

Remarques sur la modélisation

Jean Petitot, CAMS-EHESS, décembre 2024.

Introduction

J'entretiens avec Fabio Minazzi une longue et riche coopération depuis 1985. Elle a commencé lorsqu'il m'a invité cette année-là (il y a déjà quarante ans) au grand colloque de Varese en hommage à Ludovico Geymonat *La scienza tra filosofia e storia in Italia nel novecento*. Fabio m'a alors permis d'approfondir ma connaissance de la tradition italienne du rationalisme critique (Antonio Banfi, Giulio Preti, Ludovico Geymonat, Mario Dal Pra, Evandro Agazzi, ou Fulvio Papi).

L'impact du colloque de Varese fut tel, que je décidais d'organiser avec Giulio Giorello un colloque à Paris sur *La portée européenne des traditions épistémologiques italiennes* (3-5 novembre 1987, Ecole des Hautes Etudes en Sciences Sociales, Institut Henri Poincaré et Institut Culturel Italien).

C'est à partir de là que s'est mise en place une collaboration étroite, amicale et féconde avec Fabio. En 1993 nous avons écrit un long article "La connaissance objective comme valeur historique : le néo-illuminisme italien" qui est paru dans les *Archives de Philosophie* et dresse un panorama assez complet du rationalisme critique dans la Péninsule. Fabio a aussi traduit en grande partie mon ouvrage *Per un nuovo illuminismo. La conoscenza scientifica come valore culturale*, paru chez Bompiani en 2009. Je ne compte pas les nombreuses occasions que nous avons eu de collaborer, à Mendrisio, à Lecce, à Varese, et dans de très nombreux colloques internationaux.

J'aimerais donc saluer l'activité philosophique, académique et éditoriale (en particulier pour les fonds d'archives) de Fabio. C'est tout un trésor de pensée, de culture et d'engagements dont il est devenu l'héritier, le dépositaire et le chroniqueur à Varese, dans le *Centro 'Carlo Cattaneo' e 'Giulio Preti'* qui est devenu le centre de l'histoire du rationalisme critique italien.

En hommage à Fabio, j'aimerais, pour ce *Festschrift*, faire quelques remarques sur la notion de modèle.

1. Données et modèles

La notion de modèle est multiple mais nous nous focaliserons ici que sur un seul aspect, déjà extrêmement diversifié, celui du traitement de données observables par une "intelligence". L'intelligence peut être celle d'un savant héritier d'une longue tradition scientifique mais elle peut être aussi celle d'un animal ou une intelligence artificielle. L'observation des données peut aller des systèmes sensori-moteurs des organismes vivants jusqu'à des appareils ultra élaborés et dont la construction présuppose déjà elle-même des engagements théoriques énormes comme le télescope spatial James Webb ou le Grand collisionneur de hadrons (LHC) du CERN. Les modèles peuvent être hautement mathématisés comme en relativité générale, en théorie quantique des champs ou en physique statistique mais peuvent être aussi des modèles neuronaux internes de la structure de

l'environnement d'un organisme (de son "Umwelt" au sens de Jakob von Uexküll). On a donc un extraordinaire éventail de possibilités, du psychisme animal jusqu'aux plus hauts sommets de l'intelligence scientifique.

Souvent le test de l'efficacité d'un modèle est la possibilité d'anticiper le mouvement d'un "quelque chose" de mobile reconnu comme un objet par le système observateur. Un félin poursuivant sa proie, une mouche faisant un vol stationnaire, un chat sautant sur une armoire, un acrobate, un danseur, un joueur de tennis, encodent dans leur système sensori-moteur de façon implicite un ensemble sidérant de lois physiques et de calculs balistiques leur permettant d'anticiper, de contrôler, de finaliser des comportements moteurs. Ils en ont un modèle intérieur *encodé* dans des architectures fonctionnelles de connexions nerveuses qui ne sont pas explicites en formules et demeurent *implicites*. En revanche les théories physiques construisent des modèles conceptuels et mathématiques *explicites* de l'évolution des systèmes qu'elles théorisent.

