LECTURE 1 Basics on Formal Languages and Automata

$\Sigma = \{a, b, \dots\}$ a finite alphabet

$l, l' \in \Sigma$

DEF A word $u = l_1 l_2 \dots l_{|u|}$ is a finite sequence over Σ

$u \in \Sigma^*$

Empty word is ε

$u, v \in \Sigma^*$ $\quad u \leq v$ if u is a prefixe of v

Classification of formal languages

(Chomsky Hierarchy)

Focus on regular languages

DEF A finite-state automaton over Σ

is a machine $M = (Q, Q_0, \delta, F)$

where

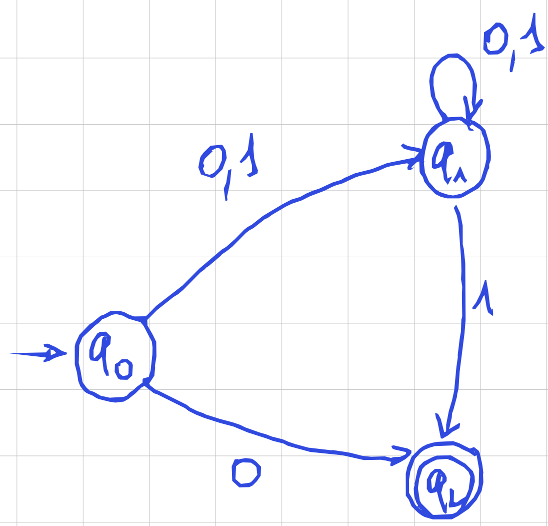
Q is a finite set of states, $Q_0 \subseteq Q$ initial states
$F \subseteq Q$ final states

$\delta \subseteq Q \times \Sigma \times Q$ transition relation

$$q \xrightarrow{a} q' \in \delta$$

Focus on regular languages

DEF A finite-state automaton over Σ is a machine $M = (Q, Q_0, \delta, F)$



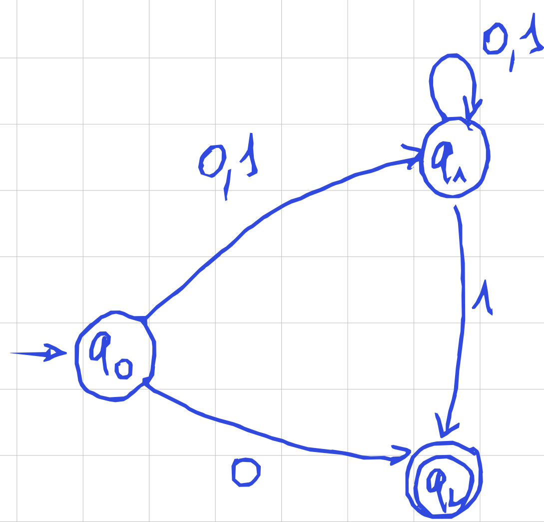
$\Sigma = \{0, 1\}$

$Q_0 = \{q_0\}$

$F = \{q_2\}$

Focus on regular languages

DEF A finite-state automaton over Σ is a machine $M = (Q, Q_0, \delta, F)$



DEF A run of M over word $u = l_1 l_2 \dots l_n$ is a sequence

$$q_0 \xrightarrow{l_1} q_1 \xrightarrow{l_2} \dots \xrightarrow{l_n} q_n$$

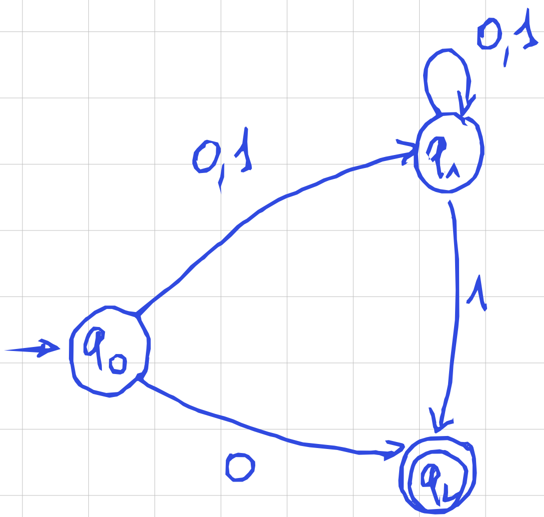
with $q_{i-1} \xrightarrow{l_i} q_i \in \delta$

