

Équilibres mixtes dans les jeux concurrents

Daniel STAN, encadré par Patricia Bouyer et Nicolas Markey, LSV

Mars-Août 2013

Le contexte général

La théorie des jeux, en particulier sur des graphes, s'avère efficace pour modéliser les interactions entre plusieurs systèmes informatiques complexes. Elle permet la modélisation d'interactions qui ne sont pas toujours antagonistes (les agents peuvent avoir des intérêts communs). La notion d'équilibre de Nash introduite par [3], pour lequel aucun joueur ne peut améliorer son gain en changeant unilatéralement de stratégie, se révèle pertinente. On s'intéresse ici plus particulièrement au calcul effectif de ces équilibres, dans différents type de jeux. Le traitement des stratégies mixtes, où les agents peuvent jouer aléatoirement, pose des problèmes de discrétisation, ainsi qu'une plus grande complexité algorithmique, lié au caractère arborescent des parties dans le cas probabiliste.

Ceci conduit par exemple à se concentrer sur le cas d'équilibres purs comme dans [1], en s'autorisant néanmoins des jeux très généraux (familles de gains très larges ω -réguliers, jeux concurrents). Notons aussi l'existence de résultats de complexité dans le cas probabiliste, par exemple dans [2] ou [4] dans des cadres (gains, types de parties, stratégies autorisées) plus restreints.

Le problème étudié

On s'intéresse ici à l'existence d'équilibres de Nash dans le cadre de jeux généraux tels que décrits dans [1], en se restreignant toutefois à des conditions de gains simples (gains terminaux et parfois conditions de Büchi avec pondération) On s'autorise cependant des stratégies randomisées ce qui nécessite à la fois d'adapter les outils et le formalisme. L'étude des jeux stochastiques n'est pas nouvelle, cependant on part ici d'un cadre déterministe que l'on cherche à étendre. D'autre part, l'étude s'est très vite tournée vers l'indécidabilité du problème d'existence d'un équilibre dans un cadre

(trop) général du problème : une réduction à l'aide d'un jeu à 14 joueurs est présentée dans [5] mais elle n'exploite pas le caractère concurrent des jeux et semble trop complexe pour fournir un nombre optimal d'agents. Ainsi, la question d'un possible « gap » entre décidabilité et indécidabilité lorsque le nombre de joueurs diminue se pose naturellement.

La contribution proposée

Nous avons essayé de simuler l'aléatoire présents dans certains modèles possédant des nœuds d'environnements stochastiques ([4]) ceci nous a conduit à l'expression d'une formule générique de calcul des équilibres, dans un cas simple. En combinant différents jeux et les équilibres associés, cette équation à l'équilibre s'est révélée très utile, pour la construction de jeux aux équilibres complexes. Ces jeux ont par la suite permis la simplification de la preuve d'indécidabilité de l'existence d'un équilibre dans un jeu à gains terminaux.

Les arguments en faveur de sa validité

La preuve d'indécidabilité se trouve simplifiée : on passe de 14 joueurs à 5 joueurs, en utilisant des modules simples (3 joueurs). Il semble d'une part que les équations à l'équilibre, dans le cadre de deux joueurs, possèdent une plus grande régularité, et d'autre part que l'exhibition des modules présentent des ensembles d'équilibres non triviaux. Ceci tend à faire penser que l'on pousse le modèle à ses limites et que l'on cerne mieux quels sont les facteurs responsables de l'indécidabilité du modèle.

Le bilan et les perspectives

L'exhibition et la preuve de correction des modules semble indiquer que l'ensemble des équilibres de Nash, dans le cas d'un faible nombre de joueurs, possède des propriétés de régularité, voire de fermeture. Par exemple, il a été impossible d'encoder dans un jeu l'ensemble des $\frac{1}{n}$ sans inclure l'équilibre limite à valeur 0. Ces propriétés de passage à la limite, si elles s'avèrent correctes, permettraient de discrétiser le problème, et d'appliquer par la suite les techniques éprouvées dans le cadre déterministe, comme le jeu des suspects présentés dans [1]. De plus, il sera intéressant de considérer des classes de stratégies intermédiaires, notamment entre les stratégies à mémoire totale et positionnelles.

Table des matières

Introduction	4
1 Jeux concurrents et équilibres	4
1.1 Définitions	4
1.1.1 Stratégies	5
1.1.2 Déroulement d'un jeu	5
1.1.3 Les équilibres de Nash	6
1.2 Exemples	7
1.3 Outils	9
2 Modules de comptage	10
2.1 Motivations	10
2.2 Constructions	10
2.3 Correction	11
2.4 Exclusivité	13
3 Indécidabilité	15
3.1 Machine à 2 compteurs	15
3.2 Encodage proposé	16
3.3 Justifications	18
3.4 Correction	18
Conclusions et perspectives	19
Bibliographie	20
Annexes	i
A Preuves des lemmes et théorèmes	i
A.1 Preuve du théorème 3	i
A.2 Preuve du lemme 2	ii
A.3 Preuve du lemme 3	ii
A.4 Preuve du lemme 4	ii
A.5 Preuve du lemme 5	iii
A.6 Preuve du lemme 6	iv
A.7 Preuve du lemme 7	iv
B Vers une réduction à trois joueurs	viii

Nous nous sommes intéressés lors de ce stage à des jeux multi-agent caractérisés par les propriétés suivantes :

- Concurrents : les agents peuvent activer des transitions simultanément, de façon indépendante
- À information parfaite : les agents connaissent les actions qui peuvent être effectuées par tous les joueurs et les gains en résultant
- Sans visibilité des actions : les agents ne connaissent pas les actions prises par les autres joueurs
- Avec mémoire : les agents connaissent l'historique des états parcourus depuis le début de la partie et choisissent leur action en conséquence
- Sans nœud stochastique : l'environnement extérieur sera nécessairement modélisé par un joueur du jeu
- On s'intéresse principalement à des conditions de gains sur les nœuds finaux, et parfois des conditions de Büchi
- Les stratégies sont randomisées : ceci est le principal ajout par rapport au cadre présenté dans [1]

On décrira dans un premier temps le formalisme employé, ainsi que les outils et exemples de base. Munis de ces outils, nous fournirons une construction de jeux simples aux équilibres non-triviaux. Ces constructions pourront ensuite être employés afin de démontrer l'indécidabilité du problème de l'existence d'un équilibre de Nash.

1 Jeux concurrents et équilibres

1.1 Définitions

Définition 1. *Un jeu concurrent fini est un n -uplet*

$$\mathcal{G} = \left\langle \text{States}, \text{Agt}, \text{Act}, \text{Mov}, \text{Tab}, (\text{Allow}_A)_{A \in \text{Agt}}, (\phi_A)_{A \in \text{Agt}} \right\rangle$$

où

- States est un ensemble fini d'états
- Agt est un ensemble fini d'agents, ou joueurs
- Act est un ensemble fini d'actions pouvant être utilisées par les joueurs dans un état
- Tab est la fonction de transition $\text{States} \times \text{Act}^{\text{Agt}} \rightarrow \text{States}$
- pour tout $A \in \text{Agt}$, $\text{Allow}_A : \text{States} \rightarrow \text{Act}$ est une fonction décrivant les actions autorisées par le joueur A dans un état donné
- pour tout $A \in \text{Agt}$ la fonction de gain de l'agent A sur les chemins (infinis) $\phi_A : \text{States}^\omega \rightarrow \mathbb{R}$

1.1.1 Stratégies

Définition 2. *En ayant connaissance de la succession d'états $h \in \text{States}^+$ ayant mené à l'état courant $\text{last}(h)$, chaque agent $A \in \text{Agt}$ va choisir une distribution sur les actions qui lui sont permises. Une stratégie pour le joueur A est une fonction σ_A satisfaisant :*

$$\forall h \in \text{States}^+ \quad \sigma_A(h) \in \text{Dist}(\text{Allow}_A(\text{last}(h)))$$

Pour $A \in \text{Agt}$, on note S_A l'ensemble des stratégies du joueur A .

On notera alors $\sigma_A(a \mid h)$ la probabilité de l'action a dans la distribution $\sigma_A(h)$. De plus, toute action $a \in \text{Act}$ peut être identifiée à une distribution, dite dégénérée, qui associe la probabilité 1 à l'action a .

Citons quelques sous-classes de stratégies :

- Les stratégies pures, lorsque pour tout h , $\sigma(h) \in \text{Act}$. On note S_A l'ensemble des stratégies pures de A .
- Stratégie constante : $\forall h, \sigma(h) = C \in \text{Dist}(\text{Act})$
- Stratégie positionnelle : $\forall h, \sigma(h) = \sigma(\text{last}(h))$

Un profil de stratégies pour \mathcal{G} est la donnée des stratégies de chaque joueur $\sigma = (\sigma_A)_{A \in \text{Agt}}$.