Dans les modèles théoriques explicites, dont nous avons l'habitude, la question centrale est de pouvoir *reconstruire*, à partir de règles nomologiques *abstraites*, les données empiriques *concrètes* observées de façon à pouvoir *anticiper* d'autres données. Dans ces procédures de reconstruction-anticipation s'intriquent tout un ensemble de niveaux cognitifs et de méthodes tant descriptives qu'explicatives, faisant intervenir inductions, déductions et abductions. Une grande partie de l'épistémologie des modèles et de la philosophie des sciences classiques a été et est toujours consacrée à la clarification de ces questions délicates.

Mais, du côté des bases neuronales de la cognition, le développement spectaculaire des neurosciences et des méthodes d'imagerie non invasives comme l'EEG (électro-encéphalographie), IRMf (imagerie par résonance magnétique fonctionnelle), TEP (tomographie par émission de positons), ont montré que les apprentissages se matérialisent par des intra-connectivités et inter-connectivités d'aires cérébrales spécifiques et dédiées. Ces connectivités ont des *architectures fonctionnelles* extrêmement précises mises en place grâce à la plasticité synaptique à travers des règles comme celles de Donald Hebb : "cells that fire together, wire together".

Ces progrès ont conduit ces dernières années à un développement considérable de l'intelligence artificielle bio-inspirée et des méthodes d'apprentissage "profond" (deep learning), ce qui a complètement renouvelé toutes les problématiques de la modélisation et a conduit à en étudier les conditions de possibilité neuro-cognitives en amont. Ces progrès d'ingénierie permettent de prendre conscience de la profondeur et de la complexité de ces conditions préalables à toute modélisation, aussi élémentaire soit elle. On cherche à comprendre ou à construire des réseaux de neurones compliqués, naturels ou artificiels, dotés de facultés de perception et de capacités d'action (cela peut être un animal ou un robot, par exemple un insecte ou un drone), qui partent des données brutes qui leurs sont accessibles et qui en dérivent des anticipations en modifiant progressivement leur connectivité interne.

Aujourd'hui, les réseaux de neurones artificiels multicouches complexes peuvent traiter un nombre considérable de données ("big data") et sont pilotés par ces données ("data-driven"), sans hypothèses préalables ou "priors". Ils sont purement *inductifs* et s'auto-entraînent à anticiper l'évolution des données. Tout un milieu technologique d'ingénierie s'est développé autour de ces

questions et, de façon récurrente, on cherche à comparer les capacités de ces intelligences artificielles avec celles de la science classique. Un test standard est de retrouver les lois de Kepler à partir des données de Tycho Brahé.

Ces systèmes neurocognitifs d'intelligence artificielle bio-inspirée sont, répétons-le, purement inductifs (même si dans un second temps on les simplifie en général en introduisant des “priors”) mais ils sont surtout *sans concepts*. Ils opèrent de façon pré-judicative et ante-prédicative, non réflexive. Les “concepts” y sont implicites, *non sémantisés*, et réalisés comme des états d'activité neurale et c'est pourquoi ils peuvent concerner des animaux ou des robots. Ils apprennent progressivement un système dynamique et son flot dans l'espace des états obtenu en itérant le générateur infinitésimal. En fait ils apprennent des *schèmes* neuro-cognitifs qui sont des *mixtes* de “pré-concepts” non sémantisés et d'algorithmes encapsulés dans les connexions. Ce n'est qu'avec l'intelligence de l'*homo sapiens* que ces schèmes se sont trouvés en quelque sorte “décompilés” (ce que l'on appelle la conceptualisation et la réflexivité) et “dé-mixés” en concepts sémantiques et en algorithmes, ce qui ouvrira plus tard, au cours de l'évolution culturelle, à la possibilité de modèles théoriques à la fois conceptuels et mathématiques.

2. La dialectique du conceptuel et du non conceptuel

Tournons-nous donc maintenant vers des modèles théoriques scientifiques et “rationnels” où, au contraire des modèles internes neuro-cognitifs non conceptuels que nous venons d'évoquer, le rôle des concepts est déterminante. Les ressources nécessaires à l'observation, les variables d'état, les dynamiques à anticiper sont toujours présentes mais des concepts y “décompilent” et rendent explicites des algorithmes internes encapsulés, cachés et implicites.

