THEORY OF COMPUTATION

Chapter7
Regular Languages
(Automata as Algebras)

A. Daneshgar
(Slides prepared by Z. Ghafouri)

Department of Mathematical Sciences
Sharif University of Technology

2020, March, 1st (1998, Esfand, 10)

An algebra

Definition: An algebra is a pair $\mathcal{A} = (A, O)$ where A is a set and O is a set of operations on A.

- An $n-$ary operation τ on $\mathcal{A}$ is a function that takes n elements of $\mathcal{A}$ and returns a single element of $\mathcal{A}$; i.e.

$$\tau : \underbrace{A \times \ldots \times A}_{n} \to A,$$

- A $0-$ary operation (nullary operation) is simply an element of A, or a constant,
- A $1-$ary operation (unary operation) is simply a function from A to A,
- A $2-$ary operation (binary operation) is simply a function from $A \times A$ to A.

A pre-automaton

Definition: A pre-automaton

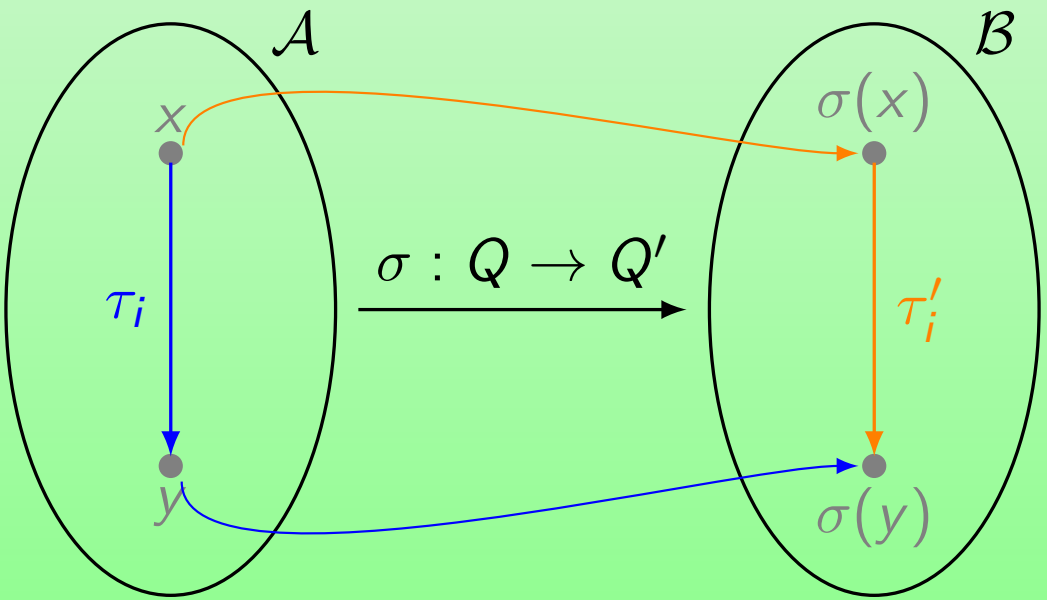
$$\mathcal{A} = \left(Q, q_0 \in Q, \{\tau_i : Q \to Q\}_{i \in \Sigma}\right)$$

with input symbols Σ of n elements, is an algebra with one nullary and n unary operations on a set of states A.

Homomorphisms

$$\mathcal{A} = \left(Q, q_0 \in Q, \{\tau_i : Q \to Q\}_{i \in \Sigma}\right)$$

$$\mathcal{B} = \left(Q', q_0{}' \in Q', \{\tau_i{}' : Q' \to Q'\}_{i \in \Sigma}\right)$$



$$\forall i \quad \tau_i' \circ \sigma = \sigma \circ \tau_i$$

$$\sigma(q_0) = q_0'$$

Isomorphisms

Definition: Two algebras

$$\mathcal{A} = \left(Q, q_0 \in Q, \{\tau_i : Q \to Q\}_{i \in \Sigma}\right)$$

and

$$\mathcal{B} = \left(Q', {q_0}' \in Q', \{{\tau_i}' : Q' \to Q'\}_{i \in \Sigma}\right)$$

are said to be isomorphic if there exists a bijective homomorphism $\sigma : \mathcal{A} \to \mathcal{B}$ such that its inverse, $\sigma^{-1} : \mathcal{B} \to \mathcal{A}$, is also a bijective homomorphism.

Subalgebras

- Let $\mathcal{A} = (A, O)$ be an algebra and let $\tau \in O$ be an $n-$ary operation of A. A subset B of A is $\tau-$ *closed* if $a_0, \ldots, a_{n-1} \in B$ imply $\tau(a_0, \ldots, a_{n-1}) \in B$.
- B is a *closed subset* of $\mathcal{A}$ if B is $\tau-$ closed for each operation τ of $\mathcal{A}$.
- A subalgebra of the algebra $\mathcal{A} = (A, O)$ is an algebra $\mathcal{B} = (B, O_B)$, where $B \subseteq A$ is a closed subset of A and

$$O_B = \{\tau|_{B^n} \ : \ \tau \in O \text{ is an } n-\text{ary operation}\}.$$

In other words, a subalgebra is a subset of an algebra which is closed under all its operations and carrying the induced operations.

