Automates probabilistes

Nathalie Bertrand*, Serge Haddad**

*Inria Rennes-Bretagne Atlantique
**LSV, ENS Cachan & CNRS & Inria

EJCIM 2015, Orléans

Plan

Un exemple introductif

Planifier ses vacances à l'étranger

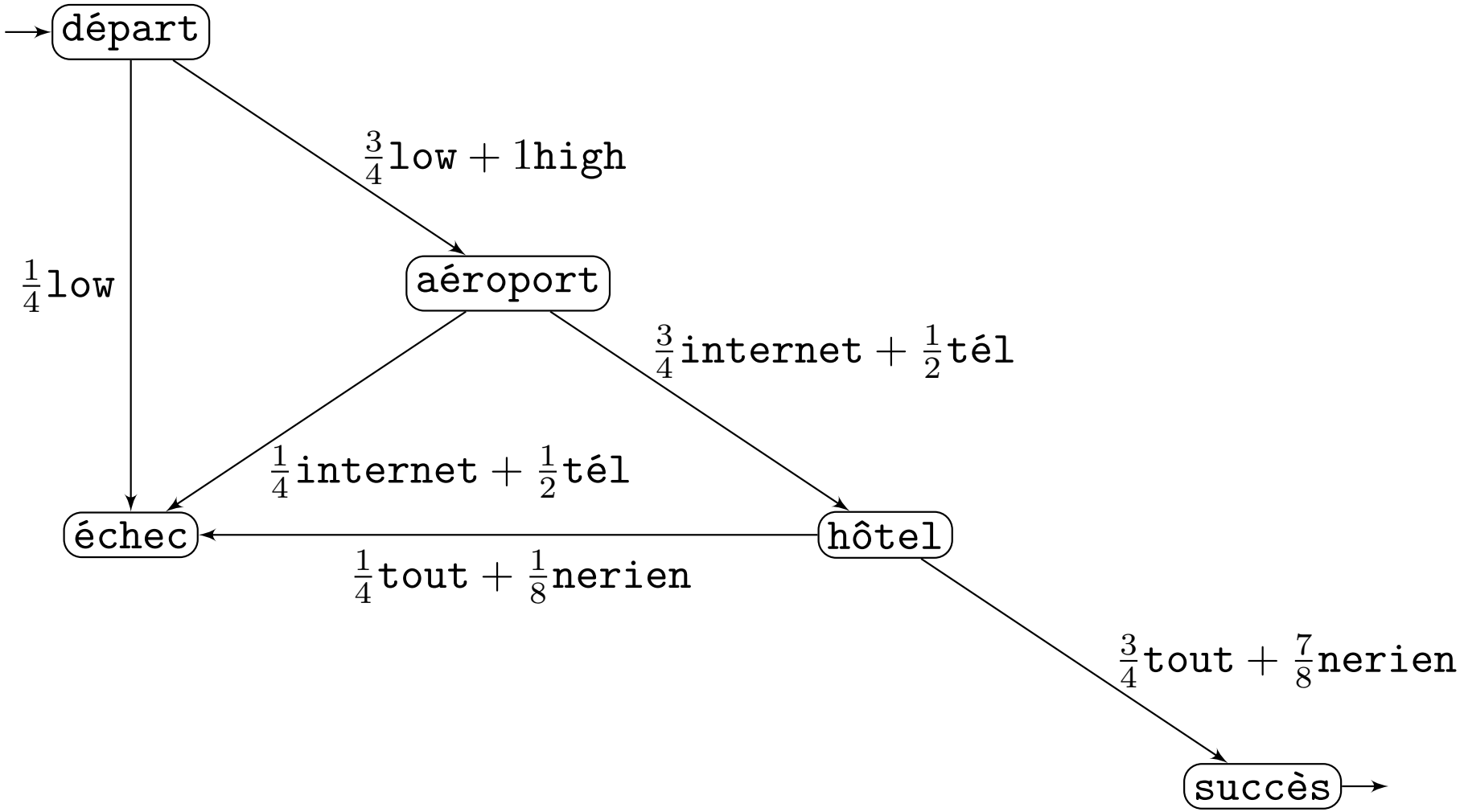
1. Choisir une compagnie d'avion `lowcost` ou `highcost` ;
2. Réserver un hôtel par `internet` ou par `téléphone` ;
3. Choisir une excursion avec l'agence `toutvoir` ou `nerienrater`.

Chacune de ces actions

1. doit être planifiée avant les vacances ;
2. a une probabilité d'échouer.

Un plan possible : `lowcost` · `internet` · `toutvoir`

Formalisation de l'exemple



La probabilité de réussite de lowcost · internet · toutvoir est égale à $\frac{27}{64}$.

Automates probabilistes (PA)

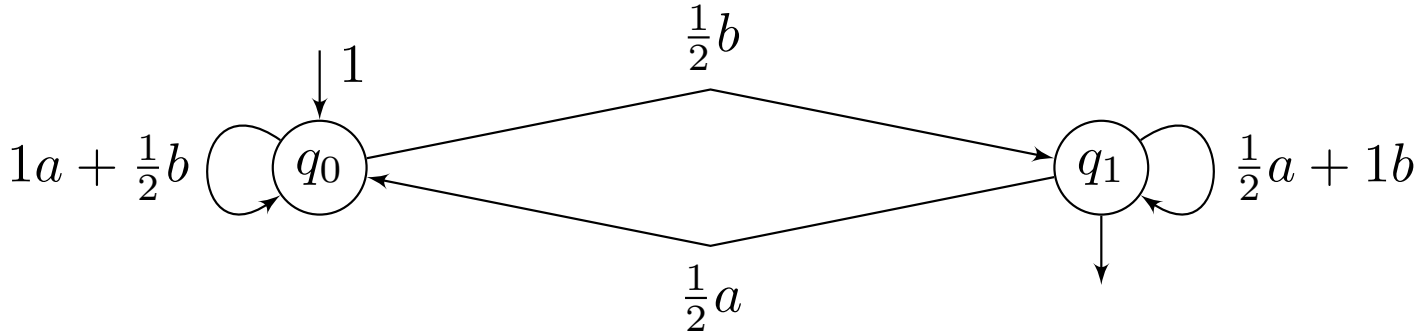
(M. O. Rabin. Information and Control 1963)

(Prix Turing 1976, avec D. Scott pour les machines non déterministes)

Un automate probabiliste $\mathcal{A} = (Q, A, \{\mathbf{P}_a\}_{a \in A}, \pi_0, F)$ est défini par :

- Q, un ensemble fini d'états ;
- A, un alphabet fini d'actions ;
- Pour tout $a \in A$, $\mathbf{P}_a$, une matrice stochastique indicée par Q i.e. pout tout $q, q' \in Q$, $\mathbf{P}_a[q, q'] \geq 0$ et $\sum_{q' \in Q} \mathbf{P}_a[q, q'] = 1$;
- π_0, la distribution initiale des états ;
- $F \subseteq Q$, un sous-ensemble d'états finals.

Illustration



- $A = \{a, b\}$;
- $Q = \{q_0, q_1\}$, $F = \{q_1\}$;
- $\pi_0[q_0] = 1$.

Un arc est étiqueté par un vecteur de probabilités indicé par A.
Ce vecteur est noté comme une somme formelle.

Par exemple, la boucle autour de q_0 est étiquetée par $1a + 0.5b$ signifiant que :

- lorsque a est choisi en q_0,
  la probabilité que le prochain état soit q_0, $\mathbf{P}_a[q_0, q_0]$, est égale à 1.
- lorsque b est choisi en q_0,
  la probabilité que le prochain état soit q_0, $\mathbf{P}_b[q_0, q_0]$, est égale à 0.5.

Stratégies dans les PA

Les stratégies sont les mots. Soit un mot $w \stackrel{\text{def}}{=} a_1 \dots a_n$.
Quelle est la probabilité d'être dans un état final en appliquant la stratégie w ?

Soient $\mathcal{A}$ un PA et $w \stackrel{\text{def}}{=} a_1 \dots a_n \in A^*$ un mot,
la *probabilité d'acceptation* de w par $\mathcal{A}$ est définie par :

$$\mathbf{Pr}_{\mathcal{A}}(w) \stackrel{\text{def}}{=} \sum_{q \in Q} \pi_{\mathbf{0}}[q] \sum_{q' \in F} \left(\prod_{i=1}^{n} \mathbf{P}_{a_i} \right) [q, q']$$

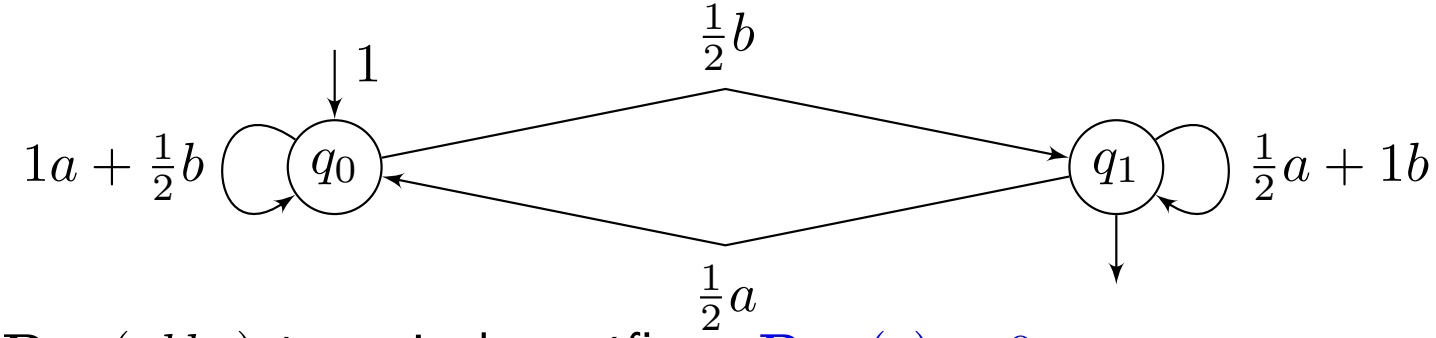
Notation. La matrice stochastique $\mathbf{P}_w$ est définie par $\mathbf{P}_w \stackrel{\text{def}}{=} \prod_{i=1}^{n} \mathbf{P}_{a_i}$.
En particulier $\mathbf{P}_\varepsilon = \mathbf{Id}$.

Avec ces notations :

$$\mathbf{Pr}_{\mathcal{A}}(w) = \pi_0 \mathbf{P}_w \mathbf{1}_F^T$$

où $\mathbf{1}_F$ est le vecteur indicateur du sous-ensemble F.

