Recall From Last Time

Definitions:

- An **alphabet** is a non-empty finite set, e.g., $\Sigma = \{0, 1\}$, etc.
- The set of **regular expressions** $R(\Sigma)$ over alphabet Σ.
- A language is regular iff it is denoted by some regular expression.
- A DFA is a tuple, $D = (Q, \Sigma, \delta, s, F)$.
- An NFA is a tuple, $N = (Q, \Sigma, \Delta, s, F)$.

Prop 1.2: Every NFA N can be translated into an NFA, N', which has the same number of states but no ϵ-transitions, s.t. $\mathcal{L}(N) = \mathcal{L}(N')$.

Proposition 1.3: For every NFA, N, with n states, there is a DFA, D, with at most 2^n states s.t. $\mathcal{L}(D) = \mathcal{L}(N)$.

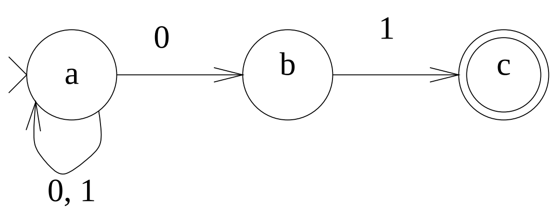
Proof: Let $N = (Q, \Sigma, \Delta, q_0, F)$. By Proposition 1.2 may assume that N has no ϵ transitions.

Let $D = (\wp(Q), \Sigma, \delta, \{q_0\}, F')$

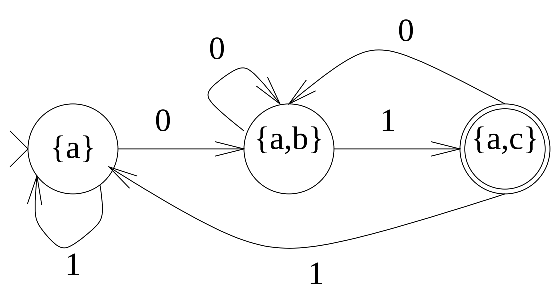
$$\delta(S, a) = \bigcup_{r \in S} \Delta(r, a)$$

$$F' = \{S \subseteq Q \mid S \cap F \neq \emptyset\}$$

N



D



Claim: For all $w \in \Sigma^\star$,

$$\delta^\star(\{q_0\}, w) \quad = \quad \Delta^\star(q_0, w)$$

By induction on $|w|$:

$|w| = 0$: $\delta^\star(\{q_0\}, \epsilon) \ = \ \{q_0\} \ = \ \Delta^\star(q_0, \epsilon)$

$|w| = k + 1$: $w = ua$.

Inductively, $\delta^\star(\{q_0\}, u) \ = \ \Delta^\star(q_0, u)$

$$\begin{aligned}
\delta^\star(\{q_0\}, ua) &= \delta(\delta^\star(\{q_0\}, u), a) \\
&= \bigcup_{r \in \delta^\star(\{q_0\}, u)} \Delta(r, a) \\
&= \bigcup_{r \in \Delta^\star(q_0, u)} \Delta(r, a) \\
&= \Delta^\star(q, ua)
\end{aligned}$$

Therefore, $\mathcal{L}(D) = \mathcal{L}(N)$.

Theorem 1.4 (Kleene's Th) Let $A \subseteq \Sigma^\star$ be any language. Then the following are equivalent:

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

Proof: Obvious that $1 \to 2 \to 3$.

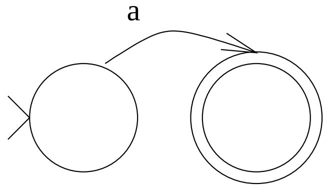
$3 \to 2$ by Prop. 1.2.

$2 \to 1$ by Prop. 1.3 (subset construction).

