
Modélisation de stratégies par Apprentissage et Anticipation génétiques

Christophe Meyer, Jean-Gabriel Ganascia et Jean-Daniel Zucker

*LIP6-IBP-CNRS, Université Paris VI
4, place Jussieu
75252 Paris Cedex 05 - France
Tel. (33-1) 44-27-37-27 ou 44-27-71-19
e-mail: {meyer,ganascia,zucker}@laforia.ibp.fr*

RÉSUMÉ : La théorie des jeux se présente essentiellement comme une théorie mathématique qui vise à substituer un calcul au pur hasard ou à la simple intuition dans la conduite à tenir dans une situation de concurrence. Cette théorie a été très largement utilisée pour la définition de programmes informatiques. Nous avons tenté de concevoir un système artificiel capable de jouer efficacement à des jeux particuliers auxquels la Théorie des Jeux ne peut s'appliquer d'une façon satisfaisante (comme pour les jeux à information non complète ou imparfaite). Notre système a la faculté, lorsqu'il ne trouve pas de solution lui assurant la victoire, de jouer par anticipation en fonction d'une modélisation du comportement de son adversaire qu'il élabore et raffine en temps réel au cours des parties qu'il joue contre lui. L'architecture proposée repose sur deux systèmes de classeurs génétiques qui permettent, l'un de modéliser le comportement des adversaires et l'autre de jouer en fonction de cette modélisation. Des expérimentations sur les capacités d'apprentissage de stratégies de notre système ont été menées sur un jeu stratégique simple. Les résultats illustrent l'intérêt de notre approche en montrant la supériorité d'un tel système sur des adversaires humains ou artificiels classiques (probabilistes simples et probabilistes adaptatifs).

1. Introduction

Malgré son nom, la théorie des jeux, branche de la recherche opérationnelle qui étudie la prise de décision en situation de concurrence, n'est pas uniquement consacrée aux activités ludiques comme le jeu d'échec ou les tournois de bridge. Toutes les activités dites stratégiques parce qu'elles portent sur la coordination et l'agencement des forces face à l'adversité, sont concernées, qu'il s'agisse de diplomatie, de politique, de choix industriels, ou de la conduite des affaires en matière économique et militaire.

Comme toutes les branches de la recherche opérationnelle, la théorie des jeux se présente essentiellement comme une théorie mathématique qui vise à substituer un calcul au pur hasard ou à la simple intuition dans la conduite à tenir. En supposant que les partenaires en concurrence cherchent tous à optimiser leurs avantages respectifs, comment déterminer les lignes de jeu les plus efficaces ? Telle est la question générique que tente de résoudre la théorie des jeux.

Ceci étant, la théorie des jeux aborde différents points de vue selon la nature des jeux et des critères d'optimisation qu'elle prend en considération.

Classiquement, les jeux se distinguent selon que l'information est complète ou incomplète, c'est-à-dire selon que les joueurs ont ou non une connaissance totale du jeu. Ainsi, le jeu d'échec est à information complète tandis que le bridge est à information incomplète. Les jeux se distinguent aussi selon que l'information est parfaite, lorsque les joueurs maîtrisent tous les paramètres du jeu quand ils jouent, ce qui est toujours le cas pour les échecs, ou que l'information est imparfaite lorsque, par exemple, deux joueurs doivent miser simultanément sans savoir ce que mise l'adversaire. Les critères d'optimisation varient aussi selon que l'on suppose que les adversaires désirent minimiser leurs risques de perte, ou qu'ils souhaitent maximiser leurs chances de gagner en évitant, autant que faire se peut, les parties nulles.

Dans le passé, les théories mathématiques sur les jeux remontent à l'origine de la théorie des probabilités, pour des jeux de pur hasard, c'est-à-dire sans autre concurrent que le sort. La théorie mathématique des jeux, quant à elle, s'est développée au XX^{ème} siècle avec les travaux d'Emile Borel, puis avec ceux de Von Neumann et de Morgenstein dans les années quarante, et de Nash, dans les années cinquante. Ces derniers sont à l'origine des notions de mini/max et de recherche de point d'équilibre qui sont au fondement de ce que l'on a coutume d'appeler aujourd'hui la Théorie des Jeux.

