

We defined Regular Expressions, Regular Sets, DFA's, and NFA's, and proved:

Kleene's Theorem: *Let $A \subseteq \Sigma^*$ be any language. TFAE*

1. $A = \mathcal{L}(D)$, for some DFA D .
2. $A = \mathcal{L}(N)$, for some NFA N w/o ϵ transitions
3. $A = \mathcal{L}(N)$, for some NFA N .
4. $A = \mathcal{L}(e)$, for some regular expression e .

\forall *intro* To prove $\forall x(\varphi(x))$: let v be arbitrary, prove $\varphi(v)$.

\rightarrow *intro* To prove $\varphi \rightarrow \psi$: assume φ , prove ψ .

or, reasoning from the contrapositive, assume $\neg\psi$, prove $\neg\varphi$.

\wedge *elim* From $\varphi \wedge \psi$ may conclude φ , ψ .

\wedge *intro* From φ, ψ may conclude $\varphi \wedge \psi$.

\perp To prove φ : assume $\neg\varphi$, prove $A \wedge \neg A$.

induction To prove $\forall x \in \mathbf{N}(\varphi(x))$: prove **base case**: $\varphi(0)$, and,
inductive step: $\forall y(\varphi(y) \rightarrow \varphi(y + 1))$.

Let $A \subseteq \Sigma^*$ be any language.

Define the **right-equivalence relation** \sim_A on Σ^* :

$$x \sim_A y \iff \forall w \in \Sigma^* (xw \in A \leftrightarrow yw \in A)$$

$x \sim_A y$ iff x and y cannot be distinguished by concatenating some string w to the right of each of them and testing for membership in A .

Example: Let $\Sigma = \{a, b\}$ and $A_1 = \{w \in \Sigma^* \mid \#_b(w) \equiv 0 \pmod{2}\}$

$$\epsilon \sim_{A_1} a \sim_{A_1} aa; \quad b \sim_{A_1} ab \sim_{A_1} bbb$$

Claim: $x \sim_{A_1} y$ iff $\#_b(x) \equiv \#_b(y) \pmod{2}$.

Proof: Suppose $x \sim_{A_1} y$. Let $w = \epsilon$.

$$xw = x \in A_1 \quad \Leftrightarrow \quad yw = y \in A_1$$

Thus, $\#_b(x) \equiv \#_b(y) \pmod{2}$.

Conversely, suppose, $\#_b(x) \equiv \#_b(y) \pmod{2}$.

Let $w \in \Sigma^*$ be arbitrary.

$$\#_b(xw) \equiv \#_b(x) + \#_b(w) \equiv \#_b(y) + \#_b(w) \equiv \#_b(yw) \pmod{2}$$

$$xw \in A_1 \quad \Leftrightarrow \quad yw \in A_1$$

$$\forall w \in \Sigma^* (xw \in A_1 \quad \Leftrightarrow \quad yw \in A_1)$$

\forall -intro

Thus,

$$x \sim_{A_1} y.$$

□

$$A_1 = \{w \in \{a, b\}^* \mid \#_b(w) \equiv 0 \pmod{2}\}$$

$$x \sim_{A_1} y \Leftrightarrow \#_b(x) \equiv \#_b(y) \pmod{2}$$

$$[u]_{\sim_A} \stackrel{\text{def}}{=} \{w \in \Sigma^* \mid u \sim_A w\}$$

$$[a]_{\sim_{A_1}} = \{w \in \{a, b\}^* \mid \#_b(w) \equiv 0 \pmod{2}\}$$

$$[b]_{\sim_{A_1}} = \{w \in \{a, b\}^* \mid \#_b(w) \equiv 1 \pmod{2}\}$$

Proposition 2.1 For any language A , \sim_A is an equivalence relation. Recall that an equivalence relation is a binary relation that is reflexive, symmetric, and transitive.

Proof: Here I am giving detailed examples of the proof rules. This is more formal than you should be in your homework.

