Le monoïde Markovien

Hugo Gimbert
CNRS, LaBRI, Bordeaux.

Réunion FREC, Mai 2011

Joint work with
Nathanaël Fijalkow (ENS Cachan)
and Youssouf Oualhadj (LaBRI)

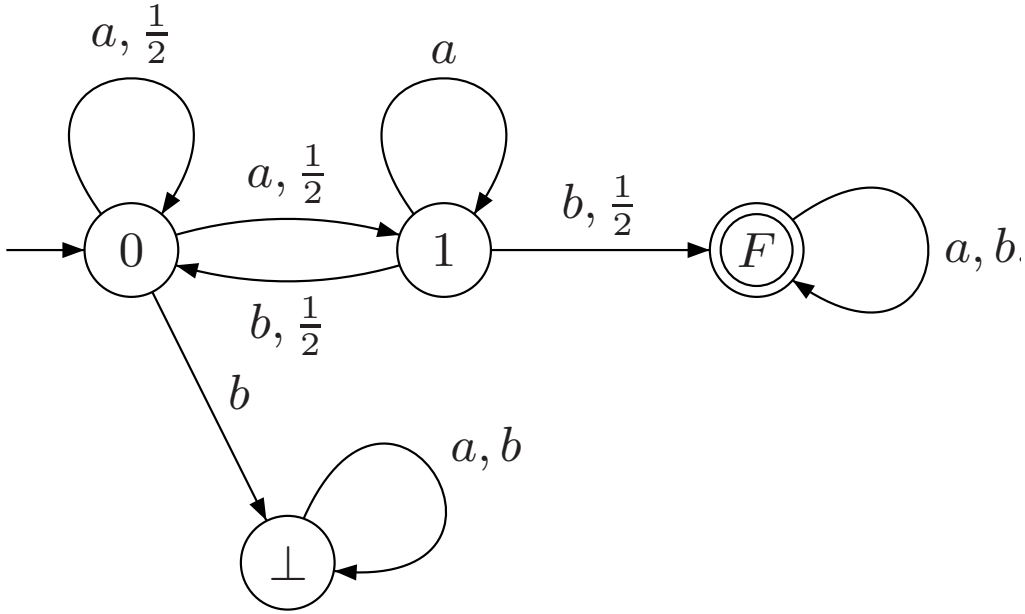
Automates probabilistes

Monoïde Markovien

Problème de la valeur 1

Leung - Simon, algebraic techniques for distance automata.

Automates probabilistes sur les mots finis (Rabin 63)



Le mot aab est accepté avec probabilité $\frac{3}{8}$.

$$= \delta_i \cdot M_a \cdot M_a \cdot M_b \cdot \delta_F$$

$$= \begin{vmatrix} 1 & 0 & 0 & 0 \end{vmatrix} \cdot \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} 0 \\ 0 \\ 1 \\ 0 \end{vmatrix}$$

Automates probabilistes sur les mots finis

Rabin (63): langage d'un automate probabiliste.

Pour $0 \leq \lambda \leq 1$, le langage $\mathcal{L}_\lambda(\mathcal{A})$ d'un automate probabiliste $\mathcal{A}$ est l'ensemble de mots acceptés avec probabilité supérieure à λ.

Automates probabilistes sur les mots finis

Rabin (63): langage d'un automate probabiliste.

Pour $0 \leq \lambda \leq 1$, le langage $\mathcal{L}_\lambda(\mathcal{A})$ d'un automate probabiliste $\mathcal{A}$ est l'ensemble de mots acceptés avec probabilité supérieure à λ.

Paz (71): pour tout $0 < \lambda < 1$, le problème du vide est indécidable.

Lien avec les automates non-déterministes?

Reformulation du résultat de Paz

Paz (71): le problème du vide est indécidable pour $\lambda = \frac{1}{2}$.

Corollaire: étant donné un automate non-déterministe, on ne peut pas décider si il existe un mot tel que au moins la moitié des calculs sur ce mot sont acceptants.

Problème du vide

Pour $\lambda = 0$ le problème (strict) du vide est décidable.

Problème du vide

Pour $\lambda = 0$ le problème (strict) du vide est décidable.

Pour $\lambda = 1$ le problème du vide est décidable.

