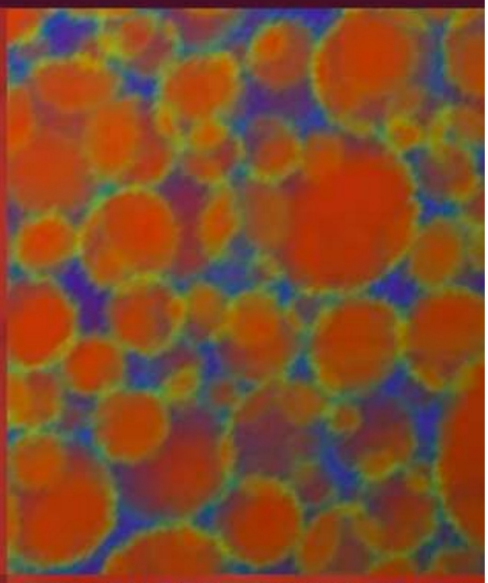
THE FRONTIERS COLLECTION



Juan Manuel Durán

COMPUTER SIMULATIONS IN SCIENCE AND ENGINEERING

Concepts–Practices–Perspectives

Springer

Juan M. Duran´

Simulations informatiques en science et ingénierie

Concepts - Pratiques - Perspectives

vendredi 17 août 2018

Springer

À maman, papa et Jo
pour leur soutien inconditionnel et leur amour.

A mon abeille
Je n'aurais pas pu faire ce voyage sans
toi.

A Mauri
En véritable amitié.

À Manuel
Avec amour et admiration.

Préface

La présence omniprésente des simulations informatiques dans toutes sortes de domaines de recherche met en évidence leur rôle en tant que nouvelle force motrice pour l'avancement de la recherche en sciences et en génie. Rien ne semble échapper à l'image de réussite que les simulations informatiques projettent auprès de la communauté des chercheurs et du grand public. Un moyen simple l'illustrer consiste à se demander comment la science contemporaine et l'ingénierie ressemble sans l'utilisation de simulations informatiques. La réponse serait s'écartent certainement de l'image actuelle que nous avons de la recherche scientifique et technique.

Autant les simulations informatiques réussissent, autant ce sont aussi des méthodes qui échouent dans leur but d'enquêter sur le monde; et autant que les chercheurs utilisent Parmi eux, les simulations informatiques soulèvent des questions importantes qui sont au cœur de la pratique contemporaine des sciences et de l'ingénierie. A cet égard, des simulations informatiques font un fantastique sujet de recherche pour les sciences naturelles, les sciences sociales, l'ingénierie et, comme dans notre cas, aussi pour la philosophie. Etudes sur simulations informatiques aborder de nombreuses facettes de la recherche scientifique et technique, et évoquer des questions d'interprétation à tendance philosophique étroitement liées aux problèmes de paramètres expérimentaux et applications d'ingénierie.

Ce livre introduira le lecteur, d'une manière accessible et autonome, à ces différents aspects fascinants des simulations informatiques. Une étude historique sur la conceptualisation des simulations informatiques au cours des soixante dernières années ouvre le vaste monde des simulations informatiques et de leurs implications. La mise au point passe ensuite à la discussion sur leur méthodologie, leur épistémologie, et les possibilités d'un cadre éthique, entre autres questions.

La portée de cet ouvrage est relativement large afin de familiariser le lecteur avec les multiples facettes des simulations informatiques. Tout au long du livre, j'ai cherché à maintenir un équilibre sain entre les idées conceptuelles associées à la philosophie des simulations informatiques d'une part, et leur pratique en science et l'ingénierie d'autre part. A cette fin, le livre a été conçu pour un large public, des scientifiques et ingénieurs, décideurs politiques et universitaires, au grand public publique. Il accueille toute personne intéressée par les questions philosophiques – et concevable réponses – aux questions soulevées par la théorie et la pratique des simulations informatiques. Il

Il faut mentionner que, bien que le livre soit écrit sur un ton philosophique, il ne s'engage pas dans des discussions philosophiques profondes. Il cherche plutôt à explorer la synergie entre les aspects techniques des simulations informatiques et la valeur philosophique qui en ressort. À cet égard, les lecteurs idéaux de ce livre sont des chercheurs de toutes disciplines travaillant sur des simulations informatiques mais ayant des penchants philosophiques.
Cela ne veut pas dire, bien sûr, que les philosophes professionnels ne trouveraient pas dans ses pages des problèmes et des questions pour leur propre recherche.

Une belle chose à propos des simulations informatiques est qu'elles offrent un champ de recherche fertile, tant pour les chercheurs utilisant les simulations que pour ceux qui y réfléchissent. À cet égard, bien que le livre puisse avoir certains mérites, il est également insuffisant à bien des égards. Par exemple, il n'aborde pas les travaux de simulation informatique en sciences sociales, un domaine de recherche très fructueux. Il ne traite pas non plus de l'utilisation des simulations informatiques dans et pour l'élaboration des politiques, de leurs utilisations pour rendre compte au grand public, ni de leur rôle dans une société démocratique où la pratique de la science et de l'ingénierie est un bien commun. C'est certainement malheureux. Mais il y a deux raisons qui, je l'espère, excusent le livre de ces défauts. La première est que je ne suis spécialiste d'aucun de ces domaines de recherche, donc ma contribution n'aurait eu que peu d'intérêt. Chacun des domaines mentionnés soulève en soi des enjeux spécifiques que les personnes impliquées dans leur étude connaissent le mieux. La deuxième raison tient au fait que, comme tous les chercheurs le savent, le temps – et aussi dans ce cas, l'espace – est un tyran. Il serait impossible de ne serait-ce qu'effleurer la surface des nombreux domaines où les simulations informatiques sont actives et florissantes.

En règle générale pour le livre, je présente un sujet donné et discute des problèmes et des solutions potentielles. Aucun sujet ne doit être traité comme sans rapport avec un autre sujet du livre, et les réponses proposées ne doivent pas non plus être considérées comme définitives. En ce sens, le livre vise à motiver de nouvelles discussions, plutôt que de fournir un ensemble fermé de sujets et les réponses à leurs problèmes fondamentaux. Chaque chapitre devrait néanmoins présenter une discussion autonome sur un thème général des simulations informatiques. Je dois également mentionner que chaque chapitre contient de nombreuses références à la littérature spécialisée, donnant au lecteur la possibilité de poursuivre ses propres intérêts sur un sujet donné.

Le livre est organisé comme suit. Dans le chapitre 1, j'aborde la question « que sont les simulations informatiques ? en donnant un aperçu historique du concept. En remontant le concept de simulation par ordinateur au début des années 1960, nous réaliserons bientôt que de nombreuses définitions contemporaines doivent beaucoup à ces premières tentatives. Une bonne compréhension de l'histoire du concept s'avérera très importante pour le développement d'une solide compréhension des simulations informatiques. En particulier, j'identifie deux traditions, une qui met l'accent sur la mise en œuvre de modèles mathématiques sur l'ordinateur, et une autre pour laquelle le trait saillant en jaune est la capacité de représentation de la simulation informatique. Selon la tradition que les chercheurs ont choisi de suivre, les hypothèses et les implications à tirer des simulations informatiques différeront. Le chapitre se termine par une discussion sur la classification désormais standard des simulations informatiques.

Le cœur du chapitre 2 est d'introduire et de discuter en détail les constituants des modèles de simulation, c'est-à-dire les modèles à la base des simulations informatiques. À

cette fin, je discute des diverses approches des modèles scientifiques et techniques avec le
le but d'ancrer les modèles de simulation comme un type plutôt différent. Une fois que c'est
accompli, le chapitre poursuit la présentation et la discussion de trois unités d'analyse
constitutifs de simulations informatiques, à savoir, la spécification, l'algorithme et
le processus informatique. Ce chapitre est le plus technique du livre, car il dessine
largement à partir d'études sur le génie logiciel et l'informatique. Pour
équilibrer cela avec une certaine philosophie, cela présente également plusieurs problèmes liés à ces
unités d'analyse - à la fois individuellement et en relation les unes avec les autres.

Le seul but du chapitre 3 est de présenter la discussion sur la question de savoir si l'ordinateur
les simulations sont épistémologiquement équivalentes à l'expérimentation en laboratoire. L'importance
d'établir une telle équivalence trouve ses racines dans une tradition qui considère l'expérimentation
comme le fondement solide de notre vision du monde. Depuis une grande partie de
le travail exigé des simulations informatiques est de fournir des connaissances et une compréhension
des phénomènes du monde réel qui autrement ne seraient pas possibles, alors le
la question de leur puissance épistémologique par rapport à l'expérimentation en laboratoire se pose
naturellement. Suivant la tradition philosophique de discuter de ces
questions, je me concentre sur le problème désormais consacré de la "matérialité" de l'informatique
simulations.

Bien que les chapitres 4 et 5 soient indépendants l'un de l'autre, ils ont en commun l'intérêt d'établir
la puissance épistémologique des simulations informatiques. Alors que le chapitre 4 le fait en discutant
des nombreuses manières dont les simulations informatiques sont fiables, le chapitre 5 le fait en
montrant les nombreuses fonctions épistémiques attachées aux simulations informatiques. Ces deux
chapitres représentent donc ma contribution aux nombreux
tente de fonder le pouvoir épistémique des simulations informatiques. Notons que
ces chapitres sont, à la base, une réponse au chapitre 2, qui traite de l'informatique
simulations vis-à-vis de l'expérimentation en laboratoire.

