

---

# Faut-il toujours suivre Nash ?

**Bruno Beaufils — Philippe Mathieu**

*Laboratoire d'Informatique Fondamentale de Lille  
Université des Sciences et Technologies de Lille  
Cité Scientifique – Bât M3  
59655 Villeneuve d'Ascq cédex  
{beaufils,mathieu}@lifl.fr*

---

*RÉSUMÉ. L'une des critiques fréquentes que l'on trouve sur la théorie des jeux est qu'elle s'appuie sur des modèles dans lesquels les coups et gains possibles sont fixés par avance. Nous présentons dans cet article un modèle de théorie des jeux dans lequel les agents doivent raisonner sur la matrice de gains pour établir une bonne stratégie. Dans ce jeu itératif modélisant des interactions entre agents, nous montrons par ailleurs que l'équilibre de Nash, notion réputée en micro-économie, n'aide pas à obtenir de bonnes stratégies dans les simulations informatiques par essence discrètes. Nous proposons une classification de stratégies dans ce modèle et montrons que ce modèle générique peut servir à construire de bonnes stratégies adaptables aux interactions entre agents lors d'échanges simultanés.*

*ABSTRACT. One of the most frequently used criticism about game theory is that it is based on models in which moves and payoff are fixed in advance. In this paper we introduce a new game theoretic model in which agents can not build good strategies without analysing payoff matrices. In this agents interactions modelling iterated game we show that a well reputed notion of micro-economics, Nash equilibrium, does not help in computer simulations, which are discrete by nature. We propose a strategy classification and show that this generic model may help building good strategies, which may be adapted to agents interactions during simultaneous exchange.*

*MOTS-CLÉS : Théorie des jeux, Simulations, Évolution*

*KEYWORDS: Game Theory, Simulations, Evolution*

---

## 1. Introduction

La théorie des jeux a été introduite par [NEU 44]. Son but originel est de permettre la modélisation des situations de conflits telles qu'on les retrouve dans les jeux de salons. Elle considère généralement que les joueurs ne sont que des agents rationnels dont le but est de maximiser leur *bien-être*. Un jeu particulier, le dilemme itéré du prisonnier classique (CIPD), introduit par [FLO 52], a permis de relativiser les techniques et résultats préconisés par cette théorie. Cependant ce même modèle a permis le développement d'une théorie de la coopération notamment suite à la publication de [AXE 84]. La théorie des jeux est depuis lors très utilisée notamment par l'informatique dès qu'il s'agit de modéliser et d'étudier la coopération entre agents et son évolution ou encore d'étudier les comportements de différents agents entre eux.

Il est intéressant de noter que la plupart des résultats ou conseils énoncés par cette théorie sont établis à partir d'un cadre continu infini (les scores sont calculés en moyenne et les rencontres entre joueurs sont de longueur infinie). Lors de la construction de systèmes multi-agents les comportements dont on étudie soit l'émergence soit l'influence ne sont que très rarement liés à des interactions infinies. L'informatique reste un monde discrétisé. Il nous semble de ce fait évident qu'une adaptation des modèles théoriques à ce cadre discret et fini ne peut rendre ces outils que plus fiables et adaptés aux besoins des informaticiens. Nous avons déjà commencé un travail d'adaptation qui est présenté dans [DEL 99, BEA 00] uniquement dans le cadre du dilemme itéré du prisonnier classique, dans lequel la matrice est fixée par avance avec uniquement deux coups possibles. Nos travaux ont permis une large utilisation des simulations informatiques pour valider, corriger ou améliorer des résultats issus du cadre classique.

Le point central de la théorie des jeux est la notion d'équilibre introduite par Nash dans [NAS 51]. Cette notion permet d'identifier dans les jeux à deux joueurs des situations de *non-regret* dans lesquelles des agents rationnels vont donc essayer de se placer. Par extension et abus de langage on parle souvent de stratégie de Nash. De telles stratégies sont des stratégies permettant de placer les joueurs dans une situation d'équilibre. La notion même de stratégie est définie en théorie de manière très formelle par les différentes suites de mouvements qui sont offertes à un joueur. En informatique on préfère souvent décrire le comportement par une approche plus fonctionnelle, si bien que deux comportements différents peuvent se ramener à une seule et même stratégie, au sens strict du terme. Les utilisations de la théorie des jeux qui sont faites par les praticiens de l'informatique se permettent souvent des approximations sur les notions utilisées. Les résultats communément admis par de telles utilisations, notamment sur Nash, peuvent de ce fait être erronés, mal adaptés ou tout simplement inutiles.