Points de coupure isolés

Rabin (63): si le point de coupure λ est isolé:

$$\exists \epsilon > 0, \forall u \in A^*, |\mathbb{P}(u) - \lambda| \geq \epsilon \ ,$$

alors le langage $\mathcal{L}_\lambda(\mathcal{A})$ est rationnel.

Points de coupure isolés

Rabin (63): si le point de coupure λ est isolé:

$$\exists \epsilon > 0, \forall u \in A^*, |\mathbb{P}(u) - \lambda| \geq \epsilon \ ,$$

alors le langage $\mathcal{L}_\lambda(\mathcal{A})$ est rationnel.

Bertoni (75): pour $0 < \lambda < 1$, le problème de l'isolation est indécidable.

Points de coupure isolés

Rabin (63): si le point de coupure λ est isolé:

$$\exists \epsilon > 0, \forall u \in A^*, |\mathbb{P}(u) - \lambda| \geq \epsilon \ ,$$

alors le langage $\mathcal{L}_\lambda(\mathcal{A})$ est rationnel.

Bertoni (75): pour $0 < \lambda < 1$, le problème de l'isolation est indécidable.

Cas ouvert: $\lambda = 1$.

1 n'est *pas* isolé
$\iff$ l'automate accepte des mots avec probabilité arbitrairement proche de 1.
$\iff$ l'automate a valeur 1.

Le problème de la valeur 1

Problème de la valeur 1. Etant donné un automate $\mathcal{A}$ est-ce que:

$$\sup_{u \in A^*} \mathbb{P}(u) = 1 \ ?$$

Procédure de décision (en temps infini):

1. Calculer l'ensemble des produits de matrices de transition

$$\mathcal{M} = \{M_a \mid a \in A\}^* \subseteq [0,1]^{Q \times Q} \ .$$

2. Calculer $\bar{\mathcal{M}}$ la fermeture topologique de $\mathcal{M}$.
3. L'automate a valeur 1 ssi $\bar{\mathcal{M}}$ contient une matrice M telle que:

$$\{q \in Q \mid M(i,q) > 0\} \subseteq F \ ,$$

où i est l'état initial et F l'ensemble des états finaux.

Un algorithme pour décider le problème de la valeur 1

Nécessité: abstraire le calcul de $\bar{M}$.

Abstraction "binaire" des probas: $(\{0,1\}, \vee, \wedge)$,

Abstraction des limites de produits de matrices de transition.
Graphes dirigés $G \in \{0,1\}^{Q \times Q}$. muni:

1. du produit de composition $G \cdot H$,
2. d'une opération d'itération $G^\sharp$ pour les idempotents.

Un algorithme pour décider le problème de la valeur 1

On calcule le monoide Markovien $\mathcal{G}$ associé à un automate $\mathcal{A}$.
Monoïde de stabilisation $(\mathcal{G}, \cdot, \sharp)$ contenant des graphes dirigés $G \in \{0,1\}^{Q \times Q}$.

Deux propriétés.

- Consistence. $\forall G \in \mathcal{G}, \exists (u_n)_{n \in \mathbb{N}}$:

$$(s,t) \in G \iff \liminf_n u_n(s,t) > 0 \ ,$$

- Complétude. $\forall (u_n)_{n \in \mathbb{N}}, \exists G \in \mathcal{G}$:

$$(s,t) \in G \iff \liminf_n u_n(s,t) = 0 \implies (s,t) \in G \ .$$

Témoin de valeur 1: graphe G tel que:

$$\{q \in Q \mid (i,q) \in G\} \subseteq F,$$

Lemme: Supposons le monoïde Markovien consistent et complet.
Alors $\mathcal{A}$ a valeur 1 si et seulement si son monoïde Markovien contient un témoin de valeur 1.

Le monoïde Markovien

Abstraction des matrices de transition: pour tout $a \in A$ on note $G_a \in \{0,1\}^{Q \times Q}$ défini par:

$$(s,t) \in G_a \iff M_a(s,t) > 0 \ .$$

Monoïde Markovien: plus petit monoïde contenant $\{G_a \mid a \in A\}$ et stable par produit $\cdot$ et itération $\sharp$.

