



Opinions Libres

le blog d'Olivier Ezratty

Que devient l'IA symbolique ?

L'Histoire scientifique de l'intelligence artificielle recèle une rivalité très ancienne entre ses deux principaux courants, celui des symbolistes et des connexionnistes. Elle perdure encore aujourd'hui, ces derniers ayant pour l'instant nettement pris le dessus après un jeu de chassé-croisé qui a duré six décennies.

Les enjeux de cette rivalité intellectuelle sont encore vivaces avec, à la clé, l'évolution de l'un des domaines les plus complexes de l'IA : le raisonnement automatique. Il fait partie des points de blocage avant de créer la mythique intelligence artificielle générale (en anglais, AGI, pour Artificial General Intelligence) capable d'imiter puis de dépasser les capacités de raisonnement généralistes de l'Homme. A vrai dire, l'IA symbolique est dans un hiver prolongé depuis plus de 15 ans. Les étés et les hivers de l'IA alternent depuis six décennies ceux de l'IA connexionniste et de l'IA symbolique !

Je vais examiner ici l'histoire de cette dualité et ses enjeux actuels en montrant que ces deux courants ont tout intérêt à se rapprocher, tout en relevant les limites d'une éventuelle AGI miracle.



résumé pour les lecteurs pressés

1. Les **deux principaux courants** de l'IA sont l'IA symbolique et l'IA connexionniste. Leurs étés et hivers alternent dans l'histoire de l'IA depuis 1955.
2. L'**IA symbolique** gère le raisonnement formel et la logique. On la retrouve notamment dans les moteurs de règles et les systèmes experts. Elle est démontrable car sa logique est décorticable. Elle est en sourdine depuis plus de 15 ans.
3. L'**IA connexionniste** est incarnée par le machine learning, les réseaux de neurones et le deep learning, qui s'entraînent à partir de données et sont des approches probabilistes, très efficaces mais moins faciles à expliquer.
4. L'**IA connexionniste domine les avancées récentes de l'IA**, pour l'analyse de données, le traitement du langage et des images. On y intègre des prémisses de raisonnement.
5. L'**IA symbolique peut exploiter l'IA connexionniste** qui aide à découvrir des règles et des faits dans les données non structurées.
6. L'**IA généraliste (AGI)** est le prochain enjeu stratégique de la recherche en IA. Il nécessite de rapprocher l'IA symbolique et l'IA connexionniste. La France y a une carte à jouer.
7. L'**IA n'est pas la solution universelle** à la résolution de tous les problèmes complexes. Les outils de simulation la complètent, l'informatique quantique pouvant élargir son champ.

Et comme inauguré dans l'article précédent sur la French Tech, je vous résume ces pages en 7 points synthétiques. Le débat peut commencer ! Je me dois aussi de préciser que ce genre d'article est le résultat, non pas d'une quelconque expertise, mais de questionnements, puis de découvertes de l'état de l'art en consultant des ressources ouvertes sur Internet et en s'appuyant sur quelques discussions sporadiques avec des spécialistes.

Symbolisme et connexionnisme

Lors des débuts de l'IA comme science à la fin des années 1950 et pendant les années 1960, les deux approches se valaient.

Les symbolistes comprenaient tout un tas de chercheurs travaillant sur la résolution de théorèmes comme **Herbert Gelernter** (créateur d'un système de démonstration de théorèmes géométriques en 1959), **Allen Newell** et **Herbert Simon** (inventeurs du General Problem Solver en 1957), **John McCarthy** (inventeur de l'appellation d'intelligence artificielle en 1955 ainsi que du langage LISP en 1958), **James Slagles** (concepteur du premier système expert, SAINT, de traitement de formules mathématiques, en 1961), **Thomas Evans** (et son programme ANALOGY de 1963 qui pouvait résoudre des problèmes de tests de QI) et aussi les Français **Alain Colmerauer** et **Philippe Roussel** (inventeurs du langage Prolog en 1972). Il s'agissait des débuts des tentatives d'émuler le raisonnement humain dans des machines.

démonstrations de théorèmes

un des premiers champs d'investigation de l'IA symbolique, au départ en géométrie

Geometry Theorem Prover

Herbert Gelernter, 1959

General Problem Solver

Allen Newell et Herbert Simon, 1959

Integration Problems Solver (SAINT)

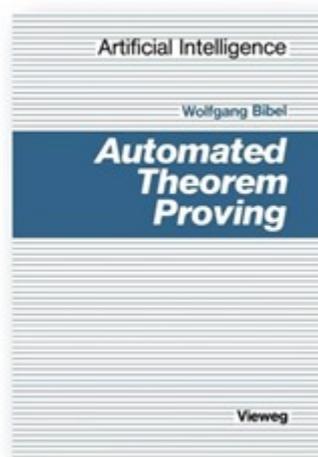
James Slagles, 1963

Algebra Problems Solver (STUDENT)

Daniel Bobrow, 1967

Geometric Analogy Problems (ANALOGY)

Tom Evans, 1968



Dans le camp des connexionnistes, on avait surtout **Frank Rosenblatt** avec ses perceptrons, les prémisses des réseaux de neurones d'aujourd'hui. Leurs limitations techniques soulevées en 1969 par Marvin Minsky couplées aux faibles capacités matérielles disponibles les empêchèrent de percer avant le 21^e siècle.