Ensuite, le chapitre 6 aborde des questions qui sont sans doute moins visibles dans la littérature
sur des simulations informatiques. La question centrale ici est de savoir si les simulations informatiques
doit être comprise comme un troisième paradigme de la recherche scientifique et technique –
la théorie, l'expérimentation et le Big Data étant le premier, le deuxième et le quatrième paradigme
respectivement. À cette fin, je discute d'abord de l'utilisation du Big Data dans la pratique scientifique et
de l'ingénierie, et de ce que cela signifie d'être un paradigme. Avec ces éléments à l'esprit,
J'entame une discussion sur les possibilités de tenir des relations causales dans la science du Big Data
ainsi que dans les simulations informatiques, et ce que cela signifie pour l'établissement de
ces méthodologies comme paradigmes de la recherche. Je termine le chapitre par une comparaison
entre les simulations informatiques et le Big Data avec un accent particulier sur ce que
les met à part.

Le dernier chapitre du livre, le chapitre 7, aborde une question pratiquement inexplorée dans la
littérature sur l'éthique de la technologie, à savoir la perspective de
une éthique des simulations informatiques. Certes, la littérature sur les simulations informatiques
s'intéresse davantage à leur méthodologie et à leur épistémologie, et beaucoup moins à
les implications éthiques qui accompagnent la conception, la mise en œuvre et l'utilisation de simulations
informatiques. En réponse à ce manque d'attention, j'aborde ce chapitre comme un
aperçu des problèmes éthiques abordés dans la littérature spécialisée.

Stuttgart, Allemagne,
août 2018

Juan M. Duran´

Remerciements

Comme c'est généralement le cas, de nombreuses personnes ont contribué à la réalisation de ce réservation possible. Tout d'abord, je voudrais remercier Marisa Velasco, P´ıo Garc´ıa et Paul Humphreys pour leur encouragement initial à écrire ce livre. Tous les trois ont eu une forte présence dans ma formation de philosophe, et ce livre certainement les possède beaucoup. A tous les trois, ma gratitude.

Ce livre a commencé en Argentine et s'est terminé en Allemagne. En tant que postdoc au Centro de Investigaciones de la Facultad de Filosof´ıa y Humanidades (CIFFyH), Universi dad Nacional de Cordoba (UNC - Argentine), financé par le Conseil national de la recherche scientifique et technique (CONICET), j'ai eu la chance d'écrire et de discuter de la premiers chapitres avec mon groupe de recherche. Pour cela, je suis reconnaissant à V´ıctor Rodr´ıguez, Jose Ahumada, Juli an Reynoso, Maximiliano Bozzoli, Penelope Lodeyro, Xavier Huvelle, Javier Blanco et Mar´ıa Silvia Polzella. Andres Il ci˜ c est un autre membre de ce groupe, mais il mérite une reconnaissance particulière. Andres a lu chaque chapitre du livre, a fait des commentaires réfléchis et a corrigé plusieurs erreurs que je n'avais pas remarquées. Il a également vérifié avec diligence bon nombre des formules que j'utilise dans le livre. Pour cela et pour les innombrables discussions que nous avons eues, merci Andy. Naturellement, toutes les erreurs sont de ma responsabilité. Un merci très sincère à CIFFyH, l'UNC, et le CONICET, pour leur soutien aux Humanités en général et à moi en particulier.

J'ai aussi de nombreuses personnes à remercier, bien qu'elles n'aient pas contribué directement au livre, ils ont montré leur soutien et leurs encouragements tout au long de bonnes et mauvais jours. Ma gratitude éternelle va à mon bon ami Mauricio Zalazar, et mon soeur Jo. Merci les gars d'être là quand j'ai le plus besoin de vous. Mes parents ont également été une source constante de soutien et d'amour, merci maman et papa, ce livre n'existerait pas sans vous. Merci également à V´ıctor Scaraf´ıa et à mon nouveau famille adoptive : les Pompers & les Ebers. Merci à tous les gars d'être si géniaux. Remerciements particuliers aux deux Omas : Oma Pomper et Oma Eber. Je vous aime mesdames. Enfin, merci à Peter Ostritsch pour son soutien dans de nombreux aspects de ma vie, la plupart d'entre eux sans rapport avec ce livre.

Le livre s'est terminé au Département de philosophie des sciences et de la technologie de simulation informatique, au Centre de calcul haute performance de Stuttgart

(HLRS), Université de Stuttgart. Le département a été créé par Michael Resch
et Andreas Kaminski et financé par le Ministerium fur Wissenschaft, Forschung und Kunst Baden-
Wurttemberg (MWK), que je remercie pour avoir fourni une atmosphère confortable pour travailler sur
le livre. J'adresse mes remerciements à tous les membres du
département, Nico Formanek, Michael Hermann, Alena Wackerbarth et Hildrun
Lampe, je n'oublierai jamais les nombreuses discussions philosophiques fondamentales que nous avons eues
au déjeuner – et sur notre nouvelle machine à expresso – sur les sujets les plus variés. Je me sens
particulièrement chanceux de partager un bureau avec Nico et Michael, bons amis et super
philosophes. Merci les gars d'avoir vérifié les formules que j'ai incluses dans le livre.
Les erreurs sont, encore une fois, entièrement de ma responsabilité. Merci également à Bjorn Schem-
bera, un vrai philosophe déguisé, pour notre temps discutant de tant de questions techniques
sur les simulations informatiques, dont certaines ont trouvé une place dans le livre. Du département
visualisation du HLRS, merci à Martin Aumuller, Thomas Obst, Wolfgang Schotte et Uwe Woessner
qui m'ont patiemment expliqué les nombreux détails de leur travail et fourni les images pour la Réalité
Augmentée et Virtuelle.
Réalité discutée dans le chapitre sur la visualisation. Les images de la tornade qui sont
également dans ce chapitre ont été fournies par le National Center for Supercomputing Applications,
University of Illinois at Urbana-Champaign. Pour cela, je suis très endetté
à Barbara Jewett pour sa patience, son temps et son aide précieuse dans la recherche des images.

Je voudrais également exprimer ma gratitude à mon ancien directeur de thèse, Ulla Pompe Alama, pour
ses encouragements et ses suggestions sur les premières ébauches. Des remerciements particuliers vont à
Raphael van Riel et l'Université de Duisburg-Essen pour leur soutien sur un court
bourse à terme. Pour plusieurs raisons, directement et indirectement liées au livre, je
Je suis endetté envers Mauricio Villasenor, Jordi Valverd u, Leandro Giri, Veronica Pedersen, Manuel
Barrantes, Itati Branca, Ramon Alvarado, Johannes Lenhard et Claus Beisbart. Merci à tous pour vos
commentaires, suggestions, encouragements lors de la
différentes étapes du livre, et pour m'avoir eu toutes sortes de conversations philosophiques avec
moi. Angela Lahee, mon éditrice chez Springer, mérite beaucoup de crédit pour elle
patience, encouragements et soutien utile dans la production de ce livre. Même si j'aurais continué à
peaufiner les idées de ce livre et, tout aussi important,
mon anglais, il est temps d'y mettre un terme.

Mon immense gratitude va à Tuncer Oren, avec qui j'ai partagé de nombreuses correspondances
sur les problèmes éthiques et moraux des simulations informatiques, aboutissant à la
dernier chapitre du livre. L'amour et le dévouement du professeur Oren pour les études philosophiques
sur la simulation informatique sont une source d'inspiration.

Enfin, à une époque où la science et la technologie sont incontestablement un élément fondamental
outil pour le progrès de la société, il est navrant de voir comment le gouvernement actuel de
l'Argentine - et de nombreux autres endroits en Amérique latine également - coupe
financement de la recherche scientifique et technologique, des sciences humaines et des sciences sociales. je
observent avec une égale horreur les décisions politiques visant explicitement la destruction du
système éducatif. Je dédie alors ce livre à la science argentine
et technologique, car ils ont montré à maintes reprises leur grandeur et leur éclat malgré des
conditions défavorables.

Ce livre doit trop à Kassandra. Elle m'a laissé son empreinte en corrigeant mon anglais, en me suggérant de réécrire tout un paragraphe, et en lâchant un rendez-vous planifié que j'avais oublié en terminant une section. Pour cela, et pour des milliers d'autres raisons, ce livre lui est entièrement dédié.

Contenu

Introduction

En 2009, un débat a éclaté autour de la question de savoir si les simulations informatiques introduisent de nouveaux problèmes philosophiques ou si elles ne sont qu'une nouveauté scientifique.
Roman Frigg et Julian Reiss, deux philosophes éminents qui ont déclenché le débat, ont noté que les philosophes ont largement assumé une certaine forme de nouveauté philosophique des simulations informatiques sans réellement se poser la question de sa possibilité.
Une telle hypothèse reposait sur une simple confusion : les philosophes pensaient que la nouveauté scientifique autorise la nouveauté philosophique. Cela a conduit à émettre un avertissement sur la croissance des affirmations exagérées et généralement injustifiées sur l'importance philosophique des simulations informatiques. Cette croissance, selon les auteurs, s'est reflétée dans le nombre croissant de philosophes convaincus que la philosophie des sciences, nourrie par les simulations informatiques, exigeait une épistémologie entièrement nouvelle, une ontologie révisée et une nouvelle sémantique.

