Tricher n'est pas jouer !

Simulations et théorie des jeux : problèmes méthodologiques

Bruno Beaufils* — **Philippe Mathieu****

Laboratoire d'Informatique Fondamentale de Lille
Université des Sciences et Technologies de Lille
Cité Scientifique – 59655 Villeneuve d'ascq

* bruno.beaufils@lifl.fr — ** philippe.mathieu@lifl.fr

RÉSUMÉ. *La théorie des jeux computationnelle est un moyen d'étudier et évaluer des comportements d'entités virtuelles en utilisant des modèles de théorie des jeux. L'un des exemples les plus connus de cette approche est le célèbre dilemme itéré du prisonnier classique (DIP). Cette technique permet d'établir non seulement des comportements sophistiqués mais aussi de les comparer et les évaluer. Néanmoins l'organisation de cette évaluation n'est pas aussi simple que certains l'imaginent. Nous montrons dans cet article un certain nombre de points à prendre en compte impérativement dans tout processus d'évaluation. Chaque point est illustré d'un exemple concret et d'une solution permettant d'y remédier.*

ABSTRACT. *Computational game theory is a way to study and to evaluate agents using models from game theory and computer simulations. One of the well known examples is the classical iterated prisoner's dilemma (CIPD). It allows the construction not only of complex behavior but also the evaluation of their* strength. *Setting up such evaluation using classical games situation is not as easy as it seems. We show methodological issues which have to be taken care of, or avoided in order to prevent trouble in simulation results interpretation. Based on some simple illustration, we exhibit two kinds of bias that could be introduced with some idea of how to fixed them.*

MOTS-CLÉS : *théorie des jeux, évaluation de comportement, dilemme itéré du prisonnier, simulation*

KEYWORDS: *game theory, behavior's evaluation, iterated prisoner's dilemma, simulation*

1. Introduction

Les sciences sociales s'occupent généralement du comportement des être humains et de l'effet de ces comportements individuels sur le groupe. Les études dans ce champ de recherche utilisent majoritairement deux types de méthodologies. La première consiste, comme dans les sciences naturelles, à observer les situations réelles avec pour objectif de déterminer les lois qui régissent les groupes observés. La seconde, qui est beaucoup plus récente, consiste à construire des modèles formels qui sont ensuite évalués puis comparés aux observations. Actuellement, ces deux méthodologies ne sont pas aussi distinctes que l'on pourrait le croire. Le plus souvent, la première méthode est utilisée comme préalable à la réalisation de la seconde, dans une boucle infinie tentant d'expliquer le monde.

De nombreux champs scientifiques utilisent de telles méthodes pour leurs études. Outre les sciences économiques, c'est notamment le cas des sciences de gestion, qui essayent de comprendre le comportement de groupes d'individus ou encore de la psychologie et de l'intelligence artificielle qui essayent de comprendre ou de mimer le comportement des individus. Des outils mathématiques ont été constitués avec cet objectif en tête. La théorie des jeux a, par exemple, été initialement créée pour comprendre et expliquer le comportement de personnes dans des jeux de société, (von Neumann *et al.*, 1944). Elle est de plus encore très largement utilisée pour tenter de comprendre le comportement économique des individus et des groupes.

Avec l'accroissement de la puissance informatique, des étapes de tests systématiques sont de plus en plus fréquemment utilisées. La raison principale est que les outils de mathématique pure ne sont pas assez puissants ni suffisamment faciles à utiliser dès qu'ils sont confrontés à une approche individu centrée. C'est notamment le cas des statistiques et des probabilités ou même de la résolution de systèmes à équations différentielles. Ces outils sont très bien adaptés si l'on souhaite avoir une vue globale du monde étudié, mais perdent très vite leur intérêt si l'on souhaite une vue au niveau microscopique. Cette inadéquation vient notamment du fait que l'approche mathématique classique s'appuie sur des modélisations *continues* pour expliquer un monde par nature *discret*. La combinaison de modèles discrets et de la puissance des ordinateurs est maintenant une vraie chance pour la simulation en *sciences sociales*.

Une autre explication de la prolifération des simulations informatiques est liée au fait que les expérimentateurs et spécialistes du domaine étudié considèrent que l'ordinateur est plus facile à utiliser que les outils mathématiques formels. Malheureusement ceci n'est pas souvent vrai. Nous allons dans cet article illustrer les difficultés rencontrées quand on utilise des simulations informatiques puis nous présenterons quelques solutions qu'il est impératif de prendre en compte dans ce type d'exercice.

Nous nous concentrons sur un exemple particulier : l'étude de comportements individuels et leurs effets non seulement sur les individus mais aussi et surtout sur la conduite du groupe. C'est exactement l'objectif initial de la théorie des jeux. Au contraire de la théorie des jeux classique qui essaie de résoudre (ou au moins d'expliquer) les situations conflictuelles et de la théorie des jeux évolutionnaire qui tente d'expliquer la dynamique des populations, la théorie des jeux computationnelle, aussi appelée théorie des jeux empirique, (Walsh *et al.*, 2002), cherche à évaluer les comportements individuels. Dans ce contexte, l'objectif de notre travail est donc de trouver de *bons* comportements encore appelés *stratégies*. Ce que *bons* signifie dépend évidemment des effets macroscopiques souhaités (coopération, coalition dynamique, etc.).

L'approche classique consiste à simuler un grand tournoi impliquant un maximum de comportements différents pour un jeu spécifique. Plus les comportements seront efficaces dans le jeu, plus ils seront considérés comme *bons*. La diversité des types de comportements dans la population d'agents étudiés est évidemment très importante pour garantir un test pertinent. Évaluer un agent dans une population dans laquelle tous les individus vont grosso-modo dans la même direction n'est pas intéressant. Pour éviter cela et assurer une diversité comportementale, les scientifiques utilisent souvent des moyens de collecte dans la population : demander, par exemple, à d'autres scientifiques comment ils se comporteraient dans ce jeu. Cette technique a connu de grands succès plus d'une fois, (Axelrod, 1984) et (Delahaye *et al.*, 1993). L'une des raisons principales est que les participants n'ont *a priori* pas la même perception du jeu. Comme leurs connaissances sont différentes, ils n'ont en général pas les mêmes idées pour construire un comportement, ce qui garantit une grande variabilité de comportements.

