

L'un des résultats annexes obtenus par la solution calculée est que le meneur de jeu (celui qui distribue les cartes) est favorisé par rapport à son adversaire dans une partie de poker HULHE : sur le long terme et si c'était toujours le même joueur qui distribuait les cartes, il serait certain de gagner de l'argent. C'est conforme à ce que les professionnels du jeu pensaient depuis longtemps : celui qui parle en premier est désavantagé.

## Des calculs massifs

L'analyse du jeu a été un énorme travail mathématique et informatique. Les diverses équipes étudiant le poker ont récemment introduit de nouvelles méthodes cruciales pour réussir. En particulier, la méthode

permettant de calculer les solutions pour des règles de jeux qui amènent un très grand nombre de situations de jeu a été mise au point il y a moins de dix ans.

Cette méthode a pour nom CFR, pour « Counterfactual Regret Minimization ». L'idée est de faire jouer un programme contre lui-même et de faire évoluer les probabilités de choix qu'il calcule pour chaque point de décision en fonction des succès et échecs obtenus lors de parties simulées. Depuis son introduction en 2007, cette méthode a permis d'étudier des jeux dont le nombre de points de décision a été de plus en plus grand. Un poker simplifié ayant  $3,8 \times 10^{10}$  situations de jeu a pu ainsi être traité en 2012, ouvrant la porte au calcul de la solution pour le jeu réel HULHE qui en comporte presque 10 000 fois plus.

Comme pour le calcul effectué afin de prouver que 16 données ne garantissent jamais une solution unique pour un problème de sudoku  $9 \times 9$  (voir *Pour la Science* de janvier 2015), ce qui a été fait est aux limites de ce qui est techniquement envisageable aujourd'hui. Pour déterminer leur solution optimale à epsilon près, l'équipe de Michael Bowling a lancé un travail informatique sur un réseau de 200 machines disposant chacune de 32 gigaoctets (32 milliards de fois 8 bits) de mémoire vive et d'un disque d'un téraoctet (mille milliards de fois 8 bits). Le jeu a été divisé en 110 565 sous-jeux. Chaque itération de l'algorithme CFR exigeait 61 minutes. Les 1 579 itérations requises ont pris deux mois pour l'ensemble du dispositif. Le calcul total équivalait à 900 années de fonctionnement d'une machine à un seul cœur.

## poker à trois cartes (poker AKQ)

### Calcul des stratégies optimales

| GAINS | $A_1/B_2$               | $A_1/B_3$     | $A_2/B_1$    |
|-------|-------------------------|---------------|--------------|
|       | $-2pq + p(1-q) - (1-p)$ | $-2p - (1-p)$ | 1            |
| A     | $A_2/B_3$               | $A_3/B_1$     | $A_3/B_2$    |
|       | -2                      | 1             | $2q + (1-q)$ |

Pour calculer l'équilibre de Nash, on étudie le gain moyen de A et de B pour des valeurs  $p$  et  $q$  fixées (tableau ci-dessus).

À titre d'exemple, expliquons l'expression  $-2pq + p(1-q) - (1-p)$  correspondant au gain de A dans le cas où A a la carte 1 et B la carte 2. Si A bluffe et B suit (probabilité  $pq$ ), A perd 2 euros (d'où le  $-2pq$ ) ; si A bluffe et B ne suit pas (probabilité  $p(1-q)$ ), A gagne 1 euro (d'où le  $+p(1-q)$ ) ; si A ne bluffe pas (probabilité  $(1-p)$ ), A perd un euro (d'où le terme  $-(1-p)$ ).

Le gain moyen de A en fonction de  $p$  et  $q$  (en supposant que les cartes sont distribuées équitablement) s'obtient en faisant la somme des gains dans chaque partie possible et en divisant par 6 (car les six parties possibles sont équiprobables). On trouve :

$$\text{Gain A}(p, q) = -(p-1/3)(q-1/3)/2 - 1/9 \\ = -\text{Gain B}(p, q)$$

L'équilibre de Nash qui donne les stratégies mixtes optimales pour A et B s'obtient par le calcul, en examinant les points où les dérivées de Gain A par rapport à  $p$  et par rapport à  $q$  s'annulent. En un tel point, si A modifie  $p$ , B peut augmenter son gain en ajustant  $q$ , et réciproquement : aucun des deux joueurs, s'il est prudent, n'a donc intérêt à changer son choix ( $p$  pour A,  $q$  pour B).

Ce calcul de dérivées donne ici :

$$\partial(\text{Gain A})/\partial p = (-q + 1/3)/2 = 0, \text{ donc } q = 1/3;$$

$$\partial(\text{Gain B})/\partial q = (-p + 1/3)/2 = 0, \text{ donc } p = 1/3.$$

En ce point  $p = q = 1/3$ ,  $\text{Gain A} = -1/9$ , ce qui signifie qu'à ce micro-poker, celui qui parle en premier (A) est perdant en moyenne de 1/9 euro à chaque partie.

D'autres détails sur ce micro-poker figurent dans le livre de B. Chen et J. Ankenman, *The Mathematics of Poker*, ConJelCo LLC, 2007 (édition en français : *Poker Maths Sup*, Micro-Application, 2013).

Le hasard semble jouer un rôle très important au poker, aussi parfois on a pensé qu'il n'y avait pas de bons joueurs, mais seulement des joueurs chanceux. Ceux qui pratiquent le poker sont persuadés du contraire : il existe des façons intelligentes de jouer et d'autres qui, lors d'une longue série

de parties, assurent de perdre. Le débat a parfois eu une importance concrète pour l'autorisation ou non du poker en ligne dans certains États où les jeux de pur hasard sont interdits. La loi française définit ainsi les jeux de hasard : « Est un jeu de hasard un jeu payant où le hasard prédomine sur l'habileté et les combinaisons de l'intelligence pour l'obtention du gain. » Plusieurs décisions de justice ont classé le poker dans la catégorie des jeux de hasard. En 2013, une nouvelle décision prise à Toulouse a décidé du contraire. La Cour de cassation a tranché en qualifiant le poker de jeu de hasard (voir <http://www.reynaud-avocat.com/la-légalité-des-jeux-d-argent/>).

Elle a tort, et plusieurs études scientifiques ont établi l'importance du savoir-faire et de l'intelligence dans la pratique du poker. Steven Levitt et Thomas Miles, de l'Université de Chicago, ont mené une étude décisive sur cette question (S. D. Levitt et T. J. Miles, *The role of skill versus luck in poker: Evidence from the world series of poker*, *Journal of Sports Economics*, vol. 15(1), pp. 31-44, 2014, article en accès libre sur <http://jse.sagepub.com/content/15/1/31.abstract>).

## Stratégies parfaites ou adaptatives ?

**L**e 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