Il est important de souligner que Frigg et Reiss ne s'opposent pas à la nouveauté des simulations informatiques dans la pratique scientifique et technique, ni à leur importance dans l'avancement de la science, mais plutôt que les simulations soulèvent peu de nouvelles questions philosophiques, voire aucune. Selon leurs propres mots, « [l]es problèmes philosophiques qui se posent en relation avec les simulations ne sont pas spécifiques aux simulations et la plupart d'entre eux sont des variantes de problèmes qui ont déjà été discutés dans d'autres contextes. Cela ne veut pas dire que les simulations ne soulèvent pas de nouveaux problèmes en elles-mêmes. Ces problèmes spécifiques sont cependant pour la plupart de nature mathématique ou psychologique, et non philosophique » (Frigg et Reiss 2009, 595).

Je partage la perplexité de Frigg et Reiss sur cette question. Il est difficile de croire qu'une nouvelle méthode scientifique – instrument, mécanisme, etc. – aussi puissante soit-elle, puisse à elle seule mettre en péril la philosophie actuelle des sciences et des techniques au point de devoir les réécrire. Mais cela n'est vrai que si l'on accepte l'affirmation selon laquelle les simulations informatiques viennent réécrire des disciplines de longue date, ce que je ne pense pas que ce soit le cas. Pour moi, si nous sommes capables de reconstruire et de donner un nouveau sens à de vieux problèmes philosophiques à la lumière de simulations informatiques, alors nous établissons fondamentalement leur nouveauté philosophique.

Posons-nous maintenant la question en quel sens les simulations informatiques sont-elles une nouveauté philosophique ? Il existe deux manières de résoudre le problème. Soit une simulation informatique

tations posent une série de questions philosophiques qui échappent aux standards philosophiques.
traitement, auquel cas ils peuvent être ajoutés à notre corpus philosophique; ou ils
défient les idées philosophiques établies, auquel cas le corpus actuel s'élargit
débats classiques dans de nouveaux domaines. Le premier cas a été proposé par (Huṃphreys
2009), alors que le second cas a été plaidé par moi-même (Duran, en cours d'examen).
Permettez-moi maintenant d'expliquer brièvement pourquoi les simulations informatiques représentent, à bien des égards,
une nouveauté scientifique et philosophique.

Le cœur de l'argument de Humphreys est de reconnaître que nous pouvons soit comprendre
simulations informatiques en se concentrant sur la façon dont la philosophie traditionnelle éclaire leur
l'étude (par exemple, à travers une philosophie des modèles, ou une philosophie de l'expérimentation), ou par
se concentrant exclusivement sur les aspects relatifs aux simulations informatiques qui constituent, dans et
à eux seuls, de véritables défis philosophiques. C'est cette deuxième façon de voir
les questions sur leur nouveauté qui accorde une importance philosophique à l'informatique
simulations.

La principale affirmation ici est que les simulations informatiques peuvent résoudre des problèmes autrement insolubles
modèles et ainsi amplifier nos capacités cognitives. Mais une telle amplification s'accompagne
un prix « pour un nombre croissant de domaines scientifiques, une épistémologie exclusivement
anthropocentrique n'est plus de mise car il existe désormais des autorités épistémiques supérieures, non
humaines » (Humphreys 2009 : 617). Humphreys appelle cela le
situation anthropocentrique comme moyen d'illustrer les tendances actuelles de la science et de l'ingénierie
où les simulations informatiques éloignent les humains du centre de
production de connaissances. Selon lui, un bref aperçu sur l'histoire de la philosophie des sciences montre
que l'homme a toujours été au centre de la production.
de la connaissance. Cette conclusion inclut la période du positivisme logique et empirique, où les sens
humains étaient l'autorité ultime (616). Une conclusion similaire découle de l'analyse d'alternatives à
l'empirisme, telles que la théorie de Quine.
et les épistémologies de Kuhn.

Confronté aux affirmations sur la nouveauté philosophique des simulations informatiques, Humphreys
souligne que le point de vue empiriste standard a empêché
une séparation complète entre les humains et leur capacité à évaluer et à produire
savoir scientifique. La situation anthropocentrique vient alors mettre en évidence
précisément cette séparation : c'est l'affirmation que les humains ont perdu leur position privilégiée en tant
qu'autorité épistémique ultime1 .
l'opinion selon laquelle la pratique scientifique ne progresse que parce que de nouvelles méthodes sont
disponibles pour traiter de grandes quantités d'informations. Traitement des informations, selon
Humphreys, est la clé du progrès de la science aujourd'hui, qui ne peut être atteint que si les humains sont
éloignés du centre de l'activité épistémique (Humphreys
2004, 8).

La situation anthropocentrique, aussi pertinente philosophiquement qu'en elle-même,
apporte également quatre nouveautés supplémentaires non analysées par la philosophie traditionnelle de
science. Ce sont l'opacité épistémique, la dynamique temporelle des simulations, le seman

---

[1] Humphreys fait une autre distinction entre la pratique scientifique entièrement réalisée par ordinateurs - celui qu'il appelle le scénario automatisé - et celui dans lequel les ordinateurs ne réaliser une activité scientifique, c'est-à-dire le scénario hybride. Il limite cependant son analyse aux scénario hybride (Humphreys 2009, 616-617).

tics, et la distinction en pratique/en principe. Tous les quatre sont de nouveaux problèmes philosophiques soulevés par des simulations informatiques; tous les quatre n'ont pas de réponse dans les récits philosophiques traditionnels des modèles et de l'expérimentation ; et tous quatre représentent un défi pour la philosophie des sciences.

La première nouveauté est l'opacité épistémique, un sujet qui attire actuellement beaucoup l'attention des philosophes. Bien que je discute de cette question en détail dans la section 4.3.2, une brève mention des hypothèses de base derrière l'opacité épistémique éclairera quelque peu la nouveauté des simulations informatiques. L'opacité épistémique est donc la position philosophique selon laquelle il est impossible pour un humain de connaître tous les éléments épistémiquement pertinents d'une simulation informatique. Humphreys présente ce point de la manière suivante : « Un processus est essentiellement épistémiquement opaque pour [un agent cognitif] X si et seulement s'il est impossible, étant donné la nature de X, pour X de connaître tous les éléments épistémiquement pertinents du processus. » (Humphreys 2009, 618). Pour mettre la même idée sous une forme différente, si un agent cognitif pouvait arrêter la simulation informatique et jeter un coup d'œil à l'intérieur, il ne serait pas en mesure de connaître les états antérieurs du processus, de reconstruire la simulation jusqu'au point d'arrêt, ou prédire les états futurs étant donné les états précédents. Être épistémiquement opaque signifie qu'en raison de la complexité et de la rapidité du processus de calcul, aucun agent cognitif ne peut savoir ce qui fait d'une simulation un processus épistémiquement pertinent.

Une deuxième nouveauté liée à l'opacité épistémique est la dynamique temporelle des simulations informatiques. Ce concept a deux interprétations possibles. Soit il fait référence au temps de calcul nécessaire pour résoudre le modèle de simulation, soit il représente le développement temporel du système cible tel que représenté dans le modèle de simulation.
Un bon exemple qui fusionne ces deux idées est une simulation de l'atmosphère : le modèle de simulation représente la dynamique de l'atmosphère pendant un an et il faut, disons, dix jours pour le calculer.

Ces deux nouveautés illustrent bien ce qui est typique des simulations informatiques, à savoir la complexité inhérente des simulations en elles-mêmes, comme c'est le cas de l'opacité épistémologique et de la première interprétation de la dynamique temporelle ; et la complexité inhérente des systèmes cibles que les simulations informatiques représentent habituellement, comme c'est le cas de la deuxième interprétation de la dynamique temporelle. Ce qui est commun entre ces deux nouveautés, c'est qu'elles consacrent toutes deux les ordinateurs comme autorité épistémique puisqu'ils sont capables de produire des résultats fiables qu'aucun humain ou groupe d'humains ne pourrait produire par lui-même. Soit parce que le processus informatique est trop complexe à suivre, soit parce que le système cible est trop complexe à comprendre, les ordinateurs deviennent la source exclusive d'informations sur le monde.

La deuxième interprétation de la dynamique temporelle est adaptée à la nouveauté de la sémantique, qui pose la question de la façon dont les théories et les modèles représentent le monde, ajustant maintenant l'image pour s'adapter à un algorithme informatique. Ainsi, la principale question ici est de savoir comment la syntaxe d'un algorithme informatique correspond au monde et comment une théorie donnée est réellement mise en contact avec des données.

Enfin, la distinction principe/pratique vise à distinguer ce qui est applicable en pratique et ce qui ne l'est qu'en principe. Pour Humphreys, c'est une fantaisie philosophique de dire qu'en principe, tous les modèles mathématiques trouvent une solution dans les simulations informatiques (623). C'est un fantasme parce que c'est clairement faux,

bien que les philosophes aient revendiqué sa possibilité – donc, en principe. Humphreys suggère plutôt qu'en abordant les ordinateurs, les philosophes doivent garder une attitude plus terre-à-terre, limitée aux contraintes techniques et empiriques que peuvent offrir les simulations.