Dès l'apparition des premiers ordinateurs, la théorie des jeux est apparue comme un formalisme approprié pour la conception de machines à jouer dans le cas de jeux à information complète et parfaite. Dans ce cadre, l'ordinateur a pour tâche de simuler les avènements prévisibles, en supposant que les adversaires jouent eux mêmes de façon totalement rationnelle, et tentent d'optimiser leurs avantages. Une telle simulation suppose le déploiement de l'ensemble des possibles, ou presque, sur un arbre de recherche. Lorsque cette simulation est rendue impossible du fait de la taille de cet arbre, des heuristiques sont introduites pour biaiser la combinatoire. La machine doit alors supposer que les concurrents adoptent des heuristiques similaires aux siennes. Ainsi, la plupart des machines qui jouent aux échecs reposent sur l'emploi de l'algorithme du mini/max et d'une fonction d'évaluation commune à tous tenant compte de l'état de l'échiquier suite à quelques demi-coups d'une exploration quasi exhaustive.

Or, en dépit des indéniables succès qu'elle a rencontrés, cette manière de faire nous semble problématique.

En effet, d'une part, les jeux à information complète et parfaite sont loin d'épuiser l'ensemble des situations de concurrence susceptibles de bénéficier de l'apport des simulations informatiques. Et, pourtant, mis à part l'exception notable du programme de jeu de poker conçu par Waterman à l'aide de systèmes de règles production, la plupart des programmes informatiques ont abordé des jeux à information complète et parfaite. À cet égard, notons que les simulations existantes de jeux comme le « master mind » ou le « scrabble » ne relèvent pas, à proprement parler, des jeux car elles ne font pas intervenir des dimensions stratégiques où jouerait la concurrence.

D'autre part, rien ne dit que lorsqu'ils sont fortement combinatoires, les jeux à information parfaite et complète bénéficient à être traités à l'aide des méthodes du min/max inspirées de la théorie des jeux. Dans ce cas, la simulation informatique est bien loin

d'épuiser l'ensemble des possibilités et, en conséquence, la rationalité supposée des adversaires se révèle être un leurre.

En somme, il serait souhaitable d'envisager d'autres types de jeux avec des méthodes différentes de celles qu'a inspiré la théorie de jeux, et ce, non seulement pour construire des programmes plus efficaces, là où les machines opèrent aujourd'hui, mais aussi, pour étendre l'empire des programmes informatiques à des situations de concurrence, comme la conception d'interfaces interactives, qui n'avaient pas vraiment été simulées auparavant.

Toutes ces considérations nous ont conduit à envisager une nouvelle perspective dans laquelle des comportements adverses sont non point calculés à l'issue d'une énumération plus ou moins longue des possibles, mais induits à partir du comportement passé des adversaires. Les progrès enregistrés ces dernières années dans la conception d'algorithmes d'apprentissage adaptatifs, à l'aide d'algorithmes génétiques, nous mettent en mesure de réaliser de telles inductions grâce auxquelles les machines se construisent une représentation interne de leurs adversaires.

Cette représentation étant construite, il reste à en tirer parti pour agir ; c'est ce qu'autorise le mécanisme d'anticipation qui est présenté dans le cœur de cet article. Plus exactement, cet article comprend trois parties. La première présente les concepts théoriques généraux d'induction et d'anticipation sur lesquels nous faisons reposer notre approche. La deuxième décrit le système SAGACE qui implémente ce modèle à l'aide algorithmes génétiques et le jeu d'ALESIA sur lequel SAGACE sera évalué. Enfin, avant la conclusion, une troisième partie fait état des expérimentations à la fois face à des joueurs humains dont on montre qu'ils sont rapidement vaincus par les capacités d'une machine qui prévoit leurs réactions, mais aussi face à différentes stratégies implantées tout spécialement. On verra, en particulier, que les comportements prévus par la théorie des jeux apparaissent comme des points fixes limites du comportement du programme SAGACE qui, en ce sens, implémente une généralisation de la théorie des jeux à des adversaires non rationnels.

2. Notre approche

2.1. Idées générales

Ces considérations nous ont amenés à tenter de concevoir un système artificiel capable de jouer efficacement à des jeux particuliers auxquels la Théorie des Jeux ne peut s'appliquer d'une façon satisfaisante.

