

Chapitre 6

Des jeux répétés et dynamiques

Le chapitre 5, considère des jeux statiques dans le sens où ils sont joués une seule fois. La théorie des jeux se focalise essentiellement en effet sur des outils pour caractériser les équilibres en présence de joueurs rationnels lorsque le jeu est joué une unique fois. Elle vise à prédire ce qu'il faut jouer si chaque joueur est rationnel. Dans ces circonstances, il est en fait difficile d'interpréter la notion de stratégie mixte : pourquoi introduire des probabilités lors que le jeu est joué une seule fois ? Plusieurs discussions ont été faites sur ce sujet [Binmore, 1999], [Osbourne and Rubinstein, 1994].

Pour introduire des aspects dynamiques en théorie des jeux, deux grandes approches ont été développées.

La première consiste à considérer des jeux répétés. Nous présentons cette approche dans la première section.

La seconde consiste à étudier certains types de comportements des joueurs. En effet, il est naturel de supposer que la réponse de chacun des joueurs est donnée par certaines règles de comportement : le choix du joueur i est une certaine fonction $f_i(Q)$ du passé Q du jeu : f_i correspond à la fonction modélisant le comportement de i .

Si les f_i correspondent à prendre la meilleure réponse à la statistique du passé du jeu, on obtient la dynamique du joueur fictif, que nous présentons dans la section 6.2.

Si les f_i correspondent à prendre la meilleure réponse au choix de l'adversaire au temps précédent, on obtient la dynamique dite myope. La dynamique pavlovienne est une de ses variantes. Elle consiste à conserver sa stratégie tant que le coût par cette stratégie est inférieur à un certain seuil, et sinon à changer de stratégie pour une meilleure réponse. Nous présenterons des résultats sur la puissance de ce type de comportements dans la section 6.3.

Nos contributions personnelles relatives à ce chapitre concerne d'une part le calcul de stratégies comportementales pour obtenir des stratégies Pareto optimale dans le contexte du routage inter-domaine dans les jeux répétés [18]. Nous avons étudié le comportement des joueurs fictifs dans ce contexte [Boussaton, 2009]. D'autre part, une caractérisation partielle de la puissance des protocoles de population utilisant des dynamiques de jeu myopes [24], [Rabie, 2009].

Les travaux présentés dans ce chapitre ont été obtenus via le travail de l'étudiant en thèse Octave Boussaton, étudiant à l'université de Nancy I, et le co-encadrement de stagiaires de l'Ecole Normale supérieure de Lyon.

6.1 Jeux répétés

Considérons le problème du dilemme du prisonnier avec la matrice de coût suivante :

		joueur 2	
		D	C
joueur 1	D	(3,3)	(0,4)
	C	(4,0)	(1,1)

Le seul équilibre de Nash pur dans ce jeu statique est (D,D). Par contre, si le jeu est répété plusieurs fois, on aimerait pouvoir dire que l'intérêt de chacun des joueurs serait de choisir C à chaque étape puisque leur coût serait plus petit que si tous les deux choisissent D. Cela impliquerait une certaine coopération entre les deux joueurs. Plusieurs questions se posent alors. Par exemple est-ce un équilibre de Nash dans un jeu répété? Existe-t-il d'autres équilibres de Nash?

Mais avant de répondre partiellement à ces questions, nous allons retranscrire ces définitions dans le cadre des jeux répétés.

Tout d'abord, nous allons rappeler les notations des jeux statiques définies dans le chapitre 5 :

- $[n] = \{1, \dots, n\}$ est l'ensemble des joueurs
- \mathcal{S}_i (resp. K_i) est l'ensemble des stratégies pures (resp. mixtes) du joueur i
- $\mathcal{K} = \prod_{i=1}^n \mathcal{K}_i$ est l'espace des stratégies mixtes.
- $\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$ est l'espace des stratégies pures.
- c_i est la fonction de coût du joueur i .

Répéter (itérer) k fois un jeu, revient à étendre l'espace des stratégies en \mathcal{S}_i^k pour chacun des joueurs i . En effet, le joueur i choisit son action $\mathbf{q}_i(t) \in \mathcal{S}_i$ au temps t pour $t = 1, 2, \dots, k$. Le jeu répété k fois est donc un jeu sur le nouvel espace de stratégies $\prod_{i=1}^n \mathcal{S}_i^k$.

A l'étape t , le profil de ce jeu est noté par $\mathbf{q}(t)$. Conformément à [Binmore, 1999], pour éviter les confusions, nous nommerons *actions* les choix $\mathbf{q}_i(t)$ de chacun des joueurs à un instant donné, et *stratégies de comportement* les suites $Q_i = \mathbf{q}_i(1), \dots, \mathbf{q}_i(k)$, c'est-à-dire les stratégies pour ce jeu global (répété).

