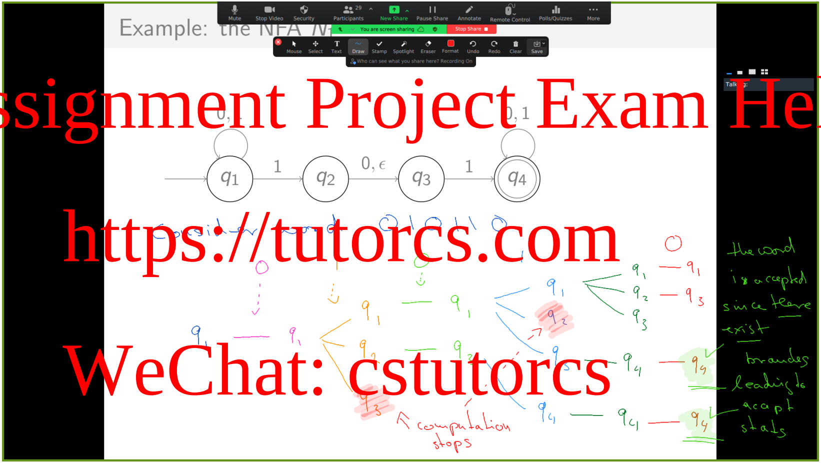
1000 X not accepted

some states have several instructions for the same letter

some states don't instructions for all letters

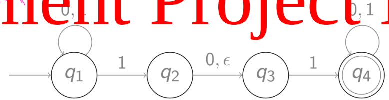
$1 = \varepsilon 1$

A word is accepted by an NFA if there is at least one way of reading the word that leads to an accept state.


Example: the NFA N
0, 1
0, 1
q1
1
q2
0, ε
q3
1
q4
the word is accepted since there exist branches leading to accept states
computation stops

Example: the NFA N_1

10001 not accepted



Reminder: For alphabet Σ, the set of all words over Σ is denoted by Σ^*.

For the machine N_1 we get

$L(N_1) = \{ w \in \{0,1\}^* \mid w$ contains 101 or 11 as a substring$\}$

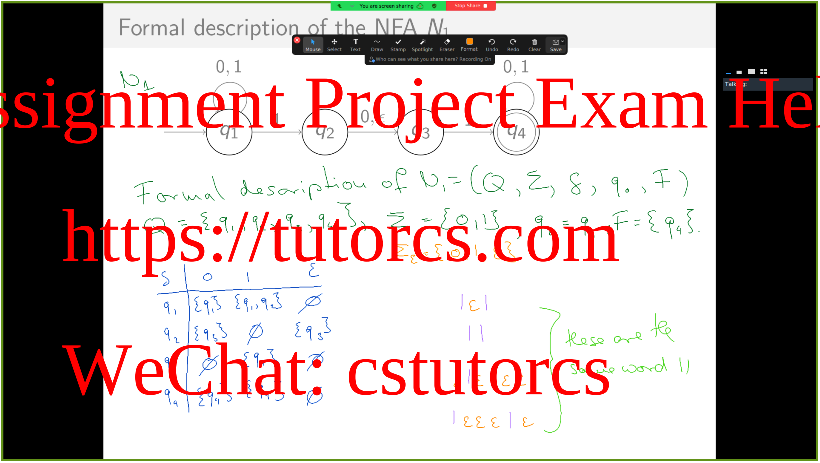
Formal definition of a non-deterministic finite automaton

For an alphabet Σ, we define: $\Sigma_\epsilon = \Sigma \cup \{\epsilon\}$.

Definition

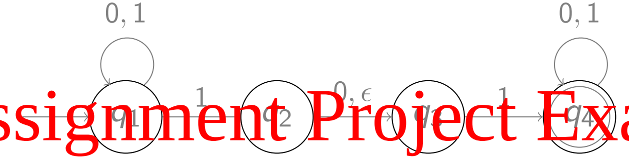
A non-deterministic finite automaton is a 5-tuple $(Q, \Sigma, \delta, q_0, F)$, where

1. Q is a finite set called the set of states,
2. Σ is a finite set called the alphabet,
3. $\delta : Q \times \Sigma_\epsilon \to \mathcal{P}(Q)$ is the transition function,
4. $q_0 \in Q$ is the start state, and
5. $F \subseteq Q$ is the set of accept states.



