

Algorithmes pour les jeux stochastiques

Hugo Gimbert

27 novembre 2009

Table des matières

1	Jeux à 0 joueur	5
1.1	Chaînes de Markov	5
1.1.1	Distributions de probabilités et leurs limites	5
1.1.2	Chaînes de Markov	5
1.2	Calcul des probabilités d'accessibilité	7
1.3	Conditions de Büchi	9
1.4	Vérification des systèmes probabilistes	11
1.4.1	Approche naïve par automates	12
1.4.2	Approche optimale	16
2	Jeux stochastiques à 2 joueurs à information parfaite	19
2.1	Jeux d'accessibilité stochastiques à information parfaite	19
2.1.1	Jeux et stratégies	19
2.1.2	Valeurs et stratégies optimales	21
2.2	Calcul des valeurs	26
2.2.1	Polynomial dans le cas à un joueur	26
2.2.2	Un problème dans $NP \cap co - NP$	27
2.2.3	Algorithme d'amélioration des stratégies	27
2.2.4	Algorithme des permutations	28

Chapitre 1

Jeux à 0 joueur

1.1 Chaînes de Markov

1.1.1 Distributions de probabilités et leurs limites

Pour mesurer les probabilités d'un nombre fini ou dénombrable d'événements, on utilise des distributions de probabilités.

Définition 1 (Distributions de probabilités). *Etant donné un ensemble \mathbf{S} fini ou dénombrable, on note $\mathcal{D}(\mathbf{S})$ l'ensemble des distributions de probabilités sur \mathbf{S} défini par $\mathcal{D}(\mathbf{S}) = \{\delta : \mathbf{S} \rightarrow [0, 1] \mid \sum_{s \in \mathbf{S}} \delta(s) = 1\}$.*

Par exemple, \mathbf{S} peut être l'ensemble $\{\text{pile}, \text{face}\}$ et $\delta(\text{pile}) = \delta(\text{face}) = \frac{1}{2}$.

Exercice 1. *Supposons qu'on lance une pièce jusqu'à obtenir face. On peut représenter cette expérience en utilisant l'ensemble d'états $\text{pile}^* \text{face} \cup \{\text{pile}^\omega\}$. Montrer que δ défini par $\delta(\text{pile}^n \text{face}) = \frac{1}{2^{n+1}}$ et $\delta(\text{pile}^\omega) = 0$ est une distribution de probabilités sur $\text{pile}^* \text{face} \cup \{\text{pile}^\omega\}$.*

Exercice 2. *Est-il possible de choisir un entier $n \in \mathbb{N}$ au hasard ?*

Exercice 3. *Supposons maintenant que $\mathbf{S} = \{\text{pile}, \text{face}\}^\omega$. Quelle particularité ont les fonctions $\delta : \mathbf{S} \rightarrow [0, 1]$ telles que $\sum_{s \in \mathbf{S}} \delta(s) = 1$?*

1.1.2 Chaînes de Markov

Pour pouvoir définir des probabilités sur des ensembles d'événements non-dénombrables comme toutes les suites possibles de pile ou face, l'exercice 3 montre que les distributions de probabilités ne sont pas un outil mathématique adapté. Il faut utiliser la notion de *mesure* de probabilité. Dans ce cours, on n'introduira pas la notion générale de mesure de probabilité, on se bornera à mesurer la probabilité d'événements dans des chaînes de Markov.

Définition 2 (Chaînes de Markov). *Soit \mathbf{S} un ensemble fini ou dénombrable. Une chaîne de Markov est un triplet $\mathcal{M} = (\mathbf{S}, s_0 \in \mathbf{S}, p : \mathbf{S} \rightarrow \mathcal{D}(\mathbf{S}))$.*

Intuitivement, une chaîne de Markov change d'état à chaque instant $n \in \mathbb{N}$. Initialement, la chaîne est dans l'état $s_0 \in \mathbf{S}$. Sachant qu'à l'instant n la chaîne est dans l'état $s \in \mathbf{S}$ alors à l'instant $n + 1$ elle sera dans l'état $t \in \mathbf{S}$ avec probabilité $p(s, t)$.

Pour représenter une suite infinie de pile ou face, on peut définir la chaîne de Markov suivante. L'ensemble d'états est $\mathbf{S} = \{\text{pret, pile, face}\}$, l'état initial est $s_0 = \text{pret}$ et les probabilités de transition sont $p(\text{face, pret}) = p(\text{pile, pret}) = \frac{1}{2}$ et $p(\text{pile, pile}) = p(\text{face, pile}) = p(\text{face, face}) = p(\text{pile, face}) = 1$. Graphiquement, on représente cette chaîne comme suit :

Il existe des notions plus générales de chaînes de Markov, par exemple les chaînes non-homogènes dont les probabilités de transition changent au cours du temps, ou les chaînes de Markov en temps continu dont les dates de changements d'états suivent une loi de Poisson.

Définition 3 (Événements). *On appelle cylindre un sous-ensemble de S^ω du type $s_0 \cdots s_n S^\omega$ avec $s_0 \cdots s_n \in S^*$. L'ensemble des événements est le plus petit sous-ensemble de S^ω qui contient les cylindres et qui est stable par complémentation (notée \neg) et union dénombrable.*

Exercice 4. *Soit s un état d'une chaîne de Markov \mathcal{M} . Montrez que l'ensemble S^*sS^ω est un événement. On notera $\{\exists n, S_n = s\}$ cet événement. Montrez que l'ensemble $(S^*s)^\omega$ est un événement. On notera $\{\exists^\infty n \in \mathbb{N}, S_n = s\}$ cet événement.*

On peut mesurer la probabilité d'un événement :

Théorème 1. *Pour toute chaîne de Markov \mathcal{M} , il existe une unique mesure de probabilité \mathbb{P}_{s_0} qui associe à tout événement $E \subseteq S^\omega$ sa probabilité $\mathbb{P}_{s_0}(E) \in [0, 1]$, avec les propriétés suivantes :*

$$\begin{aligned} \text{Unité :} & \quad \mathbb{P}_{s_0}(S^\omega) = 1 \\ \text{Complémentarité :} & \quad \mathbb{P}_{s_0}(\neg E) = 1 - \mathbb{P}_{s_0}(E) \\ \text{Additivité :} & \quad \text{si } E \cap F = \emptyset \text{ alors } \mathbb{P}_{s_0}(E \cup F) = \mathbb{P}_{s_0}(E) + \mathbb{P}_{s_0}(F) \\ \sigma\text{-additivité :} & \quad \mathbb{P}_{s_0}\left(\bigcup_{n \in \mathbb{N}} E_n\right) = \lim_{m \rightarrow \infty} \mathbb{P}_{s_0}\left(\bigcup_{n \leq m} E_n\right) \\ \text{Cylindres :} & \quad \mathbb{P}_{s_0}(t_0 \cdots t_n S^\omega) = \begin{cases} 0 & \text{si } t_0 \neq s_0 \\ p(t_0, t_1)p(t_1, t_2) \cdots p(t_{n-1}, t_n) & \text{sinon.} \end{cases} \end{aligned}$$

On ne démontrera pas ce théorème. Une preuve est disponible dans [?].

La propriété suivante nous servira très souvent.

Théorème 2 (Propriété de Markov). *Soit \mathcal{M} une chaîne de Markov avec un ensemble fini ou dénombrable d'états \mathbf{S} et E un événement. Pour toute suite finie $s_0, \dots, s_n \in \mathbf{S}^*$, on note $(s_0 \cdots s_n)^{-1}E$ l'événement :*

$$(s_0 \cdots s_n)^{-1}E = \{s_{n+1}s_{n+2} \cdots \in S^\omega \mid s_0 \cdots s_n s_{n+1}s_{n+2} \cdots \in E\} .$$

Alors :

$$\mathbb{P}_{s_0}(E \mid s_0 \cdots s_n \mathbf{S}^\omega) = \mathbb{P}_{s_n}((s_0 \cdots s_{n-1})^{-1}E) . \quad (1.1)$$

Démonstration. Soit \mathcal{E} l'ensemble des évènements $E \subseteq \mathbf{S}^\omega$ qui vérifient 1.1. Alors \mathcal{E} contient les cylindres et est clos par complémentation et union dénombrable, donc \mathcal{E} contient tous les évènements. \square

Les propriétés d'unité, de complémentarité, d'additivité et de σ -additivité sont en fait des propriétés communes à toutes les mesures de probabilités, pas seulement dans le cas particulier des chaînes de Markov.

Exercice 5. Montrer que si E est un évènement alors pour tout état $t \in \mathbf{S}$, l'ensemble $t^{-1}E = \{s_1 s_2 \cdots \in \mathbf{S}^\omega \mid t s_1 s_2 \cdots \in E\}$ est un évènement.

Exercice 6. Prouver que si l'axiome d'additivité est vérifié, alors pour toute famille d'évènements $(E_n)_{n \in \mathbb{N}}$, la limite $\lim_{m \rightarrow \infty} \mathbb{P}_{s_0} \left(\bigcup_{n \leq m} E_n \right)$ existe.