Collecter des stratégies est une chose, les évaluer en est un autre. Cela ne peut pas être fait efficacement sans des outils déterministes, permettant de valider l'évaluation quel que soit l'ensemble de stratégies à évaluer. Il faut, entre autres, s'assurer que les outils[1] utilisés sont bien construits et n'introduisent pas des biais dans les résultats de la simulation. Cela nécessite des outils spécifiques permettant de prévenir ou au minimum de mesurer les éventuelles déviations de l'objet étudié. Malheureusement, les personnes utilisant des simulations informatiques pour classifier des comportements n'ont pas toujours ce souci de prévention, soit parce qu'elles n'en sont pas conscientes, soit, parfois malheureusement, parce qu'elles en sont conscientes ! Nous montrons dans cet article quelques effets des mauvaises pratiques possibles illustrées avec des exemples les plus simples possibles.

Dans la première section, nous décrivons les méthodes de théorie des jeux computationnelle que nous utilisons. Un exemple spécifique est choisi et utilisé dans toute la suite de l'article. La section suivante décrit quelques biais qui peuvent être introduits par ces méthodes quand elles ne sont pas utilisées correctement. Ceci pourrait être résumé par la simple maxime *tricher n'est pas jouer !* Puis, dans la troisième section nous décrivons comment des ensembles de stratégies peuvent corrompre les résultats d'une évaluation simplement en utilisant des principes d'accords et de coopération. Ceci pourrait être résumé par *Comment tricher en communiquant*. Dans la dernière section, nous listons quelques débuts de solutions purement opérationnelles permettant d'éviter les problèmes précédemment illustrés.

1. On doit comprendre essentiellement algorithmes.

2. La théorie des jeux computationnelle

La théorie des jeux peut être vue comme un outil mathématique construit pour comprendre ce qu'est un comportement rationnel dans une situation donnée. Dans ce champ de recherche le terme comportement est alors nommé *stratégie*. Ces stratégies décrivent en détail ce qu'un agent doit faire dans l'ensemble des situations possibles. Ces situations sont décrites par les règles du jeu.

Dans cet article nous souhaitons montrer que l'évaluation de stratégies en théorie des jeux via des simulations informatiques n'est pas si triviale que l'on pourrait le croire. Il est indispensable de mettre en place des gardes-fous, d'être rigoureux quant aux expérimentations mais également d'être conscient des limites induites par cette méthode. Nous ne nous intéressons ici qu'aux jeux itérés. L'évaluation communément utilisée consiste à effectuer un championnat comme ce qui est fait classiquement dans les épreuves sportives (football notamment). L'objectif est d'évaluer et d'ordonner différentes stratégies pour un jeu spécifique. Dans cet article nous nous concentrons, à titre d'exemple, sur le dilemme itéré du prisonnier, cependant les réflexions menées restent valides quel que soit le jeu itéré utilisé.

Lors de telles évaluations généralement réalisées par simulations informatiques, il est malheureusement possible d'obtenir des résultats biaisés en fonction de l'implémentation utilisée, cf (Beaufils *et al.*, 2006). Ces résultats risquent alors de justifier de la qualité d'un comportement sans aucun fondement. Que l'on soit dans le cadre d'un concours avec un organisateur et des participants ou que l'on soit dans le cadre d'une simple expérience, la problématique reste identique. Dans la suite de l'article nous faisons référence à une situation de concours qui permet plus facilement de différencier les rôles de l'évaluateur et de l'évalué. Deux cas de figures se présentent :

– les erreurs d'implémentation,

– la triche organisée

Dans le premier cas il s'agit d'une implémentation souvent trop naïve et qui génère des effets pervers. La faute en incombe alors à l'évaluateur et non aux participants. On peut dire qu'il y a triche *involontaire*. Dans le second cas un participant malveillant est capable de profiter du manque d'organisation pour biaiser les résultats et, par exemple, favoriser une stratégie qu'il a soumise. On peut dire qu'il y a *triche volontaire*.

Nous montrons dans cet article non seulement les différents écueils à éviter mais aussi quelles sont les implémentations rigoureuses garantissant la qualité et surtout la robustesse des classements obtenus.

2.1. *Les jeux itérés*

Les jeux itérés que nous étudions sont basés sur la répétition d'un jeu simple qui s'appuie lui-même sur une matrice de gains.

Chaque étape d'un jeu correspond au choix par les participants d'un coup dans une liste fixe de coups disponibles. En fonction des coups choisis par les différents joueurs une situation de jeu est déterminée. Cette situation est appelée une issue du jeu simple. La matrice de gains, cf exemple du tableau 2.1, permet de représenter les gains obtenus pour un joueur en fonction de l'issue atteinte pour un jeu simple.

		B				
		C_1	$\cdots$	C_i	$\cdots$	C_n
A	C_1	$(a_{1,1}, b_{1,1})$	$\cdots$	$(a_{1,i}, b_{1,i})$	$\cdots$	$(a_{1,n}, b_{1,n})$
	$\vdots$	$\vdots$		$\vdots$		$\vdots$
	C_i	$(a_{i,1}, b_{i,1})$	$\cdots$	$(a_{i,i}, b_{i,i})$	$\cdots$	$(a_{i,n}, b_{i,n})$
	$\vdots$	$\vdots$		$\vdots$		$\vdots$
	C_n	$(a_{n,1}, b_{n,1})$	$\cdots$	$(a_{n,i}, b_{n,i})$	$\cdots$	$(a_{n,n}, b_{n,n})$

Ce jeu implique 2 joueurs, A et B. $a_{i,j}$ et $b_{i,j}$ correspondent aux gains respectifs du joueur A et du joueur B lorsque l'issue (C_i, C_j) a été atteinte, c'est-à-dire que le joueur A a choisi le coup C_i et B le coup C_j.

Tableau 1. *Une matrice de gain pour un jeu simple à 2 joueurs*

Le gain total d'un joueur est alors la somme de ses gains pour chaque jeu simple effectué. L'objectif de tous les joueurs est le même : maximiser son gain final.

Il est important de rappeler que l'objectif n'est pas de battre son adversaire mais de maximiser son *capital* en termes de score. Il est, par exemple, possible de perdre face à un adversaire mais d'obtenir un score élevé, comme le contraire.

Un joueur est identifié à la manière dont il fait les choix de ces coups lors de la suite des jeux simples effectuée. Ces choix, prédéterminés à l'avance, sont décidés par une stratégie. Cette stratégie permet, pour une itération donnée, de fixer le coup qu'un joueur va choisir en fonction des situations de jeu rencontrés jusqu'à l'itération courante et notamment en fonction des coups choisis par les adversaires. En termes informatiques on identifie une stratégie à un programme, et un comportement au cours d'un jeu à l'exécution de ce programme.