Illustration



Calcul de $\mathbf{Pr}_{\mathcal{A}}(abba)$ à partir des préfixes. $\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$.

- $\mathbf{Pr}_{\mathcal{A}}(a) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$
- $\mathbf{Pr}_{\mathcal{A}}(ab) = \mathbf{Pr}_{\mathcal{A}}(a) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(a)) = \frac{1}{2}$
- $\mathbf{Pr}_{\mathcal{A}}(abb) = \mathbf{Pr}_{\mathcal{A}}(ab) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(ab)) = \frac{3}{4}$
- $\mathbf{Pr}_{\mathcal{A}}(abba) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(abb) = \frac{3}{8}$

Plus généralement, les équations suivantes sont satisfaites :

$$\mathbf{Pr}_{\mathcal{A}}(wa) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(w) \text{ and } \mathbf{Pr}_{\mathcal{A}}(wb) = \frac{1}{2}(1 + \mathbf{Pr}_{\mathcal{A}}(w))$$

On en déduit une expression explicite de la probabilité d'acceptation :

$$\mathbf{Pr}_{\mathcal{A}}(a_1 \dots a_n) = \sum_{i=1}^{n} 2^{i-n-1} \cdot \mathbf{1}_{a_i=b}$$

Quel mot maximise la probabilité d'acceptation ?

Langages stochastiques

Caractérisation des « bonnes » stratégies.

Ceci conduit à l'introduction des *langages stochastiques*. Soit :

- $\mathcal{A}$ un automate probabiliste ;
- $\theta \in [0, 1]$ un *seuil* ;
- $\bowtie \in \{<, \leq, >, \geq, =, \neq\}$ un opérateur.

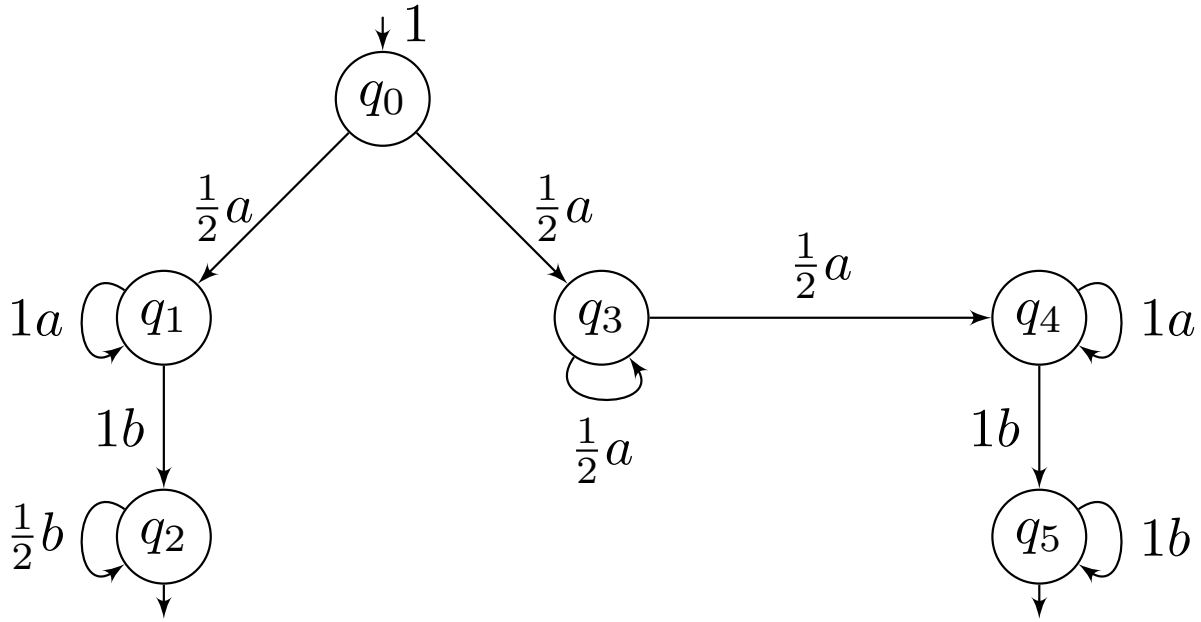
$L_{\bowtie\theta}(\mathcal{A})$ est définie par :

$$L_{\bowtie\theta}(\mathcal{A}) = \{w \in A^* \mid \mathbf{Pr}_{\mathcal{A}}(w) \bowtie \theta\}$$

Pour des questions d'expressivité et de décidabilité,
on introduit les définitions suivantes.

- Un PA *rationnel* est un PA dont les probabilités appartiennent à $\mathbb{Q}$.
- Un *langage stochastique rationnel* est un langage stochastique spécifié par un PA rationnel et un seuil rationnel.

Un PA qui compte



(une représentation succincte avec un état absorbant et rejetant omis)

Un mot z différent de $a^m b^n$ avec $m > 0, n > 0$ ne peut être accepté.

Soit $w \stackrel{\text{def}}{=} a^m b^n$ avec $m > 0, n > 0$. w peut être accepté par :

- un chemin q_0, q_1^m, q_2^n de probabilité $\frac{1}{2^n}$;
- ou une famille de chemins q_0, q_3^r, q_4^s, q_5^n avec $0 < r, s$ et $r + s = m$ de probabilité totale $\frac{1}{2} - \frac{1}{2^m}$.

En sommant : $\frac{1}{2} + \frac{1}{2^n} - \frac{1}{2^m}$.

D'où : $\mathcal{L}_{=\frac{1}{2}}(\mathcal{A}) = \{a^n b^n \mid n > 0\}$

Plan

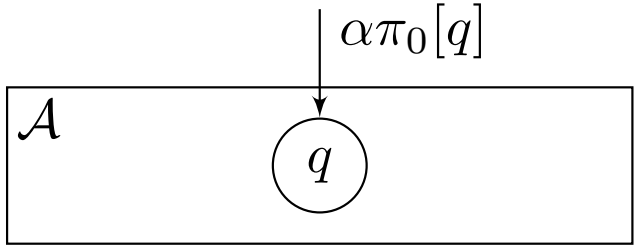
Problèmes d'expressivité

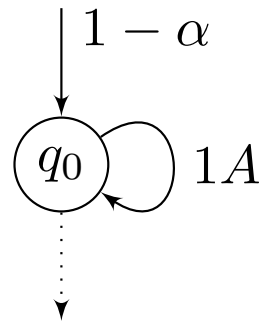
Fournir un ensemble minimal d'opérateurs de comparaison et de seuils.

Positionner les langages stochastiques vis à vis de la hiérarchie de Chomsky.

Etudier les propriétés de clôture des langages stochastiques.

Un unique seuil suffit





La valeur α dépend de $\theta \neq \frac{1}{2}$ comme suit :

- Si $\theta > \frac{1}{2}$ alors $q_0 \notin F$ et $\alpha \stackrel{\text{def}}{=} \frac{1}{2\theta}$ ce qui entraîne pour $w \in A^*$,

  $$\mathbf{Pr}_{\mathcal{A}'}(w) = \frac{1}{2\theta}\mathbf{Pr}_{\mathcal{A}}(w)$$

  Donc $w \in L_{\bowtie \frac{1}{2}}(\mathcal{A}')$ ssi $w \in L_{\bowtie \theta}(\mathcal{A})$.

- Si $\theta < \frac{1}{2}$ alors $q_0 \in F$ et $\alpha \stackrel{\text{def}}{=} \frac{1}{2(1-\theta)}$ ce qui entraîne pour $w \in A^*$,

  $$\mathbf{Pr}_{\mathcal{A}'}(w) = \frac{1-2\theta+\mathbf{Pr}_{\mathcal{A}}(w)}{2(1-\theta)}$$

  Donc $w \in L_{\bowtie \frac{1}{2}}(\mathcal{A}')$ ssi $w \in L_{\bowtie \theta}(\mathcal{A})$.

Elimination de l'égalité et de l'inégalité

Soit $\mathcal{A}$ un PA, $\mathcal{A}'$ est défini ainsi.

- L'ensemble des états est $Q' \stackrel{\text{def}}{=} Q \times Q$;
- $\mathbf{P}'_a[(q_1, q_2), (q'_1, q'_2)] \stackrel{\text{def}}{=} \mathbf{P}_a[q_1, q'_1]\mathbf{P}_a[q_2, q'_2]$;
- $\pi'_0[q_1, q_2] \stackrel{\text{def}}{=} \pi_0[q_1]\pi_0[q_2]$ et $F' \stackrel{\text{def}}{=} F \times (Q \setminus F)$.

Pour un mot w fixé,
les deux composants des états se comportent indépendamment :

$$\mathbf{Pr}_{\mathcal{A}'}(w) = \mathbf{Pr}_{\mathcal{A}}(w)(1 - \mathbf{Pr}_{\mathcal{A}}(w))$$

Aussi $\mathbf{Pr}_{\mathcal{A}'}(w) \leq \frac{1}{4}$ avec égalité ssi $\mathbf{Pr}_{\mathcal{A}}(w) = \frac{1}{2}$. D'où :

$$L_{\geq \frac{1}{4}}(\mathcal{A}') = L_{=\frac{1}{2}}(\mathcal{A})$$

Elimination des opérateurs $<$ et $\leq$

Soit $\mathcal{A}$ un PA, $\mathcal{A}'$ est obtenu en complémentant les états finals. Donc :