Maintenant, nous allons décrire comment représenter les répétitions du jeu, en particulier, en parlant d'histoire du jeu.

Définition 27 L'ensemble H_t des histoires de longueur t est $H_t = \mathcal{S}^t$. L'ensemble de toutes les histoires est $H = \cup_{t>0} H_t$.

Maintenant, nous allons introduire la notion de comportement d'un joueur qui correspond à ce qu'il va jouer à un instant donné en fonction de l'histoire du jeu (et surtout par rapport à sa connaissance de son passé).

Définition 28 Une stratégie de comportement du joueur i est une application f_i de H dans \mathcal{K}_i .

Répéter un jeu un nombre fini et fixé de fois (et connu par les joueurs) n'apporte rien de plus par rapport aux comportements des joueurs. En effet, considérons l'exemple du dilemme des prisonniers. A la dernière étape du jeu répété, la stratégie D domine la stratégie C, les deux joueurs jouent donc la stratégie D. Ensuite pour l'avant-dernière

étape, ils savent que la dernière étape sera (D,D). Donc, ils joueront la stratégie D. Un simple raisonnement de récurrence en arrière montre que chacun des joueurs a toujours intérêt à jouer D aux temps $t, t-1, \dots, 1$: voir [Binmore, 1999, page 349].

Il est plus intéressant de répéter le jeu un nombre infini de fois. Cette fois, la stratégie de coopération systématique, c'est-à-dire pour le dilemme des prisonniers jouer toujours C , devient un équilibre de Nash pour le jeu répété : en d'autres termes, coopérer devient intéressant pour chacun.

De façon similaire, si l'on suppose que l'on répète le jeu un nombre fini, mais aléatoire, de fois le jeu, avec à chaque étape une probabilité δ que le jeu continue encore une étape de plus, et $1 - \delta$ que ce soit la dernière étape, alors la probabilité que l'on joue la i ème étape vaut δ^i . En moyenne, le gain de chacun des joueurs est donc donné par l'équation 6.1 : voir [Binmore, 1999].

$$cout(H_k) = \sum_{t=1}^k \delta^{t-1} cout_i(\mathbf{q}(t)) \quad (6.1)$$

où

- $H_k = (\mathbf{q}(1), \dots, \mathbf{q}(k))$ est l'histoire du jeu.
- $cout_i$ est la fonction de coût du joueur i .

Cette façon de compter peut aussi s'interpréter comme un coût actualisé. Le paramètre δ correspond à un *taux d'actualisation* : pour un joueur, le coût δx aujourd'hui est équivalent au coût x demain. En général, le coût d'un joueur est comptabilisé avec un le taux d'actualisation $\delta > 0$ (que l'on peut prendre égal à 1, si k est infini ; on choisit $\delta < 1$ si k est fini).

Définissons deux types de jeux infinis :

Définition 29 (jeu infiniment répété et jeu escompté) *Le jeu escompté au taux d'actualisation $\delta \in]0, 1]$ est le jeu répété une infinité de fois et dont le coût final du joueur i est $\sum_{t \geq 1} (1 - \delta) \delta^{t-1} cout_i(\mathbf{q}(t))$.*

Le jeu infiniment répété est le jeu répété une infinité de fois et dont le coût final du joueur i est $\lim_{k \rightarrow \infty} \sum_{t=1}^k cout_i(\mathbf{q}(t))$ si cette limite existe.

Le problème est que la stratégie de coopération systématique n'est pas en fait le seul équilibre de Nash. Jouer la stratégie D systématiquement reste aussi un équilibre de Nash du jeu répété.

Avant d'étudier les propriétés des équilibres de Nash, considérons *l'espace des coûts réalisables* : c'est un polytope convexe représentant l'ensemble des coûts réalisables dans le jeu répété. En tant que convexe, il s'écrit $conv(\mathcal{K})$ et rappelons que \mathcal{K} correspond à l'espace des stratégies mixtes de tous les joueurs : voir 6.1 pour le jeu du dilemme des prisonniers.

Puis nous allons définir l'espace des coûts pour des stratégies coopératives pour des joueurs rationnels. En fait, coopération est lié à punition. Cela vient du fait qu'il faut inciter les joueurs rationnels à rester en coopération. Mais comment punir ? Tout simplement en jouant la situation la plus défavorable pour le joueur i (voir équation 6.2).