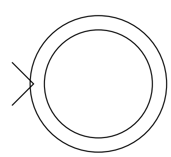
$4 \leftrightarrow 5$ by def of regular

$4 \to 3$: We show by induction on the number of symbols in the regular expression e, that there is an NFA N with $\mathcal{L}(e) = \mathcal{L}(N)$:

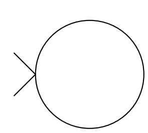
e* = *a



***e* =** ε

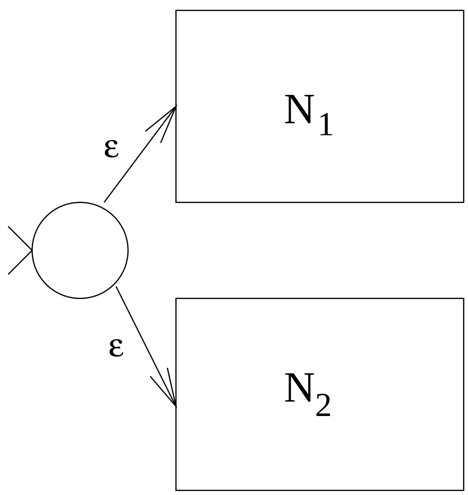


***e* =** $\emptyset$



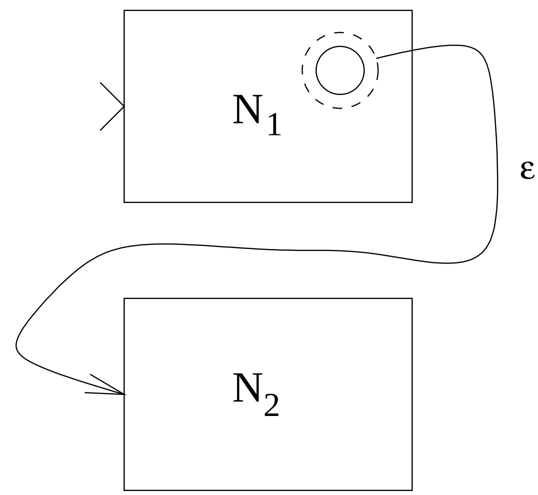
Union

$L(N) = L(N_1) + L(N_2)$



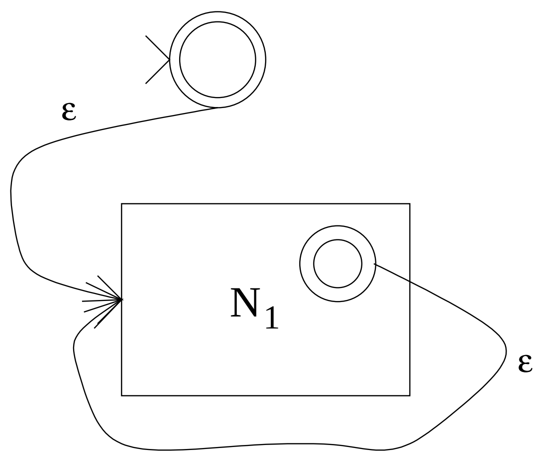
Concatenation

$L(N) = L(N_1)\,L(N_2)$



Kleene Star

$L(N) = (L(N_1))^*$



$3 \to 4$: Let $N = (\{1, \ldots, n\}, \Sigma, \Delta, 1, F)$, $F = \{f_1, \ldots, f_r\}$

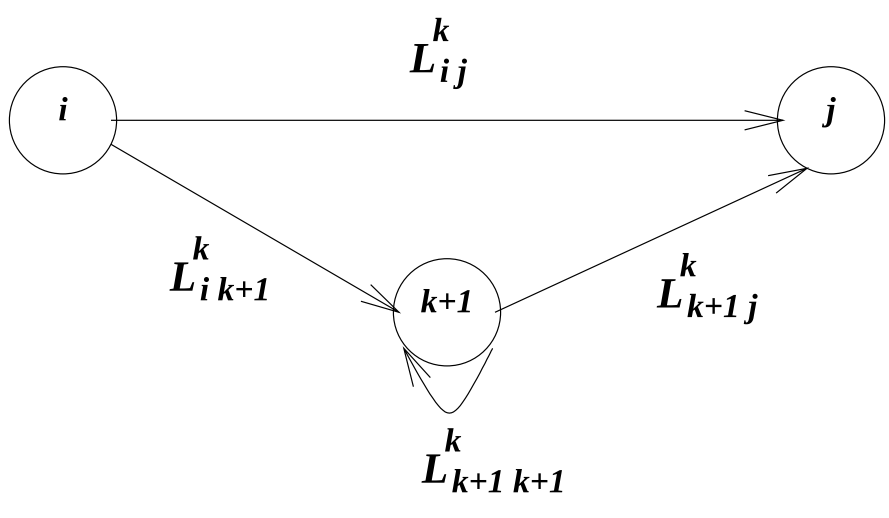
$$L_{ij}^{k} \equiv \{w \mid j \in \Delta^\star(i, w);\ \text{no intermediate state } \# > k\}$$