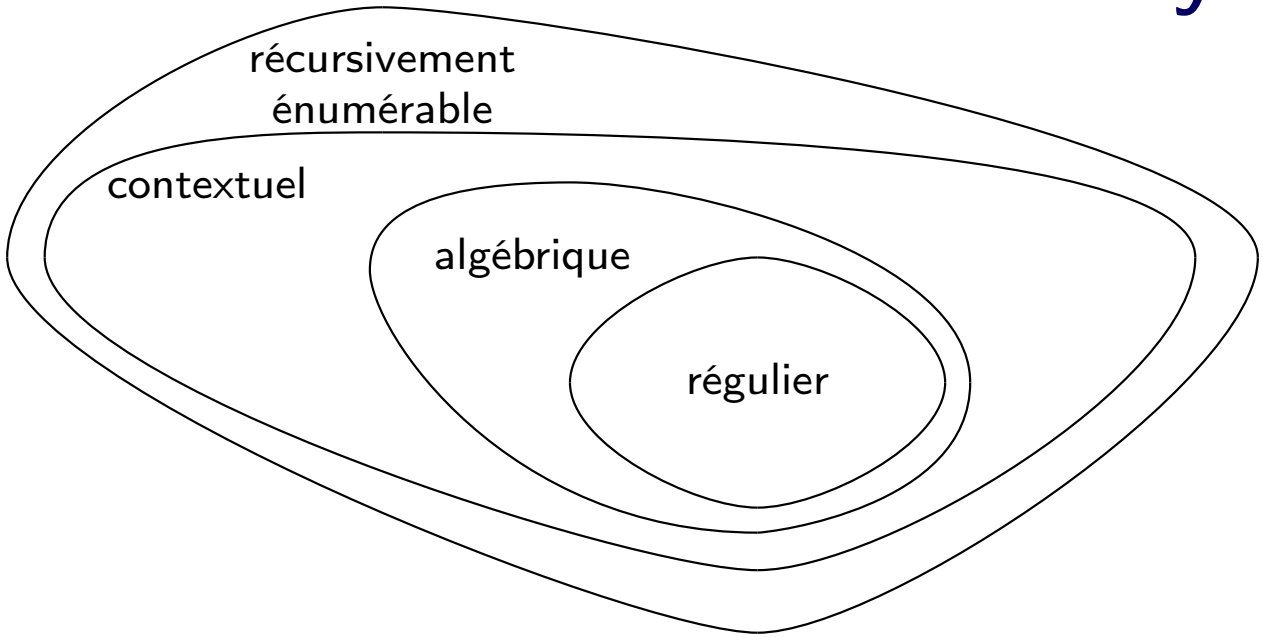
$$\mathbf{Pr}_{\mathcal{A}'}(w) = 1 - \mathbf{Pr}_{\mathcal{A}}(w)$$

D'où :

$$L_{\geq\theta}(\mathcal{A}') = L_{<\theta}(\mathcal{A})$$

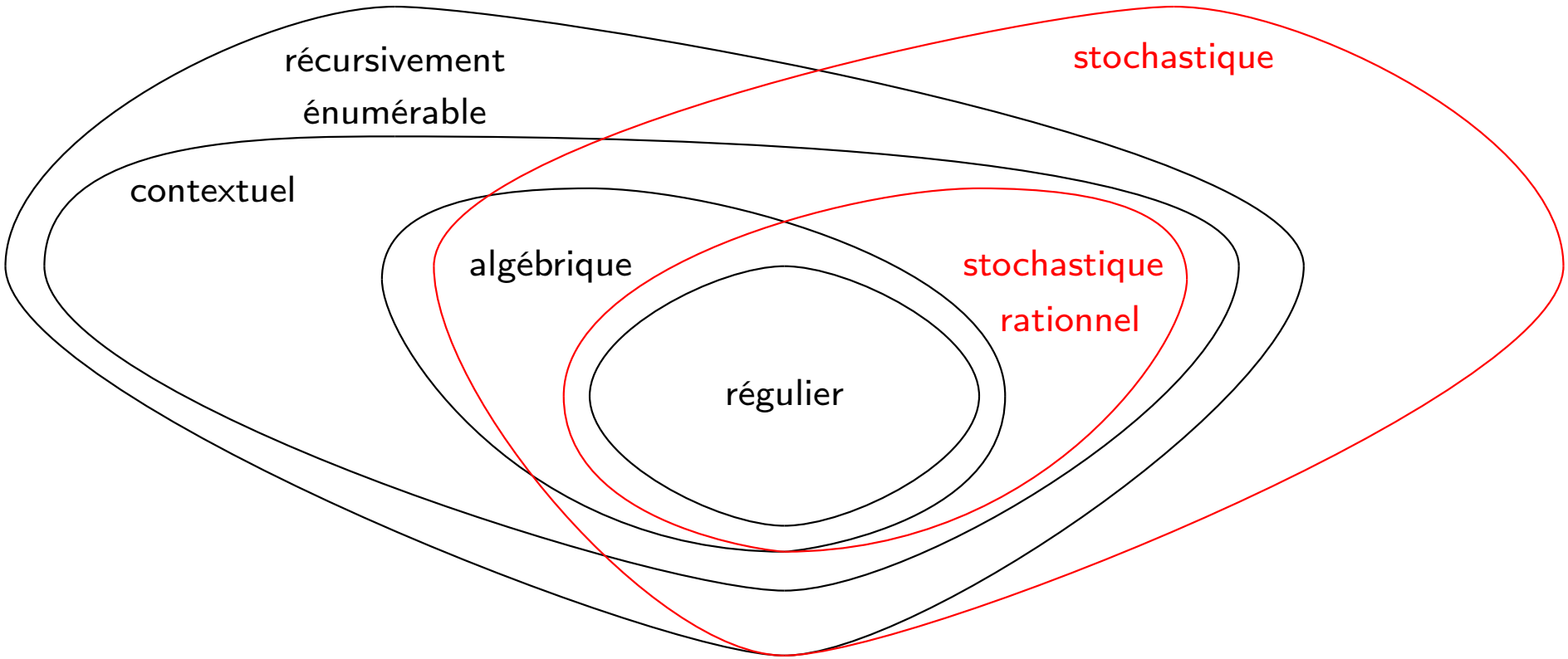
$$L_{>\theta}(\mathcal{A}') = L_{\leq\theta}(\mathcal{A})$$

La hiérarchie de Chomsky

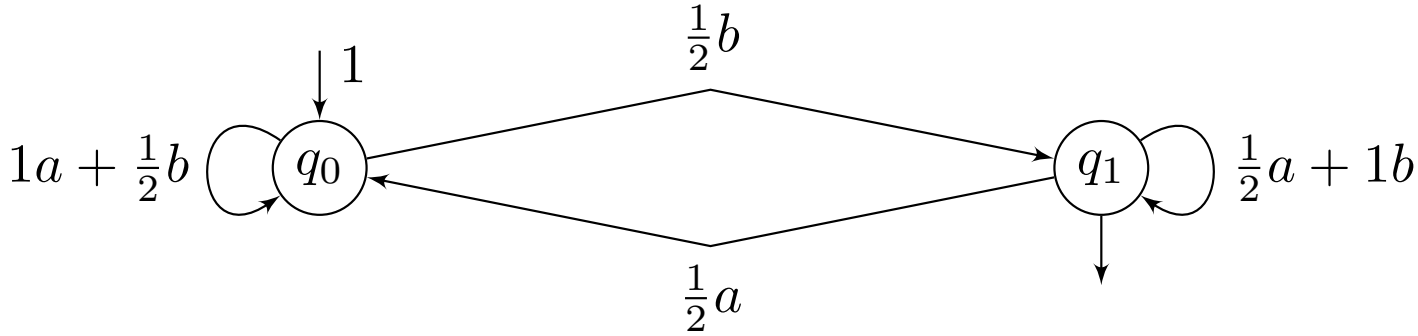


Classe	Grammaire	Machine
Langage régulier	$L \to aR \vert a \vert \varepsilon$ avec $L, R \in \Delta$ et $a \in \Sigma$	Automate fini
Langage algébrique	$L \to R_1 \dots R_n$ avec $L \in \Delta$ et $R_i \in \Delta \cup \Sigma$	Automate à pile
Langage contextuel	$L_1 \dots L_m \to R_1 \dots R_n$ $m \leq n$, $(S \to \varepsilon)$ avec $L_i, R_j \in \Delta \cup \Sigma$	Machine de Turing non déterministe opérant en espace linéaire
Langage récursivement énumérable	$L_1 \dots L_m \to R_1 \dots R_n$ avec $L_i, R_j \in \Delta \cup \Sigma$	Machine de Turing

La hiérarchie de Chomsky complétée



Langages non recursivement énumérables



Soient $v_a \stackrel{\text{def}}{=} 0$ and $v_b \stackrel{\text{def}}{=} 1$.

La probabilité d'acceptation de $w_1 \dots w_n$ est le nombre binaire $0.v_{w_n} \dots v_{w_1}$.

Aussi $\mathcal{L}_{>\theta}(\mathcal{A})$ est l'ensemble des représentations des nombres (avec développement binaire fini) plus grand que θ.

Aussi pour $0 \leq \theta < \theta' \leq 1$,

$$\mathcal{L}_{>\theta'}(\mathcal{A}) \subsetneq \mathcal{L}_{>\theta}(\mathcal{A})$$

Le cardinal des langages stochastiques est au delà du dénombrable impliquant que la « plupart » de ces langages ne sont pas recursivement énumérables.

Ce résultat n'est pas valable pour les langages stochastiques rationnels.

Langages réguliers

Un automate deterministe est un automate stochastique
avec probabilités dans $\{0, 1\}$.

Donc les langages réguliers sont des langages stochastiques rationnels.

Le langage $\{a^n b^n \mid n > 0\}$ est un langage rationnel stochastique non régulier.

Langages algébriques non stochastiques

$L \stackrel{\text{def}}{=} \{a^{n_1} b a^{n_2} b \dots a^{n_k} b a^* \mid \exists i > 1 \; n_i = n_1\}$
est un langage algébrique non stochastique.

Preuve.

L algébrique : reconnu par un automate non deterministe avec un compteur.

- On compte n_1, le nombre de a jusqu'à la première occurrence de b.
- Puis on devine une occurrence de b et on décremente le compteur par les occurrences de a jusqu'à la prochaine occurrence de b.
- Si le compteur est nul, le mot est accepté.

Supposons (1) $L = L_{>\theta}(\mathcal{A})$ ou (2) $L = L_{\geq\theta}(\mathcal{A})$.

Soit $\sum_{i=0}^{n} c_i x^i$ le polynôme minimal de $\mathbf{P}_a$.

Puisque 1 est valeur propre de $\mathbf{P}_a$, on obtient $\sum_{i=0}^{n} c_i = 0$
avec des c_i positifs et negatifs.

Par définition, $\sum_{i=0}^{n} c_i \mathbf{P}_{a^i} = 0$. Donc pour tout w,

$$\sum_{i=0}^{n} c_i \mathbf{P}_{a^i w} = \left(\sum_{i=0}^{n} c_i \mathbf{P}_{a^i} \right) \mathbf{P}_w = 0$$

Langages algébriques non stochastiques

Preuve (suite).

Soit $Pos = \{i \mid 0 \leq i \leq n \wedge c_i > 0\}$ et $NonPos = \{i \mid 0 \leq i \leq n \wedge c_i \leq 0\}$.

Notons Pos par $\{i_1, \ldots, i_k\}$.

Fixons $w \stackrel{\text{def}}{=} ba^{i_1}b \ldots ba^{i_k}b$.

Cas $L = L_{>\theta}(\mathcal{A})$**.** Soit $0 \leq i \leq n$, par définition de L,

$$\pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T > \theta \text{ ssi } i \in \{i_1, \ldots, i_k\}$$

Aussi :

$0 = \sum_{i=0}^{n} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T = \sum_{i \in Pos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T + \sum_{i \in NonPos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T$
$> (\sum_{i \in Pos} c_i)\theta + (\sum_{i \in NonPos} c_i)\theta = (\sum_{i=0}^{n} c_i)\theta = 0$

conduisant à une contradiction.

