

Une brève histoire des sciences computationnelles

REVUE MÉDECINE ET PHILOSOPHIE

Christophe Gauld*

*Psychiatre, CHU de Grenoble,
Doctorant en philosophie de la médecine, Paris 1 / Lyon 3

RÉSUMÉ

La médecine actuelle peut s'ancrer dans un paradigme nommé « computationnel ». La notion de computation vient du latin *computatio*, qui renvoie au calcul (ou opération). Selon Turing, les sciences computationnelles correspondent, au sens large, à toute science du calcul. De ce fait, l'intelligence artificielle, vaste domaine scientifique largement résumé actuellement à l'apprentissage automatique (dont fait partie l'apprentissage profond, ou *deep learning*), est une science computationnelle. Ainsi, une part non négligeable de la médecine actuelle s'intègre dans le paradigme des sciences computationnelles, dont les enjeux mêlent à la fois la philosophie, la pratique clinique, les technologies numériques, les méthodes d'analyses de données, la modélisation des maladies et des comportements et l'étude des phénomènes physiologiques. Nous cherchons dans cet article à décrire le terme de « computationnel » en le réinsérant dans le contexte historique des sciences cognitives. Nous veillerons à ne détailler aucun enjeu normatif ni éthiques, en retraçant uniquement une historiographie du concept.

MOTS-CLÉS : Intelligence artificielle ; computationnalisme ; sciences cognitives ; fonctionnalisme.

DOI : 10.51328/100

Introduction

Nous allons tenter d'éclaircir le terme de « computationnalisme » tel qu'il a été développé au sein des sciences cognitives. Nous ne chercherons pas à retracer finement l'histoire de ce terme et nous ne tenterons donc pas de cerner l'intégralité du champ des sciences cognitives. De plus, nous ne discuterons pas non plus de son ancrage dans la cybernétique, dans la logique symbolique, au sein du béhaviorisme ou de la psychologie de la première moitié du XXe siècle. Enfin, nous n'utiliserons volontairement pas le terme « d'intelligence artificielle » qui regroupe un champ large et mal défini de techniques, d'approches scientifiques, de technologies et dont le sens s'est transformé au cours du temps pour désigner aujourd'hui, dans la plupart des cas, l'apprentissage automatique (machine learning). Nous appuierons ces définitions sur la notion de syntaxe et de sémantique, que

nous allons définir.

La computation (du latin *computare*, calculer) correspond à la manipulation de données selon des règles.

Le terme s'applique aujourd'hui de manière ubiquitaire, à différents domaines scientifiques et pour désigner de nombreuses approches différentes (Miłkowski, 2018). De fait, il est rarement précisé comment le terme de « computation » a évolué au fil du temps, entraînant de nombreuses confusions. Par la suite, nous allons voir qu'il existe au moins deux principales formes de computationnalismes et qu'il existe bien souvent une confusion entre les « sciences computationnelles » et « le computationnalisme » (tout comme il peut y avoir une confusion entre le large champ des sciences cognitives et du cognitivisme) (Dietrich, 1994).

Computationalisme cognitiviste *versus* connexionnisme

Nous allons détailler une première acception du terme computationnalisme, qui possède un fort ancrage historique.

Rappelons avant tout que le champ des « sciences cognitives » décrit l'intégralité des sciences qui s'occupent du fonctionnement du cerveau et des comportements humains, tandis que le terme « cognitivisme » n'est qu'un des courants (longtemps le principal avec le béhaviorisme, au moins des années 1930 aux années 1950) de ce champ.

Dans l'histoire des sciences cognitives, le computationnalisme est une forme de cognitivisme : il désigne le fait de traiter de l'information (en concevant que le cerveau effectue des « calculs »). Nous n'entrons pas dans le détail, mais notons que le cognitivisme est lui-même une forme de fonctionnalisme (donc le computationnalisme est aussi une forme de fonctionnalisme). Il est important de comprendre que ces différents courants se distinguent du fait qu'ils intègrent 1) une syntaxe, c'est-à-dire un ensemble de règles et 2) une sémantique, c'est-à-dire des représentations du monde (Piccinini, 2004). Notons également que le « symbole » est l'unité de base de la syntaxe, et qu'il est habituel de considérer qu'il « contient » du sens, donc de la sémantique. Une science qui manie des symboles manie donc une syntaxe et une sémantique. Ainsi, le computationnalisme comme le cognitivisme sont symboliques.

