



Opinions Libres

le blog d'Olivier Ezratty

Les conséquences pratiques d'AlphaGo Zero

La performance de la troisième version d'AlphaGo, dite Zero, annoncée en octobre 2017 en a surpris plus d'un. L'IA de DeepMind, la filiale anglaise de Google, s'entraînait toute seule sans données d'origine humaine autres que les règles de base du jeu de Go. La performance semblait extraordinaire. Elle l'était au regard de la version précédente d'AlphaGo qui s'entraînait en exploitant plus de 150 000 parties de Go jouées par des maîtres du jeu et avec un double système de réseau de neurones pour l'évaluation des mouvements et des coups. Cette IA de troisième génération s'était, qui plus est, entraînée en seulement trois jours sur trois serveurs de type PC bien moins lourds que les configurations à plusieurs racks utilisées dans les deux versions précédentes, qu'elle battait à plates-coutures.

Je vais essayer de mettre cela en perspective avec d'autres avancées de l'IA, notamment en robotique. Et surtout, vis à vis de performances que DeepMind a mis en avant comme des applications potentielles d'AlphaGo Zero comme la simulation du repliement de protéines complexes, l'un des problèmes mathématiques les plus complexes à gérer qui soit aujourd'hui.

Mastering the Game of Go without Human Knowledge

David Silver¹, Julian Schrittwieser², Karen Simonyan³, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, Denis Hassabis.

DeepMind, 5 New Street Square, London EC4A 3DF.

¹These authors contributed equally to this work.

A long-standing goal of artificial intelligence is an algorithm that learns, *tabula rasa*, superhuman proficiency in challenging domains. Recently, AlphaGo became the first program to defeat a world champion in the game of Go. The tree search in AlphaGo evaluated positions and selected moves using deep neural networks. These neural networks were trained by supervised learning from human expert moves, and by reinforcement learning from self-play. Here, we introduce an algorithm based solely on reinforcement learning, without human data, guidance, or domain knowledge beyond game rules. AlphaGo becomes its own teacher: a neural network is trained to predict AlphaGo's own move selections and also the winner of AlphaGo's games. This neural network improves the strength of tree search, resulting in higher quality move selection and stronger self-play in the next iteration. Starting *tabula rasa*, our new program AlphaGo Zero achieved superhuman performance, winning 100-0 against the previously published, champion-defeating AlphaGo.

Much progress towards artificial intelligence has been made using supervised learning systems that are trained to replicate the decisions of human experts¹⁻⁴. However, expert data is often expensive, unreliable, or simply unavailable. Even when reliable data is available it may impose a ceiling on the performance of systems trained in this manner⁵. In contrast, reinforcement learning systems are trained from their own experience, in principle allowing them to exceed human capabilities, and to operate in domains where human expertise is lacking. Recently, there has been rapid progress towards this goal, using deep neural networks trained by reinforcement learning. These systems have outperformed humans in computer games such as Atari^{6,7} and 3D virtual environments⁸⁻¹⁰. However, the most challenging domains in terms of human intellect – such as the