Notre système est doté de deux capacités essentielles. D'une part, il est capable de parcourir de façon exhaustive, jusqu'à une certaine profondeur, l'ensemble des possibilités d'états futurs du jeu en fonction de ses actions et celles que puissent être celles de son adversaire (ce qui lui permet ainsi d'identifier une stratégie à court terme éventuelle lui assurant la victoire). D'autre part, il a la faculté, lorsqu'il ne trouve pas de solution lui assurant la victoire, de jouer en fonction d'une modélisation du comportement de son adversaire qu'il élabore et raffine en temps réel au cours des parties qu'il joue contre lui. Cette modélisation lui permet d'anticiper sur les actions de l'adversaire.

Le système est articulé autour d'une approche originale d'apprentissage par induction, modélisation et anticipation qui permet d'apprendre aussi bien la stratégie d'un adversaire que la façon dont celle-ci évolue : la méthode S.A.G.A.C.E.

2.2. La méthode S.A.G.A.C.E.

(Solution Algorithmique Génétique pour l'Anticipation de Comportements Evolutifs.)

Cette méthode algorithmique s'implémente sous forme d'un système de classeurs [HOL 86] dont les règles évoluent grâce à un algorithme génétique [HOL 75] & [GOL 89]. Elle repose sur une approche purement comportementale. Elle permet de prédire et d'anticiper sur le comportement d'un individu (vivant ou artificiel) mais ne prétend pas découvrir les mécanismes cognitifs ou algorithmiques sous-jacents à ce comportement [MEY 96a].

2.2.1. Architecture générale

Le système s'articule autour de plusieurs composants qui interagissent pour déterminer pour chaque coup l'action la plus judicieuse.

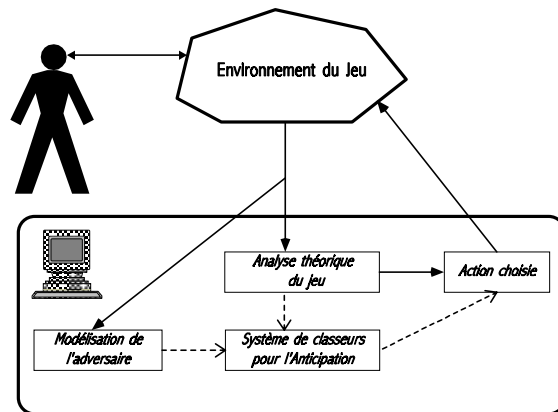


Figure 1. Architecture générale

2.2.2. Le composant d'analyse théorique

Ce composant a pour fonction de gérer les connaissances théoriques obtenues mathématiquement grâce à la Théorie des Jeux. Il permet d'analyser une situation de jeu donnée et de déterminer si celle-ci correspond à une position gagnante, c'est-à-dire s'il existe une stratégie pure garantissant la victoire. Ainsi, dans chaque situation où une stratégie gagnante est identifiée, le système l'adopte.

Si aucune solution gagnante n'est identifiée, le système tente d'utiliser son modèle de l'adversaire.

Dans l'application particulière de ALESIA, ce composant contient la bibliothèque exhaustive des coups gagnants (produite par un programme de recherche systématique).

2.2.3. Le composant d'anticipation

Le composant d'anticipation détermine à chaque prise de décision du système si la modélisation qu'il possède de son adversaire est satisfaisante ou non, et ce en fonction de certains critères liés à la précision et au taux d'erreur de prédictions passés. Si le degré de prédiction est jugé acceptable (supérieur à un taux de confiance d'anticipation constant), le

composant délivre alors une information contenant une liste des quelques actions les plus probables de l'adversaire (la LPA). En fonction de cette liste, le composant détermine alors à l'aide d'un système de classeurs quel est le coup le plus adapté.

Chaque classeur du composant d'anticipation est de la forme :

'Situation_du_jeu & prédiction_action_adversaire → action_à_effectuer'.

On notera qu'en tant que classeurs, ces règles sont génétiquement évolutives et sont donc adaptatives. La fitness de chaque classeur est ajustée d'une part après chaque déclenchement en fonction du résultat immédiat de cette application (déplacement positif ou négatif du jeton, mises comparées des deux opposants) et d'autre part en fin de partie en fonction du gain ou de la perte. [MEY 96a].