Exercice 7. On considère la chaîne de Markov dont l'ensemble des états est $\{\text{pile, face}\}$ et la probabilité de transition est $\frac{1}{2}$ de chaque état à chaque état. Définir un évènement dont la probabilité est $\frac{3}{4}$. Soit $E \subseteq \{\text{pile, face}\}^\omega$ un évènement. Quelle est la probabilité de $\text{pile} \cdot E$? Pour tout $n \in \mathbb{N}$, et $0 \leq m \leq 2^n$, définir un évènement $E_{n,m}$ dont la probabilité est $\frac{m}{2^n}$.

Exercice 8. Définir un évènement dont la probabilité est $\frac{\pi}{4}$.

Exercice 9. Pour les curieux. On peut définir un sous-ensemble de $\{0,1\}^\omega$ qui n'est pas un évènement. Pour cela on introduit la relation d'équivalence suivante sur \mathbf{S}^ω : deux suites sont équivalentes si la différence entre les deux réels de $[0,1]$ qu'elles codent (décomposition dyadique) est un nombre rationnel. A l'aide de l'axiome du choix, on peut construire un ensemble $E \subseteq \mathbf{S}^\omega$ qui contient exactement un représentant par classe d'équivalence. Pour x, y deux suites de $\{0,1\}^\omega$ on note $x + y$ la suite obtenue par addition bit à bit des deux réels correspondant à x et y . Montrer que si E est un évènement et x une suite alors $x + E$ est un évènement de même probabilité, pour la proba induite par la chaîne de Markov de pile ou face. Remarquer que $\{0,1\}^\omega = \cup_q q + E$, où q décrit l'ensemble des suites ultimement périodique. Montrer que si E est un évènement alors sa probabilité pour la chaîne de pile ou face est nulle. En déduire que E n'est pas un évènement.

1.2 Calcul des probabilités d'accessibilité

Etant donné un sous-ensemble des états de la chaîne de Markov, noté \mathbf{T} , on voudrait calculer la probabilité $\text{val}(s_0)$ de passer un jour par un ensemble d'états cibles $\mathbf{T} \subseteq \mathbf{S}$, en partant de l'état s_0 :

$$\text{val}(s) = \mathbb{P}_s(\mathbf{S}^* \mathbf{T} \mathbf{S}^\omega) = \mathbb{P}_s(\exists n, S_n \in \mathbf{T}) .$$

Cette probabilité est appelée la valeur d'accessibilité de l'état s_0 , ou tout simplement la valeur de l'état s_0 .

Proposition 1. *Soit \mathcal{M} une chaîne de Markov avec un ensemble fini ou dénombrable d'états \mathbf{S} et $\mathbf{T} \subseteq \mathbf{S}$ un ensemble d'états cible. La valeur d'accessibilité d'un état s vérifie l'équation suivante :*

$$\text{val}(s) = \begin{cases} 1 & \text{si } s \in \mathbf{T} \\ \sum_{t \in \mathbf{S}} p(s, t) \cdot \text{val}(t) & \text{sinon.} \end{cases} \quad (1.2)$$

Démonstration. Le cas $s \in \mathbf{T}$ est évident. Le cas $s \notin \mathbf{T}$ est une application de la propriété de Markov. Soit $E = \mathbf{S}^* \mathbf{T} \mathbf{S}^\omega$. Alors $\text{val}(s) = \sum_{t \in \mathbf{S}} \mathbb{P}_s(E \mid stS^\omega) \cdot \mathbb{P}_s(stS^\omega) = \sum_{t \in \mathbf{S}} \mathbb{P}_t(s^{-1}E) \cdot p(s, t) = \sum_{t \in \mathbf{S}} \text{val}(t) \cdot p(s, t)$. Où la première égalité est par définition des probabilités conditionnelles, la seconde par la propriété de Markov et la troisième parce que $s^{-1}E = E$ et par la propriété des cylindres. \square

Supposons un instant que le système d'équation (1.2) admette une unique solution, alors calculer les valeurs revient à résoudre ce système, ce qui peut se faire en temps cubique [Dix82].

Exercice 10. *Trouver un exemple de chaîne de Markov telle que l'équation (1.2) n'a pas une unique solution.*

Il faut donc rajouter des équations au système (1.2) pour caractériser les valeur d'accessibilité. Pour cela on s'intéresse aux états à partir desquels l'ensemble des états cible est accessible.

Définition 4. *On dit qu'un état t est accessible à partir d'un état s si il existe $s_0, \dots, s_n \in \mathbf{S}$ tels que $s_0 = s$, $s_n = t$ et pour tout $0 \leq m < n$, $p(s_m, s_{m+1}) > 0$. On note $\text{Acc}(s)$ l'ensemble des états accessibles à partir de s et $\text{Acc}^{-1}(\mathbf{T})$ l'ensemble des états à partir desquels \mathbf{T} est accessible.*

Exercice 11. *Quelle est la complexité en temps du calcul de $\text{Acc}(s)$ et $\text{Acc}^{-1}(\mathbf{T})$?*

Proposition 2. *Soit une chaîne de Markov \mathcal{M} avec un ensemble fini ou dénombrable d'états \mathbf{S} . Pour tout état $s \in \mathbf{S}$, $s \notin \text{Acc}^{-1}(\mathbf{T})$ si et seulement si $\text{val}(s) = 0$. Supposons \mathbf{S} fini. Si \mathbf{T} est accessible à partir de tout état $s \in \mathbf{S}$ alors $\text{val}(s) = 1$ pour tout $s \in \mathbf{S}$.*

Démonstration. Puisque l'évènement $\mathbf{S}^* \mathbf{T} \mathbf{S}^\omega$ est l'union dénombrable disjointe des événements $(E_p = \{p\mathbf{S}^\omega\})_{p \in (\mathbf{S} \setminus \mathbf{T})^* \mathbf{T}}$, La valeur de s est nulle si et seulement si $\mathbb{P}_s(E_p) = 0$ pour tout $p \in (\mathbf{S} \setminus \mathbf{T})^* \mathbf{T}$, ce qui par propriété des cylindres est équivalent à la définition de non accessibilité à \mathbf{T} .

Supposons \mathbf{S} fini et \mathbf{T} accessible à partir de n'importe quel état. Soit $U \subseteq \mathbf{S}$ l'ensemble des états de valeur minimale. On veut montrer que la

valeur des états de U est 1. Montrons que tout état accessible à partir de U est dans U . Pour tout $u \in U$, d'après (1.2), $\text{val}(u) = \sum_{t \in \mathbf{S}} p(t, u) \text{val}(t)$. Puisque u est de valeur minimale, alors $\forall t, \text{val}(t) \geq \text{val}(u)$ donc $\text{val}(u) \geq \sum_{t \in \mathbf{S}} p(t, u) \text{val}(u) = 1$. On en déduit que pour tout état t , si $p(t, u) > 0$ alors $t \in U$. Tous les états accessibles à partir de U sont donc dans U . Or par hypothèse, \mathbf{T} est accessible à partir de U , donc U intersecte \mathbf{T} . Comme les états de cible ont valeur 1, alors par définition de U tous les états de S ont valeur 1. \square

Exercice 12. *La seconde propriété de la proposition précédente reste t'elle vraie quand \mathbf{S} est infini dénombrable ?*

Proposition 3. *Soit \mathcal{M} une chaîne de Markov avec un nombre fini ou dénombrable d'états \mathbf{S} et un ensemble d'états cibles $\mathbf{T} \subseteq \mathbf{S}$. La valeur d'accessibilité des états est l'unique solution $v : \mathbf{S} \rightarrow [0, 1]$ du système d'équations :*

$$\begin{aligned} v(s) &= 1 && \text{si } s \in \mathbf{T}, && (1.3) \\ v(s) &= 0 && \text{si } \mathbf{T} \text{ n'est pas accessible à partir de } s, \\ v(s) &= \sum_{t \in \mathbf{S}} p(s, t) \cdot v(t) && \text{sinon.} \end{aligned}$$

Démonstration. D'après les propositions 1 et 2, les valeurs sont bien solutions de ce système d'équations.