- A reduced algebra is an algebra whose only subalgebras are the empty set and itself.

Reduced pre-automata

A reduced pre-automaton

$$\mathcal{A} = \left(Q, q_0 \in Q, \{\tau_i : Q \to Q\}_{i \in \Sigma}\right)$$

is a pre-automaton which is reduced as an algebra. Note that this is equivalent to the condition that each state of Q is reachable from the initial state q_0.

Equivalence relations

- An equivalence relation on X is a binary relation $\sim$ on X such that
  - $\forall x \in X \ (x \sim x)$ (reflexivity),
  - $\forall x, y \in X \ (x \sim y \rightarrow y \sim x)$ (simmetry),
  - $\forall x, y, z \in X \ (x \sim y \ \wedge \ y \sim z \rightarrow x \sim z)$ (transitivity).
- The equivalence class of an element x in X with respect to the equivalence relation $\sim$, denoted by $[x]_{\sim}$, is the set of all elements of X which are equivalent to x, i.e.

$$[x]_{\sim} = \{y \in X \mid x \sim y\}.$$

- Every two equivalence classes $[x]_\sim$ and $[y]_\sim$ are either equal or disjoint, so the set of equivalence classes is a partition of X.
- The set of all equivalence classes of X by $\sim$, denoted by $X/\sim$, is called the quotient set of X.
- For any partition $\mathcal{P}$ of a set X, the relation $\sim$ on X defined by

  $x \sim y$ if and only if x and y belong to the same element of $\mathcal{P}$,

  is an equivalence relation on X.

  Hence, talking about equivalence relations on a set is essentially the same as talking about partitions of a set: An equivalence relation determines a partition (in which the subsets are the equivalence classes), and a partition determines an equivalence relation (in which being equivalent means belonging to the same subset).

Equivalence coming from an onto-map

Any onto-map $\sigma : A \xrightarrow{onto} B$ gives rise to an equivalence relation on A according to which

$$a_1 \sim_\sigma a_2 \Leftrightarrow \sigma(a_1) = \sigma(a_1).$$

Hence, the collection of all $\sigma-$inverse images of the elements of B forms a partition of A.

Congruence relations

Definition: Let $\mathcal{A} = (A, O)$ be an algebra. A congruence relation is an equivalence relation on A which is compatible with the algebra operations. In the case of a pre-automaton where each operation is either nullary or unary, it means that for every x and y in A and each unary operation τ_i:

$$[x] = [y] \Leftrightarrow [\tau_i(x)] = [\tau_i(y)].$$

The quotient pre-automaton

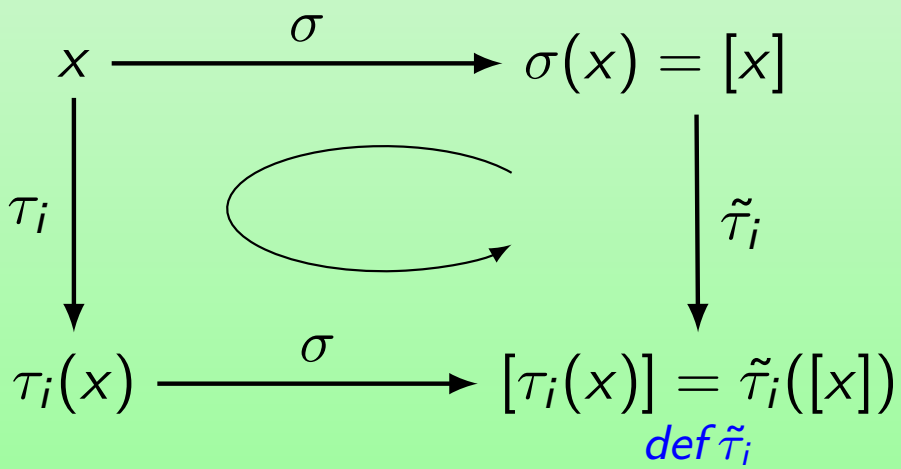
Definition: Let $\mathcal{A} = \left(Q, q_0 \in Q, \{\tau_i : Q \to Q\}_{i \in \Sigma}\right)$ be a pre-automaton and $\sim$ be a congruence relation on $\mathcal{A}$. The quotient pre-automaton of $\mathcal{A}$ by $\sim$ is

$$\mathcal{A}/\sim = \left(Q/\sim,\ [q_0],\ \{\tilde{\tau}_i : Q/\sim \to\ \ Q/\sim\}_{i \in I}\right)$$

where $\tilde{\tau}_i([x]) = [\tau_i(x)]$.