Nous nous proposons dans cet article de tenter un début de correction de ces défauts. À cet effet nous introduisons un nouveau *jeu itéré* permettant de représenter au mieux les interactions entre agents. Ces interactions se ramènent dans ce modèle à un jeu dont la matrice de gains peut être modifiée à chaque itération, c'est-à-dire

après chaque interaction. Nous ne nous intéressons pas aux raisons qui pourraient justifier ces modifications mais plus pratiquement à la manière dont les agents les gèrent. Ce modèle nous semble beaucoup plus intéressant que, par exemple, le modèle du dilemme itéré du prisonnier.

Bien qu'il soit vrai qu'un très grand nombre d'échanges commerciaux ou d'interactions réelles peuvent souvent se ramener à un simple jeu à deux joueurs et donc à une matrice de gain, ces matrices ne se ramènent cependant que rarement à une matrice de dilemme du prisonnier. On peut en revanche imaginer que ces échanges se modélisent par un refus ou une acceptation de contrats (avec d'éventuelles variantes sur les termes du contrats) et que la matrice de gains correspond alors aux profits pour chacun des intervenants. Les différents échanges ne finissent que rarement sur des profits ou des termes de contrats identiques. L'intérêt évident de ce modèle reste donc bien l'introduction d'une matrice de gain variable.

Le modèle que nous présentons est en quelque sorte un méta-modèle qui permet donc d'englober les jeux classiques. Les stratégies considérées comme efficaces dans ces jeux, [BEA 96, BEA 01], ne le restent que rarement dans d'autres modèles. On peut a contrario détecter grâce à ce modèle des stratégies efficaces globalement et qui sauront être adaptées aisément dans ces jeux. Les critères de qualité dans ce modèle peuvent aisément être appliqués dans d'autres jeux. De plus ce méta-jeu permet de montrer de manière probante la non adéquation de la théorie des jeux classique et de ses résultats à l'utilisation qui en est souvent faite en informatique.

## 2. Présentation du méta-jeu

On considère un jeu à deux joueurs qui jouent simultanément une carte parmi  $n$ . La situation de jeu, définie par le choix de chacun des deux joueurs, est considérée comme l'issue du jeu dont les scores sont déterminés par une matrice de gains de taille  $n \times n$ . Cette matrice contient les scores qu'un joueur obtient contre son adversaire. Elle est remplie aléatoirement entre 0 et  $k$  au début du jeu. Le jeu est à information complète mais imparfaite. La matrice de gain est connue des deux joueurs. Les joueurs savent que leur adversaire connaît cette matrice. Le gagnant est le joueur qui obtient le plus de points.

Le tableau 1 représente un exemple de matrice de jeu. Le jeu offre trois cartes possibles (A, B ou C). Si l'issue du jeu est l'issue [B,A], c'est-à-dire que le joueur ligne choisit la carte B pendant que le joueur colonne choisit la carte A, alors le joueur ligne remporte  $U_{21} = 8$  points tandis que le second obtient le gain  $U_{12} = 2$  points.

Lorsque ce jeu de base est itéré, une partie entre deux joueurs se déroule en  $\ell$  coups successifs. Le nombre  $\ell$  de coups n'est pas connu à l'avance par les joueurs. À chaque itération, la matrice de gains peut changer. Aussi, les joueurs sont informés de la nouvelle matrice de gains au début de chaque coup. L'information restant complète mais imparfaite ils ont également accès, après chaque coup, au coup joué par leur adversaire et donc à leur gain et à celui de leur adversaire. Le gain obtenu par un

<i>U</i>	A	B	C
A	6	2	3
B	8	5	7
C	9	8	2

**Tableau 1.** Exemple de matrice de gains (avec  $n = 3$  et  $k = 9$ )

joueur est égal à la somme des points qu'il a obtenue à chaque itération. Le gagnant est le joueur qui a accumulé le plus de points durant la partie.

Un tel type de jeu permet de représenter la quasi totalité des jeux itérés à deux joueurs, information complète et imparfaite. Le dilemme itéré classique du prisonnier peut par exemple être une instance de ce jeu si chacun des tirages aléatoires donne la même matrice de gain respectant les équations de ce jeu particulier. On nommera donc les jeux de ce type les *méta-jeux*.

Plusieurs variantes de ce méta-jeu peuvent être envisagées en fonction des contraintes que l'on impose sur le tirage aléatoire des matrices de gains. Ces contraintes peuvent porter à la fois sur la taille de la matrice de gains et sur les valeurs des scores. De façon à conserver l'aspect complet mais imparfait de l'information il faudra cependant transmettre la connaissance de ces diverses contraintes aux joueurs.

