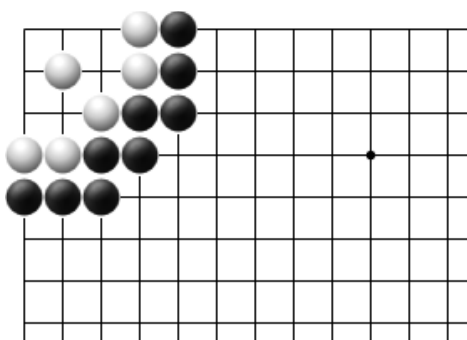


LA MÉTHODE DE MONTE-CARLO POUR BATTRE LES MEILLEURS JOUEURS DE GO

Nous avons vu la méthode de Monte-Carlo dans l'exercice 94 p.274, ce qui nous a permis de calculer l'intégrale d'une fonction dont on ne connaît pas de primitive.

La méthode de Monte-Carlo, ou méthode Monte-Carlo, désigne une famille de méthodes algorithmiques visant à calculer une valeur numérique approchée en utilisant des procédés aléatoires, c'est-à-dire des techniques probabilistes. Le nom de ces méthodes, qui fait allusion aux jeux de hasard pratiqués à Monte-Carlo, a été inventé en 1947 par Nicholas Metropolis, et publié pour la première fois en 1949 dans un article coécrit avec Stanislaw Ulam.



Le jeu de go est né en Chine il y a plusieurs milliers d'années.

Il se joue au Japon depuis 1200 ans, mais il ne s'est répandu que récemment en Occident.

Le but du jeu est la constitution de territoires en utilisant un matériel des plus simples : un plateau, appelé goban, sur lequel est tracé un quadrillage et des pions, appelés pierres, que l'on pose sur les intersections de ce quadrillage à tour de rôle.

Comme pour tous les jeux, il faut jouer un coup qui améliore sa situation et détériore celle de son adversaire. Pour estimer une situation aux échecs, une bonne estimation est de compter le nombre de pièces sur l'échiquier, en les pondérant (1 point par pion, 5 par tour...), et en ajustant la valeur trouvée par les libertés, les protections des pièces... Cela passe par le calcul d'une fonction d'évaluation, associant les scores de chacun des adversaires à chaque nœud.

Cela difficilement réalisable au go : ***on ne dispose pas de fonction d'évaluation*** (estimation des valeurs antagonistes d'une position) ne nécessitant pas des capacités « humaines » de reconnaissance de formes, l'expérience de parties déjà jouées et une très grande profondeur de calcul.

La technique d'exploration des différentes possibilités (pour chaque coup, déterminer la meilleure réponse possible, puis la meilleure réponse à celle-ci, et ainsi de suite...) échoue au go à cause de l'énorme quantité de coups plausibles, de la durée des parties et de la complexité croissante (aux échecs, la complexité est décroissante par diminution du nombre de pièces restantes).

Au jeu de go, le nombre de configurations/positions possibles, ou complexité, est de 10^{170} environ. Cette complexité est telle que les méthodes arborescentes efficaces pour le morpion (complexité 10^3), puissance 4 (complexité 10^{14}), les dames anglaises (complexité 10^{20}) ou les échecs (complexité 10^{40}) sont inefficaces pour le jeu de go.

Le nombre de combinaisons possibles est supérieur au nombre d'atomes que compte l'univers...

L'arbre du jeu couvre 10^{120} parties plausibles pour les échecs, contre 10^{600} pour le jeu de go !

Quand les méthodes exactes ne permettent pas de déterminer de manière formelle la stratégie optimale, il est ***nécessaire d'utiliser des méthodes approchées***¹.

¹ Comme par exemple le logiciel qui a créé les emplois du temps des enseignants et des élèves dans cet ensemble scolaire

Utilisée dans les programmes d'échecs ou de dames, l'approche classique consiste à **limiter la profondeur de l'arbre**, en étudiant les positions après un petit nombre de coups fixé, ce qui permet de réduire la combinatoire.

En réduisant ainsi la profondeur de l'arbre, les feuilles ne correspondent plus à la fin d'une partie.

On ne peut alors généralement pas en déterminer de manière formelle le gagnant.

On remplace alors la valeur formelle de ces positions par ce qu'on appelle une **évaluation heuristique**².

Par exemple, aux échecs, cette heuristique se base souvent sur le comptage du matériel : 9 points pour une dame, 5 points pour une tour, 3 points pour un fou ou un cavalier, 1 point pour un pion. À cette évaluation matérielle s'ajoutent des termes liés à la position des pièces, comme la mobilité ou la sécurité du roi. On peut ainsi obtenir une estimation approximative de la position en analysant la différence de force entre les deux armées.