Soit $(v_1(s))_{s \in \mathbf{S}}$ et $(v_2(s))_{s \in \mathbf{S}}$ deux solutions à (1.3). On va montrer que pour tout s , $v_1(s) \leq v_2(s)$, ce qui par symétrie prouve que $v_1 = v_2$. Soit $U \subseteq \mathbf{S}$ l'ensemble des états s qui maximisent $\phi : s \mapsto (v_1(s) - v_2(s))$. D'après (1.3), pour tout état $s \in \mathbf{S}$, $(v_1(s) - v_2(s)) = \sum_{t \in \mathbf{S}} p(s, t) \cdot (v_1(t) - v_2(t))$, donc pour tout $s \in U$ et $t \in \mathbf{S}$, si $p(s, t) \neq 0$ alors $t \in U$. Donc U contient tous les états accessibles à partir de U . Si U n'intersecte pas \mathbf{T} , alors \mathbf{T} n'est pas accessible à partir de U donc v_1 et v_2 sont nulles sur U donc le maximum de ϕ est 0. Si U intersecte \mathbf{T} , alors ϕ s'annule sur U donc le maximum de ϕ est 0. Dans les deux cas, on a montré que ϕ est négative sur \mathbf{S} c'est-à-dire que pour tout état s , $v_1(s) \leq v_2(s)$. \square

Exercice 13. *Evariste Gallois et Pescheux d'Herbinville se battent en duel, en tirant l'un sur l'autre à tour de rôle. La lutte est inégale : à chacun de ses essais Evariste a une chance sur quatre d'atteindre son adversaire, alors que Pescheux, meilleur tireur, a une chance sur trois d'atteindre Evariste. Pescheux propose la chose suivante : pour déterminer le tireur initial, on lance deux fois de suite une pièce de monnaie et Pescheux ne tirera le premier que si le résultat est deux fois pile. Cela rend-il le duel équilibré ?*

1.3 Conditions de Büchi

Dans cette section, on s'intéresse à une condition un peu plus compliquée que dans la section précédente : au lieu de s'intéresser à la probabilité d'at-

teindre l'ensemble des états cible, on s'intéresse à la probabilité de l'atteindre infiniment souvent.

On définit la valeur de Büchi d'un état par :

$$\text{valB}(s) = \mathbb{P}_s((S^*\mathbf{T})^\omega) = \mathbb{P}_s(\exists^\infty n \in \mathbb{N}, S_n = \mathbf{T}) .$$

Proposition 4. *Soit \mathcal{M} une chaîne de Markov avec un ensemble d'états \mathbf{S} fini ou dénombrable et $\mathbf{T} \subseteq \mathbf{S}$ un sous-ensemble d'états cible. La valeur de Büchi d'un état s vérifie l'équation suivante :*

$$\text{valB}(s) = \begin{cases} 0 & \text{si } \mathbf{T} \text{ n'est pas accessible à partir de } s \\ \sum_{t \in \mathbf{S}} p(s, t) \cdot \text{valB}(t) & \text{sinon.} \end{cases} \quad (1.4)$$

Démonstration. Similaire à la preuve de la proposition 1. □

On va montrer que le système (1.4) admet une unique solution. Pour cela, on introduit la notion d'état récurrent.

Définition 5. *On dit qu'un état s est récurrent si pour tout état t accessible à partir de s , s est accessible à partir de t . Dans le cas contraire, on dit que s est transient. On dit que deux états récurrents sont équivalents si et seulement si ils sont accessibles l'un à partir de l'autre. Une classe de récurrence est une classe d'équivalence pour cette relation.*

Exercice 14. *Quelle est la complexité en temps du calcul des classes de récurrences ?*

Proposition 5. *Soit $R \subseteq \mathbf{S}$ l'ensemble des états récurrents. Pour tout état $s \in \mathbf{S}$, $\mathbb{P}_s(\exists n, S_n \in R) = 1$.*

Démonstration. D'après la proposition 2, il suffit de montrer que l'ensemble R des états récurrents est accessible à partir de tous les états $s \in \mathbf{S}$. Supposons le contraire et soit U l'ensemble non vide des états à partir desquels R n'est pas accessible. Pour tout état $s \in U$, soit $N(s)$ le nombre d'états accessibles à partir de s et choisissons $u \in U$ tel que $N(u)$ est minimal. Puisque $U \cap R = \emptyset$ alors u est transient donc par définition de la transience, il existe un état $t \in \mathbf{S}$ accessible à partir de u tel que u n'est pas accessible à partir de t . Puisque tout état accessible à partir de t est accessible à partir de u , alors $N(t) < N(u)$ et $t \in U$. Cela contredit la minimalité de $N(u)$. Finalement, $U = \emptyset$. □

Exercice 15. *Quelle propriété a la suite $(p_n)_{n \in \mathbb{N}}$ définie par :*

$$p_n = \mathbb{P}_s(\exists m \leq n, S_m \in R) ?$$

La proposition suivante montre que le calcul des valeurs de Büchi d'une chaîne de Markov finie se ramène au calcul des valeurs d'accessibilité, en enlevant de l'ensemble des états cibles les états qui ne sont pas récurrents.

Proposition 6. *Soit \mathcal{M} une chaîne de Markov avec un ensemble fini d'états \mathbf{S} et un ensemble d'états cible \mathbf{T} . Soit s un état récurrent. alors la valeur de Büchi de s est $\text{valB}(s) = 1$ si la classe de récurrence de s contient un état cible et $\text{valB}(s) = 0$ sinon.*

Soit $\mathbf{T}' = \mathbf{T} \cap R$ l'ensemble des états cibles qui sont récurrents. Alors la valeur de Büchi est égale à la valeur d'accessibilité pour l'ensemble cible \mathbf{T}' .

Démonstration. Soit s récurrent dont la classe de récurrence ne contient pas d'état cible. Alors d'après la proposition 2 appliquée à la chaîne de Markov induite par la classe de récurrence de s , la valeur de s est 0.

Soit s récurrent dont la classe de récurrence contient un état cible. Soit t cet état cible. Alors d'après la proposition 2 appliquée à la chaîne de Markov induite par la classe de récurrence de s , $\mathbb{P}_s(\mathbf{S}^*t\mathbf{S}^\omega) = 1$ et $\mathbb{P}_t(t\mathbf{S}^*t\mathbf{S}^\omega) = 1$. Par application de la propriété de Markov, par récurrence, pour tout $n \in \mathbb{N}$, $\mathbb{P}_t(t(\mathbf{S}^*t)^n\mathbf{S}^\omega) = 1$. Puisque $t(\mathbf{S}^*t)^\omega = \bigcap_{n \in \mathbb{N}} t(\mathbf{S}^*t)^n\mathbf{S}^\omega$, on en déduit que $\mathbb{P}_t(t(\mathbf{S}^*t)^\omega) = 1$ et par propriété de Markov $\mathbb{P}_s((\mathbf{S}^*t)^\omega) = 1$. Puisque t est un état cible, cela prouve que $\text{val}(s) = 1$.

D'après la proposition 3, il suffit de montrer que valB vérifie bien les équations de la proposition 3, ce qui est facile à vérifier. \square

Exercice 16. *Soit \mathbf{T}'' l'ensemble des états récurrents dont la classe de récurrence contient un état cible. Montrer que la valeur de Büchi est égale à la valeur d'accessibilité à \mathbf{T}'' . Généraliser aux cas des jeux de parité.*

1.4 Vérification des systèmes probabilistes

On se donne d'une part une chaîne de Markov $\mathcal{M} = (\mathbf{S}, s_0 \in \mathbf{S}, p : \mathbf{S} \rightarrow \mathcal{D}(\mathbf{S}))$ dont chaque transition $(s, t) \in \mathbf{S}^2$ est étiqueté par une propriété $P(s, t) \in \text{Prop}$ et d'autre part une formule LTL ϕ construit sur l'ensemble des propriétés Prop .

La formule ϕ est générée par la grammaire :

$$T, \perp, \alpha \in \text{Prop}, \phi \wedge \phi, \neg \phi, \phi \vee \phi, \phi U \phi,$$

et définit un langage $L_\phi \subseteq \mathbf{S}^\omega$, qui est l'ensemble des trajectoires $s_1, s_2, \dots \in \mathbf{S}^\omega$ de la chaîne telles que $P(s_1, s_2), P(s_2, s_3), \dots$ vérifie la formule ϕ .

Notre but dans cette section est de calculer la probabilité que la formule ϕ soit vérifiée, c'est-à-dire la probabilité $\mathbb{P}_{ini}(\phi)$ définie par :

$$\mathbb{P}_{ini}(\phi) = \mathbb{P}_{s_0}(L_\phi) .$$

On décrit d'abord une approche naïve, en temps doublement exponentiel, puis une approche plus efficace, en temps exponentiel.

1.4.1 Approche naïve par automates

Une idée naturelle consiste à partir de la formule LTL, la convertir en un automate de Büchi non-déterministe généralisé, puis en automate de Büchi non-déterministe, puis en automate de Rabin déterministe. Cette chaîne de conversion est basée sur les théorèmes suivants.

Théorème 3 (Vardi, Wolper, Sistla, 1986 [CY95]). *Pour tout formule LTL ϕ dont l'ensemble des propositions atomiques est A , il existe un automate de Büchi G_ϕ sur l'alphabet $\mathcal{P}(A)$ tel qu'un mot $u \in A^\omega$ est accepté par G_ϕ si et seulement si u satisfait ϕ , et G_ϕ peut être construit en temps et espace $\mathcal{O}(2^{|\phi| \cdot \ln(|\phi|)})$.*

Faire le produit synchronisé du système initial et de l'automate de Büchi non-déterministe précédent est une idée attirante, malheureusement un tel produit synchronisé n'est plus une chaîne de Markov, à cause du non-déterminisme : on obtient ce qu'on appelle un processus de décision Markovien.

