

COMPLEXITÉ

—

Entre Mesure Scientifique et Impératifs Philosophiques

—

Mathis Reymond

janvier 2025

1 Introduction

Une pierre, un ballon, un ordinateur, un être humain, ces objets sont tous différents, pourtant nous qualifions tous par une propriété qui semble mesurable : leur *complexité*. Chacun s'accordera à dire que les objets sont ici listés par ordre croissant de complexité. Mais on emploie également le terme de complexité pour parler de problèmes ; classer des images des images de chats et de chiens est plus complexe que classer des points de deux couleurs. Bien que la complexité soit omniprésente dans notre langage naturel et en science, elle demeure majoritairement une notion intuitive qui n'a pas de définition univoque. Malgré notre difficulté à la définir, la complexité est une notion apparaissant dans l'étude de nombreux systèmes présentant des enjeux sociaux importants comme le climat, l'économie ou l'intelligence artificielle. C'est également un problème important en science comme le note [Hossenfelder 2022]. L'institu de Santa Fe a justement été fondé en 1984 pour promouvoir l'étude interdisciplinaire des systèmes complexes. Définir la complexité et en développer une théorie satisfaisante demeure donc un problème, irrésolu, qui constitue un enjeux majeur.

Dans ce texte, nous nous proposons d'étudier deux documents ; le chapitre portant sur la complexité du livre *The Oxford Handbook of Philosophy of Science* [Humphreys 2016] ainsi que le papier *Assembly theory explains and quantifies*

selection and evolution [Sharma et al. 2023]. Ce papier propose une quantification de l'assemblage d'un objet donné, c'est à dire une mesure de la difficulté à construire cet objet combinée à la rareté de ses composants. Après avoir, dans un premier temps, introduit les concepts clé de l'article, nous verrons, dans un second temps, quelles tensions ils entretiennent avec les considérations philosophiques développées dans [Humphreys 2016] en nous focalisant sur les aspects biologiques. Enfin, dans un troisième temps, nous développerons des perspectives complémentaires mettant en avant les limitations de la théorie de l'assemblage ainsi que la difficulté d'ancrer scientifiquement le concept de complexité et d'en développer un paradigme unifié.

2 Une théorie de la complexité

La définition de la complexité qu'on sait le mieux manipuler est celle proposée par Kolmogorov - ou complexité algorithmique [Kolmogorov 1963]. Elle est souvent introduite comme suit. Pour obtenir la séquence

"abababababababab"

numériquement, on peut la stocker telle qu'elle est, ce qui a un cout de 16 caractères, ou on peut utiliser un algorithme comme $8 \times ab$ qui n'est composé que de 4 caractères et permet de retrouver la séquence entière. On voudra donc dire que cette séquence a une complexité inférieure à la séquence

"akyhrukdorlsroph"

qui ne semble pas pouvoir être générée par un algorithme de moins de 16 caractères. Cette conception de la complexité est donc descriptive en ce sens qu'elle quantifie la complexité d'un objet selon les descriptions dont on en dispose. Elle fait notamment fi de la structure de l'objet, de la richesse des interactions entre ses composants ou encore des processus causaux à considérer pour expliquer son apparition.

Il existe de nombreuses autres tentatives de définition la complexité, chacune s'appliquant à une certaine classe de modèles (séquences, graphes, automates, etc.) mais aucune n'est suffisamment versatile pour parler de la complexité de nombreux autre objets. Une tentative se voulant trouver une définition de la complexité en biologie est celle de la théorie de l'assemblage [Sharma et al. 2023].

2.1 La théorie de l'assemblage

La théorie de l'assemblage est née avec l'intention de comprendre l'origine de la vie notamment pour l'identifier sur d'autres planète. Mais dans la quête de vie sur d'autres planètes, on doit accepter qu'elle peut se manifester sous des formes très différentes de celles que l'on connaît. Les astrobiologistes aimeraient pouvoir dire d'un objet donné s'il est d'origine biologique ou non, ce qui est très dur puisque cela demande de formaliser la notion de vie. C'est donc sur une prémisse

