« Mind-Reading Machines »
Modélisation des adversaires et anticipation dans les jeux à information complète et imparfaite.

Christophe Meyer (*, **) et Jean-Daniel Zucker (*)

(*) LIP6 - CNRS, Pôle IA Université Pierre et Marie Curie. Tour 46-0, 4 Place Jussieu. 75252 Paris Cedex. {Christophe.Meyer, Jean-Daniel.Zucker}@lip6.fr

(**) Mathématiques Appliquées S.A. 24, bd de l'Hôpital. 75005 Paris, France. Christophe.Meyer@animaths.com

Résumé

Il y a presque cinquante ans, Shannon a proposé une machine entièrement mécanique capable de jouer avec un succès certain contre des joueurs humains au jeu itéré « Matching penny ». Cette machine était capable de « lire dans l'esprit » de ses adversaires humains. Il la nomma : Mind-reading machine. Cet article expose plusieurs approches pour modéliser et anticiper le comportement d'un joueur humain dans des jeux à information complète et imparfaite (dans lesquels les joueurs jouent simultanément comme « Matching penny » ou « Pierre ; Ciseaux ; Papier »). En théorie, pour ces jeux, les stratégies optimales sont des répartitions probabilistes des stratégies pures (Equilibres de Nash). La détermination mathématique de ces répartitions prend pour hypothèse que tous les joueurs sont parfaitement rationnels et qu'ils sont capables de jouer aléatoirement. Nous verrons comment des « Mind-reading machines » peuvent exploiter les faiblesses des joueurs humains afin d'obtenir des performances supérieures à celles de la théorie. La méthode S.A.G.A.C.E. que nous avons développée se révèle particulièrement efficace dans ce but.

1 Introduction

Dans les années 50, au sein des laboratoires Bell, David Hagelberg a développé la première « Mind-reading machine » mécanique. Il fut suivi avec brio par Shannon qui finalisa l'approche (Shannon, 1953). Sa machine physique jouait contre des joueurs humains à un jeu répété simple à information complète et imparfaite : « Matching penny » en lisant, à sa façon, dans l'esprit de ses adversaires.

Les règles de ce jeu sont extrêmement simples et peuvent être comprises par un humain très facilement : un joueur choisit secrètement entre « pair » et « impair » et l'autre doit deviner son choix. S'il y parvient, il gagne un point sinon, son adversaire gagne un point. Le jeu est répété, le vainqueur est celui qui a marqué le plus de points à la fin d'une série de parties.

Il y a de nombreux points communs entre ce jeu et le « Dilemme itéré des prisonniers » (Rapoport, 1965). Ces deux jeux sont ce que l'on nomme des jeux à information complète et imparfaite. *Complète* parce que les joueurs connaissent parfaitement les règles du jeu : actions possibles et leurs conséquences et que ces règles n'incluent aucune composante aléatoire. *Imparfaite* parce qu'une information du jeu manque quand chaque joueur prend sa décision : le choix de son adversaire.

Pour ces jeux, les stratégies optimales déterminées par la théorie sont des répartitions probabilistes sur les stratégies pures possibles (à « Matching penny », la stratégie optimale consiste à jouer aléatoirement « pair » ou « impair » de façon uniforme). Ces stratégies optimales supposent une parfaite rationalité de l'ensemble des joueurs et garantissent l'espérance de gain (faire, en moyenne, match nul dans un jeu équitable comme « Matching penny »). Pourtant, un joueur avisé peut espérer gagner bien plus que l'espérance de gain s'il se révèle capable d'exploiter les faiblesses de ses adversaires.

La machine de Shannon obtenait un score supérieur à l'espérance de gain contre des joueurs humains. Elle était même meilleure que des joueurs humains informés de la nature uniformément aléatoire de la stratégie optimale.

Nous considérons comme une explication très plausible de cela le fait qu'il soit presque impossible à un humain de générer consciemment quoi que ce soit de véritablement aléatoire.

Quand un joueur humain joue à un tel jeu, il essaie systématiquement d'*anticiper* les choix de son adversaire, soit en tentant de se mettre à sa place, soit en tentant de découvrir des régularités dans la suite de ses stratégies pour les exploiter.