C'est cette dialectique entre concepts et algorithmes que nous nous proposons de préciser en pensant aux modèles physiques les plus connus, dont l'histoire est la plus longue et la plus riche, ceux dont la référence dans les travaux d’“automated scientific discovery” est la plus courante.

Nous partons donc du très vieux problème des relations entre la diversité des données empiriques (phénomènes observables) et l'unité des concepts théoriques dans un certain domaine de réalité. Chaque discipline va de la première à la seconde en utilisant des ressources cognitives universelles : description, abstraction, catégorisation, taxinomies, inductions, inférences, corrélations, causalité, etc. Il s'agit d'une *analyse conceptuelle* à plusieurs niveaux qui a pour fonction de “subsumer”, comme on le disait autrefois, le *divers* des phénomènes empiriques sous l'*unité* des concepts théoriques.

2.1. Analyse conceptuelle

On peut se représenter cette montée vers l'abstraction par une “colonne” ascendante d’“analyse conceptuelle” comprenant plusieurs niveaux et menant d'une conceptualité *descriptive* des données empiriques à une conceptualité de plus en plus *théorique* culminant dans des catégories

et des principes.

Relativement à cette situation initiale, les approches les plus connues de la modélisation la pensent dans le cadre de théories formelles générales des langages scientifiques (positivisme logique, Cercle de Vienne, cf. Hans Reichenbach, Rudolf Carnap, etc.) y compris de celles ayant tenu compte des critiques comme celles de Carl Hempel ou de Willard Quine, ou bien dans un cadre historiciste et relativiste (Thomas Kuhn, Gerald Holton, etc.). Notre approche de la modélisation est assez différente. Elle repose sur une dialectique entre le conceptuel et le non conceptuel.

Nous entendons ici par “concept” un contenu sémantique du langage commun au sens lexical du terme. la notion de “mouvement d'une planète” est un contenu conceptuel. En revanche, la liste des positions *continues* d'une planète est un ensemble de données numériques empiriques mesurant les différentes valeurs d'une fonction ”position” au cours d'un intervalle de temps par rapport à un certain repère spatial. Certes la “fonction” possède un contenu conceptuel mais son infinité de valeurs numériques sera considérée ici comme non-conceptuelle. La fonction d'un modèle est de les *calculer* à partir d'hypothèses générales.

Évidemment, on peut dire que de telles valeurs numériques sont mathématiquement des éléments du corps \mathbf{R} des nombres réels et que la structure algébrique de corps (avec son addition et sa multiplication dont nous apprenons les propriétés à l'école) est un concept mathématique axiomatisé. Mais là encore, les éléments quelconques du *continuum* \mathbf{R} ne sont pas conceptuels. En fait la notion même d'abstraction par généralisation est *antithétique* à la notion d'exactitude numérique dans un continuum.

2.2. Schématisme et construction

De très nombreux débats philosophiques se sont développés autour de cette alternative “conceptuel/non conceptuel” en relation avec la dialectique “discret/continu”. Nous nous inscrivons dans des traditions de pensée qui maintiennent cette alternative. On peut la faire remonter philosophiquement au moins jusqu'à Kant puisque l'opposition entre l'Esthétique transcendantale (l'espace et le temps comme intuitions pures non conceptuelles) et l'Analytique transcendantale (Catégories comme “concepts de l'entendement pur”) est fondatrice pour la *Critique de la Raison Pure* et, surtout, pour les *Premiers Principes Métaphysiques d'une Science de la Nature* qui proposent une interprétation transcendantaliste de la Mécanique newtonienne. Nous allons y revenir.

Chez Kant cette opposition mène très loin puisque c'est sur elle que repose la différence fondamentale entre la philosophie comme “connaissance rationnelle par concepts” et la science “proprement dite” comme “connaissance rationnelle par *construction* de concepts”, la “construction” étant l'attribution d'une structure mathématique à un contenu conceptuel au moyen d'une opération fondamentale de *schématisme*.

On pourrait croire cette problématique dépassée mais il n'en est rien. Le développement des sciences cognitives l'a remise au premier plan.