Reflexive: to show: $\forall x \in \Sigma^*(x \sim_A x)$

Let $x, w \in \Sigma^*$ be arbitrary ($xw \in A \leftrightarrow xw \in A$)

$\forall w \in \Sigma^*(xw \in A \leftrightarrow xw \in A)$ \forall intro

$x \sim_A x$

$\forall x \in \Sigma^*(x \sim_A x)$ \forall intro

Symmetric: to show: $\forall x, y \in \Sigma^*(x \sim_A y \rightarrow y \sim_A x)$

1. let $x, y, \in \Sigma^*$ be arbitrary

2. Suppose $x \sim_A y$.

3. $\forall w(xw \in A \leftrightarrow yw \in A)$

4. $\forall w(yw \in A \leftrightarrow xw \in A)$

5. $y \sim_A x$

6. $x \sim_A y \rightarrow y \sim_A x$ → intro 2-5

7. $\forall x, y \in \Sigma^*(x \sim_A y \rightarrow y \sim_A x)$ ∀ intro 1-6

Transitive: *to show:* $\forall x, y, z \in \Sigma^*(x \sim_A y \wedge y \sim_A z \rightarrow x \sim_A z)$

$x, y, z \in \Sigma^*$ *arb* 1. *Suppose* $x \sim_A y \wedge y \sim_A z$

2. $\forall w(xw \in A \leftrightarrow yw \in A)$ \wedge *elim 1, def of \sim_A*

3. $\forall w(yw \in A \leftrightarrow zw \in A)$ \wedge *elim 1, def of \sim_A*

$w \in \Sigma^*$ *arb* 4. $(xw \in A \leftrightarrow yw \in A)$ \forall *elim 2*

5. $(yw \in A \leftrightarrow zw \in A)$ \forall *elim 3*

6. $(xw \in A \leftrightarrow zw \in A)$

7. $\forall w \in \Sigma^*(xw \in A \leftrightarrow zw \in A)$ \forall *intro*

9. $x \sim_A z$

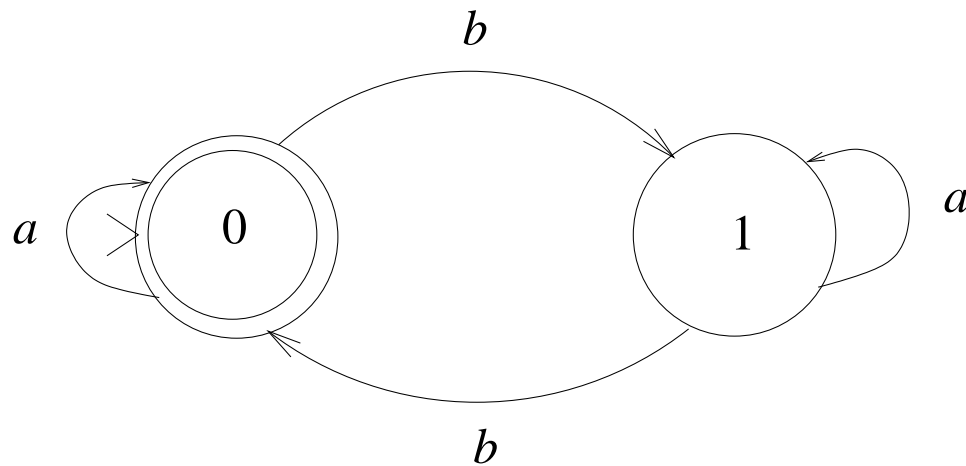
10. $x \sim_A y \wedge y \sim_A z \rightarrow x \sim_A z$ \rightarrow *intro 1-9*

11. $\forall x, y, z \in \Sigma^*(x \sim_A y \wedge y \sim_A z \rightarrow x \sim_A z)$ \forall *intro 1-10*

□

$$A_1 = \mathcal{L}(D_1) = \{w \in \{a, b\}^* \mid \#_b(w) \equiv 0 \pmod{2}\}$$

$$x \sim_{A_1} y \iff \#_b(x) \equiv \#_b(y) \pmod{2}$$



$$[a]_{\sim_{A_1}} = \{w \in \{a, b\}^* \mid \#_b(w) \equiv 0 \pmod{2}\}$$

$$[b]_{\sim_{A_1}} = \{w \in \{a, b\}^* \mid \#_b(w) \equiv 1 \pmod{2}\}$$

Myhill-Nerode Theorem: *The language A is regular iff \sim_A has a finite number of equivalence classes. Furthermore, this number of equivalence classes is equal to the number of states in the minimum-state DFA that accepts A .*