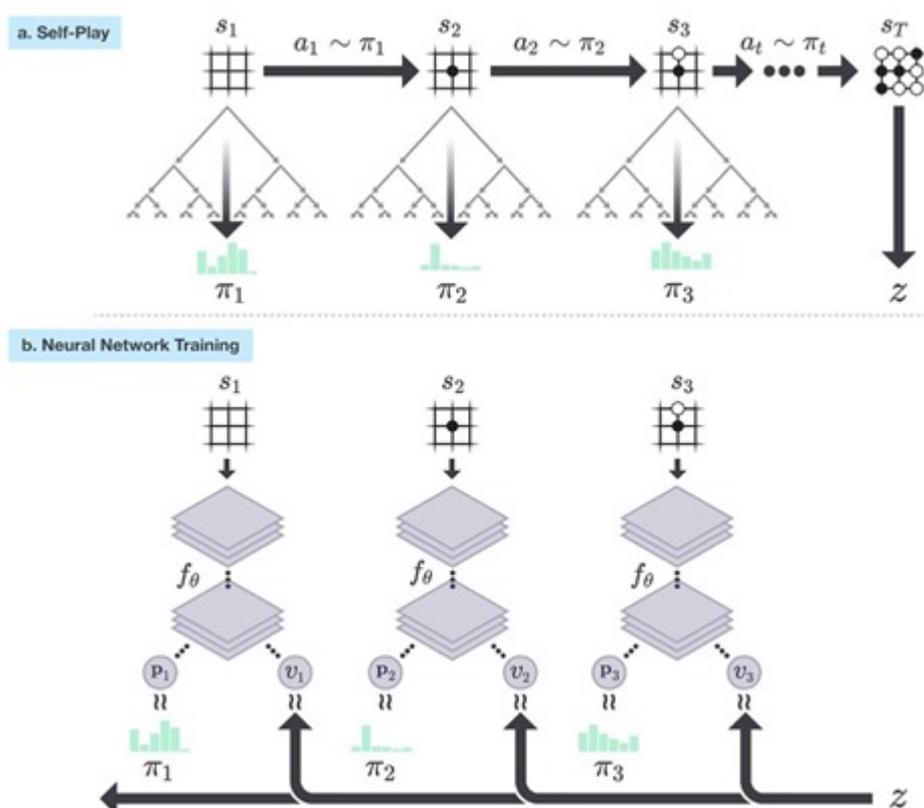
1

La performance de DeepMind AlphaGo Zero est bien documentée dans [AlphaGo Zero Learning from Scratch](#) ainsi que dans l'article de 43 pages correspondant publié dans Nature. Dans cette revue, l'accès à l'article est payant mais on le trouve aussi à l'identique sur le site de DeepMind : [Mastering the game of Go without human knowledge](#). Une analyse des parties d'AlphaGo Zero est même disponible sur YouTube.

Et voici le genre de réaction d'étonnement que la performance a pu générer : **Why AlphaGo Zero is a Quantum Leap Forward in Deep Learning** de Carlos Perez ou **La nouvelle Performance d'AlphaGo est DINGUE – Maintenant, il crée de la connaissance ex nihilo. Et ce que cela augure est profond**, de Vincent Pinte-Deregnacourt (20 octobre 2017). L'auteur se demande comment cette performance serait applicable aux lois de la physique, du droit ou des conventions sociales. Une très bonne question ! Qui alimente au passage le vieux débat sur l'Intelligence Artificielle Etroite, ad-hoc, qui n'est pas facile à répliquer dans d'autres domaines. Et nous allons voir pourquoi.

J'étais particulièrement intrigué par cette affirmation lue sur le site de DeepMind : *“While it is still early days, AlphaGo Zero constitutes a critical step towards this goal. If similar techniques can be applied to other structured problems, such as **protein folding**, reducing energy consumption or searching for revolutionary new materials, the resulting breakthroughs have the potential to positively impact society”*. Le repliement de protéines ? C'est le Graal absolu de la recherche biomédicale alors que gagner au jeu de Go ne sert pas à grand chose dans la vie courante, à part, remettre en question l'ego et le savoir pluri-millénaire des grands joueurs de Go en Asie ! Si DeepMind arrivait à faire progresser ce domaine, elle ferait faire des bonds immenses à la recherche dans la santé. Je vais y revenir en détail dans cet article.

Conceptuellement, et à haut niveau, l'approche d'AlphaGo Zero n'est pas trop complexe à comprendre. Le jeu de Go comprend des règles très simples de positionnement des jetons. Sa simulation est complexe car la combinatoire des coups est très élevée. Il est plus difficile qu'aux échecs de tester tous les coups possibles pour trouver le bon parcours de jeu. L'IA a donc pour mission d'élaguer l'arbre de décision des positions en supprimant le plus efficacement les branches mortes, celles qui ne peuvent pas mener à la victoire et en privilégiant celles qui ont le plus de chances d'y mener. Dans AlphaGo Zero, cet élagage s'appuie sur l'exécution d'un réseau de neurones d'apprentissage par renforcement pour chaque coup possible. Le système s'est d'abord entraîné contre lui-même et en jouant 4,9 millions de parties. Il a alors joué contre la version précédente d'AlphaGo Zero entraînée de son côté par une combinaison associant notamment l'exploitation de 150 000 parties existantes, d'origine humaine. Ici, le système était à la fois plus rapide et plus efficace sans exploiter ces parties préexistantes.