Le monoïde Markovien

Abstraction des matrices de transition: pour tout $a \in A$ on note $G_a \in \{0,1\}^{Q \times Q}$ défini par:

$$(s,t) \in G_a \iff M_a(s,t) > 0 \ .$$

Monoïde Markovien: plus petit monoïde contenant $\{G_a \mid a \in A\}$ et stable par produit $\cdot$ et itération $\sharp$.

Produit de matrices à coefficients dans $(\{0,1\}, \vee, \wedge)$.

$$(s,t) \in G \cdot H \iff \exists q \in Q, (s,q) \in G \wedge (q,t) \in H \ .$$

Itération?

L'opération d'itération $\sharp$

Etat récurrent: Soit $G \in \{0,1\}^{Q \times Q}$. Un état q est G-récurrent si
$\forall r \in Q$,
r est G-accessible à partir de q
$\Longrightarrow$ q est G-accessible à partir de r.

Remarque: Si G est idempotent alors q est G-récurrent ssi:

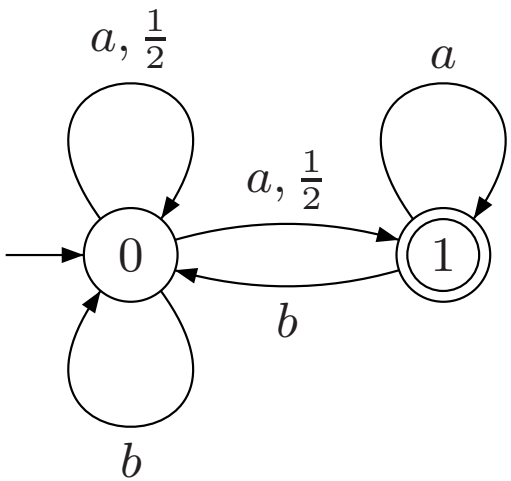
$$\forall r \in Q,\ (q,r) \in G \Longrightarrow (r,q) \in G \ .$$

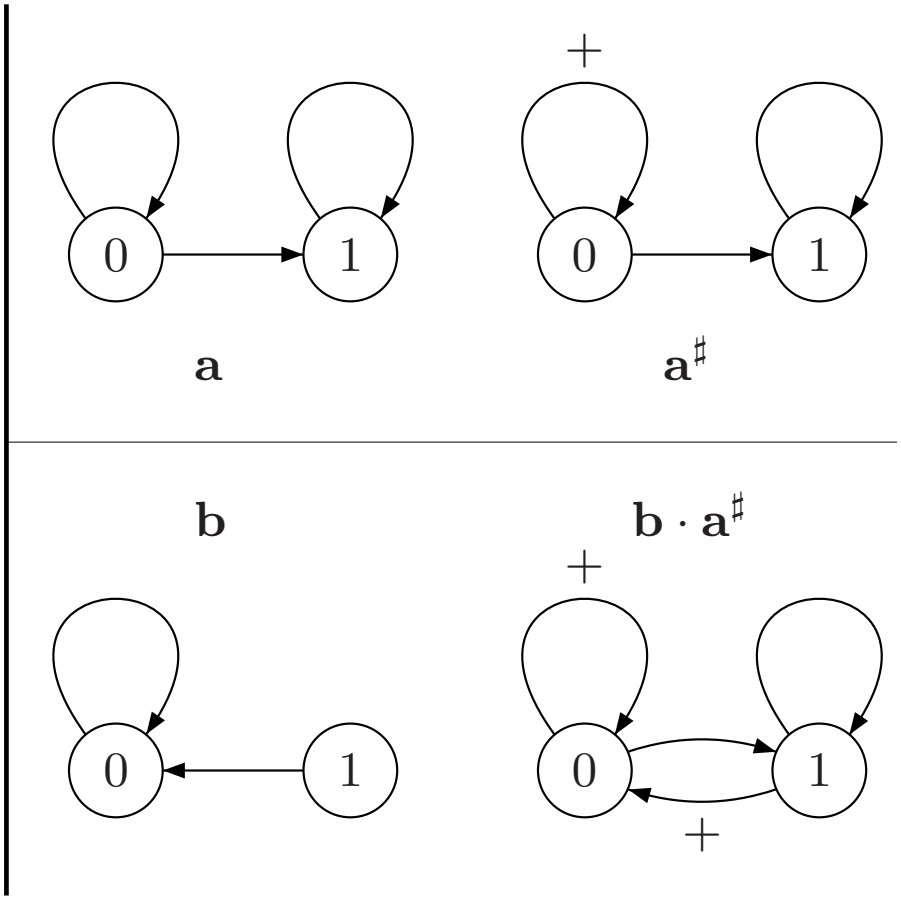
Opération d'itération: soit G idempotent.

