DM7

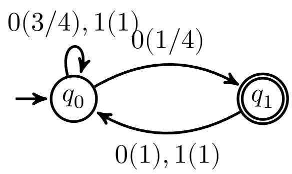
Soit $\Sigma = \{0,1\}$ un alphabet. Un *automate probabiliste* sur Σ est un quadruplet $\mathcal{A} = (Q, q_0, F, Pr)$ où :

- Q est un ensemble fini non vide dont les éléments sont appelés *états* ;
- $q_0 \in Q$ est appelé *état initial* ;
- $F \subset Q$ est un ensemble dont les éléments sont appelés *états finals* ;
- $Pr : Q \times \Sigma \times Q \to [0,1]$ est une application appelée *fonction probabiliste de transition* ; on suppose que pour tout $q \in Q$, pour tout $\alpha \in \Sigma$, $\sum_{q' \in Q} Pr(q, \alpha, q') = 1$. On note $Pr(q \xrightarrow{\alpha} q')$ pour $Pr(q, \alpha, q')$.

Une *transition* est un triplet $(q, \alpha, q') \in Q \times \Sigma \times Q$, noté $q \xrightarrow{\alpha} q'$, avec $Pr(q \xrightarrow{\alpha} q') > 0$.

On représente un automate probabiliste de manière graphique, de façon similaire à la représentation des automates non déterministes classiques. Les états sont représentés par des cercles, l'état initial par une flèche arrivant sur le cercle correspondant, les états finals par des cercles doubles (ou une flèche sortante). La fonction probabiliste de transition est représentée par une flèche entre états : si $Pr(q \xrightarrow{\alpha} q')$ est un nombre $p > 0$, on met une flèche de l'état q à l'état q', annotée par $\alpha(p)$.

Ainsi, l'automate $\mathcal{A}_0 = (\{q_0, q_1\}, q_0, \{q_1\}, Pr)$ représenté ci-dessous :



a pour fonction probabiliste de transition Pr la fonction suivante (seules les valeurs non nulles sont mentionnées)

q	α	q'	$Pr(q \xrightarrow{\alpha} q')$
q_0	0	q_0	$3/4$
q_0	0	q_1	$1/4$
q_0	1	q_0	1
q_1	0	q_0	1
q_1	1	q_0	1

Etant donné un automate probabiliste $\mathcal{A} = (Q, q_0, F, Pr)$ sur Σ, un *chemin* ρ est une suite finie de transitions $q_{i_1} \xrightarrow{\alpha_1} q_{i_2}$, ..., $q_{i_n} \xrightarrow{\alpha_n} q_{i_{n+1}}$ aussi notée $q_{i_1} \xrightarrow{\alpha_1} q_{i_2} \xrightarrow{\alpha_2} \dots \xrightarrow{\alpha_n} q_{i_{n+1}}$. On dit que ρ a pour *étiquette* le mot $\alpha_1 \dots \alpha_n \in \Sigma^*$, pour *état de départ* l'état $q_{i_1} \in Q$ et pour *état d'arrivée* l'état $q_{i_{n+1}} \in Q$. La *probabilité* de ρ, notée $Pr(\rho)$, est définie par :

$$Pr(\rho) = \prod_{k=1}^{n} Pr(q_{i_k} \xrightarrow{\alpha_k} q_{i_{k+1}})$$

Un état peut être vu comme un chemin de longueur nulle. La probabilité d'un chemin de longueur nulle est égale à 1. Un *chemin pour le mot* $u \in \Sigma^*$ est un chemin dont l'étiquette est u et l'état de départ est q_0. Ce chemin est *acceptant* si l'état d'arrivée est un état de F et *non acceptant* sinon. La *probabilité d'un mot* $u \in \Sigma^*$, notée $Pr(u)$ est par définition la somme des probabilités de tous les chemins acceptants pour le mot $u \in \Sigma^*$:

$$Pr(u) = \sum_{\rho \text{ chemin acceptant pour } u} Pr(\rho)$$

Question 1. Calculer les probabilités $Pr(\varepsilon)$, $Pr(0)$, $Pr(010)$ pour l'automate $\mathcal{A}_0$. ε désigne le mot vide.