The canonical homomorphism

Consider the map $\sigma : \mathcal{A} \to \mathcal{A}/\sim$ defined as $\sigma(x) = [x]$; then σ is called the *natural* or the *canonical* map. Obviously σ is a homomorphism, or equivalently:



In other words

$$\sigma(\tau_i(x)) = [\tau_i(x)] = \tilde{\tau}_i(\sigma(x)).$$

and

$$\sigma(q_0) = [q_0].$$

Let

$$\mathcal{A} = \left(Q, q_0 \in Q, \{\tau_i : Q \to Q\}_{i \in \Sigma}\right)$$

and

$$\mathcal{B} = \left(Q', q_0{}' \in Q', \{\tau_i{}' : Q' \to Q'\}_{i \in \Sigma}\right)$$

be two pre-automata and $\sigma : \mathcal{A} \to \mathcal{B}$ be an onto homomorphism. Since σ maps $\mathcal{A}$ onto $\mathcal{B}$, the collection of all $\sigma-$inverse images of the elements of $\mathcal{B}$ forms a partiotion of $\mathcal{A}$. Let $\sim_\sigma$ be the equivalence relation corresponding to this partition. Since σ is a homomorphism, $\sim_\sigma$ is a congruence relation on $\mathcal{A}$:

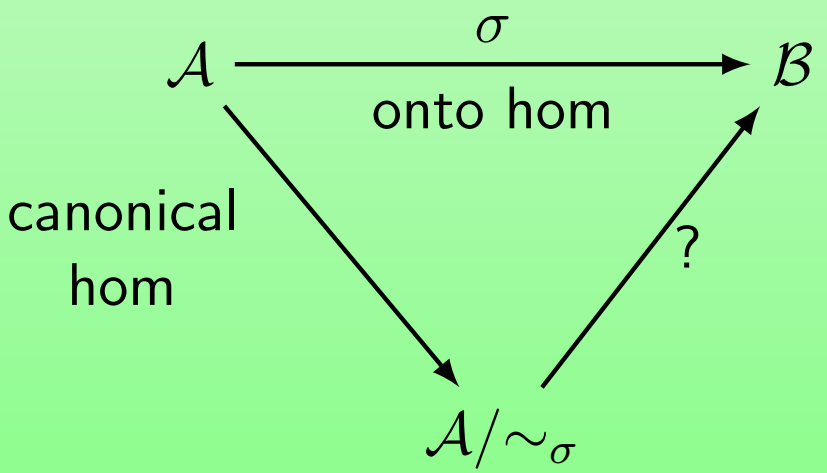
$$\begin{aligned}
[x]_{\sim_\sigma} = [y]_{\sim_\sigma} &\Leftrightarrow \sigma(x) = \sigma(y) \\
&\Leftrightarrow \tau_i{}'(\sigma(x)) = \tau_i{}'(\sigma(y)) \\
&\Leftrightarrow \sigma(\tau_i(x)) = \sigma(\tau_i(y)) \\
&\Leftrightarrow [\tau_i(x)]_{\sim_\sigma} = [\tau_i(y)]_{\sim_\sigma} \\
&\Leftrightarrow \tilde{\tau}_i([x]_{\sim_\sigma}) = \tilde{\tau}_i([x]_{\sim_\sigma}).
\end{aligned}$$

Hence, having an onto homomorphism $\sigma : \mathcal{A} \to \mathcal{B}$, one can make the quotient pre-automaton $\mathcal{A}$ by $\sim_\sigma$

$$\mathcal{A}/\sim_\sigma = \left(Q/\sim_\sigma,\ [q_0],\ \{\tilde{\tau}_i : Q/\sim_\sigma \ \to \ Q/\sim_\sigma\}_{i\in\Sigma}\right)$$

where $\tilde{\tau}_i([x]) = [\tau_i(x)]$.

Question:



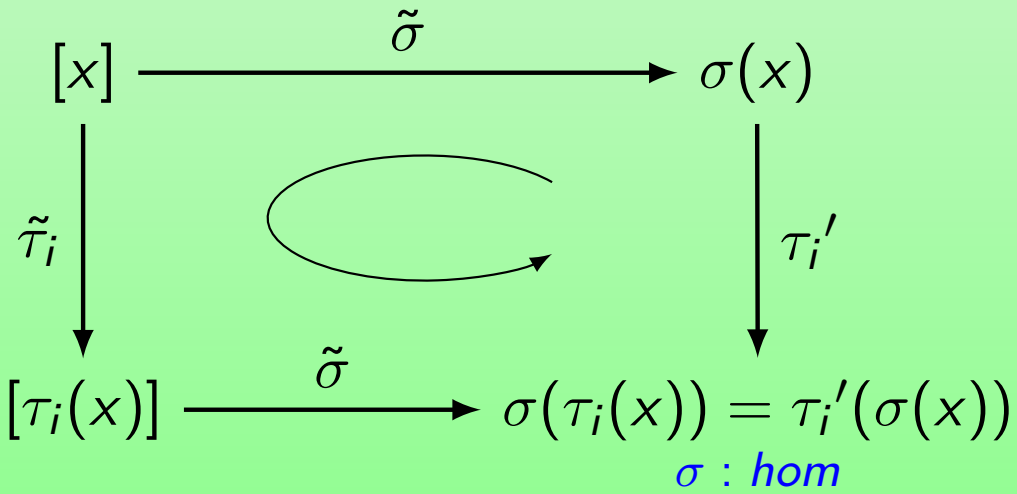
Consider the bijective map

$$\tilde{\sigma} : \mathcal{A}/\sim_\sigma \longrightarrow \mathcal{B}$$

$$[x] \mapsto \sigma(x)$$

$$[q_0] \mapsto \sigma(q_0) = {q_0}'$$

Then $\tilde{\sigma}$ is a homomorphism:



Consequently,

$$\mathcal{A}/\sim_\sigma \simeq \mathcal{B}.$$

Without loss of genelarity, assume that all pre-automata are defined on $\{0,1\}$.
So a pre-automaton is an algebra of type (0,1,1).