Ma position est complémentaire de celle de Humphreys dans le sens où elle montre comment les simulations informatiques défient les idées établies en philosophie des sciences. À cette fin, je commence par plaider pour une manière spécifique de comprendre les modèles de simulation, le type de modèle à la base des simulations informatiques. Pour moi, un modèle de simulation refond une multiplicité de modèles en un « super-modèle ». C'est-à-dire que les modèles de simulation sont un amalgame de différents types de modèles informatiques, tous ayant leurs propres échelles, paramètres d'entrée et protocoles. Dans ce contexte, je revendique trois nouveautés en philosophie, à savoir la représentation, l'abstraction et l'explication.

À propos de la première nouveauté, je prétends que la multiplicité des modèles implique que la représentation d'un système cible est plus holistique dans le sens où elle englobe tous les modèles implémentés dans le modèle de simulation. Pour mettre la même idée sous une forme assez différente, la représentation du modèle de simulation n'est pas donnée par un modèle implémenté individuel mais plutôt par la combinaison de tous.

Le défi que les simulations informatiques apportent à la notion d'abstraction et d'idéalisation est que, généralement, cette dernière présuppose une certaine forme de position de négligence. Ainsi, l'abstraction vise à ignorer les caractéristiques concrètes que possède le système cible afin de se concentrer sur leur configuration formelle ; les idéalisations, d'autre part, se présentent sous deux formes : alors que les idéalisations aristotéliciennes consistent à « supprimer » des propriétés que nous pensons ne pas être pertinentes pour nos objectifs, les idéalisations galiléennes impliquent des distorsions délibérées. Maintenant, afin de mettre en œuvre la variété requise de modèles dans un seul modèle de simulation, il est important de compter sur des techniques par lesquelles les informations sont cachées aux utilisateurs, mais non négligées des modèles (Colburn et Shute 2007). C'est-à-dire que les propriétés, structures, opérations, relations et similaires présentes dans chaque modèle mathématique peuvent être efficacement mises en œuvre dans le modèle de simulation sans indiquer explicitement comment une telle mise en œuvre est effectuée.

Enfin, l'explication scientifique est un sujet philosophique séculaire où beaucoup a été dit. En ce qui concerne l'explication dans les simulations informatiques, cependant, je propose un regard assez différent sur la question que les offres de traitement standard. Un point intéressant ici est que, dans l'idée classique que l'explication est d'un phénomène du monde réel, je m'oppose à l'affirmation selon laquelle l'explication est, d'abord et avant tout, des résultats de simulations informatiques. Dans ce contexte, de nombreuses nouvelles questions émergent en quête de réponse. Je discute de l'explication scientifique plus en détail dans la section 5.1.1.

Comme je l'ai déjà mentionné, je crois que les simulations informatiques soulèvent de nouvelles questions pour la philosophie des sciences. Ce livre est la preuve vivante de cette conviction. Mais même si nous ne croyons pas à leur nouveauté philosophique, nous devons encore comprendre les simulations informatiques comme des nouveautés scientifiques avec un œil critique et philosophique. À ces fins, ce livre présente et discute plusieurs questions théoriques et philosophiques au cœur des simulations informatiques. Après avoir dit tout cela, nous pouvons maintenant nous plonger dans leurs pages.

Les références

Colburn, Timothy et Gary Shute. 2007. "Abstraction en informatique." Esprits et Machines 17, non. 2 (juin): 169–184.

Duran, Juan M. en cours de révision. « La nouveauté des simulations informatiques : de nouveaux défis pour la philosophie des sciences. à l'étude.

Frigg, Roman et Julian Reiss. 2009. « La philosophie de la simulation : Hot New Problèmes ou même vieux ragoût ? » Synthèse 169 (3): 593–613.

Humphreys, Paul W. 2004. Nous étendre: science computationnelle, Empiri cisme et méthode scientifique. Presse universitaire d'Oxford.

————. 2009. "La nouveauté philosophique des méthodes de simulation par ordinateur." Syn ces 169 (3): 615–626.

Chapitre 1
L'univers des simulations informatiques

L'univers des simulations informatiques est vaste, florissant dans presque toutes les disciplines scientifiques, et résiste encore à une conceptualisation générale. Depuis les premiers calculs de l'orbite de la Lune effectués par des machines à cartes perforées jusqu'aux tentatives les plus récentes de simulation d'états quantiques, les simulations informatiques ont une histoire particulièrement courte mais très riche.

On peut situer la première utilisation d'une machine à des fins scientifiques en Angleterre à la fin des années 1920. Plus précisément, c'est en 1928 que le jeune astronome et pionnier dans l'utilisation des machines Leslie J. Comrie prédit le mouvement de la Lune pour les années 1935 à 2000. Au cours de cette année, Comrie fait un usage intensif d'une machine à cartes perforées Herman Hollerith. pour calculer la sommation des termes harmoniques dans la prédiction de l'orbite de la Lune. Un tel travail révolutionnaire ne resterait pas dans l'ombre et, au milieu des années 1930, il avait traversé l'océan jusqu'à l'Université Columbia à New York. C'est là que Wallace Eckert fonda un laboratoire qui utilisait des tabulatrices à cartes perforées - aujourd'hui construites par IBM - pour effectuer des calculs liés à la recherche astronomique, y compris bien sûr une étude approfondie du mouvement de la Lune.

Les utilisations par Comrie et Eckert des machines à cartes perforées partagent quelques points communs avec l'utilisation actuelle des simulations. Plus important encore, les deux implémentent un type spécial de modèle qui décrit le comportement d'un système cible et qui peut être interprété et calculé par une machine. Alors que l'informatique de Comrie rendait des données sur les mouvements de la Lune, la simulation d'Eckert décrivait le mouvement planétaire.

Ces méthodes ont certainement été les pionnières et révolutionnées leurs domaines respectifs, ainsi que de nombreuses autres branches des sciences naturelles et sociales. Cependant, les simulations de Comrie et d'Eckert diffèrent considérablement des simulations informatiques d'aujourd'hui. En y regardant de plus près, des différences peuvent être trouvées partout. L'introduction de circuits à base de silicium, ainsi que la normalisation ultérieure de la carte de circuit imprimé, ont contribué de manière significative à la croissance de la puissance de calcul. L'augmentation de la vitesse de calcul, de la taille de la mémoire et de la puissance expressive du langage de programmation a fortement remis en question les idées reçues sur la nature du calcul et de son domaine d'application. Les machines à cartes perforées sont rapidement devenues obsolètes car lentes, peu fiables dans leurs résultats, limitées dans leur programmation,

et basé sur une technologie rigide (par exemple, il y avait très peu de modules échangeables). Dans
En fait, un inconvénient majeur de la carte perforée par rapport aux ordinateurs modernes est qu'ils
sont des machines sujettes aux erreurs et chronophages, et donc la fiabilité de leur
les résultats ainsi que leur exactitude de représentation sont difficiles à justifier. Cependant, la différence peut-être
la plus radicale entre les simulations de Comrie et d'Eckert, sur le
d'une part, et les simulations informatiques modernes d'autre part, est le processus d'automatisation
qui caractérise ce dernier. Dans les simulations informatiques d'aujourd'hui, les chercheurs perdent
motifs sur leur influence et leur pouvoir d'interférer dans le processus de calcul, et
cela deviendra plus important à mesure que la complexité et la puissance de calcul augmenteront.

Les ordinateurs modernes viennent modifier de nombreux aspects de la science et de l'ingénierie
pratique avec des calculs plus précis et des représentations plus précises. Précision, puissance de calcul et
réduction des erreurs sont, comme nous le verrons, les principales clés
de simulations informatiques qui déverrouillent le monde.

À la lumière des ordinateurs contemporains, il n'est donc pas correct de soutenir que la prédiction de Comrie
du mouvement de la Lune et la solution d'Eckert des équations planétaires sont des simulations informatiques. Cela
ne veut évidemment pas dire qu'il ne s'agit pas du tout de simulations. Mais pour s'adapter à la façon dont les
scientifiques et les ingénieurs
utiliser le terme aujourd'hui, il ne suffit pas de pouvoir calculer un modèle spécial ou
pour produire certains types de résultats sur un système cible. Vitesse, stockage, langue
l'expressivité et la capacité à être (re)programmés sont des concepts majeurs pour la
notion moderne de simulation informatique.

Que sont alors les simulations informatiques ? C'est une question philosophiquement motivée qui a trouvé des
réponses différentes de la part des scientifiques, des ingénieurs et des philosophes.
L'hétérogénéité de leurs réponses rend explicite à quel point chaque chercheur
conçoit des simulations informatiques, comment leurs définitions varient d'une génération à l'autre
la suivante, et combien il est difficile d'arriver à une notion unifiée. C'est important,
cependant, pour avoir une bonne idée de leur nature. Discutons-en plus amplement.

1.1 Que sont les simulations informatiques ?

La littérature philosophique récente considère les simulations informatiques comme des aides pour surmonter
imperfections et limites de la cognition humaine. Ces imperfections et limitations sont adaptées aux contraintes
humaines naturelles de l'informatique, du traitement et de la
classer de grandes quantités de données. Paul Humphreys, l'un des premiers contemporains
philosophes d'aborder les simulations informatiques d'un point de vue purement philosophique, les prend alors
comme un « instrument d'amplification », c'est-à-dire un instrument qui accélère ce que
un être humain sans aide ne pourrait pas le faire par lui-même (Humphreys 2004, 110). Dans un sens similaire,
Margaret Morrison, énième figure centrale des études philosophiques sur ordinateur
simulations, considère que bien qu'il s'agisse d'une autre forme de modélisation, « étant donné
diverses fonctions de la simulation [...] on pourrait certainement la caractériser comme un type de
modélisation « améliorée » » (Morrison 2009, 47).