Lors d'une partie et dans chaque état, les joueurs calculeront séparément leurs distributions de probabilité sur les actions, c'est-à-dire de façon concurrente et indépendante. On étend alors les notations précédentes aux profils de stratégies de manière naturelle (indépendance des actions) :

$$\sigma(a_{A_1} \dots a_{A_x} \mid h) = \sigma_{A_1}(a_{A_1} \mid h) \times \dots \times \sigma_{A_x}(a_{A_x} \mid h)$$

1.1.2 Déroulement d'un jeu

Définition 3 (Sémantique). *Considérons un profil de stratégies $\sigma \in \mathbb{S}$. Et X_1 une variable aléatoire à valeurs dans States . On définit pour tout n ,*

$$\begin{aligned} A_n &\sim \sigma(X_n) && (A_n \text{ suit la distribution } \sigma(X_n)) \\ X_{n+1} &= X_n \cdot \text{Tab}(\text{last}(X_n), A_n) \end{aligned}$$

Notons que l'action A_n jouée à l'étape n dépend a priori de tout l'historique, mais que la transition jouée (fonction Tab) ne dépend que de l'état actuel $\text{last}(X_n)$.

La suite $(X_n)_{n \in \mathbb{N}_{>0}}$ peut par exemple être vue comme une chaîne de Markov infinie à valeur dans States^+ . Cette suite d'historiques est croissante

pour l'ordre préfixe, et possède donc une limite qui est un chemin infini aléatoire $X_\infty \in \text{States}^\omega$.

On note, pour tout historique, ou exécution infinie, $h \in \text{States}^+ \uplus \text{States}^\omega$ et $s \in \text{States}$ un état initial fixé :

$$\mathbb{P}^\sigma(h \mid s) = \mathbb{P}(X_{|h|} = h \mid X_1 = s)$$

Si cette probabilité est strictement positive, on dira que l'historique h est σ -accessible depuis s . Notons que cela impose $\text{first}(h) = s$.

Intuitivement, on initialise la partie à l'état s et les agents joueront selon le profil de stratégies σ .

On dira que h est σ -compatible lorsque h est σ -accessible depuis $\text{first}(h)$. On étend alors la notion de probabilité conditionnelle précédente :

$$\forall h' \in \text{States}^+ \quad \mathbb{P}^\sigma(h' \mid h) = \mathbb{P}(X_{|h'|} = h' \mid X_{|h|} = h)$$

Définition 4 (Espérance conditionnelle). *Dans la suite, on suppose que les fonctions de gains ϕ_A sont à valeurs dans un ensemble $\text{Img}(\phi_A) \subset \mathbb{R}$ fini de valeurs pour chaque agent A .*

On définit alors le gain moyen d'un joueur A , pour le profil de stratégies σ sachant l'historique (σ -compatible) h , par le réel suivant

$$\mathbb{E}^\sigma(\phi \mid h) = \sum_{x \in \text{Img}(\phi_A)} x \times \mathbb{P}^\sigma(\phi_A^{-1}(\{x\}) \mid h)$$

Intuitivement, on considère les différentes valeurs possibles de ϕ_A , que l'on pondère par les probabilités de leurs apparitions, dans une partie où l'historique h a déjà été observé.

On vérifie aisément que la fonction $\mathbb{E}^\sigma(\cdot \mid h)$ est linéaire, et qu'il est également possible d'étendre sa définition à toute fonction à valeur dans \mathbb{R}^{Agt} pour définir par exemple le gain moyen du profil de stratégies σ sachant l'historique h .

Dans la suite on considèrera uniquement des gains terminaux, c'est-à-dire que seuls les parties dont les chemins infinis finissent dans un état terminal (que l'on peut modéliser par une boucle) se verront récompensées :

$$\forall r \in \text{States}^\omega, \phi_A(r) > 0 \quad \Leftrightarrow \quad \exists h \in \text{States}^+, r = h.\text{last}(h)^\omega$$

1.1.3 Les équilibres de Nash

Définition 5. *Soit $\sigma \in \mathbb{S}$ un profil de stratégies. Une déviation (unilatérale) de l'agent $A \in \text{Agt}$ à partir de l'historique $h \in \text{States}^+$ est un nouveau*

profil de stratégies σ' tel qu'il existe $\sigma''_A \in \mathbb{S}_A$, une stratégie mixte pour A , vérifiant :

$$\forall h' \in \text{States}^+. \left\{ \begin{array}{l} h' \in \text{Suff}(h) \Rightarrow \sigma'_A(h') = \sigma''_A(h') \\ \forall B \in \text{Agt}. h' \notin \text{Suff}(h) \vee B \neq A \Rightarrow \sigma'_B(h') = \sigma_B(h') \end{array} \right.$$

On écrira $\sigma' = \sigma[A/\sigma''_A]^h$ ou plus simplement $\sigma' = \sigma[A/\sigma''_A]$ lorsque h sera réduit à l'état initial du jeu, de sorte que la stratégie de A sera altérée en tout historique accessible.

Forts de cette définition, il est désormais possible d'étendre la notion d'équilibre de Nash, telle qu'introduite dans [3] à des jeux à plusieurs tours comme définis ici.

Définition 6. Un équilibre de Nash (NE dans la suite) est un couple (σ, h) où

- σ est un profil de stratégies
- $h \in \text{States}^+$ est un historique initial
- Aucun joueur ne peut améliorer son gain moyen en changeant de stratégie (déviation), c'est-à-dire :

$$\forall A \in \text{Agt} \forall \sigma'_A \in \mathbb{S}_A \quad \mathbb{E}^{\sigma[A/\sigma'_A]^h}(\phi_A | h) \leq \mathbb{E}^\sigma(\phi_A | h)$$

Lorsque le profil de stratégie σ est dégénéré pour tous les joueurs, on parle d'équilibre pur. Autrement, l'équilibre est dit mixte.

1.2 Exemples de jeux finis à deux joueurs

Pour tout $x \in \mathbb{R}$ on pose dans la suite $\bar{x} = 1 - x$.

Considérons une des instances de jeu concurrent parmi les plus simples présentée en figure 1. Nous avons effectué une première analyse numérique des équilibres de Nash pouvant intervenir dans un tel jeu.

Pour le joueur $i \in \{0, 1\}$, notons $p_i = \sigma_i(a | s_0)$ la probabilité que celui-ci joue l'action a en s_0 . On peut alors exprimer les gains moyens des deux joueurs par la formule :

$$\mathbb{E}^\sigma(\Phi_i | s_0) = p_i (p_{\bar{i}}a_i + \bar{p}_{\bar{i}}c_i) + \bar{p}_i (p_{\bar{i}}b_i + \bar{p}_{\bar{i}}d_i)$$

Cette expression est fonction affine de p_i .

Ainsi, un profil de stratégies non dégénérées (pour chaque joueur), caractérisé par la paire $(p_0, p_1) \in]0, 1[^2$, est un équilibre si et seulement si les

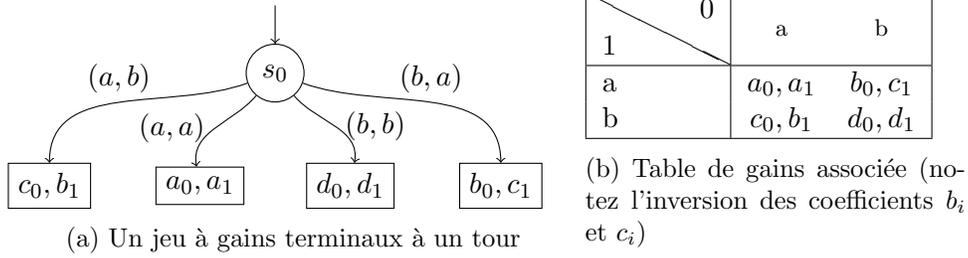


FIGURE 1

deux fonctions sont de pente nulle :

$$\begin{aligned} \forall i, \frac{\partial \mathbb{E}^\sigma(\Phi_i | s_0)}{\partial p_i}(p_i) = 0 &\Leftrightarrow \forall i, (p_i a_i + \bar{p}_i c_i) = (p_i b_i + \bar{p}_i d_i) \\ &\Leftrightarrow \forall i, [(d_i - c_i) + (a_i - b_i)] \cdot p_i = d_i - c_i \end{aligned}$$

Cette dernière équation s'est révélée particulièrement utile dans la construction intuitive de jeux concurrents pour lesquels nous souhaitons imposer les probabilités de transition. On se heurte cependant rapidement aux cas particuliers où au moins une des stratégies est dégénérée, et la condition précédente n'est plus nécessaire à l'équilibre (elle reste néanmoins une condition suffisante). En exprimant précisément les contraintes sur les pentes, on obtient une condition plus générique, valable pour tout profil de stratégie σ pour deux joueurs :

$$\forall i \begin{cases} \sigma_i(a | s_0) < 1 \Rightarrow [(d_i - c_i) + (a_i - b_i)] \cdot \sigma_{1-i}(a | s_0) \leq d_i - c_i \\ \sigma_i(a | s_0) > 0 \Rightarrow [(d_i - c_i) + (a_i - b_i)] \cdot \sigma_{1-i}(a | s_0) \geq d_i - c_i \end{cases} \quad (1)$$

Notons que dans le cas d'un profil de stratégies non dégénérées, les deux conditions d'équilibres coïncident.