Si le degré de prédiction est jugé inacceptable (inférieur au taux de confiance), le système détermine son action en considérant que son adversaire est susceptible de jouer un quelconque coup valide. Dans ce cas les classeurs utilisés ont une partie '*prédiction_action_adversaire*' indéterminée.

2.2.4. Le composant de modélisation

Le composant de modélisation enregistre en temps réel toutes les décisions qui sont prises par l'adversaire en fonction de la configuration du jeu. La modélisation s'effectue en créant des règles de comportement de l'adversaire. Celles-ci sont ajustées en permanence pour rester aussi fidèles que possible à l'attitude réelle de l'adversaire. Ces règles sont également gérées comme des classeurs. Leur forme est la suivante :

'Situation du jeu → action_effectuée'.

Quand l'adversaire effectue, dans des circonstances semblables, une action qu'il a déjà faite, le module de modélisation augmente un coefficient de confiance lié à cette action dans cette configuration. Si l'action effectuée est différente, une nouvelle règle comportementale est créée.

Aux règles comportementales sont adjoints deux coefficients qui renseignent sur leurs taux d'utilisation et de succès vis-à-vis du joueur modélisé. Par exemple, si un joueur modélisé fait très souvent une certaine action dans un contexte précis mais que le gain correspondant à cette action est négatif, alors, le coefficient de taux d'utilisation de la règle comportementale correspondante sera élevé alors que le coefficient de succès sera faible. Par ailleurs, ces coefficients ont une pondération temporelle [MEY 96a]. Le choix récent d'une action est plus important qu'un choix effectué depuis longtemps. De la même façon, si un comportement était très performant pour le joueur modélisé mais qu'il s'est avéré inadapté récemment, le coefficient de succès dépendra plus des récents résultats que des anciens et sera certainement faible. La gestion de la pondération temporelle est paramétrable en fonction de l'adéquation entre le modèle construit d'un adversaire et son comportement réel observé.

Dans toute situation, le système peut donc émettre une requête à ce composant afin d'obtenir une certaine prédiction de ce que va jouer l'adversaire.

2.2.5. Apprentissage

Il se situe à plusieurs niveaux de l'approche. D'une part, dans la gestion des règles de modélisation où il s'agit de créer les règles de comportements nouvellement observés et d'ajuster les valeurs sélectives et de confiance des règles existantes qui sont réutilisées. Dans un système de classeurs classique, seule la fitness des classeurs est modifiée en fonction des résultats de leurs applications. Dans notre approche, un coefficient de confiance, est également adjoint à chaque classeur, est ajusté. Un classeur de modélisation est déclenché pour déterminer l'action présumée de l'adversaire en fonction de ces deux coefficients. Autant il est aisé d'ajuster le coefficient de confiance (qui ne fait que renseigner sur la fréquence d'utilisation d'une action particulière de l'adversaire dans une configuration donnée du jeu) autant ajuster la fitness est délicat. Il ne s'agit pas d'une fitness classique dans la mesure où elle doit permettre d'évaluer la valeur sélective que semble accorder l'adversaire à une action particulière. En effet, l'adversaire peut avoir une fonction d'évaluation des coups différente de celle du joueur qui le modélise. Dans cette mesure, il n'est pas raisonnable de prétendre qu'on peut ajuster fidèlement la fitness des règles qu'il utilise en utilisant une autre fonction d'évaluation. Ainsi, dans notre approche, nous n'ajustons qu'une fitness particulière qui tient compte de critères globaux (comme le gain de la partie) plutôt que de considérations locales (comme l'avancement du jeton ou les mises choisies). C'est en cela que la fitness des classeurs d'anticipation et des classeurs de modélisation diffèrent.