L'approche Monte-Carlo

Alors qu'aux échecs il est simple et efficace de mesurer la force des deux armées en présence en comptant le nombre de pièces, il est très difficile d'estimer les territoires dans une position de go.

Afin de contourner cette difficulté dans la construction de fonctions d'évaluation heuristiques, l'approche de Monte-Carlo consiste à systématiquement explorer une branches de l'arbre jusqu'à sa position terminale.

Étant donné le nombre astronomiquement élevé de parties possibles, cette approche ne peut pas explorer exhaustivement toutes les possibilités : il faut donc choisir un sous-ensemble des parties possibles.

En se basant sur ce principe d'exploration partielle de l'arbre jusqu'aux feuilles, la mise au point d'un algorithme Monte-Carlo efficace repose sur deux problèmes essentiels :

- le choix des séquences à explorer : comment choisir celles qu'on explore ?
- l'évaluation des coups : étant données les séquences observées et leur résultat (gagné ou perdu, dans le cas du go), comment déterminer le meilleur coup ?

Un des algorithmes les plus primitifs consiste à simplement **choisir les coups aléatoirement**.

Pour chacun des coups possibles, la moyenne des résultats des parties obtenues après avoir joué ce coup donne une estimation de sa valeur. Il suffit ensuite de choisir le coup qui mène vers la position qui a la meilleure moyenne.

Pour donner un ordre de grandeur, les meilleurs programmes actuels (CrazyStone, Zen, Mogo...) parviennent à générer plusieurs dizaines de milliers de parties aléatoires par seconde sur un PC puissant.

Mais, parce que les coups utilisés pour l'évaluation sont choisis au hasard, il est possible qu'un coup qui serait excellent sauf pour une réponse spécifique soit de façon erronée choisi comme un bon coup.

Le résultat est un programme qui est fort d'un point de vue stratégique mais faible tactiquement.

Un Monte-Carlo plus intelligent

Une version améliorée de ces simulations dites Monte-Carlo a été proposée au milieu des années 2000.

Les idées essentielles sont les suivantes :

- ne pas choisir les coups de manière uniformément aléatoire : certains coups qui semblent meilleurs a priori (comme répondre à une menace immédiate) peuvent être choisis avec une probabilité plus élevée, ce qui améliore la précision de l'évaluation.
- augmenter la probabilité de choisir un coup qui s'est avéré être bon dans les simulations précédentes.

On parle de **méthode MCTS** pour "**Monte-Carlo Tree Search**".

2 L'heuristique (du grec ancien εὐρίσκω, eurisko, « je trouve »), parfois orthographiée euristique, est un terme de didactique qui signifie « l'art d'inventer, de faire des découvertes ». En psychologie, une heuristique de jugement désigne une opération mentale, rapide et intuitive.

Le 28 janvier 2016 : coup de maître dans l'univers de l'intelligence artificielle !

Pour la première fois, un ordinateur bat un joueur de go professionnel, comme le détaille un article de recherche publié dans la revue *Nature*. Ses vingt auteurs sont tous membres de Google DeepMind, une entreprise britannique créée en 2010 et qui a été rachetée en 2014 par le géant californien.

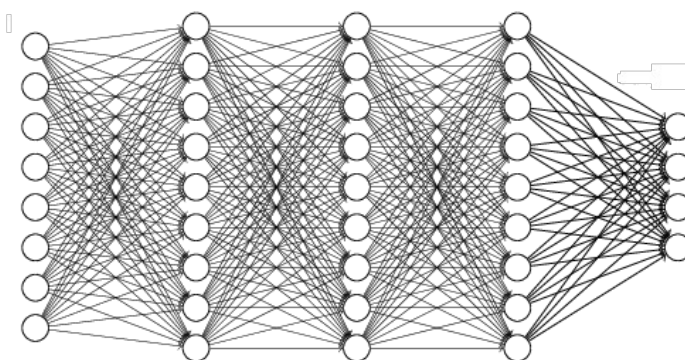
Leur algorithme, AlphaGo, a battu le champion Fan Hui par cinq victoires à zéro en octobre 2015 à Londres. Il n'a en outre perdu qu'une partie sur 500 contre les meilleurs programmes déjà sur le marché.

En mars 2016, AlphaGo bat Lee Sedol, joueur 9^e dan, alors considéré comme un des meilleurs joueurs mondiaux, en remportant successivement les trois premières parties, puis la cinquième d'un match en cinq parties. Ce match est suivi en direct sur internet, et qualifié de moment historique.

