Lecture 4

Finite Automata and Regular Languages

- Simplest computational model
- Limited computational power
- Applications: Compiler construction, program analysis. Systems biology.

Definition: A finite automaton (FA) is a 5-tuple $M = (Q, \Sigma, \delta, q_0, F)$ where

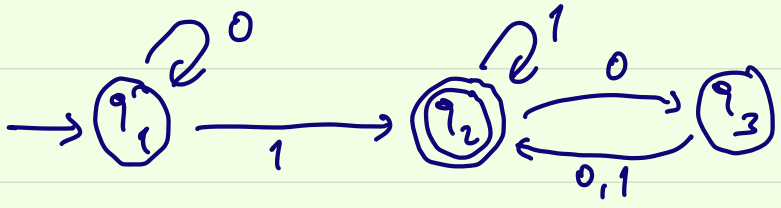
- Q is a set of states
- Σ is the input alphabet
- δ is a transition function:

$$\delta : Q \times \Sigma \longrightarrow Q \qquad \text{e.g. } \delta(q, a) = q'$$

- q_0 is the initial state (i.e. $q_0 \in Q$)
- $F \subseteq Q$ is the set of final/accepting states.

Example: $\Sigma = \{0, 1\}$

M:



q_1 : the initial state

q_2 : the final state

$Q = \{q_1, q_2, q_3\}$

$F = \{q_2\}$

$\delta(q_1, 0) = q_1$

$\delta(q_1, 1) = q_2$

$\delta(q_3, 0) = q_2$

$\delta(q_3, 1) = q_2$

δ	0	1
q_1	q_1	q_2
q_2	q_3	q_2
q_3	q_2	q_2

"deterministic"

e.g. $w = 001$ ... accepted

$q_1 \xrightarrow{0} q_1 \xrightarrow{0} q_1 \xrightarrow{1} q_2$

$w_2 = 010$ ... rejected

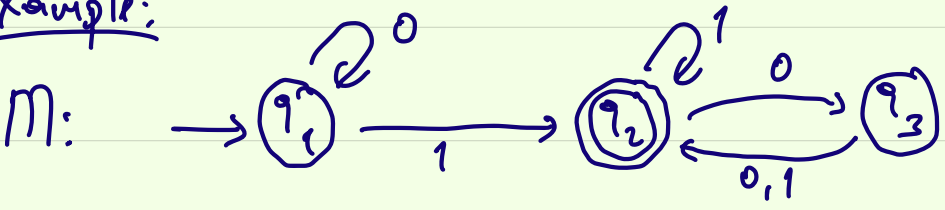
$q_1 \xrightarrow{0} q_1 \xrightarrow{1} q_2 \xrightarrow{0} q_3$

$w_3 = \lambda$ ... rejected.

Definition: A string $w \in \Sigma^*$ is **accepted** by a FA M if M starts with the initial state and enters a final state $q_f \in F$ after processing w.

Example:

M:



Q: Set of all strings accepted by M?

called $L(M)$

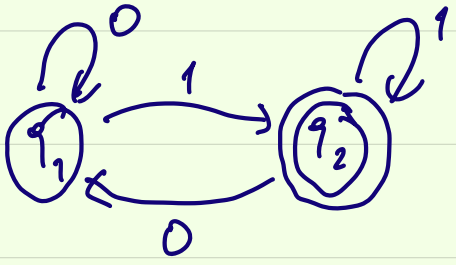
Rejected	Accepted	
λ	1	0100
0	01	01000
00	0101	010000
		0111000
		1110000

$L(M)$ = "set of all strings that contain even number of 0s after the last occurrence of 1"

Definition: For an FA M, $L(M)$ denotes the language accepted by M.

$$L(M) = \{ w \in \Sigma^* \mid M \text{ accepts } w \}$$

Example:



M_2

$L(M_2) = ?$

$\lambda \in L(M_2)$? No

$1 \in L(M_2)$? Yes

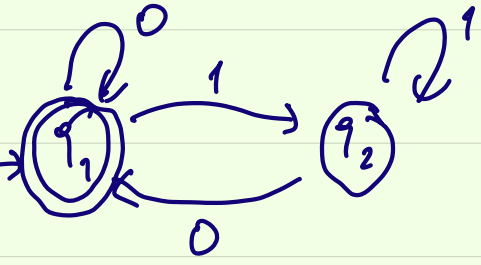
$0 \in L(M_2)$? No

$0 \ldots 0 \in L(M_2)$? No

$000\ldots01 \in L(M_2)$? Yes.

$$L(M_2) = \{ w \in \Sigma^* \mid w = w'1 \text{ , } w' \in \Sigma^* \}$$

Example:



M_3

$$L(M_3) = \{\lambda\} \cup \{ w \in \Sigma^* \mid w = w'0 \text{ , } w' \in \Sigma^* \}$$

Exercise:

- FA for only $\{w \in \Sigma^* \mid w = w'0, w' \in \Sigma^*\}$?
- FA for the set of strings that start and end with the same string. e.g. $w = 0 \ldots 0$ or $w = 1 \ldots 0$ or

Formal definition of Computation of an FA:

Given $M = (Q, \Sigma, \delta, q_0, F)$ and an input string $w = w_1 w_2 \ldots w_n$

"M accepts w if there exists a sequence of states $r_0, r_1, \ldots, r_n \in Q$ s.t.

- r_0 is the initial state $q_0 \in Q$
- $r_n \in F$
- $\delta(r_i, w_{i+1}) = r_{i+1}$ for all $i = 0, \ldots, n-1$

$$r_0 \xrightarrow{w_1} r_1 \xrightarrow{w_2} r_2 \cdots \to r_i \xrightarrow{w_{i+1}} r_{i+1} \cdots$$

Def: A language $L \subseteq \Sigma^*$ is "regular" if $L = L(M)$ for some FA M.