Dans le cas du dilemme itéré du prisonnier les joueurs sont au nombre de 2, ont à leur disposition 2 coups possibles, notés C et D. Pour chaque rencontre les 2 participants choisissent leur coup simultanément. La matrice de gain, représentée dans le tableau 2.1, identifie le gain obtenu par chacun des joueurs en fonction de l'issue atteinte. Cette matrice est connue des 2 joueurs. D'un point de vue théorie des jeux on est alors dans le cas d'un jeu à information complète et parfaite.

Afin d'utiliser des stratégies relativement connues et simples à comprendre, dans la suite de cet article nous n'utilisons, à titre d'exemple, uniquement que des stratégies du dilemme itéré du prisonnier. Ce jeu est certainement un des modèles de théorie des jeux les plus étudiés et les plus utilisés dans la littérature.

		B	
		C	D
A	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

Cette matrice correspond à une matrice du dilemme du prisonnier. Ce jeu implique 2 joueurs et 2 coups identiques (C et D) pour chacun des joueurs. La matrice de gain est symétrique : le score de A pour l'issue (C,D) est le même que celui de B pour l'issue (D,C).

Tableau 2. *Exemples de matrice de gains*

On utilisera notamment les stratégies suivantes :

`all_c` coopère toujours, quel que soit le comportement de son adversaire.

`all_d` trahis toujours, quel que soit le comportement de son adversaire.

`tit_for_tat` coopère au premier coup, ensuite à chaque coup joue le coup joué par son adversaire au coup précédent.

`spiteful` coopère jusqu'à ce que son adversaire ait trahi, après quoi elle trahit toujours quel que soit son comportement suivant[2].

`soft_majo` joue le coup que son adversaire a majoritairement jouée dans l'histoire passé de la partie. S'il a joué autant de fois `C`, que `D`, alors je coopère.

`per_X` joue périodiquement les coups de la suite X. Par exemple la stratégie `per_cdd`, joue périodiquement `C`,`D`,`D`.

Une étude rapide du dilemme du prisonnier en un coup amène comme solution que l'issue rationnelle pour chacun des joueurs est l'issue où les 2 joueurs jouent le coup `D`. En théorie des jeux on dit que cette issue du jeu est un équilibre de Nash, c'est-à-dire une situation dans laquelle aucun des joueurs ne peut regretter d'être arrivé sachant ce que l'autre a joué.

2.2. *La méthode d'évaluation*

La méthode d'évaluation d'une stratégie la plus utilisée dans ce type de jeu est la comparaison de son score à celui d'une autre stratégie. Plus une stratégie a un score élevé, mieux on la considère. Pour évaluer les stratégies d'un panel donné, on imagine donc souvent de façon naïve de faire la somme des scores de chaque stratégie de ce panel face à toutes les stratégies de ce panel puis de classer les stratégies en fonction de leur score.

Avec cette méthode on exécute en fait une sorte de championnat entre stratégies, on appelle cela un tournoi. Par exemple pour un jeu à 2 joueurs, effectuer un tournoi sur un panel $\mathcal{S}$ de n stratégies revient donc d'abord à calculer une matrice $n \times n$, puis à faire la somme des scores par ligne, pour obtenir le score de chaque stratégie.

Il est important de noter que le score d'une stratégie dans un tel tournoi n'inclut pas obligatoirement le score de la stratégie jouant face à elle-même. Afin d'éviter

2. Cette stratégie est également appelée `grim` ou `trigger` dans la littérature.

d'avoir à traiter explicitement l'apport de cette inclusion en termes de robustesse nous considérons dans la suite de cet article que les cases de la diagonale de cette matrice sont toutes nulles, c'est-à-dire qu'une stratégie n'est jamais évaluée face à elle-même. L'hypothèse oppposée pourrait être faite : une stratégie pourrait être évaluée face à elle-même soit par une rencontre normale, soit par une rencontre dans laquelle la stratégie coopère face à elle même. La simplification utilisée ne change strictement rien aux idées avancées et quasiment rien aux exemples illustrant nos propos.

En reprenant l'exemple du dilemme du prisonnier nous illustrons la méthode utilisée pour évaluer 3 stratégies `all_c`, `all_d` et `tit_for_tat`.

Pour effectuer un tournoi entre ces 3 stratégies il faut remplir la matrice correspondant aux gains de chacune des stratégies et donc effectuer 3 matchs (dans l'exemple chacun des matchs dure 10 coups) :

– `all_c` vs `all_d`

```
      all_c [    0 with   10 C] : CCCCCCCCCC
      all_d [   50 with    0 C] : DDDDDDDDDD
```

– `all_c` vs `tit_for_tat`

```
      all_c [   30 with   10 C] : CCCCCCCCCC
tit_for_tat [   30 with   10 C] : CCCCCCCCCC
```

– `all_d` vs `tit_for_tat`

```
      all_d [   14 with    0 C] : DDDDDDDDDD
tit_for_tat [    9 with    1 C] : CDDDDDDDDD
```

Grâce à ces résultats il est alors aisé de remplir la matrice de gains complète :

	all_c	all_d	tit_for_tat	
all_c	0	0	30	30
all_d	50	0	**14**	64
tit_for_tat	30	**9**	0	39

Chacune des cases de cette matrice correspond au score de la stratégie représentée en ligne face à la stratégie représentée en colonne. De ce fait chaque match permet de remplir 2 cases de la matrice. Par exemple le match `all_d` vs `tit_for_tat` rapporte 14 points à `all_d` et 9 points à `tit_for_tat`.

Le classement des 3 stratégies peut donc se faire aisément en sommant les valeurs des cases de chacune des lignes. Dans l'exemple qui nous concerne on obtient donc le classement suivant :

1) `all_d`
2) `tit_for_tat`
3) `all_c`

Avec cette méthode, plus une stratégie est haute dans le classement plus sa valeur est importante. La qualité de l'évaluation dépend alors principalement du nombre de stratégies impliquées dans le tournoi.

Classiquement une technique utilisée pour obtenir un grand nombre de stratégies est alors d'organiser un concours auprès d'individus en leur demandant de soumettre chacun une stratégie. Le tournoi est alors organisé en utilisant comme panel de stratégie l'ensemble des stratégies soumises. L'idée principale est de pouvoir rechercher de nouvelles *bonnes* stratégies pour un jeu donné en espérant ainsi garantir une hétérogénéité de celles-ci.

Cette méthode a été très régulièrement utilisée, notamment à propos du dilemme du prisonnier, par exemple par Robert Axelrod au début des années 1980, (Axelrod, 1984), par la revue *Pour la science* au début des années 1990, (Delahaye, 1992) pour l'annonce et (Delahaye *et al.*, 1993) pour les résultats, ou encore lors de différentes conférences scientifiques comme Artificial Life V en 1996 ou CEC en l'an 2000. Lors des conférences scientifiques, l'objectif est plus souvent de tester de nouvelles techniques (comme le calcul évolutionnaire) en l'appliquant à des stratégies pour le dilemme du prisonnier.