Question 2. Pour tout automate probabiliste $\mathcal{A} = (Q, q_0, F, Pr)$ et tout mot $u \in \Sigma^*$, établir l'égalité suivante en utilisant une récurrence sur la longueur du mot u :

$$Pr(u) = 1 - \sum_{\rho \text{ chemin non-acceptant pour } u} Pr(\rho)$$

Question 3. On revient à l'automate probabiliste $\mathcal{A}_0$. Quels sont les mots u dont la probabilité $Pr(u)$ pour $\mathcal{A}_0$ est égale à 0 ? Quels sont ceux dont la probabilité est égale à 1 ?

Question 4. Proposer, sans justification, une expression régulierle pour le langage des mots u dont la probabilité $Pr(u)$ pour $\mathcal{A}_0$ est non nulle.

Question 5. Montrer que pour tout automate probabiliste $\mathcal{A} = (Q, q_0, F, Pr)$, il existe un automate non nécessairement déterministe $\mathcal{A}'$ qui accepte exactement les mots u dont la probabilité $Pr(u)$ pour $\mathcal{A}$ est non nulle.

Question 6. Appliquer la construction précédente à l'automate $\mathcal{A}_0$ pour obtenir un automate non-déterministe qui accepte exactement les mots u dont la probabilité $Pr(u)$ pour $\mathcal{A}_0$ est non nulle. Déterminiser cet automate.

Soit $\mathcal{A} = (Q, q_0, F, Pr)$ un automate probabiliste sur Σ. Pour un réel $\eta \in [0, 1[$, le *η-langage reconnu par $\mathcal{A}$*, noté $\mathcal{L}_\eta(\mathcal{A})$, est défini par :

$$\mathcal{L}_\eta(\mathcal{A}) = \{u \in \Sigma^* \mid Pr(u) > \eta\}$$

On dit qu'un langage $L \subset \Sigma^*$ est *stochastique* s'il existe un automate probabiliste $\mathcal{A}$ et un réel $\eta \in [0, 1[$ tel que $L = \mathcal{L}_\eta(\mathcal{A})$.

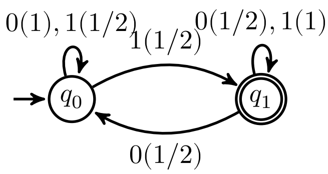
Question 7. Démontrer que tout langage régulier est stochastique.

Etant donné un mot $\alpha_1 \dots \alpha_n$ sur l'alphabet $\{0, 1\}$, on dit que l'expression $\underline{0, \alpha_1 \dots \alpha_n}_2$ est une *écriture (finie) en base* 2 du nombre réel $\sum_{i=1}^{n} 2^{-i}\alpha_i$. On note alors

$$\sum_{i=1}^{n} 2^{-i}\alpha_i = \underline{0, \alpha_1 \dots \alpha_n}_2$$

Ainsi, $\frac{1}{4} = 2^{-2} = \underline{0, 01}_2 = \underline{0, 0100}_2$.

On considère maintenant l'automate $\mathcal{A}_1 = (\{q_0, q_1\}, q_0, \{q_1\}, Pr)$ ci-dessous :



Question 8. Dans l'automate $\mathcal{A}_1$, calculer $Pr(q_0 \xrightarrow{1} q_0 \xrightarrow{1} q_1 \xrightarrow{1} q_1 \xrightarrow{0} q_0 \xrightarrow{1} q_1)$ et en donner une écriture finie en base 2.

Question 9. Dans l'automate $\mathcal{A}_1$, calculer $Pr(10)$ et en donner une écriture finie en base 2.

Question 10. Dans l'automate $\mathcal{A}_1$, calculer $Pr(1101)$ et en donner une écriture finie en base 2.

Question 11. Soit $u \in \Sigma^*$ un mot arbitraire sur Σ. Montrer que $Pr(u)$ pour $\mathcal{A}_1$ admet une écriture finie en base 2 et en donner une expression. Prouver que cette écriture est correcte.

Question 12. Soit $\eta \in [0, 1[$. Prouver l'égalité suivante :

$$\mathcal{L}_\eta(\mathcal{A}_1) = \{\alpha_1 \dots \alpha_n \in \Sigma^* \mid \underline{0, \alpha_n \dots \alpha_1}_2 > \eta\}$$

Question 13. En déduire qu'il existe des langages stochastiques qui ne sont pas réguliers.