À titre d'exemple, considérons le tableau de gain présenté en figure 2 représentant un exemple classique de jeu, appelé *matching pennies* (un joueur cherche à obtenir la même action, tandis que l'autre adversaire doit jouer une action différente pour gagner). Notons que l'inversion du sens des inégalités dans la condition d'équilibres (puisque $d_i - c_i = a_i - b_i = -1$) impose des stratégies non dégénérées pour les deux joueurs. Au vu de la condition d'équilibre, on en déduit que le seul équilibre de Nash d'un jeu *matching pennies* est obtenu lorsque les deux joueurs jouent de manière uniforme.

	0	a	b
1			
a	1, 0	0, 1	
b	0, 1	1, 0	

$$\forall i \in \{0, 1\}, \begin{cases} \sigma_i(a | s_0) < 1 \Rightarrow \sigma_{1-i}(a | s_0) \geq \frac{1}{2} \\ \sigma_i(a | s_0) > 0 \Rightarrow \sigma_{1-i}(a | s_0) \leq \frac{1}{2} \end{cases}$$

FIGURE 2 – Table de gains d'un jeu simple

1.3 Outils

En plus des quelques exemples vus précédents, nous pouvons citer le théorème suivant dû à Nash dans [3], qui traite du cas de jeux finis :

Théorème 1. *Soit \mathcal{G} un jeu fini à un tour. Alors, il existe un équilibre de Nash.*

Ce théorème s'applique dans le cas très particulier des jeux à un seul tour. Il est toutefois possible d'effectuer diverses généralisations pour montrer l'existence d'équilibres dans un jeu avec un nombre fini de tours. Cependant, dans le cas général d'un jeu infini comme étudié ici, ce théorème peut ne pas s'appliquer et pour cause : [5] montre qu'il existe des jeux avec gains terminaux qui ne possèdent pas d'équilibres de Nash.

Bien que les parties soient infinies, il est souvent possible de se réduire à des parties finies, ou tout du moins, d'« élaguer » une partie de l'arbre de la partie, grâce au théorème suivant :

Théorème 2 (Sous-jeu). *Soit $h_0, h_1 \in \text{States}^+$ deux historiques de \mathcal{G} et σ un profil de stratégies.*

On définit \mathcal{G}' à partir de \mathcal{G} en donnant de nouvelles fonctions de gains ϕ' :

$$\forall r \in \text{States}^\omega, \phi'(r) = \begin{cases} \mathbb{E}^\sigma(\phi | h_0 h_1) & \text{si } h_0 h_1 \text{ est préfixe de } r \\ \phi(r) & \text{sinon} \end{cases}$$

On suppose que

- (σ, h_0) est un équilibre de Nash de \mathcal{G}
- $h_0 h_1$ est σ -accessible depuis h_0

Sous ces hypothèses, on a :

- $(\sigma, h_0 h_1)$ est un équilibre de Nash de \mathcal{G}
- (σ, h_0) est toujours un équilibre de Nash de \mathcal{G}'

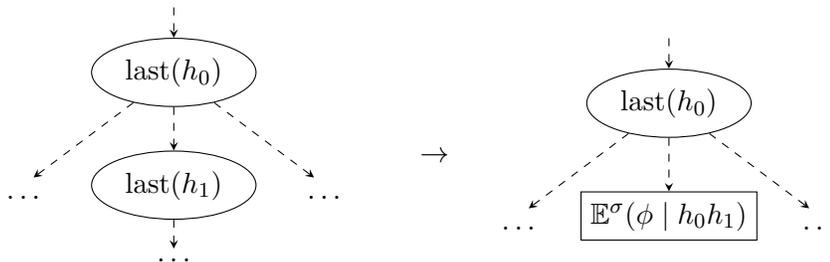


FIGURE 3 – Construction de \mathcal{G}' en remplaçant le sous-jeu en h_1

2 Modules de comptage

2.1 Motivations

On s'intéresse dans cette partie à la génération de gains décrivant des ensembles discrets mais infinis de valeurs. Plus précisément, on cherche à construire des jeux \mathcal{G} tels qu'en se restreignant aux parties où le premier joueur obtient 1, l'ensemble des gains $\mathbb{E}^\sigma(\phi)$ possibles pour σ équilibre de Nash est un ensemble $f(\mathbb{N} \uplus \{+\infty\})$ pour une fonction f donnée.

La construction de tels modules a pour objectif de coder des ensembles de gains de la forme $\{\frac{1}{n} \mid n \in \mathbb{N}\}$ ce qui permettra par la suite de modéliser des compteurs.

2.2 Constructions

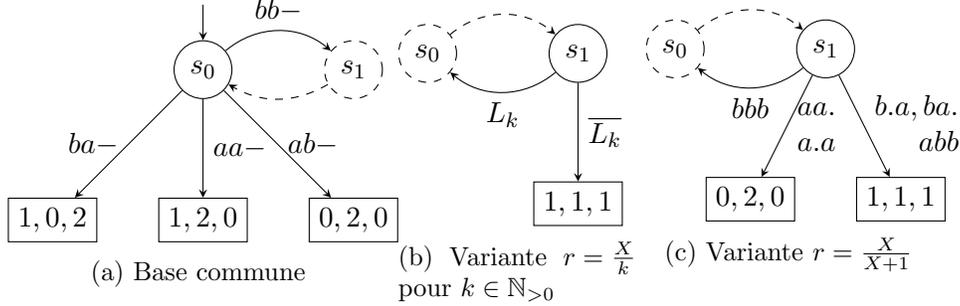
Considérons les différentes variantes d'un jeu \mathcal{G}_r , présentées en figure 4, jouées par les trois agents 0, A_1 et A_2 (dans cet ordre sur les arêtes).

La figure 4a présente ainsi la base commune, tandis que les figures 4c et 4b présentent des alternatives à la définition du nœud s_1 . \mathcal{G}_r est ainsi défini pour différentes fonctions réelles r parmi :

- $r : x \mapsto \frac{x}{x+1}$
- $r : x \mapsto \frac{x}{k}$ pour toute valeur entière de k

On aura défini les actions autorisées pour chaque joueur, de manière implicite : $a, b, -$ et $n \in \mathbb{N}$ désignent des actions et \cdot représente « n'importe quelle action ». On définit également le langage L_k par :

$$\begin{cases} L_k = \{-\} \times \{(x, x) \mid x \in \{1 \dots k\}\} \\ \overline{L}_k = \{-\} \times \{(x, y) \mid x, y \in \{1 \dots k\} \wedge x \neq y\} \end{cases}$$

FIGURE 4 – Les variantes du jeu \mathcal{G}_r

Intuitivement, l'état s_1 du jeu \mathcal{G}_r permettra de simuler la fonction r sur les gains des joueurs A_1 et A_2 . Par exemple, pour $\mathcal{G}_{X/2}$, le jeu obtenu est un *matching pennies*, où l'un des joueurs cherchera à arriver dans l'état $(1, 1, 1)$ tandis que l'autre joueur cherchera à continuer la partie dans l'état s_0 . Cette partie ne sera équilibrée (au sens de Nash) que pour une stratégie où A_1 et A_2 choisissent leurs actions de manière uniforme, c'est-à-dire ici avec probabilité $\frac{1}{2}$.

Nous verrons d'autre part que s_0 permet de coder un choix non-déterministe, résultant de la coalition des deux joueurs 0 et A_1 , pour possiblement :

- continuer la partie, c'est-à-dire réitérer la fonction r
- ou arrêter la partie immédiatement dans l'état $(1, 2, 0)$

On montrera ainsi que les gains moyens générés, en se restreignant à certains types d'équilibres de Nash, sont de la forme $(1, 1 + r^{(n)}(1), 1 - r^{(n)}(1))$ pour un certain $n \in \mathbb{N} \uplus \{+\infty\} = \bar{\mathbb{N}}$. On notera dans la suite I_r l'ensemble de ces éléments.

2.3 Correction

On peut déjà vérifier que l'état $(1, 2, 0)$ de fin de partie depuis s_0 correspond bien à $n = 0$ itérations, et même de manière plus générale :

Lemme 1. Soit (σ, hs_0) un NE de \mathcal{G}_r , avec $h \in \text{States}^+$, tel qu'il existe $x \geq 0$ vérifiant

$$\mathbb{E}^\sigma(\phi \mid hs_0) = (1, 1 + x, 1 - x)$$

Supposons de plus que (σ, h) est un NE. On a alors :

$$\mathbb{E}^\sigma(\phi \mid h) = (1, 1 + r(x), 1 - r(x))$$

TABLE 1 – Projections sur $(0, A_1)$ et $(0, A_2)$

	0		
A_1		a	b
a		0, 2	1, 1
b		1, 1	$1, 1 + x$

$$(x + 1)\sigma_0(a | h) \leq x$$

	0		
A_2		a	b
a		0, 0	1, 1
b		1, 1	$1, 1 - x$

$$(-x - 1)\sigma_0(a | h) \leq -x$$

L'encadrement de $\sigma_0(a | h)$ est obtenu en appliquant l'équation (1) vue en section 1.2, car $\sigma_1(a | h) = 0 < 1$.

La preuve utilise ainsi les équations à l'équilibre obtenues en section 1.2.

Démonstration. Premièrement, on voit que l'agent 0 peut toujours s'assurer un gain égal à 1 donc $\mathbb{E}^\sigma(\phi_0 | h) = 1$ est vérifié. On montre le résultat séparément pour les deux différentes structures de jeux. On applique le théorème 2 vu en section 1.3 pour se ramener à un jeu à un tour, où l'état s_0 est vu comme terminal à gain $(1, 1 + x, 1 - x)$.