Formal description of the NFA N_1
0, 1
0, 1
q_1
q_2
q_3
q_4
0, ε
Formal description of $N_1 = (Q, \Sigma, \delta, q_0, F)$
$\Sigma_\varepsilon = \{0, 1, \varepsilon\}$
these are the same word

Formal description of the NFA N_1



For the machine N_1 we get

$$L(N_1) = \{\mathbf{w} \mid \mathbf{w} \text{ contains 11 or 101 as a substring}\}$$

For a string $\mathbf{w} = w_1 w_2 \ldots w_n$ where $w_i \in \Sigma$ for some alphabet Σ, we let strings $\mathbf{z}_0 w_1 \mathbf{z}_1 w_2 \mathbf{z}_3 \ldots \mathbf{z}_{n-1} w_n \mathbf{z}_n$ over Σ_ϵ where each $\mathbf{z}_i$ is a sequence of 0 or more symbols ϵ represent the same word $\mathbf{w}$ as a string over Σ.

Definition
Let $N = (Q, \Sigma, \delta, q_0, F)$ a non-deterministic finite automaton (NFA) and let $\mathbf{w}$ be a string over Σ. We say that N accepts $\mathbf{w}$ if we can write $\mathbf{w} = y_1 y_2 \ldots y_m$ with $y_i \in \Sigma_\epsilon$ and there exists a sequence $s_0 s_1 s_2 \ldots s_m$ of states such that

1. $s_0 = q_0$,
2. $s_{i+1} \in \delta(s_i, y_{i+1})$ for $i = 0, 1, \ldots, m$, and
3. $s_m \in F$.

We then define the language $L(N)$ recognized by NFA N in the same way as for DFAs, namely as the set of all words that are accepted by N.

Formal definition of acceptance for NFAs

For a string $\mathbf{w} = w_1 w_2 \ldots w_n$ where $w_i \in \Sigma$ for some alphabet Σ, we let strings $\mathbf{z}_0 w_1 \mathbf{z}_1 w_2 \mathbf{z}_2 \ldots \mathbf{z}_{n-1} w_n \mathbf{z}_n$ over Σ_ϵ where each $\mathbf{z}_i$ is a sequence of 0 or more symbols ϵ represent the same word $\mathbf{w}$ as a string over Σ.

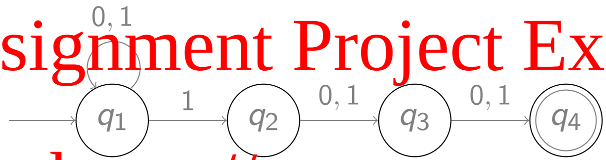
Definition

Let $N = (Q, \Sigma, \delta, q_0, F)$ a non-deterministic finite automaton (NFA) and let $\mathbf{w}$ be a string over Σ. We say that N accepts $\mathbf{w}$ if we can write $\mathbf{w} = y_1 y_2 \ldots y_m$ with $y_i \in \Sigma_\epsilon$ and there exists a sequence $s_0 s_1 s_2 \ldots s_m$ of states such that

1. $s_0 = q_0$,
2. $s_{i+1} \in \delta(s_i, y_{i+1})$ for $i = 0, 1, \ldots, m$, and
3. $s_m \in F$.

$\varepsilon|\varepsilon|$

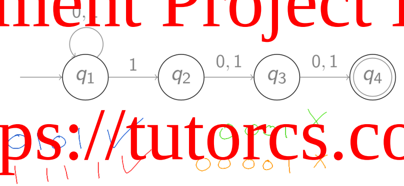
Example: the NFA N_2



For the machine N_2 we get

$$L(N_2) = \{\mathbf{w} \mid \mathbf{w} \text{ contains } 1 \text{ at the third position from the end}\}$$

Example: the NFA N_2



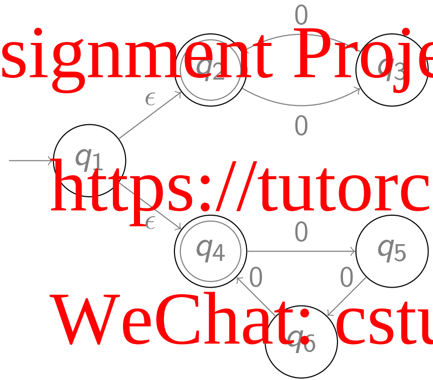
For the machine N_2 we get

$L(N_2) = \{ w \in \{0,1\}^* \mid w$ has a 1 at the third position from the end $\}$

$= \{ w = w_1 w_2 \dots w_n \in \Sigma^* \mid w_{n-2} = 1 \}$

Example: the NFA N_3

Let's consider a unary alphabet $\Sigma = \{0\}$.