Définition 30 *Pour chaque joueur i , le niveau de punition du joueur i est*

$$v_i = \max_{Q_{-i} \in \prod_{j \neq i} \mathcal{K}_j} \min_{\mathbf{q}_i \in \mathcal{K}_i} cout_i(\mathbf{q}_i, Q_{-i}) \quad (6.2)$$

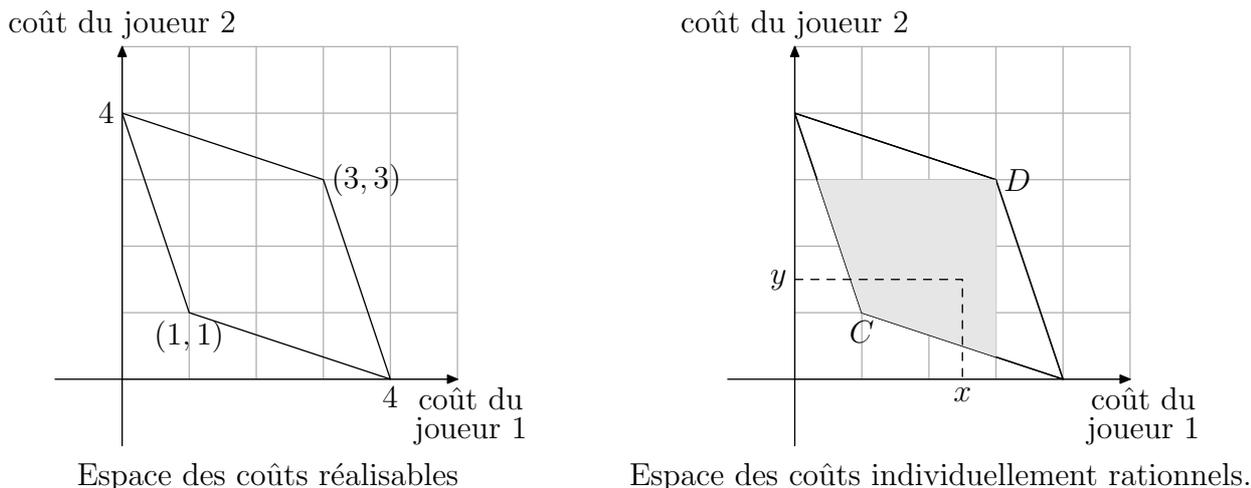


FIG. 6.1 – espace des coûts pour le problème du dilemme des prisonniers

Définissons un espace des coûts où les joueurs rationnels peuvent être sanctionnés si ils ne coopèrent pas.

Définition 31 *L'ensemble E des coûts individuellement rationnels est*

$$E = \text{conv}(\mathcal{K}) \cap \{(c_{out}_i)_{i \in [n]} : c_{out}_i \leq v_i \forall i \in [n]\}$$

En fait (voir par ex. [Binmore, 1999]) :

Théorème 24 (Le théorème de tout le monde (Folk Theorem)) *L'ensemble des coûts d'équilibre pour le jeu infiniment répété et escompté est l'ensemble des coûts individuellement rationnels.*

Intuitivement, les stratégies dans le jeu répété infini correspondent en fait aux stratégies mixtes dans le jeu statique. Considérons une stratégie mixte appartenant à E . On peut construire une stratégie de comportement correspondant à une simulation de la stratégie mixte dans le jeu statique et si un joueur i dévie, il est sanctionné.

Certains équilibres construits sont fortement critiquables en tant que solution rationnelle car ils reposent sur des punitions non-crédibles : étant donné qu'un joueur a dévié, rien assure que ses adversaires auront intérêt à punir pour toujours le cas échéant. Pour remédier à cette remarque la notion d'équilibre en sous-jeu parfait (ESJP) a été introduite par *R. J. Aumann, L. S. Shapley et A. Rubinstein* en 1976. De façon informelle, un profil de stratégies σ est un équilibre d'un sous-jeu parfait si pour toute histoire h de H , à partir de tout instant t , la suite des σ à partir t reste un équilibre. Cela permet d'oublier le passé et surtout de modéliser une sanction d'une déviation de façon limitée dans le temps. Dans ce contexte, le théorème *de tout le monde* reste valide. En effet si un joueur dévie en date t , il est puni jusqu'à la date t^2 , après quoi les joueurs oublient la déviation et se remettent à jouer la suite prévue.

Dans [18], nous nous sommes penchées sur le problème de construire un équilibre de Nash dans un jeu infiniment répété pour le routage inter-domaine. Dans un réseau, un seul sommet veut faire transiter ses informations vers un sommet destination en empruntant le chemin de plus petit coût. Chaque sommet possède un coût de transit et une fonction