Alors, comment AlphaGo fonctionne-t-il ?

L'équipe de DeepMind a utilisé deux ingrédients : les *réseaux de neurones profonds* (le *deep learning* ou *apprentissage profond*) et *l'apprentissage renforcé*.

LE DEEP LEARNING ou APPRENTISSAGE PROFOND



Le *machine learning* (*apprentissage automatique*) est en quelque sorte l'apprentissage par l'exemple : les algorithmes utilisés permettent, dans une certaine mesure, à un système piloté par ordinateur (un robot éventuellement), ou assisté par ordinateur, d'adapter ses analyses et ses comportements en réponse, en se fondant sur l'analyse de données empiriques provenant d'une base de données ou de capteurs.

Par exemple, on peut faire analyser à un algorithme des milliers d'images de chats et de chiens, afin qu'il en tire lui-même des caractéristiques : c'est la phase d'apprentissage. Puis l'algorithme apprendra de lui-même en essayant, lorsqu'on lui présente une des images, de prédire si c'est un chien ou un chat. S'il se trompe, il apprendra de son erreur et essaiera de ne plus la refaire.

Le *deep learning* est une manière particulière de faire du machine learning.

L'idée est de créer un réseau de neurones artificiels, c'est-à-dire une combinaison de milliers de variables reliées entre elles comme un réseau de neurones biologiques.

Son succès repose notamment sur des quantités énormes d'informations (images, textes, sons...) nécessaires pour « stimuler » et construire le réseau de neurones (phase d'apprentissage).

Grâce notamment au français Yann Le Cun³, une percée a été effectuée dans les années 1990 afin de calculer rapidement et efficacement les meilleures « connexions ».

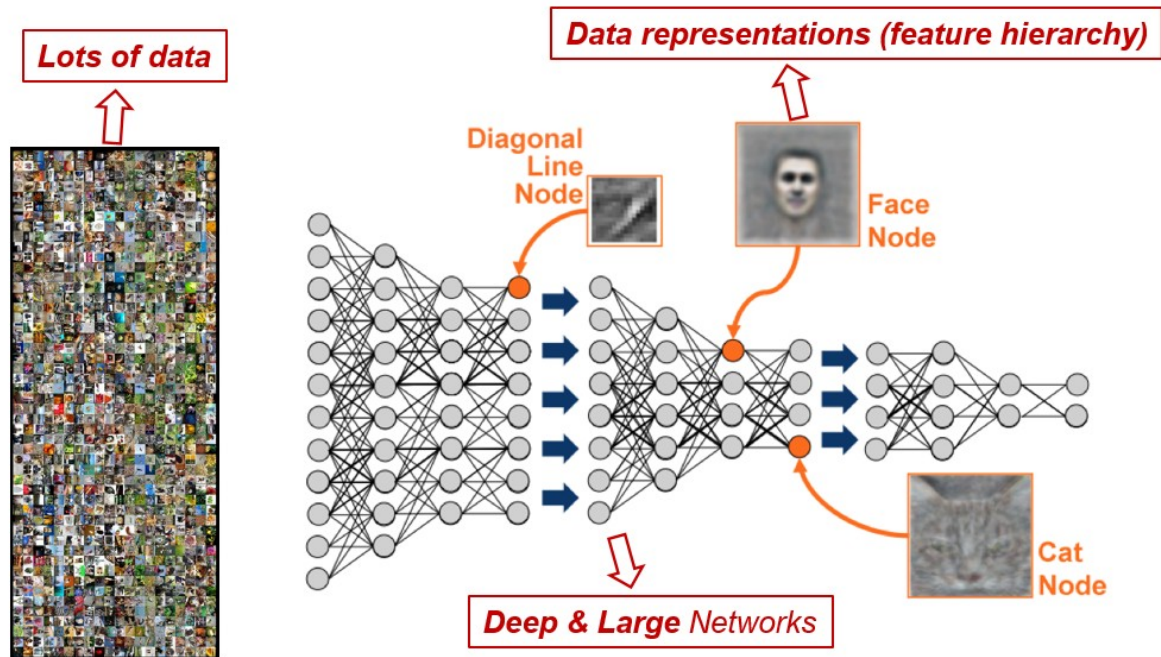
3 Il a publié plus de 130 documents et articles sur la vision artificielle, les réseaux de neurones et la reconnaissance d'images en particulier, domaine dans lequel il est considéré comme un des pionniers.

