



Des mécanismes physiques interviennent dans des étapes clés de l'embryogenèse



Cet article est initialement paru dans le magazine *Reflets de la physique* : R. Clément, « Stéphane Leduc et la physique du vivant », n° 58, pp. 26-29, juin 2018.

BIBLIOGRAPHIE

R. Clément, **Stéphane Leduc and the vital exception in the Life Sciences**, arXiv, article 1512.03660, 2015.

S. Tirard, **La biologie synthétique : de Stéphane Leduc à Craig Venter... Le retour ?**, Cahiers François Viète, série II, vol. 6-7, pp. 137-150, 2012.

J. J. Kupiec et P. Sonigo, **Ni Dieu ni gène. Pour une autre théorie de l'hérédité**, Seuil, 2003.

E. Fox Keller, **Making Sense of Life**, Harvard University Press, 2003.

S. Leduc, **Théorie physico-chimique de la vie et générations spontanées**, A. Poinat, 1910.

biaisée par la spécificité du vivant. Cette révolution se dessine enfin aujourd'hui. Ces dernières années, la biologie a connu une explosion de l'utilisation de concepts issus de différents domaines de la physique comme la physique statistique, la mécanique ou l'étude des systèmes complexes.

LE RENOUVEAU DE LA PHYSIQUE DU VIVANT

En biologie du développement, l'utilisation de modèles physiques ou mécaniques pour comprendre l'émergence macroscopique de formes à partir de comportements cellulaires est devenue courante, pas seulement dans les revues de biologie théorique ou de biophysique, mais aussi dans les revues classiques de biologie ou les revues généralistes, y compris les plus prestigieuses. Des mécanismes physiques interviennent ainsi dans des étapes clés de l'embryogenèse des vertébrés comme la gastrulation (l'apparition dans l'embryon, à un stade où il ressemble à une boule de cellules, de feuillettes qui se spécialisent) et la formation du plan d'organisation, qui définit par exemple où pousseront les membres. Et la physique explique aussi en partie la phyllotaxie, la disposition des feuilles sur la tige des plantes. Leduc, en affirmant que le problème de la biologie devait être de comprendre « le jeu des forces physiques sur la matière vivante », et en mettant en lumière les considérables « facultés d'organisation » des forces physiques ordinaires, préfigurait cette évolution.

L'héritage de Stéphane Leduc ne se limite donc pas à ses expériences osmotiques parfois naïves, mais englobe son approche théorique très novatrice des questions d'organisation en biologie. Si certaines de ses idées semblent maintenant datées ou simplistes, comme l'a par exemple souligné en 2002 l'historienne Evelyn Fox Keller dans son livre *Making Sense of Life*, il a pourtant introduit l'idée remarquable, et des preuves de principe expérimentales, que les forces physiques ordinaires ont une « faculté d'organisation » largement sous-estimée jusqu'à très récemment en biologie, préfigurant ainsi le concept moderne d'autoorganisation. Comment alors expliquer que ses contributions aient été si largement oubliées ?

Certes, il a pu manquer à son travail expérimental, qui repose essentiellement sur des analogies de forme, un support théorique plus solide. Mais son éviction témoigne aussi de la réticence à adopter en biologie une démarche non spécifique au vivant. Pourtant, de plus en plus de travaux montrent aujourd'hui comment des principes physiques hérités de Leduc ou de D'Arcy Thompson peuvent aider à appréhender les mécanismes d'organisation de la matière vivante, apportant ainsi leur pierre à la grande quête des mécanismes de l'épigenèse. ■

R

ENDEZ-VOUS

P.80 Logique & calcul
 P.86 Art & science
 P.88 Idées de physique
 P.92 Chroniques de l'évolution
 P.96 Science & gastronomie
 P.98 À picorer

LE POTENTIEL DANGER DES INTELLIGENCES SURHUMAINES

Entre l'idée de superintelligences menaçantes et celle, plus sympathique, de machines qui ne seraient que notre prolongement, laquelle est la plus vraisemblable?

L'AUTEUR



JEAN-PAUL DELAHAYE
 professeur émérite
 à l'université de Lille
 et chercheur au Centre
 de recherche en
 informatique, signal
 et automatique de Lille
 (Cristal)



Jean-Paul Delahaye a récemment publié : **Les Mathématiciens se plient au jeu**, une sélection de ses chroniques parues dans *Pour la Science* (Belin, 2017).