3. La modélisation comme problème inverse de l'abstraction

3.1. Les synthèses computationnelles (*sôzein ta phainomena* : “sauver les phénomènes”)

Si l'on admet cette dialectique conceptuel / non conceptuel on en arrive tout de suite à une difficulté. Même si une connaissance conceptuelle peut être extrêmement sophistiquée, elle ne peut pas en général *reconstruire* les données car les concepts sont des significations alors que les données sont des phénomènes possédant en général un format non conceptuel (par exemple celui de mouvements spatio-temporels). Bien sûr, les significations obtenues par subsomption conceptuelle peuvent être appliquées aux données, mais elles ne peuvent pas les reconstruire par un *calcul*.

Une caractéristique des sciences formalisées comme la physique mathématique est de pouvoir inverser la subsomption conceptuelle et de résoudre *le problème inverse de l'abstraction*. C'est selon nous le rôle fondamental de la modélisation. Elle part de concepts et de principes et, au moyen de ressources mathématiques et algorithmiques hautement *génératives*, elle reconstruit et simule des phénomènes *virtuels* qui reproduisent les phénomènes *réels*. C'est une *synthèse computationnelle* des phénomènes empiriques et non plus une analyse conceptuelle. Elle a pour fonction de rendre génératif ce qu'il y a de nomologique dans les phénomènes.

Il est essentiel de bien distinguer dans ce contexte les termes “générique” et “génératif” car leurs relations à la dialectique de l'unité et de la diversité sont inverses l'une de l'autre. Le “générique” regroupe le “divers” en classes d'équivalences et étiquette les classes par des labels conceptuels. Il “catégorise”.¹ Au contraire le “génératif” engendre (génère) du divers en itérant des règles, des procédures, des constructions. L'itération augmente progressivement la complexité. Il suffit de penser aux équations différentielles dont l'intégration consistent à itérer un générateur infinitésimal (nous allons y revenir ci-dessous) ou encore aux fractals.

L'enjeu de la modélisation est donc de *calculer* une réalité “virtuelle” qui puisse refléter au mieux la réalité empirique donnée. Comme on le dit depuis l'antiquité, il s'agit de “sauver les apparences” (*sôzein ta phainomena*), de les expliquer au moyen de constructions mathématiques génératives (très difficiles et souvent même impossibles à trouver) reposant sur des “hypothèses” théoriques.²

On peut représenter cette conception de la modélisation en ajoutant à la colonne “ascendante” (“bottom-up” comme on dit souvent dans le jargon cognitiviste) de l'analyse conceptuelle une colonne “descendante” (“top-down”) de synthèse computationnelle. Les deux colonnes se correspondent au niveau phénoménal de base (en bas des colonnes) par un isomorphisme partiel entre les phénomènes empiriquement donnés et les phénomènes computationnellement reconstruits, autrement dit entre une réalité empirique et une réalité virtuelle.

1 La notion de catégorie a un double sens. Le sens cognitif ou linguistique de label dans une classification. Et aussi le sens philosophique qui remonte à Aristote et culmine avec Kant.

2 L'idée de chercher à “reproduire” des données empiriques à partir d'une reconstruction mathématique remonte au moins à la théorie des épicycles dans l'astronomie grecque de l'antiquité.

La qualité du “fit” est la finalité de la modélisation.

3.2. Comment rendre les concepts génériques génératifs ?

Ce double mouvement ascendant / descendant rend évident que la *condition de possibilité* de la modélisation se situe en haut des colonnes et concerne la conversion de contenus conceptuels *génériques* très abstraits unifiant et *subsumant* une énorme diversité phénoménale en algorithmes très *génératifs* capables au contraire *d'engendrer* une diversité virtuelle équivalente.

Dans les modèles neuronaux purement inductifs dont nous avons parlé plus haut la question ne se pose pas car au cours de son apprentissage, le réseau encapsule dans ses connexions les algorithmes lui permettant de calculer des anticipations. Comme nous l'avons dit, ces algorithmes encapsulés dans le hardware neuronal sont “sans concept”, en quelque sorte “cachés” dans les connexions qui les compilent. En revanche, dans les modèles rationnels dont nous parlons maintenant l'induction abstractisante ascendante est conceptuelle et sa conversion en synthèse computationnelle descendante constitue le problème clé. Les concepts “décompilent” en quelque sorte des algorithmes “cachés dans les profondeurs de l'âme” comme le disait Kant à propos du schématisme.

