Automates Probabilistes

Furcy Pin

20/12/2007

1 Définition et interprétation

Pour simplifier, on se donne dès maintenant un alphabet fini A fixé dans la suite. On introduit tout d'abord quelques notions de probabilités.

Définition 1. On appelle *vecteur de probabilités* de taille n un vecteur de $([0,1])^n$ dont la somme des coordonnées vaut 1.

Définition 2. On appelle *matrice de transition* une matrice de $M_n(\mathbb{R})$ dont chaque vecteur ligne est un vecteur de probabilités i.e. à valeurs dans $([0,1])$ et le vecteur $(1,1,...,1)$ est vecteur propre pour la valeur propre 1.

Remarque : On peut facilement vérifier que le produit de deux matrices de transition est encore un matrice de transition.

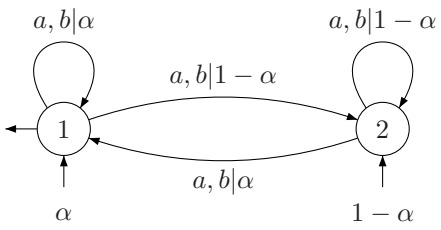
Définition 3. Un *automate probabiliste* de taille n est la donnée d'un quadruplet $\mathcal{A} = (Q, \pi, M, F)$ où :

- $Q = \{q_1, q_2, ..., q_n\}$ est un ensemble fini d'état de cardinal n
- π est un vecteur de probabilité de taille n où π_i indique la probabilité pour que l'automate démarre dans l'état q_i
- M est une application de l'alphabet A dans l'ensemble des matrices de transition. On notera $M_a = M(a)$ pour toute lettre $a \in A$
- $F \subset Q$ est un ensemble d'états finaux

Exemple et représentation : *(On prendra pour cet exemple $A = \{a, b\}$)* L'automate suivant : $\mathcal{A} = (Q, \pi, M, F)$ avec :

- $Q = \{1, 2\}$
- $\pi = (\alpha, 1-\alpha)$ où $\alpha \in [0,1]$
- $M_a = M_b = \begin{pmatrix} \alpha & 1-\alpha \\ \alpha & 1-\alpha \end{pmatrix}$
- $F = \{1\}$

sera représenté ainsi :



Interprétation :

On peut interpréter le fonctionnement de l'automate ainsi : étant donné un mot $u = u_1u_2...u_m \in A^*$, l'automate probabiliste $\mathcal{A} = (Q, \pi, M, F)$ va lire ce mot de la façon suivante :

- L'automate démarre dans l'un des états initiaux choisit aléatoirement selon la loi de distribution donné par le vecteur π.
- Il lit ensuite le mot u de gauche à droite comme un automate normal en choisissant aléatoirement un des transitions possibles. La probabilité qu'une arrête soit prise est donnée par la matrice de transition.
- Enfin le mot est accepté si l'automate arrive dans un état final après l'avoir lu.

Interprétation sur l'exemple :

Pour l'automate donné en exemple, toutes les transitions menant à l'état 1 (y compris la transition d'entrée) ont une probabilité α d'être choisie et toutes celles menant à l'état 2 ont une probabilité $1-\alpha$ d'être choisie. L'état 1 étant le seul état final, tout mot (y compris le mot vide) aura une probabilité α d'être accepté par cet automate. Autrement dit, la fonction P qui a un mot u associe la probabilité que cet automate l'accepte est constante.

Remarque : De cette façon, un automate probabiliste $\mathcal{A}$ peut accepter et rejeter un même mot u si on le lance plusieurs fois. Cependant, la probabilité $P^{\mathcal{A}}(u)$ qu'il accepte ce mot est fixe et facilement calculable, grâce à la formule suivante :

Proposition 4. *Pour tout automate probabiliste $\mathcal{A} = (Q, \pi, M, F)$ et tout mot $u = u_1u_2...u_m$, on a :*

$$P^{\mathcal{A}}(u) = \pi M_{a_1} M_{a_2} ... M_{a_m} \mathbb{1}_F$$

où $\mathbb{1}_F$ est le vecteur colonne tel que pour tout $i \in [\![1, n]\!]$, $\mathbb{1}_F = \begin{cases} 1 \text{ si } q_i \in F \\ 0 \text{ sinon} \end{cases}$

Notations :
Avant de démontrer cette proposition, nous allons introduire quelques notations utiles.