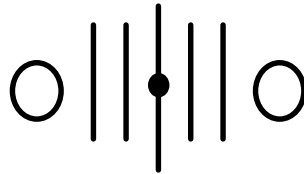
Après chaque coup, il est possible de d'évaluer les performances des deux composants. Si la prédiction de l'action de l'adversaire s'est avérée exacte, les coefficients de confiance des classeurs de modélisation correspondants est augmentée (diminuée dans le cas contraire). De plus, si l'action de l'adversaire a été bénéfique pour lui, la fitness des classeurs déclenchés qui concluaient sur cette action est augmentée (diminuée dans le cas contraire). D'autre part, dans le cas où la prédiction était exacte, la fitness des classeurs d'anticipation (qui étaient basés sur cette prédiction et qui ont été déclenchés) sont ajustées. Ainsi, l'apprentissage se situe également dans la gestion des classeurs du composant d'anticipation. Ces classeurs sont d'un type classique et l'ajustement de leurs fitness est lié aux considérations locales et globales dont nous avons parlé. L'évolution de ce système de classeur obéit à des techniques usuelles. Un algorithme génétique est déclenché après chaque partie et peut amener la création de nouveaux classeurs par croisements de classeurs existants puis mutations.

Dans les expérimentations décrites par la suite, nous n'avons pas utilisé d'algorithme génétique pour faire évoluer les classeurs du composant de modélisation. Néanmoins, nos recherches actuelles sont dirigées dans ce sens. Nous tentons en effet d'utiliser un tel algorithme pour effectuer des généralisations des classeurs de modélisation. Nous espérons ainsi pouvoir prédire le comportement de l'adversaire dans des situations nouvelles mais proches de situations déjà rencontrées.

3. Expérimentation

Afin d'évaluer notre approche, nous avons réalisé un joueur artificiel appelé Sagace qui joue contre de différents types d'adversaires à un jeu répété à information complète et imparfaite. Il s'agit du jeu ALESIA dont la description suit.

3.1. Le jeu ALESIA



Chaque joueur reçoit, au début de la partie, un capital de 50 points. Chaque coup consiste en un combat de points; Les joueurs décident simultanément et secrètement du nombre de points (au moins un) qu'ils affectent à ce coup. Le joueur qui a choisi le nombre le plus élevé bouge le marqueur (initialement au centre) vers la citadelle adverse. En cas de mises égales, le marqueur ne bouge pas. Les nombres joués sont ensuite retranchés des capitaux respectifs des joueurs. Un joueur gagne s'il parvient à déplacer le marqueur jusqu'à la citadelle adverse. Sinon, il y a partie nulle.

Ce jeu présente des caractéristiques très intéressantes. C'est un jeu dont les règles sont particulièrement simples (qu'un humain peut donc parfaitement assimiler rapidement), il est à information complète et imparfaite et il n'existe pas de stratégie générale non mixte qui garantisse la victoire (cela a été mathématiquement démontré).

3.2. Mise en œuvre

L'implémentation informatique a été réalisée en Delphi 2.0 sous Windows 95. Il est possible d'effectuer 1000 parties entre deux joueurs informatiques en 30 secondes environ sur un PC équipé d'un Pentium Pro. Cela nous a permis d'effectuer un grand nombre de séries de tests afin d'évaluer les performances du système. A chaque adversaire (humain ou simulé) du système correspond une base de modélisation et une base d'anticipation (cette base, initialement identique pour chaque adversaire, étant susceptible d'évoluer différemment).

Les classeurs contenus dans les bases de modélisation sont codés sous la forme 'Condition → action'. La partie condition étant formée d'intervalles de valeurs possibles pour : Le nombre de points des deux opposants et la situation du jeton. La partie conclusion indique la mise jouée.

Les classeurs des bases d'anticipation sont codés dans le même formalisme avec simplement une indication supplémentaire en condition codant la prédiction de l'action de l'adversaire.

Les classeurs contenus dans les bases d'anticipation ont été, pour la plupart, élaborés par un joueur humain par introspection puis codés dans le formalisme imposé par la forme des classeurs. D'autres ont été inspirés de l'observation du comportement de bons joueurs. En fait, ils ont été extraits des bases de modélisation de certains joueurs humains particulièrement performants. Nous avons déjà étudié cette idée concernant l'amorçage d'un système de classeur [MEY 96b].

4. Résultats

Nous avons défini un protocole de tests en considérant trois groupes particuliers d'adversaires pour le joueur Sagace. Cet ensemble de joueurs nous semble être suffisamment diversifié pour représenter diverses stratégies qu'un humain pourrait tenter d'adopter au jeu d'Alésia. Le joueur Sagace a donc été opposé à des humains réels puis à des joueurs probabilistes et non adaptatifs simulés et enfin à des joueurs probabilistes adaptatifs simulés (dont nous décrivons la nature exacte par la suite). Pour chacun des tests nous avons mesuré l'évolution du pourcentage de partie gagnées depuis le début du jeu par chacun des joueurs.