La vie est le seul mécanisme dont dispose l'univers pour générer des objets complexes

que la théorie de l'assemblage repose. La théorie de l'assemblage est ainsi avant tout une théorie de la complexité : si un objet est complexe, au sens de la théorie, alors il doit être vivant. Cette approche se distingue d'autres tentatives, souvent empruntées d'anthropocentrisme, de définition de la vie qui se reposent sur l'hypothèse de certaines conditions environnementales nécessaires à la vie, ou sur la présence de certains éléments particuliers identifiés dans les objets vivants connus (comme le carbone ou l'oxygène). La théorie de l'assemblage est agnostique en ce qui concerne les éléments qui composent l'objet : elle s'intéresse seulement au nombre d'étapes nécessaire à l'assemblage de ces éléments. Dans une interview [Walker 2024], les auteurs expliquent qu'ils ne font pas l'hypothèse de certains éléments nécessaires à la vie, mais reconnaissent que le carbone, par ces propriétés atomiques, est plus versatile que d'autres éléments en ce qu'il est susceptible de former de longues molécules, en se liant à lui même ou à d'autres éléments. Néanmoins, ce genre d'explication n'est simplement pas prise en charge par la théorie.

2.2 L'équation principale

L'équation la plus importante introduite par le papier définit la grandeur d'assemblage, A , qui quantifie la sélection nécessaire pour produire un ensemble donné d'objets. Elle repose sur deux autres grandeurs. D'une part, l'indice d'assemblage mesure la complexité d'un objet en fonction du nombre minimal d'étapes nécessaires pour le construire à partir de blocs de construction de base. D'autre part, le nombre de copies de l'objet dans l'ensemble des objets déjà construits avec les blocs de base est également pris en compte. L'assemblage intègre ces deux

notions et propose une mesure de la complexité qui prend en compte à la fois l'assemblage de l'objet et sa rareté. Plus l'indice d'assemblage est grand, plus le nombre d'étapes pour construire l'objet est grand et donc plus l'assemblage est grand. Plus le nombre de copies d'un objet est élevé, plus cet objet est commun, ce qui réduit la probabilité qu'il soit apparu par hasard. Il doit donc, au contraire, résulter d'un processus de sélection, ce qui conduit à un assemblage plus important. L'interprétation de l'assemblage tient en une phrase

Plus un objet requiert d'étapes pour être construit ou plus il est copié, plus il est complexe.

Un objet complexe et copié beaucoup doit être généré par un mécanisme de sélection. A partir de la mesure de l'assemblage, les chercheurs ont identifié un seuil d'assemblage que seul les objets d'origine biologique arrivent à dépasser. Ainsi, il est si peu probable qu'un objet dont l'assemblage est supérieur à 15 ait été généré par le fruit du hasard qu'on considère qu'il doit avoir été généré par de manière non abiotique, par un processus de sélection.

L'un des grands avantages de la théorie de l'assemblage est sa mise en pratique rapide. Si l'indice d'assemblage ne peut pas toujours être calculé numériquement à cause l'explosion exponentielle du nombre d'objets qu'on peut construire avec le nombre d'éléments de l'espace d'assemblage, il peut être mesuré au spectroscope de masse. Le cadre offert par la théorie se veut établir un lien entre la physique et la biologie en expliquant comment des formes complexes peuvent émerger par sélection, sans plans préétablis.

3 Tension avec les considérations philosophiques

3.1 Notions clés du chapitre

Le chapitre sur la théorie de la complexité de [Humphreys 2016] s'articule autour de deux temps. Dans un premier temps, il traite la question de l'apparition de la simplicité – c'est à dire de motifs macroscopiques résultant d'organisation microscopiques particulières – et dans un second temps, il aborde la question de l'apparition de la sophistication. La simplicité est l'ensemble des comportements macroscopiques apparaissant de l'agglomération du comportement des éléments microscopiques du système. L'ensemble de ces nombreux comportements microscopiques qui ne semblent pas pouvoir être expliqués individuelle-

ment forme la convolution. La sophistication est définie au dessus de la simplicité. Elle fait référence à l'orchestration des comportements macroscopiques simples entre eux en des dynamiques plus riches que celles de chaque comportement simple considéré individuellement. Par exemple, les systèmes chaotiques répondent à des règles macroscopiques simples mais présentent une extrême sensibilité aux conditions initiales, rendant la prédiction de leur état futur essentiellement impossible [Li and Yorke 1975]. Le chapitre tente de couvrir toutes les approches philosophiques rendant compte des processus de simplification et de sophistication des systèmes complexes.