- On notera $\mathbb{P}(\mathcal{A} \mapsto q_i) := \pi_{(i)}$ la probabilité qu'un automate $\mathcal{A}$ démarre dans l'état q_i
- On notera $\mathbb{P}(q_i \xrightarrow{u} q_j)$ la probabilité qu'un automate passe de l'état q_i à l'état q_j en lisant le mot u.
- Enfin on notera $\mathbb{P}(\mathcal{A}(u) = q_i)$ la probabilité qu'un automate $\mathcal{A}$ arrive dans l'état q_i après avoir lu le mot u.
- On notera enfin $\Pi(u)$ le vecteur de probabilité correspondant, i.e. : $\Pi(u)_{(i)} = \mathbb{P}(\mathcal{A}(u) = q_i)$ pour tout i

Preuve. On va d'abord montrer par récurrence sur m que pour tout mot $u = a_1...a_m$ non vide $\mathbb{P}(q_i \xrightarrow{u} q_j) = (M_{a_1}...M_{a_m})_{(i,j)}$.

- Si $u = a \in A$ alors c'est immédiat par définition de M.
- Si $u = a_1...a_m$ et $a_0 \in A$ alors on a :

$$\begin{aligned}\mathbb{P}(q_i \overset{a_0 u}{\to} q_j) &= \sum_{k=1}^{n} \mathbb{P}(q_i \overset{a_0}{\to} q_k)\mathbb{P}(q_k \overset{u}{\to} q_j) \\ (H.R.) &= \sum_{k=1}^{n} M_{a_0\,(i,k)}(M_{a_1}...M_{a_m})_{(k,j)} = (M_{a_0}...M_{a_m})_{(i,j)}\end{aligned}$$

On a alors pour tout mot $u = a_1...a_m$:

$$\begin{aligned}\mathbb{P}(\mathcal{A}(u) = q_j) &= \sum_{i=1}^{n} \mathbb{P}(\mathcal{A} \mapsto q_i)\mathbb{P}(q_i \overset{u}{\to} q_j) \\ &= \sum_{i=1}^{n} \pi_{(i)}(M_{a_1}...M_{a_m})_{(i,j)} \\ &= (\pi M_{a_1}...M_{a_m})_{(j)}\end{aligned}$$

Et enfin

$$P^{\mathcal{A}}(u) = \sum_{\substack{j=1 \\ q_j \in F}}^{n} \mathbb{P}(\mathcal{A}(u) = q_j) = \sum_{j=1}^{n} (\pi M_{a_1}...M_{a_m})_{(j)} \mathbb{1}_{F(j)} = \pi M_{a_1}...M_{a_m}\mathbb{1}_F$$

□

Remarque : On a au passage démontré que $\Pi(u) = (\pi M_{a_1}...M_{a_m})$ ce qui sera utile par la suite.

On sait de plus grâce à la loi des grands nombres que si on lance un automate probabiliste $\mathcal{A}$ sur un même mot u un grand nombre de fois N et qu'on note $S(N)$ le nombre de fois où ce mot est accepté, alors presque sûrement : $\frac{S(N)}{N} \underset{N \to +\infty}{\longrightarrow} P^{\mathcal{A}}(u)$.

2 Opérations sur les automates

Un automate probabiliste $\mathcal{A}$ peut être interprétés comme une fonction $P^{\mathcal{A}} : A^* \to [0,1]$. On peut alors réaliser certaines opérations naturelles sur ces automates.

Complémentaire
Soit $\mathcal{A}$ un automate de fonction associée f , alors il existe un automate de fonction associée $1-f$.

Preuve. On pose $\mathcal{A} = (Q, \pi, M, F)$ et on construit $\mathcal{A}' = (Q, \pi, M, Q \setminus F)$. Il est alors que $P^{\mathcal{A}'} = 1 - f$. □

Produit d'automates
Soient $\mathcal{A}_f$ et $\mathcal{A}_g$ des automates probabilistes de fonctions associées f et g. Alors il existe un automate probabiliste $\mathcal{A}$ tel que $P^{\mathcal{A}} = fg$.

