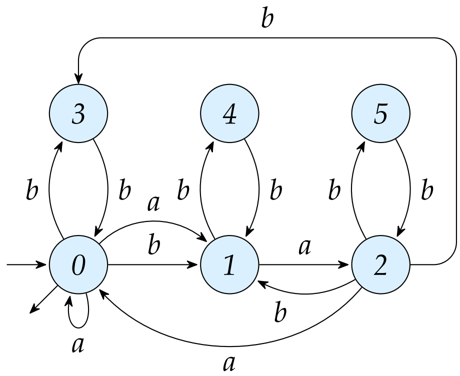
Homework 4
Automata

Version du September 20, 2020

This last [1] homework has to be returned tomorrow, on Friday, at the beginning of the tutorial.

Exercise 1 (Minimization of Brzozowski)**.** ***Take care in this exercise.*** *Count the number of a and the number of b when you copy an automaton from the draft copy to the final copy; count the incoming and outcoming arrows of each state; do not forget to mark the initial and final states. Forgeting something can be fatal when we link together such operations like in here.*

Let us denote by $\mathcal{A}$ the following non-deterministic automaton:



The transposed of $\mathcal{A}$, denoted by $T(\mathcal{A})$, is the automaton in which all the arrows of $\mathcal{A}$ have been inverted (even the initial states became final and vice versa).

The determinized of $\mathcal{A}$, denoted by $Det(\mathcal{A})$, is the DFA that we obtain from $\mathcal{A}$ by using the determinization algorithm seen during the lessons (Note that it also a theorem in the handouts ...).

1. *Build $\mathcal{A}' = Det(T(\mathcal{A}))$.*
2. *Build $\mathcal{A}'' = Det(T(\mathcal{A}'))$.*
   Note: $\mathcal{A}''$ owns 3 states. If you found something else, you made a mistake [2]
3. *Justify that $\mathcal{A}$ and $\mathcal{A}''$ recognize the same language.*

Don't be surprised if the automaton $\mathcal{A}''$ is smaller than the automaton $D(\mathcal{A})$ (that you can build on a draft copy if you want); furthermore, in our case, it is also smaller than $\mathcal{A}$. By chaining these two "co-determinizations" you have built a DFA which is equivalent to $\mathcal{A}$ of minimal size: no automaton exists with less states.

Exercise 2 (Conversion of automata into rational expressions)**.** *Let q and r be two rational expressions denoting the languages $L(q)$ and $L(r)$ respectively of $\Sigma^\star$. Let us consider the equation $X = qX + r$. A rational expression t denoting the language $L(t)$ is solution of this equation if*

$$L(t) = L(q)L(t) \cup L(r) \tag{1}$$

1. *Let prove by induction on n that if t is a solution of (1), then*

$$\forall n \in \mathbb{N},\ L(q)^n L(r) \subset L(t) \tag{2}$$

Note: by convention $L(q)^0 = \{\varepsilon\}$.

1. Courage!
2. It's sad, but it is better to make mistakes at home compared to an exam: the sofa is more confortable.

2. *Show by induction on n that if t is a solution of (1) then*

$$\forall n \in \mathbb{N},\ L(t) \subset L(q)^n L(t) \cup L(q)^{n-1} L(r) \cup \cdots \cup L(r). \tag{5}$$

   Take care not to confuse r and t in the precedent equation!

3. *If $\varepsilon \notin L(q)$ and that t is a solution of this equation, show that $L(t) \subset L(q^\star r)$.*
   Hint: if $\varepsilon \notin L(q)$ the words of $L(q)^n$ are at least of size n, then take each word of $L(t)$ and look how you can choose n in Equation (5).

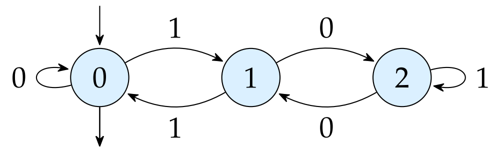
4. *Deduce based on the preceding questions the following theorem:*

   Théorème 1. *Let q and r be two rational expressions such that $\varepsilon \notin L(q)$; if the rational expression t is a solution of the equation $L(t) = L(q)L(t) \cup L(r)$, then $L(q^\star r) = L(t)$.*

   Even if several expressions t can define this same language, we will say that this solution is unique (from the language point of view) for q and r given.

5. *If $\varepsilon \in L(q)$, the equation (1) does not admit always an only solution. Give a solution t which does not depend neither from q nor from r.*

6. ***Application.*** *Let us consider the automaton $\mathcal{D}_3$ of hte preceding Homework:*



   We denote t_i the rational expression denoting the language of all the words who can be accepted by the automaton $\mathcal{D}_3$ starting from *the state i. We have for example $01 \in L(t_2)$ because it is possible to reach a final state when reading 01 starting from state 2.*
   We can state the constraints between t_0, t_1, et t_2 by reading the figure. For example, if we add a 1 at the beginning of a word recognized t_2, it will stay recognized by t_2 because of the loop on state 2. In a same manner if we add a 0 at the starting of a recognized word by t_1, this time it will be recognized by t_2. In fact, the expression t_2 satisfies the equation $t_2 = 0t_1 + 1t_2$.
   If we do this reading of the automaton for each state, we obtain the folowwing system of equations:

$$t_0 = 0t_0 + 1t_1 + \varepsilon \tag{6}$$
$$t_1 = 0t_2 + 1t_0 \tag{7}$$
$$t_2 = 0t_1 + 1t_2 \tag{8}$$

   The ε has been added into the first equation because the state 0 is final: t_0 acceptes then the empty word and the words of t_1 prefixed by 1 and also its own words prefixed by 0. The rational expression t_0, because it is associated to the initial state, denotes the language accepted by the automaton. To rebuild a rational expressionassociated to the automaton, it is sufficient to[3] to solve the system of equations (6)-(8) to find t_0.
   Let's do the first step together. By replacing (7) into (6) and (8) we eliminate t_1 from our system. Good point:

$$t_0 = (0 + 11)t_0 + 10t_2 + \varepsilon \tag{9}$$
$$t_2 = (00 + 1)t_2 + 01t_0 \tag{10}$$

   *It is now your turn: **find** t_0.*
   Hints: these two equations have the same shape: $t = qt + r$. Begin with explaining Theorem 1 at the equation (10) to express t_2 w.r.t. t_0 only, and then use your result into 9 before applying one more time the theorem.

3. The only difficulty , really, is to realize que our products are in fact concatenations). The concatenation cannot be switched and does not admit an inverse.