– **Cas** $r = \frac{x}{k}$

En jouant uniformément, A_1 (respectivement A_2) s'assure un gain supérieur à $1 + \frac{x}{k}$ (respectivement $1 - \frac{x}{k}$). On sait de plus que $\phi_1 + \phi_2 = 2$ dans chaque état terminal.

On en déduit que $\mathbb{E}^\sigma(\phi_1 | h) = 1 + \frac{x}{k} = 2 - \mathbb{E}^\sigma(\phi_2 | h)$

– **Cas** $r = \frac{x}{x+1}$

Comme $\mathbb{E}^\sigma(\phi_0 | h) = 1$, la transition vers $(0, 2, 0)$ est évitée. On peut ainsi considérer dans la suite que $x > 0$, car le cas $x = 0$ est résolu trivialement ($r(x) = x$ et les états $(1, 1, 1)$ et s_0 fournissent le même gain).

On sait de plus que $\sigma_0(a | h) > 0$, car sinon, 0 jouerait b et A_2 jouerait a pour maximiser son gain strictement ($1 > 1 - x$). Ceci signifierait que hs_0 est inaccessible par σ , ce qui contredit nos hypothèses.

Comme la transition vers $(0, 2, 0)$ est évitée, on a

$$\sigma(aa., a.a | h) = \sigma_0(a | h) (\sigma_1(a | h) + \sigma_2(a | h)) = 0 \Rightarrow \sigma_1(h) = \sigma_2(h) = b$$

On peut alors regarder le jeu en h comme un jeu à deux joueurs où l'on a fixé l'action du troisième joueur. La « projection » ainsi obtenue est toujours un équilibre de Nash. En projetant sur $(0, A_1)$ puis sur $(0, A_2)$, on obtient deux tableaux (table 1) de gains dans un jeu à deux joueurs.

On en déduit ainsi par encadrement $\sigma_0(a | h) = \frac{x}{x+1}$ puis :

$$\mathbb{E}^\sigma(\phi_1, \phi_2 | h) = 1 + \left(1 - \frac{x}{x+1}\right) \cdot x \cdot (1, -1) = (1 + r(x), 1 - r(x))$$

□

Théorème 3. *Fixons \mathcal{G}_r un des jeux précédemment défini. Pour tout $n \in \overline{\mathbb{N}}$, il existe un profil de stratégies σ^n tel que (σ^n, s_0) est un NE de \mathcal{G}_r avec gain moyen*

$$\mathbb{E}^{\sigma^n}(\phi | s_0) = \left(1, 1 + r^{(n)}(1), 1 - r^{(n)}(1)\right)$$

2.4 Exklusivité

La difficulté dans la construction des modules réside en réalité dans l'exklusivité des gains moyens obtenus : très rapidement, les jeux construits génèrent un ensemble continu de gains. On rajoute ici une condition sur les NE que l'on considère. En présence de cette restriction, on montrera que le gain moyen, de n'importe quel équilibre de Nash, dans l'un des jeux \mathcal{G}_r précédemment définis, est nécessairement dans l'ensemble

$$I_r = \left\{ \left(1, 1 + r^{(n)}(1), 1 - r^{(n)}(1)\right) \mid n \in \overline{\mathbb{N}} \right\}$$

Définition 7 (Sûreté). *Soit (σ, h) un NE de \mathcal{G}_r . On dira que σ est un profil de stratégies sûr à partir de h s'il attribue un gain de 1 au premier joueur c'est-à-dire :*

$$\mathbb{E}^\sigma(\phi_0 | h) = 1$$

Cette condition de sûreté impose que ces transitions ne peuvent être prises avec probabilité non-nulle :

Lemme 2. *Soit (σ, h) un NE sûr de \mathcal{G}_r*

Alors les actions choisies par les joueurs ne permettent pas d'atteindre l'état $(0, 2, 0)$ avec probabilité non-nulle. C'est-à-dire :

$$\text{last}(h) = s_1 \quad \Rightarrow \quad \sigma(aa. | h) = \sigma(a.a | h) = 0 \quad \text{dans le cas du jeu } \mathcal{G}_{X/(X+1)}$$

$$\text{last}(h) = s_0 \quad \Rightarrow \quad \sigma(ab- | h) = 0$$

En plus du joueur 0 que nous contraignons à un gain égal à 1, nous pouvons remarquer que le joueur A_1 peut toujours s'assurer un gain final :

Lemme 3. *Soit σ un profil de stratégie et hs_1 un historique σ -compatible. Il existe une déviation de A_1 lui assurant un gain moyen strictement positif, c'est-à-dire :*

$$\forall \sigma \in \mathbb{S} \quad \exists \sigma'_1 \in \mathbb{S}_{A_1} \quad \mathbb{E}^{\sigma[A_1/\sigma'_1]}(\phi_1 \mid hs_1) > 0$$

Lemme 4. *Soit (σ, hs_0) un NE sûr. Alors les joueurs jouent en hs_0 de manière déterministe ($\sigma(hs_0)$ est dégénérée). De plus, 0 et A_1 jouent la même action.*

Forts de ces lemmes préliminaires, nous allons pouvoir caractériser le gain moyen final, en fonction du nombre de passages par l'état s_1 dans notre jeu \mathcal{G}_r . Ce nombre de passages correspondra en effet au nombre d'itérations de la fonction r . Plutôt que de définir un nombre de passages, quantité mal définie dans ce cadre probabiliste, on introduit la notion suivante de « support d'un profil de stratégie » :

Définition 8. *On définit le support du couple $(\sigma, h) \in \mathbb{S} \times \text{States}^+$, pour h historique σ -accessible, comme l'ensemble des historiques accessibles depuis h aboutissant dans un état non-terminal :*

$$\text{Supp}(\sigma, h) = \{h' \mid \mathbb{P}^\sigma(h' \mid h) > 0 \wedge \text{last}(h') \in \{s_1, s_0\}\}$$

On s'intéresse premièrement au cas des équilibres de Nash à support fini, où l'on peut raisonner par récurrence.

Lemme 5. *Pour tout (σ, h) un NE sûr de \mathcal{G}_r à support fini, on a*

$$\mathbb{E}^\sigma(\phi \mid h) \in I_r$$

Le cas d'un support infini laisse naturellement penser qu'un nombre infini d'itérations de la fonction r a été appliqué. C'est effectivement le cas, comme l'indique le lemme suivant :

Lemme 6. *Pour tout (σ, h) un NE sûr de \mathcal{G}_r à support infini, on a*

$$\mathbb{E}^\sigma(\phi \mid h) = (1, 1, 1)$$

On conclut la construction des modules par le théorème suivant, qui résume les propriétés démontrées sur les jeux \mathcal{G}_r :

Théorème 4. *Pour \mathcal{G}_r précédemment défini. L'ensemble des gains moyens obtenus, lorsque les agents jouent un équilibre de Nash sûr, est exactement l'ensemble $I_r = (1, 1 + r^{(n)}(1), 1 - r^{(n)}(1))$.*

3 Résultat d'indécidabilité

Dans cette partie, on s'intéresse au problème de décision suivant :

Problème. *Étant donné un jeu concurrent \mathcal{G} à gain terminaux, un état initial s_0 et $\varphi : \mathbb{R}^{\text{Act}} \rightarrow \mathbb{R}$ une forme affine, existe-t-il un profil de stratégies σ tel que (σ, s_0) est un équilibre vérifiant $\varphi(\mathbb{E}^\sigma(\phi \mid s_0)) \geq 0$?*

Dans le cas de jeux tour à tour, [5] démontre que le problème est indécidable —et même qu'il n'est pas co-récurivement énumérable— dès que le jeu possède au moins 14 joueurs. La preuve proposée s'appuie sur la réduction d'une machine à deux compteurs où chaque compteur est codé par 6 joueurs, les deux derniers joueurs sont chargés de la mise à jour des compteurs et de l'état courant. La contrainte sur le gain final n'est spécifiée que pour un joueur, et modulo adaptation légère, reste valable même dans le cas où les gains sont des entiers naturels.

On cherche ici à obtenir un résultat analogue en exploitant le caractère concurrent. La motivation est double : fournir une preuve simplifiée mais aussi obtenir plus d'informations sur la borne inférieure sur le nombre de joueurs rendant le problème indécidable, et cerner les critères qui rendent le problème indécidable.

3.1 Description d'une machine à deux compteurs

Nous allons pour cela réduire le problème de l'arrêt d'une machine à deux compteurs, connu pour son indécidabilité.

Définition 9. *Une machine à deux compteurs \mathcal{M} est un triplet (Q, q_0, Δ) où*

- Q est un ensemble fini d'états
- $q_0 \in Q$ est l'état initial
- $\Delta \subseteq Q \times \Gamma \times Q$ constitue la table de transition avec $\Gamma = \{\text{inc}(j), \text{dec}(j), \text{zero}(j) \mid j \in \{1, 2\}\}$ ensemble des instructions de la machine.