Une des variantes singulières qu'il est possible de donner au jeu, est de fixer  $k = 1$ . On choisit alors de ne travailler que sur des valeurs binaires. Pour un jeu simple, le joueur gagne s'il a un gain de 1 et son adversaire de 0, il fait match nul si les deux joueurs ont le même gain et il perd s'il a 0 et son adversaire 1. Pour ce jeu itéré, le joueur qui obtient le plus de points est celui qui a remporté le plus de victoires. Cette propriété n'est pas vraie dans le cas général. En contre-exemple, supposons une partie à trois coups dans laquelle le joueur A gagne 10, 0 et 0 points et le joueur B gagne 0, 1 et 1 points. Le joueur A gagne la partie avec 10 points contre 2 points pour B, alors que B a remporté le plus de coups (2 coups victorieux pour B contre un seul pour A).

Le contenu des matrices choisies peut être soumis à des contraintes fortes sur la répartition des valeurs dans celles-ci. On peut par exemple ne permettre le choix que de matrices représentant des jeux possédant des issues en équilibre au sens de Nash. Le nombre de ces situations d'équilibre peut même être inclus dans ces contraintes et impliquer de stratégies adaptées. Nous nous intéresserons essentiellement à ce type de méta-jeu dans la suite de cet article.

Un joueur joue selon une *stratégie* qui décrit complètement son comportement durant le jeu. Cette stratégie est entièrement déterminée avant le début du jeu. Nous en présenterons quelques unes dans la prochaine section. Afin d'évaluer les stratégies et de pouvoir les comparer, on les fait s'affronter dans un tournoi. Lorsque  $p$  stratégies sont présentes, un tournoi est organisé entre chacune d'elles. On utilise  $p$  joueurs utilisant chacun une stratégie différente. Chaque stratégie joue donc, tour à tour, une partie contre chacune des autres, y compris elle-même.

Afin que toutes les stratégies soient évaluées de manière équitable, les  $\ell$  matrices de gains, utilisées pour les  $\ell$  coups de chaque partie, sont choisies une fois pour toutes au début du tournoi. Ainsi, à chaque rencontre entre deux stratégies, c'est la même  $\ell$  matrice qui est utilisée au  $i^{\text{e}}$  coup. Les différentes rencontres d'un tournoi sont donc soumises au même environnement.

Le score obtenu par une stratégie est égal à la somme des scores que cette stratégie a obtenue dans chacune des parties jouées. On pourra construire une matrice de taille  $p \times p$  des scores obtenus dans chacune des rencontres. Par exemple, avec la matrice de scores présentée dans le tableau 2 (à laquelle on a ajouté une colonne indiquant le score total pour le tournoi), la stratégie S1 obtient des scores de 510 points en jouant contre elle-même, de 608 points en jouant contre la stratégie S2 et de 332 points en jouant contre S3, soit 1450 points au total. Avec un score de 1543 points au total, c'est cependant la stratégie S2 qui remporte ce tournoi.

	S1	S2	S3	Score
S1	510	608	332	1450
S2	312	560	671	1543
S3	405	489	495	1389

**Tableau 2.** Exemple de matrice de scores d'un tournoi entre 3 stratégies

Il est assez aisé de se rendre compte avec une telle méthode d'évaluation que la rationalité que l'on considère pour les joueurs de méta-jeux n'est pas définie par la faculté pour un agent de battre son adversaire mais plutôt par celle de maximiser son score.

Le second type d'évaluation utilisé est inspiré des processus d'évolution et de sélection naturelle, et est très proche des problèmes de dynamique de population.

Considérons une population de  $N$  joueurs, chacun choisissant une stratégie dans un ensemble de  $p$  stratégies (on considère que  $p < N$ ). Un tournoi incluant tous les joueurs, et non plus seulement les stratégies, est effectué. Contrairement au tournoi présenté précédemment une même stratégie peut donc être présente plusieurs fois.

Le score collectif des joueurs utilisant une même stratégie est pris en compte. Les stratégies efficaces (cumulant de gros scores) sont favorisées par rapport aux moins bonnes. Ce favoritisme passe par une redistribution des populations de joueurs utilisant chacune des stratégies représentées. Cette redistribution se fait de manière proportionnelle. Ce schéma, aussi appelé une génération, est répété jusqu'à une éventuelle stabilisation de la population, *i.e.* pas de changements dans la population entre deux cycles.