Un exploit étonnant vient d'être réalisé dans le domaine des recherches en intelligence artificielle (IA). Le programme AlphaZero, conçu et programmé par David Silver et son équipe de la société DeepMind, apprend seul, en jouant contre lui-même et en ne connaissant que les règles, à jouer aux échecs, au go ou au shogi (échecs japonais). Il a dépassé le niveau des meilleurs joueurs humains dans ces trois jeux.

L'exploit est un formidable progrès : l'ordinateur Deep Blue, qui a battu en 1997 le champion du monde d'échecs Gary Kasparov, combinait de bons algorithmes et une connaissance experte des ouvertures de jeu, des fins de partie et de l'évaluation des configurations, ces expertises provenant des meilleurs joueurs humains qui avaient en quelque sorte « instruit le programme ».

Avec AlphaZero, plus besoin d'experts humains transmettant un savoir patiemment acquis ; pour ce type de jeux, l'IA travaille seule et obtient mieux que ce à quoi aboutissent des siècles de concentration humaine. Les chercheurs qui ont mis au point AlphaZero précisent : « Nous avons appliqué AlphaZero aux jeux d'échecs, de shogi, ainsi qu'au go, en utilisant le même algorithme et la même architecture réseau pour les trois jeux. Nos résultats démontrent qu'un algorithme polyvalent d'apprentissage par renforcement [l'algorithme favorise les choix qu'il fait quand il gagne] peut apprendre, *tabula rasa*, sans connaissances ni données humaines spécifiques. »

À mesure que l'intelligence artificielle progresse, des questions de plus en plus pressantes sont formulées à son sujet. Certains chercheurs et philosophes qualifiés parfois de technoprophètes nourrissent le débat et proposent des visions du futur, variées et contradictoires. Qu'on s'en réjouisse ou qu'on le craigne, nul ne doute que les machines intelligentes, peut-être mises en réseaux ou fusionnées aux humains, façonneront notre avenir.

Parmi ces penseurs qui tentent de déchiffrer le futur, citons Nick Bostrom, Ray Kurzweil, Ben et Ted Goertzel, Hans Moravec, Francis Heylighen, Hugo de Garis, et Clément Vidal. Tous réfléchissent à ce que pourrait être une intelligence artificielle supérieure et à son attitude face à l'humanité.

EXCÈS D'HONNEUR ET INDIGNITÉ

Comme souvent quand on comprend mal un sujet, on procède par excès et, selon son tempérament, on voit exclusivement le verre à moitié vide ou le verre à moitié plein. Pour reprendre les mots que Racine met dans la bouche d'un de ses personnages, l'IA ne mérite cependant ni l'excès d'honneur qu'on lui accorde ni l'indignité liée à la crainte qu'elle suscite.

L'étonnement face aux succès obtenus et la prise de conscience des impacts que cela aura ne doivent pas faire perdre le sens de la mesure. L'IA réussit magnifiquement dans de nombreuses tâches spécialisées, mais reste encore démunie pour tout ce qui demande du sens commun, comme la traduction automatique



1

ALPHAZERO APPREND TOUT SEUL

Aujourd'hui, les programmes de jeu d'échecs pour téléphones sont plus forts que Deep Blue, la machine qui, en 1997, a battu Gary Kasparov, alors champion du monde. Plus aucun espoir n'existe que des humains puissent reprendre le dessus sur les machines. C'est vrai aussi pour le jeu de go, pour lequel il a fallu toutefois attendre 2015 avant d'avoir des programmes dépassant les meilleurs joueurs.