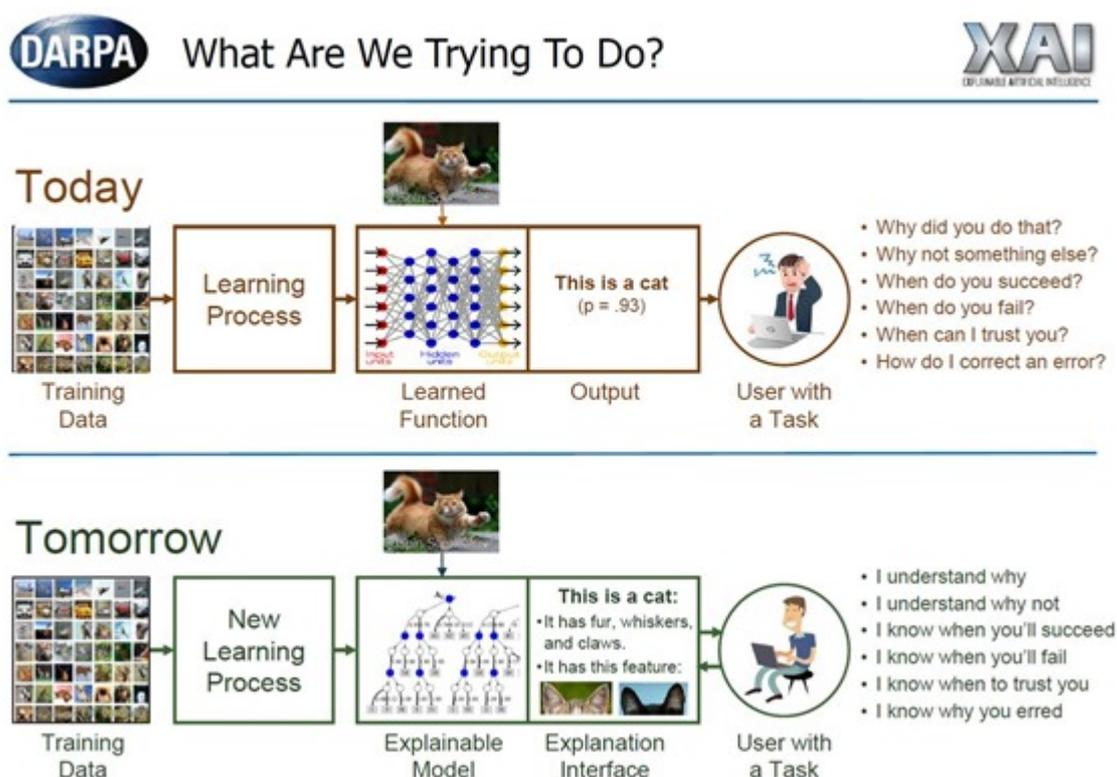
La dichotomie entre connexionnistes et symbolistes prit d'autres appellations sémantiquement voisines, comme celle séparant les **neats** des **scruffies**, littéralement "soignés vs débrillés", formulée par l'Américain **Robert Schank** pendant les années 1970.

Les neats s'appuient sur de la logique formelle et l'exploitation de connaissances existantes pour résoudre des problèmes, la base étant la logique aristotélicienne. Les scruffies utilisent des méthodes plus empiriques basées sur l'observation et l'exploitation des sens, sur des méthodes statistiques, sur la reconnaissance de formes et ce qui est devenu le machine learning.

Les débats des années 1970 portaient notamment sur les mérites comparés des deux approches dans le traitement du langage. Pouvait-on faire avancer le domaine en décrivant une structure formelle du langage ou avec des méthodes plus empiriques basées sur l'analyse de textes ? Les 10 dernières années ont clairement donné l'avantage aux méthodes empiriques du deep learning

et leurs méthodes probabilistes.

L'opposition entre symbolisme et connexionnisme porte aussi sur le fait que les approches symboliques sont documentables et prouvables tandis que les approches connexionnistes seraient empiriques et pas forcément démontrables. D'où le débat qui a court en ce moment sur l'explicabilité des algorithmes, ou plutôt sur leur inexplicabilité, relative à l'approche connexionniste exploitée dans les réseaux de neurones et dans le deep learning. De nombreux travaux de recherche contemporains visent à rendre le deep learning explicable, ou tout du moins décorticable. C'est même une priorité des projets d'IA financés par la DARPA aux USA, sous l'appellation XIA, pour Explainable IA (exemple *ci-dessous* pour la reconnaissance d'images).



L'autre appellation courante de l'**IA symbolique** est la GOFAI pour *Good Old Fashion AI*. D'un point de vue pratique, elle s'appuie notamment sur des moteurs de règles et de faits, qui permettent notamment, mais pas seulement, de créer des systèmes experts ou de faire de la programmation par contraintes. Au même titre que l'IA il y a une dizaine d'année, l'appellation de système expert est devenue un gros mot, évité par une bonne partie des spécialistes de l'IA.

La terminologie de ce domaine intègre les notions de logiques d'ordre zéro, du premier ordre (dite calcul des prédicats) et second ordre qui sont liées à la complexité des problèmes logiques à résoudre au niveau d'objets et d'ensembles d'objets. Essayer d'en comprendre les tenants et aboutissants génère un beau mal de crâne !