Bref, nous avons une IA qui apprend toute seule sans avoir besoin de données d'entraînement. Est-ce généralisable facilement ? Est-ce le début d'une IA généraliste ? En fait, on en est encore loin. Comment d'ailleurs fonctionne une IA qui n'a pas besoin de données d'entraînement ? Quelles sont les règles et les informations utilisées par l'IA ? Comment le monde est-il modélisé et simulable ? C'est une question d'univers fermé ou ouvert, et de problèmes discrets ou continus, bref, une question dépendant de la nature des règles et éléments mathématiques qui régissent le fonctionnement d'un système complexe.

Pour simplifier les choses, nous allons illustrer la question avec trois types de règles : les règles discrètes avec une combinatoire finie, les règles continues avec une combinatoire infinie, et les règles déduites implicitement des systèmes car ceux-ci ne sont pas modélisables par des règles explicites, ce qui est le cas pour ce qui relève du vivant en général et du comportement des consommateurs en particulier.

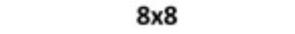
Règles discrètes et combinatoire finie

Ces règles discrètes sont exploitées dans les jeux de société (dames, échecs, Go). Le principe de base des solutions d'IA pour gagner à ces jeux consiste à élagage d'une manière ou d'une autre un arbre de décision de taille finie.

Pour sortir d'un labyrinthe parfait, on tourne toujours vers la gauche ou vers la droite et à un moment donné, la topologie du labyrinthe fait que l'on trouve la sortie. Ce n'est pas forcément optimal mais cela marche. La règle d'élagage est simple et même binaire. Dans le cas du jeu de Go, la règle est celle du placement des pions et de la manière de battre l'adversaire en entourant ses positions.

Dans les jeux de société comme le tic-tac-toe, les dames, les échecs, ou le Go, les règles définissent la grille de jeu, la variété des pions et les règles de leur déplacement. Le jeu de Go a des règles plus simples parce qu'il n'existe que deux types de pions, les blancs et les noirs, et une règle d'encerclément des pions adverses, mais la combinatoire est plus élevée qu'aux échecs, simplement parce que l'aire de jeu est plus grande, et les règles de positionnement des pions plus libres.

<u>COMPLEXITY NUMBERS</u>	
	Go 40
	Chess 14
	Scrabble 10
	Poker 10
	Backgammon 8
	Checkers 8
	Hearts 5
	Blackjack 2
	Craps 0.001
	Lotteries 0.0000001
	Roulette 0



8x8



19x19

Yann LeCun l'explique bien dans [Facebook's head of AI wants us to stop using the Terminator to talk about AI](#) (octobre 2017) : *“So, AlphaGo is using reinforcement learning. And reinforcement learning works for games; it works for situations where you have a small number of discrete actions, and it works because it requires many, many, many trials to run anything complex. AlphaGo Zero has played millions of games over the course of a few days or few weeks, which is possibly more than humanity has played at a master level since Go was invented thousands of years ago. This is possible because Go is a very simple environment and you can simulate it at thousands of frames per second on multiple computers. [...] But this doesn't work in the real world because you cannot run the real world faster than real time.”*

Certains s'amuse parfois à comparer le nombre de parties possibles du jeu de Go avec le nombre d'atomes dans l'Univers, ce qui est impressionnant mais n'a aucun sens. Ce ne sont pas des données comparables ! Une combinatoire de positions n'est pas homothétique à un nombre d'objets ou d'atomes ! L'association d'idée a

tendance à gonfler les capacités d'AlphaGo. Or une IA ne peut pas simuler des systèmes ultra-complexes à l'échelle atomique, comme un être vivant ou l'ensemble d'une planète. La combinatoire est alors plusieurs ordres de grandeur d'ordres de grandeur supérieurs à celle du jeu de Go !