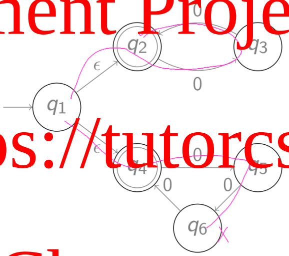


For the machine N_3 we get

$$L(N_3) = \{\mathbf{w} \mid \text{The number of 0 symbols in } \mathbf{w} \text{ is divisible by 2 or by 3}\}$$

Example: the NFA N_2

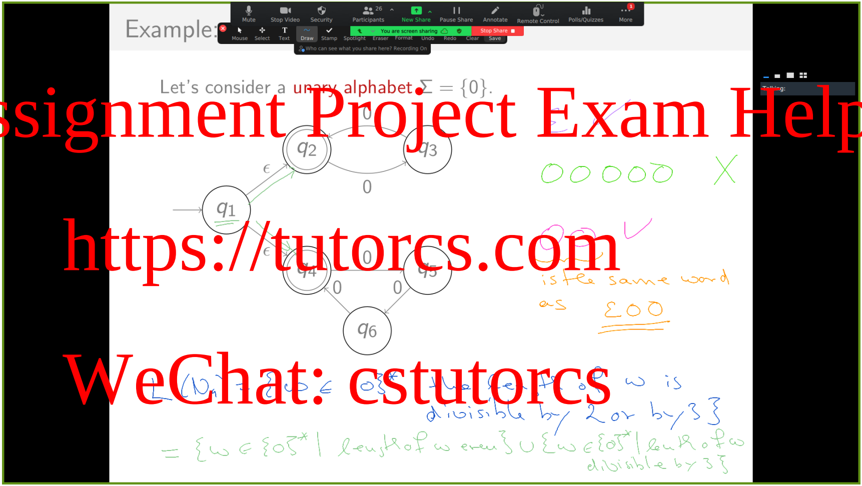
Let's consider a unary alphabet $\Sigma = \{0\}$.



00000 X

00 ✓

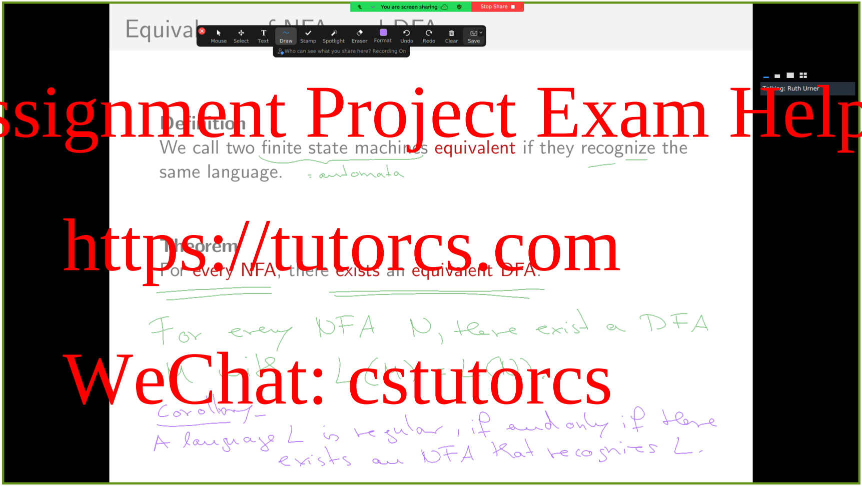
$L(N_2) = \{w \in 0^*$ the length of w is divisible by 2 or by 3$\}$


Example:
Let's consider a unary alphabet Σ = {0}.
q1
q2
q3
q4
q5
q6
ε
0
00000
is the same word as
ε00
L(N) = {w ∈ {0}* | the length of w is divisible by 2 or by 3}
= {w ∈ {0}* | length of w even} ∪ {w ∈ {0}* | length of w divisible by 3}

Definition
We call two finite state machines equivalent if they recognize the same language.