Cet article compare quatre types de « Mind-reading machine » virtuelles (des programmes) sur un plan théorique et expérimental dans des parties les opposant les unes aux autres et contre des joueurs humains. Les résultats donnent des indications sur les raisons possibles du succès de ces machines.

L'approche S.A.G.A.C.E. que nous avons développée (Meyer et Ganascia, 1996), (Meyer et al., 1997) se révèle être la plus efficace des méthodes comparées. Elle repose sur un modèle explicite du comportement des adversaires et surtout sur un modèle de la façon dont il évolue au cours des parties. Ce modèle prend en considération les choix passés de l'adversaire, les limites supposées de sa mémoire (quand il s'agit d'un adversaire humain), la façon dont il réagit à un succès ou un échec.

L'architecture de la méthode S.A.G.A.C.E. se compose de deux systèmes de classeurs (Holland, 1986), (Wilson, 1989), dont les règles s'adaptent par apprentissage et par évolution génétique (Holland, 1975), (Goldberg, 1989). Le premier système construit le modèle de l'adversaire, le second détermine la stratégie en exploitant (ou pas) les informations fournies par le premier.

Les résultats montrent les avantages d'une telle approche d'apprentissage adaptative prenant en compte explicitement les faiblesses potentielles des joueurs humains (mémoire limitée, irrationalité , etc.) sur des approches théoriques (Théorie des jeux) ou traditionnelles en Intelligence Artificielle (automates à états finis ou anticipation par reconnaissance de patterns par exemple).

Les seules machines véritablement efficaces à de tels jeux construisent des modèles de leurs adversaires : certaines (comme la machine de Minasi) se contentent d'analyser le comportement passé de leurs adversaires pour y repérer des séquences répétées afin de les exploiter. D'autres (comme la machine de Shannon ou S.A.G.A.C.E.) prennent également en compte l'influence présumée des résultats des coups joués sur leurs adversaires (comment ils réagissent à un succès, à un échec).

La modélisation dans les jeux à information complète et parfaite est un terrain de recherche qui suscite actuellement de vifs intérêts scientifiques (Carmel et Markovitch, 1996), (Meyer, et al., 1997), (Stone et Veloso, 1996).

L'analyse de la modélisation des adversaires dans le contexte du jeu « Matching penny » conduit à un grand nombre de questions fondamentales sur la nature de l'anticipation des comportements humains dans les jeux, voire dans toute autre activité humaine.

Pour construire des « Mind-reading machines » efficaces, il nous semble indispensable de considérer les aspects particuliers des choix des humains : leur mémoire (sa structure, ses limites, sa pérennité), leur irrationalité et, dans un contexte de jeu, leur incapacité à manipuler correctement le hasard (Delahaye, 1995).

La section 2. de cet article présente le problème de l'anticipation dans les « Mind-reading machines » (désormais notées : MRM). La section 3. compare différentes MRM. La section 4. présente certaines expérimentations et les résultats correspondants.

2. Mind-Reading Machines (MRM)

2.1. Problème de l'anticipation dans les MRM

Afin de comparer différentes stratégies pour la construction de MRM, nous proposons d'utiliser différents critères. Un premier critère concerne le fait que la MRM construise, ou pas, un modèle de son adversaire. Nous appellerons « véritables MRM » celles qui gèrent un tel modèle et « fausses MRM » les autres. Pour simplifier nos propos, nous considérerons que la MRM est toujours le joueur A et l'adversaire le joueur B.

Une notion fondamentale pour un jeu tel que « Matching penny » est celle de pattern. Un pattern correspond à une séquence de n-uplets de la forme (a_i, b_i …, n_i) où les x_i sont les actions effectuées par le joueur x au $i^{ème}$ coup. Pour les jeux à deux joueurs, il s'agit donc de couples de la forme (a_i, b_i). Pour « Matching penny », si $a_i = b_i$ alors A a gagné le coup.