Preuve. On pose $\mathcal{A}_f = (Q^f, \pi^f, M^f, F^f)$ et $\mathcal{A}_g = (Q^g, \pi^g, M^g, F^g)$. On définit alors $\mathcal{A}$ comme suit :
$\mathcal{A} = (Q = Q^f \times Q^g, \pi = \pi^f \otimes \pi^g, M = M^f \otimes M^g, F^f \times F^g)$ où :
- $\pi_{(i,i')} = \pi^f_{(i)}\pi^g_{(i')}$ pour tous $1 \leq i \leq |Q^f|, 1 \leq i' \leq |Q^g|$
- $M_{a(i,i')(j,j')} = M^f_{a\,(i,j)} M^g_{a\,(i',j')}$ pour toute lettre a et tous $1 \leq i, j \leq |Q^f|$, $1 \leq i', j' \leq |Q^g|$

On vérifie qu'on a alors par définition de $\mathcal{A}$:
- $\mathbb{P}(\mathcal{A} \mapsto (q_i, q_{i'})) = \mathbb{P}(\mathcal{A}_f \mapsto q_i)\mathbb{P}(\mathcal{A}_g \mapsto q_{i'})$ pour tous $q_i \in Q^f, q_{i'} \in Q^g$
- $\mathbb{P}((q_i, q_{i'}) \xrightarrow[\mathcal{A}]{a} (q_j, q_{j'})) = \mathbb{P}(q_i \xrightarrow[\mathcal{A}_f]{a} q_j)\mathbb{P}(q_{i'} \xrightarrow[\mathcal{A}_g]{a} q_{j'})$ pour toute lettre a

d'où pour tout mot u : $\mathbb{P}(\mathcal{A}(u) = (q_i, q_{i'})) = \mathbb{P}(\mathcal{A}_f(u) = q_i)\mathbb{P}(\mathcal{A}_g(u) = q_{i'})$ pour tout i, i.e. $\Pi(u) = \Pi^f(u) \otimes \Pi^g(u)$ et finalement on vérifie que l'on a bien

$$\begin{aligned} P^{\mathcal{A}}(u) &= \Pi(u)\mathbb{1}_F = \left(\Pi^f(u) \otimes \Pi^g(u)\right)\left(\mathbb{1}_{F^f} \otimes \mathbb{1}_{F^g}\right) = \left(\Pi^f(u)\mathbb{1}_{F^f}\right)\left(\Pi^g(u)\mathbb{1}_{F^g}\right) \\ &= f(u)g(u). \end{aligned}$$

□

Somme pondérée
Soient $\mathcal{A}_f$, $\mathcal{A}_g$ et $\mathcal{A}_h$ des automates probabilistes de fonctions associées f, g et h. Alors il existe un automate probabiliste $\mathcal{A}$ tel que $P^{\mathcal{A}} = fh + g(1-h)$.

Preuve. On pose $\mathcal{A}_f = (Q^f, \pi^f, M^f, F^f)$, $\mathcal{A}_g = (Q^g, \pi^g, M^g, F^g)$ et
$\mathcal{A}_h = (Q^h, \pi^h, M^h, F^h)$. On définit alors $\mathcal{A}$ comme suit :
$\mathcal{A} = (Q^f \times Q^g \times Q^h, \pi^f \otimes \pi^g \otimes \pi^h, M^f \otimes M^g \otimes M^h, F)$
où $F = F^f \times Q^g \times F^h \sqcup Q^f \times F^g \times (Q^h \setminus F^h)$
On alors pour tout mot u :

$$\begin{aligned} P^{\mathcal{A}}(u) &= \mathbb{P}(\mathcal{A}(u) \in F) \\ &= \mathbb{P}(\mathcal{A}(u) \in F^f \times Q^g \times F^h) + \mathbb{P}(\mathcal{A}(u) \in Q^f \times F^g \times (Q^h \setminus F^h)) \\ &= \mathbb{P}(\mathcal{A}_f(u) \in F^f)\mathbb{P}(\mathcal{A}_h(u) \in F^h) + \mathbb{P}(\mathcal{A}_g(u) \in F^g)\mathbb{P}(\mathcal{A}_h(u) \notin F^h) \\ &= f(u)h(u) + g(u)(1-h(u)) \end{aligned}$$

□

Remarque : L'automate donné en exemple réalisant une fonction constante, si f et g sont des fonctions associées à des automates probabilistes, on peut d'après ce qui précède réaliser un automate probabiliste de fonction associée $\frac{1}{2}(f+g)$ en prenant $h(u)=\frac{1}{2}$. On pourra aussi réaliser la fonction $\frac{1}{2}(f+1-g)$.