Une bonne stratégie est alors une stratégie qui reste dans la population le plus longtemps possible, et dans la plus grande proportion possible. Ce type d'évaluation demande des qualités supplémentaires aux stratégies et notamment la capacité de bien se comporter face à elle-même. Ce type d'évaluation se nomme *évolution écologique*

et a notamment été popularisé par [AXE 84]. Classiquement on considère qu'au départ de l'évolution chaque stratégie de l'ensemble test est représentée dans la population et que toutes les sous-populations de stratégies ont la même taille.

### 3. Stratégies

Contrairement aux jeux classiques le terme de *stratégie* est utilisé dans les méta-jeux pour définir le comportement d'un agent dans l'absolu. Une même stratégie doit pouvoir être utilisée dans plusieurs instances de méta-jeux, sans connaître a priori les différentes matrices de gain et donc les différentes issues du jeu itéré. Une stratégie ne peut donc pas ici être une simple énumération de choix en fonction des choix de l'adversaire.

Dans les méta-jeux, les stratégies correspondent beaucoup plus à des comportements décrits de manière générique. Elles sont plus proches de processus descriptifs qu'énumératifs. Deux stratégies complètement différentes peuvent amener à des déroulements de jeux totalement identiques. En ce sens une stratégie est ici facilement identifiable à un programme informatique.

Une particularité des méta-jeux est que les stratégies ne peuvent se contenter de se baser uniquement sur le jeu de l'adversaire. La nature même de l'interaction est modifiée à chaque itération. Les stratégies doivent donc non seulement prendre en compte le comportement de l'adversaire mais également analyser la matrice à laquelle elles ont affaire. En d'autres termes elles doivent être capables d'inclure dans leur processus de choix l'analyse de la matrice de gain à chaque itération.

D'un point de vue opérationnel une stratégie se décompose en plusieurs opérations qui correspondent à chacune des phases d'une étape d'un jeu :

- 1) une phase d'initiation durant laquelle le joueur est informé de la nouvelle matrice de gains, libre à lui de la stocker ou de la pré-analyser ;
- 2) une phase de choix durant laquelle le joueur est interrogé sur le coup qu'il désire jouer et doit alors impérativement donner une réponse ;
- 3) une phase de terminaison durant laquelle le joueur est informé du coup joué par son adversaire, libre à lui de le stocker afin de le prendre en compte pour la prochaine étape du jeu. Cela lui permet également de connaître son score ainsi que celui de son adversaire pour cette étape du jeu.

Une stratégie précise donc quels sont les calculs et décisions à effectuer pour chacune de ces trois phases. Pour les phases d'initiation et de terminaison, il existe un comportement par défaut et le joueur peut ne définir que l'opération de la phase de choix. Pour l'initiation, le comportement par défaut se contente de stocker la matrice de gains actuelle uniquement pour la durée de cette étape. Pour la terminaison, le comportement par défaut ne fait rien.

Un certain nombre de stratégies est décrite dans la suite de cette section. Si aucune précision n'est donnée, la description correspond à la phase de choix, les phases d'initiation et de terminaison utilisent le comportement par défaut. Ainsi, la stratégie de la première ligne : «*jouer la première ligne de la matrice de gains*» définit une stratégie qui joue toujours la carte numéro 1, la première ligne de la matrice de gains.

Certains comportements sont qualifiés d'*incomplets*. Un comportement incomplet n'est capable de choisir une carte que dans certaines circonstances. Par exemple, le comportement suivant est incomplet : «*si la matrice de gains comporte au moins un équilibre de Nash, jouer le premier équilibre de Nash trouvé*». Pour définir une stratégie, il faut en plus associer un comportement pour la close *sinon*. C'est pourquoi les stratégies *composites* permettent d'associer un comportement incomplet à une stratégie. Ainsi, dans notre cas, on peut définir une stratégie composite en disant : «*jouer le comportement incomplet de Nash, sinon, jouer la stratégie de la première ligne*».

Nous allons maintenant donner un certain nombre de définitions de stratégies. Seules les idées des stratégies seront décrites. Pour de plus amples détails sur le comportement ou l'implémentation de celles-ci on peut se référer à [DEL 00]. Ces stratégies sont celles qui sont utilisées dans les expérimentations décrites dans la prochaine section.

Les stratégies *aveugles* choisissent une ligne à l'aveugle, c'est-à-dire sans même prendre en compte les scores de la matrice de gains :

**Bêta** choisit toujours de jouer la première ligne de la matrice de gains.

REMARQUE. — la taille des matrices de gains étant variable, il n'est pas possible de jouer toujours une ligne d'un certain numéro autre que 1 ou 2 (la taille est supérieure à 2).