$$L_{ij}^{0} = \{a \mid j \in \Delta(i, a)\} \cup \{\epsilon \mid i = j\}$$

$$L_{ij}^{k+1} = L_{ij}^{k} \cup L_{i\,k+1}^{k}(L_{k+1\,k+1}^{k})^\star L_{k+1,j}^{k}$$

$$e = L_{1\,f_1}^{n} \cup \cdots \cup L_{1\,f_r}^{n}$$

$$\mathcal{L}(e) = \mathcal{L}(N)$$



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

Define the **right-equivalence relation** $\sim_A$ on $\Sigma^\star$:

$$x \sim_A y \quad \Leftrightarrow \quad (\forall w \in \Sigma^\star)(xw \in A \;\leftrightarrow\; yw \in A)$$

$x \sim_A y$ 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: $A_1 = \{w \in \{a,b\}^\star \mid \#_b(w) \equiv 0 \pmod 2\}$

$$\epsilon \sim_{A_1} a \sim_{A_1} aa \qquad b \sim ab \sim bbb$$

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

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

$$xw = x \in A_1 \qquad \leftrightarrow \qquad yw = y \in A_1$$

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

Suppose, $\#_b(x) \equiv \#_b(y) \pmod 2$.

$$(\forall w)\#_b(xw) \equiv \#_b(yw) \pmod 2 \ .$$

$$(\forall w)(xw \in A_1 \quad \leftrightarrow \quad yw \in A_1)$$

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

$$\begin{aligned} [u]_{\sim_A} &= \{w \in \Sigma^\star \mid u \sim_A w\} \\ [a] &= \{w \in \{a,b\}^\star \mid \#_b(w) \equiv 0 \pmod 2\} \\ [b] &= \{w \in \{a,b\}^\star \mid \#_b(w) \equiv 1 \pmod 2\} \end{aligned}$$

Exercise: Show that 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: Reflexive: $(\forall x \in \Sigma^\star)(x \sim_A x)$

Let $x, w \in \Sigma^\star$ be arbitrary.

$$(xw \in A \;\leftrightarrow\; xw \in A)$$

$(\forall w \in \Sigma^\star)(xw \in A \;\leftrightarrow\; xw \in A)$ because w was arbitrary.

$$x \sim_A x$$

$(\forall x \in \Sigma^\star)(x \sim_A x)$ because x was arbitrary.

Symmetric: $(\forall x, y \in \Sigma^\star)(x \sim_A y \rightarrow y \sim_A x)$

Let $x, y, \in \Sigma^\star$ be arbitrary.

Suppose $x \sim_A y$.

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

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

$$y \sim_A x$$

$$x \sim_A y \rightarrow y \sim_A x$$

$$(\forall x, y \in \Sigma^\star)(x \sim_A y \rightarrow y \sim_A x)$$

Transitive:

$$(\forall x, y, z \in \Sigma^\star)((x \sim_A y \wedge y \sim_A z) \rightarrow x \sim_A z)$$

Let $x, y, z \in \Sigma^\star$ be arbitrary.

Suppose $x \sim_A y \wedge y \sim_A z$.

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

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

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

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

$$(yw \in A \ \leftrightarrow \ zw \in A)$$

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

$(\forall w \in \Sigma^\star)(xw \in A \ \leftrightarrow \ zw \in A)$ because w was arbitrary.