Cette notion de pattern permet de décrire des stratégies simples telles que :

$$A:\ \textit{SI}\ (\ (a_i,b_i)\ /\ a_i \neq b_i)\ \textit{ALORS}\ a_{i+1} \leftarrow -a_i$$

Qui signifie pour le joueur A : « changer de stratégie ($-a_i$ est la stratégie inverse de a_i) si le dernier coup a été perdant » (jouer « pair » si on avait joué « impair » et réciproquement).

Cette stratégie ne modélise pas son adversaire mais se base uniquement sur le précédent coup (et son résultat).

Le tableau 1 expose les autres critères que nous avons retenus pour décrire et comparer les différentes MRM que nous avons analysées pour le jeu « Matching penny ».

Stratégie	▪ implicite (le modèle de l'adversaire n'est pas explicite) ▪ adaptation (par apprentissage) ou pas ▪ utilisation du hasard
Résultats des coups	▪ pas pris en compte, implicites ou explicites
Mémoire des coups	▪ taille, pérennité de la mémoire
Longueur des Patterns	▪ combien de coups sont concernés
Taille de la mémoire	▪ nombre d'éléments conservés
Contenu d'un élément de mémoire	▪ type de données stockées dans la mémoire

Table 1: Critères de comparaisons de MRM.

2.2 Exemples de véritables MRM

Cette section présente un ensemble de véritables MRM suivant les critères précités. En tant que MRM, ces machines prennent en compte de façon explicite ou implicite le comportement de leurs adversaires.

Nous appellerons CS (Coup Suivant) le prédicat signifiant, sous la forme CS(B,M), que le prochain coup de B est M (« pair » ou « impair »).

2.2.1. TitAfterTit

Il s'agit d'une allusion humoristique à la célèbre stratégie 'TitForTat' utilisée au jeu du « Dilemme des prisonniers » (Axelrod, 1984).

Cette stratégie consiste à toujours jouer (après le premier coup) comme l'adversaire au coup précédent quel qu'ait pu être le résultat.

Dans cette stratégie, la prédiction du coup de l'adversaire est qu'il reproduit toujours son dernier coup sans tenir compte de ses conséquences.

Cette prédiction peut s'exprimer de la façon suivante :

SI (*,M) *ALORS* CS(B,M)

Les propriétés de la stratégie TitAfterTit peuvent être résumées dans le tableau suivant :

Stratégie	implicite, non adaptative, pas d'aléatoire
Résultats des coups	non
Mémoire des coups	1 (seulement le dernier coup est mémorisé)
Longueur des Patterns	1
Taille de la mémoire	1 (le dernier coup de l' adversaire)
Contenu d'un élément de mémoire	un coup (« pair » ou « impair »)

Tableau 2. TitAfterTit

2.2.2. MRM de Shannon (Shannon, 1953).

Le type de patterns considérés par « Shannon » concerne le succès (ou l'échec) de deux coups successifs et l'attitude de l'adversaire en fonction de ces deux coups (a-t-il persévéré ou changé de stratégie).

Il existe donc huit situations possibles :

- Le joueur gagne, rejoue la même chose et gagne à nouveau. Il peut alors jouer la même chose ou changer.
- Le joueur gagne, rejoue la même chose et perd. Il peut alors jouer la même chose ou changer.
- Le joueur gagne, joue différemment et gagne à nouveau. Il peut alors jouer la même chose ou changer.
- Le joueur gagne, joue différemment et perd. Il peut alors jouer la même chose ou changer.
- Le joueur perd, rejoue la même chose et gagne. Il peut alors jouer la même chose ou changer.
- Le joueur perd, rejoue la même chose et perd à nouveau. Il peut alors jouer la même chose ou changer.
- Le joueur perd, joue différemment et gagne. Il peut alors jouer la même chose ou changer.
- Le joueur perd, joue différemment et perd à nouveau. Il peut alors jouer la même chose ou changer.

Ces huit situations correspondent aux huit mémoires de « Shannon » qui contiennent uniquement deux informations :

La première indiquant si le joueur a joué la même chose ou changé sa stratégie la dernière fois que chacune de ces situations s'est présentée.

La seconde indiquant si cette attitude a été répétée à l'issue de la dernière situation antérieure identique.