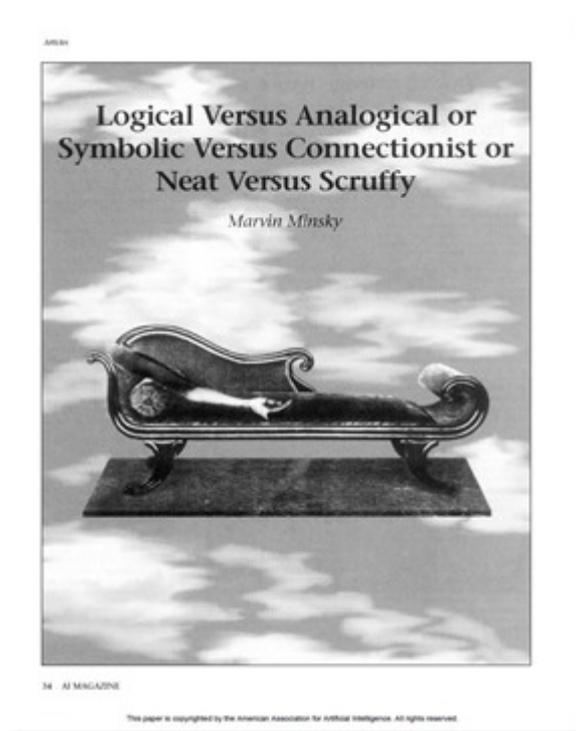
On peut aussi y associer le concept de **logique floue** de Lotfi Zadeh, élaborée dans les années 1960, même si celui-ci présente la particularité d'être associable aussi bien à du raisonnement formel et à du machine learning (cf. **Fuzzy Logic in Machine Learning** de Eyke Hüllermeier, 2017).

L'IA symbolique exploite la manipulation logique de symboles. Ces opérations servent à répondre à des questions de logique formelle pour résoudre des problèmes associant des règles, faits et

raisonnements connus du savoir humain. Les solveurs exploitent des méthodes de chaînage arrière ou de chaînage avant pour trouver la solution à des problèmes logiques. On parle aussi de logique déductive et explicite. Les systèmes de raisonnement non monotoniques servent de leur côté à résoudre des problèmes complexes pour lesquels l'information disponible est incomplète ou changeante. Ce sont des domaines de recherche assez abscons et encore plus difficiles à appréhender que les réseaux de neurones !

Du côté de l'industrie, l'IA symbolique a connu son heure de gloire pendant les années 1980 avec la mode des systèmes experts, qui servaient notamment à créer des solutions de maintenance ou de résolution de problèmes complexes. Puis ce fut le déclin, explicable par l'inadéquation de ces méthodes pour le traitement du langage qui était et est toujours un grand champ de recherche de l'IA, et aussi parce qu'il était difficile de collecter et structurer les règles du savoir humain.

Pour Marvin Minsky dans **Logical vs Analogical or Symbolic vs Connectionist or Neat vs Scruffy** publié en 1991 (19 pages), les méthodes symboliques manquaient de souplesse à force de vouloir absolument être prouvables ! Depuis une dizaine d'années, les systèmes experts sont des "RBMS" pour Rules Based Management Systems et intégrés dans les systèmes d'aide à la décision d'entreprises. A l'échelle du grand public, c'est du marketing et du branding de poussière rangée sous le tapis !



Pour sa part, l'**IA connexionniste** décompose les problèmes en sous-éléments connectés entre eux. Elle exploite par biomimétisme des neurones biologiques et le fonctionnement empirique et expérimental de la nature. La discipline a connu une renaissance depuis une vingtaine d'années avec les réseaux convolutionnels, le deep learning et notamment, son application dans le traitement du langage et de la vision. Elle met en place des techniques parallélisant les problèmes et utilisant de la logique inductive mais peu explicitable.

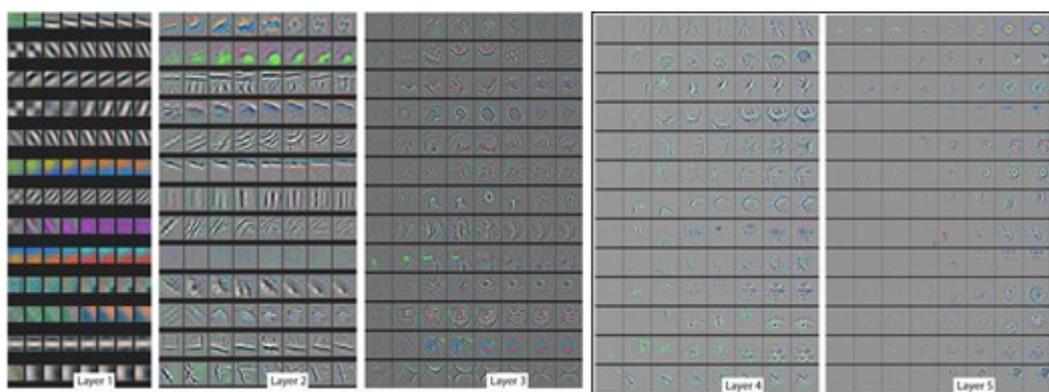
Les solutions d'IA connexionnistes exploitant de gros volumes de données cherchent notamment à découvrir les règles implicites contenues dans les données, par exemple pour reconnaître la nature des objets présents dans une image. Le connexionnisme mis en œuvre dans les réseaux de neurones et le deep learning est assimilable à des méthodes probabilistes. Lorsque ses solutions

donnent un résultat, il n'est pas formel mais exprimé le plus souvent sous la forme d'un pourcentage de véricité. Ainsi, un chat n'est jamais reconnu formellement dans une image, mais avec un pourcentage de 90% ou plus. Le pourcentage est acceptable lorsqu'il est inférieur aux erreurs humaines, voir nettement inférieur à ces erreurs quand il s'agit de systèmes critiques comme la conduite autonome.

Il en va de même pour la traduction d'une phrase, qui est d'ailleurs souvent approximative car elle repose aussi sur des approches également probabilistes. Si les données qui alimentent ces outils ne représentent pas bien l'espace du possible, leurs réponses pourront être inexactes ou, tout du moins, statistiquement ambiguës ou biaisées.

Certaines méthodes connexionnistes permettent d'extraire des règles explicites de contenus. C'est le cas dans le traitement du langage qui peut extraire des règles ou des données structurées contenues dans des textes.