Au sens du chapitre sur la complexité de [Humphreys 2016], la théorie de l'assemblage – et les perspectives sur la complexité qui en découlent – proposée par [Sharma et al. 2023] s'inscrit dans les approches de modularité et presque-décomposabilité du processus de simplification de la convolution des systèmes complexes. La presque-décomposabilité est une propriété des systèmes complexes qui se caractérise par une organisation modulaire mais dont le fonctionnement des modules n'est pas tout à fait indépendant de celui des autres. En plus d'interagir entre eux pour former les dynamiques macroscopiques du système, les modules s'influencent mutuellement, dans une moindre mesure, ce qui donne au système des propriétés accommodantes pour son étude comme une certaine stabilité et, *a fortiori*, sa simplicité. Ce point de vue est très combinatoire. Il encapsule la perspective développée par l'assemblage dans [Rosas et al. 2024]. Cependant, dans l'article, des notions philosophiques semblent manquantes comme nous nous attacherons à le montrer par la suite.

3.2 Emergence, complexité, contraintes

Mais d'abord, nous voulons insister sur l'importance de trois notions qui nous semblent inextricables : la complexité (ou de manière duale, la simplicité), l'émergence et les contraintes.

Un maçon n'a pas besoin de connaître la physique des particules pour construire une maison. C'est un fait frappant d'évidence, mais qui a une implication profonde. Cela signifie que les propriétés macroscopiques d'un système peuvent être étudiées indépendamment de ses propriétés microscopiques. C'est à dire que sous l'effet de contraintes physiques, les innombrables particules composant

la matière s'organisent de telle sorte qu'à l'échelle du maçon, elles se comportent comme des objets simples, formant des unités, qu'il peut appréhender individuellement. Cette apparente simplicité est en contraste avec d'autres systèmes, des systèmes complexes, dans lesquelles à la fois les composants microscopiques et macroscopiques du système ont des effets macroscopiques qui se mêlent, ce qui rend l'étude du système nettement plus difficile.

Dans un contexte biologique, cette idée peut se reformuler comme suit : de la possible complexité de tous les arrangements de cellules possibles émergent des structures au fonctionnement simple – dont le comportement descriptible par des équations dérivables empiriquement et permettant de faire prédictions – sous l'effet des contraintes imposées à la fois par l'environnement et par le corps tel qu'il est déjà construit.

En somme, l'apparition de comportements macroscopiques simples n'est rendue possible que par un phénomène d'émergence se produisant sous l'effet de contraintes spécifiques.

3.3 Nature descriptif de la théorie de l'assemblage

Notons ensuite que la théorie de l'assemblage est purement descriptive. Le calcul de l'indice d'assemblage fonctionne sur le même principe que la théorie de Kolmogorov : pour chaque objet, elle se donne un ensemble de possibilités (des chemins d'assemblage) qu'elle ordonne au regard d'une mesure (la longueur des chemins) et choisit de retenir celle qui a la plus petite mesure. La différence est donc qu'elle s'applique à des graphes dirigés au lieu de s'appliquer à des séquences et qu'elle intègre le nombre de répétitions des composants identifiés de l'objet.

Si cette métrique est une approche intéressante pour détecter ce qu'on voudrait considérer vivant, sa nature descriptive ne permet pas de comprendre comment et pourquoi ou comment les objets sont générés. C'est une théorie très peu explicative, et chaque objet qu'elle classe comme vivant ou non doit encore être interprété. De plus, la notion de contrainte et la nature des contraintes conduisant ou non un objet à être généré n'est pas développée ce qui est pourtant, semble-t-il, un maillon manquant capital pour non pas détecter la vie, mais prétendre comprendre son apparition. Par opposition, [Humphreys 2016] nous rappelle que

dans le cadre de la théorie de Lotka, les organismes s'adaptent dans le but de maximiser leur consommation et leur utilisation d'énergie de l'environnement. Sans discuter le crédit à accorder à cette théorie, elle a notable qu'elle est beaucoup plus explicative que la théorie de l'assemblage et se repose sur un concept plus naturel à apprécier : l'énergie, qui est perçue comme une contrainte à laquelle les objets doivent se plier.