Dans tous les cas le *modus operandi* est souvent le même : un organisateur demande à des participants de fournir des stratégies qu'il se charge de tester de la manière la plus objective possible en organisant un tournoi grâce à des simulations informatiques. Les participants fournissent leur stratégie via un programme informatique, une description ou plus simplement en remplissant un formulaire web décrivant le comportement qu'il propose. L'organisateur s'assure généralement de l'objectivité du concours en organisant les simulations et en étudiant minutieusement les résultats de celles-ci.

Des biais dans l'interprétation des résultats de tels tournois peuvent facilement apparaître. Cette apparition peut être volontaire ou involontaire.

3. Tricher sans communiquer

3.1. *Le problème de la normalisation*

L'un des premiers problèmes dans la version itérée du dilemme du prisonnier consiste à éviter le paradoxe du *pendu*. Si les joueurs connaissent la longueur des parties alors ils peuvent en arriver à considérer chacun des coups d'une partie comme le dernier et, consécutivement, toujours jouer le coup `D` qui est le coup de l'équilibre de Nash du jeu simple, voir (Binmore, 1999, 350) pour une idée de la preuve de cet écueil. Pour éviter cela, il faut faire en sorte qu'aucun des joueurs ne connaisse la longueur de la partie afin d'éviter cette convergence malheureuse.

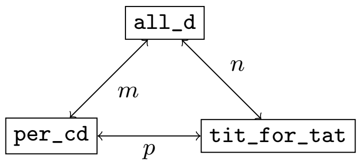
Une des solutions théoriques généralement mises en œuvre est d'utiliser un facteur d'escompte. Ce facteur permet de réduire l'importance des coups à venir par rapport au coup passé. Cette solution est envisageable tant qu'il ne s'agit que de faire des études formelles avec des parties de longueur infinie sans expérimentation. La validité du taux d'escompte n'étant établie que dans le cas d'horizon de jeux infinis. Dans le cas de simulations calculées, la difficulté du choix de la valeur du taux d'escompte et, surtout, l'impossibilité d'observer la fin d'un programme comportant une boucle infinie rend son utilisation impossible.

Pour cela certains expérimentateurs préfèrent tirer au sort, pour chaque rencontre, la longueur de la partie avec un écart moyen et un écart type fixé, et cela sans communiquer ce résultat aux joueurs.

Malheureusement il existe de nombreux cas où un joueur malchanceux peut se retrouver toujours avec des parties courtes par rapport à la moyenne. Dans de telles conditions il perd, alors qu'avec un peu de chance il aurait pu gagner.

En guise d'illustration on peut utiliser 3 stratégies `all_d`, `tit_for_tat`, `per_cd`. Considérons les longueurs de parties suivantes :

Joueurs	**Longueur**
`per_cd` vs `all_d`	m
`all_d` vs `tit_for_tat`	n
`tit_for_tat` vs `per_cd`	p



Lors d'un tournoi les scores obtenus par chacune de ces stratégies en fonction des longueurs de parties sont alors les suivants :

	`all_d`	`per_cd`	`tit_for_tat`
`all_d`		$6(m/2)$	$n+4$
`per_cd`	$m/2$		$(5p/2)+3$
`tit_for_tat`	$n-1$	$(5p/2)-2$	

Avec $n = 50$, $m = 100$, $p = 60$, donc un écart moyen de 20, on obtient comme classement :

1) `all_d` avec 354 points ;
2) `per_cd` avec 203 points ;
3) `tit_for_tat` avec 197 points.

Si par contre $n = 50$, $m = 60$, $p = 100$, donc toujours avec un écart moyen de 20, on obtient alors le classement :

1) `tit_for_tat` avec 297 points ;
2) `per_cd` avec 283 points ;
3) `all_d` avec 234 points.

Il est important de noter que ces 2 classements sont exactement l'opposé l'un de l'autre en ce qui concerne la tête et la queue du classement. Il faut également remarquer que l'écart moyen et l'écart type de ces distributions de longueurs de parties sont toujours identiques, à savoir 20 et 26,46.

Si on normalise la longueur des parties $n = m = p$ on obtient alors le classement suivant :

1) `all_d` avec $4(n+1)$ points ;

2) `per_cd` avec $3(n+1)$ points ;

3) `tit_for_tat` avec $\frac{7n}{2} - 3$ points.

Il faut également remarquer que la normalisation des gains à la valeur du coup moyen de ces deux tournois donne évidemment un résultat identique qui est différent des deux précédents :

1) `all_d` avec un gain moyen de 4 points ;

2) `tit_for_tat` avec un gain moyen de 3 points ;

3) `per_cd` avec un gain moyen de 3,5 points.

La longueur des parties doit être inconnue des joueurs mais on doit, soit choisir aléatoirement, au départ de la compétition, la longueur de la partie qui sera alors la même pour tous, soit normaliser les gains en fonction du nombre de coups joués par chacun. Dans un tournoi on ne doit pas jouer contre soi-même même si dans une compétition écologique on joue évidemment contre ses semblables.

Dans la suite de l'article nous utilisons, pour simplifier la présentation, des parties de longueur 10 pour les différents tournois présentés.

3.2. *Plusieurs fois la même stratégie*

Dans un contexte particulier, 2 stratégies, a priori différentes, peuvent avoir le même comportement. En illustration on peut citer `tit_for_tat` et `spiteful`, qui sont clairement des stratégies différentes, mais qui ont cependant exactement le même comportement face à 2 autres stratégies aussi différentes que `all_c` et `all_d`.

D'une manière générale on rencontre le même cas de figure quand on compare un génotype et son phénotype. Dans un vocabulaire informatique on dit que 2 programmes différents peuvent produire le même résultat.

Vu différemment mais dans le même ordre d'idée il se peut que plusieurs stratégies de noms différents soient au bout du compte complètement identiques. Dans ce cas la multiplication de la présence d'une telle stratégie dans l'ensemble des participants est un biais à l'évaluation de sa qualité.

Si on a, par exemple, n fois la même stratégie `S1` (au sens du même code mais pas forcément du même nom) avec un ensemble d'autres stratégies toutes différentes deux à deux, `S1` est souvent la meilleure. En revanche, si on ne met qu'un seul exemplaire de S1 c'est une autre stratégie qui est considérée comme la meilleure en remportant le tournoi.