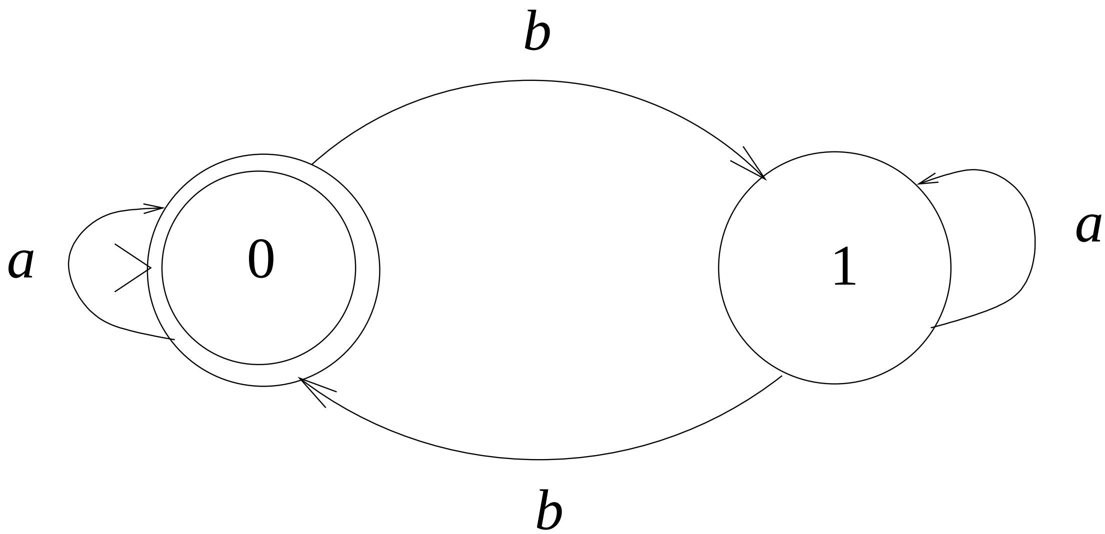
$$x \sim_A z$$

$$(x \sim_A y \wedge y \sim_A z) \rightarrow x \sim_A z$$

$(\forall x, y, z \in \Sigma^\star)(x \sim_A y \wedge y \sim_A z) \rightarrow x \sim_A z$ because x, y, z were arbitrary.

- To prove $(\forall x)\varphi$: let x be arbitrary, prove φ, conclude $(\forall x)\varphi$.
- To prove $\varphi \to \psi$: assume φ, prove ψ, conclude $\varphi \to \psi$.
- From $\varphi \wedge \psi$ may conclude φ, ψ.
- From φ, ψ may conclude $\varphi \wedge \psi$.
- To prove φ: assume $\neg\varphi$, prove $A \wedge \neg A$, conclude φ.

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



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,

$$D = (\{q_1, q_2, \ldots, q_n\}, \Sigma, \delta, q_1, F)$$

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

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

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

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

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

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

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

$$x \sim_A y$$

Thus, there are at most n equivalence classes!

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

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

Define $D = (\{E_1, \ldots, E_m\}, \Sigma, \delta, [\epsilon], F)$ 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: $\delta^\star([\epsilon], x) = [x]$.

Proof: by induction on $|x|$ [exercise].

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

Example: Prove that the following language is regular and its minimal DFA has seven states:

$$A_7 \;=\; \{w \in \{0,1,\ldots,9\}^\star \mid 7|w\}$$

$$D_7 \;=\; (\{0,1,\ldots,6\},\Sigma,\delta_7,0,\{0\})$$
$$\delta_7(q,d) \;=\; (10q+d)\text{mod } 7 \quad = \quad (3q+d)\text{mod } 7$$

Must show $\mathcal{L}(D_7) = A_7$ [exercise]; and,

$$(\forall i \neq j \in \{0,1,\ldots,6\})(i \not\sim_{A_7} j)$$

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

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

$$3i+d \;\equiv\; 3j+d \text{ (mod } 7)$$
$$3i \;\equiv\; 3j \text{ (mod } 7)$$
$$15i \;\equiv\; 15j \text{ (mod } 7)$$
$$i \;\equiv\; j \text{ (mod } 7)$$

$\Rightarrow\Leftarrow$

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

Example: Show $E = \{a^n b^n \mid n \in \mathbf{N}\}$ is not regular.

pf: Let $i \neq j \in \mathbf{N}$ be arbitrary.