L'ajustement à l'entraînement des paramètres des neurones dans un réseau de deep learning permet de générer de bons résultats, comme pour l'étiquetage d'objets dans des images (ou "labellisation", qui est un anglicisme, ou "tagging" en anglais). Leur inexplicabilité tient à ce que l'étiquetage n'est pas facilement décomposable en règles et sous-règles compréhensibles par l'Homme. Les règles d'un réseau de neurones sont déduites empiriquement des données d'entraînement et sont enfouies dans ses paramètres !



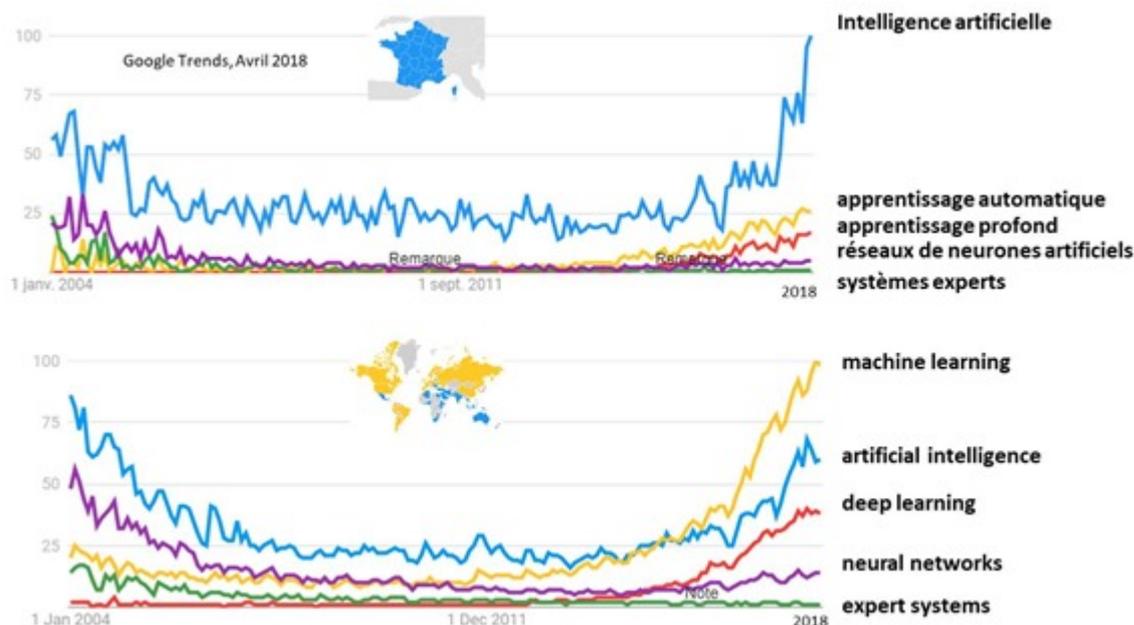
Dans les réseaux convolutifs (ou convolutionnels), ces règles sont réparties dans des filtres et des couches de neurones qui peuvent contenir des milliards de paramètres numériques. Ces filtres qui évoluent pendant la phase d'entraînement des réseaux ont une forme qui n'est généralement pas reconnaissable par l'œil humain. On peut le constater dans l'exemple *ci-dessus* de cette évolution pendant la phase d'entraînement sur les cinq couches d'un réseau convolutionnel.

Pour en savoir plus sur ce débat historique entre connexionnisme et symbolisme, vous pourrez vous référer à **Artificial Intelligence : Connectionist and Symbolic Approaches**, de Ron Sun, publié en 2000 (17 pages). Il positionne bien le débat, même si le document date d'avant le tsunami actuel du deep learning. Vous pouvez aussi consulter cette belle histoire de l'IA pré-deep learning de Nils Nillson, **Quest for Artificial Intelligence**, de 2010 (707 pages, en particulier des pages 199 à 207, 285 à 303, 485 à 469).

Google Trends de l'IA

Dans une analyse temporelle exploitant l'historique des recherches sur Google via Google Trends, on peut observer l'évolution de l'intérêt pour ces différentes techniques. Dans le monde (en bas), le machine learning a pris le dessus de l'IA et du deep learning. Les systèmes experts

sont très bas, même il y a plus de 10 ans, leur période dorée s'étant éteinte à la fin des années 1980. En France, l'appellation d'intelligence artificielle a toujours eu le dessus sur les autres. Je me rappelle pourtant qu'il y a moins de 10 ans, l'IA était mal vue, c'était un truc de loser !



Comme toutes les données, celles-ci présentent peut-être un biais. La population utilisant Internet en 2004 n'avait pas la même distribution qu'aujourd'hui. Le nombre d'Internautes est passé de moins d'un milliard de personnes en 2004 à plus de 3,5 milliards en 2018.

Quand le symbolisme exploite le connexionnisme

Dans le traitement du langage, l'IA connexionniste peut servir à identifier des règles présentes dans les textes et dans les données et l'IA symbolique peut ensuite les exploiter pour aider à décortiquer un problème complexe et prendre une décision. La nature et la vie humaine fonctionnent de la sorte.

Nous les humains procédons souvent par tâtonnement pour trouver la bonne méthode pour résoudre tel ou tel problème de la vie comme pour trouver un travail, exercer un métier, communiquer, éduquer ses enfants ou encore faire fonctionner son ordinateur. On apprend aussi par transmission de règles. C'est d'ailleurs le rôle des religions comme des lois, des constitutions ou de l'éducation dans l'organisation de la vie humaine.

Par analogie, pour entreprendre, on peut apprendre par la lecture (symbolisme) ou par l'expérience (connexionnisme). Les deux sont complémentaires. Les livres et le savoir formalisé permettent de gagner du temps en évitant de faire trop d'erreurs par l'expérience, comme avec les fameux pivots des startups. L'expérience permet d'explorer l'inconnu qui est le propre de l'innovation. Dans les startups, le savoir déjà collectivement acquis porte sur les fondamentaux connus comme la création d'une équipe, la structuration d'une idée, d'un business plan et d'une approche marketing. Le savoir à créer est lié à la confrontation de la nouveauté au marché. Ce dernier est difficile à modéliser au préalable parce qu'il est trop complexe et imprévisible malgré le meilleur des bons sens.