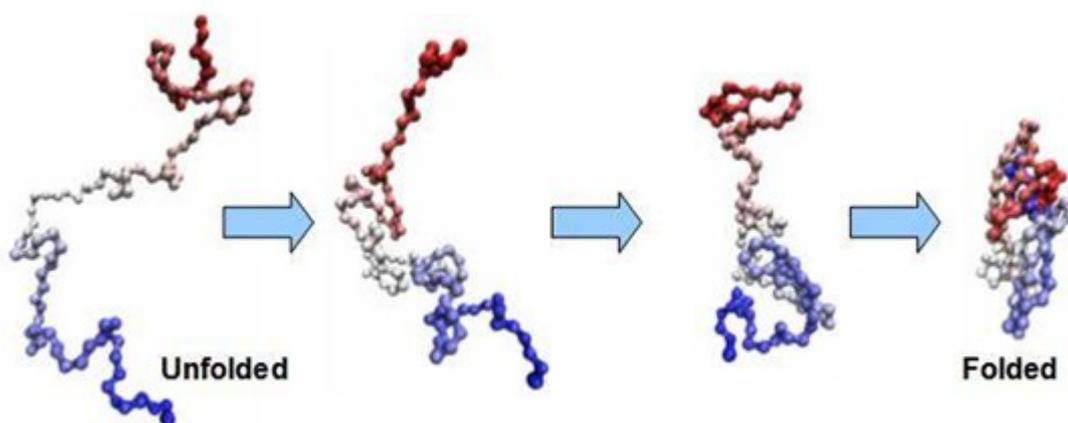
Il existe une variante de cette catégories de règles, toujours avec des règles discrètes et une combinatoire finie, mais intégrant une incertitude liée à une information incomplète sur le jeu des adversaires. C'est le cas du jeu de Poker. La performance de deux IA ayant gagné au Poker est de ce point de vue plus étonnante que celle d'Alpha Go Zero, avec d'un côté **Libratus** et de l'autre **DeepStack**. Toutes les deux ont été annoncées en 2017 mais leurs auteurs sont moins efficaces en marketing que Google et DeepMind !

Les problèmes de ce type ne sont pas nombreux dans la vie pratique. Il s'agit essentiellement des jeux de société, les premiers systèmes auxquels se sont attaquées des générations entières de chercheurs en IA.

Il en existe cependant qui font partie du champ de l'IA depuis sa naissance, et surtout depuis les années 1980 : les systèmes experts à base de moteurs de règles et de solveurs ! Ceux-ci exploitent des règles en général discrètes et une combinatoire finie de ces règles pour trouver une solution optimale à partir d'une question posée. Les mécanismes de résolution de problèmes font partie des solveurs, qui exploitent des techniques dites de chaînage avant ou arrière. Ils n'utilisent pour l'instant pas encore les réseaux de neurones ou le deep learning car il semble que ce n'est pas efficace. Certains moteurs de règles peuvent être alimentés par des règles identifiées de manière empirique à partir de l'analyse de données terrain et de techniques de machine learning pour les données et de deep learning pour ce qui est du traitement du langage.

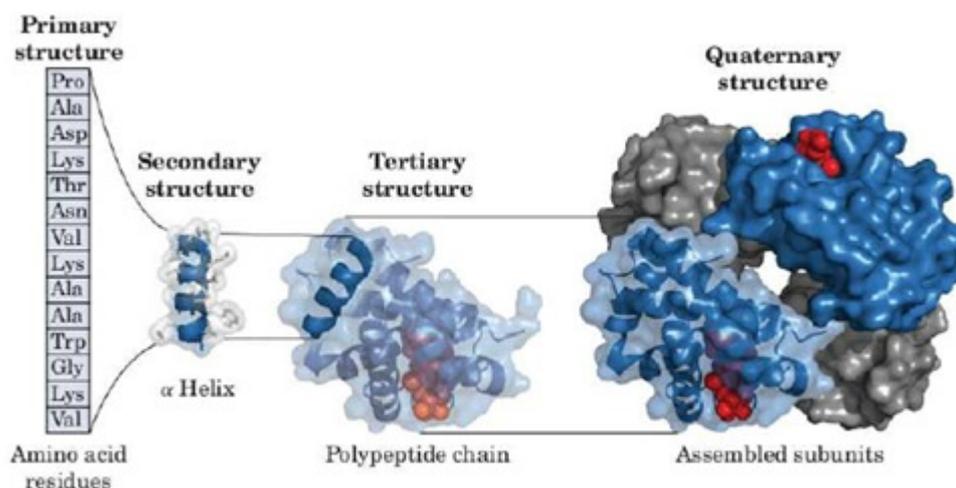
Règles continues et combinatoire infinie