Définition 10 (Sémantique). *Une configuration de la machine est un triplet $C = (q, c_1, c_2) \in Q \times \mathbb{N}^2$.*

On note alors $q\Delta = \{(\gamma, q') \mid (q, \gamma, q') \in \Delta\}$. L'ensemble des transitions possibles depuis C est alors défini pour tout $(\gamma, q') \in q\Delta$ par :

- $C \xrightarrow{\gamma} (q', c_1 + (2 - j), c_2 + (j - 1))$ si $\gamma = \text{inc}(j)$
- $C \xrightarrow{\gamma} (q', c_1 - (2 - j), c_2 - (j - 1))$ si $\gamma = \text{dec}(j)$ et $c_j > 0$

– $C \xrightarrow{\gamma} (q', c_1, c_2)$ si $\gamma = \text{zero}(j)$ et $c_j = 0$

Une configuration $C = (q, c_1, c_2)$ est dite terminale si $|q\Delta| = 0$.

Notons que l'ensemble $q\Delta$ peut contenir plus d'un élément et que la machine peut donc être non déterministe. Cette construction non déterministe est notamment utile à la réalisation simultanée d'un test à zéro sur le compteur c_j (instruction $\text{zero}(j)$) et un test de non-égalité à zéro sur le même compteur (via l'instruction $\text{dec}(j)$) ce qui permet de réaliser un branchement déterministe en fonction de l'état d'un compteur.

3.2 Encodage proposé

Nous allons réduire le problème du non-arrêt d'une machine à deux compteurs au problème de l'existence d'un équilibre de Nash avec contraintes, au sens large, sur les gains moyens. Ce dernier problème sera prouvé indécidable, et plus exactement non récursivement énumérable. On considère dans la suite une machine à deux compteurs $\mathcal{M} = (Q, q_0, \Delta)$ fixée. On cherche à coder ses exécutions infinies dans un équilibre de Nash mixte σ dans un jeu à cinq joueurs. On indiquera sur le joueur 0 une contrainte large sur son gain moyen : celui-ci devra être supérieur ou égal à 1. Ceci correspond à la contrainte de sûreté vue dans les modules d'itérations précédents. Ces modules seront en effets réutilisés et joués alternativement par le joueur 0, A_1^0 et A_2^0 puis par A_1^1 et A_2^1 . Les modules d'itération permettront d'effectuer les opérations de vérification des compteurs, en s'assurant que ces derniers possèdent bien la forme adéquate. L'introduction des deux joueurs supplémentaires permettra de répliquer l'état des compteurs depuis un état de la machine vers le suivant, à l'aide de pondérations qui permettront la mise à jour.

Pour tout $q \in Q$, $\gamma \in \Gamma$ et $t \in \{0, 1\}$, on définit un état de « début de transition vers q » (en appliquant l'opération γ), $T_{\gamma q}^t$, suivi d'un état « de mise à jour de compteurs » $C_{\gamma q}^t$ qui se termine par un état « de branchement non-déterministe » en q^t . La succession de ces trois états est schématisée en figure 5 et peut être vue comme un cycle de calcul de la machine à deux compteurs.

On désignera par $\{\delta_1 = (q_1, \gamma_1) \dots \delta_k = (q_k, \gamma_k)\} = q\Delta$ l'ensemble des transitions qui peuvent être prises depuis q , c'est-à-dire le prochain état q' qui sera visité à partir de q , en appliquant l'opération γ' sur ses compteurs.

Pour $t \in \{0, 1\}$ fixé, l'état $C_{\gamma q}^t$ fait appel à un nouvel état γ^t , défini en table 2, qui sera joué par les joueurs 0, A_1^t et A_2^t , les joueurs $A_1^{\bar{t}}$, $A_2^{\bar{t}}$ étant uniquement observateurs de la partie. On complète le jeu \mathcal{G}_r en définissant

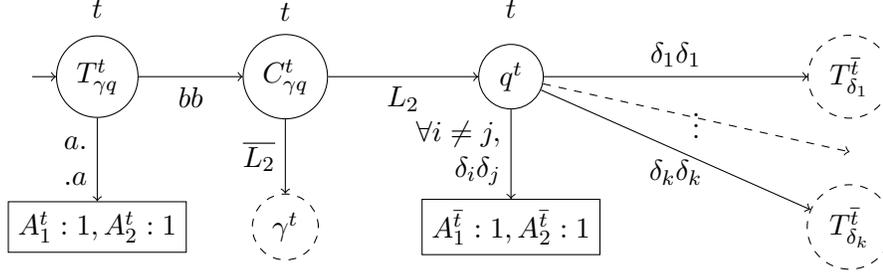


FIGURE 5 – Cycle de mise à jour de l'opération γ , de parité t , avec arrivée dans l'état q . Le label au dessus de chaque état indique le couple de joueurs acteurs de la partie.

TABLE 2 – Définition du module γ^t en fonction de l'opération γ

opérateur γ	module	$K(\gamma)$	ϕ_1^t	ϕ_2^t
init	état-terminal	1	2	0
zero(i)	$\mathcal{G}_{\frac{X}{3-i}}$	1	$1 + \frac{1}{(3-i)^n}$	$1 - \frac{1}{(3-i)^n}$
inc(i)	$\mathcal{G}_{X/(X+1)}$	$i + 1$	$1 + \frac{1}{n+1}$	$1 - \frac{1}{n+1}$
dec(i)	$\mathcal{G}_{X/(X+1)}$	$\frac{1}{i+1}$	$1 + \frac{1}{n+1}$	$1 - \frac{1}{n+1}$

On réutilise les jeux \mathcal{G}_r augmentés de deux joueurs spectateurs A_0^t et A_1^t . Les gains moyens, dans le cas d'équilibres sûrs, sont exprimés en fonction de $n \in \bar{\mathbb{N}}$.

le gain des deux derniers joueurs par l'équation linéaire suivante :

$$\forall j \in \{1, 2\}, \phi_j^t = \phi_0 + 2K(\gamma) (\phi_{1-j}^t - \phi_0)$$

où $K(\gamma)$ est une constante non nulle dépendant de l'opération γ considérée. Notons l'ajout d'une instruction supplémentaire à Γ , qui initialisera les compteurs en début d'exécution.

En pratique, on se restreindra à des équilibres purs, et on aura donc $\mathbb{E}^\sigma(\phi_0, \phi_1^t, \phi_2^t) = (1, 1+1/n, 1-1/n)$. Le gain moyens des joueurs en \bar{t} deviendra par linéarité $(1 - 2K(\gamma)/n, 1 + 2K(\gamma)/n)$. On note ainsi que si un joueur obtient un gain supérieur à 1 lors d'un tour, il obtiendra un gain inférieur à 1 au tour suivant, et vice-versa. Cette alternance permettra d'opposer continuellement les deux joueurs d'un même couple, qui n'auront jamais d'intérêt commun en $C_{\gamma q}^t$ à terminer la partie ou à la continuer.

Le choix de la fonction r d'itération dans γ^t permettra de forcer la forme du compteur, tandis que la valeur de $K(\gamma)$ influencera la mise à jour du compteur.

3.3 Justifications

Intuitivement, le codage de la paire de compteurs (c_1, c_2) se fera en l'état γ^t par :

$$\mathbb{E}^\sigma(\phi_0, \phi_{A_1^t}, \phi_{A_2^t} \mid h\gamma^t) = \left(1, 1 + \frac{1}{2c_1 3^{c_2}}, 1 - \frac{1}{2c_1 3^{c_2}}\right)$$

En chaque état du jeu, le gain moyen obtenu par les couples de joueurs (A_1^t, A_2^t) sera égal à 2. Lors du début de transition de parité t , l'état $T_{\gamma q}^t$ assure un gain égal à 1 aux deux joueurs t , les joueurs \bar{t} pouvant avoir des gains plus variables. De plus, cet état devra être joué de manière pure afin d'éviter à tout prix l'état terminal $(A_i^t : 1)$ perdant pour 0 (sûreté). En $C_{\gamma q}^t$, la topologie de la partie obligera les deux joueurs à jouer de manière uniforme (*matching pennies*), ce qui permettra d'activer les deux transitions avec une probabilité connue. L'état q^t devra également être joué de manière pure pour assurer la sûreté de la partie (tout désaccord des joueurs sera perdant).

Finalement, la propagation du compteur depuis γ^t vers le prochain module $\gamma^{\bar{t}}$ sera forcée par le gain strictement égal à 1 des joueurs A_1^t et A_2^t en $T_{\gamma q}^t$.

3.4 Correction

Premièrement, il est aisé de voir que l'on peut encoder toute exécution dans un équilibre de Nash sûr.

Théorème 5. *Si \mathcal{M} possède une exécution r infinie de la forme :*

$$q_0, \underbrace{c_1^{(0)}, c_2^{(0)}}_{=(0,0)} \xrightarrow{i_0} q_1, c_1^{(1)}, c_2^{(1)} \xrightarrow{i_1} q_2, c_1^{(2)}, c_2^{(2)} \xrightarrow{i_3} \dots$$

Alors il existe un profil de stratégies σ tel que $(\sigma, T_{\text{init}q}^0)$ est un équilibre de Nash sûr.

Dans la suite, on considère un profil de stratégie σ fixé, et on s'intéresse à la réciproque (l'existence d'une exécution infinie si existence d'un équilibre pur).