**Nigaud** choisit toujours de jouer la dernière ligne de la matrice.

**Boucle** commence par choisir la première ligne, puis la seconde, la troisième et ainsi de suite. Elle incrémente le numéro choisi à chaque étape du jeu.

**RandomStrat** joue un numéro de ligne qu'elle choisit aléatoirement entre 1 et la taille de la matrice.

Les stratégies *calculatrices* s'appuient sur les scores de la matrice de gains et vont l'analyser à l'aide de calculs :

**Naive** joue la ligne contenant le gain de valeur maximale.

**SemiNaive** joue la ligne dont la somme des gains est maximale. Si plusieurs lignes ont une somme maximale, elle joue la première qu'elle rencontre (celle dont le numéro est le plus petit).

**AggressiveNaive** calcule les scores maxima de chaque colonne et choisit la carte de la ligne symétrique à la colonne pour laquelle cette valeur est la plus basse.

Lorsque le joueur choisit la  $i^{eme}$  ligne, son adversaire va obtenir un des gains de la  $i^{eme}$  colonne. Minimiser les gains de l'adversaire revient donc à jouer la ligne pour laquelle la colonne correspondante offre les gains les plus faibles. Cette stratégie est dite *agressive* parce qu'elle ne cherche pas à obtenir le meilleur score possible, mais cherche à minimiser celui de l'adversaire.

**AgressiveSemiNaive** calcule la somme des scores de chacune des colonnes de la matrice des gains. Elle joue ensuite la ligne symétrique à la colonne pour laquelle cette somme est la plus basse.

**Equitable** joue la ligne qui a la plus grande valeur sur sa diagonale. Elle est équitable dans le sens où elle cherche à ce que son adversaire ait le même score (et soit le plus élevé possible).

**Prudente** calcule les scores minima de chaque ligne et choisit la ligne de plus grand minimum. Elle est prudente au sens où elle s'assure un score minimum le plus élevé possible.

**MaxGap** cherche à maximiser l'écart entre les scores des deux joueurs (en sa faveur bien sûr). Elle revient donc à jouer la ligne contenant le maximum pour la matrice  $U - U^t$ , si  $U$  est la matrice de gains du joueur.

**MaxLineGap** Tout comme la stratégie MaxGap, cette stratégie vise à maximiser l'écart entre les scores des deux joueurs (toujours en sa faveur). Ici, cependant, on calcule la somme des gains sur les lignes de la matrice  $U - U^t$  et on joue la ligne pour laquelle cette somme est maximale.

**IntCommunEgoiste** recherche les coups pour lesquels le gain commun est maximal. S'il y a plusieurs maxima, elle joue celui qui offre la plus grande valeur pour son propre gain. Cette stratégie cherche l'intérêt commun, c'est à dire qu'elle cherche à maximiser le gain commun.

**FirstNash** cherche les équilibres de Nash de la matrice de gains. Elle joue la carte permettant d'atteindre le premier équilibre qu'elle rencontre en parcourant la matrice colonne par colonne. C'est une stratégie incomplète.

**MaxNash** cherche les équilibres de Nash de la matrice de gains et joue celui qui offre le gain le plus grand pour le joueur. C'est une stratégie incomplète.

**NashIntCommun** cherche les équilibres de Nash de la matrice de gains et joue celui qui offre le gain commun le plus grand. C'est une stratégie incomplète.

Les stratégies adaptatives cherchent à modifier leur comportement au cours du jeu en fonction du comportement de leur adversaire :

**SelfBest** cherche à jouer le mieux possible contre elle-même. Elle vise donc à maximiser ses gains ainsi que ceux de son adversaire quand elle joue contre elle-même. Pour cela elle met en place un processus permettant à deux joueurs l'utilisant de se retrouver dans la situation de gain collectif maximal en utilisant un protocole de synchronisation particulier.

**SelfBestIB** est basée sur la stratégie SelfBest et en reprend le principe, à savoir jouer le mieux possible contre soi-même. Cependant, elle corrige le défaut principal



de celle-là. En effet, alors que SelfBest joue très bien contre elle-même, elle joue de manière inefficace contre les autres joueurs, et obtient de piètres résultats en tournoi.