Theorem
For every NFA, there exists an equivalent DFA.

Corrolary
A language L is regular if and only if there exists an NFA that recognizes it.

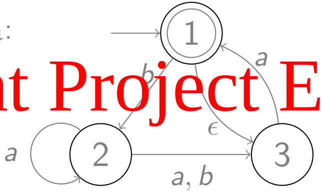

You are screen sharing
Stop Share
Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save
Who can see what you share here? Recording On
Talking: Ruth Urner
Equival
Definition
We call two finite state machines equivalent if they recognize the same language.
= automata
Theorem
For every NFA, there exists an equivalent DFA.
For every NFA N, there exist a DFA
M s.t. L(N) = L(M).
Corollary-
A language L is regular, if and only if there exists an NFA that recognizes L.

To prove the theorem from the previous slide, we need to show that for every NFA N, there exists a DFA M with $L(M) = L(N)$.

Before we see the general proof for this, we go through one concrete example. We transform NFA N_4 into an equivalent DFA.

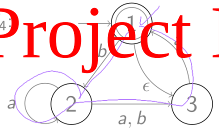
Transforming an NFA into a DFA–example

Consider the NFA N_4:



Transforming an NFA into a DFA–example

Consider the NFA N_4:



Examples of accepted words

a

''

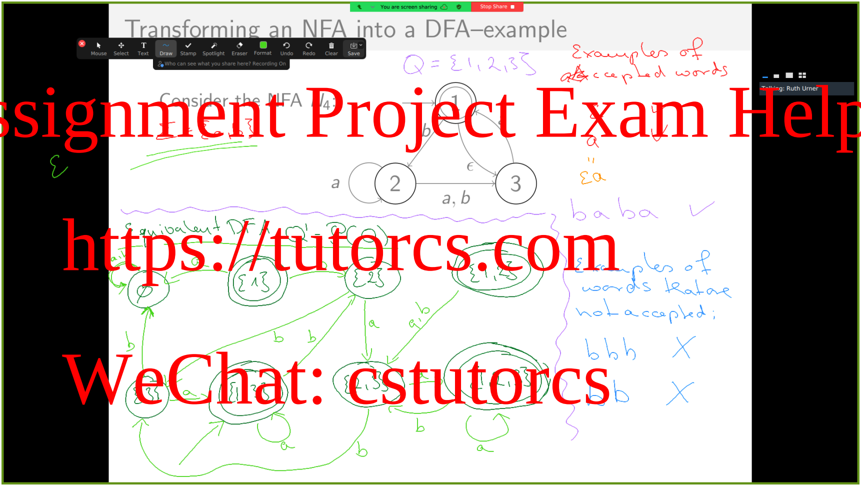
εa

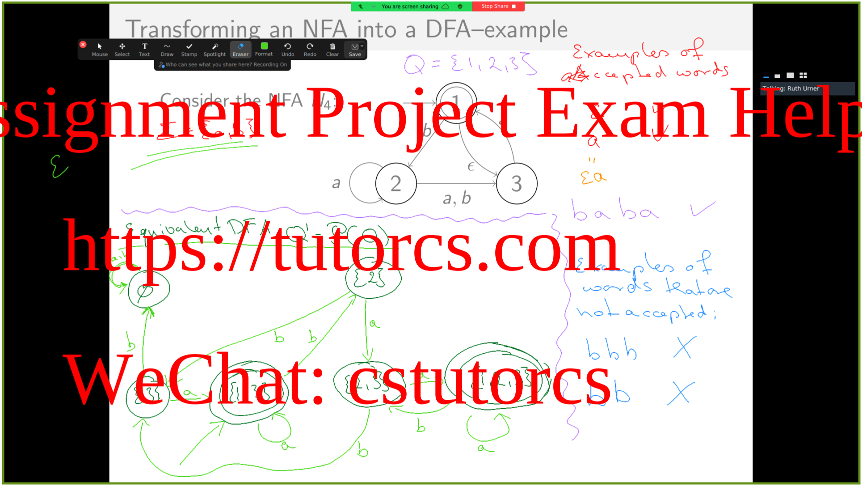
baba ✓

Examples of words that are not accepted;