Fin 2013, il rejoint Facebook pour créer et diriger le laboratoire d'intelligence artificielle FAIR à New York, Menlo Park et depuis 2015 à Paris, notamment pour travailler sur la reconnaissance d'images et de vidéos.

Par exemple, pour identifier une image, un algorithme va analyser chaque pixel et va arriver à extraire lui-même les caractéristiques principales de cette image. Les différentes couches de neurones effectuent un apprentissage hiérarchique : elles cherchent d'abord à catégoriser les éléments les plus simples de l'information avant de passer aux plus compliqués. Le système cherchera à reconnaître les lignes avant les visages...

Lorsque le programme a accumulé les informations sur les éléments de base, il peut alors les réorganiser de différentes manières en blocs plus complexes.

En apprenant ainsi, *l'algorithme va apprendre lui-même*, à partir de milliers d'exemples, à classifier et identifier des objets, des images, des mots ou des phrases, des sons, etc.



Vous connaissez, sans le savoir, des exemples pratiques de *machine learning* : reconnaissance d'images dans Google Images, la voiture autonome de Google, la classification des emails dans Gmail, la traduction en temps réel de Skype, la reconnaissance vocale Siri d'Apple (ou Cortana de Microsoft), la détection de fraude dans le monde de la banque, la reconnaissance faciale...

Pour mieux comprendre ce qu'est le deep learning, voir l'excellente vidéo de D. Louapre :
<https://youtu.be/trWrEWfhTVg>

L'APPRENTISSAGE RENFORCÉ

Une seconde technique, spécialité de DeepMind, est l'apprentissage renforcé.

Cette technique consiste à *faire jouer un système contre lui-même* afin qu'il apprenne tout seul de ses échecs et de ses victoires.

L'apprentissage renforcé imite là encore le cerveau humain qui a besoin de répétition pour apprendre (c'est le renforcement neuronal). Il suffit d'observer un bébé tenter des centaines de fois de se saisir d'un objet pour comprendre comment fonctionne notre cerveau : par le renforcement.

En faisant jouer l'algorithme contre lui-même, les chercheurs améliorent leur réseau neuronal.

D'accord, mais... et pour AlphaGo ?

Pour réussir ce coup de maître, DeepMind a donc combiné les meilleurs techniques d'intelligence artificielle et la puissance de calcul mise à sa disposition par la plateforme de *cloud computing* de Google.

Au cœur du dispositif, **deux réseaux de deep learning ayant chacun une tâche différente.**

Le premier, appelé *policy network* (quelque chose comme « réseau politique » ou « réseau de décision ») travaille à prédire le prochain coup.

Le second, appelé *value network* (« réseau de valeur »), estime l'issue favorable d'un mouvement à partir de la configuration du plateau de jeu.

Les deux réseaux ont été associés à un arbre binaire de recherche qui utilise **la méthode algorithmique probabiliste de Monte-Carlo** : jouer toutes les fins de parties possibles à partir d'une position en comptabilisant les parties gagnantes et perdantes. Les réseaux neuronaux sont là pour affiner la recherche, réduire sa complexité et sa profondeur, en adoptant une forme de raisonnement plus proche de l'imagination humaine.

Chaque réseau est composé de douze couches contenant chacune des millions de connexions neuronales.

Les « réseaux politiques » ont été entraînés avec **une base de données de 30 millions de coups provenant de parties jouées par des experts** jusqu'à ce qu'ils puissent prédire un mouvement 57 % du temps.

Mais pour qu'AlphaGo puisse développer ses propres stratégies de jeu, DeepMind a eu recours à un apprentissage renforcé en lui faisant jouer des milliers de parties contre lui-même.

Puis, les *policy networks* ont été utilisés pour entraîner les *value networks*, toujours par apprentissage renforcé.

« Ces réseaux de valeur peuvent évaluer n'importe quelle position de go et estimer l'éventuel gagnant ».

Google.

AlphaGo a d'ores et déjà fait accomplir un pas de géant à l'intelligence artificielle. Mais au-delà du jeu de go, la grande question est de savoir ce que Google compte faire de ce formidable outil.

« Même si les jeux sont la plateforme idéale pour développer et tester rapidement des algorithmes d'intelligence artificielle, au bout du compte, nous voulons appliquer ces techniques à d'importants problèmes du monde réel », commente l'entreprise qui cite comme exemple la modélisation climatique et le diagnostic de maladies complexes. On peut aussi penser que Google exploitera cette technologie pour perfectionner ses propres services, notamment son moteur de recherche et son assistant vocal pour smartphones.

