CSE206 Homework : Language of a non-deterministic automaton

quentin.guilmant@lix.polytechnique.fr

March 2020

Abstract

The purpose of this homework is to give another method to compute the language of non-deterministic automaton using linear algebra methods. We first prove a useful theorem that ensures existence and uniqueness of the solution, then we explicit it. You can answer either in English or in French.

We fix a finite alphabet Σ for this homework.

1 Reminders and notations

Notation. We let ***Reg*** be the set of the regular languages over Σ. If $A, B \in$ ***Reg***, AB is the concatenation of languages and $A + B$ is the union of regular languages.

The notation is in accordance with the fact that these operations behave as the usual operations over numbers.

Remark 1.1. $\varnothing$ is the neutral element of $+$ and $\{\varepsilon\}$ is the neutral for the concatenation.

Definition 1.2. Let $m, n \in \mathbb{N}$, we denote $\mathcal{M}_{m,n}$ the matrix with m lines and n columns such that the coefficients are regular languages. We also denote $\mathcal{M}_n = \mathcal{M}_{n,n}$ the set of square matrices.

Definition 1.3. Addition and multiplication extend as usual to matrices. If $A, B \in \mathcal{M}_{m,n}$, with $A = (a_{i,j})_{(i,j)\in[\![1\ ;\ m]\!]\times[\![1\ ;\ n]\!]}$ and $B = (b_{i,j})_{(i,j)\in[\![1\ ;\ m]\!]\times[\![1\ ;\ n]\!]}$, then

$$A + B = (a_{i,j} + b_{i,j})_{(i,j)\in[\![1\ ;\ m]\!]\times[\![1\ ;\ n]\!]}$$

In a similar way, if $A \in \mathcal{M}_{m,n}$ with $A = (a_{i,j})_{(i,j)\in[\![1\ ;\ m]\!]\times[\![1\ ;\ n]\!]}$, and $B \in \mathcal{M}_{n,p}$ with $B = (b_{i,j})_{(i,j)\in[\![1\ ;\ n]\!]\times[\![1\ ;\ p]\!]}$, then

$$AB = (c_{i,j})_{(i,j)\in[\![1\ ;\ m]\!]\times[\![1\ ;\ p]\!]}$$

where

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Definition 1.4. If $A, B \in \mathcal{M}_{m,n}$, with $A = (a_{i,j})_{(i,j)\in[\![1\ ;\ m]\!]\times[\![1\ ;\ n]\!]}$ and $B = (b_{i,j})_{(i,j)\in[\![1\ ;\ m]\!]\times[\![1\ ;\ n]\!]}$, then we write

$$A \subseteq B \qquad \textit{iff} \qquad \forall i,j \quad a_{i,j} \subseteq b_{i,j}$$

Question 1. What is the identity matrix $I_n \in \mathcal{M}_n$ such that if $A \in \mathcal{M}_n$, $I_nA = AI_n = A$? What is the zero-matrix $0_n \in \mathcal{M}_n$ such that if $A \in \mathcal{M}_n$, $A + O_n = O_n + A = A$?

Definition 1.5. For $A \in \mathcal{M}_n(\textbf{\textit{Reg}})$, we use the convention $A^0 = I_n$.

2 A useful lemma

In this section we prove the following theorem :

Theorem 2.1. *Let $A \in \mathcal{M}_n$ and $B \in \mathcal{M}_{n,1}$ with $A = (a_{i,j})_{i,j\in[\![1\ ;\ n]\!]}$ and $B = (b_i)^T_{i\in[\![1\ ;\ n]\!]}$. Let $X = AX + B$ be a matricial linear equation over regular languages. Then the smallest (for inclusion) solution to this equation is*

$$A^*B = \sum_{k\in\mathbb{N}} A^kB = \bigcup_{k\in\mathbb{N}} A^kB$$

This solution is a column-vector of regular languages. Finally if for all $i, j \in [\![1\ ;\ n]\!]$, $\varepsilon \notin a_{i,j}$ then it is the unique solution.