Accepting if $q_n \in F$

Focus on regular languages

DEF A finite-state automaton over Σ

is a machine $M = (Q, Q_0, \delta, F)$

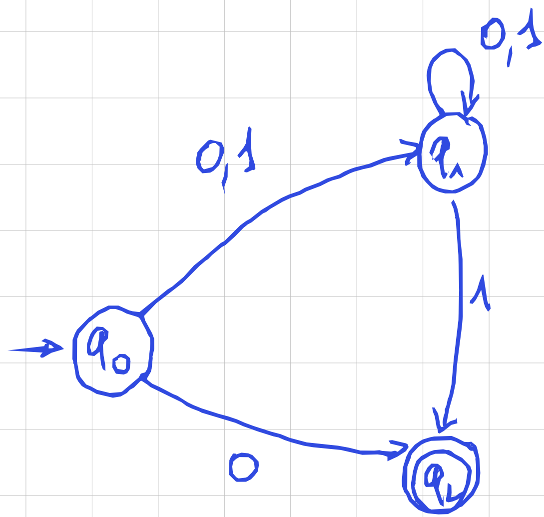


DEF $\mathcal{L}(M) = \{ u \in \Sigma^* \mid$ there is an accepting run over $u \}$

Focus on regular languages

DEF A finite-state automaton over Σ

is a machine $M = (Q, Q_0, \delta, F)$



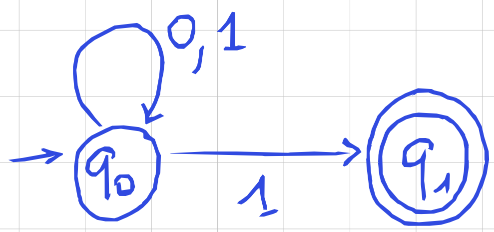
DEF $\mathcal{L}(M) = \{ u \in \Sigma^* \mid$ there is an accepting run over $u \}$

$\mathcal{L}(M) = \{0\} \cup \{ u \in \Sigma^* \mid$ the last letter of u is a $1 \}$

Non-determinism of finite-state automata

Another example: An automaton for the language

$$\{ u \in \{0,1\}^* \mid u \text{ ends with a } 1 \}$$



$$\delta \subseteq Q \times \Sigma \times Q$$

$$\Psi$$

$$(q_0, 1, q_0)$$

$$(q_0, 1, q_1)$$

Exercices

(1) Define M over $\Sigma = \{0, 1\}$ such that

$\mathcal{L}(M) = \{ u \in \Sigma^* \mid u$ has an even number of letter $0 \}$

(2) Define M over $\Sigma = \{0, 1\}$ such that

$\mathcal{L}(M) = \{ u \in \Sigma^* \mid u$ has an odd number of 0 and every 1 is immediately followed by a $0 \}$

DEF Regular language

$L \subseteq \Sigma^*$ is regular if there exists a finite state automaton M s.t. $\mathcal{L}(M) = L$.

PROP (pumping lemma)

For every regular language there exists $m \in \mathbb{N}$ such that for every word $w v w' \in \Sigma^*$ with $|v| \geq m$, there exists a factorisation $v_0 v_1 v_2$ of v with $v_1 \neq \varepsilon$ such that

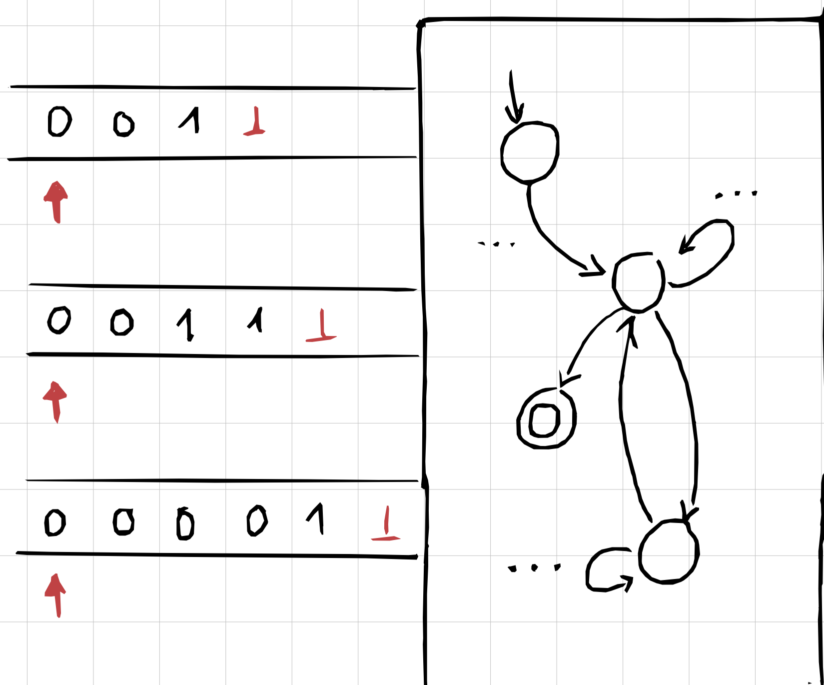
$$w v w' \in L \text{ iff } w v_0 v_1^k v_2 w' \in L, \text{ for all } k$$