Les deux affirmations sont fondamentalement correctes. Les simulations informatiques calculent, analysent,
rendre et visualiser les données de nombreuses façons qui sont inaccessibles pour n'importe quel groupe d'humains.

hommes. Comparez, par exemple, le temps nécessaire à un humain pour identifier les potentiels antibiotiques pour les maladies infectieuses telles que l'anthrax, avec une simulation du ribosome en mouvement au détail atomique (Laboratoire 2015). Ou, si vous préférez, comparez n'importe quel ensemble de capacités de calcul humain avec les superordinateurs utilisés au High Performance Computing Center Stuttgart, domicile du Cray XC40 Hazel Hen avec une performance maximale de 7,42 pétaflops et une capacité de mémoire de 128 Go par nœud.1

Comme l'ont souligné Humphreys et Morrison, il existe différents sens dans lesquels les simulations informatiques renforcent nos capacités. Cela pourrait être en amplifiant nos compétences en calcul, comme le suggère Humphreys, ou en améliorant notre modélisation. capacités, comme le suggère Morrison.

On serait naturellement enclin à penser que les simulations informatiques amplifient notre capacité de calcul ainsi que d'améliorer nos capacités de modélisation. Cependant, un un rapide coup d'œil sur l'histoire du concept montre le contraire. Pour certains auteurs, une bonne La définition doit souligner l'importance de trouver des solutions à un modèle. Aux autres, la bonne définition centre l'attention sur la description des modèles de comportement d'un système cible. Selon la première interprétation, la puissance de calcul de la machine nous permet de résoudre des modèles qui, autrement, seraient analytiquement insolubles. Dans ce respect, une simulation informatique « amplifie » ou « renforce » nos capacités cognitives en fournir une puissance de calcul à ce qui est au-delà de notre portée cognitive. La notion de la simulation informatique dépend alors de la physique de l'ordinateur et fournit l'idée que le changement technologique repousse les limites de la science et recherche en ingénierie. Une telle affirmation est également historiquement fondée. Chez Hollerith machines à cartes perforées à l'ordinateur à base de silicium, l'augmentation de la physique la puissance des ordinateurs a permis aux scientifiques et aux ingénieurs de trouver différentes solutions à une variété de modèles. Permettez-moi d'appeler cette première interprétation le point de vue de la résolution de problèmes sur les simulations informatiques.

Dans la seconde interprétation, l'accent est mis sur la capacité de la simulation pour décrire un système cible. Pour cela, nous avons un langage puissant qui représente, à certains degrés de détail acceptables, plusieurs niveaux de description. À cet égard, une simulation informatique « amplifie » ou « améliore » nos capacités de modélisation en fournissant une représentation plus précise d'un système cible. Ainsi comprise, la notion de la simulation informatique est adaptée à la manière dont ils décrivent un système cible, et donc sur le langage informatique utilisé, les méthodes de modularisation, les techniques de génie logiciel, etc. J'appelle cette deuxième interprétation la description des motifs du point de vue du comportement sur des simulations informatiques.

Étant donné que les deux points de vue mettent l'accent sur des interprétations différentes - bien que pas nécessairement incompatibles - des simulations informatiques en tant qu'amplificateurs, certaines distinctions peut être dessiné. Pour commencer, du point de vue de la résolution de problèmes, les simulations informatiques ne sont pas des expériences au sens traditionnel du terme, mais plutôt la manipulation d'un

---

1 Il convient de noter que l'activité de notre réseau neuronal est, dans certains cas spécifiques, plus rapide que n'importe quel supercalculateur. Selon une publication relativement récente, l'ordinateur japonais Fujitsu K, composé de 82 944 processeurs, prend environ 40 minutes pour simuler une seconde d'activité du réseau neuronal en temps réel, biologique. Afin de simuler partiellement l'activité neuronale humaine, les chercheurs créent environ 1,73 milliard de cellules nerveuses virtuelles connectées à 10,4 billions de synapses virtuelles (Himeno 2013).

structure abstraite et formelle (c.-à-d. modèles mathématiques). En fait, pour de nombreux partisans
De ce point de vue, la pratique expérimentale est confinée au laboratoire traditionnel comme
les simulations informatiques sont plus une pratique de calcul des nombres, plus proche des mathématiques et
de la logique. La description des schémas de comportement du point de vue, d'autre part
d'autre part, nous permet de traiter les simulations informatiques comme des expériences d'une manière simple
sens. L'intuition sous-jacente est qu'en décrivant le comportement d'un
système cible, les chercheurs sont capables de réaliser quelque chose de très similaire à la pratique
expérimentale traditionnelle, comme mesurer des valeurs, observer des quantités et
détection d'entités.

Comprendre les choses de cette manière a quelque parenté avec la méthodologie des simulations
informatiques. Comme je le dis plus loin, le point de vue de la technique de résolution de problèmes considère
qu'une simulation est l'implémentation directe d'un modèle sur un ordinateur physique. C'est-à-dire que des
modèles mathématiques sont implémentés sur le sim pliciter informatique. La description des modèles de point
de vue du comportement, au contraire, soutient que les simulations informatiques ont une méthodologie
appropriée qui est plutôt différente de tout ce que nous avons vu dans le domaine scientifique et technique. Ces
méthodologies
les différences entre les deux points de vue s'avèrent être au centre des différends ultérieurs sur
la nouveauté des simulations informatiques dans la recherche scientifique et technique.

Une autre différence entre ces deux points de vue réside dans les raisons d'utiliser
simulations informatiques. Alors que le point de vue de la résolution de problèmes affirme que l'utilisation
des simulations informatiques ne se justifie pragmatiquement que lorsque le modèle ne peut être
Résolue par des méthodes plus traditionnelles, la description des schémas de comportement du point de vue
considère que les simulations informatiques offrent des informations précieuses sur le système cible malgré son
intraitabilité analytique. Notons que ce qui est également en jeu ici
est la priorité épistémique d'une méthode sur une autre. Si l'utilisation de simulations informatiques n'est justifiée
que lorsque le modèle ne peut pas être résolu analytiquement - autant de
partisans du point de vue de la résolution de problèmes – alors les méthodes analytiques sont épistémiquement
supérieures aux calculatoires. Cela configure un point de vue spécifique concernant
la place que les simulations informatiques occupent dans l'agenda scientifique et technique.
En particulier, il minimise considérablement la fiabilité des simulations informatiques pour
recherche en territoire inconnu. Nous aurons plus à dire à ce sujet tout au long de ce
livre.2

Enfin, l'approche des problèmes dans les simulations informatiques peut être très différente selon le point
de vue adopté. Pour le point de vue de la résolution de problèmes, tous les problèmes liés
aux résultats de la simulation (par exemple, précision, calculabilité, représentabilité, etc.)
peut être résolu pour des raisons techniques (c'est-à-dire en augmentant la vitesse et la mémoire, en changeant
l'architecture sous-jacente, etc.). Au lieu de cela, pour l'avocat de la description
des modèles de comportement, les mêmes questions ont un traitement entièrement différent. Les résultats
incorrects, par exemple, sont abordés en analysant des considérations pratiques
au niveau de la conception, telles que de nouvelles spécifications pour le système cible, des évaluations
alternatives des connaissances d'expertise, de nouveaux langages de programmation, etc. Dans le même ordre d'idées,
des résultats incorrects peuvent être dus à une fausse représentation du système cible au niveau de

---

2 De nombreux philosophes ont essayé de comprendre la nature des simulations informatiques. Ce que j'ai
proposé ci-dessus n'est qu'une des caractéristiques possibles. Pour plus d'informations, le lecteur pourra se référer à ce qui suit
auteurs (Winsberg 2010 ; Vallverdu 2014 ; Morrison 2015 ; Winsberg 2015 ; Saam 2016).

les étapes de signe, de spécification et de programmation (voir section 2.2), ou de fausses représentations à l'étape de calcul (par exemple, des erreurs pendant le temps de calcul – voir section 4.3). La compréhension générale du chercheur ainsi que la solution à ces problèmes changent considérablement selon le point de vue adopté.

Pour illustrer certains des points soulevés jusqu'à présent, prenons une simple simulation informatique basée sur une équation de la dynamique d'un satellite en orbite autour d'une planète sous l'effet des marées. Pour simuler une telle dynamique, les chercheurs commencent généralement par un modèle mathématique du système cible. Un bon modèle est fourni par la mécanique newtonienne classique, telle que décrite par MM Woolfson et GJ Pert dans (Woolfson et Pert 1999b).