Par ailleurs, la nature descriptive de la théorie de l'assemblage s'observe aussi en remarquant sa déconnexion aux réalités physiques. La façon Kolmogorovienne de retenir le plus court chemin d'assemblage a un sens strictement combinatoire, ce qui convient bien pour une théorie computationnelle, mais beaucoup moins pour une théorie de la vie se voulant paver la voie entre physique et biologie [Walker 2024]. Le plus court chemin d'assemblage n'est, par exemple, pas nécessairement le moins coûteux en énergie, ce qui est une grandeur bien souvent minimisée lors d'opérations dans le monde physique. Et a fortiori, le plus court chemin d'assemblage n'est pas non plus nécessairement la façon retenue par la nature pour générer l'objet. De plus, ce qui échappe à cette vue combinatoire, c'est la nature profondément statistique de nombreux systèmes physiques. Dans ceux-ci, la rareté d'un objet n'est pas un argument pour sa complexité, mais simplement le reflet de moindres contraintes poussant à son apparition. Un environnement exerçant des contraintes légèrement différentes sur les objets qu'il contient pourrait voir la proportion de ces objets varier drastiquement. Le modèle d'Ising et les automates cellulaires sont une bonne illustration de la sensibilité aux contraintes (capturées par les règles de proximités) des objets d'un environnement.

Enfin, cette vue descriptive s'oppose à une vue fonctionnelle, et elle conduit à développer une conception non standard de certaines notions biologiques. On attendra d'une théorie de la complexité en biologie qu'elle se rapporte aux composants macroscopiques structurant un corps, comme les organes ou les os. Mais les auteurs [Sharma et al. 2023] précisent dans leur interview [Walker 2024] que la théorie de l'assemblage ne s'applique en pratique qu'aux molécules de petite taille. Comme elle est descriptive et combinatoire, elle peut théoriquement se prononcer quant à l'assemblage de cellules, mais une cellule est déjà trop grosse pour pouvoir explorer les différents chemins qui pourraient amener à son assemblage et en isoler le plus court. Les organes ou les os sont encore davantage hors de portée. Les auteurs font aussi référence aux trois principes de

l'évolution selon Lewontin et avance que leur théorie est compatible avec ceux-ci. Pour Lewontin [Lewontin 1970] pour que l'évolution contrainte par la sélection naturelle se produise, il faut que

- 1. Différents individus d'une population aient différentes morphologies, physiologies et comportements.*
- 2. Différents phénotypes aient des chances différentes de survie et de reproduction selon l'environnement.*
- 3. Il y ait une corrélation entre les parents et leurs descendants dans la contribution de chacun aux générations futures.*

Tant que ces trois conditions sont remplies, une population peut subir des changements évolutifs. Ces trois principes ont des contreparties beaucoup plus pauvres pour les auteurs de la théorie de l'assemblage [Walker 2024] qui mentionnent l'aléatoire, les pressions sélectives et l'hérédité. Cette dernière est matérialisée dans la théorie de l'assemblage sous la forme d'une équation différentielle ; mais il nous apparaît en revanche moins clairement comment cet aléatoire ou ces pressions sélectives sont intégrées à la théorie

De manière plus anecdotique, on peut également noter que, contrairement à la définition Kolmogorovienne de l'indice d'assemblage, les mécanismes d'adaptation n'incarnent pas toujours un optimal. Le nerf laryngé récurrent chez la girafe est un exemple de mécanisme d'adaptation qui n'est pas optimal. Ce nerf prend sa source au niveau du cerveau et atteint le larynx. Chez la girafe, le trajet de ce nerf est très long, il descend du cerveau, passe autour de l'aorte, remonte le long de l'oesophage pour atteindre le larynx. Le trajet de ce même nerf était direct chez un ancêtre de la girafe. D'un point de vue fonctionnel, un nerf laryngé récurrent formant une boucle n'est pas la façon la plus simple d'assembler un cou long, mais c'était la façon accessible, compte tenu des contraintes au moment de l'élongation du cou des girafes.

Finalement, il nous semble qu'une bonne proposition pour une théorie de la complexité en biologie ne doit pas seulement permettre de reconnaître les assemblages combinatoires longs, mais plutôt de nous permettre d'appréhender comment et pourquoi se produit la réduction de cette complexité dans des systèmes organisés, fonctionnels, interdépendants et résiliants à certains changements microscopiques, mais pas à d'autres. Et une théorie sans connection au sens

physique des réalités modélisées n'est peut-être simplement pas suffisante comme le note [Humphreys 2016]

"Il n'est pas clair que la vie artificielle a des leçons générales à enseigner sur la complexité, au delà de ce qui est déjà connu de la biologie évolutive."