Proof: *Suppose $A = \mathcal{L}(D)$ for some DFA, $D = (\{q_1, q_2, \dots, q_n\}, \Sigma, \delta, q_1, F)$*

Let $S_i = \{w \mid \delta^(q_1, w) = q_i\}$*

Claim: *Each S_i contained in single \sim_A equivalence class.*

Let $x, y \in S_i, w \in \Sigma^$ be arbitrary.*

$$\delta^*(q_1, xw) = \delta^*(\delta^*(q_1, x), w) = \delta^*(\delta^*(q_1, y), w) = \delta^*(q_1, yw)$$

$$\mathcal{L}(D) = \{z \mid \delta^*(q_1, z) \in F\}$$

$$xw \in A \leftrightarrow \delta^*(q_1, xw) \in F \leftrightarrow \delta^*(q_1, yw) \in F \leftrightarrow yw \in A$$

$$\forall w (xw \in A \leftrightarrow yw \in A)$$

$$x \sim_A y$$

Thus, there are at most n equivalence classes!

Conversely, suppose that there are finitely many equivalence classes of \sim_A : E_1, \dots, E_m .

Let $[x]$ be the equivalence class that x is in.

Define $D = (\{E_1, \dots, E_m\}, \Sigma, \delta, [\epsilon], F)$ where

$$F = \{[x] \mid x \in A\}$$

$$\delta([x], a) = [xa]$$

Must show that δ is well defined, i.e.,

$$([x] = [y]) \quad \Rightarrow \quad ([xa] = [ya])$$

Suppose $x \sim_A y$.

$$\forall w (xw \in A \leftrightarrow yw \in A)$$

$$\forall w (xaw \in A \leftrightarrow yaw \in A)$$

Thus, $xa \sim_A ya$.

Claim: $\delta^*([\epsilon], x) = [x]$.

Proof: *by induction on $|x|$*

base case: $\delta^*([\epsilon], \epsilon) = [\epsilon]$

inductive case: *let x be arbitrary, with $|x| = r + 1$*

so $x = wa$ for some $a \in \Sigma, w \in \Sigma^r$

inductive hypothesis tells us $\delta^([\epsilon], w) = [w]$, thus,*

$$\delta^*([\epsilon], x) = \delta^*([\epsilon], wa) = \delta(\delta^*([\epsilon], w), a) = \delta([w], a) = [wa] = [x]$$

□

From the claim we see that,

$$x \in \mathcal{L}(D) \leftrightarrow \delta^*([\epsilon], x) \in F \leftrightarrow [x] \in F \leftrightarrow x \in A$$

Thus as desired, $\mathcal{L}(D) = A$.

□

Example: Prove that the following language is regular and its minimal DFA has seven states: $A_7 = \{w \in \{0, 1, \dots, 9\}^* \mid 7|w\}$.

Let $D_7 = (\{0, 1, \dots, 6\}, \Sigma, \delta_7, 0, \{0\})$; $\delta_7(q, d) = (10q + d) \bmod 7 = (3q + d) \bmod 7$

Show $\mathcal{L}(D_7) = A_7$ [exercise]; and, $\forall i \neq j \in \{0, 1, \dots, 6\} (i \not\sim_{A_7} j)$

Let $i \neq j \in \{0, 1, \dots, 6\}$ be arbitrary.

