DM7 (éléments de réponses)

Question 1. Le mot vide n'est l'étiquette d'aucun chemin acceptant et donc : $Pr(\varepsilon) = 0$.
Puisque $q_0 \xrightarrow{0} q_1$ est le seul chemin acceptant d'étiquette 0 et que sa probabilité vaut $1/4$, on a : $Pr(0) = \frac{1}{4}$.
Il existe deux chemins acceptants d'étiquette 010 qui sont :

$$q_0 \xrightarrow{0} q_0 \xrightarrow{1} q_0 \xrightarrow{0} q_1 \quad \text{et} \quad q_0 \xrightarrow{0} q_1 \xrightarrow{1} q_0 \xrightarrow{0} q_1$$

Ainsi : $Pr(010) = \frac{3}{4} \times 1 \times \frac{1}{4} + \frac{1}{4} \times 1 \times \frac{1}{4} = \frac{1}{4}$.

Question 2. Prouvons le résultat plus général suivant. Si $q \in Q$, notons *q-chemin pour u* un chemin d'étiquette u d'état initial q. On parle aussi de q-chemin acceptant (ou non-acceptant) pour u selon que l'état final du chemin est (ou n'est pas) dans F. Posons ensuite :

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

Alors, par récurrence sur la taille de u, on a :

$$\forall q \in Q,\ Pr_q(u) = 1 - \sum_{\rho\ q\text{-chemin non-acceptant pour } u} Pr(\rho)$$

- *Initialisation* : ε est l'unique mot vide. Soit $q \in Q$. Le seul q-chemin pour ε est le chemin de longueur nulle q.
  Si ce chemin est q-acceptant ($q \in F$) on a $Pr_q(u) = 1$ et la somme des probabilités des q-chemin non-acceptant pour ε est bien nulle (il n'y a pas de tel chemin).
  Sinon, $Pr_q(u) = 0$ et la somme des probabilités des q-chemin non-acceptant pour ε est vaut 1 (probabilité du chemin q).
  Dans les deux cas, on a la formule demandée.
- *Hérédité* : supposons le résultat vrai pour les mot de longueur $n \geq 0$ donné. Considérons un mot u de longueur $n+1$. On peut l'écrire $u = xv$ avec v mot de longueur n et $x \in \{0, 1\}$. Soit $q \in Q$; les q-chemins pour u sont du type

  $$q \xrightarrow{x} q' \xrightarrow{v} q''$$

  et on a donc

  $$Pr_q(u) = \sum_{q' \in Q} Pr(q \xrightarrow{x} q') Pr_{q'}(v)$$

  Par hypothèse de récurrence, on a :

  $$\forall q' \in Q,\ Pr_{q'}(v) = 1 - \sum_{\rho\ q'\text{-chemin non-acceptant pour } v} Pr(\rho)$$

  On en déduit que :

  $$\begin{aligned} Pr_q(u) &= \sum_{q' \in Q} Pr(q \xrightarrow{x} q') - \sum_{q' \in Q} \left(Pr(q \xrightarrow{x} q') \sum_{\rho\ q'\text{-chemin non-acceptant pour } v} Pr(\rho) \right) \\ &= 1 - \sum_{\substack{q' \in Q \\ \rho\ q'\text{-chemin non-acceptant pour } v}} (Pr(q \xrightarrow{x} q') Pr(\rho)) \\ &= 1 - \sum_{\rho\ q\text{-chemin non-acceptant pour } u} Pr(\rho) \end{aligned}$$

  Ceci prouve le résultat au rang $n+1$.

Il suffit d'appliquer le résultat avec $q = q_0$.

Question 3. Si u est un mot se terminant par 1 alors un chemin dans $\mathcal{A}_0$ de q_0 à q_1 et d'étiquette u a une dernière transition de probabilité nulle et donc une probabilité nulle. Ainsi, $Pr(u) = 0$. Si u est un mot se terminant par 0 alors un chemin bouclant sur q_0 et passant en q_1 lors de la lecture de la dernière lettre est composé de transitions de probabilités toutes non nulles et a donc une probabilité non nulle. On a alors $Pr(u) > 0$. Quant au mot vide, il est de probabilité nulle dans $\mathcal{A}_0$. Ainsi, les mots u tels que $Pr(u) = 0$ pour $\mathcal{A}_0$ sont ceux de $\{\varepsilon\} \cup \Sigma^* 1$.