De même, on n'apprend pas aux enfants en bas âge à traverser la rue en les laissant y aller au hasard en espérant qu'ils ne se fassent pas écraser et que les adultes découvrent par

l'observation les erreurs des enfants (connexionnisme puis apprentissage par renforcement). On transmet donc aux enfants les règles apprises consistant à regarder des deux côtés de la route, à traverser sur les passages piétons et à respecter les feux. Cela permet de gagner du temps et de sauver des vies (symbolisme) !

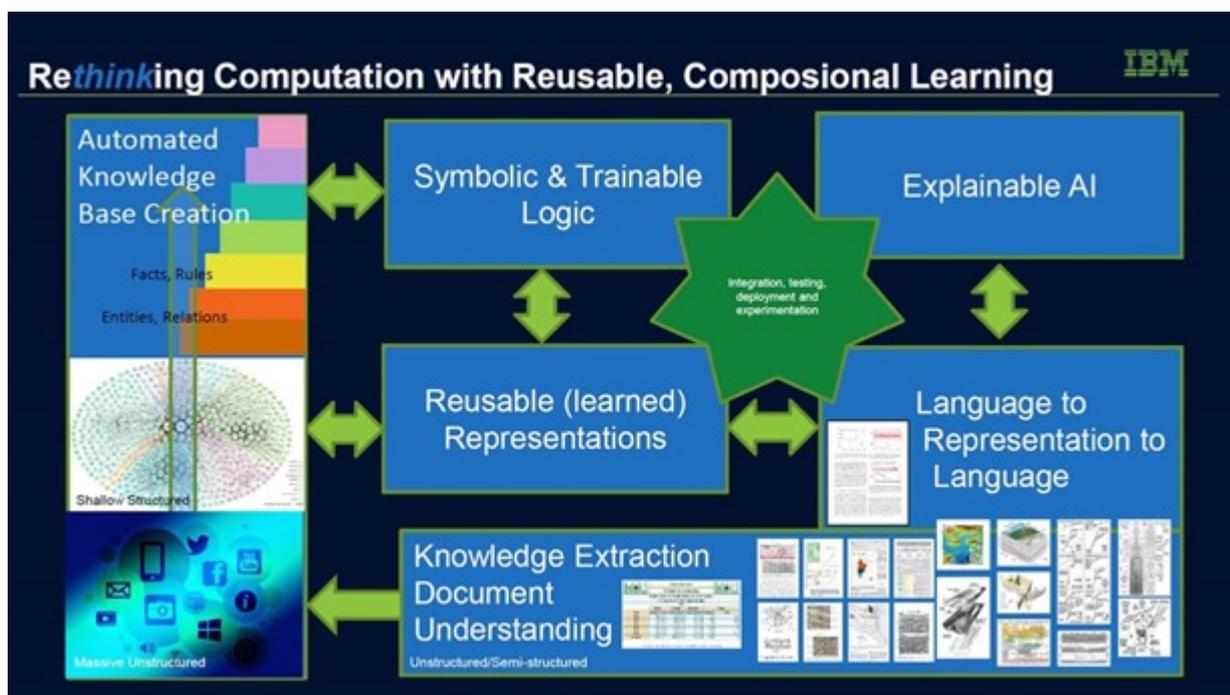
La question se pose aussi pour ce qui est des fameux dilemmes des véhicules autonomes qui font l'objet de débats sur le rôle de l'IA. En cas d'accidents inévitables, doivent-ils écraser une vingtaine de repris de justice ou la maman avec son enfant ? Il s'agit d'un jugement moral qui ne peut être réglé que par des règles (symbolisme). On peut éventuellement faire voter des citoyens ou leurs représentants pour décider de la règle à appliquer (une forme de connexionnisme). Mais le nombre de cas possibles peut être très élevé, rendant la formalisation difficile. Heureusement, les chances que cela puisse arriver sont bien rares.

La connaissance humaine symbolique est faite d'accumulation de faits et règles, parfois collectées laborieusement et sur une longue durée. Un peu comme lorsque les chasseurs cueilleurs ont dû apprendre à distinguer les baies, fruits et champignons tueurs de ceux qui étaient comestibles. L'exploitation de règles connues est théoriquement plus économe que leur recherche expérimentale permanente lorsqu'elles sont déjà connues. Ce qui rappelle parfois que les méthodes de machine learning d'analyse de données sont parfois à ranger dans la catégorie des marteaux pour écraser des mouches, quand leurs résultats relèvent d'un savoir commun déjà acquis. Au minimum, cependant, elles permettent de quantifier les résultats.

Le connexionnisme humain sert donc à traiter les flux de l'apprentissage tandis que le symbolisme exploite le stock de connaissances et expertises acquises, qu'il peut intégrer dans les raisonnements, même si ceux-ci sont très empiriques et d'une rationalité limitée par les émotions et autres pulsions. Comment apprend-on ? Soit de manière symbolique, en apprenant directement des règles connues, soit en observant le monde et en déduisant des règles de l'observation et l'expérimentation.

L'apprentissage passe aussi par l'action. Un enfant apprend à tenir une fourchette en tâtonnant. Il déforme son cerveau, plus précisément son cervelet, via l'apprentissage qui a pour effet de modifier les relations entre les neurones. C'est une forme de connexionnisme.