Le passage d'un automate de Büchi non-déterministe à un automate déterministe peut se faire, en utilisant la construction de Safra, au prix d'une explosion du nombre d'états et d'une classe de condition d'acceptation plus générale.

Théorème 4 (Safra [Saf88]). *Tout automate de Büchi non-déterministe avec n états peut être transformé en un automate de Rabin équivalent avec $2^{\mathcal{O}(n \ln(n))}$ états*

Rappelons la définition d'un automate de Rabin.

Définition 6. *Un automate de Rabin déterministe sur l'alphabet Σ est un automate $(Q, \Sigma, \delta : Q \times \Sigma \rightarrow Q)$ sur les mots infinis muni d'une condition d'acceptation donnée par des paires d'ensembles d'états $(E_1, F_1), \dots, (E_n, F_n)$, et un calcul est acceptant si et seulement si pour une de ces paires (E, F) , un des états de F est visité infiniment souvent et tous les états de E sont visités finiment souvent.*

On fait le produit synchronisé de l'automate de Rabin et de la chaîne de Markov de la manière suivante.

Définition 7 (Produit synchronisé). *Le produit synchronisé d'une chaîne de Markov $\mathcal{M} = (\mathbf{S}, s_0 \in \mathbf{S}, p : \mathbf{S} \rightarrow \mathcal{D}(\mathbf{S})$ étiquetée par $\lambda : \mathbf{S}^2 \rightarrow A$ et d'un automate déterministe de Rabin $\mathcal{A} = (Q, A, \delta : Q \times \mathcal{P}(\text{Prop}) \rightarrow Q)$ sur l'alphabet A est la chaîne de Markov $\mathcal{M}_{\mathcal{A}}$ dont l'ensemble des états est $\mathbf{S} \times Q$ et les probabilités de transition sont données par $p((s, q), (t, r)) = p(s, t)$ si $(q, \lambda(s, t), r)$ est une transition de \mathcal{A} et $p((s, q), (t, r)) = 0$ sinon.*

Finalement, le calcul de $\mathbb{P}_{s_0}(\phi)$ est basé sur la proposition suivante :

Proposition 7. Soit \mathcal{M}_ϕ le produit synchronisé de la chaîne de Markov \mathcal{M} et d'un automate de Rabin déterministe sur $\mathcal{P}(\text{Prop})$ qui reconnaît le langage défini par ϕ .

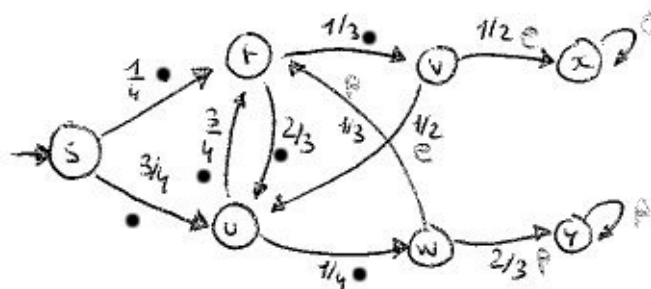
On dit qu'un état récurrent s de \mathcal{M}_ϕ est un bon état récurrent si il existe une paire (E, F) d'acceptation de l'automate de Rabin tel que la classe de récurrence de s n'intersecte pas F et intersecte E .

Alors $\mathbb{P}_{s_0}(\phi)$ est égale à la valeur d'accessibilité aux bons états récurrents dans la chaîne \mathcal{M}_ϕ à partir de l'état initial (s, q) où s est l'état initial de la chaîne de Markov et q l'état initial de l'automate.

Démonstration. Puisque l'automate de Rabin est déterministe, l'ensemble des trajectoires dans \mathcal{M}_ϕ est en bijection avec l'ensemble des trajectoires dans \mathcal{M} . De plus, une trajectoire dans \mathcal{M} satisfait ϕ si et seulement si sa projection dans l'automate de Rabin est un calcul acceptant. Puisque chaque état d'une classe de récurrence dans laquelle passe une trajectoire est visité infiniment souvent avec probabilité 1, on en déduit le théorème. \square

Exercice 17. On reprend l'exemple du duel, en rajoutant un pouvoir supplémentaire à Evariste et Pecheux : à chaque fois que ceux-ci sont touchés, ils ont respectivement une chance sur deux et deux chances sur trois d'y rester, et respectivement une chance sur deux et une chance sur trois de se relever. Dans ce cas, le duel recommence du début. On étiquette la chaîne de Markov avec deux propriétés : e et p qui étiquettent respectivement les transitions où Pecheux et Evariste sont blessés. Calculer la probabilité que la formule $\phi = (G\neg e \wedge EGp) \vee EGe$ soit satisfaite.

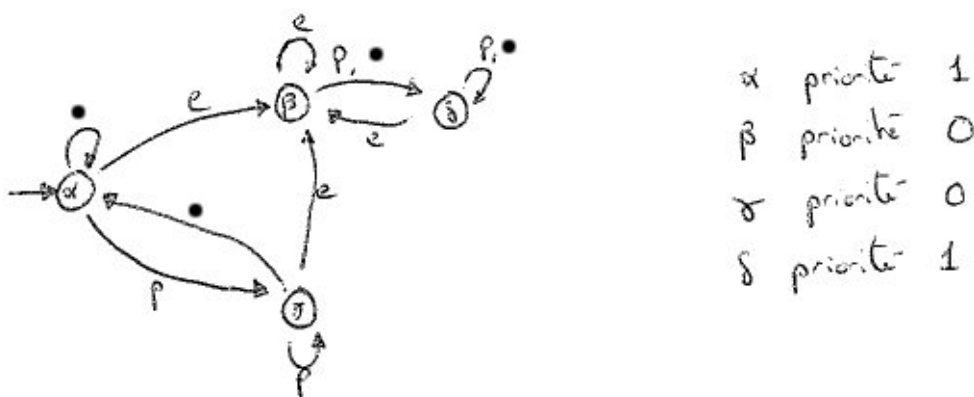
Correction de l'exercice 17 On a la chaîne de Markov suivante, qu'on appelle \mathcal{M} .



Chaque transition est étiquetée soit par \cdot , soit par p soit par e .

Le langage L_ϕ des mots de l'alphabet $\{\cdot, p, e\}$ satisfaisant ϕ est $\{p, \cdot\}^* p^\omega \cup \{p, \cdot\}^* (e\{p, \cdot\}^*)^* e^\omega$. Pour calculer un automate \mathcal{A}_ϕ acceptant L_ϕ , on peut ap-

plier l'algorithme de Safra, mais dans la majorité des cas il n'est pas possible de l'exécuter sans l'aide d'un ordinateur. On peut aussi tout simplement tatonner, sous la contrainte que l'automate résultant doit être déterministe. On trouve le candidat suivant.

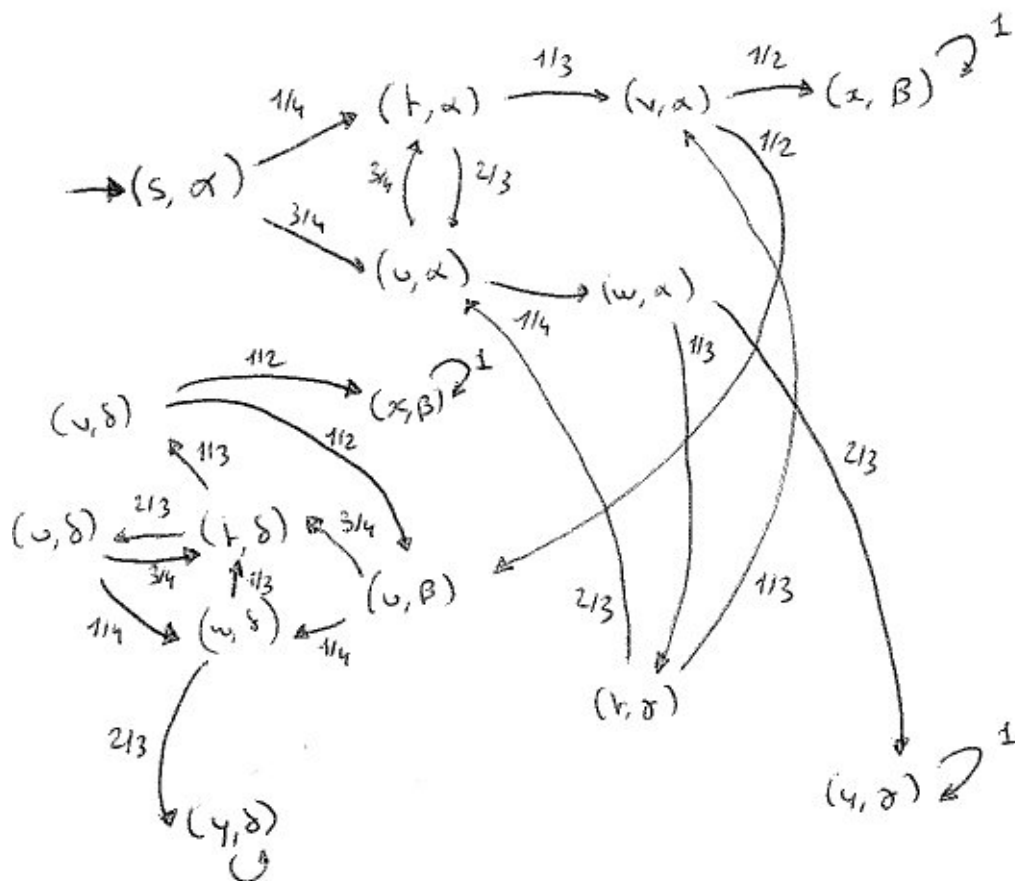


Cet automate, qu'on appelle \mathcal{A} , est un automate à parité avec deux priorités, 0 et 1.

