**AlphaZero, algorithme prodige de l’échiquier**

Vingt ans après la défaite de l’homme face à la machine, le jeu d’échecs a donné lieu à un nouveau duel  : Stockfish, champion du monde 2016 des logiciels « classiques », a été battu à plate couture par l’intelligence artificielle d’une filiale de Google.

LE MONDE | 19.12.2017 à 12h00 | Par [Pierre Barthélémy](http://abonnes.lemonde.fr/journaliste/pierre-barthelemy/)

L’annonce a été faite le 5 décembre, en plein milieu du London Chess Classic, un des tournois d’échecs les plus relevés de la saison : l’algorithme d’intelligence artificielle AlphaZero de la société britannique DeepMind – rachetée en 2014 par Google – avait éreinté Stockfish, champion du monde des programmes d’échecs 2016 et très supérieur à n’importe quel champion humain. Non seulement, AlphaZero a remporté 28 victoires sur 100 parties sans concéder la moindre défaite – les 72 autres rencontres s’étant soldées par la nulle –, non seulement il n’évaluait que 80 000 coups par seconde contre 70 millions à son adversaire, mais cette humiliation est intervenue alors que l’entité de DeepMind ne connaissait rien aux échecs quelques heures seulement avant cette confrontation.

Interrogés par [Chess.com](https://www.chess.com/fr), les grands maîtres présents à Londres semblaient pour le moins éberlués par le ratatinage en règle subi par le programme qui est leur compagnon quotidien d’entraînement et de préparation. Le meilleur joueur français (et n° 5 mondial) Maxime Vachier-Lagrave déclarait ainsi : « C’est un résultat remarquable. Même en rêve, je serais incapable de battre Stockfish ne serait-ce qu’une seule fois. Le score d’AlphaZero est particulièrement impressionnant. » Futur vainqueur du London Chess Classic, l’Américain Fabiano Caruana (n °4 mondial), disait : « Je suis stupéfait. Je pense que jamais un programme d’échecs n’avait autant dominé les autres. » Son compatriote Wesley So (n °6 mondial) renchérissait : « Je suis choqué. C’est un gros événement, qui va changer le visage du jeu. » Etc.

## Apprentissage par renforcement

Même s’il débarque de manière spectaculaire dans le monde des 64 cases, AlphaZero n’est pas un parfait inconnu. Il s’agit de la version générique d’AlphaGo, l’algorithme qui, au cours des deux dernières années, a établi la supériorité de la machine sur l’Homo sapiens au go. Mais, alors qu’AlphaGo s’était formé au jeu en « analysant » une base de données contenant de nombreuses parties jouées par des hommes, son héritier s’est passé de cette étape : il a lui-même forgé sa théorie des échecs à partir des seules règles du jeu, comme l’explique l’équipe de DeepMind dans un article mis en ligne sur le site de prépublications scientifiques [arXiv](https://arxiv.org/).

Lire aussi :   [Intelligence artificielle : toujours plus puissant, AlphaGo apprend désormais sans données humaines](http://abonnes.lemonde.fr/pixels/article/2017/10/18/intelligence-artificielle-toujours-plus-puissant-alphago-apprend-desormais-sans-donnees-humaines_5202931_4408996.html)

Pour ce faire, AlphaZero a utilisé ce que les informaticiens appellent l’« apprentissage par renforcement ». L’algorithme a joué des millions de parties contre lui-même, tout d’abord en choisissant ses coups au hasard – puisqu’on ne lui avait rien appris de la stratégie, de la théorie des ouvertures, etc. Puis, au fur et à mesure que les parties s’accumulaient, il engrangeait des statistiques monumentales sur les positions rencontrées, les coups joués. Au bout d’un moment, il a fini par savoir que, dans telle configuration, tel coup avait la meilleure probabilité de conduire à la victoire.

Comme l’explique Tristan Cazenave, professeur d’intelligence artificielle au Laboratoire d’analyse et modélisation de systèmes pour l’aide à la décision (Lamsade, CNRS/Université Paris-Dauphine) et spécialiste des jeux, « cette approche fournit une évaluation très précise et permet d’explorer les coups les plus intéressants ». Elle s’apparente à ce qui se passe dans la tête d’un champion qui sait, de manière instinctive, en raison des milliers des parties qu’il a jouées ou analysées, quels sont les meilleurs « coups candidats » dans une position donnée. Ce filtre permet de se concentrer sur eux uniquement et de les étudier le plus loin possible.