« Shannon » joue en fonction du contenu des huit mémoires. Si la seconde information indique que l'attitude de l'adversaire a été répétée, « Shannon » considère qu'il va de nouveau répéter son attitude et joue le coup correspondant. Si l'information indique que l'attitude de l'adversaire n'a pas été répétée, « Shannon » joue au hasard (uniformément « pair » ou « impair »). Si une prédiction s'avère non fondée, la mémoire correspondante est simplement effacée même si elle avait été juste de nombreuses fois précédemment.

Les propriétés de la stratégie « Shannon » peuvent être résumées dans le tableau suivant :

Stratégie	implicite, non adaptative, utilisation de l'aléatoire
Résultats des coups	implicite
Mémoire des coups	totale
Longueur des Patterns	2
Taille de la mémoire	**8**
Contenu d'un élément de mémoire	3 (conséquence du coup, attitude au coup suivant, répétition)

Tableau 3. « Shannon »

2.2.3. **Minasi** (Minasi, 1991)

« Minasi » est une MRM qui a été développée, à l'origine pour le célèbre jeu « Pierre / Ciseaux / Papier ». Nous l'avons adaptée à « Matching penny ».

L'algorithme correspondant est extrêmement simple. Il repose sur une analyse des patterns composés par les coups successifs des deux joueurs. Dans cet algorithme, un pattern est une séquence de paires de coups (un coup par joueur). Pendant une partie, l'ensemble des paires de coups est enregistré en continu :

Ex : R s P r P r S r P s P r R r.

→ Avec R pour Pierre (Rock) ; S pour Ciseaux (Scissors) et P pour Papier (Paper).

Les coups en majuscules correspondent aux coups du joueur « Minasi », les coups en minuscules à son adversaire.

Quand le système doit effectuer son choix pour le coup en cours, il recherche dans l'enregistrement la plus longue série de coups (en partant des derniers coups) répétée au moins une fois qui correspond à la situation actuelle. Cette série est utilisée pour déterminer le coup à jouer.

Exemple :

Si l'enregistrement est le suivant :

P r P r R p R s P r P p S r R r P r S r R p S s R s R r P r

En partant de la fin de la séquence (les coups les plus récents) l'algorithme effectue la recherche de la façon suivante :

Coups suivant / Séquence	**r**	**p**	**s**	
r	5	3	0	❶
P r	2	2	0	❷
r P r	1	1	0	❸
R r P r	1	0	0	❹

La ligne ❶ correspond à la recherche du pattern 'r' : Il a été localisé 8 fois dans l'enregistrement (la dernière, n'ayant pas de successeur, ne compte pas). 5 fois il a été suivi d'un 'r' et 3 fois d'un 'p'.

La ligne ❷ correspond à la recherche du pattern 'P r' : Il a été localisé 4 fois dans l'enregistrement. 2 fois il a été suivi d'un 'r' et 2 fois d'un 'p'.

La ligne ❸ correspond à la recherche du pattern 'r P r' : Il a été localisé 2 fois dans l'enregistrement. 1 fois il a été suivi d'un 'r' et 1 fois d'un 'p'.

La ligne ❹ correspond à la recherche du pattern 'R r P r' : Il a été localisé 1 seule fois dans l'enregistrement et a été suivi d'un 'r'.

Comme le pattern 'R r P r' n'a été localisé qu'une fois, il est impossible de trouver une séquence répétée plus grande qui soit suivie éventuellement d'un autre coup. La recherche s'interrompt et statue sur le fait que l'adversaire va jouer 'r'. L'algorithme de « Minasi » va donc choisir de jouer 'P'. Si deux séries répétées statuent sur deux coups différents, l'algorithme indique de jouer aléatoirement.

L'algorithme de « Minasi » suppose que ses adversaires réitèrent un choix quand les circonstances sont approximativement les mêmes.

« Minasi » ne tient aucun compte des conséquences des choix effectués par les deux joueurs. Il n'adapte pas sa stratégie en fonction des résultats qu'elle obtient et ne considère pas que l'adversaire puisse le faire pour lui-même.