We will show that $a^i \not\sim_E a^j$.

Let $w = b^i$

$$a^i w \in E; \qquad a^j w \notin E$$

Thus $\sim_E$ has infinitely many equivalence classes.

Thus by the Myhill-Nerode Theorem, E is not regular.

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

$$(\forall x, y \in \Sigma^\star)(h(xy) = h(x)h(y)) \qquad (2.0)$$

Examples:

$$\begin{array}{rcl} h : \{0,1,2,3\}^\star & \to & \{a,b\}^\star \\ h(0) = aa,\; h(1) = b, & & h(2) = aba,\; h(3) = \epsilon \\ h(012310) & = & aababab aa \end{array}$$

$$\begin{array}{rcl} g : \{a,b\} & \to & \{a,b,c\} \\ g(a) = a, & & g(b) = cbc \\ g(baa) & = & cbcaa \end{array}$$

Notation: for function $f : A \to B$, 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:

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

$$\begin{array}{rcl} h^{-1}(A_1) & = & \{w \in \{0,1,2,3\}^\star \mid \#_1(w) + \#_2(w) \equiv 0 \pmod 2\} \\ g(A_1) & = & \{w \in \{a,b,c\}^\star \mid \#_{cbc} \equiv 0 \pmod 2;\ \text{no other b or c}\} \end{array}$$

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

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

Proof: (1,2): Let $\mathcal{L}(e) = A$, $\mathcal{L}(f) = B$.

Thus $\mathcal{L}(e \cup f) = A \cup B$; $\mathcal{L}(e \circ f) = AB$

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

Let $\overline{D} = (Q, \Sigma, \delta, s, Q - F)$.

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

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

(5): Let $A = \mathcal{L}(e)$.

Thus $h(A) = \mathcal{L}(h(e))$.

Example:

$$
\begin{aligned}
g(a) &= a, \quad g(b) = cbc \\
A &= \mathcal{L}(a^\star(ba^\star ba^\star)^\star) \\
g(A) &= \mathcal{L}(a^\star(cbca^\star cbca^\star)^\star)
\end{aligned}
$$

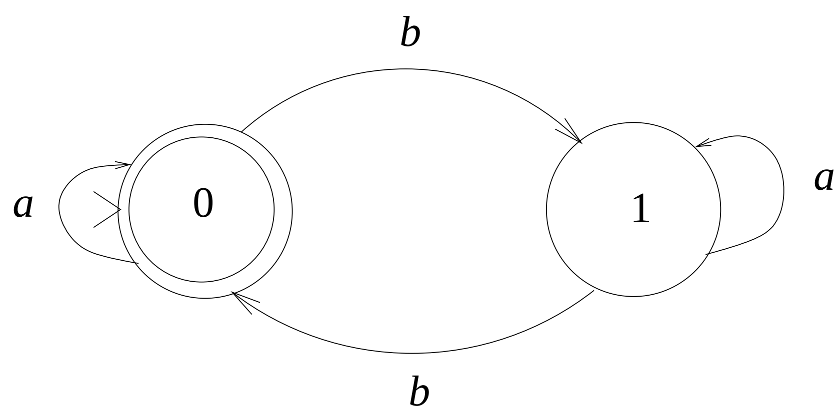
(6): Let $A = \mathcal{L}(D)$, DFA, $D = (Q, \Sigma, \delta, s, F)$.
Let $D' = (Q, \Gamma, \delta', s, F)$.

$$\delta'(q, \gamma) \quad = \quad \delta^{\star}(q, h(\gamma))$$

Example:

$$h(0) = aa, \quad h(1) = b, \quad h(2) = aba, \quad h(3) = \epsilon$$

D



D'