PROP The class of regular languages is closed under union and complementation

Proof We use finite-state automata

on the board.

n-tape finite-state automata



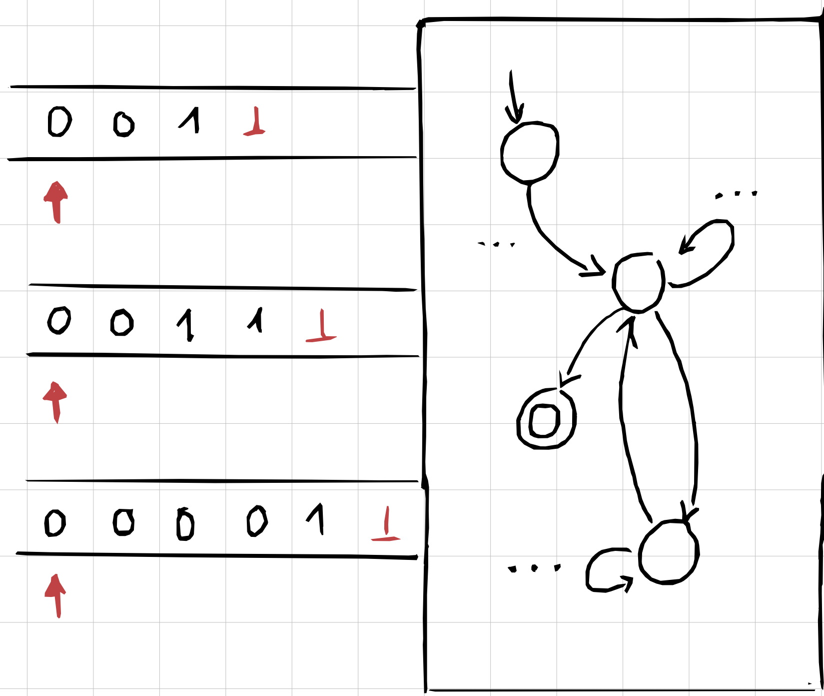
States ...

Head on each tape

A blank symbol

$\perp \notin \Sigma$

n-tape finite-state automata



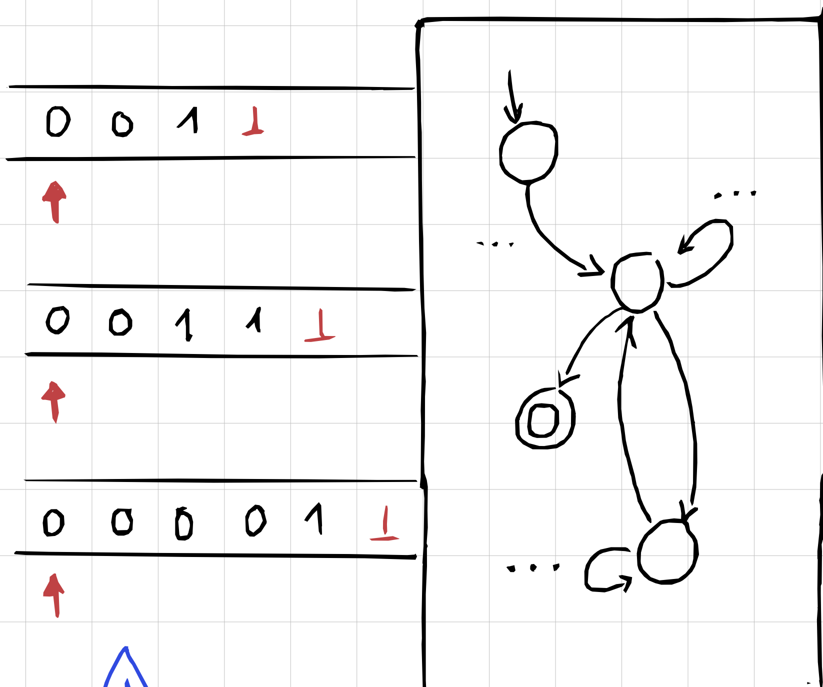
States ...