« Minasi » possède, en quelque sorte, un modèle de son adversaire, consistant uniquement en l'enregistrement des coups. Contre un adversaire déterministe (objet de plusieurs séries d'expérimentations que nous avons faites) ce modèle est idéal et « Minasi » offre d'excellents résultats. Etrangement, ce modèle est suffisant pour vaincre (en moyenne) presque systématiquement un humain sur une série de 100 parties.

Les propriétés de la stratégie « Minasi » peuvent être résumées dans le tableau suivant :

Stratégie	simple, non adaptative, utilisation de l'aléatoire
Résultats des coups	implicite
Mémoire des coups	totale
Longueur des Patterns	De 1 au nombre de coups joués.
Taille de la mémoire	**Le nombre de coups joués**
Contenu d'un élément de mémoire	un coup

Tableau 4. Minasi

2.2.4. S.A.G.A.C.E. (simplifiée) (Meyer et al., 1997)

Le principe de la version de S.A.G.A.C.E. utilisée dans ce contexte est similaire à « Minasi » avec une considération supplémentaire pour chaque coup joué : a-t-il conduit à un succès ou à un échec pour l'adversaire.

L'idée est, comme le fait « Shannon », de modéliser la façon dont l'adversaire réagit à une victoire ou une défaite.

De plus, la mémoire associée à chaque pattern prend en considération l'ancienneté de chaque coup. Un coup récent a ainsi plus d'importance qu'un coup plus ancien. Il s'agit là d'une autre différence fondamentale avec « Minasi ». Elle permet d'exploiter la faiblesse de la mémoire humaine (Stern, 1985).

Les propriétés de la stratégie S.A.G.A.C.E. peuvent être résumées dans le tableau suivant :

Stratégie	simple, **adaptative**, utilisation de l'aléatoire
Résultats des coups	**explicite**
Mémoire des coups	Totale
Longueur des Patterns	De 1 au nombre de coups joués.
Taille de la mémoire	Le nombre de coups joués
Contenu d'un élément de mémoire	**Un coup et son ancienneté**

Tableau 5. S.A.G.A.C.E.

2.2.5. Humains

Si les machines tentent de « lire dans l'esprit » des joueurs humains, l'inverse est également vrai. Les stratégies de modélisation humaines sont en partie inconnues mais les expérimentations que nous avons conduites nous incitent à penser qu'un joueur humain prend en compte des critères similaires à ceux qui définissent S.A.G.A.C.E. : les coups joués, leurs éventuels succès, la conscience d'être soit même modélisé, etc. Quand un joueur humain veut tromper son adversaire, il tente de jouer au hasard (au moins en partie). Ces tentatives sont souvent infructueuses parce que les humains s'avèrent incapables de générer un aléatoire correct (des séquences apparaissent naturellement dans les choix censés être aléatoires).

Le tableau suivant tente de décrire les caractéristiques des stratégies de modélisation de joueurs humains à « Matching penny ».

Stratégie	complexe, **adaptative**, Pseudo-aléatoire
Résultats des coups	Variable[1]
Mémoire des coups	Variable (en moyenne les 3 ou 4 derniers)
Longueur des Patterns	? (mais semble-t-il très court)
Taille de la mémoire	variable (mais limitée)
Contenu d'un élément de mémoire	?

Tableau 6. Humains

2.3. Exemples de fausses MRM

Une machine utilisant une stratégie purement aléatoire est un exemple de fausses MRM. Ces stratégies consistent à jouer aléatoirement « pair » ou « impair » suivant une répartition probabiliste uniforme ou pas (avec un poids p associé à « pair » et (1-p) associé à « impair »). Certaines stratégies utilisant des automates à états finis sont également de fausses MRM. Par exemple, une telle stratégie ne prenant en compte que ses propres coups passés appartient à cette catégorie : (I : « impair », P : « pair ») II→P, PP→I, IP→P, PI->I. (conduit au cycle IIPP si elle est initialisée avec II). Cette stratégie ignore les coups de l'adversaire et les conséquences des coups des deux joueurs.

Ces fausses MRM perdent systématiquement contre les véritables si elles sont déterministes. Quand elles sont uniformément aléatoires, elles obtiennent naturellement l'espérance de gain (elles perdent, en moyenne, aussi souvent qu'elles gagnent). Si elles obéissent à des répartitions non-uniformes, elles perdent contre les véritables MRM qui sont capables d'exploiter cette dissymétrie.