Cas $L = L_{\geq\theta}(\mathcal{A})$**.** Soit $0 \leq i \leq n$, par définition de L,

$$\pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T \geq \theta \text{ ssi } i \in \{i_1, \ldots, i_k\}$$

Aussi : $0 = \sum_{i=0}^{n} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T = \sum_{i \in Pos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T + \sum_{i \in NonPos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T$
$>(\sum_{i \in Pos} c_i)\theta + (\sum_{i \in NonPos} c_i)\theta = (\sum_{i=0}^{n} c_i)\theta = 0$

conduisant à une contradiction.

Langages stochastiques non algébriques

$$L \stackrel{\text{def}}{=} \{a^n b^n c^n \mid n > 0\}$$

est un langage stochastique rationnel non algébrique.

Preuve.

A l'aide du lemme d'Ogden, on prouve facilement que L n'est pas algébrique.

Observons que $L = L_1 \cap L_2$ avec $L_1 \stackrel{\text{def}}{=} \{a^n b^n c^+ \mid n > 0\}$ et $L_2 \stackrel{\text{def}}{=} \{a^+ b^n c^n \mid n > 0\}$.

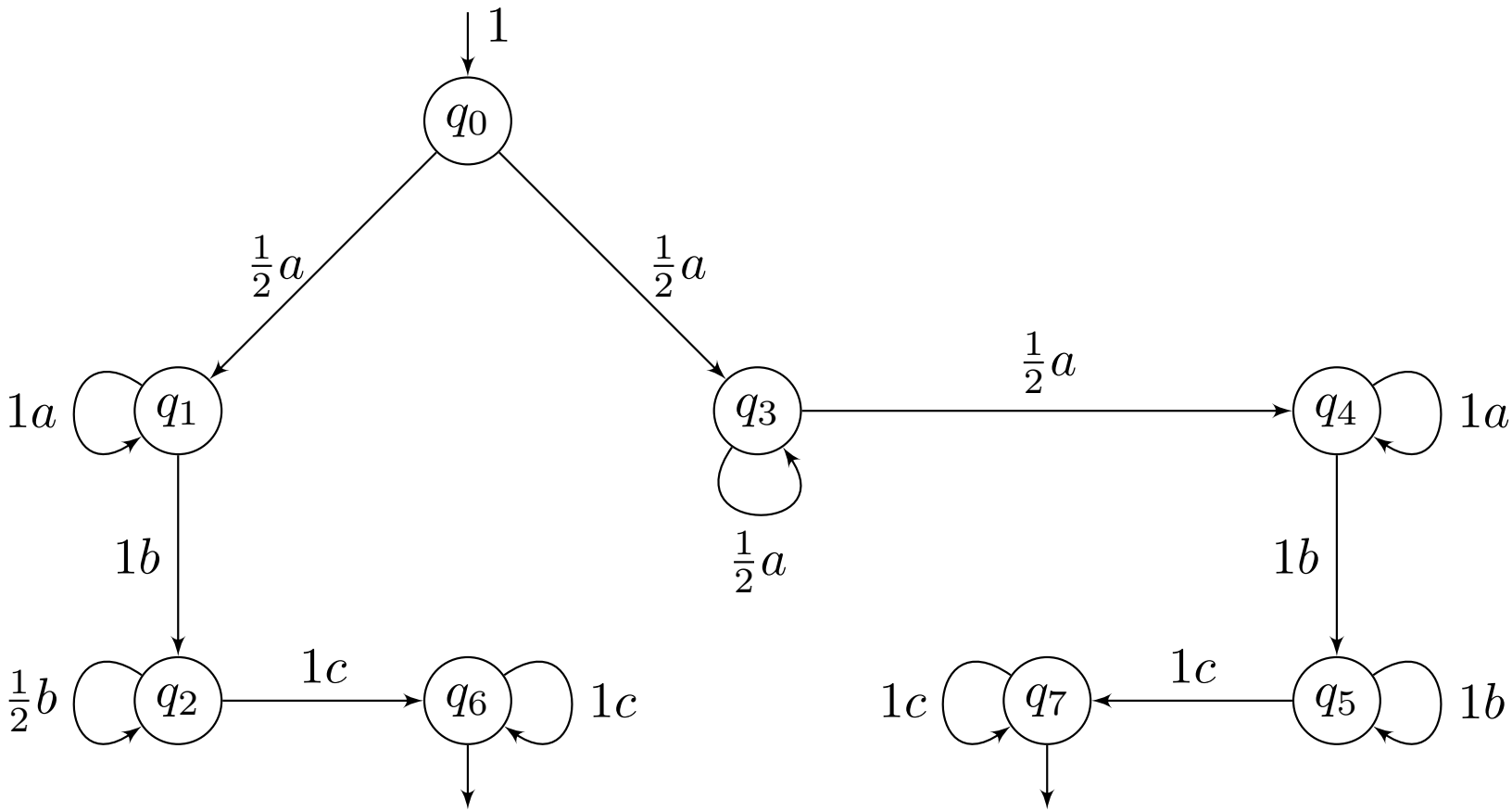
Aussi nous établissons que :

- pour $i \in \{1, 2\}$, $L_i = L_{=\frac{1}{2}}(\mathcal{A}_i)$ pour un certain $\mathcal{A}_i$;
- la famille des langages $\{L_{=\frac{1}{2}}(\mathcal{A})\}_{\mathcal{A}}$ est close par intersection.

Langages stochastiques non algébriques

Preuve (suite).

$$L_{=\frac{1}{2}}(\mathcal{A}) = \{a^n b^n c^+ \mid n > 0\}$$



Langages stochastiques non algébriques

Preuve (fin).

Soient $L_{=\frac{1}{2}}(\mathcal{A}_1)$ et $L_{=\frac{1}{2}}(\mathcal{A}_2)$ deux langages arbitraires.

A l'aide de la construction « produit-complément », soient $\mathcal{A}'_1$ et $\mathcal{A}'_2$ tels que :

- pour tout w, $\mathbf{Pr}_{\mathcal{A}'_i}(w) \leq \frac{1}{4}$;
- $L_{=\frac{1}{2}}(\mathcal{A}_i) = L_{=\frac{1}{4}}(\mathcal{A}'_i)$.

On construit $\mathcal{A}$ ainsi :

- L'ensemble des états est $Q \stackrel{\text{def}}{=} Q'_1 \times Q'_2$;
- $\mathbf{P}_a[(q_1, q_2), (q'_1, q'_2)] \stackrel{\text{def}}{=} (\mathbf{P}'_1)_a[q_1, q'_1](\mathbf{P}'_2)_a[q_2, q'_2]$;
- $\pi'_0[q_1, q_2] \stackrel{\text{def}}{=} \pi'_{1,0}[q_1]\pi'_{2,0}[q_2]$ et $F \stackrel{\text{def}}{=} F'_1 \times F'_2$.

Par construction, $\mathbf{Pr}_{\mathcal{A}}(w) = \mathbf{Pr}_{\mathcal{A}'_1}(w)\mathbf{Pr}_{\mathcal{A}'_2}(w)$.

Pour tout w, $\mathbf{Pr}_{\mathcal{A}}(w) \leq \frac{1}{16}$ et $\mathbf{Pr}_{\mathcal{A}}(w) = \frac{1}{16}$ ssi $\mathbf{Pr}_{\mathcal{A}'_1}(w) = \mathbf{Pr}_{\mathcal{A}'_2}(w) = \frac{1}{4}$.

Ainsi,

$$L_{=\frac{1}{16}}(\mathcal{A}) = L_{=\frac{1}{2}}(\mathcal{A}_1) \cap L_{=\frac{1}{2}}(\mathcal{A}_2)$$

Langages contextuels

La classe des langages stochastiques rationnels est strictement incluse dans la classe des langages contextuels.

Preuve. Une procédure déterministe en espace linéaire (loin d'être optimale).

Pré-calcul en espace constant.

- Calculer le ppcm, b, des dénominateurs de :
  θ, des éléments des matrices $\{\mathbf{P}_a\}_{a\in A}$ et du vecteur π_0.
- Construire les matrices entières $\mathbf{P}'_a \stackrel{\text{def}}{=} b\mathbf{P}_a$ et le vecteur $\pi'_0 \stackrel{\text{def}}{=} b\pi_0$.

Pour $w \stackrel{\text{def}}{=} a_1 \dots a_n$, l'appartenance devient $\pi'_0(\prod_{i=1}^n \mathbf{P}'_{a_i})\mathbf{1}_F^T \bowtie \theta b^{n+1}$?

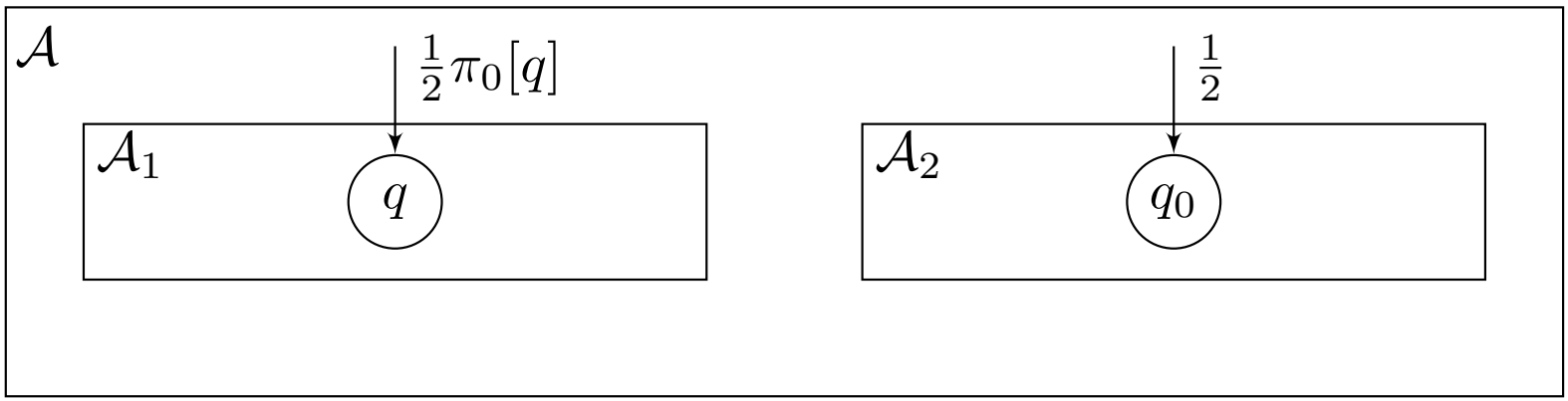
- Calculer θb^{n+1} en espace $O(n)$.
- Calculer $\mathbf{v} \stackrel{\text{def}}{=} \pi'_0(\prod_{i=1}^n \mathbf{P}'_{a_i})$
  en initialisant $\mathbf{v}$ à π'_0 et le multipliant iterativement par $\mathbf{P}'_{a_i}$.
  Les vecteurs sont bornés par b^{n+1}. D'où un calcul en espace $O(n)$.
- La somme et la comparaison se font aussi en espace $O(n)$.