4.1. Adversaires humains

Nous avons fait jouer notre système contre différents joueurs humains et avons dégagé le résultat le plus représentatif de cette série de tests, chacune d'elles étant composée d'une centaine de parties (Figure 2).

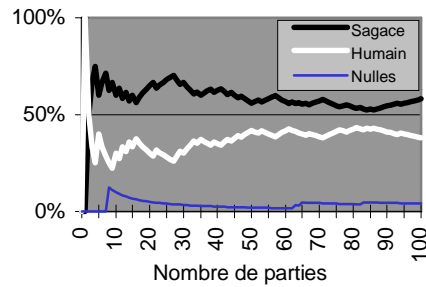


Figure 2. Adversaires humains.

Comme le montre la figure 2, les progressions du taux de parties gagnées par les joueurs n'est pas monotone. Les deux joueurs semblent s'adapter au jeu l'un de l'autre sans arrêt. Après environ 85 parties, le joueur humain ne semble plus capable d'imaginer de nouvelles stratégies adaptées au joueur Sagace et le taux de parties gagnées par celui-ci augmente légèrement. Des séries de parties contre différents humains et plus longues montrent que les courbes finissent par se stabiliser. En général autour des mêmes valeurs que celles de la Figure 2. (55% pour le joueur Sagace et 45% pour le joueur humain). Ce type de courbes est difficile à interpréter mais il souligne cependant la grande adaptabilité du jeu humain face à celui de la machine et réciproquement. Il montre également que l'intensité des variations diminue avec le nombre de parties.

4.2. Adversaires probabilistes

Dans un second type d'expérimentations, nous avons fait jouer notre système contre des joueurs simulés ayant des stratégies probabilistes. Ces stratégies faisant intervenir le hasard viennent naturellement à l'esprit d'un humain cherchant à contrecarrer la modélisation dont

il sait qu'il pourrait faire l'objet. Le premier de ces adversaires est un joueur probabiliste jouant aléatoirement un nombre compris entre un et le nombre de points qu'il lui reste.

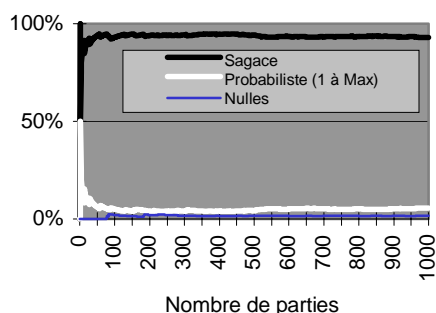


Figure 3. Adversaire probabiliste (1 à Max).

Très rapidement, le joueur Sagace parvient à générer par son analyse une modélisation très fiable de cet adversaire et s'y adapte parfaitement (cf. figure 3).

Le second adversaire simulé est également probabiliste mais il choisit une mise toujours comprise entre 1 et 50. Si la mise calculée est supérieure au nombre de points qu'il lui reste, il joue tous ses points.

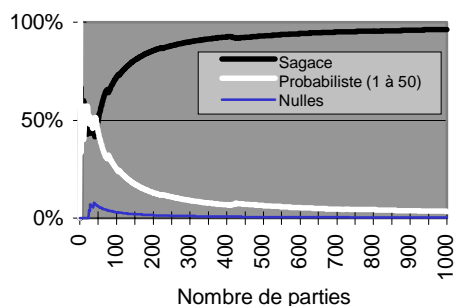


Figure 4. Adversaire probabiliste (1 à 50).

Ce joueur demande un temps plus long que le joueur précédent pour être modélisé mais obtient finalement de moins bons résultats (c.f. Figure 4). La modélisation est plus délicate (la convergence est moins rapide) parce qu'il n'y a pas équiprobabilité dans les choix d'action. Les coups les plus hauts étant naturellement favorisés. Le modèle élaboré par le joueur Sagace est néanmoins très fiable après une cinquantaine de parties.