Au cours de la conférence de presse qui a suivi la présentation d'AlphaGo, Demis Hassabis a été questionné sur les risques potentiels à long terme d'une intelligence artificielle aussi performante. Il a répondu que son entreprise collaborait sur ce sujet avec des universitaires, organisait des conférences et s'était dotée d'un comité d'éthique interne. « Par ailleurs, nous nous sommes mis d'accord avec Google pour que ces technologies ne soient pas utilisées à des fins militaires », a-t-il assuré.

BILAN : CE QU'IL FAUT RETENIR

AlphaGo s'est nourri de **30 millions de mouvements de joueurs professionnels** pour sa phase d'apprentissage, qu'il a complété par une **phase d'apprentissage renforcé**.

Il utilise la **méthode de Monte-Carlo**, guidée par **deux réseaux de neurones profonds**, contenant chacun **12 couches** qui tendent à imiter le fonctionnement du cerveau humain, capable de traiter plusieurs informations en même temps, de les associer puis d'en déduire une conclusion.

Pourquoi autant de battage et de milliards investis autour de l'intelligence artificielle ?

C'est que Google, Facebook, IBM et consorts espèrent un jour atteindre le Graal : imiter au mieux le fonctionnement du cerveau humain. Un « système » capable de gérer des milliers d'informations en même temps, d'apprendre par lui-même aussi bien de ses erreurs que de son environnement, de s'adapter et même, parfois, de se réparer. Autant d'atouts qui ont un poids économique fort.

Pour les sociétés informatiques, c'est la promesse de machines plus rapides, plus précises, plus fiables.

Sur son blog, Google met en avant des utilisations « bien sous tous rapports » de son IA : modélisation à destination de l'environnement (modélisation du climat), de la science et de la médecine...

Certes... Mais les principales motivations de Google ou Facebook sont à aller chercher ailleurs que dans leur amour pour la science : l'IA doit servir à **améliorer leurs services de reconnaissance d'image mais aussi de marketing**.

Pour ces entreprises, l'IA est la promesse de **prédire précisément nos décisions d'achats, nos préférences ou encore nos déplacements**.

Bref, de nous proposer des produits parfaitement adaptés à nos habitudes, nos goûts...

IBM travaille ainsi sur **Watson**⁴, programme informatique d'intelligence artificielle dont le but est de répondre à des questions formulées en langage naturel.



Il peut donc **analyser le contenu de ce que nous postons sur les réseaux sociaux** et ainsi comprendre si vous vous apprêtez à partir en vacances, à déménager, à vous marier ou à avoir un enfant. Pour ensuite vous proposer des couches, de la crème solaire ou des boîtes en carton. En identifiant le contenu des photos que vous postez, la machine pourra déterminer que vous êtes fan de vélo, de chat ou de soirées déguisées. Là encore, de précieuses informations pour mieux cibler le consommateur que vous êtes.

Aux dernières nouvelles, IBM aurait décidé de **restreindre Watson à un usage médical**...

Un centre de recherche de la faculté de médecine de l'université de Tokyo dirigé par le professeur Arinobu Tojo a utilisé Watson en 2016 pour une aide au diagnostic d'un cas rare de leucémie. En dix minutes le supercalculateur a trouvé la véritable cause ; les médecins estiment qu'il leur aurait fallu deux semaines, rendant le traitement beaucoup plus incertain.

Sources :

- un article de M. Zaffagni : <http://www.futura-sciences.com/tech/actualites/technologie-jeu-go-nouvelle-version-AlphaGo-battu-n-1-mondial-61409/>
- <https://www.contrepoints.org/2016/02/02/237404-alphago-nouvelle-victoire-google-pour-lintelligence-artificielle>
- http://mobile.lemonde.fr/pixels/article/2016/01/27/premiere-defaite-d-un-professionnel-du-go-contre-une-intelligence-artificielle_4854886_4408996.html
- les vidéos de Science étonnante (David Louapre)
- Wikipédia
- https://www.sciencesetavenir.fr/high-tech/intelligence-artificielle/intelligence-artificielle-le-deep-learning-battu-par-l-humain-ouf_108424
- <http://statistique.blogs.sciencesetavenir.fr/archive/2015/06/04/apprentissage-automatique-le-deep-learning-23278.html>
- http://www.lemonde.fr/pixels/article/2015/07/24/comment-le-deep-learning-revolutionne-l-intelligence-artificielle_4695929_4408996.html
- <http://brain.kaist.ac.kr/research.html>

4 Le nom « Watson » fait référence à Thomas J. Watson, dirigeant d'IBM de 1914 à 1956, avant même que cette société ne s'appelle ainsi.