SCIENCES SORBONNE UNIVERSITÉ

Année 2024–2025

LU2IN005 - 2 février 2025

devoir maison supplémentaire

On se fixe dans un alphabet $A = \{0, 1\}$.

Un *automate probabiliste* sur l'alphabet A est donné par $\mathcal{A} = (S, \mathrm{Pr}, s_0, F)$ où :

(i) S est un ensemble fini non vide dont les élements sont appelés *états*,

(ii) $s_0 \in S$ est appelé *état initial*,

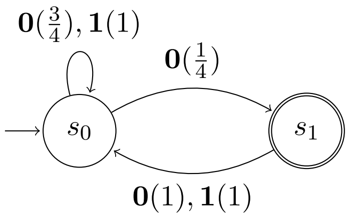
(iii) $F \subseteq S$ est un ensemble dont les éléments sont appelés états finals,

(iv) $\mathrm{Pr} : S \times A \times S \to [0, 1]$ est une application appellée *fonction probabiliste de transition*,

(v) On suppose de plus que pour tout $s \in S$ et tout $a \in A$,

$$\sum_{s' \in S} \mathrm{Pr}(s, a, s') = 1$$

On peut adapter les notations du cours à ce nouveua genre d'automates. Ainsi, l'automate $\mathcal{A}_0 = (\{s_0, s_1\}, s_0; \{s_1\}, \mathrm{Pr})$ représenté ci-dessous a pour fonction probabiliste de transition Pr la fonction suivante (seules les valeurs non nulles sont mentionnées) :



s	a	s'	$\mathrm{Pr}(s, a, s')$
s_0	0	s_0	3/4
s_0	0	s_1	1/4
s_0	1	s_0	1
s_1	0	s_0	1
s_1	1	s_1	1

Étant donné un automate probabiliste $\mathcal{A} = (S, \mathrm{Pr}, s_0, F)$ sur A, une *execution* e est une suite finie de transitions $s_{i_1} \xrightarrow{u_1} s_{i_2}, \ldots, s_{i_n} \xrightarrow{u_n} s_{i_{n+1}}$ aussi notée $s_{i_1} \xrightarrow{u_1} s_{i_2} \xrightarrow{u_1} \ldots \xrightarrow{u_n} s_{i_{n+1}}$, on dit que e a comme *étiquette* le mot $u_1 \ldots u_n \in A^*$ pour état de départ $s_{i_1} \in S$ et état d'arrivé $s_{i_n} \in S$.

La *probabilité* d'une execution noté $\mathrm{Pr}(e)$ est définie par

$$\mathrm{Pr}(e) = \prod_{k=1}^{n} Pr(s_{i_k} \xrightarrow{a_k} s_{i_{k+1}})$$

Un état peut être vu comme une execution de longueur nulle, sa probabilité est égale à 1.

Une *execution pour le mot* $u \in A^*$ est une execution dont l'étiquette est u et l'état de départ est s_0. Cette execution est *acceptante* si l'état d'arrivé est un état de F, *non-acceptante* sinon.

La *probabilité d'un mot* $u \in A^*$, noté $\Pr(u)$, est par définition la somme des probabilité de tous les executions acceptantes pour le mot $u \in A^*$:

$$\Pr(u) = \sum_{e \text{ execution acceptante pour } u} \Pr(e)$$

1. Calculer les probabilités $\Pr(\varepsilon)$, $\Pr(\mathbf{0})$, $P(\mathbf{010})$ de l'automate $\mathcal{A}_0$. (ici ε est le mot vide).

2. Soit $\mathcal{A} = (S, \Pr, s_0, F)$ un automate probabiliste. Dans l'énoncé, on a utilisé des notations du langage des probabilités sans montrer que les objets que nous manipulons sont des probabilités. Le but de cette question est de montrer la propriétée de *masse-unitaire* pour $\Pr(e)$ où e est une execution de longueur fixée.

   (a) ($\star$) Montrer par induction sur n la propriétée de *masse unitaire*.

$$1 = \sum_{e \text{ execution de longueur } n} \Pr(e)$$

   (b) En déduire

$$\Pr(u) = 1 - \sum_{e \text{ execution non-acceptante pour } u} \Pr(e)$$

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

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

5. Montrer que pour tout automate probabiliste $\mathcal{A} = (S, \Pr, s_0, F)$, 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.

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

Soit $\mathcal{A} = (S, \Pr, s_0, F)$ un automate probabiliste sur A. Pour un réel $x \in [0, 1[$ le *x-langage reconnu par* $\mathcal{A}$, noté $\mathcal{L}_x(\mathcal{A})$, est défini par :

$$\mathcal{L}_x(\mathcal{A}) := \{u \in A^* \mid \Pr(u) > x\}$$

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

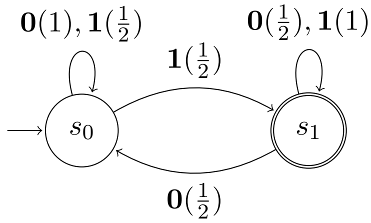
7. Démontrer que tout langage rationnel est stochastique.

Étant donné un mot $a_1 \cdots a_n$ sur l'alphabet $\{0, 1\}$, on dit que l'expression $\underline{0, a_1 \cdots a_n}_2$ est une *écriture (finie) en base 2* du nombre réel $\sum_{i=1}^{n} 2^{-i} a_i$. On note alors :

$$\sum_{i=1}^{n} 2^{-i} a_i = \underline{0, a_1 \cdots a_n}_2$$

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

8. On considère maintenant l'automate $\mathcal{A}_1 = (\{s_0, s_1\}, \mathrm{Pr}, s_0, \{s_1\})$ ci-dessous :



- **(a)** Dans l'automate $\mathcal{A}_1$, calculer $Pr(s_0 \xrightarrow{\mathbf{1}} s_0 \xrightarrow{\mathbf{1}} s_1 \xrightarrow{\mathbf{1}} s_1 \xrightarrow{\mathbf{0}} s_0 \xrightarrow{\mathbf{1}} s_1)$ et en donner une écriture finie en base 2.
- **(b)** Dans l'automate $\mathcal{A}_1$, calculer $\mathrm{Pr}(\mathbf{10})$ et en donner une écriture finie en base 2.
- **(c)** Dans l'automate $\mathcal{A}_1$, calculer $\mathrm{Pr}(\mathbf{1101})$ et en donner une écriture finie en base 2.
- **(d)** Soit $u \in A^*$ un mot arbitraire sur A. Montrer que $\mathrm{Pr}(u)$ pour $\mathcal{A}_1$ admet une écriture finie en base 2, et en donner une expression. Prouver que cette écriture est correcte.
- **(e)** Soit $x \in [0, 1[$. Prouver l'égalité suivante :

$$\mathcal{L}_x(\mathcal{A}_1) = \left\{a_1 \cdots a_n \in A^* \mid \underline{0, a_n \cdots a_1}_2 > x\right\}$$

- **(f)** En déduire qu'il existe des langages stochastiques qui ne sont pas rationnels. *Indication : On rappelle que l'ensemble des langages rationnels est dénombrable.*