En considérant tit_for_tat pour S1 et les 2 stratégies all_d, et per_cd on peut aisément se convaincre d'une telle situation sur un tournoi avec des parties de 10 tours :

```
TOURNAMENT RANK                TOURNAMENT RANK
1 :       all_d = 44           1 : tit_for_tat = 62
2 :      per_cd = 33           2 : tit_for_tat = 62
3 : tit_for_tat = 32           3 :      per_cd = 61
                               4 :       all_d = 58
```

Nous avons montré comment la présence répétée d'une même stratégie influence les évaluations. Cette situation fréquente se retrouve essentiellement dans 2 cas de figures :

– lorsque 2 stratégies ont des codes identiques mais des noms différents. Il est particulièrement aisé pour un participant à un concours de volontairement favoriser sa stratégie en inondant le concours avec celle-ci en lui donnant des noms tous différents. En théorie des jeux, comme dans d'autre contexte, des problèmes identiques ont déjà été identifiés. Il s'agit de la possibilité pour un agent de manipuler le résultat d'un processus de choix en cachant certaines informations en sa possession. Il s'agit alors d'identifier les mécanismes manipulables ou de construire des méthodes résistantes aux manipulations, quand cela est possible. Le cœur de ces problèmes est la notion de *strategy-proof* en théorie des jeux, (Vickrey, 1961), ou plus récemment le concept de *false bid names* (Yokoo *et al.*, 2004), plus particulièrement sur les systèmes d'enchères combinées ;

– lorsque 2 stratégies ont des codes différents, des noms différents mais une exécution identique due au contexte de la simulation. C'est d'ailleurs un problème classique en informatique puisqu'il existe une infinité de manières d'écrire la même chose.

On peut remarquer que déterminer que deux stratégies sont identiques est un problème indécidable dans le cas général (c'est un résultat fort d'informatique théorique) mais parfaitement décidable dans les cas contraints du genre remplissage de formulaire de codage de comportement. C'est d'ailleurs une technique utilisable facilement pour des saisies interactives de stratégies sur le web. Implicitement cette technique de définition de stratégies est moins intéressante car on peut souvent effectuer un tournoi avec toutes les instanciations possibles, voir par exemple les travaux de (Beaufils *et al.*, 1998a). Dans un tel cas, l'évaluation est alors complètement décidable.

4. Tricher en communiquant

Il est communément admis lors d'une rencontre que les joueurs ne connaissent a priori pas les stratégies de leurs adversaires. Ils peuvent essayer de la déduire en fonction de leur comportement. C'est la difficulté et un des intérêts majeurs du problème, qui rentre dans le cadre de l'inférence comportementale. En effet si on connaît d'avance la stratégie de l'adversaire on peut déterminer la meilleure réponse possible qui permet alors d'obtenir le gain maximum.

Bien qu'il soit impossible de communiquer directement avec son adversaire, afin par exemple de s'accorder sur un comportement particulier, il est néanmoins possible d'utiliser le passé pour communiquer *indirectement*. L'astuce principale réside simplement dans le fait d'utiliser une amorce reconnaissable, dans les premiers coups du jeu, pour être identifié par l'adversaire.

La communication peut alors s'instaurer par un système de codage qui a été décidé au préalable par le concepteur. Pour tricher deux types de cas sont envisageables :

– une unique stratégie utilisée par plusieurs joueurs qui tentent de se reconnaître ;

– plusieurs stratégies dont l'objectif de certaines est de favoriser les autres.

4.1. *Les stratégies qui se reconnaissent*

On peut illustrer ce cas de figure par une stratégie qui joue une amorce puis qui si elle reconnaît son amorce coopère toujours, sinon trahit toujours.

Dans un tel cas il se peut que si la moitié de la population est constituée de joueurs utilisant cette stratégie, un joueur de cette sous-population devienne le gagnant. En revanche si un seul joueur utilise cette stratégie alors ce joueur ne le sera plus puisqu'il ne pourra plus tirer avantage de cette communication.

Considérons par exemple la stratégie `cheater` qui joue l'amorce (`CDDC`) puis observe les coups joués par son adversaire lors des 4 premiers coups. Si l'adversaire a joué cette même suite de coups alors `cheater` coopère toujours, sinon elle trahit toujours.

Si on place 3 joueurs utilisant cette stratégie face à `tit_for_tat`, `spiteful` et `soft_majo` le vainqueur d'un tournoi est `cheater`. Si on fait la même expérience avec une seule occurrence de `cheater` alors il est dernier du classement :

```
TOURNAMENT RANK                TOURNAMENT RANK
1 :    cheater2 = 109          1 :    spiteful = 75
2 :    cheater3 = 109          2 : tit_for_tat = 74
3 :     cheater = 109          3 :   soft_majo = 73
4 :    spiteful = 105          4 :     cheater = 57
5 : tit_for_tat = 102
6 :   soft_majo = 99
```

Une fois que 2 joueurs utilisant `cheater` se sont reconnus ils prennent 3 points à chaque tour alors que le reste du temps, face aux autres, ils en font perdre au moins 2 points à leur adversaire.

En général ce genre de comportement peut s’identifier aisément sur le résultat car les joueurs trichant de cette manière sont tous groupés dans le classement obtenu avec le même score (en tout cas dans le cas de stratégies déterministes).

4.2. *Les maître/esclaves*

Sur le même principe que précédemment on peut concevoir une stratégie qui profite d’un ensemble d’autres stratégies qui se laissent exploiter. Celle qui profite des autres est dite la maîtresse et les autres sont considérés comme des esclaves.

Par rapport à la situation précédente, une fois que les joueurs se sont reconnus, le maître prend 5 points à chaque tour en trahissant et les esclaves se laissent exploiter en coopérant uniquement face au maître. Face aux autres joueurs les esclaves peuvent adopter n’importe quel autre comportement agressif.

Si on considère, par exemple, que la stratégie maîtresse `master` qui joue (`CDDC`) puis, si elle a reconnu son amorce joue toujours `D`, alors que si elle ne la reconnaît pas elle joue comme `tit_for_tat`.

La stratégie esclave fait elle comme `cheater` : elle joue (`CDDC`) puis si l’adversaire a fait la même chose elle coopère toujours sinon elle trahit toujours.