Cette première acception du terme « computationnalisme » réfère donc uniquement aux processus de calcul du cerveau.

Le connexionnisme désigne un courant des sciences cognitives qui fait référence aux réseaux de neurones. Il s'agit de comprendre le cerveau comme un ensemble d'entités interconnectés et dont le calcul émerge de l'intégralité de ce réseau (sans avoir à le décomposer ou à le localiser au sein du cerveau). Le connexionnisme ne souhaite pas renoncer à la représentation, mais il la considère comme distribuée et non pas dépendante d'une seule entité physique. Le connexionnisme est né dans les années 1950 et ce courant s'oppose fortement au cognitivisme. Il n'est pas un fonctionnalisme (Churchland, 1981). Le connexionnisme datant d'avant les années 1990 – 2000, est uniquement syntaxique, car il manie des règles mais il ne manie pas de représentations (de sémantique). Cette première forme de connexionnisme est ainsi nommée subsymbolique, car il s'applique à des plus bas niveaux que le symbole : il concerne les réseaux et non de supposés « contenus mentaux ».

À noter cependant que les modèles computationnels ne sont pas connexionnistes, qui nécessitent au moins deux fonctions spécifiques (la rétropropagation de l'erreur et le *parallel processing*).

Computationalisme connexionnisme *versus* cognitivisme

Après les années 2000, le terme de computationnalisme « change de bord ». Avec notamment les travaux sur les réseaux convolutionnels (et l'arrivée de l'apprentissage profond, ou *deep learning*), le connexionnisme manie des représentations (c'est-à-dire devient sémantique en plus d'être syntaxique). On ne parle pas de « symbole » cependant, car le traitement de l'information se

fait à travers des réseaux et non par manipulation de symboles (de supposés contenus mentaux). Du fait de traitement de l'information à la fois syntaxique et sémantique, cette deuxième forme de connexionnisme (ou « néo-connexionnisme ») devient une forme de computationnalisme (et donc un fonctionnalisme).

Pour le dire autrement, le terme de « computationnalisme » désignait un traitement de l'information basé sur des symboles (le « cognitivisme ») : du fait des progrès du connexionnisme, le terme de « computationnalisme » s'applique désormais au traitement de l'information qui est basé sur des réseaux (le « néo-connexionnisme »).

Mais pourquoi ce terme de computationnalisme a-t-il pu « changer de bord » à ce point ?

Computationalismes, cognitivisme et connexionnisme ne sont pas au même niveau

En fait, le terme de « computationnel » est insuffisant pour désigner une approche à lui seul. Il ne se réfère qu'au traitement de l'information permettant la réalisation de la fonction (d'où son accointance avec le terme de « fonctionnalisme ») (Marr, 1982).

Or, il peut exister différents « mécanismes » sous-jacents à cette réalisation d'une fonction (par exemple, le vol d'un oiseau). Le niveau de la « computation » est le niveau de la fonction tandis que le niveau du « mécanisme » est le niveau des règles qui rendent possible cette fonction (ces règles peuvent être nommées « algorithmes »). Depuis les années 2000, ces règles sont le plus souvent des réseaux, donc « connexionnistes », dans la plupart des modèles computationnels actuels. Mais ces règles peuvent correspondre à n'importe quel autre mécanisme (par exemple celles d'un automate cellulaire).

Ce constat permet de comprendre que le niveau computationnel, de la fonction, est indépendant des règles qui permettent sa réalisation (les algorithmes), et donc de l'implémentation physique de ces algorithmes. Il existe donc de nombreuses réalisations physiques d'une même fonction, problématique appelée « réalisabilité multiple » en sciences cognitives.

Computationalisme et sciences computationnelles

Il existe une autre confusion, régulièrement retrouvée dans la littérature, entre « computationnalisme » et « computationnel ». Les sciences computationnelles développent des modèles computationnels, dont le but est d'étudier des systèmes complexes grâce à l'utilisation d'algorithmes (et donc de modèles mathématiques formels), de nombreuses variables et de ressources informatiques.