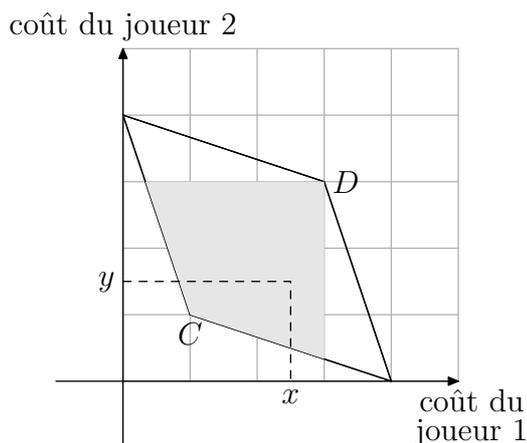


FIG. 6.2 – Pour tout point (x, y) de la zone grisée, il y a une paire q_1, q_2 de stratégies qui garantit le gain x pour le joueur 1, y pour le joueur 2, et tel que (q_1, q_2) soit un équilibre de Nash pour le dilemme des prisonniers itéré. En particulier pour le point C qui correspond à la coopération systématique, et pour le point D qui correspond à la défection systématique.

de prix pour transiter les données. Cette dernière peut fluctuer pour attirer le trafic. Et l'objectif est de maximiser la différence moyenne entre le prix de transit et le coût de faire transiter les messages. Dans ce contexte, pour certaines instances, nous avons trouvé un algorithme polynomial qui permet de construire une stratégie coopérative correspondant à un équilibre de Nash Pareto optimal.

Observons qu'en général, le nombre d'équilibres de Nash du jeu répété n'est pas dénombrable puisque toute stratégie mixte dans le jeu statique dans l'espace E a son pendant dans le jeu répété. Dans ce contexte, un jeu répété possède une multitude d'équilibres de Nash.

Le problème du dilemme des prisonniers répété a été beaucoup étudié. En particulier, [Axelrod, 1984], a décrit le résultat de tournois de règles de comportements pour le dilemme des prisonniers itéré. Axelrod argumente dans [Axelrod, 1984] qu'il existe un meilleur comportement, nommée *TIT – FOR – TAT*. Le comportement *TIT – FOR – TAT* consiste à coopérer à la première étape, puis ensuite à faire la même chose que son adversaire aux temps ultérieurs.

6.2 Vers la dynamique des joueurs fictifs

Maintenant, nous allons présenter des approches étudiant des types de comportement. Tout d'abord, nous allons présenter l'approche des dynamiques des joueurs fictifs en étudiant si elle permet d'apprendre un équilibre de Nash. Ensuite, nous retranscrivons la notion des jeux répétés dans les systèmes informatiques répartis plus particulièrement dans les protocoles de population définis dans [Angluin et al., 2004] inspirés par les réseaux de capteurs.

Dans cette section, nous allons introduire l'apprentissage d'équilibre de Nash que nous développerons plus en détail dans le chapitre 7. Le jeu est répété à chaque étape. En pratique, le joueur i est confronté au problème suivant à chaque instant t : quelle action jouée au temps t étant donnée l'histoire du jeu ?

Il est naturel de supposer que la réponse de chacun est donnée par certaines règles de comportement : pour chaque joueur i , $q_i(t) = f_i(Q)$, où f_i est une fonction qui détermine le comportement de i en fonction du passé Q .

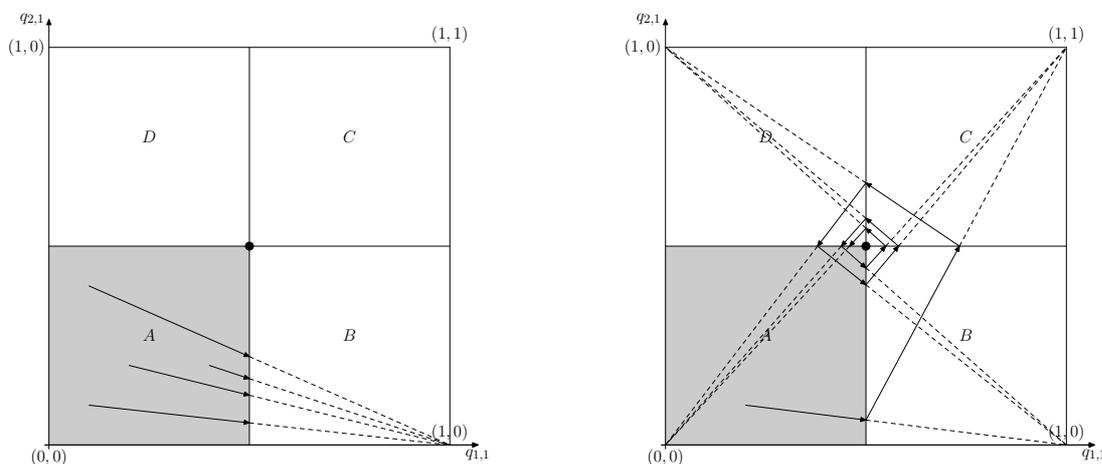
Nous allons étudier la dynamique des joueurs fictifs qui permet dans certains cas de garantir que le système converge vers un équilibre de Nash.