Montrons que le langage de \mathcal{A} est bien L_ϕ . Un mot est accepté par \mathcal{A} si et seulement si le calcul sur ce mot est de priorité 0, c'est à dire si et seulement si le calcul reste ultimement soit dans l'état β soit dans l'état γ . Dans le premier cas, le mot est dans $\{p, \cdot\}^*(e\{p, \cdot\}^*)^*e^\omega$ et dans le second cas le mot est dans $\{p, \cdot\}^*p^\omega$. Cela montre que le langage de \mathcal{A} est inclus dans L_ϕ . Réciproquement, il est aisé de construire un calcul acceptant pour tout mot dans $\{p, \cdot\}^*p^\omega$, le calcul sera du type $(\alpha\gamma)^*\gamma^\omega$ et pour tout mot dans $\{p, \cdot\}(e\{p, \cdot\}^*)^*e^\omega$, le calcul sera du type $(\alpha\gamma)^*(\beta\delta)^*\beta^\omega$.

Pour calculer la probabilité $\mathbb{P}_s(\phi)$ on effectue le produit synchronisé \mathcal{M}_ϕ de la chaîne de Markov \mathcal{M} et de l'automate déterministe \mathcal{A} , et on calcule la probabilité d'accès aux bons états récurrents dans \mathcal{M}_ϕ à partir de (s, α) . On obtient la chaîne de Markov suivante, dans laquelle on a représenté seulement les états accessibles à partir de (s, α) .

1.4. VÉRIFICATION DES SYSTÈMES PROBABILISTES



Il y a trois états récurrents, (x, β) , (y, δ) et (y, γ) qui forment trois classes de récurrence distinctes. Puisque β et γ sont de priorité 0 et δ de priorité 1, les bons états récurrents sont (x, β) et (y, γ) .

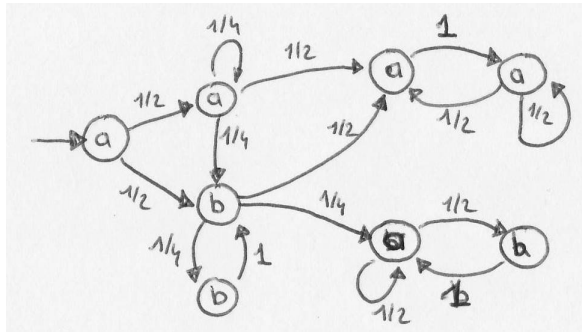
On doit calculer la probabilité d'atteindre $\{(x, \beta), (y, \gamma)\}$ à partir de (s, α) . On note sa la valeur de (s, α) , ta celle de (t, α) etc... Il y a un seul état à partir duquel $\{(x, \beta), (y, \gamma)\}$ n'est pas accessible, c'est (y, δ) .

Il suffit de résoudre le système suivant, dont on sait qu'il a une unique solution d'après la Proposition 3.

$$\left\{ \begin{array}{l} xb = 1 \\ yg = 1 \\ yd = 0 \\ sa = \frac{1}{4}ta + \frac{3}{4}ua \\ ta = \frac{1}{3}va + \frac{2}{3}ua \\ ua = \frac{3}{4}ta + \frac{1}{4}wa \\ va = \frac{1}{2}xb + \frac{1}{2}ub \\ wa = \frac{1}{3}tg + \frac{2}{3}yg \\ tg = \frac{2}{3}ua + \frac{1}{3}va \\ ub = \frac{1}{4}wd + \frac{3}{4}td \\ vd = \frac{1}{2}xb + \frac{1}{2}ub \\ ud = \frac{3}{4}td + \frac{1}{4}wd \\ td = \frac{1}{3}vd + \frac{2}{3}ud \\ wd = \frac{1}{3}td + \frac{2}{3}yd \end{array} \right.$$

On en déduit $td = \frac{12}{13}$, $ub = \frac{10}{13}$, $va = \frac{23}{26}$ et sa vaut environ 0,924, la formule est donc vérifiée avec environ 92% de chances.

Exercice 18. Calculer la probabilité que la formule LTL $FGa \vee G(b \implies Xa)$ soit vérifiée par la chaîne de Markov suivante :



On commencera par chercher un automate de co-Büchi déterministe qui reconnaît le langage de mots infinis sur l'alphabet $\{a, b\}$ spécifié par ϕ . Puis on fera le produit synchronisé de cette chaîne et de l'automate. Enfin on calculera la probabilité d'accès aux composantes fortement connexes acceptantes.

1.4.2 Approche optimale

Une meilleure approche, plus efficace algorithmiquement, est décrite dans l'article "An Optimal Automata Approach to LTL Model Checking of Pro-

1.4. VÉRIFICATION DES SYSTÈMES PROBABILISTES

probabilistic Systems", par Jean-Michel Couvreur, Nasser Saheb et Grégoire Sutre [CSS03].

Chapitre 2

Jeux stochastiques à 2 joueurs à information parfaite

Dans ce chapitre on étudie des jeux d'accessibilité stochastiques dans lesquels deux joueurs ont une influence sur les probabilités de transition. Les joueurs ont une information parfaite à propos de l'état du jeu.

2.1 Jeux d'accessibilité stochastiques à information parfaite

Ces jeux ont été étudiés pour la première fois par Anne Condon dans l'article [Con92, Con93], sous le nom de "Simple Stochastic Games".

2.1.1 Jeux et stratégies

Deux joueurs appelés Max et Min s'affrontent au cours d'un jeu, en déplaçant un jeton sur un plateau de jeu. L'objectif du joueur Max est que le jeton atteigne l'ensemble des états cibles, alors que le but du joueur Min est l'exact opposé. La partie se déroule en plusieurs tours de jeu, à chaque tour de jeu le jeton est sur un certain état, et en fonction de cet état soit le joueur Max, soit le joueur Min, soit le hasard, choisit le prochain état du jeu.

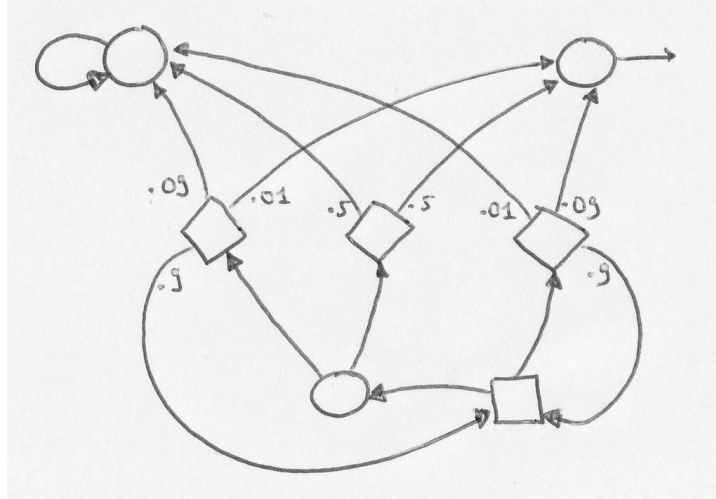
Formellement,

Définition 8 (Jeu d'accessibilité stochastique). *Un jeu d'accessibilité stochastique est un tuple $\mathcal{M} = (\mathbf{S}, \mathbf{S}_{\text{Max}}, \mathbf{S}_{\text{Min}}, \mathbf{S}_R, E, \mathbf{T}, p : \mathbf{S}_R \rightarrow \mathcal{D}(\mathbf{S}))$ où \mathbf{S} est l'ensemble des états partitionné en l'ensemble $\mathbf{S}_{\text{Max}} \subseteq \mathbf{S}$ des états contrôlés par Max, l'ensemble $\mathbf{S}_{\text{Min}} \subseteq \mathbf{S}$ des états contrôlés par Min et l'ensemble \mathbf{S}_R des états aléatoires. Les transitions déterministes sont décrites par $E \subseteq (\mathbf{S}_{\text{Max}} \cup \mathbf{S}_{\text{Min}}) \times \mathbf{S}$, chaque état a au moins un arête sortante. L'ensemble des états cible est \mathbf{T} , on suppose que ce sont des états absorbants contrôlés par Max. Comme dans une chaîne de Markov, pour tout état aléatoire $s \in \mathbf{S}_R$*

CHAPITRE 2. JEUX STOCHASTIQUES À 2 JOUEURS À
INFORMATION PARFAITE

et tout état $t \in \mathbf{S}$, on appelle $p(t, s)$ la probabilité de passer de l'état s à l'état t , avec la propriété $\sum_{t \in \mathbf{S}} p(t, s) = 1$.