Par exemple, les neurosciences computationnelles sont nées très tard dans l'histoire des sciences cognitives (que l'on pourrait faire remonter au début du pragmatisme, du fonctionnalisme, du béhaviorisme et de la machine de Turing aux alentours des années 1930). Les neurosciences computationnelles sont nées avec les modèles de neurones biologique avec un article séminal en 1907 (LIF) (Brunel et Van Rossum, 2017), et introduisant donc les systèmes dynamiques non linéaires (à noter que le terme de computationnalisme, qui n'avait rien à voir avec le cerveau jusque dans les années 1980, est en réalité né avec McCulloch et Pitts en 1943, qui s'intéressaient à la

plausibilité biologique des réseaux, dont les thèses ont été confirmées par Hebb). Les neurosciences computationnelles sont devenues connexionnistes lors de l'essor de ce dernier en 1980 (notamment à la suite de la Conférence de 1985 organisée par Schwartz qui a introduit le terme de « neurosciences computationnelles »). Après les années 2000, elles se rattachent à l'apprentissage profond (et donc aux réseaux de neurones artificiels), aux systèmes complexes, et intègrent les modèles bayésiens du codage prédictif de Friston (dès 2005). Après les années 2010, les neurosciences computationnelles se rapprochent de l'apprentissage par renforcement. À noter également que la société des neurosciences a été créée relativement tard, dans les années 1970.

Du fait de l'évolution des connaissances, les sciences computationnelles sont actuellement principalement axées autour d'approches connexionnistes (et non cognitivistes). Le terme de « computationnel » est donc appliqué à une science en particulier, et semble plus restreint que le « computationnalisme ». Il correspond cependant au cadre théorique du computationnalisme, puisque les sciences computationnelles expliquent les fonctions (computationnalisme) par des algorithmes (connexionnisme).

Les neurosciences computationnelles se distinguent fortement des neurosciences « classiques » (cognitivistes) parce qu'elles cherchent à expliquer comment se déroule le traitement de l'information (niveau computationnel), autrefois réservé à la psychologie (tandis que les neurosciences classiques s'occupaient des mécanismes à l'origine de ce traitement de l'information).

Conclusion

Ainsi, nous avons que le terme de computationnel est plus restreint que le computationnalisme et que ce dernier a connu au moins deux acceptions au fil du temps (un computationnalisme « cognitiviste – symbolique » et un second « connexionniste – subsymbolique »). Ainsi, le cognitivisme renvoie à la computation selon des représentations et des règles maniant des symboles (les contenus mentaux), tandis que le connexionnisme renvoie à la computation selon des réseaux de neurones et des règles inscrites elles-mêmes dans ces réseaux. Nous avons donc isolé deux formes de computationnalismes et deux formes de connexionnismes. Avec certains auteurs comme Piccinini (2009), on pourrait aller plus loin en affirmant que toute forme de traitement de l'information est nécessairement connexionniste (selon une troisième forme de connexionnisme, au sens large), puisqu'il fait intervenir des réseaux d'entités.

Enfin, nous avons vu que l'évolution de ces acceptions a fortement été influencée par le niveau algorithmique, permettant la réalisation de la fonction par le biais de différents mécanismes. Une telle constatation permet d'affirmer que la question du traitement de l'information doit nécessairement se pencher sur les mécanismes sous-jacents – autrement dit, que la psychologie devrait pouvoir comprendre les neurosciences.

RÉFÉRENCES

Brunel, N., Van Rossum, M. C. (2007). Lapicque's 1907 paper: from frogs to integrate-and-fire. *Biological cybernetics*, 97(5-6), 337-339.

Dietrich, E. (1994). Computationalism. In *Thinking Computers and Virtual Persons* (pp. 109-136). Academic Press.

Churchland, P. M., Churchland, P. S. (1981). Functionalism, qualia, and intentionality. *Philosophical Topics*, 12(1), 121-145.

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences*, 360(1456), 815-836.

McCulloch, W. S., Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.

Miłkowski, M. (2018). From computer metaphor to computational modeling: the evolution of computationalism. *Minds and Machines*, 28(3), 515-541.

Piccinini, G. (2004). Functionalism, computationalism, and mental contents. *Canadian Journal of Philosophy*, 34(3), 375-410.

Piccinini, G. (2009). Computationalism in the Philosophy of Mind. *Philosophy Compass*, 4(3), 515-532.