4 Perspectives complémentaires

On peut noter qu'au sens de la théorie de l'assemblage, un crayon à papier doit avoir une complexité inférieure à celle d'un être humain. Pourtant, aucun crayon à papier n'aurait jamais été assemblé sans être humain. C'est un exemple quelque peu cynique mais qui nous rappelle un aspect capital de la complexité : les complexités de deux objets doivent pouvoir être comparées – c'est d'ailleurs ce qui a motivé le développement de la théorie de l'assemblage. Celle d'un crayon de papier et d'un humain ne peut pas l'être. Et comme nous l'avons déjà noté, la théorie de l'assemblage ne prend en fait en charge que les molécules de petite taille. Mais même dans ce cadre restreint, elle demeure descriptive, et, en particulier, elle est agnostique quant aux éléments constitutifs d'un objet. La complexité mesurée n'est donc pas celle de l'objet en lui-même, mais seulement celle de son processus d'assemblage en un sens combinatoire (longueur de la chaîne d'assemblage). Néanmoins, la définition de l'assemblage est dépendante – implicitement dans l'article – de l'espace d'assemblage. Donc, soit nous ne pouvons comparer que des objets ayant le même espace d'assemblage – ce qui autorise très peu de comparaisons –, soit, comme la théorie ne prend pas en compte les spécificités de l'espace d'assemblage (nature des éléments qu'il contient, taille), des objets de complexité similaire peuvent afficher des propriétés très différentes et une comparaison phénotypique de ces objets, basée sur la complexité de leur comportement ou leur caractère macroscopiques, n'est pas permise par la théorie de l'assemblage – ce qui est pourtant une perspective fonctionnelle, qu'il nous intéresserait d'avoir en biologie.

5 Conclusion

Finalement, dans la théorie de l'assemblage, nous nous attendions à trouver une théorie de la complexité spécifique à la biologie, et éclairant les concepts

de sélection et d'évolution comme son titre le laissait entendre. Mais le travail réalisé relève en fait davantage d'une théorie combinatoire assez classique en complexité et plutôt déconnectée de la biologie, bien qu'elle soit exploitable pour distinguer les petites molécules d'origine biologiques des autres molécules.

Nous nous sommes montrés plutôt critiques à l'endroit de cette théorie. En plus d'être descriptive et de manquer d'explicativité, elle ne satisfait pas tout fait aux impératifs philosophiques consensuels d'une théorie de la complexité. Néanmoins, en ce que cette théorie a de lacunaire, nous avons pu identifier plus clairement ce qu'on doit attendre d'une théorie de la complexité en biologie :

- Fondamentalement, la complexité est une notion qui doit être comparée.
- Elle ne doit pas faire fi d'un certain sens physique et intégrer des contraintes, pour avoir une portée explicative.
- Elle doit passer la barrière de l'émergence, c'est à dire parler d'objet de différentes échelles.
- En particulier, elle doit s'appliquer à des objets macroscopiques (molécules, cellules, organes etc.) pour permettre de parler de phénotype ou de comportement.

References

- Hossenfelder, Sabine (2022). *Existential Physics: A Scientist's Guide to Life's Biggest Questions*. New York: Viking. ISBN: 978-1984879455.
- Humphreys, Paul (Sept. 2016). *The Oxford Handbook of Philosophy of Science*. Oxford University Press. Chap. Complexity Theory. ISBN: 9780199368815. DOI: 10.1093/oxfordhb/9780199368815.001.0001.
- Kolmogorov, Andrei (1963). *On Tables of Random Numbers*. Vol. 25. Sankhyā Ser. A., pp. 369–375.
- Lewontin, R. C. (1970). "The Units of Selection". In: *Annual Review of Ecology and Systematics* 1, pp. 1–18. DOI: 10.1146/annurev.es.01.110170.000245. URL: <https://doi.org/10.1146/annurev.es.01.110170.000245>.
- Li, Tien-Yien and James A. Yorke (1975). "Period Three Implies Chaos". In: *American Mathematical Monthly* 82.10, pp. 985–992.

- Rosas, Fernando E. et al. (2024). *Software in the natural world: A computational approach to hierarchical emergence*. arXiv: 2402.09090 [nlin.AO]. URL: <https://arxiv.org/abs/2402.09090>.
- Sharma, Abhishek et al. (Oct. 2023). "Assembly theory explains and quantifies selection and evolution". In: *Nature* 622.7982, pp. 321–328. ISSN: 1476-4687. DOI: 10.1038/s41586-023-06600-9.
- Walker, Sara (Oct. 2024). *Sara Walker: "We Could Find Aliens In The Lab"*. YouTube video. Interview by Dr. Brian Keating. URL: https://www.youtube.com/watch?v=zWh-LBkq_9Y.