Le machine learning est connexionniste parce qu'il permet d'identifier des règles ou des patterns dans les données. Ainsi, dans une base de clients bien documentée, on va pouvoir identifier ceux des paramètres descriptifs qui permettent de prédire leur comportement. Le machine learning non supervisé peut servir à distinguer des segments clients et leurs caractéristiques, ou les paramètres discriminants. Les marketeurs vont ensuite relier cette information nouvelle à leur savoir formel.



Même chez **IBM** qui est assez connexionniste dans l'ensemble, l'approche cognitiviste consiste à extraire des règles de textes via du deep learning appliqué au traitement du langage puis à les exploiter dans des systèmes symboliques et de la logique formelle. Le schéma ci-dessus vient de la présentation "**AI for Complex Situations: Beyond Uniform Problem Solving**" de Michael Witbrock du Cognitive Computing Research d'IBM (2017). Il s'agit donc bien d'une fusion des approches connexionnistes et symboliques.

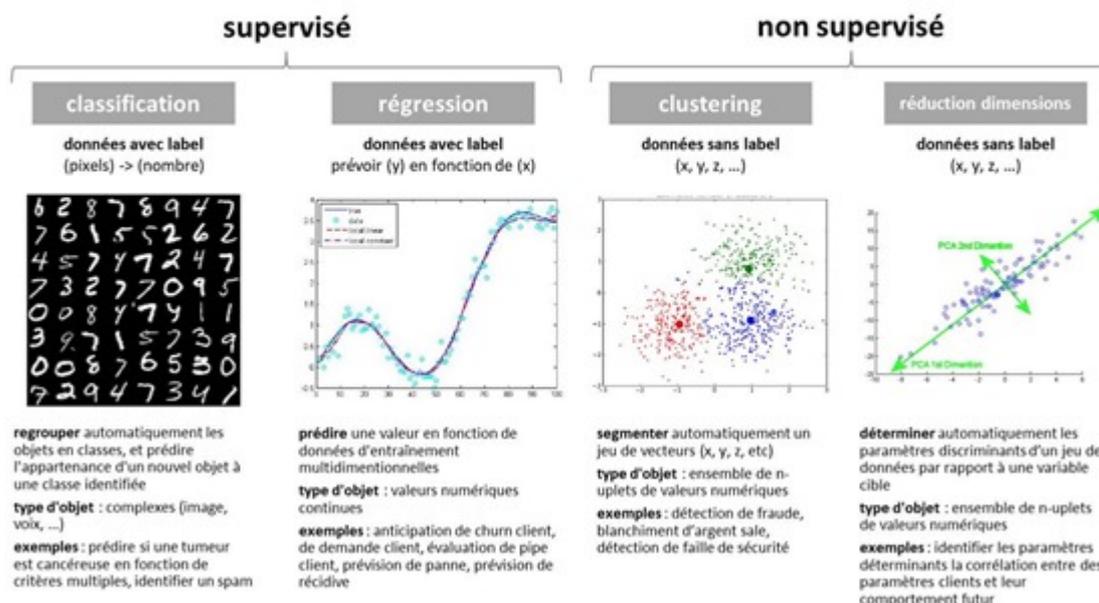
Dans un autre style, la startup franco-américaine **Cosmo Tech** associe machine learning, symbolisme, modélisation de systèmes complexes avec un langage spécifique (CoSML) et simulation pour créer des solutions d'entreprises.

Nous sommes en pleine ère connexionniste

Le renouveau relativement récent de l'IA est lié à l'approche connexionniste, avec le boom des réseaux de neurones et du deep learning depuis au moins 2006. Cette vague a relancé des mythes nombreux autour des risques de l'IA, souvent liés à l'image que l'on s'en fait au travers de la science-fiction dystopique.

Une partie de ces mythes viennent d'une incompréhension de ce que font les solutions à base de deep learning. Elles ne raisonnent pas et n'appliquent pas de logique formelle, tout du moins pour l'instant. Le deep learning ne comprend pas ni ne réfléchit. Il ne gère pas de représentation globale des connaissances. Il a une mémoire limitée à la détection de formes. Il génère une vision mathématique et probabilistique du monde, y compris dans le traitement du langage qui est réalisé en transformant les mots et phrases en vecteurs et en faisant des calculs de distance entre vecteurs pour identifier des mots ou phrases voisins, mais sans en comprendre le sens.

principales méthodes de machine learning



De même, ce que l'on appelle par abus de langage apprentissage automatique n'est pas du raisonnement, même en mode dit non supervisé. Dans le machine learning, le "non supervisé" n'est pas non plus le versant "automatique" du "supervisé". Le machine learning supervisé sert à classer des objets ou à faire des prédictions en s'appuyant sur des données qui sont déjà étiquetées par la nature ou par des humains, comme les noms d'objets présents dans des images. Tandis que le machine learning non supervisé sert à faire de la segmentation automatique et à déterminer des paramètres discriminants de segmentation (cf. le schéma *ci-dessus* qui résume cela).

Quant à l'apprentissage par renforcement, il ajuste les paramètres d'un système comme les paramètres d'un réseau de neurones en fonction de la réaction du monde extérieur à son fonctionnement. On peut l'exploiter pour apprendre à un robot à manipuler des pièces par tâtonnement, comme l'enfant qui apprend à marcher... en formant étape par étape les bonnes liaisons dans son cervelet, le coprocesseur de mouvement du cerveau des mammifères.

D'où vient cette mode connexionniste ? De la conjonction connue entre algorithmes, abondance de capteurs et données et de puissance machine. Les solutions de deep learning sont devenues plus efficaces dans leur domaine (langage, vision) que les méthodes formelles antérieures qui étaient trop laborieuses.