Nous n'avons pas reproduit les expériences correspondantes dans cet article parce qu'elles présentent peu d'intérêt.

3. Experimentation

Toutes les véritables MRM que nous avons décrites ont été historiquement développées pour affronter des joueurs humains. Une méthode logique pour comparer ces machines est de les confronter chacune à un ensemble de joueurs humains puis de comparer leurs performances respectives afin de les classer.

[1] Parfois le joueur prend le temps de la réflexion et quand on lui demande à quoi il pense, il dit qu'il essaie de deviner les conséquences des coups précédents pour l'adversaire.

Cependant, à l'heure où les humains inter-agissent de plus en plus souvent et naturellement avec des systèmes artificiellement intelligents, il devient important sinon impératif d'être capable de « lire dans des esprits artificiels ». Pour cette raison, nous avons également confronté les MRM entre elles comme Axelrod l'avait fait pour le « Dilemme des prisonniers ». (Axelrod, 1984), (Rapoport, 1965). Pour les expérimentations, nous avons implémenté les MRM décrites et effectué un tournoi sur des séries de 10 parties de 500 coups en incluant des joueurs humains. Nous avons conservé les courbes les plus significatives. Pour chaque série de parties, nous avons enregistré le pourcentage de prédictions exactes lorsque l'adversaire n'a pas joué aléatoirement et le pourcentage de coups aléatoires.

Le tableau suivant indique les résultats obtenus.

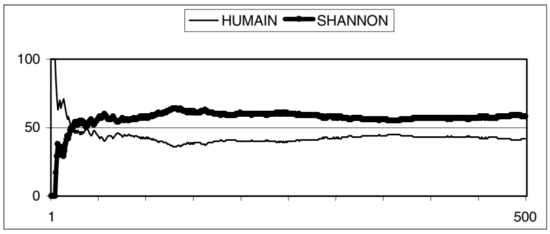
	Prédictions correctes	Coups aléatoires	Parties gagnées
SHANNON	60%	65%	58%
HUMAN	?	?	42%
MINASI	57%	**10**%	56%
HUMAN	?	?	44%
SAGACE	**66**%	20%	66%
HUMAN	?	?	34%
SAGACE	**57**%	**17**%	54%
SHANNON	40%	75%	46%
SAGACE	**57**%	17%	57%
MINASI	42%	**6**%	43%
MINASI	**54**%	**7**%	51%
SHANNON	45%	76%	49%

Tableau 7. Résultats du tournoi

Ce tableau montre qu'un joueur humain est systématiquement battu par les MRM que nous avons implémentées. Les courbes présentées dans les sections suivantes donnent des détails plus précis sur la dynamique des confrontations.

On peut également constater, à la lecture du tableau, que S.A.G.A.C.E. est la machine la plus performante et qu'elle obtient le meilleur taux de prédictions exactes. « Shannon » est la MRM qui a la plus recours au hasard et « Minasi » la moins (elle ne joue que par prédictions, fussent-elles très souvent fausses).

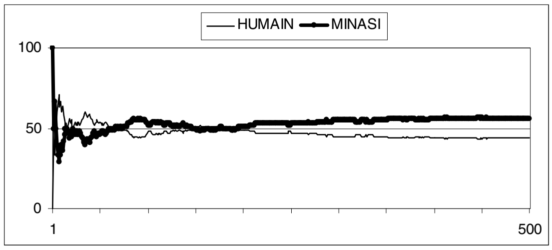
SAGACE apparaît comme la meilleure MRM bien qu'elle ait parfois recours au hasard (en fait, à chaque fois que ses prédictions sont contradictoires ou pas assez tranchées).



