

Stratégies parfaites ou adaptatives ?

Le calcul mathématique et informatique d'un équilibre de Nash donnant une stratégie mixte optimale pour jouer au poker HULHE a conduit en 2015 au programme de jeu *Cepheus*, qui ne perd jamais quel que soit son adversaire rencontré en face à face sur un grand nombre de parties. Cependant, ce programme de jeu n'est pas conçu

pour s'adapter au jeu particulier des joueurs rencontrés, et donc ne les exploitera pas toujours au mieux.

Un autre programme, *Polaris*, aussi conçu à l'Université de l'Alberta au Canada, a été élaboré avec un objectif d'ajustement à l'adversaire. L'idée n'a pas été de modéliser le joueur adverse, ce que font d'autres programmes de poker, car cette modélisation ne fonctionne bien que lorsqu'un très grand

nombre de parties consécutives sont connues.

Le principe adopté pour *Polaris* a été de prédéfinir une série limitée d'assez bonnes stratégies de jeux (en utilisant des calculs d'équilibre de Nash pour des versions simplifiées du jeu) et de comparer ce qu'elles auraient donné quand quelques parties ont été jouées contre l'adversaire. Dès que 50 parties sont disponibles pour faire cette comparaison, le choix opéré

parmi la liste des bonnes stratégies fixées au départ, en comparant leurs résultats sur les 50 parties, se révèle judicieux et conduit à une bonne exploitation des particularités du joueur qu'affronte le programme.

En 2008, le programme *Polaris* a affronté six joueurs professionnels dans des séries de 500 parties en *duplicate* (chaque partie est jouée deux fois en inversant les distributions). Il a gagné trois séries, en a perdu deux et a fait un match nul (M. Bowling et al., *A demonstration of the Polaris poker system*, Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems, 2009).



L'ÉQUIPE DE CEPHEUS.
De gauche à droite :
Mike Johanson, Michael
Bowling et Neil Burch.

Université de l'Alberta

Plusieurs questions que les professionnels se posaient sur le jeu, concernant par exemple l'intérêt de bluffer dans certaines situations, ont pu être résolues en examinant la solution calculée. Ainsi, même les professionnels du jeu peuvent tirer de précieuses informations de telles recherches.

Une fois les calculs faits, la stratégie optimale calculée, nommée *Cepheus*, s'applique assez rapidement (*pour la tester et s'informer sur ses propriétés, voir le site <http://poker.srv.ualberta.ca>*).

Pour comprendre le travail des chercheurs, revenons sur ce qu'est une stratégie optimale dans un jeu à information incomplète avec hasard. La compréhension de ce type de jeux repose sur l'analyse qu'en a faite John von Neumann dans son célèbre livre *Theory of Games and Economic Behavior* écrit avec Oskar Morgenstern en 1944. Cette théorie, complétée entre autres par John Nash (qui reçut pour cela le prix Nobel d'économie en 1994, et qui est décédé en mai dernier), permet dans les bons cas de définir et de calculer les meilleures stratégies probabilistes, aussi dénommées stratégies mixtes.

Envisageons d'abord le jeu pierre-feuille-ciseaux (P-F-C) : chacun des deux joueurs choisit au même instant P, F ou C. Le choix P gagne sur C (la pierre casse les ciseaux) ; F gagne sur P (la feuille enveloppe la pierre) ; C gagne sur F (les ciseaux coupent la feuille). La meilleure façon de jouer à ce jeu pour un joueur ne peut pas consister à toujours choisir P (ou F, ou C) car alors son adversaire observant cette stratégie constante gagnerait à chaque coup : il choisira de toujours jouer F (ou toujours C, ou toujours P, respectivement). Le calcul, ici rendu facile par des considérations de symétrie, montre que la stratégie optimale pour un joueur est de jouer aléatoirement C avec une probabilité 1/3, P avec une probabilité 1/3 et F avec une probabilité 1/3. Avec cette stratégie, votre adversaire ne vous battra jamais et si lui-même l'adopte, vous ferez jeu égal, aux fluctuations statistiques près.

Pour nous approcher du poker HULHE, considérons le poker à trois cartes (parfois appelé poker AKQ). Deux joueurs A et B s'affrontent. Le jeu ne comporte que trois cartes qui sont la carte 1, la carte 2, la carte 3

et dont la hauteur détermine la force (la carte 3 gagne sur les deux autres, etc.). Chaque joueur engage un euro. On distribue une carte à chacun. Chaque joueur regarde sa carte. Le joueur A doit choisir : enchérir en ajoutant un euro au pot, ou renoncer et donc perdre l'euro engagé. Si le joueur A a enchéri, le joueur B doit lui-même choisir de suivre (ajouter un euro au pot), ou renoncer (et donc perdre l'euro engagé). S'il suit, les joueurs montrent leur carte et la plus forte emporte le pot.

Théorie des jeux pour micro-poker

La plupart des situations de ce jeu très simple sont évidentes à analyser. Si par exemple A dispose de la carte 3, il doit enchérir puisqu'il ne peut pas perdre. Il y a cependant deux situations de jeu qui méritent une analyse plus fine. Si le joueur A dispose de la carte 1, il peut espérer que B a la carte 2 et tenter de le bluffer en enchérissant. Si A a enchéri et que B a la carte 2, B se demandera si A bluffe ou s'il a vraiment

la carte 3, et ne saura donc pas avec certitude quoi faire : enchérir ou renoncer ?