Le cas des réseaux convolutionnels qui servent à détecter des objets dans les images est intéressant. La détection porte sur la détection de formes de bas niveau, puis de formes de niveau intermédiaire comprenant les formes du niveau du dessous, puis de relations entre les formes de plus haut niveau et de classes d'objets. Ces réseaux présentent l'inconvénient de ne pas trop se soucier de la position des formes intermédiaires dans les images.

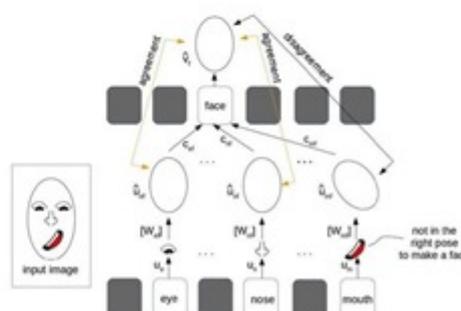
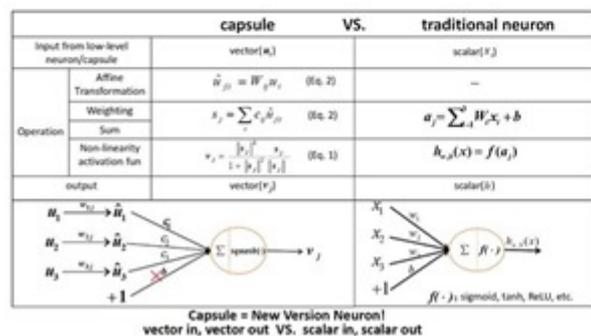
Malgré ces défauts, ils fonctionnent relativement bien pour des usages courants et qui plus est, en exploitant des images d'assez basse résolution, souvent inférieure à 224x224 pixels. La raison ? Les limitations machine et notamment les capacités mémoire des GPU qui ne peuvent pas emmagasiner suffisamment de paramètres de ces réseaux de neurones pendant les phases d'entraînement. Aujourd'hui, la mémoire la plus proche d'un GPU Nvidia de la série Volta, au

standard HBM2, a une capacité de 16 Go.

capsule networks

variante de CNN qui gère la position relative des objets et évite les défauts actuels des CNNs
décembre 2017, Geoff Hinton
ajoute des couches de convolution "vectorielles"
pas encore testé à grande échelle

The following pictures may fool a simple CNN model in believing that this a good sketch of boat, human face, etc.



D'où l'intérêt des "capsule networks", inventés par Geoff Hinton de l'Université de Toronto et Google, et dont les travaux ont été publiés en décembre 2017 ! C'est du tout neuf ! Ils tiennent compte des défauts des réseaux convolutionnels et intègrent la notion de positionnement relatif des formes détectées dans les images, via des vecteurs. C'est une manière d'ajouter un peu de structuration aux méthodes de détection, sans ajouter de sens pour autant.

L'enjeu de l'AGI

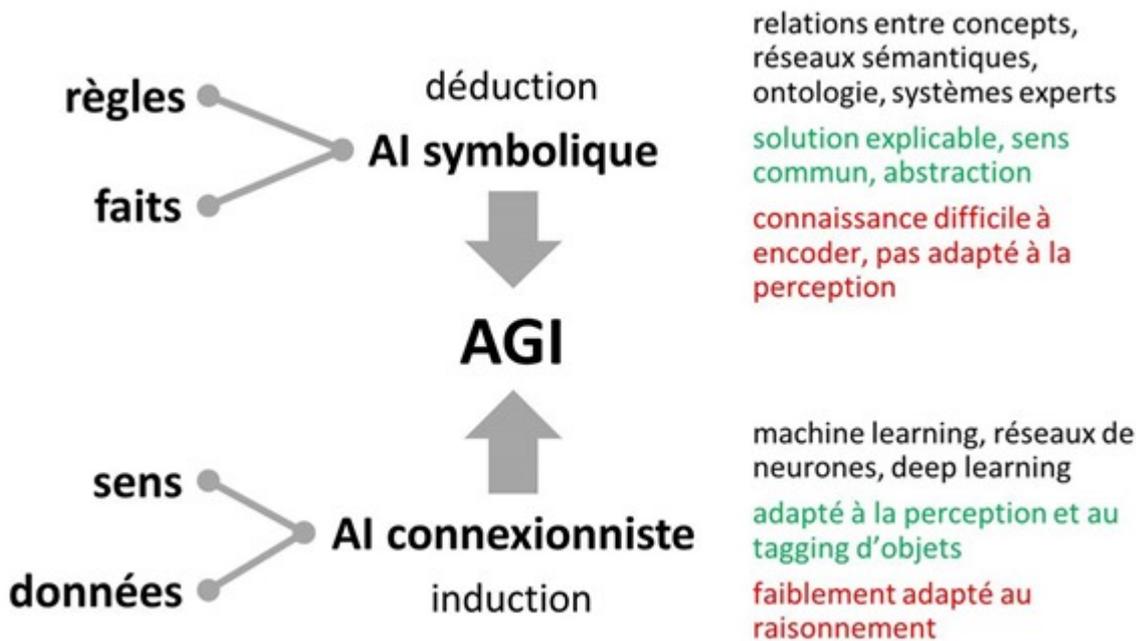
Suivant l'effet de mode du moment, le Rapport de la **Mission Villani** fait la part belle au deep learning et aux méthodes connexionnistes. L'approche symbolique semble y être mise en sourdine. C'est un peu lié au background du chercheur en IA qui faisait partie de la Mission Villani, Marc Schoenauer, Directeur de Recherche à l'INRIA. C'est en fait un spécialiste des approches évolutionnaires qui consistent, pour résoudre un problème, à créer un grand nombre de solutions alternatives puis à les comparer pour choisir la meilleure, en quelque sorte, du darwinisme logiciel accéléré. C'est une voie complémentaire à celle du connexionnisme.