Head on each tape

A blank symbol ⊥

Heads move only to the right

m-tape finite-state automata



States ...

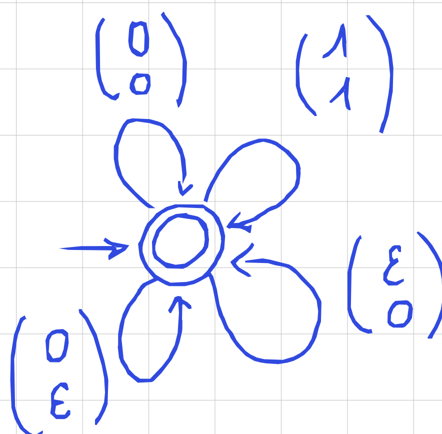
Head on each tape

A blank symbol ⊥

Heads move only to the right

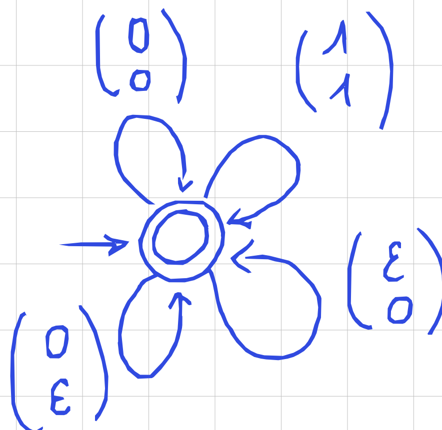
⚠ In general, heads may not move synchronously

Example A two-tape automaton that accepts $(u, v) \in \{0,1\}^* \times \{0,1\}^*$ whenever $\mathrm{proj}_1(u) = \mathrm{proj}_1(v)$



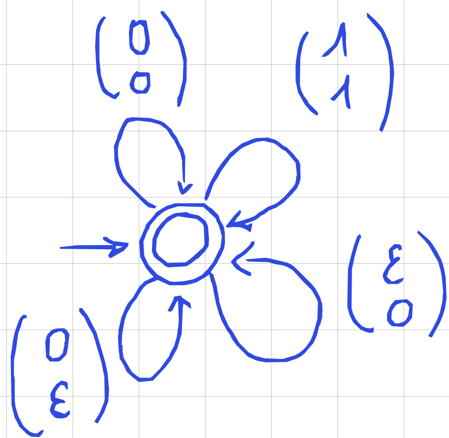
where $\mathrm{proj}_1(u)$ is obtain by removing 0's in u.

Example A two-tape automaton that accepts $(u, v) \in \{0, 1\}^* \times \{0, 1\}^*$ whenever $|u|_1 = |v|_1$



ε means head does not move

Example



ε means head does not move

0 0 1 0 1 ⊥ $\to$ 0 0 1 0 1 ⊥ $\to$ 0 0 1 0 1 ⊥

1 0 0 1 ⊥ $\quad$ 1 0 0 1 ⊥ $\quad$ 1 0 0 1 ⊥

$\to$ 0 0 1 0 1 ⊥ $\to$ 0 0 1 0 1 ⊥ $\quad$ ...