bbb X

bb X


You are screen sharing
Stop Share
Transforming an NFA into a DFA–example
Mouse
Select
Text
Draw
Stamp
Spotlight
Eraser
Format
Undo
Redo
Clear
Save
Who can see what you share here? Recording On
Talking: Ruth Urner
Consider the NFA N_4:
Q = {1, 2, 3}
Examples of accepted words
a
εa
baba
Equivalent DFA
Examples of words that are not accepted:
bbb
bb
a, b
ε
b
a


You are screen sharing
Stop Share
Transforming an NFA into a DFA–example
Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save
Who can see what you share here? Recording On
Consider the NFA N_4:
Q = {1,2,3}
Examples of accepted words
1
2
3
a
b
a, b
ε
baba
Examples of words that are not accepted;
bbb

General proof of Theorem

Proof

Let $N = (Q, \Sigma, \delta, q_0, F)$ be some NFA.

We need to construct a DFA $M = (Q', \Sigma, \delta', q_0', F')$ with $L(M) = L(N)$.

Assuming that N does not contain any ϵ-transitions, we construct M as follows:

1. $Q' = \mathcal{P}(Q)$

2. For $R \in Q'$ and $a \in \Sigma$, we let

$$\delta'(R, a) = \{q \in Q \mid q \in \delta(r, a) \text{ for some } r \in R\}$$

3. $q_0' = \{q_0\}$

4. $F' = \{R \in Q' \mid \text{there exists an } r \in F \text{ with } r \in R\}$.

This completes the construction for NFAs that don't contain ϵ-transitions.

General proof of Theorem

Proof

Let $N = (Q, \Sigma, \delta, q_0, F)$ be some NFA.

We need to construct a DFA $M = (Q', \Sigma, \delta', q_0', F')$ with $L(M) = L(N)$.

Assuming that N does not contain any ϵ-transitions, we construct M as follows:

$\in Q' = \mathcal{P}(Q)$

1. $Q' = \mathcal{P}(Q)$
2. For $R \in Q'$ and $a \in \Sigma$ we let

$$\delta'(R, a) = \{q \in Q \mid q \in \delta(r, a) \text{ for some } r \in R\}$$

3. $q_0' = \{q_0\}$
4. $F' = \{R \in Q' \mid \text{there exists an } r \in R \text{ with } r \in F\}$.

This completes the construction for NFAs that don't contain ϵ-transitions.

Proof of Theorem

Proof continued

If N does contain ϵ-transitions, we extend the above construction as follows:

For a state $R \in Q'$, we let $E(R)$ denote the set of states that can be reached via 0 or more ϵ-transitions from some state $r \in R$ (in N).

Then we modify the transition function of M to be

$$\delta'(R, a) \;=\; \{q \in Q \;\mid\; q \in E(\delta(r, a)) \text{ for some } r \in R\}$$

and the start state of M to be

$$q_0' \;=\; E(q_0).$$

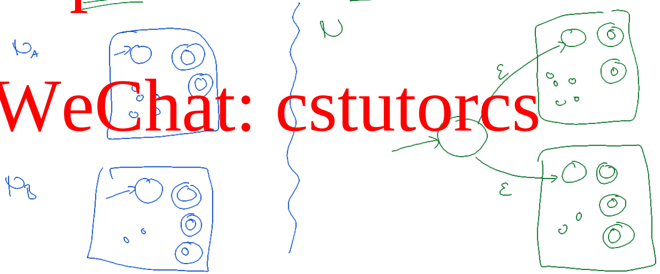
Regular languages are closed under regular operations

Theorem

Regular languages are closed under unions.

"Proof" by illustration

Given A and B and NFAs N_A and N_B with $L(N_A)=A$ and $L(N_B)=B$. We need to construct an NFA N that recognizes $A \cup B$.



Regular languages are closed under regular operations

Theorem

Regular languages are closed under unions.