De même, l'IA symbolique et le raisonnement automatique sont assez maigrement représentés dans la feuille de route de l'INRIA de 2016 (Cf **Intelligence artificielle, défis actuels et l'action d'Inria**, 82 pages). On en trouve au laboratoire IRIT de Toulouse ainsi qu'au LIRIS du CNRS à Lyon et probablement plein d'autres laboratoires français que je ne connais pas.

Pourtant, en prenant du recul, on peut se demander si l'enjeu n'est que dans le deep learning. Les débats font toujours rage pour savoir si l'on réussira un jour à créer une intelligence artificielle générale (AGI) rivalisant avec le côté généraliste de l'intelligence humaine, mais restant rationnelle, ce qui n'est pas toujours le cas de cette dernière. Quelle forme scientifique prend cette quête ? C'est en fait probablement un mélange des approches connexionnistes et symboliques. A l'instar de la méthode des ensembles qui est utilisée dans le machine learning et qui associe les résultats de plusieurs méthodes pour en tirer le meilleur, par moyenne ou moyenne pondérée.

On peut réexaminer la dimension éthique du Rapport Villani au travers de ce prisme entre

connexionnisme et symbolisme. Les solutions connexionnistes étant moins explicables que les solutions symboliques, elles génèrent des inquiétudes plus fortes. D'où les nombreuses recherches sur l'explicabilité des algorithmes du deep learning. Les règles symboliques sont plus claires, comme les fameuses lois de la robotique d'Asimov.



De son côté, **Yann LeCun** pense que l'on pourra un jour gérer le raisonnement avec des réseaux de neurones et du deep learning, ce qui relève de la branche connexionniste du raisonnement automatique. Il existe d'ailleurs une branche symbolique des réseaux de neurones, promue notamment par le chercheur brésilien **Artur d'Avila Garcez** qui est basé à Londres. Ce sont des voies importantes pour avancer dans le raisonnement automatique. Et depuis environ 2016 émerge la **programmation différentiable** qui vise à exploiter des réseaux de neurones à mémoire associative pour résoudre des problèmes complexes à la frontière du raisonnement.

Aux USA, divers laboratoires et startups travaillent sur l'AGI avec notamment **Numenta** et ses techniques plutôt connexionnistes de HTM (Hierarchical Temporal Memory), **DeepMind**, lui aussi plutôt connexionniste et **Cycorp** de Doug Lenat avec sa base géante de faits et règles, donc plutôt symbolique. En Suisse, **Nnaisense** adopte aussi une approche connexionniste, créée par Jürgen Schmidhuber, le père des réseaux de neurones LSTM à mémoire utilisés dans le traitement du langage. Mais aucune de ces entités n'avance significativement dans la voie de l'AGI. Elles ont bien du mal à intégrer le raisonnement généraliste dans leurs approches. Et ce n'est pas faute d'avoir un long historique de philosophes et logiciens qui ont décrit dans le détail les différentes formes de raisonnement.

Reste à savoir à quoi servirait une AGI. Nombreux sont les prophètes de l'AGI qui ne creusent pas assez l'étendue des problèmes que l'AGI permettrait de résoudre. Il pourrait s'agir de résoudre des problèmes très complexes qui associent des données sur le monde réel et le savoir humain déjà accumulé, tout en intervenant dans un monde incomplet et non entièrement documenté. Cela peut par exemple passer par la recherche de solutions dans la santé, dans l'économie et la finance ou dans les mécanismes de lutte contre le réchauffement climatique. Une AGI bien construite pourrait suggérer des pistes de travail à des chercheurs parmi un large ensemble de possibilités préétablies. Mais de là à expliquer certains mystères de l'Univers tels

que l'origine exacte de la gravitation ou bien l'existence ou non de la matière noire ?

Il manque en fait une composante clé aux concepts actuels de l'AGI : les capacités d'expérimentation dans le monde réel et de simulation quand l'expérimentation n'est pas possible, est dangereuse ou est trop lente. La connaissance humaine est le fruit de l'expérimentation dans le monde physique. L'imagination humaine permet parfois d'élaborer des théories, via des expériences de pensée comme pour Albert Einstein, qui ne sont vérifiées que longtemps après expérimentalement. Ce fut le cas pour l'existence des ondes gravitationnelles (1916-2015), du boson de Higgs (1964-2015) ou du fermion de Majorana (1937-2017), ce dernier ayant une utilité potentielle pour créer des ordinateurs quantiques.



L'IA connexionniste est adaptée à l'interprétation d'observations. Elle peut aussi servir à simuler des phénomènes relativement simples comme ce fut le cas avec le jeu de Go dans AlphaGo Zero en 2017. Elle ne peut expérimenter par elle-même qu'avec un lien avec le monde physique. Qui plus est, elle est relativement démunie, toute seule, pour simuler les phénomènes complexes comme ceux de la chimie et du vivant. Les lois naturelles sont en effet trop complexes et "continues", pas "discrètes". C'est là où, un jour, l'informatique quantique prendra peut-être le relai de l'IA en ajoutant des capacités de simulation aux IA connexionnistes et symboliques. L'IA n'est en effet pas la réponse à tous les problèmes de ce bas monde !

PS : ce mois d'avril 2018 marque la douzième année révolue du blog "Opinions Libres". Quand j'ai démarré, j'étais un newbie du web 2.0. Maintenant, j'ai l'impression d'être un dinosaure d'un monde qui délaisse progressivement l'écriture, et je résiste à l'extinction ! ☐

Cet article a été publié le 17 avril 2018 et édité en PDF le 6 septembre 2020.
(cc) Olivier Ezratty - "Opinions Libres" - <https://www.oezratty.net>