Pour tout mot u, il existe un chemin non-acceptant pour u et de probabilité > 0 (celui qui boucle sur q_0). Ainsi, la somme des longueur de ces chemins est > 0. Avec la question précédente, $Pr(u) < 1$ et il n'existe pas de mot u tels que $Pr(u) = 1$ pour $\mathcal{A}_0$.

Question 4. La question précédente montre que le langage cherché est Σ^*0.

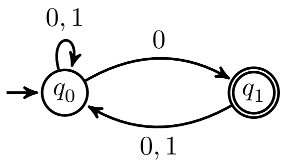
Question 5. Notons $\mathcal{A}' = (Q, q_0, F, \gamma)$ où $\gamma \subset Q \times \Sigma \times Q$ est défini par :

$$\gamma = \{(q, \alpha, q')/\ Pr(q, \alpha, q') > 0\}$$

On a un automate non déterministe dont on va montrer qu'il reconnaît exactement les mots u dont la probabilité $Pr(u)$ pour $\mathcal{A}$ est non nulle.

- Soit u un mot reconnu par $\mathcal{A}'$. Il existe alors un chemin dans $\mathcal{A}'$ d'origine q_0 et d'extrémité dans F d'étiquette u. Par définition de γ, le chemin similaire dans $\mathcal{A}$ est composé de transitions de probabilités > 0 et a donc une probabilité > 0. Ainsi $Pr(u) > 0$ (c'est la somme de quantités ≥ 0 dont une au moins est > 0).
- Réciproquement, soit u tel que $Pr(u) > 0$. Il y a alors au moins un chemin acceptant pour u dans $\mathcal{A}$ et de probabilité > 0 et donc composé de transitions de probabilités > 0. Par définition de γ, le même chemin existe dans $\mathcal{A}'$ et va de q_0 à un élément de F. Ainsi, $u \in \mathcal{A}'$.

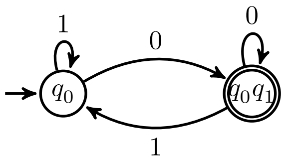
Question 6. Dans le cas de $\mathcal{A}_0$, on obtient :



La table du déterminisé de cet automate est :

	q_0	q_0, q_1
0	q_0, q_1	q_0, q_1
1	q_0	q_0

et sa représentation graphique est :



Question 7. Soit L un langage rationnel. Il existe un automate déterministe complet $\mathcal{A} = (Q, q_0, F, \delta)$ qui reconnaît ce langage. Considérons la fonction Pr définie par :

$$\forall q \in Q,\ \forall \alpha \in \Sigma,\ Pr(q, \alpha, \delta(q, \alpha)) = 1$$

les autres valeurs prises par Pr étant nulles. Ceci revient à donner à chaque transition qui existe dans $\mathcal{A}$ la probabilité 1. Notons $\mathcal{A}' = (Q, q_0, F, Pr)$. Le fait que $\mathcal{A}$ est déterministe nous permet de voir que $\mathcal{A}'$ est déterministe (la somme des probabilités des transitions issues d'un état donné q et étiquetée par une lettre donnée α vaut 1).
Par construction, les chemins dans $\mathcal{A}'$ sont tous de probabilité 1. $Pr(u) > 0$ équivaut à l'existence d'un chemin acceptant pour u dans $\mathcal{A}'$ et donc au fait que le chemin dans $\mathcal{A}$ d'origine q_0 et d'étiquette u se termine dans un élément de F, c'est à dire au fait que u est reconnu par $\mathcal{A}$. On a ainsi $L = \mathcal{L}_0(\mathcal{A}')$ et L est stochastique.