Le troisième adversaire simule un théoricien expert. Il est en effet capable de repérer mathématiquement quand une situation permet une stratégie qui assure la victoire et la joue dans ce cas. Sinon, son comportement est déterminé par un ensemble de règles probabilistes introduites d'après une expertise humaine du jeu (par exemple 'ne jamais jouer plus de un point de plus que ce qu'il reste à l'adversaire').

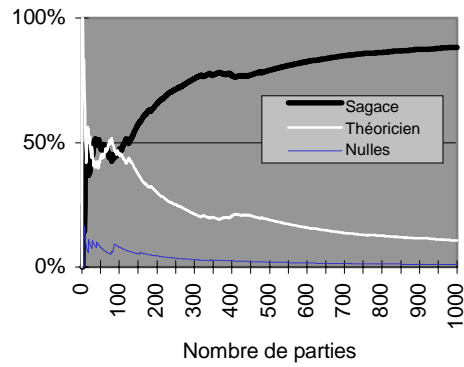


Figure 5. *Adversaire théoricien.*

Au moins 150 parties sont nécessaires au joueur Sagace avant d'avoir un modèle fiable de ce joueur (cf. figure 5). Avant, il y a une période pendant laquelle les joueurs se valent. Ensuite, pendant les 150 parties suivantes, le joueur Sagace se révèle très performant puis, le pourcentage de parties gagnées finit se stabiliser. En continuant la série de parties, ce pourcentage atteint la valeur limite 85%.

4.3. Adversaires adaptatifs

Dans un troisième type d'expérimentations, nous avons fait jouer notre système contre des joueurs simulés ayant des stratégies adaptatives. Ces stratégies viennent naturellement à l'esprit d'un humain cherchant à contrecarrer la modélisation dont il sait qu'il pourrait faire l'objet en modifiant dynamiquement ses stratégies. L'adversaire simulé est un joueur qui va faire évoluer ses stratégies en fonction de leur performances. Ce joueur a été implémenté sous forme d'un système de classeurs. Ce joueur adaptatif apprend de façon classique par rétribution positive et négative de ses règles de choix d'action en fonction de leurs performances [HOL 86].

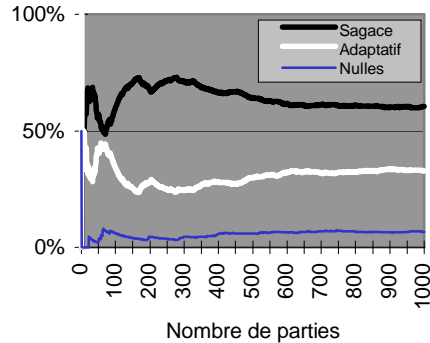


Figure 6. *Adversaire Adaptatif.*

Durant les cinquante premières parties, le joueur Sagace parvient relativement bien à modéliser le joueur adaptatif et son pourcentage de gain augmente (cf. figure 6). Pendant les cinquante parties suivantes, le joueur adaptatif s'adapte (!) puis, le joueur Sagace parvient à déterminer un nouveau modèle cohérent et son pourcentage de gain ré-augmente. Après un certain nombre de ces oscillations 'adaptatives', le pourcentage de gain des deux joueurs se stabilise (60% pour le joueur Sagace, 33% pour le joueur adaptatif). On remarquera que ce type de joueur se révèle moins performant que les bons joueurs humains.

5. Discussion

L'architecture de notre système permet de s'affranchir des limites de prudence imposées par la Théorie des Jeux tout en profitant de ses résultats mathématiques quand ils sont maximaux. Pour autant, notre méthode n'apporte théoriquement aucune garantie de résultat, contrairement à la Théorie des Jeux. Il s'agit d'une appréhension différente de l'inéquation entre la sécurité, le risque et les performances.

Contre tout adversaire à stratégie figée, cette approche est redoutable parce qu'elle permet de retrouver ses règles comportementales et permet alors de choisir les stratégies adaptées.