On introduit pour cela une propriété de « codage valide » d'un pas d'exécution de la machine :

Définition 11 (Réduction). *Soit $(h, \gamma, q, t, c_1, c_2) \in \text{States}^* \times \Gamma \times \{0, 1\} \times \mathbb{N}^2$.*

On introduit la propriété $P(h, \gamma, q, t, c_1, c_2)$ définie par la conjonction des conditions suivantes :

$$(\sigma, hT_{\gamma q}^t) \text{ est un NE sûr} \quad (2)$$

$$\mathbb{P}^\sigma (hT_{\gamma q}^t C_{\gamma q}^t \gamma^t \mid hT_{\gamma q}^t) = \frac{1}{2} \quad (3)$$

$$\mathbb{E}^\sigma (\phi_1^t, \phi_2^t \mid hT_{\gamma q}^t C_{\gamma q}^t \gamma^t) = \left(1 + \frac{1}{2c_1 3c_2}, 1 - \frac{1}{2c_1 3c_2}\right) \quad (4)$$

Intuitivement, on exprime ainsi le codage de l'état (q, c_1, c_2) après l'opération γ , dans le jeu à l'historique initial h . Il nous reste alors à montrer que cette propriété de codage local se propage correctement, ce que le lemme suivant affirme :

Lemme 7 (Progression). *Supposons la propriété $P(h, \gamma, q, t, c_1, c_2)$ vraie.*

Alors il existe $q'\gamma' \in q\Delta$, ainsi que $(c'_1, c'_2) \in \mathbb{N}^2$ tels que $P(hT_{\gamma q}^t C_{\gamma q}^t q^t, \gamma', q', \bar{t}, c'_1, c'_2)$ est vraie. De plus, la transition suivante est valide :

$$(q, c_1, c_2) \xrightarrow{\gamma'} (q', c'_1, c'_2)$$

Le lemme 7 précédent fournit ainsi le théorème d'indécidabilité :

Théorème 1. *Le problème de l'existence d'un équilibre de Nash avec contrainte affine sur les gains n'est pas décidable (en réalité, il n'est pas co-récurivement énumérable.)*

Démonstration. On réduit le problème de l'arrêt d'une machine de Turing : pour \mathcal{M} donnée, on construit le jeu concurrent associé \mathcal{G} . Posons φ la fonction affine renvoyant le gain du joueur 0 soustrait de 1. Ainsi :

$$\varphi(\mathbb{E}^\sigma (\phi \mid h)) \geq 0 \quad \Leftrightarrow \quad (\sigma, h) \text{ est un NE sûr}$$

- Si \mathcal{M} admet une exécution infinie depuis q , alors le théorème 5 montre l'existence d'un profil de stratégies σ tel que $(\sigma, T_{\text{init}q}^0)$ est un NE sûr.
- Réciproquement, si $(\sigma, T_{\text{init}q}^0)$ est un NE sûr. Alors on montre aisément (initialisation) que $P(\epsilon, \text{init}, q, 0, 0, 0)$ est vérifiée. Le lemme 7 permet alors de construire une exécution infinie de \mathcal{M} (hérédité.)

□

L'exhibition et la preuve de correction des modules semble indiquer que l'ensemble des équilibres de Nash, dans le cas d'un faible nombre de joueurs, possède des propriétés de régularité, voire de fermeture. Par exemple, il a été

impossible d'encoder dans un jeu l'ensemble des $\frac{1}{n}$ sans inclure l'équilibre limite à valeur 0. Ces propriétés de passage à la limite, si elles s'avèrent correctes, permettraient de discrétiser le problème, et d'appliquer par la suite les techniques éprouvées dans le cadre déterministe, comme le jeu des suspects présentés dans [1]. De plus, il sera intéressant de considérer des classes de stratégies intermédiaires, notamment entre les stratégies à mémoire totale et positionnelles.

Références

- [1] Romain Brenguier. *Équilibres de Nash dans les Jeux Concurrents – Application aux Jeux Temporisés*. Thèse de doctorat, Laboratoire Spécification et Vérification, ENS Cachan, France, November 2012.
- [2] Vincent Conitzer and Tuomas Sandholm. New complexity results about nash equilibria. *Games and Economic Behavior*, 63(2) :621 – 641, 2008. Second World Congress of the Game Theory Society.
- [3] John F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1) :48–49, 1950.
- [4] Michael Ummels. *Stochastic Multiplayer Games : Theory and Algorithms*. Ph.D. Thesis, Department of Computer Science, RWTH Aachen, Germany, January 2010.
- [5] Michael Ummels and Dominik Wojtczak. The complexity of Nash equilibria in limit-average games. In Joost-Pieter Katoen and Barbara König, editors, *Proceedings of the 22nd International Conference on Concurrency Theory (CONCUR'11)*, volume 6901 of *Lecture Notes in Computer Science*, pages 482–496, Aachen, Germany, September 2011. Springer.

Appendices

A Preuves des lemmes et théorèmes

A.1 Preuve du théorème 3

Démonstration. Nous allons définir pour chaque $n \in \overline{\mathbb{N}}$, un profil de stratégie σ^n et nous allons démontrer que (σ^n, s_0) est bien un NE avec gain moyen $(1, 1 + r^{(n)}, 1 - r^{(n)})$

– **Pour** $n = +\infty$

On construit un profil de stratégie σ^∞ qui joue $bb-$ en chaque hs_0 , ce qui continue la partie vers hs_0s_1 .

En hs_0s_1 , on fait jouer A_1 et A_2 de façon uniforme. Si $r = X/(X+1)$, on spécifie de plus que 0 joue b avec probabilité 1.

Dans les deux cas : la probabilité de terminer dans l'état $(1, 1, 1)$ en un tour est constante et strictement positive. Avec probabilité 1, cet état sera atteint, ainsi :

$$\mathbb{E}^{\sigma^\infty}(\phi \mid s_0) = (1, 1, 1)$$

On vérifie de plus qu'il s'agit bien d'un équilibre de Nash : en hs_0 , 0 a déjà obtenu son gain maximal et A_1 ne peut qu'obtenir un gain nul en déviant. En hs_1 , les deux alternatives pour A_1 et A_2 fournissent le même gain.

– **Si** $n = 0$

$$\sigma^1(s_0) = aa-$$

Ce profil de stratégies donne un gain moyen égal à $(1, 2, 0)$ ce qui est la valeur maximale pour les deux premiers joueurs, seuls à pouvoir dévier dans l'état s_0 .

– **Si** $n = n' + 1 > 0$ et que l'on suppose avoir construit $\sigma^{n'}$.

Par le théorème de Nash 1, appliqué au jeu à un tour commençant en s_0s_1 et donnant un gain de $(1, 1 + r^{(n)}(1), 1 - r^{(n)}(1))$ en $s_0s_1s_0$, on sait qu'il existe une stratégie $\rho \in \text{Dist}(\text{Act})^{\text{Agt}}$ qui est un équilibre de Nash local. On pose ainsi

$$\left\{ \begin{array}{l} \sigma^n(s_0) = bb- \\ \sigma^n(s_0s_1) = \rho \\ \forall h \in \text{States}^+, \sigma^n(s_0s_1h) = \sigma^{n'}(h) \end{array} \right.$$

A.2 Preuve du lemme 2

Par le théorème 2 (sous-jeu), on sait que $(\sigma^n, s_0 s_1)$ est un NE, car c'est un NE local et que $(\sigma^n, s_0 s_1 s_0)$ est aussi un NE (par hypothèse de récurrence). D'autre part, on en déduit de plus que

$$\mathbb{E}^{\sigma^n}(\phi \mid s_0) = \left(1, 1 + r^{(n')}(1), 1 - r^{(n')}(1)\right)$$

Le lemme 1 précédent impose ainsi

$$\begin{aligned} \mathbb{E}^{\sigma^n}(\phi \mid s_0) &= \left(1, 1 + r \left(r^{(n')}(1)\right), 1 - r \left(r^{(n')}(1)\right)\right) \\ &= \left(1, 1 + r^{(n)}(1), 1 - r^{(n)}(1)\right) \end{aligned}$$

□

A.2 Preuve du lemme 2

Démonstration. Supposons que cela ne soit pas le cas. C'est-à-dire $\mathbb{P}^\sigma(\overline{h(0, 2, 0)} \mid h) > 0$. Exprimons le gain moyen de 0 :

$$\mathbb{E}^\sigma(\phi_0 \mid h) = \mathbb{P}^\sigma(\overline{h(0, 2, 0)} \mid h) \mathbb{E}^\sigma(\phi_0 \mid h \wedge \overline{h(0, 2, 0)})$$

Or ϕ_0 est une fonction majorée par 1 donc $\mathbb{E}^\sigma(\phi_0 \mid h \wedge \overline{h(0, 2, 0)}) \leq 1$ d'où $\mathbb{E}^\sigma(\phi_0 \mid h) < 1$ ce qui contredit l'hypothèse de sûreté. □

A.3 Preuve du lemme 3

Démonstration. Dans le cas d'un jeu $\mathcal{G}_{X/k}$, il suffit de faire jouer A_1 de façon uniforme en hs_1 ($\sigma'_1(hs_1) \sim \mathcal{U}(\{1 \dots k\})$) ce qui lui assure un gain au moins égal à $1 - \frac{1}{k}$.