Opérations avec les langages réguliers

La classe des langages stochastiques (rationnels) est close par intersection et union avec les langages reguliers.

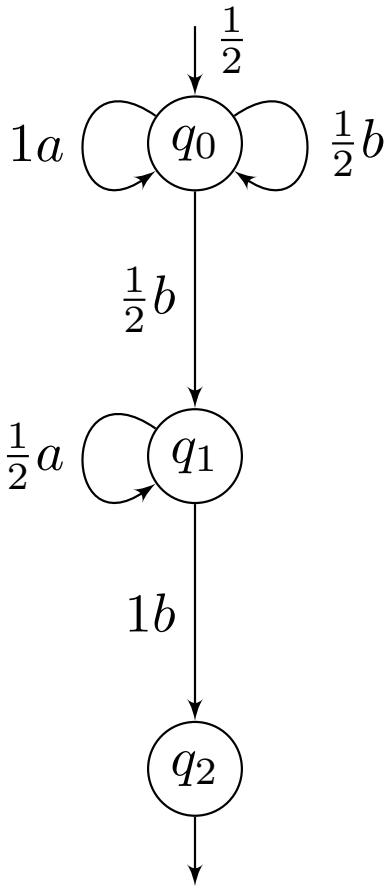
Preuve.

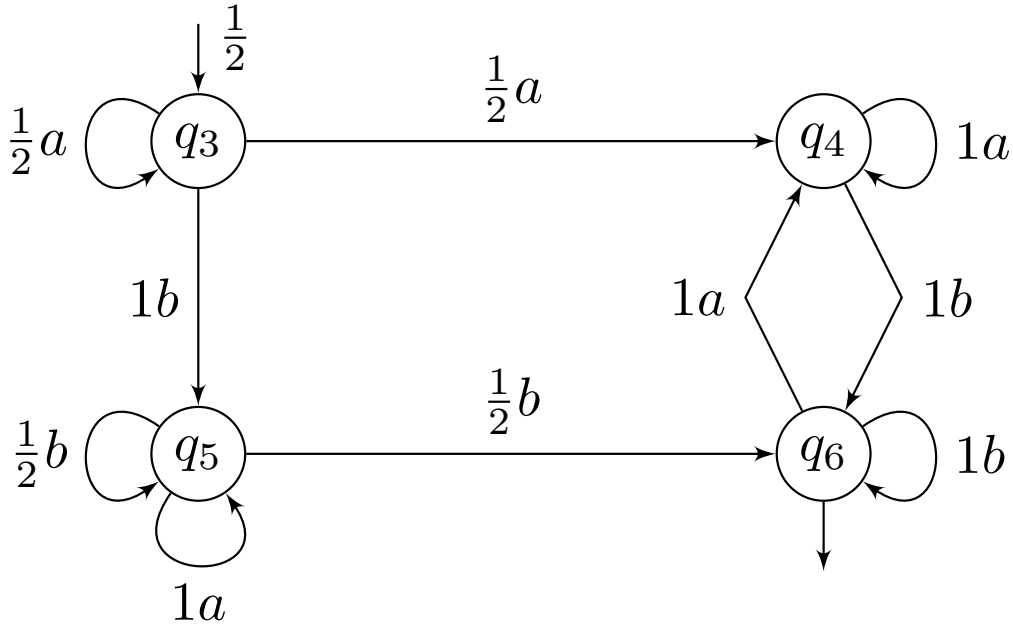
Soit $L_{\bowtie\theta}(\mathcal{A}_1)$ un langage stochastique (rationnel) (with $\bowtie \in \{>, \geq\}$) et $L_{=1}(\mathcal{A}_2)$ langage régulier.



$L_{\bowtie \frac{\theta}{2}}(\mathcal{A}) = L_{\bowtie\theta}(\mathcal{A}_1) \cup L_{=1}(\mathcal{A}_2)$ and $L_{\bowtie \frac{1+\theta}{2}}(\mathcal{A}) = L_{\bowtie\theta}(\mathcal{A}_1) \cap L_{=1}(\mathcal{A}_2)$

Un langage stochastique





$$L_{=\frac{1}{2}}(\mathcal{A}) = \{a^{m_1}b \dots ba^{m_k}b \mid 1 < k \wedge m_1 = m_k\}$$

car $\mathbf{Pr}_{\mathcal{A}}(a^{m_1}b \dots ba^{m_k}b) = \frac{1}{2}\left(\left(\frac{1}{2}\right)^{k+m_k-1} + 1 - \left(\frac{1}{2}\right)^{k+m_1-1}\right)$

Concaténation

La classe des langages stochastiques (rationnels) n'est pas close par concaténation avec un langage régulier.

Preuve.

Soit $L \stackrel{\text{def}}{=} \{a^{m_1} b \dots b a^{m_k} b \mid 1 < k \wedge m_1 = m_k\}$
le language stochastique précédent.

Alors $LA^* = \{a^{m_1} b a^{m_2} b \dots a^{m_k} b a^* \mid \exists i > 1 \; m_i = m_1\}$
n'est pas stochastique.

Itération

La classe des langages stochastiques (rationnels) n'est pas close par itération.

Preuve.

Soit $L \stackrel{\text{def}}{=} \{a^{m_1} b \dots b a^{m_k} b \mid 1 < k \wedge m_1 = m_k\}$ le language stochastique précédent. Supposons que $L^* = L_{\bowtie\theta}(\mathcal{A})$ with $\bowtie \in \{>, \geq\}$.

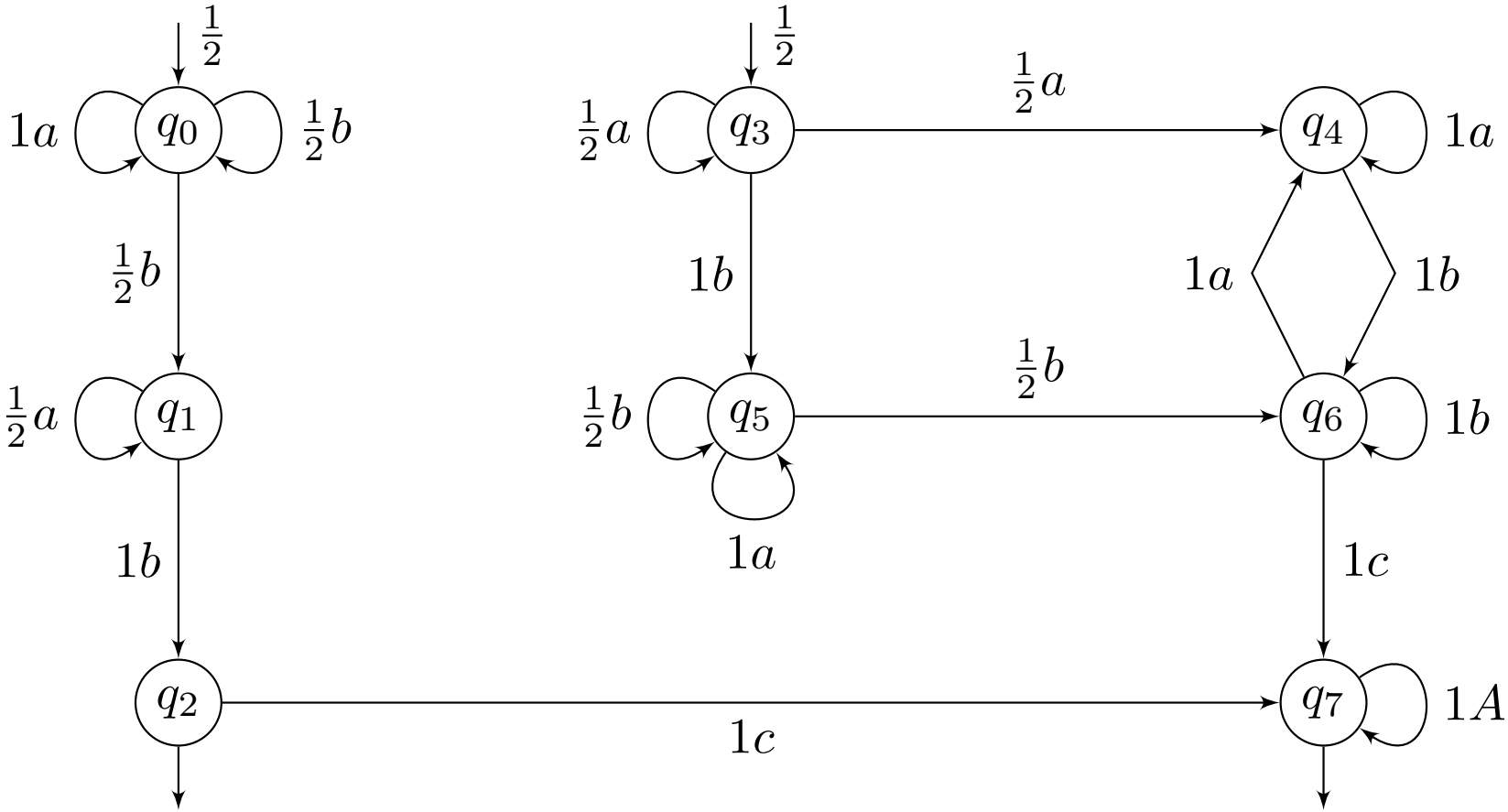
Soit $\sum_{i=0}^{n} c_i x^i$ le polynôme minimal de $\mathbf{P}_a$.
Puisque 1 est valeur propre de $\mathbf{P}_a$, on a $\sum_{i=0}^{n} c_i = 0$
et il y a des c_i négatifs et positifs.

Par définition, $\sum_{i=0}^{n} c_i \mathbf{P}_{a^i} = 0$. Pour tout w, $\sum_{i=0}^{n} c_i \mathbf{P}_{a^i w} = 0$.
Soient $c_{i_1}, \dots, c_{i_k}$ les coefficients positifs du polynôme.
Soit $w \stackrel{\text{def}}{=} b a^{i_1} b (a^{i_2} b)^2 \dots (a^{i_k} b)^2$. $a^i w \in L^*$ ssi $i \in \{i_1, \dots, i_k\}$.