Courbe 1. Humain / « Shannon »

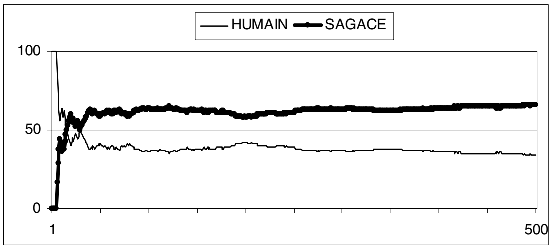
« Shannon » exige au moins 40 parties pour obtenir un modèle cohérent de ses adversaires humains. Ensuite, le pourcentage de parties gagnées par « Shannon » se stabilise autour de 58% (+/- 3% suivant les parties) (courbe 1).

Notons que « Shannon » effectue 65% de ses coups au hasard (voir tableau 7).



Courbe 2. Humain / « Minasi »

La cohérence du modèle est plus longue à obtenir pour « Minasi ». 200 parties sont nécessaires pour cela. « Minasi » s'adapte correctement à un joueur humain mais très lentement (courbe 2).

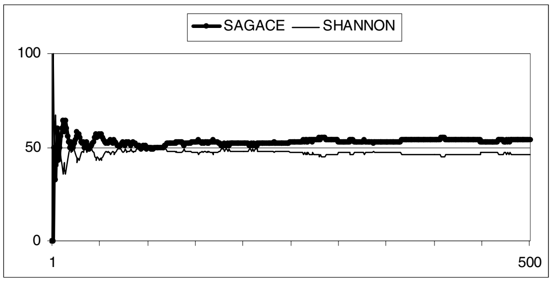


Courbe 3. Humain / S.A.G.A.C.E.

Contre un joueur humain, S.A.G.A.C.E. s'adapte très rapidement en atteignant son score moyen final de 66% (+/- 4%) au bout d'une trentaine de parties (courbe 3).

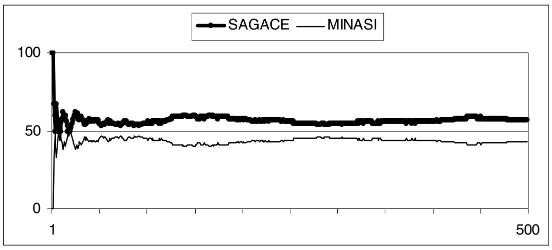
Contre un joueur humain, S.A.G.A.C.E. semble être la plus appropriée des trois approches décrites. Nous pensons que cela est dû au fait qu'elle prend explicitement en compte les caractéristiques intuitives des stratégies humaines : irrationalité, mémoire limitée, adaptation (réaction aux succès et aux échecs).

« Minasi » ne tient pas compte de la mémoire limitée des humains ni de leur adaptation potentielle, c'est, à notre sens, ce qui explique ses performances moindres.



Courbe 4. S.A.G.A.C.E. / « Shannon »

Comme on peut le voir sur la courbe 4, lorsque S.A.G.A.C.E. et « Shannon » sont confrontées l'une à l'autre, la progression du pourcentage de parties gagnées n'est pas monotone. Les deux adversaires semblent s'adapter continuellement l'un à l'autre. Après 100 parties, « Shannon » choisit de plus en plus souvent aléatoirement (cf. tableau 7). Il devient incapable de modéliser S.A.G.A.C.E. (40% de prédictions correctes). Evidemment, comme « Shannon » joue très souvent aléatoirement, S.A.G.A.C.E. ne peut qu'anticiper rarement.



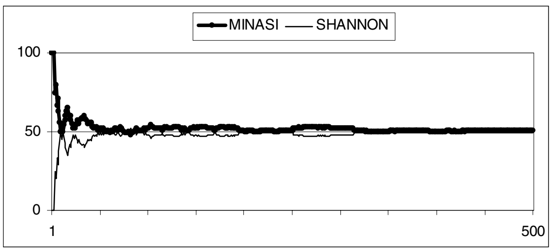
Courbe 5. S.A.G.A.C.E. / « Minasi »

La courbe 5 montre à quel point il est difficile de modéliser S.A.G.A.C.E. « Minasi » recourt au hasard seulement 6% du temps et ne parvient toutefois pas à obtenir un bon score contre S.A.G.A.C.E. (seulement 42% de parties gagnées) (cf. tableau 7).

S.A.G.A.C.E. est très efficace contre « Minasi » (57% de parties gagnées).