Pour une planète de masse M et un satellite de masse m (M), sur une orbite de demi-grand axe a et d'excentricité e, l'énergie totale est

$$E = -\frac{GMm}{2a} \tag{1.1}$$

et le moment cinétique est

$$H = \{GMa(1-e^{2})\}m \tag{1.2}$$

Si E doit diminuer, alors a doit devenir plus petit ; mais si H est constant, alors e doit devenir plus petit, c'est-à-dire que l'orbite doit s'arrondir. La quantité qui reste constante est a(1−e est décrit par une masse ^{2}), le semi-latus rectum comme le montre la figure 1.1. La planète ponctuelle, P, et le satellite par une distribution de trois masses, chacune m/3, aux positions S1,S2 et S3, formant un triangle équilatéral lorsqu'il est libre de contrainte. Les masses sont reliées, comme indiqué, par des ressorts, chacun de longueur non contrainte l et de même constante de ressort, k (figure 1.2). Ainsi, un ressort constamment étiré à une longueur l ^{0} exercera une force vers l'intérieur égale à

$$F = k(l^{0} - l) \tag{1.3}$$

Il est également important d'introduire un élément dissipatif dans le système en faisant dépendre la force du taux de dilatation ou de contraction du ressort, donnant la loi de force suivante :

$$F = k(l^{0} - l) - c\,\frac{dl0}{dt} \tag{1.4}$$

où la force agit vers l'intérieur aux deux extrémités. C'est le deuxième terme de l'équation 1.4 qui donne la simulation des pertes par hystérésis dans le satellite (18-19).

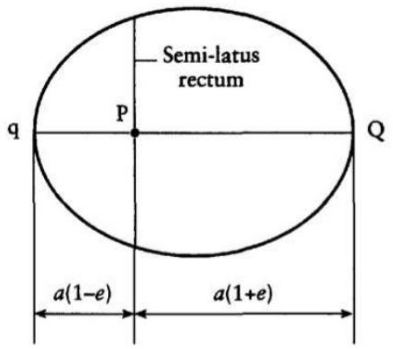


Fig. 1.1 L'orbite elliptique d'un satellite par rapport à la planète à un foyer. Les points q et Q sont respectivement les points les plus proches et les plus éloignés de la planète. (Woolfson et Pert 1999b, 19)

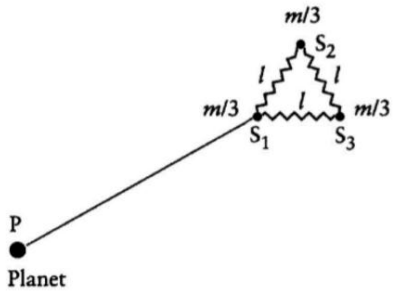


Fig. 1.2 Le satellite est décrit par trois masses, chacune m/3, reliées par des ressorts chacun de même longueur libre, l. (Woolfson et Pert 1999b, 19)

Il s'agit d'un modèle de simulation informatique d'un satellite en orbite autour d'une planète soumise au stress des marées. Le satellite s'étire le long du rayon vecteur de façon périodique, à condition que l'orbite soit non circulaire. Etant donné que le satellite n'est pas parfaitement élastique, il y aura des effets d'hystérésis et une partie de l'énergie mécanique sera convertie en chaleur et rayonnée. À toutes fins utiles, néanmoins, la simulation spécifie entièrement le système cible.

Les équations 1.1 à 1.4 sont une description générale du système cible. Puisque l'intention est de simuler un phénomène spécifique du monde réel avec des caractéristiques concrètes, il doit être isolé en définissant les valeurs des paramètres de la simulation. Dans le cas de Woolfson et Pert, ils utilisent l'ensemble suivant de valeurs de paramètres (Woolfson et Pert 1999b, 20) :

1. nombre de corps = 4

2. masse du premier corps (planète) = 2 x 1027 kg 3.
masse du satellite = 3 x 1022 kg 4. pas
de temps initial = 10 s 5.
temps total de simulation = 125000 s 6.
corps choisi comme origine = 1
7 .tolérance = 100 m 8.
distance initiale du satellite = 1 x 108 m 9.
longueur non tendue du ressort = 1 x 106 m 10.
excentricité initiale = 0,6

Ces paramètres définissent une simulation informatique d'un satellite de la taille de Triton, la plus grande lune de Neptune en orbite autour d'une planète avec une masse proche de celle de Jupiter - y compris, bien sûr, une contrainte de marée spécifique, des effets d'hystérésis, etc. Si les paramètres ont été modifiés, alors naturellement la simulation est celle d'un autre phénomène - bien que toujours d'une interaction à deux corps utilisant la mécanique newtonienne.

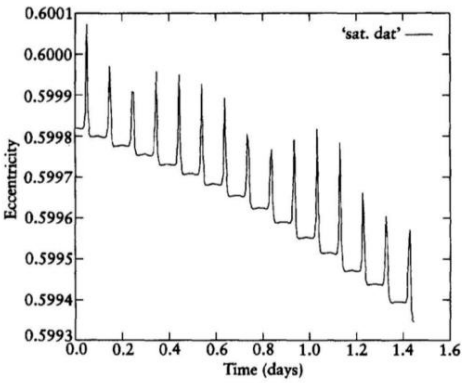


Fig. 1.3 L'excentricité orbitale en fonction du temps (Woolfson et Pert 1999b, 20)

Voici un extrait du code correspondant aux modèles mathématiques ci-dessus tel que programmé en FORTRAN par Woolfson et Pert.3

```
      PROGRAMME NBODY

[...]

      C LES VALEURS DE A ET E SONT CALCULÉES TOUTES LES 100 ÉTAPES
      C ET SONT MÉMORISÉS AVEC L'HEURE.
      C

            IST=IST+1 SI((IST/
            100)*100.NE.IST)ALLER A 50 IG=IST/100 SI(IG.GT.1000)ALLER
            A 50


      C
      C TROUVEZ D'ABORD LA POSITION ET LA VITESSE DU CENTRE DE MASSE
      SATELLITE C.
```

3 Pour le code complet, voir (Woolfson et Pert 1999a).

```
      C
            DO 1 K=1,3
            POS(K)=0
            VEL(K)=0
            DO 2 J=2,NB
            POS(K)=POS(K)+X(J,K)
            VEL(K)=VEL(K)+V(J,K)
            2 CONTINUER
            POS(K)=POS(K)/(NB-1.0)
            VEL(K)=VEL(K)/(NB-1.0)
            1 CONTINUER

      C
      C CALCULER LA DISTANCE ORBITALE
      C
            R=SQRT(POS(1)**2+POS(2)**2+POS(3)**2)

      C
      C CALCULER V**2
      C
            V2=VEL(1)**2+VEL(2)**2+VEL(3)**2
      C
      C CALCULER L'ÉNERGIE INTRINSÈQUE
      C
      TOTM=CM(1)+CM(2)+CM(3)+CM(4)
            EN=-G*TOTM/R+0.5*V2

[...]

            DÉMARRAGE DU SOUS-PROGRAMME
            DIMENSION X(20,3),V(20,3),STOCK(1000),STSTORE(1000),
            +STORE(1000),XTEMP(2,20,3),VTEMP(2,20,3),CM(20),XT(20,3),
            +VT(20,3), DELV(20,3)
[...]
            LIRE(5,*)NORIG

[...]

      C
      C LES VITESSES INITIALES SONT CALCULEES POUR LES TROIS COMPOSANTES
      C SATELLITE POUR QUE LA VITESSE ANGULAIRE DE SPIN DU SATELLITE SOIT
      C APPROXIMATIVEMENT ÉGALE À LA VITESSE ANGULAIRE ORBITALE.

      C VV=SQRT(G*(CM(1)+CM(2)+CM(3)+CM(4))*(1-ECC)/D)
      V(2,2)=VV*DIS2/
      DV(3,2)=VV*DIS3*COS(ANGLE)/D [...]


RETOUR
FIN
```

Comme discuté précédemment, l'une des caractéristiques du point de vue de la résolution de problèmes est que le modèle mathématique ci-dessus peut être directement implémenté sur l'ordinateur. C'est-à-dire que le code présenté ici correspond dans toute sa mesure au modèle mathématique. En principe, rien – pertinent – n'est ajouté et rien – pertinent – n'est éliminé. Ainsi, la seule raison d'utiliser un ordinateur est de trouver l'ensemble des solutions du modèle de manière plus rapide et moins coûteuse. La description des modèles de point de vue comportemental, d'autre part, reconnaît l'existence d'un processus de mise en œuvre du modèle sous forme de simulation informatique, qui comprend la transformation du modèle mathématique en un type de modèle plutôt différent.

Comme nous pouvons le voir, les deux positions sont fondées. Le point de vue de la résolution de problèmes a raison de prétendre que les simulations informatiques doivent refléter le modèle mathématique mis en œuvre, sinon des problèmes de représentation, de fiabilité, etc. apparaissent. La description des modèles de point de vue du comportement, d'autre part, reflète la pratique scientifique et d'ingénierie plus sans ambiguïté.

Avant de continuer, voici un bon endroit pour introduire une nouvelle terminologie. Appelons « modèles mathématiques » les modèles utilisés dans les domaines scientifiques et techniques qui utilisent le langage mathématique. Les équations ci-dessus en sont des exemples. Appelons « modèles de simulation » ces modèles mis en œuvre sur un ordinateur - comme

une simulation informatique – qui utilise un langage de programmation.4 Un exemple
d'un modèle de simulation est le code ci-dessus.