Cas $L^* = L_{>\theta}(\mathcal{A})$**.** Soit $0 \leq i \leq n$, $\pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T > \theta$ ssi $i \in \{i_1, \dots, i_k\}$.
Aussi : $0 = \sum_{i=0}^{n} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T > (\sum_{i=0}^{n} c_i)\theta = 0$
conduisant à une contradiction.

Cas $L^* = L_{\geq\theta}(\mathcal{A})$**.** Soit $0 \leq i \leq n$, $\pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T \geq \theta$ ssi $i \in \{i_1, \dots, i_k\}$.
Aussi : $0 = \sum_{i=0}^{n} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T > (\sum_{i=0}^{n} c_i)\theta = 0$
conduisant à une contradiction.

Un langage stochastique



$$L_{=\frac{1}{2}}(\mathcal{A}) = \{a^{m_1}b \dots ba^{m_k}bcA^* \mid 1 < k \wedge m_1 = m_k\}$$

Homomorphisme

La classe des langages stochastiques (rationnels) n'est pas close par homorphisme.

Preuve.
Soit $L \stackrel{\text{def}}{=} \{a^{m_1}b \dots ba^{m_k}bcA^* \mid 1 < k \wedge m_1 = m_k\}$
le précédent langage stochastique.

Définissons l'homomorphisme h de A vers $A' \stackrel{\text{def}}{=} \{a, b\}$ par :

$$h(a) \stackrel{\text{def}}{=} a \qquad h(b) \stackrel{\text{def}}{=} b \qquad h(c) \stackrel{\text{def}}{=} \varepsilon$$

Alors $h(L) = \{a^{m_1}ba^{m_2}b \dots a^{m_k}ba^* \mid \exists i > 1 \; m_i = m_1\}$
n'est pas un langage stochastique.

Plan

Deux problèmes de décision

Soient $\mathcal{A}$ et $\mathcal{A}'$ des automates probabilistes.

Premier problème

$\mathcal{A}$ et $\mathcal{A}'$ sont-ils équivalents ?

$$\forall w \in A^* \ \mathbf{Pr}_{\mathcal{A}}(w) = \mathbf{Pr}_{\mathcal{A}'}(w)$$

Second problème

$L_{\bowtie\theta}(\mathcal{A})$ est-il égal to $L_{\bowtie'\theta'}(\mathcal{A}')$?

Pour des automates déterministes, il s'agit du même problème.
Il est résolu en temps polynomial par produits d'automates
qui fournissent un témoin de non équivalence de taille inférieure à $|Q||Q'|$.

Rappels d'algèbre linéaire

Soient $\mathbf{v_0} \in \mathbb{R}^n$ et $\mathbf{v_1}, \ldots, \mathbf{v_k}$ des vecteurs linéairement indépendants de $\mathbb{R}^n$.

Comment décider si $\mathbf{v}_0$ est une combinaison linéaire de $\mathbf{v_1}, \ldots, \mathbf{v_k}$?

- Résoudre en $O(k^2 n)$

$$\begin{pmatrix} \mathbf{v}_1[1] & \ldots & \mathbf{v}_k[1] \\ \ldots & \ldots & \ldots \\ \mathbf{v}_1[n] & \ldots & \mathbf{v}_k[n] \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} \mathbf{v}_0[1] \\ \vdots \\ \mathbf{v}_0[n] \end{pmatrix}$$

- Quand $\mathbf{v_1}, \ldots, \mathbf{v_k}$ sont orthogonaux

(i.e. pour tout $a \neq b$, $\mathbf{v}_a \cdot \mathbf{v}_b \stackrel{\text{def}}{=} \sum_{i=1}^{n} \mathbf{v}_a[i]\mathbf{v}_b[i] = 0$)

Calculer en $O(kn)$ la projection orthogonale

$$\mathbf{w}_0 = \sum_{i=1}^{k} \frac{\mathbf{v}_0 \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \mathbf{v}_i$$

Décider en $O(n)$ si $\mathbf{v}_0 = \mathbf{w}_0$.

Principe de l'algorithme d'équivalence

Enumération des mots

Recherche d'un contre-exemple par longueur croissante en démarrant avec ε.

Une pile

Gestion d'une pile de mots w pour chercher des contre-exemples aw avec $a \in A$.

Par souci d'efficacité, la pile contient des paires $\langle(\mathbf{P}_w\mathbf{1}_F, \mathbf{P}'_w\mathbf{1}_{F'}), w\rangle$.

Une famille orthogonale pour restreindre l'énumération

Gen est un ensemble de vecteurs orthogonaux non nuls de $\mathbb{R}^{Q\cup Q'}$.

Si w n'est pas un contre-exemple, vérifier si $\mathbf{v} \stackrel{\text{def}}{=} (\mathbf{P}_w\mathbf{1}_F, \mathbf{P}'_w\mathbf{1}_{F'})$ est combinaison linéaire de vecteurs de Gen.

- Calcul de $\mathbf{v}'$, la projection de $\mathbf{v}$ sur le sous-espace engendré par Gen ;
- Comparaison de $\mathbf{v}'$ avec $\mathbf{v}$.

Si $\mathbf{v}' \neq \mathbf{v}$ alors :

- $\langle(\mathbf{P}_w\mathbf{1}_F, \mathbf{P}'_w\mathbf{1}_{F'}), w\rangle$ est empilé ;
- $\mathbf{v} - \mathbf{v}'$ est ajouté à Gen.

L'algorithme

If $\pi_0 \cdot \mathbf{1}_F \neq \pi'_0 \cdot \mathbf{1}_{F'}$ **then return**$(\mathbf{false}, \varepsilon)$

$Gen \leftarrow \{(\mathbf{1}_F, \mathbf{1}_{F'})\}$; $\mathbf{Push}(Stack, \langle(\mathbf{1}_F, \mathbf{1}_{F'}), \varepsilon\rangle)$

Repeat

  $\langle(\mathbf{v}, \mathbf{v}'), w\rangle \leftarrow \mathbf{Pop}(Stack)$

  For $a \in A$ **do**

    $\mathbf{z} \leftarrow \mathbf{P}_a\mathbf{v}$; $\mathbf{z}' \leftarrow \mathbf{P}'_a\mathbf{v}'$

    If $\pi_0 \cdot \mathbf{z} \neq \pi'_0 \cdot \mathbf{z}'$ **then return**$(\mathbf{false}, aw)$

    $\mathbf{y} \leftarrow \mathbf{0}$; $\mathbf{y}' \leftarrow \mathbf{0}$

    For $(\mathbf{x}, \mathbf{x}') \in Gen$ **do**

      $\mathbf{y} \leftarrow \mathbf{y} + \frac{\mathbf{z}\cdot\mathbf{x}}{\mathbf{x}\cdot\mathbf{x}}\mathbf{x}$

      $\mathbf{y}' \leftarrow \mathbf{y}' + \frac{\mathbf{z}'\cdot\mathbf{x}'}{\mathbf{x}'\cdot\mathbf{x}'}\mathbf{x}'$

    If $(\mathbf{z}, \mathbf{z}') \neq (\mathbf{y}, \mathbf{y}')$ **then**

      $\mathbf{Push}(Stack, \langle(\mathbf{z}, \mathbf{z}'), aw\rangle)$

      $Gen \leftarrow Gen \cup \{(\mathbf{z} - \mathbf{y}, \mathbf{z}' - \mathbf{y}')\}$

Until $\mathbf{IsEmpty}(Stack)$

return$(\mathbf{true})$

Complexité

Complexité temporelle

Un élément est empilé ssi un élément est ajouté à Gen.

Il y a au plus $|Q| + |Q'|$ éléments dans Gen.

Donc il y a au plus $|Q| + |Q'|$ itérations de la boucle externe.

L'indice de la boucle intermédiaire parcourt A
tandis que l'indice de la boucle interne parcourt Gen.

Les opérations au sein de cette boucle se font en $O(|Q| + |Q'|)$.

Ceci conduit à une complexité en $O((|Q| + |Q'|)^3|A|)$.

Longueur des témoins

La longueur du témoin est au plus $|Q| + |Q'|$.
(valide aussi pour les automates déterministes)

Correction

Supposons que les automates ne sont pas équivalents et l'algorithme renvoie **true**.

Soit u un mot non examiné tel que $\mathbf{Pr}_{\mathcal{A}}(u) \neq \mathbf{Pr}_{\mathcal{A}'}(u)$.

Soit $u \stackrel{\text{def}}{=} w'w$ avec $w(\neq u)$ le plus grand suffixe testé par l'algorithme.

Parmi les u possibles, choisissons-en un tel que $|w'|$ est minimal.

Affirmation. Il existe un w'' inséré dans la pile avant w tel que : $\mathbf{Pr}_{\mathcal{A}}(w'w'') \neq \mathbf{Pr}_{\mathcal{A}'}(w'w'')$

Soit $Gen = \{w_1, \ldots, w_k\}$ à l'examen de w. Il existe $\lambda_1, \ldots, \lambda_k$ tels que :

So : $\mathbf{P}_w \mathbf{1}_F = \sum_{i=1}^{k} \lambda_i \mathbf{P}_{w_i} \mathbf{1}_F$ et $\mathbf{P}'_w \mathbf{1}_{F'} = \sum_{i=1}^{k} \lambda_i \mathbf{P}'_{w_i} \mathbf{1}_{F'}$

$$\mathbf{Pr}_{\mathcal{A}}(w'w) \stackrel{\text{def}}{=} \pi_0 \mathbf{P}_{w'} \mathbf{P}_w \mathbf{1}_F = \sum_{i=1}^{k} \lambda_i \pi_0 \mathbf{P}_{w'} \mathbf{P}_{w_i} \mathbf{1}_F = \sum_{i=1}^{k} \lambda_i \mathbf{Pr}_{\mathcal{A}}(w'w_i)$$

De manière similaire : $\mathbf{Pr}_{\mathcal{A}'}(w'w) = \sum_{i=1}^{k} \lambda_i \mathbf{Pr}_{\mathcal{A}'}(w'w_i)$

Il existe donc i, avec $\mathbf{Pr}_{\mathcal{A}}(w'w_i) \neq \mathbf{Pr}_{\mathcal{A}'}(w'w_i)$.

Soit $w' \stackrel{\text{def}}{=} w'''a$. aw_i est examiné par l'algorithme.