3 Langages reconnus par un A.P.

Définition 5. Soit $\mathcal{A}$ un automate probabiliste et $\lambda \in [0,1]$. On définit $L(\mathcal{A},\lambda) := \{u \in A^* \mid P^{\mathcal{A}}(u) > \lambda\}$ l'ensemble des mots u acceptés avec une probabilité plus grande que λ. λ est appelé point de coupure.

Remarque : D'après la loi des grands nombres, si on prend un mot u tel que $P^{\mathcal{A}}(u) > \lambda$ et qu'on lance l'automate probabiliste $\mathcal{A}$ sur u un grand nombre de fois alors u sera presque sûrement reconnu par $\mathcal{A}$ comme appartenant à $L(\mathcal{A},\lambda)$.

Exemple : L'automate qui réalise la fonction $\frac{1}{2}(f+1-g)$ reconnaîtra avec le point de coupure $\lambda=\frac{1}{2}$ l'ensemble des mots u tels que $f(u) > g(u)$

Le théorème de Rabin
On commence par définir la notion de point de coupure ε-isolé.

Définition 6. Soient $\mathcal{A}$ un automate probabiliste et $\lambda \in [0,1]$. On dit que λ est un *point de coupure ε-isolé* pour $\mathcal{A}$ si il existe un $\varepsilon > 0$ tel que pour tout $u \in A^*$, $|P^{\mathcal{A}}(u) - \lambda| \geq \varepsilon$.

Théorème 7 (Rabin)**.** *Soit $\mathcal{A}$ un automate probabiliste et $\lambda \in [0,1]$ un point de coupure ε-isolé pour $\mathcal{A}$, alors le langage $L(\mathcal{A},\lambda)$ est rationnel. De plus il existe un automate fini déterministe reconnaissant $L(\mathcal{A},\lambda)$ avec moins de $\left(1+\frac{1}{2\varepsilon}\right)^{n-1}$ états.*

Pour la démonstration de ce théorème, nous aurons besoin d'un théorème vu en cours dû à Nerode ainsi que d'un petit lemme.

Théorème 8 (Nerode)**.** *Un langage L est rationnel si et seulement si l'ensemble $\{w^{-1}L \mid w \in A^*\}$ est fini. De plus les automates déterministes minimaux reconnaissant L ont exactement $\left|\{w^{-1}L \mid w \in A^*\}\right|$ états.*

Lemme 9. *Soit $\mathcal{P}_n$ l'ensemble des vecteurs probabilités de taille n et soit U_ε un sous-ensemble de $\mathcal{P}_n$ tel que pour tous vecteurs ξ et ξ' de U_ε distincts, $\sum_{i=1}^{n} |\xi_i - \xi'_i| \geq \varepsilon$. Alors U_ε possède au plus $\left(1+\frac{2}{\varepsilon}\right)^{n-1}$ éléments.*

Preuve. (La figure 1 ci-après illustre la démonstration dans le cas $n=3$) Le polytope $\mathcal{P}_n$ est contenu dans un hyperplan de dimension $n-1$, on pose Σ son volume. On considère à présent le polytope $(1+\frac{\varepsilon}{2})\mathcal{P}_n$, image de $\mathcal{P}_n$ par l'homothétie centrale de rapport $(1+\frac{\varepsilon}{2})$, son volume vaut donc $(1+\frac{\varepsilon}{2})^{n-1}\Sigma$. Pour chaque $\xi \in U_\varepsilon$, on considère le petit polytope V_ξ image de $\mathcal{P}_n$ par homothétie de rapport $\frac{\varepsilon}{2}$ puis translation de vecteur ξ : $V_\xi = \{\, \eta \mid \forall i \in [\![1,n]\!], \eta_i \geq \xi_i \text{ et } \sum_{i=1}^{n} \eta_i - \xi_i = \frac{\varepsilon}{2}\}$.