Un second type de problèmes de combinatoires à traiter est celui de règles que l'on peut qualifier de "continues". On les retrouve en particulier dans la physique des matériaux, la chimie, et le fameux repliement des protéines dans leur forme tertiaire voire quaternaire qui est évoqué à la fin de l'article de DeepMind sur AlphaGo Zero. Je vais creuser un peu ce sujet.



Les protéines sont fabriquées par les ribosomes dans les cellules et à partir de l'assemblage en continu de 21 acides aminés différents. L'encodage de ces acides aminés provient de la lecture de brins d'ARN messagers issus de la copie de morceaux de l'ADN des chromosomes que l'on appelle les gènes. Chaque acide aminé est relié au précédent par une liaison chimique avec un degré de liberté qui se mesure par une orientation spatiale et un angle de rotation. Un nouvel acide aminé va former un angle avec le précédent dans un sens ou dans l'autre en fonction des forces d'attraction ou de répulsion des autres atomes de la protéine sachant que celles-ci vont évoluer au gré de la création de la protéine après l'ajout de cet acide aminé. Les acides aminés ont naturellement tendance à créer des structures géométriques comme des hélices dans la structure secondaire des protéines. Dans la structure tertiaire, la protéine atteint son équilibre géométrique et dans la structure

quaternaire, des entités séparées sont assemblées pour créer une molécule complexe [voir cette intéressante explication dans **How to Fold (and Misfold) a Protein** et sa seconde partie, de Sarah Kearns, février 2017].



C'est le cas des histones autour desquelles l'ADN s'enroule dans les chromosomes, des ribosomes qui sont un assemblage de plus de 70 protéines et de quatre brins complexes d'ARN et de tout un tas de protéines complexes que l'on trouve dans les cellules ou de l'ATPase, une protéine complexe située dans les mitochondries qui convertit l'ADP en ATP, cette dernière étant la source énergétique principale du métabolisme des cellules.

Le repliement d'une protéine va conditionner sa fonction, à savoir son interaction avec les autres protéines et enzymes des cellules. Un déficit de certaines protéines peut être à l'origine de la maladie d'Alzheimer. Savoir comment créer des gènes qui fabriqueront des protéines pouvant s'attaquer à différentes pathologies extra- ou intra-cellulaires est donc très intéressant. On peut se définir une cible de protéine à attaquer dont on connaît la forme primaire mais pas forcément tertiaire. Il faut donc pouvoir la déduire de sa formule qui est une suite d'acides aminés conditionnée par le code ADN de ses gènes, pour déterminer les zones à toucher et ensuite concevoir une protéine de synthèse via du génie génétique, et l'implanter éventuellement dans les chromosomes de cellules vivantes grâce aux fameux ciseaux à ADN CRISP-Cas9. Pour fabriquer cette nouvelle protéine, il faut pouvoir en déterminer sa composition et son repliement. Donc, en gros, pour créer une thérapie à base de protéine de synthèse, il faut idéalement pouvoir simuler et la ou les protéines ciblées et la protéine d'attaque. Aujourd'hui, on ne sait pas le faire et tout relève d'expérimentations laborieuses et de bio-mimétisme.

Les lois physiques associées au repliement des protéines sont connues au niveau atomique avec les liaisons covalentes où des atomes partagent des électrons comme dans les liaisons hydrogène-hydrogène ou hydrogène-oxygène, les attractions-répulsions ioniques ou les forces de Van der Vals qui régissent la manière dont les parties des protéines interagissent entre elles. Ces forces sont antagonistes et leur équilibre est difficile à déterminer par un simple calcul. La forme que prend une protéine est déterminée par la recherche d'un équilibre énergétique optimum sachant qu'elle est baignée dans un environnement en général aqueux, qui ajoute des interactions ioniques diverses avec les molécules d'eau, certaines parties des protéines étant hydrophobes et d'autres hydrophiles [voir **Solving for Minimum Energy Structures with Neural Networks** de Brian Barch and Norman Tubman, 2017].