Le mot $u' \stackrel{\text{def}}{=} w'w_i$ a une décomposition $u' \stackrel{\text{def}}{=} z'z$ où z le plus grand suffixe examiné par l'algorithme a pour suffixe aw_i. Aussi $|z'| < |w'|$: une contradiction.

Indecidabilité du problème d'égalité

Soit $\mathcal{A}$ un automate stochastique rationnel,
la question $L_{=\frac{1}{2}}(\mathcal{A}) = \{\varepsilon\}$? est indécidable.

Preuve.

Par réduction du problème (indécidable) de correspondance de Post (PCP) :

Soit un alphabet A et deux morphismes φ_1, φ_2 de A vers $(0+1)^+$,
existe-t-il un mot $w \in A^+$ tel que $\varphi_1(w) = \varphi_2(w)$?

Déjà indécidable si les images des lettres appartiennent à $(10+11)^+$.

Insérer un 1 avant chaque lettre d'une image
réduit le problème original à cette variante.

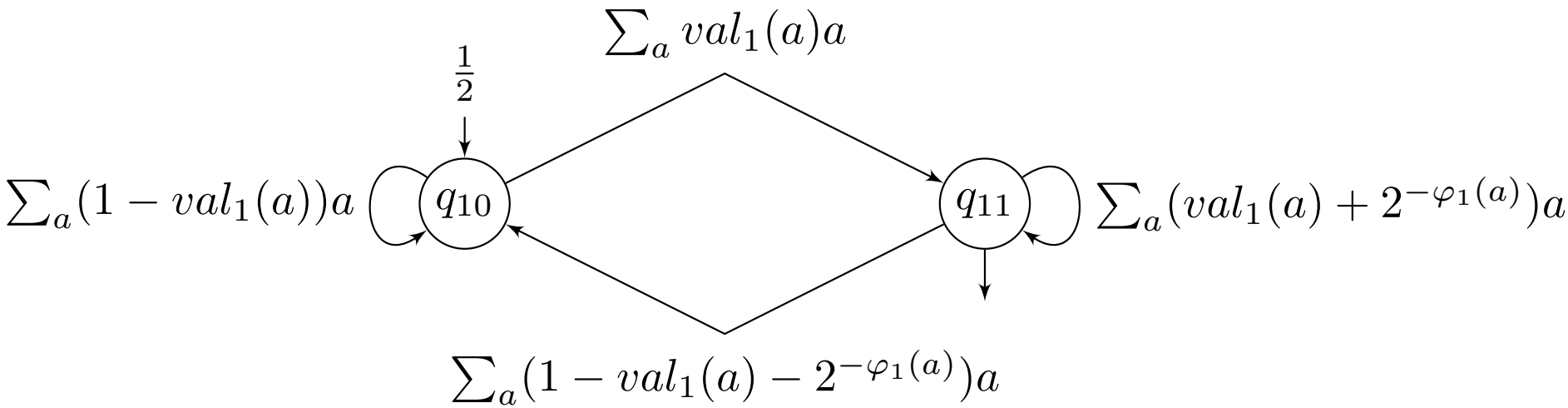
Un mot $w \stackrel{\text{def}}{=} a_1 \dots a_n \in (10+11)^+$ définit une valeur $val(w) \in [0,1]$ par :

$$val(w) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \frac{a_i}{2^{n+1-i}}$$

Puisque tout mot débute par un 1, $val(w) = val(w')$ implique $w = w'$.

Réduction de PCP

Pour $w \in A^+$ et $i \in \{1, 2\}$, soit $val_i(w) \stackrel{\text{def}}{=} val(\varphi_i(w))$.



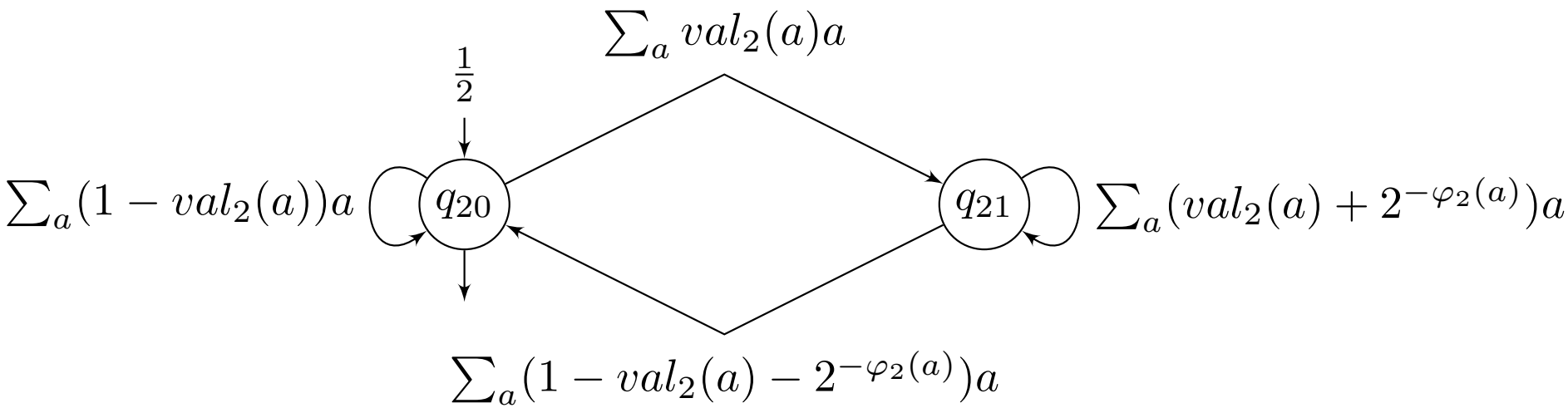
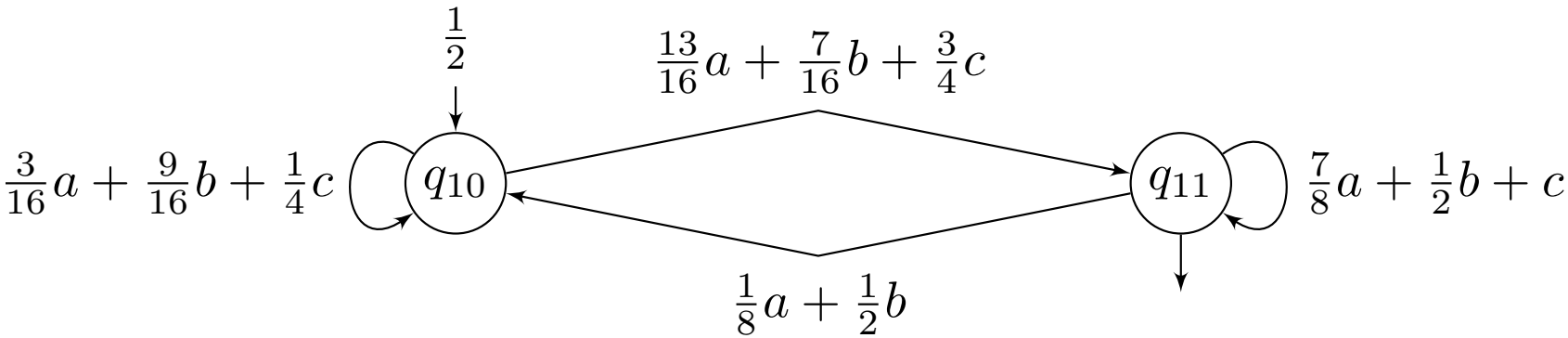
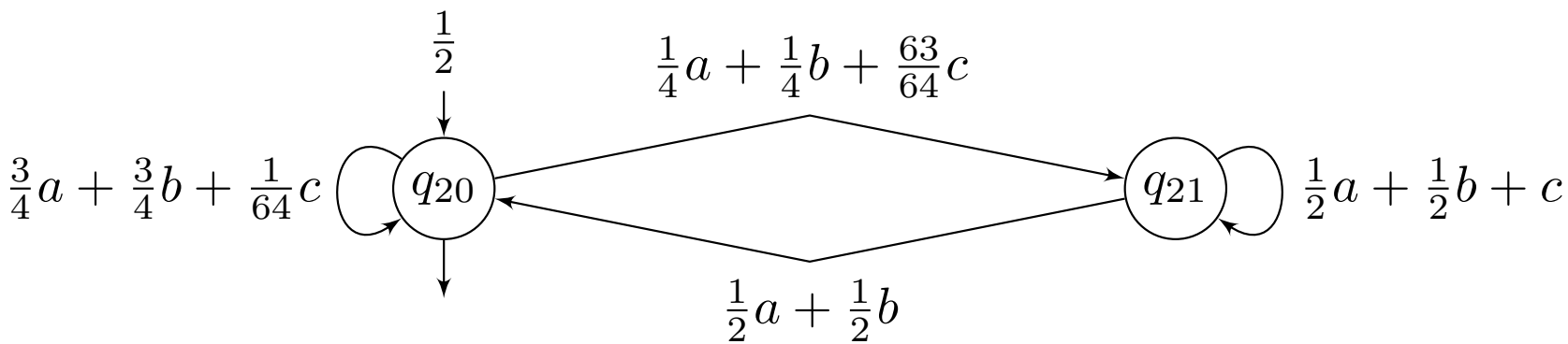


Illustration de la réduction

A	a	b	c
φ_1	$(1)0(1)1$	$(1)0(1)0$	$(1)1$
φ_2	$(1)0$	$(1)0$	$(1)1(1)1(1)1$

A	a	b	c
val_1	$\frac{13}{16}$	$\frac{7}{16}$	$\frac{3}{4}$
val_2	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{63}{64}$





Correction de la réduction

L'équation de récurrence :

$$\mathbf{1}_{q_{i0}}\mathbf{P}_{wa}\mathbf{1}_{q_{i1}}^T = \mathbf{1}_{q_{i0}}\mathbf{P}_w\mathbf{1}_{q_{i1}}^T(val_i(a) + 2^{-|\varphi_i(a)|}) + (1 - \mathbf{1}_{q_{i0}}\mathbf{P}_w\mathbf{1}_{q_{i1}}^T)val_i(a)$$

$$= val_i(a) + 2^{-|\varphi_i(a)|}\mathbf{1}_{q_{i0}}\mathbf{P}_w\mathbf{1}_{q_{i1}}^T$$

Par induction nous obtenons que pour tout $w \stackrel{\text{def}}{=} a_1 \dots a_n$:

$$\mathbf{1}_{q_{i0}}\mathbf{P}_w\mathbf{1}_{q_{i1}}^T = \sum_{j=1}^{n} val_i(a_j)2^{-\sum_{j<k\leq n}|\varphi_i(a_k)|} = val_i(w)$$

Ainsi pour $w \in A^+$: $\mathbf{Pr}_{\mathcal{A}}(w) = \frac{1}{2}(val_1(w) + 1 - val_2(w))$.