Ce polytope est inclus dans $(1+\frac{\varepsilon}{2})\mathcal{P}_n$, puisque pour tout $\eta \in V_\xi$:
$\sum_{i=1}^{n} \eta_i = \sum_{i=1}^{n} \xi_i + \sum_{i=1}^{n} \eta_i - \xi_i = 1 + \frac{\varepsilon}{2}$. De plus si $\xi, \xi' \in U_\varepsilon$, alors $V_\xi \cap V_{\xi'} = \emptyset$. En effet si on suppose par l'absurde qu'il existe $\eta \in V_\xi \cap V_{\xi'}$, alors on a : $\sum_{i=1}^{n} |\xi_i - \xi'_i| < \sum_{i=1}^{n} \eta_i - \xi_i + \sum_{i=1}^{n} \eta_i - \xi'_i = \varepsilon$: absurde. Ainsi les polytopes V_ξ pour $\xi \in U_\varepsilon$ sont deux à deux disjoints contenus dans $(1+\frac{\varepsilon}{2})\mathcal{P}_n$de volume $(1+\frac{\varepsilon}{2})^{n-1}\Sigma$ et ils sont tous de volume $(\frac{\varepsilon}{2})^{n-1}\Sigma$. Il y a donc au plus $\frac{(1+\varepsilon/2)^{n-1}\Sigma}{(\varepsilon/2)^{n-1}\Sigma} = (1+\frac{2}{\varepsilon})^{n-1}$ et U_ε possède donc au plus $(1+\frac{2}{\varepsilon})^{n-1}$ éléments. □

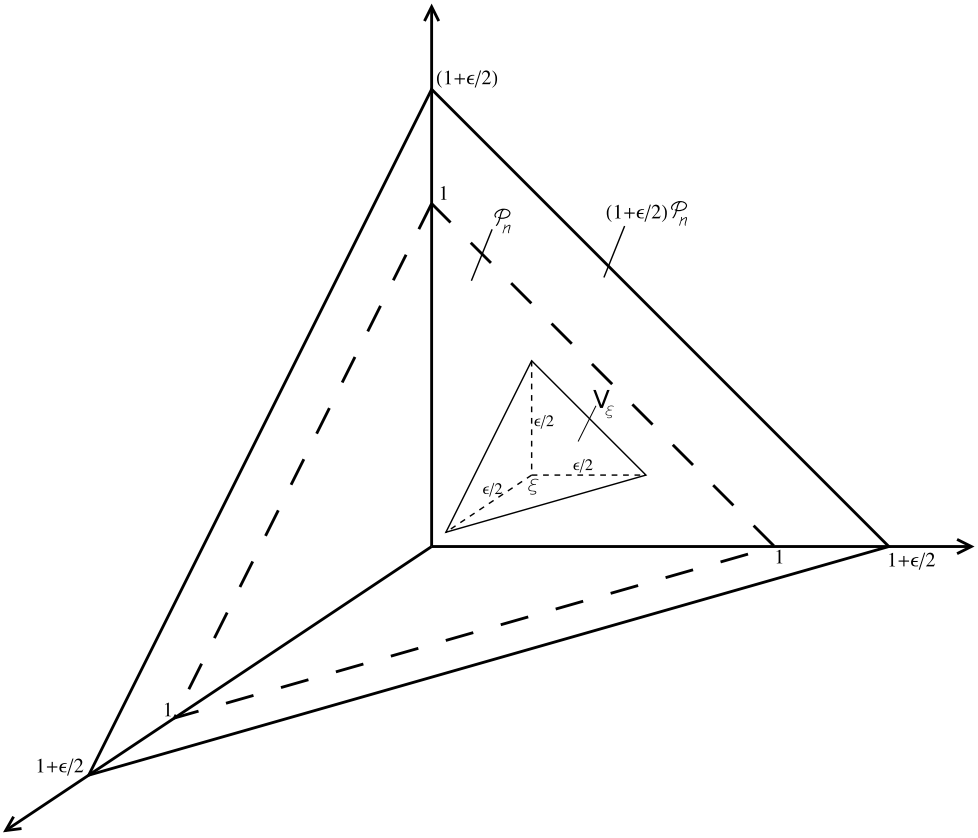


FIG. 1 – Exemple dans le cas $n = 3$

Remarque : L'inégalité précédente est bien stricte, en effet l'inégalité triangulaire $|a-b| < |a-c|+|c-b|$ peut être légèrement améliorée lorsque $c \leq a$ et $c \leq b$ avec au moins une des deux inégalités strictes. En effet on peut supposer sans nuire à la généralité que $a \leq b$ et on a alors : $|a-b| = b-a \leq c-a \leq c-a+b-a$ avec au moins une des deux égalités strictes même quand $a = b < c$ ou $a < b = c$. Ainsi, si dans la preuve l'inégalité n'est pas nécessairement vraie pour chacun des termes de la somme, elle l'est pour au moins un terme.