Contre un adversaire adaptatif, les performances du système dépendent naturellement du degré d'adaptabilité de cet adversaire. L'humain est, a priori, un adversaire très adaptatif. Pourtant, les joueurs humains sont très souvent battus par le système. Quand un joueur humain est informé du type d'apprentissage utilisé par le système il tente souvent d'adopter une stratégie très différente de sa stratégie d'origine. Soit il tente de se rendre imprévisible en jouant plus ou moins aléatoirement (mais cette tentative lui fait très souvent jouer des coups sous-optimaux et le fait perdre) ou en tentant de tromper le système en jouant plusieurs fois un coup particulier et en changeant dès qu'il pense que le système va jouer en fonction de cette information volontairement donnée. En fait, ces joueurs humains modélisent le système qui tente de les modéliser. Cette stratégie humaine demande une grande perspicacité et une grande concentration mais peut s'avérer très performante.

6. Conclusions et perspectives

Cette approche des jeux répétés à information complète et imparfaite ne remet pas en cause des résultats mathématiquement validés tout en offrant une solution pour appréhender des situations complexes dans lesquelles il est important sinon impératif d'avoir un modèle fiable des individus ou agents avec lesquels on est en interaction.

Nous pensons améliorer le système en le dotant de moyens de s'auto-modéliser afin de lui permettre d'estimer quand il deviendra lui-même prédictible et de changer de comportement en fonction. Egalement, cela permettra d'analyser plus rigoureusement l'apprentissage effectué par le système.

Par ailleurs, nous comptons étudier de quelle manière il serait envisageable d'ajuster les coefficients probabilistes d'une stratégie mixte optimale déterminée grâce à la Théorie des Jeux en fonction de la modélisation du comportement de ses adversaires effectuée grâce à S.A.G.A.C.E.

Enfin, nous tenterons de mettre en application certaines de nos idées concernant la généralisation de règles génétiques dans les classeurs afin d'élargir les prédictions des actions d'un adversaire à des situations non réellement observées mais similaires.

Références

- [AUM 87] Aumann R. (1987). « Game theory » *The new Palgrave* : a Dictionary of Economics. The MacMillan Press.
- [AXE 92] Axelrod R. (1992). *Donnant, donnant Théorie du comportement coopératif*, Edition Odile Jacob.
- [DAM 87] Van Damme E. (1987). *Stability and Perfection of Nash Equilibrium*, Springer Verlag, Berlin.
- [DEL 92] Delahaye J.P. (1992) « L'altruisme recomposé ? » *Pour la science*, N° 181.
- DuncanLuce R. et Raiffa H. (1957) *Games and decision*. John Willey and Sons, Inc.
- [FIN 71] Findler N. ; Klein H. ; Gould W. ; Kowal A. et Mening J. (1971) *Studies on Decision Making Using the Game of Poker in Information Processing*. Proceedings of IFIP Congress 71 Vol 2.
- [FUD 91] Fudenberg D. et Tirole J. (1991). *Game theory*, The MIT Press.
- [GOL 89] Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine learning*. Addison-Wesley.
- [HAR 67] Harsanyi J.C. (1967) « Games with incomplete information played by 'bayesian' players ». *Management science* 14.
- [HOL 75] Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press.
- [HOL 86] Holland, J.H ; Holyoak, K.J. Thagard, P.R. and Nisbet, R.E. (1986). *Induction : Processes of Inference, Learning, and Discovery*. MIT Press, Cambridge
- [KRE 91] Kreps D. (1991) *Games theory and Economic Modelling*, Oxford University Press.
- [MAY 82] Maynard Smith J. (1982) *Evolution and theory of Games*, Cambridge University Press.
- [MEY 96a] Meyer C. et Ganascia J.G. (1996a) *S.A.G.A.C.E. Solution Algorithmique Génétique pour l'Anticipation de Comportements Evolutifs*. Paris VI, LAFORIA N°96/32.
- [MEY 96b] Meyer C. et Ganascia J.G. (1996b) *Utilization of imitation and anticipation mechanisms to bootstrap an evolutive distributed artificial intelligence system*. Animal Societies as an Alternative Metaphorical Basis for DAI - Workshop ICMAS 1996.
- [MOU 91] Moulin R.B. (1991) *Théorie des Jeux pour l'économie et la politique*, Hermann.
- [MYE 91] Myerson R.B. (1991) *Game Theory Analysis of Conflict*, Harvard University Press.
- [NEU 44] Neumann J. von et O. Morgenstern (1944). *Theory of Games and Economic Behavior*, Princeton University Press.