Donc $w \in L_{=\frac{1}{2}}(\mathcal{A})$ ssi $val(\varphi_1(w)) = val(\varphi_2(w))$ impliquant $\varphi_1(w) = \varphi_2(w)$.

Problèmes qualitatifs

Un problème de PA est dit *qualitatif*
si les probabilités qui apparaissent dans l'enoncé appartiennent à $\{0, 1\}$.

Le problème indécidable précédent n'est pas un problème qualitatif
car le seuil peut être différent de 0 et 1.

Observation. Très souvent, les procédures de décision des problèmes qualitatifs ne comparent les probabilités qu'à 0 ou 1.

Exemple. Pour tester si $L_{>0}(\mathcal{A}) \neq \emptyset$, on cherche un chemin d'un état initial à un état final dans le graphe sous-jacent de l'automate.

Deux problèmes qualitatifs

Recherche d'une stratégie sûre

Etant donné un automate probabiliste $\mathcal{A}$,
existe-t-il un mot w tel que $\mathbf{Pr}_{\mathcal{A}}(w) = 1$?

- On complémente les états finals.
- On considère $\mathcal{A}'$ l'automate non déterministe obtenu en oubliant les probabilités.
- On décide si $L(\mathcal{A}') \neq \Sigma^*$.

Ce problème est PSPACE-complet.

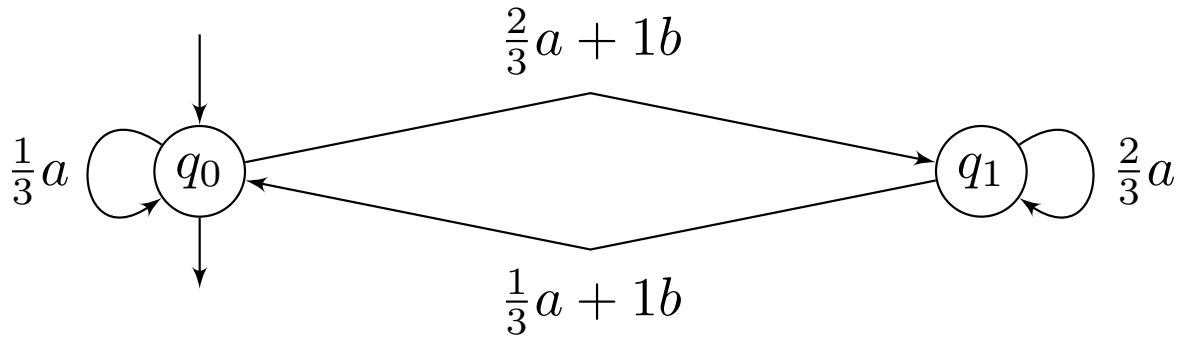
Recherche de stratégies presque sûres

Etant donné un automate probabiliste $\mathcal{A}$,
existe-t-il une suite de mots $\{w_n\}_{n\in\mathbb{N}}$ telle que $\lim_{n\to\infty} \mathbf{Pr}_{\mathcal{A}}(w_n) = 1$?

Ce problème appelé *problème de la valeur 1* est indécidable.
(H. Gimbert et Y. Oualhadj, ICALP 2010)

Plan

Probabilités des mots



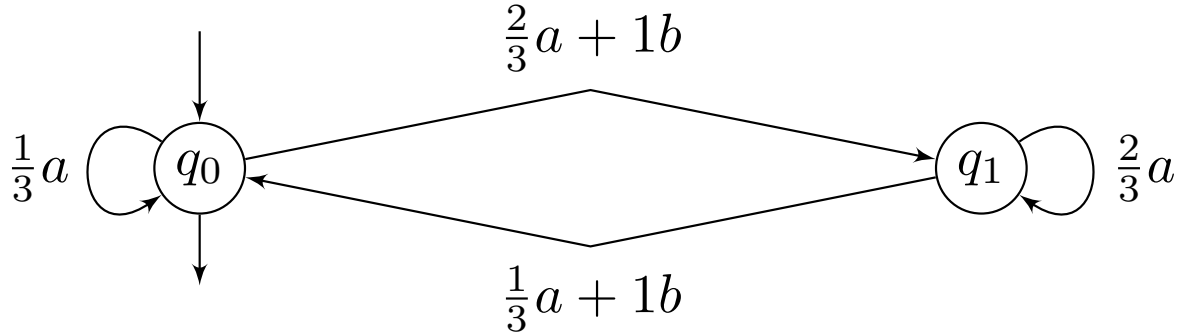
Calculer $\mathbf{P}_{a^2}$ puis $\mathbf{P}_{a^n}$ pour $n \in \mathbb{N}$.

Calculer $\mathbf{P}_{b^2}$ puis $\mathbf{P}_{b^n}$ pour $n \in \mathbb{N}$.

Calculer $\mathbf{P}_{ab}$ et $\mathbf{P}_{ba}$.

Calculer $\mathbf{P}_w$ pour w quelconque.

Caractérisation du langage



Décrire $L_{\leq\theta}(\mathcal{A})$ en fonction de θ.

Stochastique ?

Soit θ fixé. Le langage $\{w \mid 0.\overline{w} \geq \theta\}$
où $\Sigma = \{z(ero), u(n)\}$ et $\overline{w}$ est le nombre binaire représenté par w, est-il stochastique ?

Le langage $\{a^m b a^{m_1} b a^{m_2} b \cdots a^{m_k} b a^* \mid \exists j \leq k \ \sum_{i \leq j} m_i = m\}$
est-il stochastique ?

Corrigés (1)

$$\mathbf{P}_{a^2} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \mathbf{P}_a$$

D'où $\mathbf{P}_{a^n} = \mathbf{P}_a$.

$$\mathbf{P}_{b^2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \mathbf{Id}$$

D'où si b est impair, $\mathbf{P}_{b^n} = \mathbf{P}_b$ sinon $\mathbf{P}_{b^n} = \mathbf{Id}$.

$$\mathbf{P}_{ba} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \mathbf{P}_a$$

Si $w = b^{n_1} a \dots a b^{n_k}$ alors $\mathbf{P}_w = \mathbf{P}_{ab^{n_k}}$.

D'où si n_k est pair alors $\mathbf{P}_w = \mathbf{P}_a$ sinon

$$\mathbf{P}_w = \mathbf{P}_{ab} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Corrigés (2)

$\pi_0 \cdot \mathbf{P}_b \cdot \mathbf{1}_F^T = 0$

$\pi_0 \cdot \mathbf{P}_a \cdot \mathbf{1}_F^T = \frac{1}{3}$

$\pi_0 \cdot \mathbf{P}_{ab} \cdot \mathbf{1}_F^T = \frac{2}{3}$

$\pi_0 \cdot \mathbf{Id} \cdot \mathbf{1}_F^T = 1$

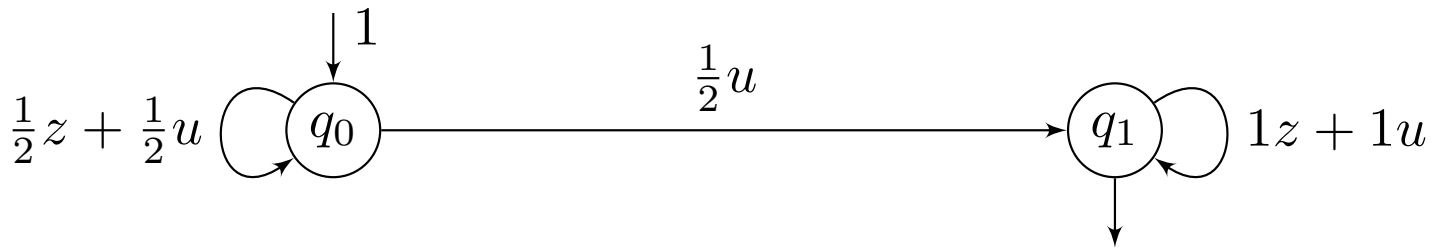
D'où :

$L_{\leq\theta}(\mathcal{A}) = \{b^n \mid n \text{ impair}\}$ si $\theta < \frac{1}{3}$

$L_{\leq\theta}(\mathcal{A}) = L_{\leq 0}(\mathcal{A}) \cup \{b^{n_1} a \dots a b^{n_k} \mid n_k \text{ pair}\}$ si $\theta < \frac{2}{3}$

$L_{\leq\theta}(\mathcal{A}) = L_{\leq \frac{1}{3}}(\mathcal{A}) \cup \{b^{n_1} a \dots a b^{n_k} \mid n_k \text{ impair}\}$ si $\theta < 1$

Corrigés (3)



$$\Pr(zw) = \frac{1}{2}\Pr(w) \qquad \Pr(uw) = \frac{1}{2}(1 + \Pr(w))$$

Corrigés (4)

Soit $\sum_{i=0}^{n} c_i x^i$ le polynôme minimal de $\mathbf{P}_a$.

Soit $Pos = \{i \mid 0 \leq i \leq n \wedge c_i > 0\}$ et $NonPos = \{i \mid 0 \leq i \leq n \wedge c_i \leq 0\}$.

Notons Pos par $\{i_1 < \ldots < i_k\}$.

Fixons $w \stackrel{\text{def}}{=} ba^{i_1} ba^{i_2 - i_1} \ldots ba^{i_k - i_{k-1}} b$.

Cas $L = L_{>\theta}(\mathcal{A})$**.** Soit $0 \leq i \leq n$, par définition de L,

$$\pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T > \theta \text{ ssi } i \in \{i_1, \ldots, i_k\}$$

Aussi :

$0 = \sum_{i=0}^{n} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T = \sum_{i \in Pos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T + \sum_{i \in NonPos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T$
$> (\sum_{i \in Pos} c_i)\theta + (\sum_{i \in NonPos} c_i)\theta = (\sum_{i=0}^{n} c_i)\theta = 0$

conduisant à une contradiction.

Cas $L = L_{\geq\theta}(\mathcal{A})$**.** Soit $0 \leq i \leq n$, par définition de L,

$$\pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T \geq \theta \text{ ssi } i \in \{i_1, \ldots, i_k\}$$

Aussi : $0 = \sum_{i=0}^{n} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T = \sum_{i \in Pos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T + \sum_{i \in NonPos} c_i \pi_0 \mathbf{P}_{a^i w} \mathbf{1}_F^T$
$>(\sum_{i \in Pos} c_i)\theta + (\sum_{i \in NonPos} c_i)\theta = (\sum_{i=0}^{n} c_i)\theta = 0$

conduisant à une contradiction.