Pick d s.t. $3i + d \equiv 0 \pmod{7}$. **Suppose** $3j + d \equiv 0 \pmod{7}$.

$$3i + d \equiv 3j + d \pmod{7}$$

$$3i \equiv 3j \pmod{7}$$

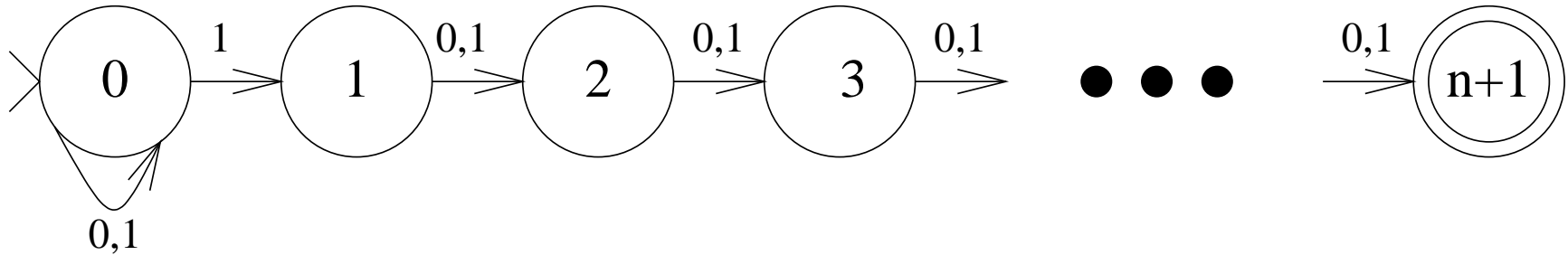
$$15i \equiv 15j \pmod{7}$$

$$i \equiv j \pmod{7}$$

⊥

Thus, $i \circ d \in A_7$, $j \circ d \notin A_7$, $i \not\sim_{A_7} j$.

Recall: $NFA, N_n, \text{ s.t. } \mathcal{L}(N_n) = \mathcal{L}((0 \cup 1)^*1(0 \cup 1)^n)$



Claim: *The minimal DFA that accepts $\mathcal{L}(N_n)$ has exactly 2^{n+1} states.*

Proof: $\sim_{\mathcal{L}(N_n)}$ has exactly the 2^{n+1} equivalence classes: $[w]$, for $w \in \{0, 1\}^{n+1}$.

These equivalence classes are all distinct: let $u, v \in \{0, 1\}^{n+1}$, with $u \neq v$.

Thus, u and v must differ at some position, i . WLOG say that $u = \alpha 0 \beta$ and $v = \alpha' 1 \beta'$ with $|\alpha| = |\alpha'| = i - 1$.

Thus, $u 1^{i-1} \notin \mathcal{L}(N_n)$ but $v 1^{i-1} \in \mathcal{L}(N_n)$ and thus as claimed, $u \not\sim_{\mathcal{L}(N_n)} v$.

These are all the equivalence classes: let $w \in \{0, 1\}^*$. If $|w| \geq n + 1$, let $T(w)$ be the last $n + 1$ characters of w . If $|w| < n + 1$, let $T(w)$ be $0^{n+1-|w|}w$. In either case it is easy to see that $w \sim_{\mathcal{L}(N_n)} T(w)$.

The claim thus follows from the Myhill-Nerode Theorem. □

A language **homomorphism** is a function $h : \Sigma^* \rightarrow \Gamma^*$ s.t.

$$\forall x, y \in \Sigma^* (h(xy) = h(x)h(y)) \quad (2.1)$$

Examples:

$$h : \{0, 1, 2, 3\}^* \rightarrow \{a, b\}^* \quad g : \{a, b\} \rightarrow \{a, b, c\}$$

$$h(0) = aa \quad g(a) = a$$

$$h(1) = b \quad g(b) = cbc$$

$$h(2) = aba$$

$$h(3) = \epsilon$$

$$h(012310) = aabababaa \quad g(baa) = cbcaa$$

Notation: for function $f : A \rightarrow B$, sets $S \subseteq A, T \subseteq B$,

$$f(S) = \{f(a) \mid a \in S\}; \quad f^{-1}(T) = \{a \in A \mid f(a) \in T\}$$

Example:

$$A_1 = \{w \in \{a, b\}^* \mid \#_b(w) \equiv 0 \pmod{2}\}$$

$$h^{-1}(A_1) = \{w \in \{0, 1, 2, 3\}^* \mid \#_1(w) + \#_2(w) \equiv 0 \pmod{2}\}$$

$$g(A_1) = \{w \in \{a, b, c\}^* \mid \#_{cbc} \equiv 0 \pmod{2}; \text{ no other } b \text{ or } c\}$$