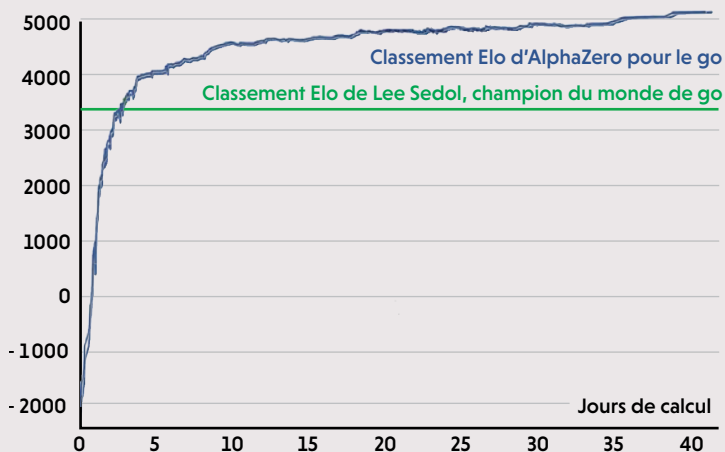
Les progrès de l'intelligence artificielle concernant ces jeux ne s'arrêtent pas là. Le programme AlphaZero, contrairement à Deep Blue et bien d'autres programmes, n'a pas besoin d'utiliser l'expertise humaine. Partant seulement des règles du jeu pour les échecs, le go et le shogi, le programme joue contre lui-même. Au départ, il joue au hasard, mais en prenant en compte les résultats des parties jouées, il apprend et se perfectionne petit à petit en utilisant un mécanisme d'apprentissage construit sur un réseau de neurones formels. Au bout d'une période de quelques heures dans le cas des échecs et du shogi, et de quelques jours pour le go, le programme atteint un niveau de jeu dépassant celui des meilleurs joueurs humains et aussi celui des meilleurs programmes. Gary Kasparov a examiné la façon dont joue le programme

d'échecs résultant de ce processus d'apprentissage par renforcement et a fait quelques commentaires intéressants :

« Contrairement aux programmes traditionnels tels que Stockfish et Fritz, qui utilisent des fonctions d'évaluation prédéfinies par des humains, ainsi que d'énormes bibliothèques d'ouvertures et de fin de partie, AlphaZero commence en ne connaissant que les règles du jeu d'échecs, sans la moindre stratégie humaine associée. En quelques heures, il joue plus de parties contre lui-même qu'on en a mémorisées dans l'histoire des échecs humains. Il apprend seul les meilleures façons de jouer, en particulier en réévaluant les pièces du jeu. Il est devenu assez puissant pour battre les meilleurs joueurs mondiaux d'échecs. Il a par exemple obtenu 28 victoires, 72 nuls et aucune défaite contre Stockfish. Je dois dire que j'ai été ravi de voir qu'AlphaZero avait un style dynamique et ouvert comme le mien. On pense en général que les machines jouent bien en choisissant des coups techniques, conduisant le plus souvent à des parties nulles. J'ai observé à l'inverse qu'AlphaZero donne la priorité à la dynamique des pièces plutôt qu'aux pièces elles-mêmes, préférant des positions qui, à mes yeux, semblaient

risquées et agressives. Les programmes reflètent généralement les priorités et les préjugés des programmeurs, mais dans le cas d'AlphaZero, qui n'est victime d'aucun préjugé, je crois que son style correspond simplement à la réalité profonde du jeu. Sa compréhension supérieure lui permet de surclasser le meilleur programme du monde et cela bien qu'il opère le calcul d'un nombre beaucoup moins important de positions par seconde. Il illustre l'adage classique : "Travaillez intelligemment, plutôt que beaucoup." »

Ainsi, AlphaZero joue Fg5 dans la position de sa partie contre Stockfish (ce qu'aucun joueur humain ni aucun programme classique ne trouvent) et gagne. Pour une discussion du coup, voir <https://youtu.be/thZzaS-noSo>. Nous avons également appris lors de son match contre Stockfish qu'AlphaZero préférerait avoir les blancs : sur les 28 victoires d'AlphaZero, 25 ont été remportées avec les pièces blanches. La machine a également compilé la fréquence à laquelle elle a utilisé les différentes ouvertures mises au point depuis des siècles par les humains, et qu'elle a retrouvées ! L'ouverture du gambit dame est finalement sa favorite.



2

ANTICIPATIONS SINGULIÈRES

Peter Diamandis est l'un des fondateurs de l'Université de la Singularité (<https://su.org>) une organisation californienne d'inspiration transhumaniste. Il a récemment interrogé divers spécialistes à propos de ce qui se passerait dans les vingt prochaines années. Il a effectué la synthèse des réponses (voir <https://www.diamandis.com/blog/countdown-to-the-singularity>)

Voici quelques-unes de ces anticipations. Précisons qu'elles ne font pas l'unanimité et qu'elles semblent plutôt optimistes.

- En 2020, des systèmes de diagnostic utilisant de l'IA et proposant des recommandations seront utilisés dans la plupart des établissements de santé américains.