Preuve du théorème de Rabin. Soit $\mathcal{A}$ un automate probabiliste et $\lambda \in [0,1]$ un point de coupure ε-isolé pour $\mathcal{A}$. On pose $L = L(\mathcal{A}, \lambda)$. On dira que deux mots u et v sont équivalents si $u^{-1}L = v^{-1}L$. L'idée consiste à montrer grâce au lemme qu'il existe un nombre fini de classes d'équivalences pour en déduire par le théorème dû à Nerode que L est rationnel. Soient u et v deux mots non équivalents, c'est-à-dire qu'il existe un mot z tel que $uz \in L$ et $vz \notin L$ ou inversement $uz \notin L$ et $vz \in L$ (sans nuire à la généralité, on se placera dans le premier cas). Dans le cas d'un automate probabiliste cela se traduit par $P^{\mathcal{A}}(uz) > \lambda$ et $P^{\mathcal{A}}(vz) \leq \lambda$.

On pose alors $\chi(z) = M_{z_1}...M_{z_r}\mathbb{1}_F$, de sorte que $P^{\mathcal{A}}(uz) = \Pi(u)\chi(z)$ et $P^{\mathcal{A}}(vz) = \Pi(v)\chi(z)$. On a alors $\Pi(u)\chi(z) > \lambda$ et $\Pi(v)\chi(z) \leq \lambda$ d'où $(\Pi(u) - \pi(v))\chi(z) \geq 2\varepsilon$ puisque λ est un point de coupure ε-isolé. On a donc :

$$\sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))\chi_i(z) \geq 2\varepsilon$$

Or,

$$\begin{aligned}
\sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))\chi_i(z) &\leq \left(\sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^+\right) \max_{i\in[\![1,n]\!]}\chi_i(z) \\
&\qquad + \left(\sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^-\right) \min_{i\in[\![1,n]\!]}\chi_i(z) \\
&= \left(\sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^+\right)(\max\chi_i(z) - \min\chi_i(z)) \\
&\leq \sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^+ \\
&= \frac{1}{2}\sum_{i=1}^{n}|\Pi_i(u) - \Pi_i(v)|
\end{aligned}$$

[|Explications : On a ici noté $x = (x)^+ + (x)^-$ avec $(x)^+ \geq 0$ et $(x)^- \leq 0$.
On a de plus utilisé deux fois la relation suivante :

$$\begin{aligned}
0 &= 1 - 1 \\
&= \sum_{i=1}^{n}\Pi_i(u) - \sum_{i=1}^{n}\Pi_i(v) \\
&= \sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v)) \\
&= \sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^+ + \sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^-
\end{aligned}$$

d'où :

$$\textstyle\sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^+ = -\sum_{i=1}^{n}(\Pi_i(u) - \Pi_i(v))^-$$

et donc :

$$\sum_{i=1}^{n} |\Pi_i(u) - \Pi_i(v)| = \sum_{i=1}^{n} (\Pi_i(u) - \Pi_i(v))^+ - \sum_{i=1}^{n} (\Pi_i(u) - \Pi_i(v))^-$$
$$= 2 \sum_{i=1}^{n} (\Pi_i(u) - \Pi_i(v))^+ \qquad |]$$

Finalement, on obtient :

$$2\epsilon \leq \frac{1}{2} \sum_{i=1}^{n} |\Pi_i(u) - \Pi_i(v)| \quad \text{i.e.} \quad \sum_{i=1}^{n} |\Pi_i(u) - \Pi_i(v)| \geq 4\epsilon$$

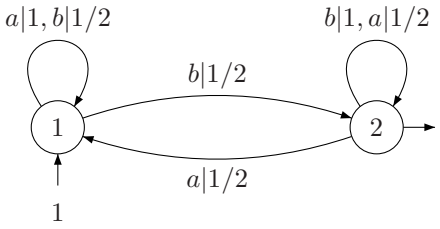
On remarque alors que l'ensemble des $\Pi(u)$ ou u parcours l'ensemble des classes d'équivalence des mots de A^* est un ensemble $U_{4\varepsilon}$ qui vérifie les hypothèses du lemme, son application nous indique donc que cet ensemble est fini de cardinal $\leq (1 + \frac{1}{2\varepsilon})^{n-1}$, puis que l'ensemble des classes d'équivalence lui même est fini de cardinal $\leq (1 + \frac{1}{2\varepsilon})^{n-1}$, et enfin par application du théorème que L est rationnel et qu'il est reconnu par un automate avec moins de $(1 + \frac{1}{2\varepsilon})^{n-1}$ états. □

Conclusion :
Le théorème de Rabin est un des théorèmes les plus intéressants de la théorie des automates probabilistes. On aurait pu démontrer que L est rationnel très rapidement en prenant des majorations grossières, mais la borne du nombre d'états de l'automate minimal aurait été plus grande. Celle que nous venons de donner n'est cependant pas optimale, et une amélioration de la borne du lemme permettrait de la diminuer. En 1971, la recherche d'une borne optimale était un problème ouvert, tout comme la recherche d'un algorithme permettant de déterminer si une probabilité est un point de coupure isolé.