Consider the following fixed pre-automaton:

$$U_r = \left(\{0,1\}^*, \lambda \in \{0,1\}^*, \{r_i : \{0,1\}^* \to \{0,1\}^*\}_{i=0,1}\right)$$

where $r_i(w_1 \dots w_n) = w_1 \dots w_n i$.

Claim: For any reduced pre-automaton

$$\mathcal{B} = \left(P, p_0 \in P, \{\tau_i : P \to P\}_{i=0,1}\right)$$

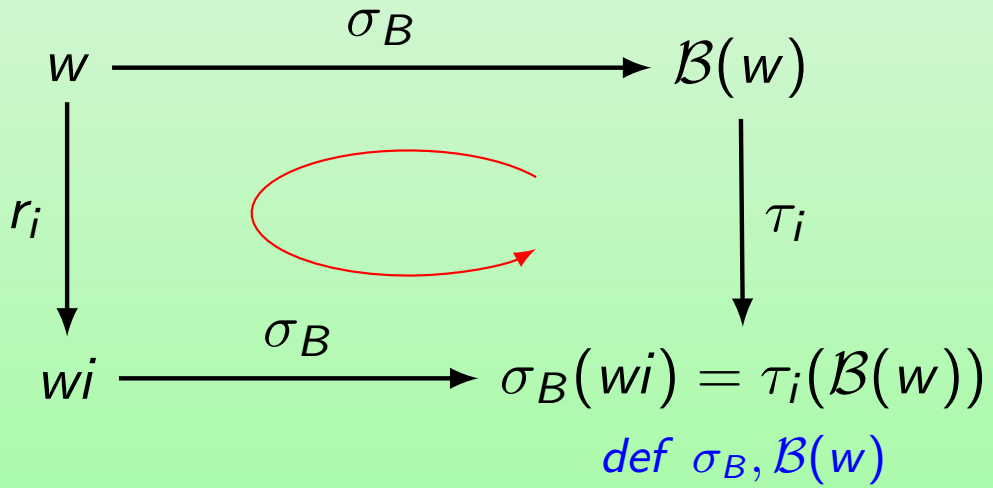
there exists an onto homomorphism $\sigma_B : U_r \xrightarrow{onto} \mathcal{B}$ such that $\sigma_B(w) = \mathcal{B}(w)$. In other words

$$\sigma_B(w_1 \dots w_n) = \tau_{w_n} \circ \dots \circ \tau_{w_2} \circ \tau_{w_1}(p_0).$$

Sketch of Proof.

$\mathcal{B}$: a reduced pre-automaton $\Rightarrow$ σ_B : an onto-map

and



Hence,

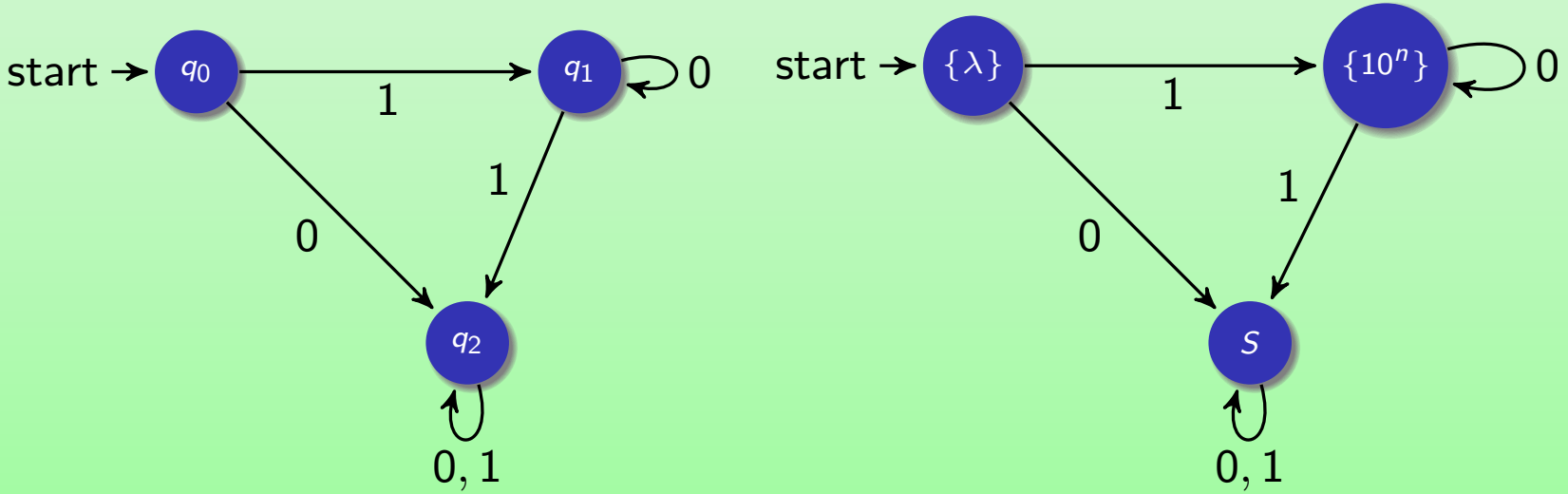
σ_B is a homomorphism.