On arrive aujourd'hui à simuler numériquement le repliement d'à peine quelques acides aminés de peptides, qui sont des petites protéines. Or les protéines qui constituent le gros du fonctionnement de nos cellules et de leurs

membranes peuvent comprendre plusieurs centaines voire milliers d'acides aminés.

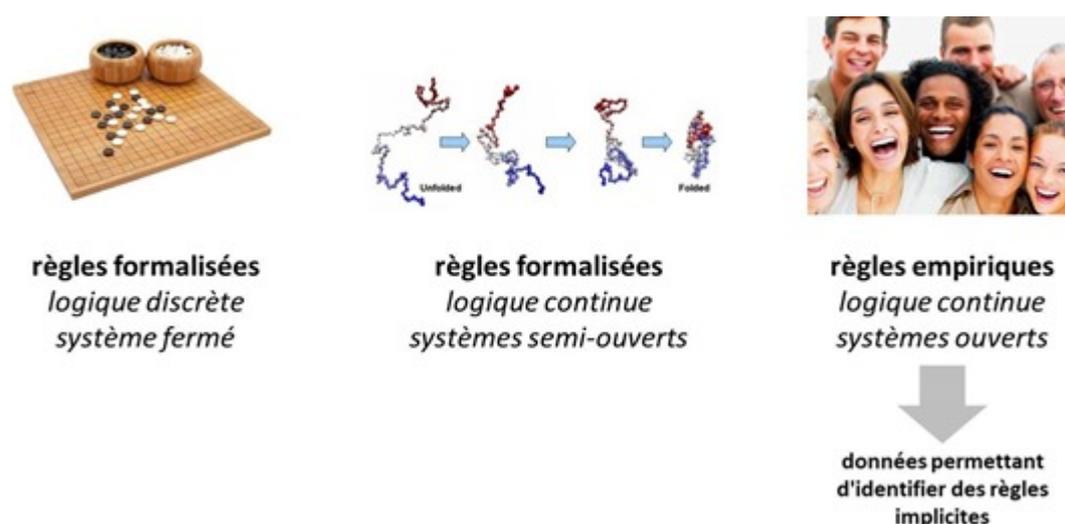
La détermination du repliement des protéines ne relève pas de la recherche dans un arbre de décision comme dans le jeu de Go. On peut éventuellement s'appuyer sur un réseau de neurones de deep learning, mais on ne sait pas forcément comment l'entraîner car on manque de données d'exemples. De nombreux chercheurs travaillent sur cette épineuse question et depuis des années, avec des progrès réguliers.

Il existe une autre voie, plus long terme. Ce genre de simulation sera peut-être un jour à la portée d'ordinateurs quantiques qui présentent la spécificité de pouvoir simuler plusieurs états simultanément dans un système complexe, et pas nécessairement avec des modèles mathématiques du type de ceux du deep-learning. Nous aurons l'occasion d'en reparler en 2018 car le sujet est florissant et passionnant.

Bref, le repliement des protéines est un sujet épineux et stratégique. Si DeepMind arrive à le traiter, ce sera très bien et même révolutionnaire avec des implications incommensurables dans le secteur de la santé. Mais DeepMind ne le fera pas forcément avec la même méthode que celle qui est mise en oeuvre dans AlphaGo Zero.

Règles déduites implicitement dans des systèmes ouverts et complexes

Dans le monde du vivant et dans l'environnement, les forces élémentaires de la physique se combinent avec des niveaux de complexité et de combinatoire très élevés. Que ce soit pour conditionner les comportements des individus, les évolutions de la météo, les maladies ou de nombreux autres phénomènes naturels.



Les logiciels et l'IA ne peuvent pas facilement simuler ce monde réel car il est trop complexe. Ils se contentent donc d'identifier des règles inconnues par l'exploitation des données issues de l'observation de l'environnement et avec divers capteurs. Le machine learning et le deep learning identifient ces règles implicites qui associent par un exemple un profil type de client à un comportement, ou une image d'objet à son nom.