Si la démonstration du théorème n'est pas constructive, on peut cependant en déduire facilement un algorithme (très long) qui permettrait de calculer un automate déterministe équivalent : Il suffit de trouver toutes les classes d'équivalence en faisant une recherche exhaustive sur tout les mots de longeur $\leq 2(1 + \frac{1}{2\varepsilon})^{n-1}$ puis de construire un automate en appliquant la démonstration du théorème dû à Rabin.

4 Compléments

Pour terminer, on va donner un exemple d'automate probabiliste fort intéressant qui permettra de répondre à toutes les questions que l'on peut se poser après avoir énoncé le théorème de Rabin : Existe-t-il des langages non rationnels reconnus par des automates probabilistes ? Est-ce qu'un point de coupure non isolé engendre un langage non rationnel ? Existe-t-il des points de coupure non isolés ? Pour répondre à tout cela, considérons l'automate probabiliste suivant : $\mathcal{A} = (\{1, 2\}, (1, 0), M, \{1\})$ avec $Ma = \begin{pmatrix} 1 & 1/2 \\ 1 & 1/2 \end{pmatrix}$ et $Mb = \begin{pmatrix} 1/2 & 1 \\ 1/2 & 1 \end{pmatrix}$.

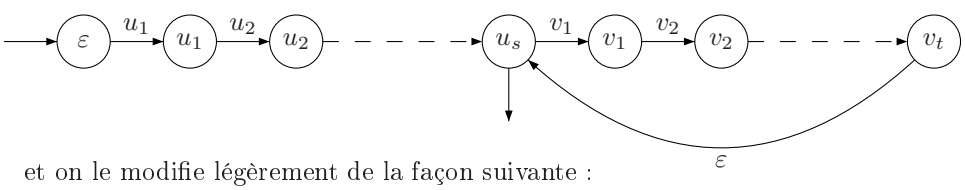


Fonctionnement de l'automate : On remarque que pour tout mot u : $P^{\mathcal{A}}(ua) = \frac{1}{2}P^{\mathcal{A}}(u)$ et $P^{\mathcal{A}}(ub) = P^{\mathcal{A}}(u) + \frac{1}{2}(1 - P^{\mathcal{A}}(u)) = \frac{1}{2} + \frac{1}{2}(P^{\mathcal{A}}(u))$. On constate alors que si on identifie a avec 0 et b avec 1, on a si $u = a_1a_2...a_m$: $P^{\mathcal{A}}(u) = \sum_{i=0}^{m-1} \frac{a_{m-i}}{2^{i+1}} = a_1a_2...a_m0_{\frac{1}{2}}$: le mot ub est l'écriture en base 1/2, poids faibles à droite, de $P^{\mathcal{A}}(u)$: c'est un nombre dyadique. On obtient alors une bijection entre les mots de A^* et les nombres dyadiques de $[0, 1]$. Or, on sait que l'ensemble des nombres dyadiques est dense dans $[0, 1]$, ce qui entraîne de nombreuses conclusions.

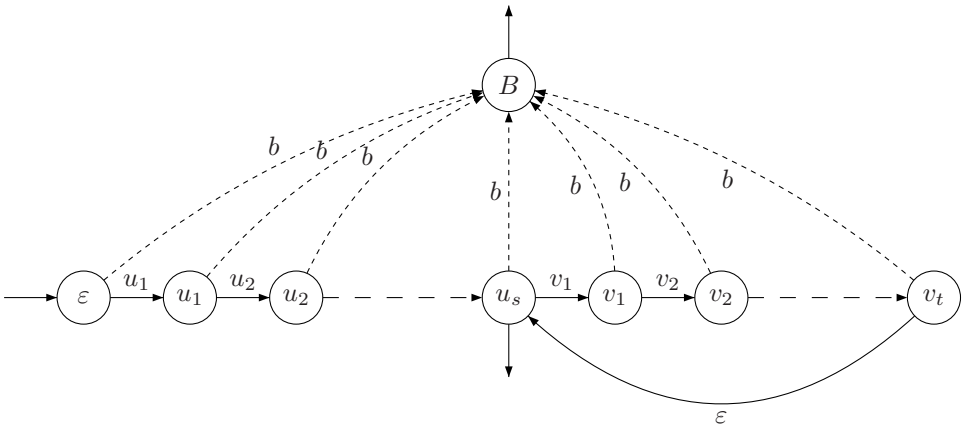
Conclusions : La densité des nombres dyadiques dans $[0, 1]$ entraîne que pour cet automate :