Therefore,

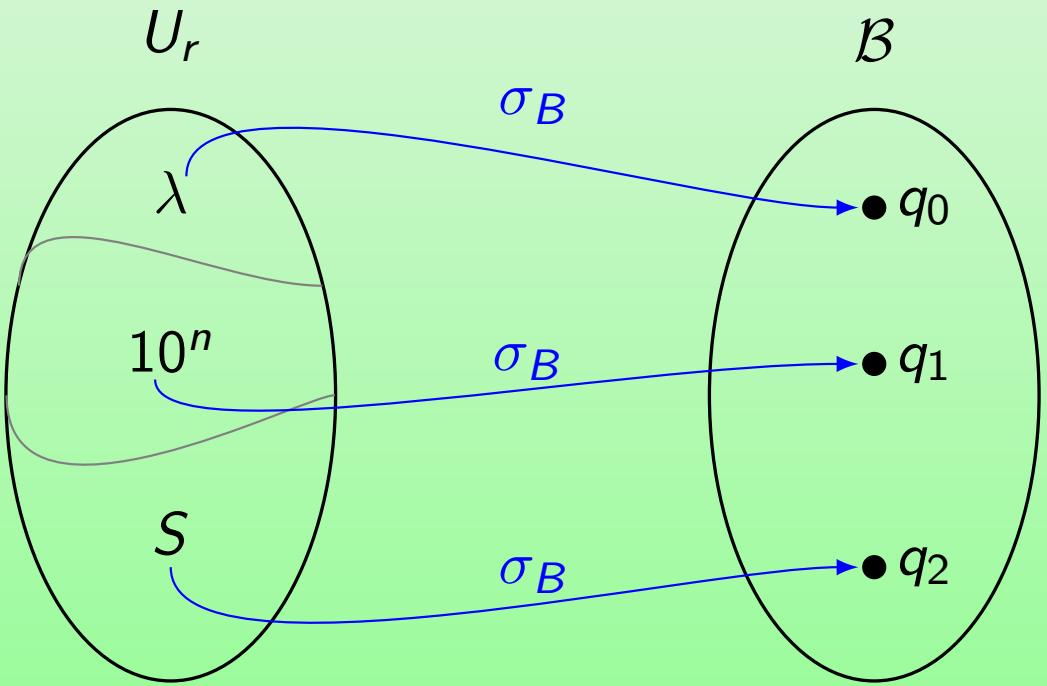
$$U_r/\sim_{\sigma_B} \simeq \mathcal{B}.$$

Example

The following two pre-automata are isomorphic:



where $S = \{0,1\}^* \setminus (\{\lambda\} \cup \{10^n \mid n \in \mathbb{N}\})$.


U_r
$\mathcal{B}$
σ_B
λ
q_0
10^n
σ_B
q_1
S
σ_B
q_2

The minimal automaton

Let $L \subseteq \{0,1\}^*$ be a language. We are trying to construct a minimal pre-automaton $\mathcal{M}_L$ to accept L. We will be needing some ingredients as follows:

- Definition: For a language $L \subseteq \{0,1\}^*$ and any $x \in \{0,1\}^*$, define:

$$x\backslash L = \{u \in \{0,1\}^* \mid xu \in L\}.$$

From this definition, it follows that for every $x, y \in \{0,1\}^*$:

$$\begin{aligned} x\backslash(y\backslash L) &= \{u \in \{0,1\}^* \mid xu \in y\backslash L\} \\ &= \{u \in \{0,1\}^* \mid yxu \in L\} \\ &= (yx)\backslash L. \end{aligned}$$

- Definition: For a language $L \subseteq \{0,1\}^*$, define a relation $\sim$ on $\{0,1\}^*$ as follows: for $x, y \in \{0,1\}^*$,

$$x \sim y \quad \Leftrightarrow \quad x\backslash L = y\backslash L$$
$$\Leftrightarrow \quad \{u \in \{0,1\}^* \mid xu \in L\} = \{u \in \{0,1\}^* \mid yu \in L\}.$$

In other words, $x \sim y$ if and only if for any $u \in \{0,1\}^*$, xu and yu are either both in L or both are not in L.

- Note: The relation $\sim$ is indeed an equivalence relation; It is easy to see that the relation $\sim$ is reflexive, symmetric, and transitive, because the equality relation has these properties.

Now we are ready to construct the minimal pre-automaton accepting L.