La notion de joueur fictif a été introduite dans [Brown, 1951]. Le joueur est un joueur *fictif* si le joueur a le comportement suivant : en fonction du passé, et de la statistique des stratégies de ses adversaires, le joueur jouera une meilleure réponse. A la première répétition du jeu, il se comportera comme un joueur myope, ensuite il constituera ses propres croyances sur la stratégie de ses adversaires. Cette croyance est basée sur les statistiques des stratégies jouées par les adversaires. Plus formellement, le processus de prédiction des joueurs dépend des actions passées : pour prédire la stratégie du joueur au temps t , il estimera que chaque joueur adverse jouera l'action i avec probabilité $n_i/(t-1)$ au temps t si ce dernier a joué l'action i n_i fois au temps $1, 2, \dots, t-1$.

Pour simplifier l'étude, en suivant le raisonnement de [Binmore, 1999] nous allons prendre l'exemple d'un jeu à deux joueurs symétrique ayant la forme stratégique suivante :

		joueur 2	
		1	2
joueur 1	1	(3,1)	(0,3)
	2	(1,2)	(2,0)

Si le joueur 2 a utilisé n_i fois l'action numéro i entre le temps 1 et $t-1$, alors le joueur 1 estimera que le joueur 2 jouera au temps t l'action i avec probabilité $q_{2,i}(t) = n_i/(t-1)$. De façon analogue, le joueur 2 évaluera la probabilité $q_{1,i}(t)$ que le joueur 1 joue l'action i .



direction de la dynamique dans la zone A exemple de comportement de la dynamique

FIG. 6.3 – Convergence vers un équilibre mixte dans un jeu à 2 joueurs

Pour étudier la dynamique du système, il suffit de passer d'un temps discret à un temps continu (voir [Binmore, 1999] pour plus de détails). Le système peut être décrit par le couple $(q_{1,1}, q_{2,1})$ avec $q_{i,1}$ correspondant à la probabilité que le joueur i joue la stratégie 1.

Une analyse simple sur les meilleures réponses pour chaque joueur (voir [Binmore, 1999]) montre que aussi longtemps que $(x_2(t), y_2(t))$ reste dans la zone A de la partie gauche de

la figure 6.3, le joueur 1 voudra utiliser la stratégie pure 2, et le joueur 2 la stratégie pure 1.

Ici, la dynamique $(q_{1,1}, q_{2,1})$ restera donc dans cette zone jusqu'à l'instant $t + \tau$ pour $\tau > 0$ suffisamment petit. Puisqu'on connaît le choix du joueur 2 entre le temps t et $t + \tau$, on peut donc évaluer $y_2(t + \tau)$ comme

$$q_{2,1}(t + \tau) = \frac{tq_{2,1}(t)}{t + \tau}. \quad (6.3)$$

Par réécriture, on a $\frac{q_{2,1}(t+\tau) - q_{2,1}(t)}{\tau} = -q_{2,1}(t)$. En faisant tendre τ vers 0, on obtient

$$q'_{2,1}(t) = \frac{q_{2,1}(t)}{t}.$$

De façon similaire, on obtient $q'_{1,1}(t) = \frac{1 - q_{1,1}(t)}{t}$.

Les points qui satisfont ces deux équations sont sur le segment entre $(q_{1,1}(t), q_{2,1}(t))$ et le point $(1, 0)$. Sur les zones B , C , et D , une étude similaire peut être réalisée. La figure 6.3 montre comment la dynamique se comporte dans les quatre différentes zones. Une simple analyse sur les distances entre la dynamique et le point d'équilibre à chaque entrée de la dynamique dans une zone permet de constater que elle converge donc vers l'équilibre de Nash mixte du jeu.

Dans [Boussaton, 2009], nous avons analysé le comportement de la dynamique des joueurs fictifs de cette manière sur le problème de routage inter-domaine dans des jeux à deux joueurs suite au modèle proposé dans les travaux de [18]. Malheureusement, nous n'avons pas pu étendre le résultat de convergence au problème à trois joueurs.

En effet, la convergence n'a pas lieu sur tous les jeux : on peut facilement considérer des jeux où les trajectoires ne convergent pas, ou avec des cycles limites [Binmore, 1999], [Monderer and Sela, 1996], [Shapley, 1964]. Par exemple, pour le jeu généralisé "papier"- "ciseau"- "caillou" la dynamique du joueur fictif ne converge pas : voir par exemple [Binmore, 1999], [Shapley, 1964].