- Quel que soit λ dans $[0, 1]$ point de coupure, λ n'est pas isolé.
- Si λ et λ' sont des points de coupure avec $\lambda > \lambda'$, alors il existe un mot u tel que $\lambda > P^{\mathcal{A}}(u) > \lambda'$ et donc $L(\mathcal{A}, \lambda) \neq L(\mathcal{A}, \lambda')$
- On obtient ainsi autant de langages deux à deux distincts reconnaissables par un automate probabiliste que de réels dans $[0, 1]$, soit une quantité indénombrable. Or, les langages rationnels sont dénombrables, donc cet automate reconnaît des langages non rationnels.
- On va enfin montrer que $L(\mathcal{A}, \lambda)$ est un langage rationnel si et seulement si λ est rationnel, ce qui assurera que notre automate reconnaît des langages rationnels et des langages non rationnels avec des points de coupure non isolés, et qui justifiera un peu plus le fait que les langages rationnels ont de bonnes raisons d'être "rationnels".

Preuve : On sait qu'un langage L est rationnel ssi son langage miroir $\tilde{L}$ l'est, on va donc montrer que λ est rationnel ssi $\widetilde{L(\mathcal{A}, \lambda)}$ l'est. On rappelle qu'un nombre réel est rationnel ssi sont développement dyadique est ultimement périodique (résultat classique).
Si λ est rationnel : On sait alors que son développement dyadique est ultimement périodique : il existe donc des mots u,v sur l'alphabet $\{0, 1\}$ tels que $\lambda = {}_{\frac{1}{2}}uv^*$. On considère alors l'automate trivial reconnaissant uv^*, i.e. si $u = u_1...u_s$ et $v = v_1...v_t$, on considère l'automate suivant :



et on le modifie légèrement de la façon suivante :



en ajoutant un état final B et une transition $u_i \xrightarrow{b} B$ ou $v_i \xrightarrow{b} B$ pour chaque i tel que $u_{i+1} = a$. De cette façon l'automate reconnaîtra tous les mots w tels que $_{\frac{1}{2}}w >_{\frac{1}{2}} uv^* = \lambda$, et il reconnaîtra donc $\widetilde{L(\mathcal{A}, \lambda)}$.
Réciproquement : si λ n'est pas rationnel, alors son développement dyadique s'écrit $\lambda =_{\frac{1}{2}} v$ avec $v = a_1 a_2 ... a_m ...$ un mot infini qui n'est pas ultimement périodique. On voit alors que pour tout $i \neq j$, on a $a_{i+1}a_{i+2}... \neq a_{j+1}a_{j+2}...$, car sinon v serait ultimement périodique. Sans nuire à la généralité, on peut supposer $a_{i+1}a_{i+2}... < a_{j+1}a_{j+2}...$ et donc il existe un entier k tel que $a_{i+1}a_{i+2}...a_{i+k} < a_{j+1}a_{j+2}...a_{j+k}$ et tel que $a_{i+k+1} = a$ (Sinon, v serait de la forme ub^*). On voit alors en posant $L = L(\mathcal{A}, \lambda)$ que $a_{i+1}...a_{i+k}b \in (a_1...a_i)^{-1}L$ mais que $a_{i+1}...a_{i+k}b < a_{j+1}...a_{j+k}$ et donc que $a_{i+1}...a_{i+k}b \notin (a_1...a_j)^{-1}L$. Finalement, pour tous $i \neq j$, $(a_1...a_i)^{-1}L \neq (a_1...a_j)^{-1}L$, et donc L n'est pas rationnel.

□

Références

[1] Azaria Paz. *Introduction to Probabilistic Automata.* Computer Science and Applied Mathematics. Academic Press, 1971.