La problématique est par conséquent tout sauf évidente puisque, répétons-le, l'unification générique conceptuelle et la générativité algorithmique semblent être au prime abord opposées. C'est d'ailleurs pourquoi les vraies modélisations sont si rares et ont toujours constitué un triomphe de la science où se sont croisées d'extraordinaires innovations longtemps purement mathématiques et maintenant également informatiques.³

3. Le tournant transcendantal

Le chemin fut donc extrêmement long pour combler le fossé entre l'analyse conceptuelle (catégories et principes, compréhension théorique) et la synthèse computationnelle (lois et équations différentielles, explication opérationnelle). Il fut mathématique mais aussi métaphysique et, selon nous, particulièrement bien thématiqué par les *Metaphysische Anfangsgründe der Naturwissenschaft* (*Premiers Principes Métaphysiques de la Science de la Nature*) de Kant (1786) qui sont le premier texte à avoir explicité les fondements métaphysiques de la Mécanique newtonienne.

En 1955 Jules Vuillemin, dans *Physique et métaphysique kantienne* expliquait qu'en physique le transcendantal signifie simplement “relativité et groupes de symétrie”. Son analyse admirable des *Metaphysische Anfangsgründe* est à l'origine de notre défense et illustration de la philosophie transcendantale depuis 1970. Ce texte kantien est selon nous le plus grand texte philosophique sur la mécanique newtonienne. Il a fait l'objet d'une analyse non moins admirable que celle de Vuillemin par Michael Friedman dans *Kant's Construction of Nature* en 2013. Nous l'avons commenté dans de nombreux articles.

On sait que Kant a corrélé très précisément la structure de la théorie de Newton avec les

³ Nous avons étudié de nombreux exemples de telles réussites modélisatrices, par exemple dans notre intervention “Conceptual Analysis and Computational Synthesis in Mathematical Physics” à un colloque de l'AIPS édité en 2017 par Evandro Agazzi.

quatre classes de Catégories et de Principes mis en place dans la *Critique de la Raison pure*. Cela est rendu possible à cause de l'apport métaphysique crucial de l'*Esthétique transcendantale*. Cette dernière explique que les concepts ne sont pas la seule condition de possibilité d'une connaissance des objets et qu'une autre condition de possibilité vient du fait que la *donation* des phénomènes empiriques possède un *format*.

Le lien entre les concepts et le calcul devient alors possible. C'est le *schématisme transcendantal* comme “mixte”. Pour les mouvements formalisés par la mécanique newtonienne l'*Esthétique transcendantale* est celle de l'espace-temps newtonien avec le groupe de la relativité galiléenne. Cela est admirablement expliqué philosophiquement dans la *Phoronomie*. La démarche kantienne peut être résumée de la façon suivante.

1. Il existe une mathématisation du format de donation des phénomènes c'est-à-dire des “formes de l'intuition”. Elles sont alors converties en “intuitions formelles” appartenant à un univers mathématique spécifique (et non à une logique générale). Cet univers mathématique est *constitutivement* associé à la classe de phénomènes considérés.
2. Les catégories peuvent dès lors être schématisées en relation avec les mathématiques spécifiques de l'*Esthétique transcendantale* qui gouverne donc la schématisation dont découlent les Principes.
3. Les principes mathématisés en accord avec la structure des intuitions formelles peuvent être appliqués à la physique.
4. Lorsqu'elles sont appliquées, elles interprètent très précisément les structures théoriques de la Mécanique Newtonienne (Phoronomie, Dynamique, Mécanique, Phénoménologie) : les structures d'arrière plan (background structure), les lois et les équations différentielles, l'épistémologie de l'“hypothèses non fingo”.
5. La synthèse computationnelle devient ainsi ce que Kant appelle une “construction mathématique” de la Nature.