J'ai commencé ce chapitre en distinguant deux points de vue de l'informatique
simulations. D'une part, le point de vue de la résolution de problèmes qui met l'accent sur la
côté informatique des simulations ; d'autre part, la description des modèles de
point de vue comportemental qui met l'accent sur les représentations des systèmes cibles.
Comme mentionné précédemment, le point de vue de la résolution de problèmes ne néglige pas la
représentation du système cible, pas plus que la description des modèles de comportement.
ne considèrent pas le calcul comme un problème central des simulations informatiques. Aucun point de vue
ne reflète une position « tout ou rien » - un bon exemple en est la définition donnée
par Thomas H. Naylor, Donald S. Burdick et W. Earl Sasser où ils défendent une
modèles de comportement à la page 1361 de (Naylor et al. 1967) et abonnez-vous également au point de
vue de la résolution de problèmes plus loin à la page 1319. La différence entre
ces deux points de vue résident, encore une fois, dans les principales caractéristiques mises en évidence par chaque compte.
Voyons si nous pouvons rendre cette distinction plus claire.5

1.1.1 Les simulations informatiques comme techniques de résolution de problèmes

Du point de vue de la résolution de problèmes, les simulations informatiques présentent généralement certains
des fonctionnalités suivantes. Tout d'abord, des simulations sont adoptées pour les cas où la cible
système est trop complexe pour être analysé par lui-même - appelez-le la caractéristique de complexité.
Deuxièmement, les simulations sont utiles dans les cas où le modèle mathématique sous-jacent
ne peut pas être résolu analytiquement - appelez cela la caractéristique d'absence d'analyse. Troisièmement, mathématique
les modèles sont directement implémentés sur l'ordinateur - appelez cela l'implémentation directe
fonctionnalité. Les caractéristiques de complexité et de manque d'analyse soulignent nos limites humaines
pour analyser certains types de modèles mathématiques, tout en améliorant
les simulations de puissance de calcul comme une vertu. La fonctionnalité de mise en œuvre directe
accompagne ces idées en affirmant qu'il n'y a pas de méthodologie médiatrice entre le modèle mathématique
et l'ordinateur physique. Plutôt, les équations de
le modèle mathématique sont implémentés – ou résolus – simpliciter sur la physique
ordinateur sous la forme d'une simulation informatique.

La littérature ancienne du point de vue de la résolution de problèmes présente une vision assez uniforme
point de vue sur la question. Pour la plupart, des philosophes professionnels, des scientifiques,
et les ingénieurs considèrent la puissance de calcul des simulations comme la clé débloquant leur

---

4 Au sens strict, un modèle de simulation est une structure plus complexe constituée, entre autres,
d'une spécification codée dans un langage de programmation comme un algorithme et finalement implémentée comme
un processus informatique. Bien que la même spécification puisse être écrite dans différents langages de
programmation et mise en œuvre par différentes architectures informatiques, elles sont toutes considérées comme identiques.
modèle de simulation. Ainsi compris, le langage de programmation ne détermine pas à lui seul
notion de « modèle de simulation ». J'aborderai ces questions plus en détail dans le chapitre 2.

5 Une introduction intéressante à l'histoire de l'informatique peut être trouvée dans les travaux de
(Ceruzzi 1998 ; De Mol et Primiero 2014, 2015), et notamment sur des simulations informatiques (Oren 2011b, 2011a).

pouvoir épistémique. Un bon premier exemple est la définition donnée par Claude McMillan et Richard Gonzales en 1965. Dans leur travail, les auteurs énoncent quatre points caractéristiques des simulations, à savoir

1. La simulation est une technique de résolution de problèmes.
2. C'est une méthode expérimentale.
3. L'application de la simulation est indiquée dans la solution des problèmes de (a) conception de systèmes (b) analyse des systèmes.
4. La simulation est utilisée lorsque les systèmes considérés ne peuvent pas être analysés utilisant des méthodes analytiques directes ou formelles. (McMillan et Gonzalez 1965) ´

Cette définition est, dans la mesure où j'ai pu la trouver, la première qui conçoit ouvertement simulations informatiques comme techniques de résolution de problèmes. Ce n'est pas seulement parce que le auteurs le disent explicitement dans leur premier point, mais parce que la définition adopte deux des trois caractéristiques standard de ce point de vue. Le point 4 est explicite sur l'utilisation de simulations pour trouver des solutions à des modèles mathématiques autrement insolubles, tandis que le point 3 suggère l'adoption de simulations pour la conception du système et le système car ils sont trop complexes pour être analysés par eux-mêmes (c'est-à-dire la complexité fonctionnalité).

Un an plus tard, Daniel Teichroew et John Francis Lubin présentent leur propre définition. Fait intéressant, cette définition rend trois caractéristiques de ce point de vue plus visible que toute autre définition dans la littérature. Les auteurs commencent par identifier ce qu'ils appellent des « problèmes de simulation », c'est-à-dire des problèmes qui sont traités par des techniques de simulation – nous verrons ensuite ce que sont ces techniques. Une simulation problème est essentiellement un problème mathématique avec de nombreuses variables, paramètres et fonctions qui ne peuvent pas être traitées analytiquement (c'est-à-dire la caractéristique de complexité) et donc les simulations informatiques sont la seule ressource disponible pour les chercheurs (c'est-à-dire la fonction d'absence d'analyse). La troisième caractéristique, l'implémentation directe d'un modèle mathématique, peut être trouvé à plusieurs endroits dans l'article. En fait, les auteurs classent deux types de modèles, à savoir, des modèles à changement continu (c'est-à-dire ceux qui utilisent des équations aux dérivées partielles ou des équations différentielles ordinaires) et des modèles à changement discret (c'est-à-dire, les modèles où les changements dans l'état du système sont discrets) (Teichroew et Lubin 1966, 724). Pour les auteurs, les deux types de modèles sont mis en œuvre directement sous forme de simulation informatique. Selon les propres mots des auteurs,

> Les problèmes de simulation sont caractérisés par le fait qu'ils sont mathématiquement insolubles et qu'ils ont résisté à une solution par des méthodes analytiques. Les problèmes impliquent généralement de nombreuses variables, de nombreuses paramètres, fonctions qui ne se comportent pas bien mathématiquement et variables aléatoires.
> Ainsi, la simulation est une technique de dernier recours. Pourtant, beaucoup d'efforts sont maintenant consacrés à « l'informatique simulation' car c'est une technique qui donne des réponses malgré ses difficultés, ses coûts et temps requis. (724)

Il y a une autre affirmation intéressante à souligner ici. Remarquons que les auteurs mettent en évidence le sentiment que les tenants de ce point de vue ont à l'égard de simulations informatiques : il s'agit d'une technique de dernier recours.6 C'est-à-dire que l'utilisation

6 En ce qui concerne ce dernier point, le professeur Oren a organisé en 1982 un institut d'études avancées de l'OTAN à Ottawa axé sur l'étude du contexte d'utilisation des simulations informatiques (communication personnelle). Voir par exemple les articles publiés dans (Oren, Zeigler et Elzas 1982 ; Oren 1984).

les simulations informatiques ne se justifient que lorsque les méthodes analytiques ne sont pas disponibles. Mais il s'agit plus d'un préjugé épistémologique contre les simulations informatiques que d'une vérité établie. Des travaux récents effectués par des philosophes montrent que, dans de nombreux cas, les chercheurs préfèrent les simulations informatiques aux méthodes analytiques. Ceci, bien sûr, pour les cas évidents où le système cible est insoluble – comme Teichroew et Lubin indiquer correctement - et où les solutions analytiques ne sont pas disponibles. Vincent Ar Dourel et Julie Jebeile soutiennent que les simulations informatiques pourraient même être supérieures à solutions analytiques dans le but de faire des prédictions quantitatives. Selon pour ces auteurs, « certaines solutions analytiques rendent les applications numériques difficiles ou impossible (...) les solutions analytiques sont parfois trop sophistiquées par rapport à le problème en jeu (...) [et] les méthodes analytiques n'offrent pas une approche générique pour résoudre des équations comme [les simulations informatiques le font] »7 (Ardourel et Jebeile 2017, 203).

Or, les partisans du point de vue de la résolution de problèmes sont également présents dans la littérature contemporaine. Une définition largement contestée – qui, malgré le changement d'auteur d'esprit a en quelque sorte réussi à devenir un standard dans la littérature - est Humphreys définition de travail : « Une simulation informatique est toute méthode mise en œuvre par ordinateur pour explorer les propriétés des modèles mathématiques où les méthodes analytiques sont indisponible » (Humphreys 1990, 501).

Ici, Humphreys nous donne deux caractéristiques des simulations informatiques en tant que solutionneurs de problèmes. Ce sont, les simulations sont des modèles mathématiques mis en œuvre sur un ordinateur, et ils sont utilisés lorsque les méthodes analytiques ne sont pas disponibles. Jusqu'à présent, Humphreys est un défenseur classique de la technique de résolution de problèmes. Cependant, un examen plus approfondi de la définition montre que les inquiétudes de Humphreys incluent également la nature du calcul. Voici pourquoi.