Pour $\mathcal{G}_{X/(X+1)}$, il suffit de faire purement jouer A_1 en a : $\sigma'_1(hs_1) = a$. Ceci lui fournit un gain moyen $2\sigma_0(a \mid hs_1) + 1\sigma_0(b \mid hs_1) \geq \sigma_0(a \mid hs_1) + \sigma_0(b \mid hs_1) = 1$. □

A.4 Preuve du lemme 4

Démonstration. Nous distinguons le cas où la transition $bb-$ est jouée du cas où elle ne l'est pas.

- Cas 1 : $\sigma(bb- \mid hs_0) = 0 = \sigma_0(b \mid hs_0)\sigma_1(b \mid hs_0)$
D'autre part (lemme 2) $\sigma(ab- \mid hs_0) = 0 = (1 - \sigma_0(b \mid hs_0))\sigma_1(b \mid hs_0)$
D'où en sommant $\boxed{\sigma_1(b \mid hs_0) = 0}$ (la stratégie de A_1 est pure).

A.5 Preuve du lemme 5

Afin de prouver que la stratégie du joueur 0 est pure (en a), on construit une déviation σ'_1 de A_1 :

$$\sigma'_1 = \sigma_1[A_1/b]^h[A_1/\widetilde{\sigma}'_1]^{hs_1}$$

où $\widetilde{\sigma}'_1$ est la stratégie construite au lemme précédant (assurant un gain non-nul à A_1), c'est-à-dire que A_1 joue désormais b (avec probabilité 1) puis joue en hs_0s_1 une stratégie lui assurant un gain non-nul par le lemme 3.

En notant $\sigma' = \sigma[1/\sigma'_1]$, on a

$$\begin{aligned} \mathbb{E}^{\sigma'}(\phi_1 | hs_0) &= 2\sigma_0(a | hs_0) + \sigma_0(b | hs_0) \underbrace{\mathbb{E}^{\sigma'}(\phi_1 | hs_0s_1)}_{>0} \\ &\geq 2\sigma_0(a | hs_0) \\ &\geq \mathbb{E}^{\sigma}(\phi_1 | hs_0) \end{aligned}$$

Avec égalité si et seulement-si $\boxed{\sigma_0(b | hs_0) = 0}$.

Cette égalité est donc nécessaire à l'équilibre, ce qui montre que le joueur 0 joue a avec probabilité 1.

– Cas 2 : hs_0s_1 est accessible, c'est-à-dire $\sigma(bb | hs_0) > 0$

On a toujours $\sigma(ab | hs_0) = \sigma_0(a | hs_0)\sigma_1(b | hs_0) = 0$

Ceci impose désormais que $\boxed{\sigma_0(a | hs_0) = 0}$ c'est-à-dire que le joueur 0 joue b (sinon on aurait $\sigma_1(b | hs_0) = 0$ puis $\sigma(bb | hs_0) = 0$).

Le gain de A_1 devient ainsi

$$\mathbb{E}^{\sigma}(\phi_1 | hs_0) = \sigma_1(b | hs_0)\mathbb{E}^{\sigma}(\phi_1 | hs_0s_1)$$

Or (σ, hs_0s_1) est un équilibre, donc A_1 obtient un gain strictement positif (autrement il pourrait dévier par le lemme précédent), donc $\mathbb{E}^{\sigma}(\phi_1 | hs_0s_1) > 0$.

La quantité $\mathbb{E}^{\sigma}(\phi_1 | hs_0)$ est donc affine en $\sigma_1(b | hs_0)$ de coefficient strictement positif. La condition d'équilibre en hs_0 impose donc

$$\boxed{\sigma_1(b | hs_0) = 1} \text{ (} A_1 \text{ joue } b \text{.)}$$

□

A.5 Preuve du lemme 5

Démonstration. On montre le résultat par récurrence sur $|\text{Supp}(\sigma, h)|$.

- Si $|\text{Supp}(\sigma, h)| = 1$
Si $\text{last}(h) = s_0$, alors la seule transition possible est $aa-$ car s_0 est « déterministe » (lemme 4) et ne possède pas de successeur non-terminal. On arrive alors dans l'état $(1, 2, 0) \in I_r$ (aucune itération).
Sinon $\text{last}(h) = s_1$, alors l'état terminal atteint est nécessairement $(1, 1, 1) \in I_r$ (une infinité itérations), l'état $(0, 2, 0)$ étant interdit par sûreté (lemme 2).
- Si $|\text{Supp}(\sigma, h)| = n + 1$
 Supposons le résultat vrai pour tout équilibre à support de taille n .
Si $\text{last}(h) = s_0$, alors la seule transition possible est $bb-$ avec probabilité 1, car s_0 est « déterministe » et le support contient plus d'un élément. Ainsi $\mathbb{E}^\sigma(\phi \mid h) = \mathbb{E}^\sigma(\phi \mid hs_1)$ et on applique l'hypothèse de récurrence à (σ, hs_1) qui est un NE sûr à support de taille n .
Sinon, $\text{last}(h) = s_1$. Le support contient plus d'un élément, donc $\mathbb{P}^\sigma(hs_0 \mid h) > 0$.
 D'une part, on peut appliquer l'hypothèse de récurrence au NE sûr (σ, hs_0) à support de taille n , c'est-à-dire qu'il existe $i \in \bar{\mathbb{N}}$ tel que

$$\mathbb{E}^\sigma(\phi \mid hs_0) = \left(1, 1 + r^{(i)}(1), 1 - r^{(i)}(1)\right)$$

D'autre part, on peut appliquer le lemme 1 en h ($\text{last}(h) = s_1$) afin d'itérer et obtenir :

$$\mathbb{E}^\sigma(\phi \mid h) = \left(1, 1 + r^{(i+1)}(1), 1 - r^{(i+1)}(1)\right)$$

□

A.6 Preuve du lemme 6

Démonstration. Comme $\mathbb{E}^\sigma(\phi_0 \mid h) = 1$, la partie se termine presque sûrement. Cet état ne peut être $(0, 2, 0)$, par sûreté. Il ne peut pas s'agir de $(1, 0, 2)$ ou $(1, 2, 0)$, car sinon, le support serait fini par le lemme 4 (déterminisme de s_0).

Le seul état terminal possible est donc $(1, 1, 1)$ qui est atteint presque sûrement. □

A.7 Preuve du lemme 7

Démonstration. Comme $(\sigma, hT_{\gamma q}^t)$ est sûr, on sait que la partie est jouée de façon pure sans tomber dans l'état puits ($A_i^t : 1$). C'est-à-dire $\mathbb{P}^\sigma(hT_{\gamma q}^t C_{\gamma q}^t \mid h) = 1$. Ainsi, $(\sigma, hT_{\gamma q}^t C_{\gamma q}^t)$ est également un équilibre pur, et comme les

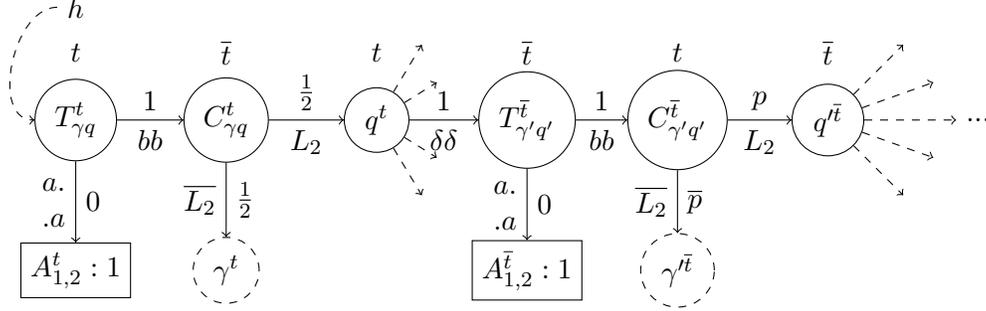


FIGURE 6 – Probabilités de transition connues, sur deux cycles t puis \bar{t} consécutifs.

deux transitions sont activées, $(\sigma, hT_{\gamma q}^t C_{\gamma q}^t \gamma^t)$ et $(\sigma, hT_{\gamma q}^t C_{\gamma q}^t q^t)$ sont des NE purs.

On pose $h' = hT_{\gamma q}^t C_{\gamma q}^t q^t$. Le branchement non-déterministe en h' doit donc être joué de manière pure par les deux joueurs (afin d'éviter l'état perdant pour le joueur 0). Il existe donc $\delta = \gamma' q' \in q\Delta$ transition prise avec probabilité 1 (accord des deux joueurs). La propriété de sûreté a ainsi été propagée à $(\sigma, h'T_{\gamma q}^{\bar{t}})$ ce qui montre (2).

Comme précédemment, la transition bb depuis $h'T_{\gamma' q'}^{\bar{t}}$ est prise avec probabilité 1 (accord des joueurs), ce qui propage la propriété de sûreté à $h'T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}}$.

Posons :

$$p = \mathbb{P}^\sigma \left(h'T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \mid h' \right)$$

On cherche désormais à montrer $p = \frac{1}{2}$ (propriété (3)). La situation connue à ce point de la preuve est présentée en figure 6.