Voici un exemple de jeu stochastique. Les sommets triangulaires sont les sommets de \mathbf{S}_R , les sommets ronds sont ceux de \mathbf{S}_{Max} , les sommets carrés ceux de \mathbf{S}_{Min} .



Exercice 19. Un exemple de jeu d'accessibilité stochastique est le "421". Donner d'autres exemples de jeux d'accessibilité stochastiques. Le poker et le Scrabble peuvent-ils être représentés avec ce formalisme ?

Les joueurs jouent à l'aide de *stratégies*. Pour le joueur Max une stratégie est une fonction $\sigma : \mathbf{S}^* \mathbf{S}_{\text{Max}} \rightarrow \mathbf{S}$ telle que pour toute suite $s_0, \dots, s_n \in \mathbf{S}^*$, $(s_n, \sigma(s_0, \dots, s_n)) \in E$ est une transition. La définition d'une stratégie pour le joueur Min est symétrique.

Une fois fixée deux stratégies σ et τ , on veut définir pour tout état $s \in \mathbf{S}$ la probabilité $\mathbb{P}_s^{\sigma, \tau} (\exists n, S_n \in \mathbf{T})$ d'atteindre \mathbf{T} . Pour cela on définit la chaîne de Markov dont l'ensemble des états est $s\mathbf{S}^*$, et les probabilités de transition p^* sont données par :

$$p^*(s_0 \cdots s_n, s_0 \cdots s_n s_{n+1}) = \begin{cases} 1 & \text{si } s_n \in \mathbf{S}_{\text{Max}} \text{ et } s_{n+1} = \sigma(s_0 \cdots s_n) , \\ 1 & \text{si } s_n \in \mathbf{S}_{\text{Min}} \text{ et } s_{n+1} = \tau(s_0 \cdots s_n) , \\ p(s_{n+1}, s_n) & \text{si } s_n \in \mathbf{S}_R , \\ 0 & \text{sinon.} \end{cases}$$

Par définition, $\mathbb{P}_s^{\sigma, \tau} (\exists n, S_n \in \mathbf{T})$ est la valeur d'accessibilité à $\mathbf{S}^* \mathbf{T}$ dans cette chaîne de Markov.

Un cas particulier qui nous intéresse plus particulièrement est celui du calcul des valeurs d'accessibilité quand les stratégies sont positionnelles, au sens suivant.

2.1. JEUX D'ACCESSIBILITÉ STOCHASTIQUES À INFORMATION PARFAITE

Définition 9 (Stratégies positionnelles). Une stratégie $\sigma : \mathbf{S}^* \mathbf{S}_{\text{Max}} \rightarrow \mathbf{S}$ est dite positionnelle si elle ne dépend que de l'état courant c'est-à-dire si pour toutes parties s_0, \dots, s_n et s'_0, \dots, s'_m qui terminent dans le même état $s_n = s'_m \in \mathbf{S}_{\text{Max}}$, alors $\sigma(s_0, \dots, s_n) = \sigma(s'_0, \dots, s'_m)$. Alors σ est parfaitement spécifiée par sa restriction $\sigma : \mathbf{S}_{\text{Max}} \rightarrow \mathbf{S}$. La définition d'une stratégie positionnelle pour le joueur Min est symétrique.

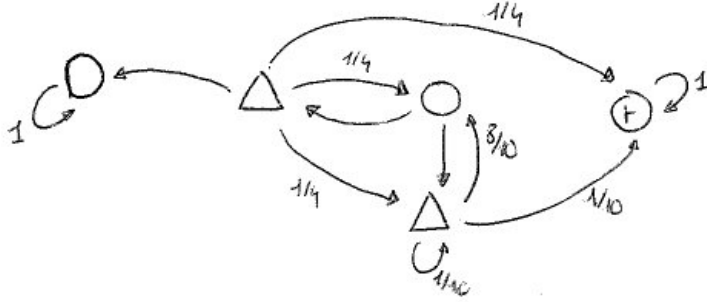
Dans le cas de stratégies positionnelles, le calcul des valeurs d'accessibilité est particulièrement facile.

Proposition 8. Soit \mathcal{M} un jeu et $\sigma : \mathbf{S}_{\text{Max}} \rightarrow \mathbf{S}$ et $\tau : \mathbf{S}_{\text{Min}} \rightarrow \mathbf{S}$ deux stratégies positionnelles. Soit $\mathcal{M}[\sigma, \tau]$ la chaîne de Markov d'états \mathbf{S} et de probabilités de transition p^* définies par :

$$\begin{cases} p^*(s, t) = p(t, s) & \text{si } s \in \mathbf{S}_R, \\ p^*(s, t) = 1 & \text{si } s \in \mathbf{S}_{\text{Max}} \text{ et } t = \sigma(s), \\ p^*(s, t) = 0 & \text{si } s \in \mathbf{S}_{\text{Max}} \text{ et } t \neq \sigma(s), \\ p^*(s, t) = 1 & \text{si } s \in \mathbf{S}_{\text{Min}} \text{ et } t = \tau(s), \\ p^*(s, t) = 0 & \text{si } s \in \mathbf{S}_{\text{Min}} \text{ et } t \neq \tau(s), \end{cases}$$

Alors $\mathbb{P}_s^{\sigma, \tau}(\exists n, S_n \in \mathbf{T})$ est la valeur d'accessibilité à \mathbf{T} dans $\mathcal{M}[\sigma, \tau]$.

Exercice 20. Calculer les probabilités d'accès à t pour les deux stratégies positionnelles de l'exemple suivant.



2.1.2 Valeurs et stratégies optimales

Le but du joueur Max est de choisir des stratégies qui rendent $\mathbb{P}_s^{\sigma, \tau}(\exists n, S_n \in \mathbf{T})$ aussi proche de 1 que possible, quelque soit la réponse τ de l'adversaire, alors que le but du joueur Min est contraire. Cela est formalisé par les notions de valeurs inférieures et supérieures.

Définition 10. Soit $s \in \mathbf{S}$ un état d'un jeu d'accessibilité stochastique. Les

valeurs inférieures et supérieures d'un état $s \in \mathbf{S}$ sont définies par :

$$\begin{aligned} \text{val}_*(s) &= \sup_{\sigma} \inf_{\tau} \mathbb{P}_s^{\sigma, \tau} (\exists n, S_n \in \mathbf{T}) \quad , \\ \text{val}^*(s) &= \inf_{\tau} \sup_{\sigma} \mathbb{P}_s^{\sigma, \tau} (\exists n, S_n \in \mathbf{T}) \quad . \end{aligned}$$

Intuitivement, la valeur inférieure $\text{val}_*(s)$ est la valeur d'un jeu où Max choisit une stratégie σ et annonce σ au joueur Min. Le joueur Min choisit alors une contre-stratégie τ aussi bonne que possible. Les deux joueurs transmettent alors à un arbitre les deux stratégies qu'ils ont choisies, puis l'arbitre calcule la probabilité d'atteindre \mathbf{T} en partant de s en utilisant σ et τ . Finalement, Min doit donner à Max la somme correspondante.

La valeur supérieure correspond à un jeu similaire, mais dans ce cas c'est Min qui choisit d'abord sa stratégie et Max qui choisit une contre-stratégie aussi bonne que possible.

Evidemment, il est avantageux de connaître la stratégie de l'autre avant de choisir sa propre stratégie, ce qui est énoncé dans la proposition suivante.

Proposition 9. *Pour tout état $s \in \mathbf{S}$, $\text{val}_*(s) \leq \text{val}^*(s)$.*

Démonstration. Soit $\epsilon > 0$. Par définition de $\text{val}_*(s)$ il existe une stratégie σ_ϵ telle que pour toute stratégie τ , $\mathbb{P}_s^{\sigma_\epsilon, \tau} (\exists n S_n \in \mathbf{T}) \geq \text{val}_*(s) - \epsilon$. Par définition de $\text{val}^*(s)$,

$$\text{val}^*(s) = \inf_{\tau} \sup_{\sigma} \mathbb{P}_s^{\sigma, \tau} (\exists n S_n \in \mathbf{T}) \geq \inf_{\tau} \mathbb{P}_s^{\sigma_\epsilon, \tau} (\exists n S_n \in \mathbf{T}) \geq \text{val}_*(s) - \epsilon.$$

Puisque c'est vrai pour toute ϵ , cela prouve l'énoncé de la proposition 9. \square

Une question naturelle est la suivante : l'inégalité de la Proposition 9 est-elle une égalité ? Dans le cas des jeux d'accessibilité, la réponse est oui. Cela reste vrai pour des conditions de gain plus complexes, telles que les conditions de parité, les évaluations de performances. Intuitivement, cela reste vrai pour tous les jeux définis de manière constructive, à somme nulle et à information parfaite, le lecteur intéressé trouvera plus de détails dans [Mar75, Mar98].