Claim: For a language $L \subseteq \{0,1\}^*$, $\mathcal{M}_L = (Q\ ,\ q_0\ ,\ \tau_0\ ,\ \tau_1)$ is the minimal pre-automaton w.r.t. right concatenation accepting L, where

1. $Q = \{x\backslash L \mid x \in \{0,1\}^*\}$,
2. $q_0 = \lambda\backslash L$,
3. $\tau_i(x\backslash L) = i\backslash(x\backslash L) = (xi)\backslash L, \qquad i = 0, 1.$

Example

$L = \{1\,0^n \mid n \in \mathbb{N}\}$

$\lambda \backslash L = L$

$1 \backslash L = \{0^n \mid n \in \mathbb{N}\} = Z$

$1\,0^j \backslash L = Z \qquad j \in \mathbb{N}$

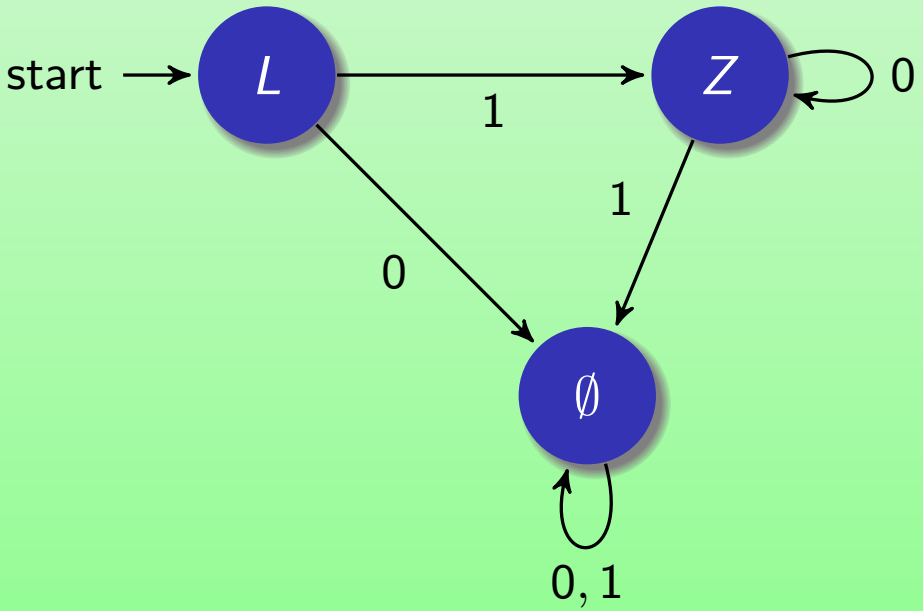
$0 \backslash L = \emptyset$

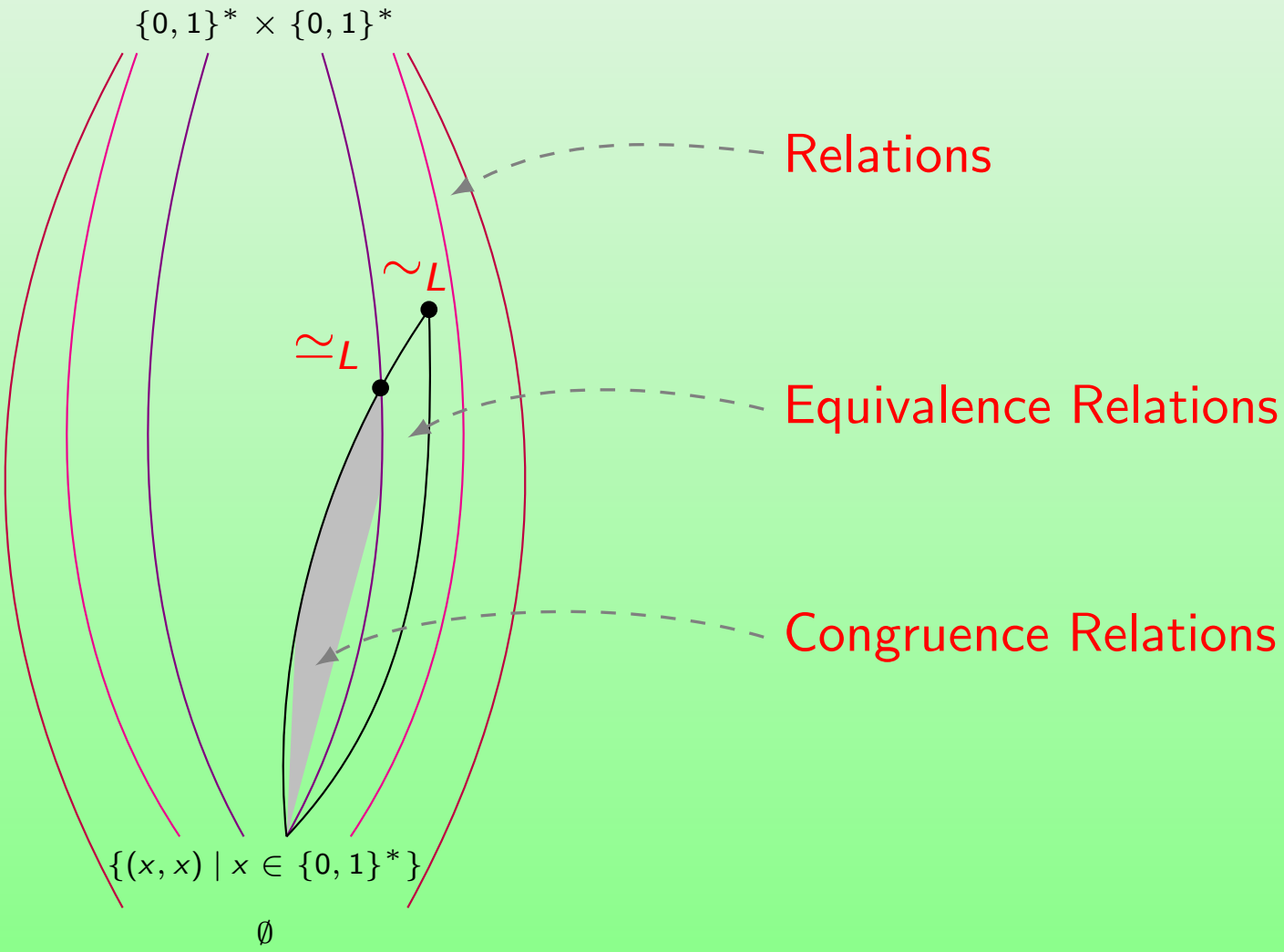
$0u \backslash L = \emptyset \qquad u \in \{0,1\}^*$