Si on évalue la stratégie `master` face à `cheater`, `tit_for_tat`, `soft_majo` et `spiteful`, elle est alors clairement en tête d’évaluation alors que l’esclave se retrouve à la fin de ce classement. S’il n’y a pas d’esclave elle est dernière du classement :

```
TOURNAMENT RANK                TOURNAMENT RANK
1 :      master = 105          1 : tit_for_tat = 84
2 : tit_for_tat = 98           2 :    soft_majo = 83
3 :   soft_majo = 96           3 :     spiteful = 75
4 :    spiteful = 90           4 :       master = 67
5 :     cheater = 65
```

Un tel exemple pourrait même être affiné en considérant un maître et n esclaves tous avec des amorces différentes. Le travail du maître serait alors juste de reconnaître les différents types d’esclaves et de jouer `tit_for_tat`, `spiteful` ou toute autre *bonne* stratégie face aux autres joueurs.

On peut remarquer que le pouvoir de nuisance de l’esclave peut être sélectif et ne nuire qu’à certains. Par exemple changer le comportement de `cheater` de manière à ce qu’elle joue `tit_for_tat` au lieu de `all_d` aura pour effet de ne nuire qu’aux stratégies agressives.

Dans une vraie compétition il faut cependant trouver un participant qui accepte d’utiliser un comportement qui sera mal classé. Si on peut identifier ces joueurs coalisés, par exemple parce que la règle de concours l’autorise explicitement, alors on peut imaginer leur attribuer la moyenne des gains des membres de la famille, ce qui corrigerait ce problème. Malheureusement cette identification est en général difficile. En effet étant donné que l'objectif des compétitions n'est pas d'évaluer des groupes mais des individus, leurs règles interdisent implicitement ce genre d'actions.

Dans un concours il faut éviter qu'un même participant puisse présenter plusieurs stratégies. Il reste vrai que dans une compétition ouverte ce genre de collusion peut aussi s'obtenir par un accord entre plusieurs participants. Il faut donc également surveiller le classement final obtenu et étudier en détail les stratégies obtenant des scores identiques.

5. Les *solutions*

Après avoir présenté quelques-unes des difficultés que l'évaluation de stratégies peut entraîner nous présentons quelques méthodes rendant cette évaluation plus objective et surtout plus fiable, notamment en termes de robustesse. Nous présentons trois méthodes. La première consiste essentiellement à fixer l'environnement de comparaison de façon à pouvoir comparer efficacement les stratégies, les deux autres au contraire tentent d'évaluer un comportement dans des environnements différents et mesurent donc plus précisément la robustesse de la valeur d'une stratégie.

5.1. *Évaluation de chaque stratégie contre un panel de base*

Hormis les problèmes de normalisation et d'implémentation des outils de simulations, dont nous avons présenté quelques solutions simples, le problème principal de l'évaluation des comportements semble être aisément identifiable : il manque un référentiel de comparaison. Une solution possible à ce problème est donc de créer artificiellement ce référentiel. On peut par exemple créer un environnement de base dans lequel chaque stratégie peut être plongée aux fins d'évaluations.

La comparaison de 2 stratégies `A` et `B` revient alors simplement à comparer la valeur de chacune des stratégies face à l'environnement spécifié. La méthode utilisable peut se résumer à :

– fixer un ensemble de stratégies $\mathcal{S}$;

– exécuter un tournoi impliquant toutes les stratégies de $\mathcal{S}$ plus `A` pour en déterminer le rang de `A` dans ce tournoi ;

– exécuter un tournoi impliquant toutes les stratégies de $\mathcal{S}$ plus `B` pour en déterminer le rang de `B` dans ce tournoi ;

– comparer le rang de `A` et `B`.

Ce type de comparaison à deux peut évidemment se généraliser à n. Classiquement des panels de base sont reconnus et construits au fur et à mesure du temps. On peut en retrouver quelques-uns notamment dans les premières expériences d'Axelrod ou dans celles que nous avons effectuées. Une autre technique est d'utiliser de très grands ensembles de stratégies différents construits de manière automatique à partir d'un modèle particulier. C'est par exemple ce qui est fait dans (Beaufils *et al.*, 1998b).

On peut néanmoins remarquer que la véritable difficulté est simplement déportée sur le choix des stratégies composant cet ensemble de base. En effet, la valeur de la stratégie est alors très dépendante de l'environnement choisi. La valeur d'une stratégie dans ce cas est relative à l'environnement en question. On est cependant dans une évaluation plus objective que la simple comparaison de n stratégies entre elles.

Il est d'ailleurs possible d'exploiter la relative faiblesse de cette approche. On peut, par exemple, imaginer construire des environnements étalons spécifiques aux qualités recherchées pour un comportement. Chaque panel étalon peut, par exemple, correspondre à un environnement spécifique dans lequel on veut tester la qualité d'un comportement. On peut imaginer quelques exemples simples : environnement homogène de comportements agressifs, ou au contraire de comportements bienveillants mais réactifs, etc. L'utilisation de chacun de ces panels étalons pour déterminer la valeur d'une stratégie permet de qualifier cette valeur.

Finalement une mesure encore plus absolue peut alors être une composition (pondérée ou non) de valeurs obtenues dans des panels étalons différents. Cette composition est alors un gage de qualité de la valeur calculée puisqu'elle prend en compte la force d'une stratégie face à différents environnements.

Il reste néanmoins vrai que l'étalonnage de cette mesure reste un problème crucial. En effet la robustesse des valeurs obtenues dépend grandement des choix faits pour les panels de base.

5.2. *Compétitions écologiques*

Pour parer au manque de robustesse inhérent à la solution précédente, une seconde méthode d'évaluation des stratégies est de simuler les principes de la sélection naturelle : moins un individu est fort, moins il a de chance de survivre. Cette simplification des mécanismes d'évolution est utilisée dans le cadre d'évolution de populations d'agents utilisant les stratégies d'un panel.

L'idée est de faire en sorte que chaque agent choisisse une stratégie dans le panel et rencontre un à un tous les autres agents. Il y a donc plusieurs représentants d'une même stratégie. Une fois que toutes les rencontres ont eu lieu chacun est alors capable de connaître la *valeur* de la stratégie qu'il utilise en cumulant les scores obtenus face à tous les autres. Les agents ayant choisi une stratégie ramenant peu de points ont tendance à disparaître, alors que ceux ayant choisi une stratégie rapportant beaucoup de points vont au contraire *croître et multiplier*. Le cycle recommence ensuite et on peut voir fluctuer la représentation des différentes stratégies dans la population globale.

L'évolution est reproduite tant que ces représentations ne se sont pas stabilisées, c'est-à-dire tant qu'entre deux cycles elles évoluent.

Le mécanisme évolutionnaire utilisé est une sorte de reproduction parthénogénétique dans laquelle un individu seul peut se reproduire. La population est polymorphique dans le sens où chaque individu a un comportement différent. On applique ainsi à la théorie des jeux les principes biologiques de l'évolution, lui donnant du même coup une nouvelle utilité, non pas d'étude des comportements économiques humains, mais d'étude de la dynamique des populations d'animaux.