Dans [Robinson, 1951], il a été prouvé que la dynamique du joueur fictif converge vers un équilibre de Nash dans les jeux à somme nulle. Ce résultat a été aussi prouvé pour les jeux 2×2 [Miyasawa, 1961] non dégénérés c'est-à-dire tels qu'aucun joueur a des utilités identiques dans sa matrice. La condition de non dégénérescence est nécessaire [Monderer and Sela, 1996]. Ce résultat a été étendu pour les jeux non dégénérés à deux joueurs ayant 3 stratégies [Shapley, 1964] tout en donnant un exemple où la dynamique du joueur fictif ne converge pas sans cette hypothèse. Ce résultat a été étendu à des jeux à deux joueurs ayant n stratégies [Berger, 2003].

Beaucoup de travaux ont été consacrés à savoir quand le comportement de la dynamique du joueur fictif converge ou pas voir [Fudenberg and Levine, 1996] pour un survol. Il est à noter que :

Théorème 25 *La dynamique des joueurs fictifs converge vers des équilibres de Nash pour les jeux de potentiel exacts [Monderer and S., 1996] et pour les jeux de potentiel ordinal [Berger, 2007].*

Ces dynamiques du joueur fictif nécessitent une vision globale du jeu : chaque joueur doit connaître sa meilleure réponse, sa matrice de coût en fonction des stratégies de ses adversaires. L'hypothèse de connaître la matrice complète de coût est forte puisque cela implique que le joueur a une grande connaissance du système.

6.3 Vers les protocoles de population

La puissance de calcul de réseaux d'agents mobiles, anonymes, ayant des ressources limitées a été étudié récemment. En particulier, Angluin et al. ont proposé un modèle de calcul réparti dans [Angluin et al., 2004]. Le modèle a été proposé à l'origine pour modéliser les calculs réalisés par les réseaux de capteurs dans lesquels les agents sont passivement mobiles. Le premier exemple cité dans [Angluin et al., 2004] correspond à des capteurs attachés à des oiseaux et programmés pour vérifier certaines propriétés : par exemple comme déterminer si plus de 5 % de la population satisfait une propriété (avoir de la fièvre dans l'exemple de [Angluin et al., 2004]).

6.3.1 Description d'un protocole de population

Dans ce modèle, un *protocole de population* correspond à un ensemble fini d'agents ayant un nombre fini d'états. Les agents interagissent par paire choisie par un adversaire. Chaque interaction a pour conséquence une actualisation des deux agents selon une règle de transition δ . Comme les agents sont anonymes (agents non distinguables), une configuration d'un système à un instant donné est caractérisée par le nombre n_i d'individus étant dans l'état i , pour $1 \leq i \leq k$.

Plus formellement, un *protocole de population* est défini par un uplet $(Q, \Sigma, \iota, \omega, \delta)$. Les agents possèdent un ensemble fini d'états internes Q . Σ correspond à un ensemble de symboles d'entrée. Les fonctions ι, ω correspondent à des fonctions d'encodage des entrées et de décodage des sorties : $\iota : \Sigma \rightarrow Q$ est la fonction qui traduit les symboles d'entrée en états, et $\omega : Q \rightarrow \{0, 1\}$ est la fonction qui décode la sortie associée à un état. La relation $\delta \subseteq Q^4$ décrit comment les paires d'agents interagissent, en fait elle peut être vue comme une liste d'interactions possibles. Par abus de notation, une interaction peut être décrite par la notation $(q_1, q_2) \rightarrow (q'_1, q'_2)$, ou éventuellement par la notation $q_1 q_2 \rightarrow q'_1 q'_2$.

L'exécution du protocole correspond à une séquence d'interactions entre deux agents à partir d'une configuration initiale. À chaque étape, une interaction se déroule entre une paire d'agents. À l'issue d'une interaction entre u et v , les états de ces deux agents évoluent en fonction des règles de transition : l'agent u (resp. v) devient dans l'état u' (resp. v') si $(u, v, u', v') \in \delta$.

Notons par la suite $C \rightarrow C'$ si C et C' sont deux configurations telles que C' peut être obtenu par C par une interaction entre paire d'agents de C .

Une exécution est *équitable* si pour toutes les configurations C , si à partir de la configuration C , il existe une interaction permettant d'atteindre la configuration C' en une étape (c'est-à-dire $C \rightarrow C'$) alors pour toute dérivation $C_0 C_1 \dots$, avec $C_i \rightarrow C_{i+1}$ pour tout i , si C apparaît infiniment souvent, alors C' aussi.

À tout moment, pendant l'exécution d'un protocole de population, l'état q de chaque agent détermine la sortie par la valeur $\omega(q)$ à cet instant : une configuration C s'interprète par 1 (respectivement 0) si tous ses agents sont dans des états tels que $\omega(q) = 1$ (resp. 0).

La majorité des travaux sur ces modèles et leurs variantes portent sur la caractérisation des prédicats calculables $\psi : \mathbb{N}^m \rightarrow \{0, 1\}$.