## « Effet horizon »

Depuis des décennies, les programmateurs des logiciels d’échecs – depuis Deep Blue, le supercalculateur d’IBM qui a battu en 1997 le champion du monde de l’époque, le Russe Garry Kasparov, jusqu’à Stockfish – ont privilégié une philosophie bien différente, celle dite de la « force brute ». L’idée consiste à établir l’arborescence des coups possibles dans une position donnée, de les analyser et de les évaluer tous, puis de passer au coup suivant et ainsi de suite. Le nombre de positions obtenues augmente de façon exponentielle mais ce handicap a été en grande partie compensé par la puissance sans cesse croissante des ordinateurs sur lesquels les programmes tournent.

Par ailleurs, les informaticiens ont élaboré des stratégies pour couper les branches de l’arbre visiblement pourries. Ils ont également intégré à leurs programmes des bases d’ouvertures et de finales, nourries par les dernières découvertes des joueurs d’échecs. Enfin, ils ont raffiné à l’extrême la fonction qui leur sert à évaluer les positions, pour tenir compte par exemple de la bonne ou mauvaise santé de la structure de pions, de la possession ou non de la paire de fous, de la sécurité du roi, etc.

Lire aussi :   [L’espèce humaine, échec et mat](http://abonnes.lemonde.fr/jeux-d-esprit/article/2014/09/11/l-espece-humaine-echec-et-mat_4486139_1616889.html)

L’avantage de la force brute est que le logiciel ne ratera aucune combinaison rapide, qu’il verra instantanément ou presque un mat en dix coups. L’inconvénient réside dans le fait que, en raison des contraintes temporelles d’une partie, les programmes du commerce sont contraints de choisir une ligne sans jamais en avoir fait une analyse complète et peuvent être victimes de ce que l’on appelle « l’effet d’horizon ». Tout comme on ne peut voir ce qui est au-delà de l’horizon, un logiciel basé sur la force brute n’a pas l’assurance que, même si la ligne qu’il sélectionne est la plus prometteuse sur les 20 coups qu’il a calculés, elle ne va pas se transformer en bourde au 21e.

Les dix parties fournies dans son article par l’équipe d’AlphaZero montrent que le bébé de Deep Mind développe un jeu très positionnel qui tente de restreindre au maximum les possibilités de Stockfish. Ainsi que l’a remarqué Maxime Vachier-Lagrave, cité par Chess.com, « son plus gros avantage, c’est la quasi-absence d’effet d’horizon. Dans la plupart des parties, AlphaZero était gagnant et Stockfish a continué à évaluer la position comme équilibrée sur près de 20 coups parfois ! (…) Mais à la fin de la séquence, sa position se détériorait très rapidement. »

## Vieille version de Stockfish

Stockfish s’est tout de même plaint des conditions de l’expérience en la personne d’un de ses développeurs historiques, le Norvégien Tord Romstad. Celui-ci a remarqué que « la version de Stockfish utilisée était vieille d’un an, jouait avec plus de tâches en arrière-plan qu’au cours de tous les tests que nous avons effectués ». Il a aussi noté que la cadence de jeu imposée – une minute maximum par coup – réduisait à néant les efforts des programmateurs pour que le logiciel consacre plus de temps de réflexion dans les positions complexes. Surtout, Tord Romstad a souligné que, pour son apprentissage, AlphaZero avait tourné sur « un système dédié, ultrapuissant » : 5 000 processeurs spécialement conçus par Google pour le machine learning.

Tristan Cazenave regrette pour sa part que « le programme ne soit pas disponible, qu’on ne puisse pas jouer avec et que l’on n’ait que quelques parties ». Le chercheur rappelle que « DeepMind a fourni, pour les joueurs de go, plein d’ouvertures avec les différentes évaluations de la machine ». Peut-être en fera-t-il autant pour les échecs ?

En savoir plus sur http://www.lemonde.fr/sciences/article/2017/12/19/alphazero-algorithme-prodige-de-l-echiquier\_5231860\_1650684.html#03QzWL0re6A2oDm7.99