Courbe 6. « Shannon » / « Minasi »

La courbe 6 illustre le problème de



l'incompréhension ! Les stratégies « Shannon » et « Minasi » ont des façons si différentes de s'adapter qu'elles se révèlent incapables de se modéliser l'une l'autre. Après 200 parties, « Shannon » 'renonce' à modéliser « Minasi » et devient uniformément aléatoire. Cela entraîne naturellement que « Minasi » soit totalement incapable de deviner ses coups et joue tout aussi aléatoirement. En ce sens, elles convergent simultanément vers la solution optimale au sens de Nash : une répartition aléatoire uniforme de choix « pair » ou « impair ». Cette convergence simultanée s'apparente tout à fait à la méthode du « fictitious play » de Brown [Brown, 1951].

4. Conclusion

Cet article a analysé différentes approches pour l'anticipation de joueurs humains dans des jeux à information complète et imparfaite et a montré que les « Mind-reading machines » tiraient avantage de la difficulté qu'ont les humains à être véritablement aléatoires et de leur manque de rationalité (liée à leur faible mémoire, leur indétermination, leur forme d'adaptation).

Les résultats des séries d'expérimentations que nous avons conduites et commentées dans cet article donnent une explication possible au succès de ces machines. Ils montrent également les avantages de notre approche (S.A.G.A.C.E.) contre des joueurs humains ou d'autres « Mind-reading machines » traditionnelles.

Références

[Axelrod, 1984] Robert Axelrod *The evolution of cooperation*. Basic Books, NY, 1984.

[Brown, 1951] G.W. Iteratives Solutions of Games by Fictitious Play. *In Brown Activity analysis of Production and Allocation*, NY, Wiley. 1951.

[Carmel and Markovitch, 1996] David Carmel and Shaul Markovitch. Incorporating opponent models into adversary search. *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Portland, Oregon, 1996.

[Delahaye, 1995] Jean-Paul Delahaye. *Logique, Informatique et paradoxes*. Pour la science, Diffusion Belin. 1995.

[Goldberg, 1989] David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine learning*. Addison-Wesley, 1989.

[Holland *et al.*, 1986] John H. Holland, K.J. Holyoak, P.R. Thagard, and R.E. Nisbet. *Induction : Processes of Inference, Learning, and Discovery*. MIT Press, Cambridge, 1986.

[Lorenz and Markovitch,1993] David H. Lorenz and Shaul Markovitch. *Derivative evaluation function learning using genetic operators*. In Epstein and Levinson, editors. Proceedings of the AAAI Fall Symposium on Intelligent

Games: Planning and Learning, Menlo Park, CA. The AAAI Press, 1993.

[Meyer and Ganascia, 1996] Christophe Meyer and Jean-Gabriel Ganascia. *S.A.G.A.C.E. Genetic Algorithmic Solution for the Anticipation of Evolving Behaviors*. Paris VI, LAFORIA N°96/32, 1996

[Meyer et al., 1997] C. Meyer, J-G Ganascia, J-D Zucker. Learning Strategies in Games by Anticipation. *Proc of the fifteenth International Joint Conference on Artificial Intelligence, IJCAI'97*. Morgan Kaufman editor 1997.

[Minasi, 1991] M. Minasi "Recognizing Patterns", *AI Expert*, February 1991.

[Neumann, 1944] John von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*, Princeton University Press, 1944.

[Rapoport, 1965] Anatol Rapoport and A. Chammah. *Prisoner's Dilemma*. Ann Arbor, University of Michigan Press, 1965.

[Shannon ,1953] C.E. Shannon ; *"A Mind-Reading (?) Machine"*, Bell Laboratories Memorandum, March, 1953.

[Stone and veloso, 1996] Peter Stone and Manuela Veloso. *Towards Collaborative and Adversarial Learning: A Case Study in Robotic Soccer*. IJHCS. 1996

[Stern, 1985] Leonard Stern. *Structures and strategies of human memory*. Dorsey Press. 1985.

[Waterman and Hayes-Roth, 1978] D. Waterman and F. Hayes-Roth. *Pattern-Directed Inference Systems.* Academic Press, 1978.