- En 2020, de nouveaux catalyseurs seront découverts grâce à l'utilisation de processeurs à portes quantiques, ce qui marquera le début de la fin de la chimie traditionnelle.

- En 2022, les gens pourront légalement voyager dans des voitures autonomes partout aux États-Unis. [Si cette prédiction signifie que des voitures parfaitement autonomes existeront en 2022, elle est en contradiction avec l'avis d'autres experts, qui ne les pensent pas possibles avant une cinquantaine d'années.]

- En 2024, des missions humaines privées auront été lancées avec comme objectif la surface de Mars.

- En 2024, les premiers contrats « un centime par kilowattheure » pour l'énergie solaire et éolienne seront signés. [Aujourd'hui, les meilleurs prix dans le monde pour un kilowattheure d'électricité sont de l'ordre de 5 centimes, très rarement 2 centimes, et cela résulte d'une surproduction locale d'énergie électrique qui, sinon, serait jetée. En France, le kilowattheure est vendu autour de 15 centimes.]

- En 2030, l'IA réussira à passer le test de Turing, ce qui signifie qu'elle pourra égaler et dépasser l'intelligence humaine dans tous les domaines. [Là encore, cette prédiction semble particulièrement optimiste.]

- En 2034, les robots agiront comme des domestiques, des majordomes, des infirmières et des nourrices et deviendront des compagnons à part entière. Ils rendront possible une autonomie prolongée des personnes âgées chez elles.

- En 2036, des traitements de longévité seront disponibles et pris en charge par les assurances de santé, prolongeant ainsi la durée de vie moyenne de 30 à 40 ans.

➤ d'œuvres littéraires. La conduite automatisée de véhicules a récemment donné lieu à beaucoup de rêves, alors que les spécialistes annoncent prudemment qu'on ne fera de voitures totalement autonomes, où vous vous asseyez à l'arrière et lisez votre journal, que dans cinquante ans au mieux (voir S. Shladover, « La voiture en quête d'autonomie », *Pour la Science* n°466, août 2016).

Dans un article de cette rubrique, je mentionnais aussi les méthodes conduisant à flouer les réseaux de neurones profonds destinés à la classification d'images, ce qui conduit à ce qu'ils croient reconnaître une autruche quand on leur présente une photo où tout humain voit clairement un camion (*Pour la Science* n°488, juin 2018).

Aujourd'hui, ce domaine de recherche qui n'a pas une définition et une délimitation précises, réussit magnifiquement dans des tâches spécialisées de plus en plus nombreuses, mais rencontre deux obstacles principaux. D'une part, l'IA obtient ses résultats sans procéder comme le font les humains pour les mêmes tâches, mais en s'aidant de dispositifs techniques et de calculs dont se passent les humains : radars et lasers pour les voitures partiellement autonomes, capacités colossales de calculs et de mémorisation qui n'ont rien d'équivalent dans les cerveaux humains. D'autre part, les succès obtenus ne se combinent pas, et l'IA ne sait pas aujourd'hui former des systèmes cohérents associant les capacités spécialisées maîtrisées pour constituer des êtres autonomes doués de sens commun et dont on puisse raisonnablement soutenir qu'ils ont conscience d'eux-mêmes, possèdent une forme de sensibilité, de l'humour et des personnalités structurées et robustes.

ANNONCES PRÉMATURÉES ET SINGULARITÉ

En 1958, Herbert Simon, qui a reçu en 1978 le prix d'économie en mémoire d'Alfred Nobel, annonçait que l'ordinateur battrait le champion humain aux échecs dans les dix années à venir. Il avait raison sur la possibilité d'un tel exploit, mais pas sur la durée d'attente : il a fallu quarante ans pour que sa prédiction se réalise.

En 1966, Irving Good, statisticien reconnu qui conseilla Stanley Kubrick pour son film 2001, *l'Odyssée de l'espace*, annonçait l'avènement vers l'an 2000 d'une superintelligence artificielle au moins équivalente à l'intelligence humaine. Elle n'est pas encore là et ceux qui prévoient sa venue ont reculé l'échéance. Raymond Kurzweil, aujourd'hui directeur de l'ingénierie chez Google, et Peter Diamandis, cofondateur et président de la Singularity University (une société privée californienne), évoquent 2030, alors que Ted et Ben Goertzel,