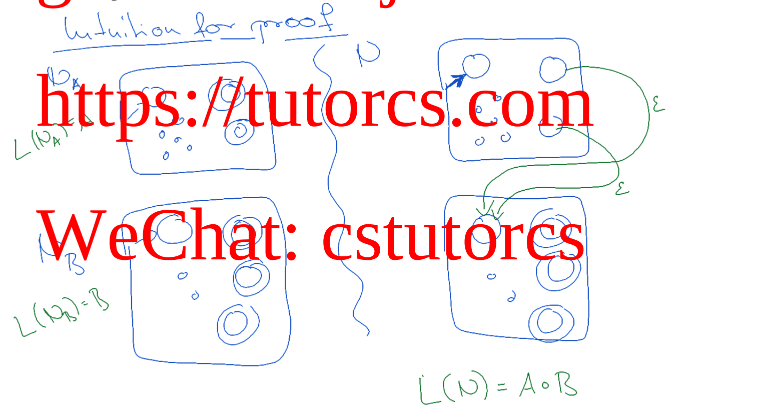
Proof (formal version)

We need to show that if two languages A and B are regular, then so is $A \cup B$. If A and B are regular, then there exist NFAs $N_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ and $N_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$ such that $L(N_1) = A$ and $L(N_2) = B$. We need to show that there exists an NFA M with $L(M) = A \cup B$. The picture on the previous slide illustrates a construction for $M = (Q, \Sigma, \delta, q_0, F)$. Formally, we set:

1. $Q = Q_1 \cup Q_2 \cup \{q_0\}$, where q_0 is a new state, that is $q_0 \notin Q_1 \cup Q_2$ and this state q_0 will then also be the start state of M.
2. $\delta(q, a) = \begin{cases} \{q_1, q_2\} & \text{if} \quad q = q_0 \quad \text{and} \quad a = \epsilon \\ \delta_i(q, a) & \text{if} \quad q \in Q_i \\ \emptyset & \text{else} \end{cases}$
3. $F = F_1 \cup F_2$.

Theorem
Regular languages are closed under concatenations.

"Proof" by Illustration



Theorem

Regular languages are closed under concatenations.

Proof (formal version)

We need to show that if two languages A and B are regular, then so is $A \circ B$. If A and B are regular, then there exist NFAs $N_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ and $N_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$ such that $L(N_1) = A$ and $L(N_2) = B$. We need to show that there exists an NFA M with $L(M) = A \circ B$. The picture on the previous slide illustrates a construction for $M = (Q, \Sigma, \delta, q_0, F)$. Formally, we set:

1. $Q = Q_1 \cup Q_2$
2. $q_0 = q_1$
3. $\delta(q, a) = \begin{cases} \delta_1(q, a) \cup \{q_2\} & \text{if} \quad q \in F_1 \text{ and } a = \epsilon \\ \delta_1(q, a) & \text{if} \quad q \in Q_1 \text{ and } (q \notin F_1 \text{ or } a \neq \epsilon) \\ \delta_2(q, a) & \text{if} \quad q \in Q_2 \end{cases}$
4. $F = F_2$.

Regular languages are closed under regular operations

Theorem

Regular languages are closed under the star-operation.

“Proof” by illustration

Regular languages are closed under regular operations

Theorem

Regular languages are closed under the star-operation.

Proof (formal version)

We need to show that if languages A is regular, then so is A^*. If A is regular, then there exist an NFA $N_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ such that $L(N_1) = A$. We need to show that there exists an NFA M with $L(M) = A^*$. The picture on the previous slide illustrates a construction for $M = (Q, \Sigma, \delta, q_0, F)$. We set:

1. $Q = Q_1 \cup \{q_0\}$, where q_0 is a new state, that is $q_0 \notin Q_1 \cup Q_2$ and this state q_0 will then also be the start state of M.

2. $\delta(q,a) = \begin{cases} \{q_1\} & \text{if} \quad q = q_0 \quad \text{and} \quad a = \epsilon \\ \delta_1(q,a) & \text{if} \quad q \in Q_1 \text{ and } (q \notin F_1 \text{ or } a \neq \epsilon) \\ \delta_1(q,a) \cup \{q_1\} & \text{if} \quad q \in F_1 \text{ and } a = \epsilon \end{cases}$

3. $F = F_1 \cup \{q_0\}$.