La population de nos agents ne sera jamais infinie, un agent est indivisible, et à un moment donné, il faut pouvoir lui dicter son comportement. C'est pourquoi contrairement aux biologistes, et aux théoriciens des jeux, nous restons volontairement dans un cadre discret, déterministe, et calculable. C'est le cadre idéal pour les simulations informatiques.

Nous utilisons donc un algorithme basé sur la même idée que celle de (Smith, 1982), en y apportant les quelques adaptations nécessaires.

Tout d'abord, nous limitons les stratégies utilisables par les agents à celles d'un panel particulier. Soit $\mathcal{S}$ ce panel, et $|\mathcal{S}|$ la taille de ce panel. Les stratégies de $\mathcal{S}$ sont notées s_i avec $i \leq |\mathcal{S}|$. On considère de plus que le nombre d'individus est fixé, et invariable du début à la fin de l'évolution. Soit Π la taille de cette population. On note $W_t(\texttt{A})$ le nombre d'agents utilisant la stratégie `A` à la génération t. Autrement dit W_t(A) est la taille de la sous-population d'agents utilisant `A`. On dit que c'est la taille de la population de `A`.

Pour toute valeur de t on a :

$$\Pi = \sum_{i=1}^{i=|\mathcal{S}|} W_t(s_i)$$

Le score d'un individu utilisant la stratégie s_i à la génération t est $V_t(s_i)$ avec :

$$V_t(s_i) \quad = \quad \left(\sum_{j=1}^{j=|\mathcal{S}|} W_t(s_j) V(s_i|s_j) \right) - V(s_i|s_i)$$

Un individu a pour score la somme des scores qu'il obtient en jouant contre tous les autres joueurs, sauf lui.

Finalement la population de toutes les stratégies s_i est redistribuée de manière proportionnelle au score par :

$$W_{t+1}(s_i) \quad = \quad \alpha_t W_t(s_i) V_t(s_i)$$

α_t est le coefficient de normalisation de la population

$$\alpha_t = \frac{\Pi}{\sum_{j=1}^{j=|\mathcal{S}|} W_t(s_j)V_t(s_j)}$$

Pratiquement tous les calculs se font sur des nombres entiers, on se contente donc de travailler dans un espace discret. La seule exception étant pour la normalisation : on arrondit à l'entier le plus proche pour le calcul des nouvelles tailles de population.

Dans les évolutions écologiques de bases, on choisira pour tout i

$$W_0(s_i) = \frac{\Pi}{|\mathcal{S}|}$$

Ce genre d'évaluation pose évidemment beaucoup plus de questions que les tournois, comme par exemple, la question essentielle de la stabilité des répartitions de populations. Existe-t-il des stratégies qui, sous certaines conditions, entraînent toujours les évolutions vers une stabilité des populations, un peu comme les stratégies évolutionnairement stables en théorie des jeux évolutionnaires, (Smith, 1982) ? Existe-t-il des situations d'instabilité chronique, des oscillations des répartitions de populations, des situations au bord du chaos, (Mathieu *et al.*, 1999) ?

On peut imaginer une approche purement mathématique. Elle nécessite d'être dans un cadre continu pour pouvoir faire de l'intégration notamment sur le calcul de la dynamique des populations. Pour arriver à cette approche on fait souvent une simplification du modèle de la dynamique en gérant des représentations en nombres réels (pourcentage) de la taille des populations. Dans notre cas nous utilisons des nombres entiers (donc des calculs exacts) grâce à l'utilisation d'arrondi.

Pour faciliter leur intégration, il est d'ailleurs aussi tentant de supprimer le terme $-V(s_i|s_i)$ du calcul de $V_t(s_i)$. Ceci permettrait ainsi à un joueur de se rencontrer lui même. De cette manière, on obtient un système d'équations simples continues intégrables facilement. Calculer de manière exacte avec une méthode d'arrondi entier est cependant important, parce que dans la réalité les individus ne sont pas divisibles.

La différence fondamentale entre l'approche continue et discrète est que dans le cas continu une stratégie ne disparaît jamais, alors que dans le réalité ça n'est pas le cas : un individu peut disparaître d'une population. Ceci étant posé, nos expériences montrent que, arrondi ou pas, les dynamiques sont sensiblement identiques.

Avec ce genre de solution la robustesse de la mesure effectuée est grandement améliorée. Les cas de triches potentiels présentés dans les sections précédentes deviennent très vite évitables.

5.3. *Sous-classes*

Un autre moyen d'éviter le phénomène volontaire ou involontaire de stratégies qui s'entraident consiste à ne pas calculer les évaluations sur l'ensemble des stratégies fournies, mais sur des sous-ensembles de celles-ci. Pour peu que chaque stratégie soit présente dans le même nombre de sous-ensembles, la répartition est alors équitable et il devient possible de classer ces stratégies par la somme cumulée des résultats qu'elles ont obtenu dans tous les sous-ensembles dans lesquels elles ont participé. Un moyen efficace pour obtenir des sous-ensembles équitables pour chaque stratégie consiste à calculer toutes les combinaisons de p stratégies parmi les n possibles. Nous appelons `sous-classes(p,n)` les sous-ensembles obtenus par cette technique des combinaisons (calcul des C_n^p ensembles de p stratégies parmi n). Par exemple avec $n = 10$ stratégies et $p = 9$ on obtient 10 sous-classes de 9 stratégies. De même `sous-classes(8,10)=45` sous-ensembles possibles. Dans le cas général une stratégie participera à exactement C_{n-1}^{p-1} sous-ensembles, ce qui est fortement représentatif quand p est proche de n (rapport de $\frac{n}{p}$), mais surtout, si une stratégie en aide une autre, elle ne pourra le faire que dans C_{n-2}^{p-2} évaluations (rapport de $\frac{n(n-1)}{p(p-1)}$. celle qui est aidée affichera alors un score pitoyable un grand nombre de fois dans les évaluations réalisées affichant donc un cumul déplorable.

n	p	évaluations	présence d'1	rencontre entre 2
		C_n^p	$C_n^p * \frac{p}{n}$	$C_n^p * \frac{p(p-1)}{n(n-1)}$
n	$n-1$	n	$n-1$	$n-2$
n	$n-2$	$\frac{n(n-1)}{2}$	$\frac{(n-1)(n-2)}{2}$	$\frac{(n-2)(n-3)}{2}$
n	$n-3$	$\frac{n(n-1)(n-2)}{6}$	$\frac{(n-1)(n-2)(n-3)}{6}$	$\frac{(n-2)(n-3)(n-4)}{6}$
10	9	10	9	8
10	8	45	36	28
10	7	120	84	56
20	19	20	19	18
20	18	190	171	153
20	17	1140	969	816
50	49	50	49	48
50	48	1225	1176	1128
50	47	19600	18424	17296
100	99	100	99	98
100	98	4950	4851	4753
100	97	161700	1656849	152096
1000	999	1000	999	998
1000	998	499500	498501	497503
1000	997	166167000	165668499	165170996