Dans la biologie, le machine learning permet de déterminer d'éventuelles corrélations entre les mutations du génome et l'apparition de cancers divers, ou encore, entre les habitudes alimentaires, la pollution et l'apparition de ces mêmes cancers. Ce sont des corrélations mais pas forcément des causalités, ce qui explique les limites de la médecine dite prédictive. Elle prévoit éventuellement mais ne permet pas forcément de traiter les pathologies. Pour guérir les patients, il faut inventer des traitements et les corrélations ne suffisent pas. On peut juste limiter les facteurs de risques.

Pour revenir à mon obsession du repliement des protéines, des méthodes d'évaluation du repliement des

protéines en structure tertiaire utilisent des réseaux de neurones récurrents et à mémoire qui utilisent des règles empiriques (RNN, LSTM, voir **Deep Recurrent Neural Network for Protein Function Prediction from Sequence** de Xueliang Leon Liu, 2017). Ces réseaux de neurones sont entraînés avec des bases de protéines dont on connaît déjà la structure tertiaire, qui a été identifiée avec des techniques d'observation utilisant la cristallographie à rayons X, la même technique utilisée par **Rosalind Franklin** en 1953 pour déterminer la structure en hélice de l'ADN, expliquée dans **cette vidéo**.

Ce sont des méthodes probabilistes et pas de la simulation des lois physiques gouvernant la position des atomes des molécules les uns vis à vis des autres. Dans ce cas-là, les règles de repliement des protéines sont donc déduites de manière implicite par l'observation. La science est construite à partir d'une combinaison d'observations, d'élaboration de règles et théories élaborées a priori ou vérifiées a posteriori par l'observation, et affinées si besoin est.

Les IA peuvent apprendre par elles-même dans un environnement ouvert en interagissant avec le monde physique, ce que font certains robots avancés qui tâtonnent et apprennent de manière autonome à manipuler des pièces. C'est de l'apprentissage par renforcement applicable à un monde ouvert.

Enfin, il existe des problèmes tellement complexes qu'ils n'ont pas de solution formelle, relevant du principe d'indécidabilité du théorème de Gödel. C'est par exemple le cas de l'optimisation parfaite d'un mix marketing. Cette optimisation parfaite est impossible et indécidable car l'environnement est très dynamique, il comporte trop d'inconnues et l'univers du problème n'est pas facilement simulable. Les outils et leurs règles changent sans cesse. Cela impose de se mettre en permanence en posture d'expérimentation. On peut mesurer l'efficacité d'une solution prise isolément, mais pas optimiser l'usage de l'ensemble des solutions dans leur environnement.

L'IA, l'Homme et les règles complexes

Arriver à distinguer le niveau de performance de l'IA comparée à l'Homme est un épineux débat. Il existe des domaines où elle dépasse déjà largement les capacités humaines (*ci-dessous*, en vert) et d'autres où elle est encore en-deca.