Closure Theorem for Regular Sets: *Let $A, B \subseteq \Sigma^*$ be regular languages and let $h : \Sigma^* \rightarrow \Gamma^*$ and $g : \Gamma^* \rightarrow \Sigma^*$ be homomorphisms. Then the following languages are regular:*

- | | | |
|-------------------------------|---------------|----------------|
| 1. $A \cup B$ | 2. $A \cap B$ | 3. AB |
| 4. $\bar{A} = (\Sigma^* - A)$ | 5. $h(A)$ | 6. $g^{-1}(A)$ |

Proof: (1,3): Let $\mathcal{L}(e) = A$, $\mathcal{L}(f) = B$. Thus $\mathcal{L}(e \cup f) = A \cup B$; $\mathcal{L}(e \circ f) = AB$

(2): $A \cap B = \overline{\bar{A} \cup \bar{B}}$

(4): Let $\mathcal{L}(D) = A$, DFA $D = (Q, \Sigma, \delta, s, F)$. Let $\bar{D} = (Q, \Sigma, \delta, s, Q - F)$.

Thus $\mathcal{L}(\bar{D}) = \bar{A}$

(5) Let e be a regular expression for A . Then we can inductively define $H(e)$ and then prove by induction on e that $\mathcal{L}(H(e)) = h(A)$.

base cases: for $a \in \Sigma$, $H(a) = h(a)$; $H(\emptyset) = \emptyset$

inductive cases: $H(e \cup f) = H(e) \cup H(f)$; $H(e \circ f) = H(e) \circ H(f)$; $H(e^*) = (H(e))^*$

(6): Let $D = (Q, \Sigma, \delta, q_0, F)$ be a DFA accepting A . Define $D' = (Q, \Gamma, \delta', q_0, F)$ as follows: $\delta'(q, \gamma) = \delta^*(q, g(\gamma))$. Thus,

$w \in \mathcal{L}(D') \Leftrightarrow \delta'^*(q_0, w) \in F \Leftrightarrow \delta^*(q_0, g(w)) \in F \Leftrightarrow g(w) \in A \Leftrightarrow w \in g^{-1}(A) \square$

Example Use Of Closure Theorem for Regular Sets

Prop. *If $A \subseteq \Sigma^*$ is regular, then so is $A/\Sigma^* = \{w \in \Sigma^* \mid \exists u \in \Sigma^*(wu \in A)\}$.*

Proof: *Let $\Gamma = (\Sigma \cup \widehat{\Sigma})$ and define the homomorphisms $t, h : \Gamma^* \rightarrow \Sigma^*$ as follows:
 $t : \sigma \mapsto \sigma; t : \widehat{\sigma} \mapsto \sigma; d : \sigma \mapsto \sigma; d : \widehat{\sigma} \mapsto \epsilon$.*

Thus t takes hats off, and d deletes characters with hats on.

By the Closure Theorem for Regular Sets, the following sets are regular:

$$\begin{aligned}
 t^{-1}(A) &= \{w \in \Gamma^* \mid t(w) \in A\} \quad (\text{characters may have hats}) \\
 t^{-1}(A) \cap \Sigma^* \widehat{\Sigma}^* &= \{w \in \Gamma^* \mid t(w) \in A \text{ and all characters with hats are at end of } w\} \\
 d(t^{-1}(A) \cap \Sigma^* \widehat{\Sigma}^*) &= A/\Sigma^*
 \end{aligned}$$

Note: *there is also a simple proof of this proposition by modifying the DFA, D , accepting A . If $D = (Q, \Sigma, \delta, q_0, F)$, let $D' = (Q, \Sigma, \delta, q_0, F')$, where F' is the set of states in Q for which there is a path of length zero or more to a state in F . Then $A/\Sigma^* = \mathcal{L}(D')$. □*