On dit que le protocole *reconnaît* le prédicat $\psi : \mathbb{N}^m \rightarrow \{0, 1\}$ si pour tout uplet (n_1, \dots, n_m) (configuration initiale), toute exécution débutant par le codage de cet uplet mène à des configurations dont les interprétations sont égales à 1 lorsque $\psi(n_1, \dots, n_m) = 1$, et aussi mène à des configurations dont les interprétations sont égales à 0 lorsque

$\psi(n_1, \dots, n_m) = 0$.

En particulier, dans [Angluin et al., 2004], [Angluin et al., 2006], la puissance de ce modèle a été caractérisée :

Théorème 26 ([Angluin et al., 2006]) – *Tout prédicat $\psi : \mathbb{N}^m \rightarrow \{0, 1\}$ définissable en arithmétique de Presburger est calculé par un protocole de population.*
 – *Réciproquement tout prédicat $\psi : \mathbb{N}^m \rightarrow \{0, 1\}$ calculé est définissable en arithmétique de Presburger.*

Comme les ensembles définissables en arithmétique de Presburger coïncident avec les ensembles semi-linéaires sur les entiers, le prédicat “5% d’agents dans l’état 1” est définissable en arithmétique de Presburger et par conséquent il existe un protocole pour décider s’il existe 5% d’agents dans l’état 1.

Pour plus d’informations sur les résultats sur ces modèles et ses variantes, nous renvoyons notre lecteur à [Angluin et al., 2007], [Aspnes and Ruppert, 2007].

6.3.2 Des jeux dans les protocoles de population

Nous avons vu que les protocoles de population correspondent à des interactions entre paires d’agents. Il est naturel de considérer que chacune de ces interactions est le fruit d’un jeu.

Nous nous sommes intéressées au comportement *pavlovien* et avons cherché à caractériser la puissance des protocoles correspondant.

Nous notons $BR_{\neq q}$ une meilleure réponse parmi toutes les stratégies privées de la stratégie q .

Définition 32 (Protocole de population associé à un jeu) *Considérons un jeu à deux joueurs ayant M_1 et M_2 comme matrices de coût avec $\mathcal{S}_1 = \mathcal{S}_2$. Soit Δ un réel que l’on nommera seuil.*

Le protocole associé à ce jeu correspond au protocole de population dont l’ensemble des états Q est l’ensemble des stratégies ($Q = \mathcal{S}_1 = \mathcal{S}_2$) et dont les règles de transition δ sont définies de la façon suivante :

$$(q_1, q_2, q'_1, q'_2) \in \delta$$

où

- $q'_1 = q_1$ quand $M_1(q_1, q_2) \leq \Delta$
- $q'_1 \in BR_{\neq q_1}(q_2)$ quand $M_{q_1, q_2} > \Delta$

et

- $q'_2 = q_2$ quand $M_1(q_2, q_1) \leq \Delta$
- $q'_2 \in BR_{\neq q_2}(q_1)$ quand $M_{q_2, q_1} > \Delta$,

Définition 33 (Protocole de population pavlovien) *Un protocole de population est pavlovien si on peut le construire à partir d’un jeu à deux joueurs et d’un seuil Δ*

Remarquons que lorsque le jeu n’est pas symétrique, il faut distinguer les deux entités : celui qui a initié l’interaction (correspondant au joueur 1) et celui qui subit cette interaction (correspondant au joueur 2) .

Il existe des protocoles pavloviens qui calculent le *OU* (resp. *ET*) logique sur les bits d'entrée. Il suffit simplement de considérer des protocoles à deux états qui calculent ces opérations : tout protocole à deux états est pavlovien.

Par exemple, voici un protocole calculant le *OU* logique :

$$\left\{ \begin{array}{l} 01 \rightarrow 11 \\ 10 \rightarrow 11 \\ 00 \rightarrow 00 \\ 11 \rightarrow 11 \end{array} \right. \text{ et les matrices de coût } M_1 = M_2 \text{ avec } \Delta = 0 :$$

		Adversaire	
		0	1
Joueur	0	0	+1
	1	+1	0

Par exemple pour la règle $01 \rightarrow 11$, lorsque un joueur jouant 0 rencontre un autre jouant 1, alors celui jouant 0 a un coût positif. Comme il a un comportement pavlovien et que son coût est supérieur au seuil, pour la prochaine interaction, il changera de stratégie en jouant 1.

Une remarque tout simple, est que tous les protocoles de population ne sont pas pavloviens [24]. En effet, il suffit de considérer un protocole de population symétrique à trois états et de construire un ensemble de règles afin de faire une permutation entre les états. Cette propriété empêche de construire une matrice de jeu.