Conclusion

En tant que “mixte”, le schématisme transcendantal est l'opérateur dialectisant la part conceptuelle et la part computationnelle de la procession de modèles physiques ayant conduit de l'astronomie de l'antiquité (Eudoxe, Hipparque, Ptolémée, Proclus) à la physique moderne, procession dont un tournant reste les lois de Kepler. Mais si nous bouclons la boucle et revenons à notre introduction concernant l'apprentissage profond dans les réseaux de neurones de l'intelligence artificielle contemporaine, nous retrouvons la notion de schème, mais cette fois-ci comme schème *empirique*, et même comme schème empirique mixte non encore “dé mixé” de “pré-concepts” non sémantisés et d'intuitions sensori-motrices et perceptives. De l'intelligence animale aux sommets de l'intelligence scientifique, le long de l'évolution biologique et de l'évolution culturelle se sont développées des architectures schématiques d'abord implicites, puis explicitées, thématiques par réflexivité, sémantisées et formalisées. Kant avait profondément raison en affirmant que

“Le schématisme de l'entendement pur, en vue des phénomènes et de leur simple forme, est un art caché dans les profondeurs de l'âme humaine, et dont nous aurons de la peine à arracher à la nature les secrets du fonctionnement pour les mettre à découvert sous les

yeux.”

De Kant à Husserl, l'idéalisme transcendantal a concerné toutes ces architectures fonctionnelles neuronales (déjà entrevues par Descartes) que l'on devait supposer être enfouies “dans la profondeur de l'âme humaine” sans pouvoir pour autant n'en rien dire puisque le cerveau était une “boîte noire”. On parlait alors de *synthétique a priori*. La déconstruction du synthétique a priori a consisté à éliminer la référence à la boîte noire d'une subjectivité interne empiriquement inconnaissable pour lui substituer une cognition socialisée “transparente”. Mais les neurosciences cognitives contemporaines, surtout grâce aux méthodes révolutionnaires d'imagerie *in vivo*, ont rendu la boîte noire en partie transparente. D'où ce retour massif de Kant dont nous parlions dans notre introduction. Bref, nous pourrions dire que les structures synthétiques a priori supposées par l'idéalisme transcendantal sont confirmées et relayées par le matérialisme neuronal des architectures fonctionnelles et que notre conception de la modélisation doit en tenir le plus grand compte.

C'est cette profonde actualité de la philosophie critique que je partage avec Fabio Minazzi depuis tant d'années. Ce *Festschrift* attestera de l'immense travail qu'il a accompli pour réhabiliter le rationalisme critique dans la compréhension de l'objectivité de la connaissance, avec toutes ses dimensions culturelles, axiologiques et civiques. Son opus de 2022 *Historical Epistemology and European Philosophy of Science: Rethinking Critical Rationalism and Transcendentalism* en est une superbe synthèse.

Bibliographie

Friedman, M., *Kant's Construction of Nature*, Cambridge University Press, 2013.

Kant, I., 1786. *Metaphysische Anfangsgründe der Naturwissenschaft*, Kants gesammelte Schriften, Band IV, Preussische Akademie der Wissenschaften, Berlin, Georg Reimer, 1911.

Minazzi, F., *Historical Epistemology and European Philosophy of Science: Rethinking Critical Rationalism and Transcendentalism*, (trad. R. Sadleir), Springer, Cham, 2022.

Minazzi, F., Zanzi, L. (a cura di), *La scienza tra filosofia e storia in Italia nel Novecento*, Atti del Convegno Internazionale Varese, 24-26 ottobre 1985, Presidenza del Consiglio dei Ministri, 1987.

F. Minazzi, J. Petitot, “La connaissance objective comme valeur historique : le néo-illuminisme italien, in *Philosophies en Italie*”, *Archives de Philosophie*, 56/4, 621-660, 1993. Trad. it. in F. Minazzi, *L'epistemologia come ermeneutica della ragione*, Erga Edizioni, C.N.R.-Genova, 1998.

J. Petitot, *Per un nuovo illuminismo. La conoscenza scientifica come valore culturale*, Préface de F. Minazzi, Bompiani, Milano, 2009.

Petitot, J., 2016. “Conceptual Analysis and Computational Synthesis in Mathematical

Physics”, *Varieties of Scientific Realism* (a cura di E. Agazzi), Springer, 2016.

Vuillemin, J., *Physique et Métaphysique kantienne*, Presses Universitaires de France, Paris, 1955.