Les conditions d'équilibre en $hT_{\gamma q}^t$ imposées par A_1^t et A_2^t (qui ne doivent pas dévier) s'expriment alors par

$$\begin{cases} 1 \leq \mathbb{E}^\sigma (\phi_1^t \mid hT_{\gamma q}^t C_{\gamma q}^t) = \frac{1}{2} \left(1 + \frac{1}{2c_1 3c_2} \right) + \frac{1}{2} \mathbb{E}^\sigma (\phi_1^t \mid h'T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}}) \\ 1 \leq \mathbb{E}^\sigma (\phi_2^t \mid hT_{\gamma q}^t C_{\gamma q}^t) = \frac{1}{2} \left(1 - \frac{1}{2c_1 3c_2} \right) + \frac{1}{2} \mathbb{E}^\sigma (\phi_2^t \mid h'T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}}) \end{cases}$$

Supposons premièrement par l'absurde que $p = 1$. Comme les A_i^t peuvent tout deux s'assurer un gain supérieur à 1 dans $q^{\bar{t}}$ et que la somme de leur

gain vaut 2, l'équilibre vérifie :

$$\mathbb{E}^\sigma \left(\phi_1^t, \phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \right) = \mathbb{E}^\sigma \left(\phi_1^t, \phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} q^{\bar{t}} \right) = (1, 1)$$

L'équation précédente devient ainsi :

$$\left\{ 1 \leq \frac{1}{2} \left(1 + \frac{1}{2^{c_1 3^{c_2}}} \right) + \frac{1}{2} \wedge 1 \leq \frac{1}{2} \left(1 - \frac{1}{2^{c_1 3^{c_2}}} \right) + \frac{1}{2} \right\} \Leftrightarrow \frac{1}{2^{c_1 3^{c_2}}} = 0$$

Ce qui est impossible pour des valeurs finies des compteurs : $p < 1$.

Ainsi, le module $\gamma^{\bar{t}}$ est activé (c'est un NE sûr accessible) : il existe donc $m \in \bar{\mathbb{N}}$ tel que $\mathbb{E}^\sigma \left(\phi_1^{\bar{t}}, \phi_2^{\bar{t}} \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \gamma^{\bar{t}} \right) = \left(1 + \frac{1}{m}, 1 - \frac{1}{m} \right)$ d'où par linéarité : $\mathbb{E}^\sigma \left(\phi_0, \phi_1^t, \phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \gamma^{\bar{t}} \right) = \left(1 - \frac{2K(\gamma')}{m}, 1 + \frac{2K(\gamma')}{m} \right)$. On vérifie ainsi que la somme des gains moyens des joueurs t vaut 2.

Supposons maintenant par l'absurde que $p = 0$. Comme A_t^1 et A_t^2 peuvent chacun s'assurer un gain supérieur à 1 en jouant vers $q^{\bar{t}}$, et que la somme totale de leur gain (en $\gamma^{\bar{t}}$) vaut 2, on en déduit cette fois $\mathbb{E}^\sigma \left(\phi_1^t, \phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \right) = \mathbb{E}^\sigma \left(\phi_1^t, \phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \gamma^{\bar{t}} \right) = (1, 1)$ qui est encore une fois impossible : Ainsi, $0 < p < 1$.

La transition vers $q^{\bar{t}}$ est donc activée : $\mathbb{E}^\sigma \left(\phi_1^t, \phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} q^{\bar{t}} \right) = (1, 1)$.

On sait aussi que $\frac{2K(\gamma')}{m} \neq 0$ ($m < \infty$), sinon on aurait encore une fois $\mathbb{E}^\sigma \left(\phi_1^t, \phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \right) = (1, 1)$.

Finalement, on a

$$\begin{cases} \mathbb{E}^\sigma \left(\phi_1^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \gamma^{\bar{t}} \right) < \mathbb{E}^\sigma \left(\phi_1^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} q^{\bar{t}} \right) \\ \mathbb{E}^\sigma \left(\phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} \gamma^{\bar{t}} \right) > \mathbb{E}^\sigma \left(\phi_2^t \mid h' T_{\gamma' q'}^{\bar{t}} C_{\gamma' q'}^{\bar{t}} q^{\bar{t}} \right) \end{cases}$$

Ce qui impose $p = \frac{1}{2}$.

Il reste maintenant à montrer la propriété de codage des compteurs (4) et la validité de la transition. La condition d'équilibre en $hT_{\gamma q}^t$ se réduit à

$$\begin{cases} 1 \leq \frac{1}{2} \left(1 + \frac{1}{2^{c_1 3^{c_2}}} \right) + \frac{1}{2} \left[\frac{1}{2} \left(1 - \frac{2K(\gamma')}{m} \right) + \frac{1}{2} \right] \\ 1 \leq \frac{1}{2} \left(1 - \frac{1}{2^{c_1 3^{c_2}}} \right) + \frac{1}{2} \left[\frac{1}{2} \left(1 + \frac{2K(\gamma')}{m} \right) + \frac{1}{2} \right] \end{cases} \Leftrightarrow m = K(\gamma') 2^{c_1} 3^{c_2}$$

A.7 Preuve du lemme 7

$K(\gamma') \in \{\frac{1}{3}, \frac{1}{2}, 1, 2, 3\}$, il existe donc $(c'_1, c'_2) \in \mathbb{Z}^2$ tels que $m = 2^{c'_1} 3^{c'_2}$ $(c'_1, c'_2) \in \mathbb{N}^2$ ce qui montre (4).

La validité de la transition γ' est vérifiée trivialement pour les mises à jour de compteurs en observant la valeur de $K(\gamma')$. De plus, si $\gamma' = \text{zero}(j)$ ($j \in \{1, 2\}$), m est de la forme $\frac{1}{(3-j)^n}$ c'est-à-dire $c'_j = 0 = c_j$ (car $K(\text{zero}(j)) = 1$). \square

TABLE 3

		0	
	1	a	b
a		$1, (k+1)$	$0, 0$
b		$0, k$	$0, \frac{1}{c}$

TABLE 4

		0	
	1	a	b
a		$1, 1$	$0, 0$
b		$0, \frac{k+1}{c+1}$	$0, \frac{l+1}{c+1}$

B Vers une réduction à trois joueurs

Considérons à titre d'exemple le tableau 3, où $k \in \mathbb{N}$ et $c \in \mathbb{R}_{>0}$

Notons premièrement que les stratégies aa et bb sont les deux seuls équilibres de Nash purs. Une condition suffisante sur l'existence d'un équilibre mixte est

$$\begin{cases} p_0 = \frac{\frac{1}{c}}{\frac{1}{c} + 1} = \frac{1}{c+1} \\ p_1 = 0 \end{cases}$$

On obtient alors le gain moyen suivant pour les deux joueurs :

$$\mathbb{E}^\sigma(\Phi \mid s_0) = \left(0, \frac{k+1}{c+1}\right)$$

Considérons maintenant pour $k, l \in \mathbb{N}$, la table ??.

Une condition suffisante d'équilibre :

$$\begin{cases} p_0 = \frac{\frac{l+1}{c+1}}{\frac{l+1}{c+1} + 1 - \frac{k+1}{c+1}} = \frac{l+1}{c+1 + (l-k)} \\ p_1 = 0 \end{cases}$$

On a alors

$$\begin{aligned}
\mathbb{E}^\sigma(\Phi_1 \mid s_0) &= \frac{l+1}{c+1+(l-k)} \frac{k+1}{c+1} + \left(1 - \frac{l+1}{c+1+(l-k)}\right) \frac{l+1}{c+1} \\
&= \frac{l+1}{c+1+(l-k)} \frac{k+1}{c+1} + \frac{c+1+l-k-(l+1)}{c+1+(l-k)} \frac{l+1}{c+1} \\
&= \frac{l+1}{(c+1+(l-k))(c+1)} (k+1+c+1+l-k-(l+1)) \\
&= \frac{l+1}{(c+1+(l-k))(c+1)} (1+c) \\
&= \frac{l+1}{c+1+(l-k)}
\end{aligned}$$

Pour $l = k - 1$, on obtient ainsi $\mathbb{E}^\sigma(\Phi_1 \mid s_0) = \frac{k}{c}$. Cette expression est valable tant que $p_0 \in [0, 1]$, c'est-à-dire si et seulement si $k \leq c$. Autrement, les seuls équilibres possibles sont les équilibres purs.

Notons que les conditions vues précédemment étaient uniquement suffisantes, et non nécessaires, car $p_1 = 0$. Il serait tout de même envisageable de forcer uniquement cet équilibre en utilisant un troisième joueur qui jouerait un rôle antagoniste (comme vu dans les preuves des modules \mathcal{G}_r , $\phi_2 = 1 - \phi_1$).

Il resterait de plus à appliquer le théorème 2.

On aurait ainsi fabriqué, depuis un état final (ou suite de partie) où l'un des joueurs obtient un gain $\frac{1}{c}$, un nouveau jeu où le joueur obtient un gain $\frac{k}{c}$. Ceci permettrait ainsi d'éviter de dupliquer les joueurs $A_{1,2}$ et ainsi obtenir une réduction à trois joueurs (le joueur 0 et les deux joueurs antagonistes).