$11u \backslash L = \emptyset \qquad u \in \{0,1\}^*$

$101u \backslash L = \emptyset \qquad u \in \{0,1\}^*$

$\vdots$




$\{0,1\}^* \times \{0,1\}^*$
$\sim_L$
$\sim_L$
$\{(x,x) \mid x \in \{0,1\}^*\}$
$\emptyset$
Relations
Equivalence Relations
Congruence Relations

- A language $L \subseteq \{0,1\}^*$ determines an equivalence relation $\sim_L$ on $\{0,1\}^*$ in a usual way, i.e. for $x, y \in \{0,1\}^*$

$$x \sim_L y \;\Leftrightarrow\; x, y \in L \;\vee\; x, y \in L^c.$$

- Now we define a congruence relation $\simeq_L$ on $\{0,1\}^*$ as follows: for $x, y \in \{0,1\}^*$,

$$x \simeq_L y \;\Leftrightarrow\; x\backslash L = y\backslash L.$$

- Claim: $\simeq_L$ is the greatest congruence relation less than $\sim_L$.

Claim1: The relation $\simeq_L$ is a congruence relation on U_r.

Proof. It is easy to see that $\simeq_L$ is an equivalence relation on $\{0,1\}^*$, so it suffices to show that $\simeq_L$ is compatible with r_i, for i=0,1:

$$
\begin{aligned}
x \simeq_L y &\Leftrightarrow x\backslash L = y\backslash L \\
&\Leftrightarrow \forall u \in \{0,1\}^* \ (xu, yu \in L \ \vee \ xu, yu \notin L) \\
&\Leftrightarrow \forall u \in \{0,1\}^* \ \forall i \in \{0,1\} \ (xiu, yiu \in L \ \vee \ xiu, yiu \notin L) \\
&\Leftrightarrow xi\backslash L = yi\backslash L \\
&\Leftrightarrow xi \simeq_L yi \\
&\Leftrightarrow r_i(x) \simeq_L r_i(y).
\end{aligned}
$$

Claim2: $\simeq_L \subseteq \sim_L$.

Proof.

$$
\begin{aligned}
x \simeq_L y &\Rightarrow x\backslash L = y\backslash L \\
&\Rightarrow \forall u \in \{0,1\}^* \ (xu, yu \in L \ \vee \ xu, yu \notin L) \\
&\Rightarrow \text{for } \lambda \in \{0,1\}^* \ (x\lambda, y\lambda \in L \ \vee \ x\lambda, y\lambda \notin L) \\
&\Rightarrow x, y \in L \ \vee \ x, y \notin L \\
&\Rightarrow x, y \in L \ \vee \ x, y \in L^c \\
&\Rightarrow x \sim_L y.
\end{aligned}
$$

...

Claim3: For any congruence relation ρ on $\{0,1\}^*$,

$$\rho \subseteq \sim_L \Longrightarrow \rho \subseteq \simeq_L .$$

Proof.

$$\begin{aligned}
(x,y) \in \rho &\Rightarrow \forall i \in \{0,1\}\ (r_i(x), r_i(y)) \in \rho \\
&\Rightarrow \forall i \in \{0,1\}\ (xi, yi) \in \rho \\
&\Rightarrow \forall u \in \{0,1\}^*\ (xu, yu) \in \rho \\
&\Rightarrow \forall u \in \{0,1\}^*\ xu \sim_L yu \\
&\Rightarrow \forall u \in \{0,1\}^*\ (xu, yu \in L \ \vee \ xu, yu \in L^c) \\
&\Rightarrow \forall u \in \{0,1\}^*\ (xu, yu \in L \ \vee \ xu, yu \notin L) \\
&\Rightarrow x\backslash L = y\backslash L \\
&\Rightarrow x \simeq_L y.
\end{aligned}$$

Now we can form the quotient pre-automaton $U_r/\simeq_L$ and show that

$$U_r/\simeq_L \ \simeq \ \mathcal{M}_L.$$

Therefore, it follows from the construction of $U_r/\simeq_L$ that

$\mathcal{M}_L$ is the minimal pre-automaton w.r.t. right concatenation accepting L.