$$(q,r) \in G^{\sharp} \Longleftrightarrow (q,r) \in G \wedge r \text{ est } G\text{-récurrent}.$$

Un exemple

En oubliant les $+$:





The limitedness problem for distance automata

Hashigushi 82, 90. Limitedness theorems.

Distance automata: meilleur calcul $(\mathbb{N}, \min, +)$.
Probabilistic automata: calculs parallèles $(\mathbb{R}, +, *)$.

Approche algébrique.
Simon et Leung. *On semigroups of matrices over the tropical semiring*.

Automate à distance: le minimum de deux petites distances est petit.
Automate probabiliste: la somme de deux petites probabilités n'est pas forcément petite.

Indécidabilité

Théorème [G., Oualhadj, 09]: le problème de la valeur 1 est indécidable.

Preuve: construction ad-hoc. Technique de [Baier, Bertrand, Groesser, 08]. Réduction à partir du problème du vide.

Indécidabilité

Théorème [G., Oualhadj, 09]: le problème de la valeur 1 est indécidable.

Preuve: construction ad-hoc. Technique de [Baier, Bertrand, Groesser, 08]. Réduction à partir du problème du vide.

Corollaire: on ne peut pas décider si, étant donné un automate non-déterministe, il existe des mots dont la proportion des calculs acceptant est arbitrairement proche de 1.

Décidabilité pour les automates étanches

Lemme: Si le monoïde Markovien d'un automate $\mathcal{A}$ contient un témoin de valeur 1 alors $\mathcal{A}$ a valeur 1.

Théorème: [Fijalkow, G., Oualhadj 11] si l'automate $\mathcal{A}$ est étanche alors la réciproque du lemme précédent est vraie.

Corollaire: le problème de la valeur 1 est décidable pour les automates probabilistes étanches.

Décidabilité pour les automates étanches

Une fuite d'un état $r \in Q$ à un état $q \in Q$ est une suite $(u_n)_{n\in\mathbb{N}}$ de mots idempotents tels que:

1. pour tout état $s, t \in Q$, la suite $(u_n(s,t))_{n\in\mathbb{N}}$ converge vers une valeur $u(s,t)$.On dénote $\mathcal{M}_u$ la chaîne de Markov avec états Q et probabilités de transition $(u(s,t))_{s,t\in Q}$,
2. l'état r est récurrent dans $\mathcal{M}_u$,
3. $\forall n \in \mathbb{N}, u_n(r,q) > 0$,
4. et r n'est pas accessible depuis q dans $\mathcal{M}_u$.

Définition: un automate est étanche si il est sans fuite.

Remarque: on peut décider si un automate est étanche ou non.

Décidabilité pour les automates étanches

Proposition: [Fijalkow, G., Oualhadj 11] si le monoïde Markovien d'un automate $\mathcal{A}$ ne contient pas de témoin de valeur 1 et si $\mathcal{A}$ est étanche alors $\mathcal{A}$ a valeur inférieure à $1 - p^{3 \cdot 2^{4|Q|^2}}$, où p est la probabilité de transition minimale non-nulle.

Preuve:
adaptation de la preuve algébrique de Imre Simon pour les automates à distance.

Forêt de factorisation.

Un idempotent instable $G \neq G^\sharp$ correspond à une décroissance stricte dans l'ordre $\leq_{\mathcal{J}}$.

Conclusion

Le monoïde Markovien permet d'abstraire l'ensemble des calculs d'un automate probabiliste.

Les automates étanches: une classe d'automates probabilistes dont la valeur est calculable.

Automates probabilistes = jeux à un joueur et demi sans observation.
Extension aux jeux avec observation partielle.