Question 8.

$$Pr(q_0 \xrightarrow{1} q_0 \xrightarrow{1} q_1 \xrightarrow{1} q_1 \xrightarrow{0} q_0 \xrightarrow{1} q_1) = \frac{1}{2} \times \frac{1}{2} \times 1 \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16} = \underline{0{,}0001}_2$$

Question 9. $q_0 \xrightarrow{1} q_1 \xrightarrow{0} q_1$ est le seul chemin acceptant pour 10 et sa probabilité est $\frac{1}{4}$. Ainsi :

$$Pr(10) = \frac{1}{4} = \underline{0{,}01}_2$$

Question 10. On peut représenter de manière arborescente les différents chemins acceptants pour 1101 dans $\mathcal{A}_1$:
On a alors :

$$Pr(1101) = \frac{1}{8} + \frac{1}{16} + \frac{1}{8} + \frac{1}{8} + \frac{1}{4} = \frac{1}{2} + \frac{1}{8} + \frac{1}{16} = \underline{0,1011}_2$$

Question 11. On a $Pr(\varepsilon) = 0$ dont une écriture finie en base 2 est 0. Montrons maintenant que :

$$\forall u = \alpha_1 \dots \alpha_n \in \Sigma^n,\ Pr(u) = \underline{0, \alpha_n \dots \alpha_1}_2$$

On procède pour cela par récurrence sur n. Si $u = \alpha_1 \dots \alpha_n$, je note $N_u = \underline{0, \alpha_n \dots \alpha_1}_2$.

- On a $Pr(0) = 0 = \underline{0,0}_2$ et $Pr(1) = \frac{1}{2} = \underline{0,1}_2$ ce qui montre le résultat pour $n = 1$.
- Supposons le résultat vrai jusqu'à un rang $n \geq 1$. Soit $v = \alpha_1 \dots \alpha_n \alpha_{n+1}$; on pose $u = \alpha_1 \dots \alpha_n$. Par hypothèse de récurrence, la lecture de u depuis q_0 amène en q_1 avec probabilité N_u et donc en q_0 avec probabilité $1 - N_u$. On distingue deux cas.
  - Si $\alpha_{n+1} = 0$, la lecture de v mènera en q_1 si et seulement celle de u à amené en q_1 et qu'on emprunte la transition de q_1 vers lui même pour la dernière lettre. La probabilité est alors de $\frac{1}{2}N_u = N_v$ (multiplier par $1/2$ décale toutes les « décimales »).
  - Si $\alpha_{n+1} = 1$, les bon chemins sont ceux qui mènent par v à q_0 puis passent à q_1 ou qui mènent par v à q_1 et y restent. Ceci advient avec probabilité $\frac{1}{2}(1 - N_u) + N_u = \frac{1}{2} + \frac{1}{2}N_u = N_v$ (on décale les “décimales” par le facteur $1/2$ et on ajoute une première « décimale » égale à 1 par l'ajout de $1/2$).

  On obtient ainsi le résultat au rang $n + 1$.

Question 12. Le mot vide est de probabilité nulle et donc n'est pas dans $\mathcal{L}_\eta(\mathcal{A}_1)$.
Soit $u = \alpha_1, \dots \alpha_n$ un mot non vide. Il est dans $\mathcal{L}_\eta(\mathcal{A}_1)$ si $Pr(u) > \eta$ et la question précédente montre que ceci équivaut à $\underline{0, \alpha_n \dots \alpha_1}_2 > \eta$.

Question 13. Pour conclure, il reste à montrer qu'il existe $\eta \in [0, 1[$ tel que $\mathcal{L}_\eta(\mathcal{A}_1)$ n'est pas rationnel.
Or, le nombre de langages rationnels sur Σ est dénombrable (en effet, $Rat(\Sigma)$ est défini par récurrence à partir d'un nombre fini de cas de base et par application de trois règles ; l'ensemble R_n des langages rationnels obtenus par application d'au plus n fois une règle est fini ; $Rat(\Sigma)$ est la réunion, dénombrable, des R_n).
On pourra conclure si on montre qu'il existe un nombre infini non dénombrable de $\mathcal{L}_\eta(\mathcal{A}_1)$. Ceci est vrai car si $\eta < \eta'$ alors $\mathcal{L}_{\eta'}(\mathcal{A}_1) \subset \mathcal{L}_\eta(\mathcal{A}_1)$ et ceci est une inclusion stricte puisque l'ensemble des nombres ayant une écriture finie en base 2 est dense dans $[0, 1]$ (même preuve que dans le cs décimal) et qu'on peut donc trouver un nombre ayant une écriture finie en base 2 dans $]\eta, \eta'[$.