Théorème 5. *Pour tout état $s \in \mathbf{S}$, $\text{val}_*(s) = \text{val}^*(s)$ et cette valeur commune est appelée la valeur de l'état s , notée $\text{val}(s)$. De plus il existe des stratégies $\sigma^\#$ et $\tau^\#$ telles que pour toutes stratégies σ et τ ,*

$$\mathbb{P}_s^{\sigma, \tau^\#} (\exists n, S_n \in \mathbf{T}) \leq \text{val}(s) \leq \mathbb{P}_s^{\sigma^\#, \tau} (\exists n, S_n \in \mathbf{T}) \quad ,$$

les stratégies $\sigma^\#$ et $\tau^\#$ sont appelées des stratégies optimales

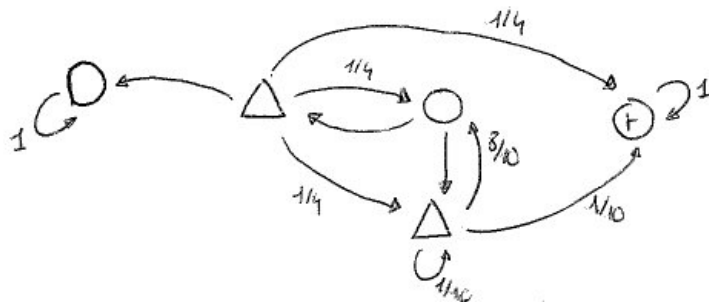
Démonstration. C'est un corollaire du théorème 6. \square

2.1. JEUX D'ACCESSIBILITÉ STOCHASTIQUES À INFORMATION PARFAITE

Appliquons le théorème 5 au 421. Si vous jouez avec une stratégies optimale vous pouvez avant de commencer à jouer, entièrement décrire votre stratégie à votre adversaire sans diminuer vos chances de gagner.

Cela reste vrai pour le poker (heads-up), mais il faut utiliser un théorème un peu plus général, qui couvre le cas des jeux à information partielle.

Exercice 21. Calculer la valeur et une stratégie optimale pour le jeu suivant.



Exercice 22. Que pensez-vous du jeu suivant ? Deux joueurs posent chacun sur la table une somme non nulle d'argent, $s_1 \in \{1, 2, 3, \dots\}$ pour le joueur 1 et $s_2 \in \{1, 2, 3, \dots\}$ pour le joueur 2. La somme $s_1 + s_2$ est ensuite distribuée de la façon suivante :

1. si $s_1 = s_2$, chacun récupère la somme qu'il a mise sur la table
2. si $s_1 < s_2$ alors le joueur 1 empoche $s_1 + s_2$,
3. sauf dans le cas où s_1 divise s_2 , dans ce cas c'est le joueur 2 qui empoche $s_1 + s_2$.
4. La règle est symétrique si $s_2 < s_1$.

Exercice 23. Si toutes les probabilités de transition sont 0 ou 1, comment calculer les valeurs ?

Comme dans le cas des chaînes de Markov, les valeurs sont solutions d'un système d'équation assez naturel.

Proposition 10. Les valeurs inférieures $val_* : \mathbf{S} \rightarrow [0, 1]$ et supérieures $val^* : \mathbf{S} \rightarrow [0, 1]$ sont solutions du système d'équation :

$$v(s) = \begin{cases} 1 & \text{si } s \in \mathbf{T} , \\ \max_{(s,t) \in E} v(t) & \text{si } s \in \mathbf{S}_{\text{Max}} , \\ \min_{(s,t) \in E} v(t) & \text{si } s \in \mathbf{S}_{\text{Min}} , \\ \sum_{t \in \mathbf{S}} p(t, s) \cdot v(t) & \text{si } s \in \mathbf{S}_R . \end{cases} \quad (2.1)$$

Démonstration. Soit $\sigma : \mathbf{S}^* \rightarrow \mathbf{S}_{\text{Max}}$ et $\tau : \mathbf{S}^* \rightarrow \mathbf{S}_{\text{Min}}$ deux stratégies optimales. On commence par montrer que pour tout $s \in \mathbf{S}_{\text{Max}}$, $\text{val}^*(s) = \max_{(s,t) \in E} \text{val}^*(t)$. Si s est un état cible c'est évident car les états cibles sont absorbants. Si s n'est pas un état cible alors,

$$\begin{aligned} \mathbb{P}_s^{\sigma, \tau}(\exists n, S_n \in \mathbf{T}) &= \mathbb{P}_s^{\sigma, \tau}(\exists n > 0, S_n \in \mathbf{T}) \\ &= \mathbb{P}_{\sigma(s)}^{\sigma[s], \tau[s]}(\exists n \geq 0, S_n \in \mathbf{T}) \quad , \end{aligned}$$

où $\sigma[s]$ est la stratégie définie par $\sigma[s](s_0 s_1 \cdots s_n) = \sigma(s s_0 s_1 \cdots s_n)$ et la seconde égalité est par application de la propriété de Markov. Quand τ parcourt l'ensemble des stratégies possibles du joueur 2 alors $\tau[s]$ également, donc

$$\text{val}^*(\sigma(s)) \geq \inf_{\tau} \mathbb{P}_{\sigma(s)}^{\sigma[s], \tau[s]}(\exists n \geq 0, S_n \in \mathbf{T}) = \inf_{\tau} \mathbb{P}_s^{\sigma, \tau}(\exists n > 0, S_n \in \mathbf{T}) = \text{val}^*(s) \quad ,$$

ce qui prouve que $\text{val}^*(s) \leq \max_{(s,t) \in E} \text{val}^*(t)$. Maintenant, supposons qu'il existe t tel que $\text{val}^*(t) > \text{val}^*(s)$ et $(s, t) \in E$, alors Max pourrait s'assurer de gagner avec probabilité $\text{val}^*(t)$ en partant de s , et ce en choisissant d'abord d'aller en t puis en suivant une stratégie optimale à partir de t , ce qui contredit la définition de la valeur supérieure de s . Finalement, $\text{val}^*(s) = \max_{(s,t) \in E} \text{val}^*(t)$.

La preuve que $\text{val}^*(s) = \min_{(s,t) \in E} \text{val}^*(t)$ quand $s \in \mathbf{S}_{\text{Min}}$ est similaire.

Maintenant, soit s un état random, la preuve que $\text{val}^*(s) = \sum_{t \in \mathbf{S}} p(t, s) \cdot \text{val}^*(t)$ est similaire au cas des chaînes de Markov, et basée sur la propriété de Markov. \square

Exercice 24. *Le système de la proposition précédente admet-t'il une unique solution en général ?*

Les stratégies optimales peuvent être choisies d'un type très simple.

Théorème 6 (Existence de stratégies optimales positionnelles). *Pour tout état $s \in \mathbf{S}$ d'un jeu d'accessibilité stochastique, avec \mathbf{S} fini, les valeurs inférieures et supérieures sont égales. De plus il existe des stratégies optimales $\sigma^\#$ et $\tau^\#$. De plus ces stratégies peuvent être choisies positionnelles.*

Elementaire, inachevée. Pour chaque état s du joueur Min, on choisit un état successeur $\tau^\#(s)$ dont la valeur inférieure est minimale :

$$\text{val}_*(\tau^\#(s)) = \min_{(s,t) \in E} \text{val}_*(t) \quad .$$

Pour chaque état s du joueur Max on note $\sigma^\#(s)$ la distribution de probabilité uniforme sur l'ensemble $A(s)$ des états t tel que $(s, t) \in E$ et

$$\text{val}^*(\sigma^\#(s)) = \max_{(s,t) \in E} \text{val}^*(t) \quad .$$

2.1. JEUX D'ACCESSIBILITÉ STOCHASTIQUES À INFORMATION
PARFAITE

On pose $\sigma^\#(s)(t) = \frac{1}{|A(s)|}$ si $t \in A(s)$ et 0 sinon.

On note $v(s) = \mathbb{P}_s^{\sigma^\#, \tau^\#} (\exists n, S_n \in \mathbf{T})$.

On va montrer que :

$$v(s) = \text{val}_*(s) = \text{val}^*(s) \quad (2.2)$$

et que $\sigma^\#$ et $\tau^\#$ sont des stratégies optimales.