Pour remédier à ce problème, SelfBestIB va essayer d'identifier son adversaire. Tant que son adversaire joue le comportement attendu, elle joue comme SelfBest. Au premier écart de l'adversaire par rapport au comportement attendu, SelfBestIB passe la main à une autre stratégie. Ceci est facilité par le fait que SelfBestIB est un comportement incomplet. Il doit donc être associé à une autre stratégie, cette dernière prenant le relais quand l'adversaire ne joue pas suivant la stratégie SelfBest.

Par exemple, si l'on associe SelfBestIB à SemiNaive, le joueur va commencer par utiliser la stratégie SelfBest, puis dès qu'il se rend compte que son adversaire ne joue pas suivant une stratégie SelfBest, il joue suivant la stratégie SemiNaive. Ainsi, on profite à la fois de la stratégie SelfBest, qui joue bien contre elle-même, et de la stratégie SemiNaive, qui maximise l'espérance du gain, contre les autres joueurs.

#### 4. Expériences

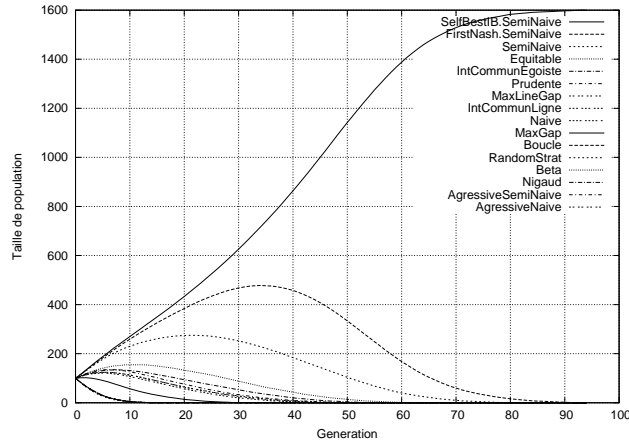
Afin, entre autres, de mettre à l'épreuve les résultats classiques de la théorie des jeux et de son cadre continu nous avons conduit un certain nombre d'expériences à partir d'un logiciel de simulation adapté aux méta-jeux permettant de calculer des tournois mais aussi des évolutions écologiques. Pour toutes les simulations dont les résultats sont présentés ici, les parties durent 1000 itérations, les matrices de gains de chaque itération ont des tailles comprises entre 2 et 5 et la variable  $k$  est à valeur dans  $[0, 10]$ . Afin de pouvoir reproduire les tests à l'identique le jeu de matrices de gains est toujours le même, résultat d'un tirage avec la clé 960021934641 dans notre simulateur.

Tout d'abord nous avons conduit un certain nombre d'expérimentations en imposant comme contrainte que chaque matrice de gain comporte une et une seule issue d'équilibre au sens de Nash. Ceci revient à dire qu'à chaque coup la rationalité théorique imposerait les stratégies choisissant la carte favorisant cette issue.

Étant donné qu'il n'y a qu'un seul équilibre de Nash, les trois stratégies FirstNash, MaxNash et NashIntCommun (qui jouent respectivement le premier équilibre de Nash trouvé, celui de gain maximum et celui qui maximise le gain commun) vont toujours jouer de la même manière. Les tests reprennent donc l'ensemble des stratégies décrites précédemment mais avec une seule de ces trois stratégies, ici FirstNash.

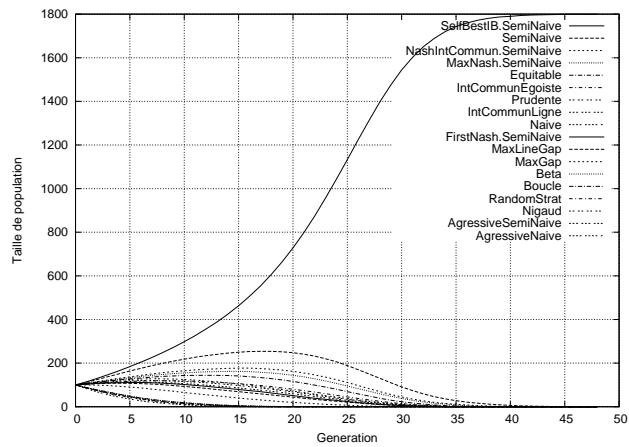
La figure 1 illustre une évolution incluant toutes ces stratégies. La stratégie SelfBestIB.SemiNaive l'emporte devant FirstNash devançant également SemiNaive.

FirstNash semble donc être battue. En étudiant les scores en tournoi, on s'aperçoit que SemiNaive domine les scores au niveau des stratégies aveugles, alors que FirstNash domine au niveau des stratégies calculatrices.