1 0 0 1 ⊥ $\quad$ 1 0 0 1 ⊥

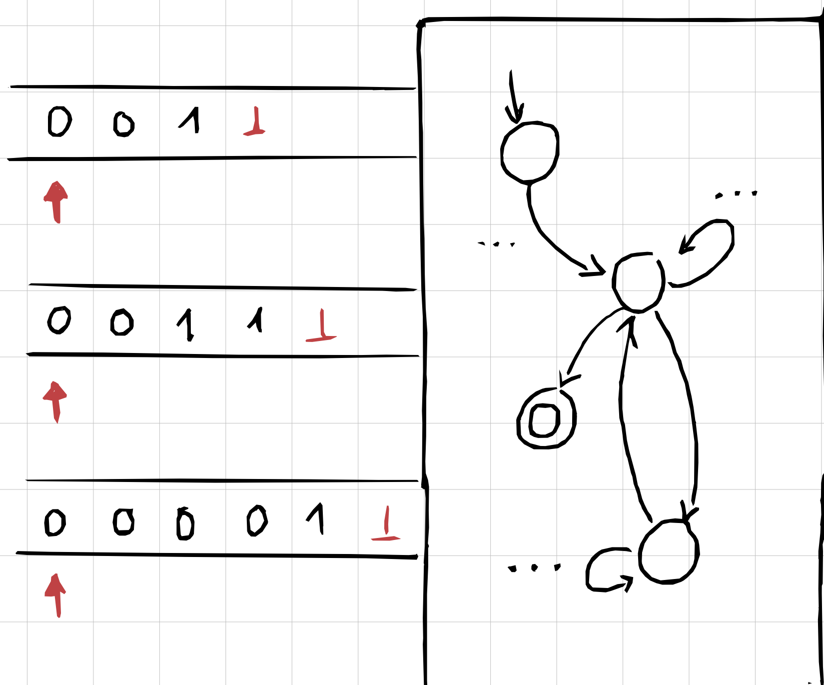
m-tape automata characterize m-ary rational relations over Σ^*

THM Rational relations are closed under union, but not under intersection (and complementation)

Proof see Berstel 1979 "Transductions and context-free languages"

n-tape finite-state automata

Synchronous moves of heads



THM Rational relations are closed under union, but not under intersection (and complementation)

THM Regular relations are closed under union and complementation

Zoo on regular relations

Since heads move synchronously, n-tuples of words over Σ can be seen as a single word on alphabet $\underbrace{(\Sigma \cup \{\perp\})^n}_{\text{write it } \Sigma_\perp}$ $(\perp \notin \Sigma)$

Convolution of triple of words

a a b

a

a a

Convolution of triples of words

$$\begin{pmatrix} a \\ a \\ a \end{pmatrix} \begin{pmatrix} a \\ \perp \\ a \end{pmatrix} \begin{pmatrix} b \\ \perp \\ \perp \end{pmatrix} = \otimes(aab, a, aa)$$

is a word over alphabet $(\Sigma_\perp)^3$

Convolution n-tuples of words

DEF The convolution of n words

$$(u_1, \ldots, u_n) \in (\Sigma^*)^n \text{ is}$$

the word $\otimes(u_1, \ldots, u_n)$ over $(\Sigma_\perp)^n$

of length $\max_i |u_i|$, and k-th letter is

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \text{ with } a_i = \begin{cases} \text{the k-letter of } u_i & \text{if } k < |u_i| \\ \perp & \text{otherwise} \end{cases}$$

DEF The convolution of a relation $R \subseteq (\Sigma^*)^m$ is

$$\otimes R = \{ \otimes(u_1, \ldots, u_m) \mid (u_1, \ldots, u_m) \in R \}$$

DEF A relation $R \subseteq (\Sigma^*)^m$ is regular (or synchronous rational) if $\otimes R$ is a regular language

Exercices Take a finite alphabet Σ

Show that the following relations are regular

1. The prefix relation $\leqslant$ over Σ^*
2. The equal length relation over Σ^*
3. The longest common prefix relation $u \sqcap v = w$
4. Take a total order $<$ over Σ : $\leqslant_{lex}$

- The lexicographic order over Σ^*
- The length-lexicographic order over Σ^*

$u \leqslant_{llex} v$ if $|u| \leqslant |v|$ or $|u| = |v|$ and $u \leqslant_{lex} v$

THM Let $R, S \subseteq \Sigma^{*m}$ be regular relations. The following relations are also regular.

1. $R \cup S$, $R \cap S$, $R \setminus S$

2. Projection $\{\bar{v} \mid$ there is $u \in \Sigma^*$ with $(u, \bar{v}) \in R\}$

3. Instanciation $\{\bar{v} \mid (u, \bar{v}) \in R\}$ with fixed $u \in \Sigma^*$

4. Cylindrification $\{(u, \bar{v}) \mid u \in \Sigma^*$ and $\bar{v} \in R\}$

5. Permutation of coordinates of R

Moreover, there is an effective way that given automata for $\otimes R$ and $\otimes S$ constructs an automaton for the convolution of 1.-5.