In general we use this theorem in the 1-dimensional case :

Corollary 2.2. *If A and B are regular languages then A^*B is the smallest solution to $X = AX + B$ and is $\varepsilon \notin A$ then it is the unique condition.*

2.1 A minimum solution

We now fix $A \in \mathcal{M}_n$ and $B \in \mathcal{M}_{n,1}$ with $A = (a_{i,j})_{i,j\in[\![1\ ;\ n]\!]}$ and $B = (b_i)^T_{i\in[\![1\ ;\ n]\!]}$. The following question is optional. It is just to help you to think about the problem.

Question 2. Show that A^*B is a solution to the equation.

Question 3. Let X be a solution of $X = AX + B$. Show that $A^*B \subseteq X$.

2.2 The smallest solution is regular

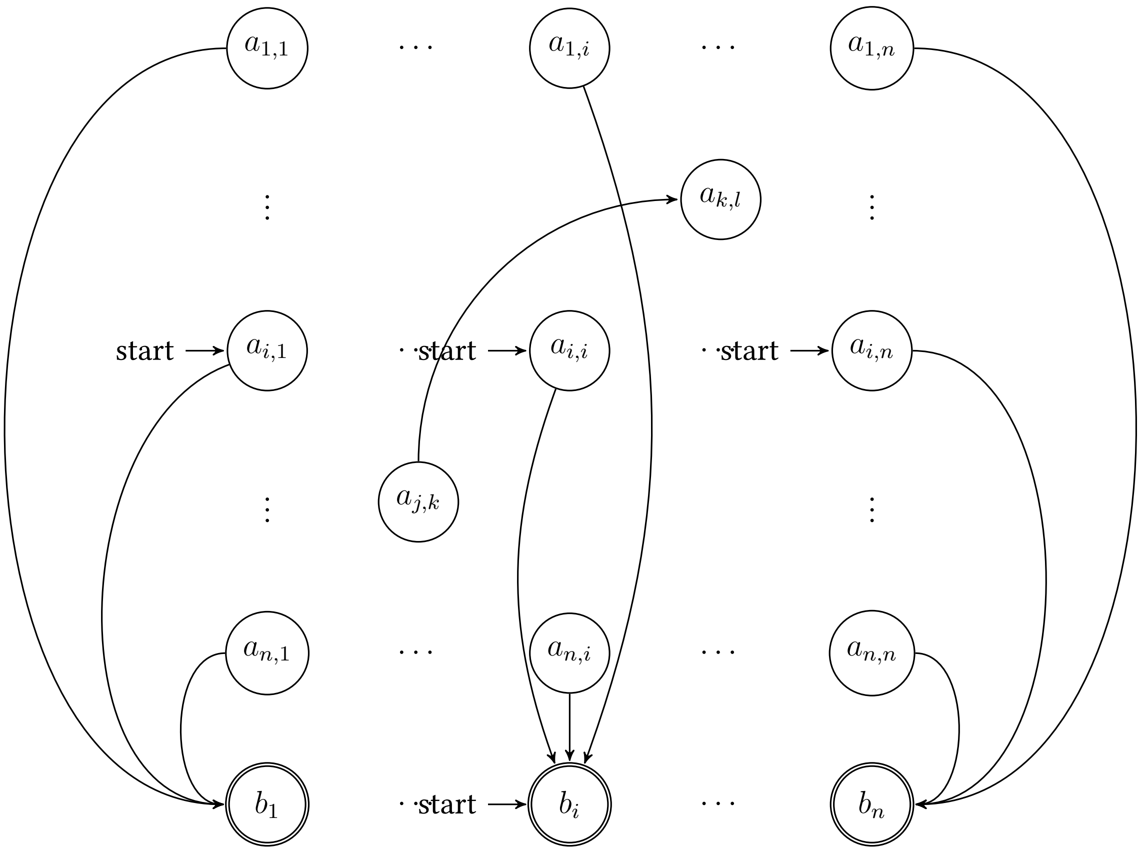
We have shown that A^*B is the smallest solution to the equation, but it is for the moment just a set. Indeed, A^*B is defined as an infinite union of regular languages : none of the coordinate of A^* is an $a^*_{i,j}$ so it must not be trivial to be sure that the languages are indeed regular. We now show that it is a vector of regular languages. For $k \in \mathbb{N}$ we let

$$A^kB = \begin{pmatrix} L_1^{(k)} \\ \vdots \\ L_n^{(k)} \end{pmatrix}$$

Question 4. What is the language $L_i^{(0)}$ for $i \in [\![1 \,;\, n]\!]$?