In order to show this isomorphism, consider the map $\sigma : U_r \to \mathcal{M}$ by $\sigma([x]_{\simeq_L}) = \mathcal{M}(x) = x\backslash L$ for any $x \in \{0,1\}^*$.

- σ is a well-defined map:

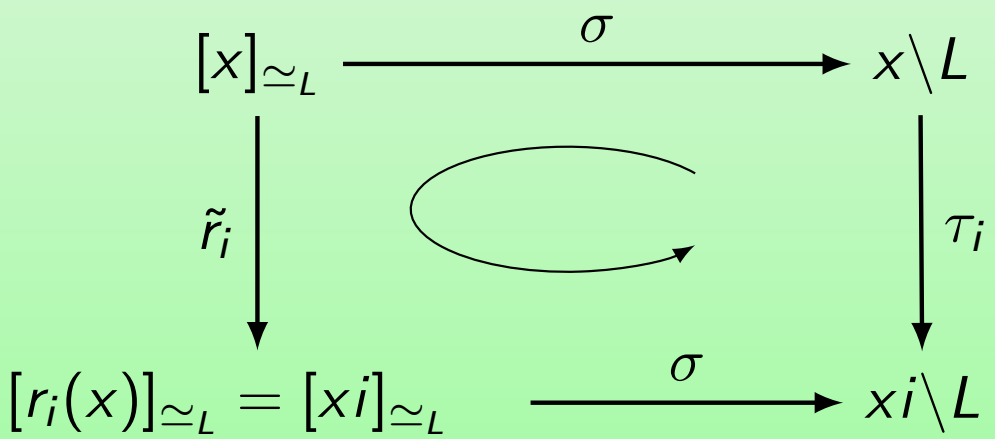
  1. For any $x = x_1 \dots x_n$:

$$\begin{aligned}
\mathcal{M}(x) &= \mathcal{M}(x_1 \dots x_n) \\
&= \tau_{x_n} \circ \dots \circ \tau_{x_2} \circ \tau_{x_1}(\lambda \backslash L) \\
&= \tau_{x_n} \circ \dots \circ \tau_{x_2}(x_1 \backslash L) \\
&= \tau_{x_n} \circ \dots \circ \tau_{x_3}(x_1 x_2 \backslash L) \\
&\vdots \\
&= x_1 \dots x_n \backslash L \\
&= x \backslash L.
\end{aligned}$$

  2. For any $x, y \in \{0,1\}^*$:

$$[x]_{\simeq_L} = [y]_{\simeq_L} \Leftrightarrow x \simeq_L y \Leftrightarrow x \backslash L = y \backslash L.$$

- σ is a homomorphism:



$$\tau_i \circ \sigma = \sigma \circ \tilde{r}_i.$$

- Obviously, σ is a bijective map.

In a similar way, we can treat L w.r.t. left concatenation as follows:

Definition: For a language $L \subseteq \{0,1\}^*$ and any $x \in \{0,1\}^*$, define:

$$L\backslash x = \{u \in \{0,1\}^* \mid ux \in L\}.$$

It follows from the definition that for every $x, y \in \{0,1\}^*$:

$$\begin{aligned} (L\backslash x)\backslash y &= \{u \in \{0,1\}^* \mid uy \in L\backslash x\} \\ &= \{u \in \{0,1\}^* \mid uyx \in L\} \\ &= L\backslash(yx). \end{aligned}$$

Hence, for a language $L \subseteq \{0,1\}^*$ the minimal pre-automaton w.r.t. left concatenation accepting L is $\mathcal{N}_L = (Q\ ,\ q_0\ ,\ \tau_0\ ,\ \tau_1)$, where

1. $Q_L = \{L\backslash x \mid x \in \{0,1\}^*\}$,
2. $q_0 = L\backslash\lambda$,
3. $\tau_i(L\backslash x) = (L\backslash x)\backslash i = L\backslash(ix)$. $\qquad i = 0, 1$.

Example

$L = \{1\,0^n \mid n \in \mathbb{N}\}$

$$L\backslash\lambda = L$$
$$L\backslash 0 = L$$
$$L\backslash 0^j = L \qquad j \in \mathbb{N}$$
$$L\backslash 1 = \{\lambda\}$$
$$L\backslash 1\,0^j = \{\lambda\} \qquad j \in \mathbb{N}$$
$$L\backslash u = \emptyset$$
$$u \neq \lambda,\ u \neq 0^j, 1\,0^j \qquad j \in \mathbb{N}$$