Soit σ quelconque pour le joueur Max. Pour tout $n \in \mathbb{N}$, on note

$$\mathbb{E}_s^{\sigma, \tau^\#} [\text{val}_*(S_n)] = \sum_{s \in \mathbf{S}} \text{val}_*(s) \cdot \mathbb{P}_s^{\sigma, \tau^\#} (S_n = s) .$$

D'après la définition de $\tau^\#$ et la proposition 10,

$$\mathbb{E}_s^{\sigma, \tau^\#} [\text{val}_*(S_n)] \geq \mathbb{E}_s^{\sigma, \tau^\#} [\text{val}_*(S_{n+1})] .$$

On montre ensuite que

$$\mathbb{P}_s^{\sigma, \tau^\#} (\exists n, \text{val}_*(S_n) = 0 \mid \forall n, \text{val}_*(S_n) < 1) = 1 .$$

On en déduit que :

$$\mathbb{P}_s^{\sigma, \tau^\#} (\exists n, S_n \in \mathbf{T}) \leq \mathbb{E}_s^{\sigma, \tau^\#} \left[\limsup_n \text{val}_*(S_n) \right] \leq \limsup_n \mathbb{E}_s^{\sigma, \tau^\#} [\text{val}_*(S_n)] \leq \text{val}_*(s),$$

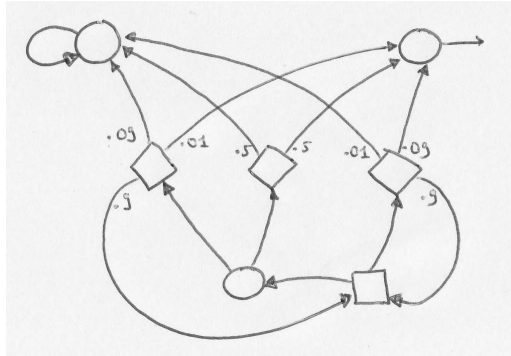
et en particulier,

$$v(s) \leq \text{val}_*(s) ,$$

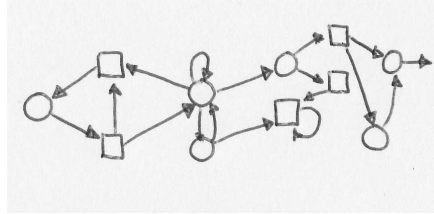
on a donc fait la moitié du travail pour prouver (2.2).

Soit τ quelconque pour le joueur Min. Alors contre $\sigma^\#$ l'espérance de la valeur supérieure au n -ième tour est au moins la valeur supérieure de l'état initial. Une partie conforme à $\sigma^\#$ soit reste ultimement dans l'ensemble des états de valeur supérieure zéro, soit visite presque-sûrement tous les états de valeur supérieure un, y compris la cible. \square

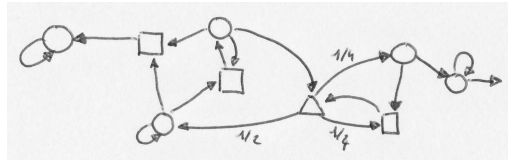
Exercice 25. Trouver des stratégies optimales au jeu suivant en se basant entre autres sur une énumération des stratégies positionnelles.



Exercice 26. Trouver des stratégies optimales au jeu déterministe suivant.



Et au jeu suivant.



Exercice 27 (Duel à trois). Evariste affronte Pecheux et Blaise en duel à trois. Pecheux et Blaise sont de fins tireurs : ils atteignent leur cible deux fois sur trois, alors qu'Evariste atteint sa cible une fois sur trois. Les joueurs tirent à tour de rôle jusqu'à ce qu'il n'y ait qu'un survivant. A chaque fois que c'est à lui de tirer, Evariste a le choix entre viser Pecheux, viser Blaise ou tirer en l'air. Pecheux et Blaise tirent systématiquement l'un sur l'autre jusqu'à être seuls contre Evariste, ou bien morts. Quelle est la stratégie optimale d'Evariste ?

2.2 Calcul des valeurs

2.2.1 Polynomial dans le cas à un joueur

Les valeurs des états sont les solutions du programme linéaire suivant,

$$\text{minimiser } \sum_{s \in \mathbf{S}} v(s)$$

sous les contraintes :

$$\begin{aligned} \forall s \in \mathbf{S}, & & v(s) &\geq 0 \\ \forall s \in \mathbf{S}_{\text{Max}}, \forall (s, t) \in E, & & v(s) &\geq v(t) \\ \forall s \in \mathbf{S}_R, & & v(s) &\geq \sum_{t \in \mathbf{S}} p(s, t) \cdot v(t) \\ \forall s \in \mathbf{T}, & & v(s) &= 1 \end{aligned}$$

Démonstration. D'après la Proposition 10, les valeurs satisfont bien les contraintes de ce programme linéaire. On peut montrer que si v est une solution du programme linéaire, qui est différente des valeurs, alors l'ensemble des sommets où la différence entre v et les valeurs est maximale est un ensemble de sommets de valeur 0, donc les valeurs sont bien la plus petite solution du programme linéaire. \square

2.2.2 Un problème dans $NP \cap co - NP$

Théorème 7. *Le problème "étant donné un jeu d'accessibilité stochastique à deux joueurs, est-ce que sa valeur est strictement supérieure à $\frac{1}{2}$?" est dans $NP \cap co - NP$.*

Démonstration. On devine un couple de stratégies (σ, τ) . On utilise le programme linéaire pour vérifier que σ est une meilleure réponse à τ et τ une meilleure réponse à σ . \square

Question ouverte ardue : ce problème peut-il être résolu en temps polynomial ? Pour attaquer le problème, on peut commencer par le cas des jeux dont le graphe est de largeur (path-width, tree-width, clique-width) bornée. Ca particuliers résolus par Philippe Rannou, élève de l'ENS Ker-Lann, lors d'un stage de L3 au LaBRI. Une autre question ouverte est décrite dans la sous-section suivante.

2.2.3 Algorithme d'amélioration des stratégies

Deviner une stratégie optimale peut être fastidieux, car il y en a un nombre exponentiel. Un algorithme plus sympathique est l'algorithme d'amélioration des stratégies.

A l'étape $n + 1$ de l'algorithme, on a une stratégie σ_{n+1} qui est "mieux" que σ_n en un sens qu'on précisera. La stratégie σ_0 est choisie au hasard. La stratégie σ_{n+1} est calculée comme suit.

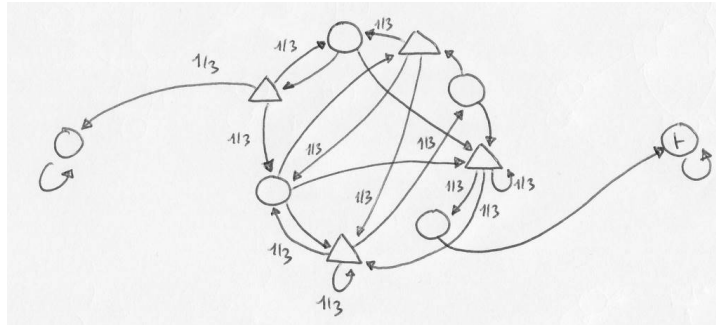
1. On calcule une meilleure réponse positionnelle τ_n à σ_n
2. On calcule les valeurs v_n pour les deux stratégies σ_n et τ_n .
3. Si pour tout sommet s du joueur Max, $\sigma_n(s)$ est de v_n maximale parmi les successeurs de s , alors σ_n est optimale et l'algo se termine. Sinon σ_{n+1} est obtenue à partir de σ_n en choisissant un successeur de v_n maximal partout où ce n'est pas le cas et en ne modifiant pas σ_n si c'est le cas.

Proposition 11. *L'algorithme d'amélioration des stratégies termine.*

Démonstration. (A compléter) La suite $(v_n)_{n \in \mathbb{N}}$ est croissante. \square

Exercice 28. *Donner une borne sur la complexité de l'algorithme d'amélioration des stratégies.*

Exercice 29. Appliquer l'algorithme d'amélioration des stratégies au jeu suivant.



Question ouverte : cet algorithme termine-t'il en temps polynomial ?

Question préliminaire : cet algorithme termine t'il en temps polynomial pour les jeux à un joueur ?

2.2.4 Algorithme des permutations

Utile pour les jeux où il y a peu de sommets random. Polynomial pour un nombre fixé de tels sommets.

Bibliographie

- [Con92] A. Condon. The complexity of stochastic games. *Information and Computation*, 96 :203–224, 1992.
- [Con93] A. Condon. On algorithms for simple stochastic games. In *Advances in computational complexity theory*, volume 13 of *DIMACS series in discrete mathematics and theoretical computer science*, pages 51–73, 1993.
- [CSS03] Jean-Michel Couvreur, Nasser Saheb, and Grégoire Sutre. An optimal automata approach to ltl model checking of probabilistic systems. In *LPAR*, pages 361–375, 2003.
- [CY95] Costas Courcoubetis and Mihalis Yannakakis. The complexity of probabilistic verification. *J. ACM*, 42(4) :857–907, 1995.
- [Dix82] J. D. Dixon. Exact solution of linear equations using p -adic expansions. *Numerische Mathematik*, 40 :137–141, 1982.
- [Mar75] D. Martin. Borel determinacy. *Annals of Mathematics*, 102 :363–371, 1975.
- [Mar98] D.A. Martin. The determinacy of Blackwell games. *Journal of Symbolic Logic*, 63(4) :1565–1581, 1998.
- [Saf88] Shmuel Safra. On the complexity of omega-automata. In *FOCS*, pages 319–327, 1988.