Question 5. For any $k \in \mathbb{N}$ and $i \in [\![1 \,;\, n]\!]$, give a relation between the languages $L_i^{(k+1)}$, the $L_j^{(k)}$ for $k \in [\![1 \,;\, n]\!]$ and the $a_{i,j}$ for $k \in [\![1 \,;\, n]\!]$.

We consider the following automaton $\mathcal{A}_i$.



$$a_{i,j} = (\Sigma, Q_{i,j}, q_{0,i,j}, \{f_{i,j}\}, \Delta_{i,j}) \qquad \text{(we identify } a_{i,j} \text{ with its automaton)}$$
$$b_i = (\Sigma, Q_i, q_{0,i}, \{f_i\}, \Delta_i) \qquad \text{(we identify } b_i \text{ with its automaton)}$$
$$\mathcal{A}_i = \left(\Sigma, \bigcup_{i,j} (Q_{i,j} \cup Q_i), \{q_{0,i,j} \mid j \in [\![1 \,;\, n]\!]\} \cup \{q_{0,i}\}, \{f_{0,i,j} \mid j \in [\![1 \,;\, n]\!]\} \cup \{f_{0,i}\}, \Delta \right)$$

with
$$\Delta = \left(\bigcup_{i,j} (\Delta_{i,j} \cup \Delta_i) \right) \cup \{(f_{0,i,j}, \varepsilon, q_{0,j}) \mid i, j \in [\![1 \,;\, n]\!]\}$$
$$\cup \{(f_{0,j,k}, \varepsilon, q_{0,k,l}) \mid j, k, l \in [\![1 \,;\, n]\!]\}$$

(We assume, up to renaming,that the $Q_{i,j}$s and Q_is are disjoint.)

Question 6. Let $w \in [\![\mathcal{A}_i]\!]$. Show that $w \in L_i := \bigcup_{k \in \mathbb{N}} L_i^{(k)}$.

Question 7. Let $w \in \bigcup_{k \in \mathbb{N}} L_i^{(k)}$. Show that $w \in [\![\mathcal{A}_i]\!]$.

Question 8. Conclude that the entries of A^*B are regular languages.

2.3 Uniqueness of the solution

In this subsection we consider that none of the languages $a_{i,j}$ contains ε. Let $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ be a solution to $X = AX + B$. We define $X^{(k)} = A^k X = \begin{pmatrix} X_1^{(k)} \\ \vdots \\ X_n^{(k)} \end{pmatrix}$.

Question 9. Show that if $w \in X_i^{(k)}$ then $|w| \geq k$.

Question 10. Show that $X \subseteq A^*B$ and conclude that A^*B is the unique solution to the equation.

3 Regular expression of a non-deterministic automaton in polynomial time

Let $\mathcal{A} = (\Sigma, Q, I, F, \Delta)$ be a finte state automaton with no ε-transition. For $q \in Q$ we let $\mathcal{A}_q = (\Sigma, Q, \{q\}, F, \Delta)$ the automaton $\mathcal{A}$ such that the only initial state is q.

Question 11. Give a matricial equation of the form $X = AX + B$ that must satisfy the vector $X = ([\![\mathcal{A}_q]\!])_{q \in Q}$. Explain why this solution is the unique solution. (Hint : B is the vector $(b_q)_{q \in Q}$ with $b_q = \varnothing$ if $q \notin F$ and $b_q = \{\varepsilon\}$ if $b_q \in F$.)

Question 12. Give an algorithm to compute a solution to this matricial equation. (Hint : proceed as you were solving it for numbers (Gauss algorithm) and use corollary 2.2)

Question 13. Apply your algorithm to the automaton we saw in tutorial :