**Figure 1.** Évolution : variante avec un unique équilibre de Nash

Dans un autre jeu de simulations les matrices de gains contiennent toutes au moins un équilibre de Nash. Ainsi, les trois stratégies FirstNash, MaxNash et NashIntCommun ont des comportements différents. Les expérimentations portent donc sur un ensemble de stratégies un peu plus grand.



**Figure 2.** Evolution : variante avec au moins un équilibre de Nash

L'évolution, illustrée figure 2, est remportée assez largement par SelfBestIB.SemiNaive, suivie de SemiNaive. NashIntCommun et MaxNash sont respectivement troisi`eme et quatri`eme. Entre les trois stratégies tirant partie des équilibres de Nash, NashIntCommun s'en tire le mieux, suivie de MaxNash puis FirstNash.

Finalement, les résultats obtenus dans cette variante sont très proches de ceux obtenus dans le cas général, en dehors d'une légère amélioration pour NashIntCommun et MaxNash.

Les expérimentations ont permis de mettre à jour plusieurs résultats.

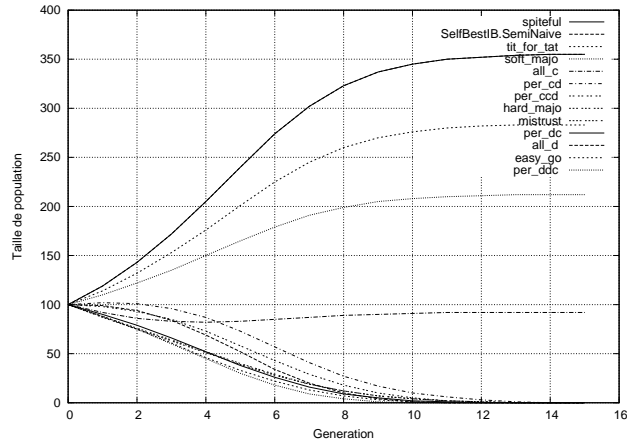
Tout d'abord, il y a un écart entre deux groupes de stratégies. D'un côté un groupe qui obtient des résultats faibles, constitué des stratégies aveugles et des stratégies agressives. De l'autre côté, un groupe qui obtient d'assez bons résultats. Il est composé des stratégies calculatrices non agressives. De plus les stratégies agressives, qui cherchent uniquement à diminuer les points de l'adversaire, obtiennent de piètres résultats. Même si les stratégies agressives gagnent certaines de leur parties, elles obtiennent des scores très faibles et finissent en dernières positions en tournoi et en évolution. Une bonne stratégie ne doit donc pas forcément chercher à gagner une partie contre son adversaire, mais doit chercher à obtenir le score le plus élevé possible, quitte à perdre sa partie. Elle peut perdre une bataille pour gagner la guerre. Cela conforte certains des résultats classiques notamment ceux répandus sur le dilemme itéré du prisonnier, [AXE 84, BEA 00].

Dans le cas où toutes les matrices de gains ont un et un seul équilibre de Nash, même si la stratégie jouant l'équilibre obtient de bons résultats, elle est dominée par SelfBestIB.SemiNaive. Dans le cas où les matrices de gains ont au moins un équilibre de Nash, les stratégies jouant l'équilibre sont dominées de manière plus marquée, et par deux autres stratégies. Bien que ceci ne constitue pas réellement une preuve, ce sont des cas où il y a toujours un équilibre de Nash, et où il existe une meilleure stratégie que celle qui joue l'équilibre.

Les simulations montrent donc bien que jouer pour atteindre l'équilibre de Nash n'est pas forcément le meilleur comportement. De manière intuitive, ceci s'explique par le fait que dans les populations utilisées tous les agents n'ont pas forcément un comportement rationnel. De ce fait essayer d'obtenir une issue d'équilibre peut se retourner contre le joueur dans un cadre de stratégies hétérogènes non forcément rationnelles.

La figure 3 représente une évolution d'un dilemme itéré du prisonnier classique. Cette évolution implique 12 stratégies historiquement utilisées dans la littérature auxquelles on a ajouté la stratégie SelfBestIB. Cet ajout permet de vérifier la solidité de cette stratégie dans un cadre plus simple et également de la comparer à la stratégie `tit_for_tat` mise en avant dans [AXE 84]. C'est la seule manière de comparer ces deux stratégies puisqu'il est possible à un comportement défini dans le cadre des méta-jeux de s'adapter à un jeu particulier alors qu'une stratégie du CIPD ne peut, elle, pas être utilisée dans un méta-jeux.