Par la suite, nous noterons les prédicats de la façon suivante [*formule logique*], où une *formule logique* du type $\sum_i a_i \sigma_i \geq b_i$ (respectivement : $\sum_i a_i \sigma_i \equiv b_i \pmod{c_i}$) est vraie si la somme des nombres de symboles σ_i multipliés par a_i est supérieure à b_i (resp. congrue à b_i modulo c_i) : les protocoles calculant le *OU* (resp. *ET*) logique correspondent aussi à des prédicats de comptage d'agents : $[x.0 \geq 1]$ (resp. $[x.1 = 0]$), où x est l'entrée. Il suffit simplement changer les notations pour définir des protocoles de comptage $[x.\sigma \geq 1]$, et $[x.\sigma = 0]$ pour toutes les lettres σ de l'alphabet Σ .

Malheureusement, avec les définitions classiques des protocoles de population, le protocole ci-dessous ne calcule pas un *XOR* logique sur les bits en entrée et donc pas le prédicat de comptage équivalent $[x.1 \equiv 1 \pmod{2}]$.

$$\left\{ \begin{array}{l} 01 \rightarrow 01 \\ 10 \rightarrow 10 \\ 00 \rightarrow 00 \\ 11 \rightarrow 00 \end{array} \right. \quad (6.4)$$

Ce protocole permet de calculer le *XOR* logique dans le sens faible dans lequel au bout d'un certain temps, soit tous les agents sont dans l'état 0 si le *XOR* logique des bits d'entrée est 0. Sinon, seul un seul agent sera dans l'état 1 mais pas tous les agents.

Protocoles associés à un jeu symétrique. Nous conjecturons qu'il n'existe pas de protocoles pavlovien symétrique qui calculent le *XOR* logique (i.e. le prédicat de comptage équivalent à $[x.1 \equiv 1 \pmod{2}]$). De plus, dans [24], une caractérisation partielle de la puissance des calculs des protocoles pavloviens a été étudiée :

Théorème 27 Soit x l'entrée du protocole,

- le prédicat $[x.\sigma \geq k]$, qui est vrai quand le nombre d'occurrences du symbole σ en entrée est plus grand que k avec $k \leq 4$ ou k une puissance de 2.
 - le prédicat $[x.\sigma \geq x.\sigma']$, avec σ, σ' deux symboles en entrée,
- est calculable par un protocole pavlovien associé à un jeu symétrique.

On peut remarquer que le prédicat $[x.\sigma \geq x.\sigma']$ correspond au problème de la majorité qui détermine si le symbole σ apparaît plus que le symbole σ' en entrée.

Le problème d'élection d'un chef peut être résolu par un protocole pavlovien. Il suffit simplement de rendre le protocole classique non-symétrique par un protocole symétrique. Dans le protocole symétrique deux états L_1 et L_2 représentent le fait qu'un agent peut être chef. Voici le protocole résolvant un problème d'élection avec une taille de la population supérieure à 3 :

$$\left\{ \begin{array}{l} L_1L_2 \rightarrow L_1N \\ L_1N \rightarrow NL_2 \\ L_2N \rightarrow NL_1 \\ NN \rightarrow NN \\ L_2L_1 \rightarrow NL_1 \\ NL_1 \rightarrow L_2N \\ NL_2 \rightarrow L_1N \\ L_1L_1 \rightarrow L_2L_2 \\ L_2L_2 \rightarrow L_1L_1 \end{array} \right. \text{ avec sa matrice de jeu } (\Delta = 0) :$$

		Adversaire		
		L_1	L_2	N
Joueur	L_1	+3	0	+3
	L_2	+1	+3	+3
	N	+2	+3	0

Protocoles asymétriques. Dans [Rabie, 2009], la puissance de calcul des protocoles pavloviens asymétriques a été étudiée. Elle semble plus grande :

Théorème 28 Soit $x = (x_1, \dots, x_z)$ l'entrée du protocole,

- le prédicat $[x \cdot \sigma \geq k]$, avec $\sigma = (\sigma_1, \dots, \sigma_z)$, et n'importe quel entier k ;
 - le prédicat $[x \equiv r \pmod{m}]$ avec n'importe quels entiers r et m ;
 - le prédicat $[x.\sigma \geq x.\sigma']$, avec σ, σ' deux symboles en entrée,
- est calculable par un protocole pavlovien asymétrique.

Pour que la puissance de calcul corresponde à l'arithmétique de Presburger, il faudrait construire un protocole calculant des combinaisons linéaires booléennes. Nous conjecturons qu'il n'y en existe pas. Dans ce modèle, il semble difficile d'obtenir des preuves d'impossibilité. Par contre, avec l'ajout de l'hypothèse que les entités peuvent exécuter en parallèle des protocoles pavloviens, alors, on peut prouver que les protocoles pavloviens reconnaissent tout prédicat de l'arithmétique de Presburger.

Des premiers pas ont été amorcé pour les protocoles à population infinie [8], [Aupy, 2009].