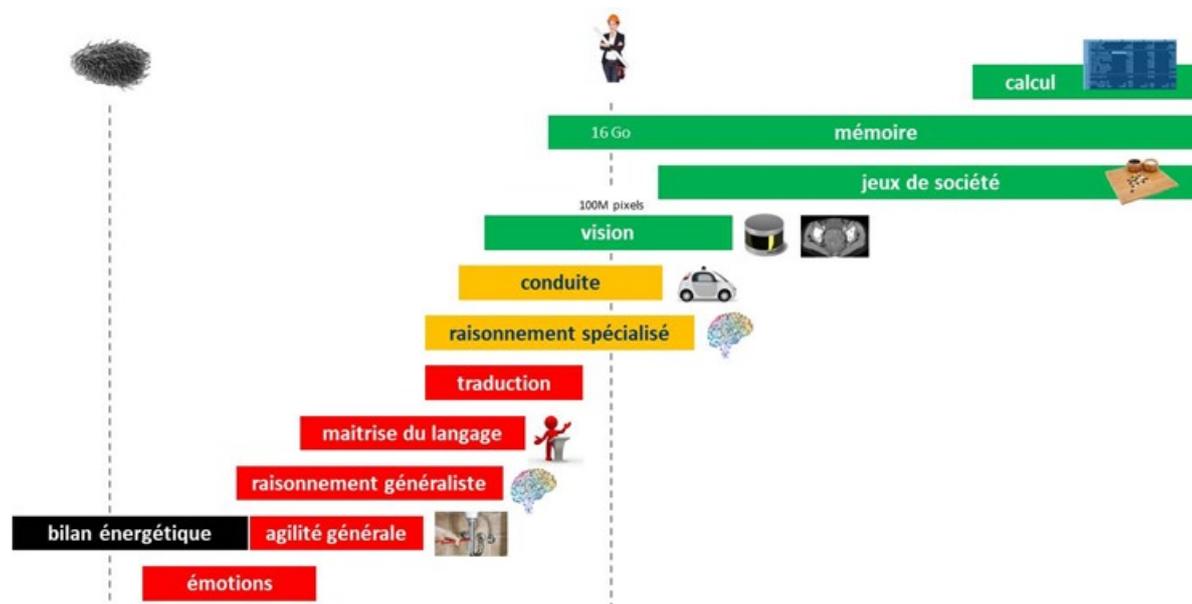
Les jeux sont un domaine où les IA sont déjà devant l'Homme car ils fonctionnent sur des règles simples où la combinatoire est certes élevée mais finie. La combinaison de la force brute des machines et des techniques probabilistes d'élagage d'arbre de décision font l'affaire.

La vision est un domaine où l'IA dépasse déjà les capacités humaines, notamment dans la vision spécialisée que constitue l'interprétation de résultats d'imagerie médicale. Et encore, ce niveau de performance est atteint, comme dans celui de la conduite autonome, qu'avec des images dont la résolution est généralement réduite à 224×224 pixels. L'œil capte de son côté 100 millions de pixels par œil ! Cela crée un énorme décalage qualitatif. Lorsque la puissance des prochaines générations de processeurs neuromorphiques augmentera, cela permettra de faire tourner des réseaux de neurones convolutionnels (ou convolutifs, les deux appellations étant couramment utilisées) avec une bien meilleure résolution. Et la vision artificielle s'améliorera encore !

Le traitement du langage est le domaine qui est encore en retrait par rapport aux capacités humaines, même la traduction automatique qui est encore très imparfaite. Mais si vous ne connaissez pas le japonais, le traducteur français-japonais est évidemment meilleur que vous ! Il faut donc distinguer la comparaison entre l'IA et un humain en particulier et avec l'ensemble des humains en général. Ce d'autant plus que le savoir ingéré par les IA est presque toujours un savoir composite issu d'un grand nombre d'humains.

Côté énergétique, l'Homme et la vie en général ont encore quelques points d'avance. Le cerveau ne consomme que 20 W et le corps environ 100 W au repos. Une IA ne fait pas grand chose avec seulement 20 W, qui correspondent à la consommation électrique d'un laptop convertissant des photos RAW en JPEG sous

Lightroom (*c'est précis isn't it ?*). Les trois serveurs d'AlphaGo Zero doivent probablement consommer au minimum 1000 à 2000 W. Les premières moutures consommaient plus de 20 KW.



Ce qui est indiqué en rouge dans le schéma n'est cependant pas *théoriquement* inaccessible aux IA. Il est probable que des progrès significatifs seront réalisés dans les années à venir, au moins dans la traduction, la maîtrise du langage et le raisonnement généraliste. L'agilité physique générale en robotique fera aussi des progrès. Ceux-ci ne dépendent pas que de l'IA. Ils sont aussi liés aux progrès dans la mécanique, les capteurs, les moteurs et les batteries, qui sont plus lents à émerger que dans l'IA.

Ces domaines où l'IA est encore en retrait par rapport à l'Homme correspondent finalement aux aspects de la vie qui ne reposent pas sur des règles simples ! Ce qui permet de boucler en beauté ce petit papier de circonstance en vous rassurant un peu sur votre devenir !

Cet article a été publié le 9 novembre 2017 et édité en PDF le 23 mars 2024.
(cc) Olivier Ezratty – “Opinions Libres” – <https://www.oezratty.net>