Cette évolution et une étude de l'adaptation du comportement de SelfBestIB montre que cette stratégie reste *bonne* dans ce cadre. Il est à noter que dans le dilemme classique le comportement qu'adopte SelfBestIB est le même que celui de la stratégie



**Figure 3.** Evolution : *SelfBestIB* dans un cadre classique (CIPD)

*spiteful*<sup>1</sup>, à savoir qu'elle ne joue la solution proposée par Nash que quand elle est certaine que l'adversaire ne cherche pas à maximiser l'intérêt collectif.

## 5. Conclusion

Nous avons introduit dans cet article un nouveau modèle issu de la théorie des jeux. Ce méta-jeu permet de décrire un très grand nombre d'interactions entre agents. Il englobe notamment les modèles les plus utilisés tel le dilemme itéré du prisonnier. Un certain nombre de stratégies adaptées à ce méta-jeu ont été décrites.

Des méthodes d'évaluations adaptées au cas discret et fini ont été décrites et utilisées lors de simulations informatiques. Ces expérimentations ont notamment permis de mettre en lumière l'inadéquation de la rationalité telle qu'elle est décrite en théorie des jeux, au travers la notion d'équilibre de Nash, à l'utilisation que l'on en fait classiquement en informatique. En effet certes les stratégies basées sur cette notion d'équilibre se comportent bien, mais il a été relativement aisé de mettre en place des comportements différents encore plus efficaces. La qualité de ces comportements essentiellement construits sur la notion d'adaptativité se retrouve également dans d'autres modèles. Les résultats qui ont commencé à apparaître tendent à remettre en question la notion de rationalité et de solution des jeux sont donc renforcés par ces expériences.

Il reste cependant que ces résultats ne peuvent masquer les vrais problèmes de la théorie des jeux et de son utilisation. Un grand nombre de résultats admis sont souvent soit mal interprétés soit mal appliqués. La description des stratégies composites

1. *spiteful* coopère tant que l'adversaire coopère puis trahi toujours ensuite

qui sont mis en place pour ce méta-jeu et les simulations effectuées permettent de réaffirmer la différence fondamentale entre la prédiction économique basée sur les statistiques et la compréhension informatique basé essentiellement sur la simulation. Il faut donc savoir adapter les outils de la théorie des jeux aux utilisations que l'on en fait.

Les simulateurs utilisés pour les résultats présentés dans cet article sont disponibles auprès des auteurs et ont été écrits en JAVA.

## 6. Bibliographie

- [AXE 84] AXELROD R., *The evolution of cooperation*, Basic Books, New-York, USA, 1984.
- [BEA 96] BEAUFILS B., DELAHAYE J.-P., MATHIEU P., « Our Meeting with Gradual, A Good Strategy for the Iterated Prisoner's Dilemma », LANGTON C. G., SHIMOHARA K., Eds., *Proceedings of Artificial Life V*, Cambridge, MA, USA, 1996, The MIT Press/Bradford Books, p. 202–209, *Artificial Life V*, Nara, Japan, May 16-18 1996.
- [BEA 00] BEAUFILS B., « Modèles et simulations informatiques des problèmes de coopération entre agents », Thèse de doctorat, Université des Sciences et Technologies de Lille, janvier 2000.
- [BEA 01] BEAUFILS B., DELAHAYE J.-P., MATHIEU P., « Adaptive Behaviour in the Classical Iterated Prisoner's Dilemma », KUDENKO D., ALONSO E., Eds., *Proceedings of AISB'01 symposium on Adaptive Agents and Multi-agent systems*, The Society for the Study of Artificial Intelligence and the Simulation of Behaviour, 2001, p. 65–72.
- [DEL 99] DELAHAYE J.-P., MATHIEU P., « Le monde agité de la coopération », *Pour La Science*, vol. Hors Série, n° 24, 1999, p. 58–66.
- [DEL 00] DELABRE F., « Évaluation de comportements adaptatifs entre agents dans des méta-jeux », Mémoire de DEA, Laboratoire d'Informatique Fondamentale de Lille, juillet 2000.
- [FLO 52] FLOOD M. M., « Some Experimental Games », Research Memorandum n° RM-789-1, 1952, The RAND Corporation, Santa-Monica, CA, USA.
- [NAS 51] NASH J., « Non-Cooperative Games », *Annals of Mathematics*, vol. 54, n° 2, 1951, p. 286–295.
- [NEU 44] VON NEUMANN J., MORGENSTERN O., *Theory of Games and Economic Behavior*, Princeton University Press, 1944, The standard reference is the revised edition of 1947.