L'analyse de ce jeu (voir l'encadré pages 80-81) conduit à la conclusion assez peu évidente que lorsque A dispose de la carte 1, il doit bluffer dans un tiers des cas au hasard [comme pour le jeu pierre-feuille-ciseaux]. Lorsque B a la carte 2 et que A vient d'enchérir, B doit lui aussi enchérir dans un tiers des cas au hasard. En jouant ainsi, un joueur qui, lors d'une série de parties, sera parfois en position A et parfois en position B ne pourra pas être battu.

Cette stratégie mixte est optimale et correspond à ce qu'on nomme un équilibre de Nash. Il est remarquable ici que la théorie des jeux conseille au joueur A de bluffer parfois quand il a la carte 1. Le bluff, contrairement à ce qu'on pourrait penser, n'est pas une question de psychologie, mais est le comportement rationnel optimal d'un bon joueur qui a étudié mathématiquement le jeu : la théorie ordonne de bluffer et indique précisément avec quelle probabilité. Dans le cas du poker HULHE, la situation est comparable... mais beaucoup plus compliquée.

Vraiment optimale ?

Il faut noter ici un point un peu subtil qui concerne ce type de résultats, et en particulier le calcul de la stratégie optimale *Cepheus*. Lorsqu'un joueur joue mal en commettant toujours les mêmes erreurs, son adversaire s'en apercevra parfois et tentera alors d'en tirer avantage. Au jeu pierre-feuille-ciseaux, il est clair que si votre adversaire joue toujours *pierre*, vous en profiterez en jouant toujours *feuille*. Vous ne jouerez plus la stratégie optimale probabiliste, vous prendrez le risque que votre adversaire se réveille et se mette à exploiter votre jeu constant, mais s'il s'obstine à toujours jouer *pierre*, vous gagnerez bien plus en jouant toujours *feuille* qu'avec la stratégie mixte « optimale ».

Les situations peuvent d'ailleurs être bien plus complexes. Cette rubrique dans *Pour la Science* de septembre 1999 décrivait les succès de méthodes de jeu proposées par Claude Shannon et par un programme d'intelligence artificielle qui, au jeu

de type pierre-feuille-ciseaux, battaient nettement la plupart des joueurs humains. Ces programmes repéraient et exploitaient les régularités des joueurs, régularités dont eux-mêmes n'avaient pas conscience et dont ils n'arrivaient pas à se libérer. Ces programmes ne jouaient bien sûr pas la stratégie optimale de la théorie des jeux, mais une méthode plus subtile fondée sur l'analyse de la succession des choix de l'adversaire. Les méthodes de la théorie des jeux permettent de trouver les stratégies mixtes qui ne perdent jamais (statistiquement) ; c'est très bien. Mais on peut espérer mieux et vouloir tirer le plus possible des faiblesses identifiables de son adversaire, grâce à l'analyse des coups passés et à des algorithmes adaptatifs.

Il n'existe pas de méthode mathématique pour définir et calculer ce que serait une stratégie exploitant aussi rapidement que possible toute faiblesse repérable d'un adversaire. Toutefois, il est clair que dans une vraie partie de pierre-feuille-ciseaux ou de poker, bien jouer implique une telle prise en compte.

Soyons clairs : dans un tournoi qui opposerait des joueurs humains jouant obstinément de mauvaises stratégies (donc analysables et exploitables), la récente stratégie optimale *Cepheus* ou d'autres stratégies prenant quelques risques pour exploiter les joueurs faibles, il n'est pas certain que la stratégie *Cepheus* soit la gagnante. Dit encore autrement : la stratégie optimale *Cepheus* ne se fait battre par personne dans une compétition à deux joueurs, mais dans une compétition où plusieurs joueurs joueraient successivement par deux, les gains moyens de la stratégie *Cepheus* seraient probablement inférieurs à ceux du programme *Polaris* (par exemple) évoqué plus haut, et qui, lui, est conçu pour s'ajuster aux joueurs rencontrés (voir l'encadré page ci-contre).

Quand l'information est incomplète et que le hasard intervient, tout devient compliqué et subtil, et la théorie mathématique n'a pas encore trouvé le moyen d'avoir le dernier mot face aux méthodes heuristiques et indéfiniment améliorables de l'intelligence artificielle.

L'AUTEUR



J.-P. DELAHAYE est professeur émérite à l'Université de Lille et chercheur

au Centre de recherche en informatique, signal et automatique de Lille (CRISTAL).

BIBLIOGRAPHIE

M. Bowling et al., *Heads-up limit hold'em poker is solved*, *Science*, vol. 347, pp. 145-149, 2015.

Cepheus Project, pages web d'information sur le programme optimal calculé par l'équipe de Michael Bowling, 2015 : <http://poker.srv.ualberta.ca/>

O. Tammelin, *Solving large imperfect information games using CFR+*, prépublication arXiv:1407.5042, 2014.

J. Rubin et I. Watson, *Computer poker : A review*, *Artificial Intelligence*, vol. 175(5), pp. 958-987, 2011.

J.-P. Delahaye, *La fin des dames anglaises ?*, *Pour la Science* n° 363, janvier 2008 ; *L'intelligence humaine à nouveau dominée*, *Pour la Science* n° 263, septembre 1999.



Retrouvez la rubrique Logique & calcul sur www.pourlascience.fr