L'avantage de cette technique de sous-classes est qu'elle peut être utilisée aussi bien pour des évaluations en tournoi que dans des évaluations en compétitions écologiques. Le cumul des résultats peut aussi se faire de deux manières différentes, soit en cumulant les scores obtenus dans chaque sous-classe, mais dans un jeu à somme non nulle cela n'a pas vraiment de sens, soit en cumulant les classements d'une stratégie, ce qui est sémantiquement plus intéressant. Pour reprendre notre exemple, si une stratégie obtient un cumul de 10 cela signifie alors qu'elle a été classée première dans les 10 sous-classes calculées. Bien sûr, C_n^p croît très vite, on se contente donc en général de p très proche de n ($p = n - 1$ ou $p = n - 2$), mais cette technique est parfaitement équitable et en général très informative et très efficace pour repérer *qui aide qui*.

Cette technique des sous-classes est par contre peu efficace pour les cas où il y a un grand nombre d'esclaves puisque chaque sous-ensemble calculé contiendra toujours des esclaves, mais elle apporte une nouvelle couche de robustesse dans l'évaluation des stratégies. Elle est sans doute insuffisante, un maître rencontrant souvent son esclave, mais elle a au moins l'avantage de l'objectivité du choix de ses rencontres.

6. Conclusion

L'évaluation de stratégies dans le cadre de la théorie des jeux et plus précisément des jeux itérés rentre dans un cadre de recherches que l'on pourrait nommer *théorie des jeux computationnelle*. Comme il est en général difficile voire impossible d'apporter la preuve formelle qu'une stratégie est meilleure qu'une autre dans l'absolu, la compétition informatique de stratégies devient indispensable.

De nombreux concours ont d'ailleurs lieu régulièrement à travers le monde sur des modèles de jeux bien connus (dilemme itéré du prisonnier, Roshambo, etc.). Dans cet article nous avons montré à travers le cas du dilemme itéré du prisonnier que cette évaluation ne doit pas être faite à la légère et qu'il est très facile d'obtenir des résultats biaisés. Consciemment ou inconsciemment il est facile de favoriser une stratégie ou même une famille de stratégies dans un concours. Nous avons décrit dans cet article plusieurs biais dont il faut être conscient et auxquels il faut prendre garde comme la méthode d'évaluation, la répétition de stratégies, les maître-esclaves, etc.

Nous avons pris un soin particulier à fournir de nombreux exemples et à les décrire pour qu'ils soient reproductibles. Nous avons proposé plusieurs méthodes permettant de les éviter comme les compétitions écologiques, les classes complètes et les sous-classes. Bien sûr, aucune de ces méthodes n'est à elle seule ni parfaite ni même suffisante, mais chacune d'entre elles apporte un peu plus de robustesse au classement des stratégies. Une bonne évaluation consiste alors à utiliser toutes ces méthodes et à les prendre toutes en compte à l'aide une fonction paramétrique. Selon le type de robustesse souhaité, les paramètres de la fonction devront être ajustés.

Cet article se veut donc un *survey* de ce qu'il faut faire et ne pas faire lors de la comparaison informatique de stratégies dans le cadre des jeux itérés.

7. Bibliographie

Axelrod R., *The evolution of cooperation*, Basic Books, New-York, USA, 1984.

Beaufils B., Delahaye J.-P., Mathieu P., « Complete Classes of Strategies for the Classical Iterated Prisoner's Dilemma », *in* V. W. Porto, N. Saravanan, D. Waagen, A. E. Eiben (eds), *Evolutionnary Programming VII*, vol. 1447 of *Lecture Notes in Computer Science*, Springer-Verlag, p. 33-41, 1998a. Evolutionnary Programing VII, San Diego, CA, USA, March 25-27, 1998.

Beaufils B., Delahaye J.-P., Mathieu P., « Complete Classes of Strategies for the Classical Iterated Prisoner's Dilemma », *in* V. W. Porto, N. Saravanan, D. Waagen, A. E. Eiben (eds), *Evolutionnary Programming VII (EP'7)*, vol. 1447 of *Lecture Notes in Computer Science*, Springer, p. 33-41, 1998b.

Beaufils B., Mathieu P., « Cheating is not playing: Methodological Issues of Computational Game Theory », *in* G. Brewka, S. Coradeschi, A. Perini, P. Traverso (eds), *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI'06)*, IOS Press, p. 185-189, 2006.

Binmore K., *Jeux et théorie des jeux*, DeBoeck, 1999.

Delahaye J., « L'altruisme récompensé ? », *Pour La Science (French Edition of Scientific American)*, vol. 181, p. 150-156, 1992.

Delahaye J., Mathieu P., « L'altruisme perfectionné », *Pour La Science (French Edition of Scientific American)*, vol. 187, p. 102-107, 1993.

Mathieu P., Beaufils B., Delahaye J.-P., « Studies on Dynamics in the Classical Iterated Prisoner's Dilemma with Few Strategies: Is There Any Chaos in the Pure Dilemma ? », *in* C. Fonlupt, J.-K. Hao, E. Lutton, E. Ronald, M. Schoenauer (eds), *Proceedings of the 4th european conference on Artificial Evolution (AE'99)*, vol. 1829 of *Lecture Notes in Computer Science*, Springer, p. 177-190, 1999.

Smith J. M., *Evolution and the Theory of Games*, Cambridge University Press, 1982.

Vickrey W., « Counterspeculation, auctions, and competitive sealed tenders », *Journal of Finance*, vol. 16, p. 8-37, 1961.

von Neumann J., Morgenstern O., *Theory of Games and Economic Behavior*, Princeton University Press, 1944. The standard reference is the revised edition of 1947.

Walsh W. E., Das R., Tesauro G., Kephart J. O., « Analyzing complex strategic interactions in multi-agent games », *AAAI-02 Workshop on Game Theoretic and Decision Theoretic Agents*, 2002.

Yokoo M., Sakurai Y., Matsubara S., « The effect of false-name bids in combinatorial auctions: new fraud in internet auctions », *Games and Economic Behavior*, vol. 46, p. 174-188, 2004.