Plus tôt, j'ai mentionné que Naylor, Burdick et Sasser ont déclaré que les simulations informatiques sont des méthodes numériques mises en œuvre sur l'ordinateur. Humphreys, à la place, conçoit les simulations informatiques comme des méthodes mises en œuvre par ordinateur. La distinction n'est pas inutile, puisqu'elle en dit long sur la nature du calcul d'un modèle.
En fait, Humphreys insiste pour séparer trois notions différentes : les mathématiques numériques, les méthodes numériques et l'analyse numérique. Les mathématiques numériques sont branche des mathématiques concernée par l'obtention des valeurs numériques des solutions à un problème mathématique donné. Les méthodes numériques, quant à elles, sont des mathématiques numériques qui visent à trouver une solution approximative au modèle.
Enfin, l'analyse numérique est l'analyse théorique des méthodes numériques et la solutions calculées (502). Les méthodes numériques, par elles-mêmes, ne peuvent pas être directement liées aux simulations informatiques. Au moins deux fonctionnalités supplémentaires doivent être incluses. Premièrement, les méthodes numériques doivent être appliquées à un problème scientifique spécifique. C'est important car le modèle mis en œuvre n'est pas n'importe quel modèle, mais d'un genre spécifique (c.-à-d. modèles scientifiques et techniques). De cette façon, il n'y a pas de place pour confondre simulations informatiques réalisées dans une installation scientifique avec simulation informatique réalisées à des fins artistiques. Deuxièmement, la méthode doit être implémentée sur un vrai

7 Les auteurs identifient les « méthodes numériques » avec les « simulations informatiques » (Ardourel et Jebeile 2017, 202). Comme je le montre ensuite, ces deux concepts doivent rester séparés. Cependant, cela ne représentent une objection à leur demande principale.

informatique et calculable en temps réel. Cette deuxième caractéristique garantit que le modèle est adapté au calcul et conforme aux normes minimales de la recherche scientifique (par exemple, que le calcul se termine dans un délai raisonnable, que les résultats sont précis dans une certaine plage, etc.)

Bien qu'il n'ait été suggéré que comme définition de travail, Humphreys a reçu de vives objections qui l'ont pratiquement forcé à changer sa position d'origine. L'un des principaux critiques était Stephan Hartmann , qui a objecté que la définition de Humphreys manquait la nature dynamique des simulations informatiques. Hartmann a ensuite proposé sa propre définition :

> Les simulations sont étroitement liées aux modèles dynamiques. Plus concrètement, une simulation se produit lorsque les équations du modèle dynamique sous-jacent sont résolues. Ce modèle est conçu pour imiter l'évolution temporelle d'un système réel. Autrement dit, une simulation imite un processus par un autre processus. Dans cette définition, le terme « processus » désigne uniquement un objet ou un système dont l'état change dans le temps. Si la simulation est exécutée sur un ordinateur, on l'appelle une simulation informatique (Hartmann 1996, 83 - emphase dans l'original).

En simplifiant cette définition, on pourrait dire qu'une simulation informatique consiste à trouver l'ensemble des solutions d'un modèle dynamique à l'aide d'un ordinateur physique. Soulignons quelques hypothèses intéressantes. Premièrement, le modèle dynamique est conçu comme ne présentant aucune différence par rapport à un modèle mathématique. Ainsi compris, le modèle utilisé par MM Wolfson et GJ Pert pour simuler la dynamique d'un satellite orbitant autour d'une planète sous contrainte de marée est le modèle implémenté sur le calculateur physique. Deuxièmement, Hartmann ne s'inquiète pas trop des méthodes utilisées pour résoudre le modèle dynamique. Le papier et le crayon, les méthodes numériques et les méthodes mises en œuvre par ordinateur semblent tout aussi bien convenir. Cette préoccupation découle du fait que le même modèle dynamique est résolu par un agent humain ainsi que par l'ordinateur. À l'instar de Naylor, Burdick et Sasser, une telle hypothèse soulève des questions sur la nature du calcul.

Il est intéressant de noter que la définition de Hartmann a été chaleureusement accueillie par la communauté philosophique. La même année, Jerry Banks, John Carson et Barry Nelson ont présenté une définition similaire à celle de Hartmann, mettant également l'accent sur l'idée de la dynamique d'un processus dans le temps et de la représentation comme imitation. Ils la définissent de la manière suivante : « [une] simulation est l'imitation du fonctionnement d'un processus ou d'un système du monde réel au fil du temps. Qu'elle soit effectuée à la main ou sur un ordinateur, la simulation implique la génération d'une histoire artificielle d'un système et l'observation de cette histoire artificielle pour tirer des conclusions concernant les caractéristiques de fonctionnement du système réel » (Banks et al. 2010, 3) . Francesco Guala suit également Hartmann dans la distinction entre les modèles statiques et dynamiques, l'évolution temporelle d'un système et l'utilisation de simulations pour résoudre mathématiquement le modèle mis en œuvre (Guala 2002). Plus récemment, Wendy Parker y a fait explicitement référence en caractérisant une simulation comme « une séquence d'états ordonnée dans le temps qui sert de représentation d'une autre séquence d'états ordonnée dans le temps » (Parker 2009, 486).

Maintenant, malgré les différences entre Humphreys et Hartmann, ils sont également d'accord sur quelques points. En fait, ils considèrent tous les deux les simulations informatiques comme des équipements de calcul à grande vitesse capables d'améliorer notre capacité d'analyse pour résoudre des modèles mathématiques autrement insolubles. Après les objections initiales de Hartmann, Humphreys a inventé une nouvelle définition, cette fois basée sur la notion de calcul

modèle. Je discuterai des modèles dans la section suivante, car je pense que cette nouvelle conceptualisation des simulations informatiques se qualifie mieux pour les descriptions de modèles.
du point de vue du comportement.

Un résumé éclairant se trouve dans les travaux de Roman Frigg et Julian Reiss.
Selon les auteurs, il existe deux sens dans lesquels la notion de simulation informatique est définie dans la littérature actuelle. Il y a un sens étroit, où « 'simulation'
fait référence à l'utilisation d'un ordinateur pour résoudre une équation que nous ne pouvons pas résoudre analytiquement, ou plus généralement pour explorer les propriétés mathématiques des équations où l'analyse
les méthodes échouent ». Il existe également un sens large, où le terme « simulation » désigne
l'ensemble du processus de construction, d'utilisation et de justification d'un modèle qui implique des mathématiques analytiquement insolubles » (Frigg et Reiss 2009, 596). Pour moi, les deux sens
pourrait être inclus dans le cadre des techniques de résolution de problèmes du point de vue de l'ordinateur simulations.

Les deux catégories sont certainement méritoires et éclairantes. Les deux capturent généralement les nombreux sens dans lesquels les philosophes du point de vue de la résolution de problèmes définissent
la notion de simulation informatique. Alors que le sens étroit se concentre sur l'heuristique
capacité de simulations informatiques, le sens large met l'accent sur l'aspect méthodologique,
aspects épistémologiques et pragmatiques des simulations informatiques en tant que solutionneurs de problèmes. Passons maintenant à une manière différente de conceptualiser les simulations informatiques.

1.1.2 Les simulations informatiques comme description des modèles de comportement

La vision des simulations informatiques comme des techniques de résolution de problèmes contraste avec la vue des simulations comme description des modèles de comportement. Selon ce point de vue, les simulations informatiques visent principalement à décrire le comportement d'une cible
système dans lequel ils se développent ou se déploient. Comme mentionné précédemment, cela ne veut pas dire que la puissance de calcul des simulations est minimisée dans tous les sens. Les simulations informatiques en tant que résolveurs de problèmes ont bien compris ce point dans le sens où la vitesse, la mémoire,
et le contrôle sont des facteurs essentiels qui soulignent la nouveauté des simulations dans le domaine scientifique. et la pratique de l'ingénierie. Cependant, de ce point de vue, la puissance de calcul
des simulations est considérée comme une caractéristique de second niveau. En ce sens, au lieu de situer la valeur épistémologique des simulations informatiques dans leur capacité à résoudre
un modèle mathématique, leur valeur vient de la description des modèles de comportement des systèmes cibles.

Maintenant, qu'est-ce que les modèles ? Je les considère comme des descriptions qui reflètent les structures, les hommages, les performances et le comportement général du système cible dans un contexte spécifique. langue. Plus précisément, ces structures, attributs, etc. sont interprétés comme
concepts utilisés dans les sciences (par exemple, H2O, masse, etc.), les relations causales (par exemple, la collision de deux boules de billard), des nécessités naturelles et logiques (par exemple, qu'aucune sphère d'uranium a une masse supérieure à 100 000 kilogrammes8), lois, principes et

---

[8] Une précision s'impose ici. Un ordinateur n'est pas technologiquement altéré pour simuler un enrichissement sphère d'uranium d'une masse supérieure à 100 000 kilogrammes. Au contraire, le type de contraintes que nous trouver dans les ordinateurs sont liés à leurs propres limites physiques et à celles indiquées par les théories de

constantes de la nature. En bref, les modèles sont des descriptions d'un système cible qui utilisent le vocabulaire scientifique et technique. Naturellement, ces schémas s'appuient également sur des connaissances spécialisées, des « trucs du métier », des expériences passées et des préférences individuelles, sociétales et institutionnelles. En ce sens, pour ce point de vue, les simulations informatiques sont un conglomérat de concepts, de formules et d'interprétations qui facilitent la description des modèles de comportement d'un système cible.

La différence dans la conceptualisation des simulations informatiques de cette manière, par opposition au point de vue de la résolution de problèmes, est que les caractéristiques physiques de l'ordinateur ne sont plus la principale valeur épistémique des simulations informatiques. C'est plutôt leur capacité à décrire les modèles de comportement d'un système cible qui porte le fardeau. Imitant la section précédente, commençons par quelques premières définitions.

En 1960, Martin Shubik définissait une simulation de la manière suivante :

> Une simulation d'un système ou d'un organisme est le fonctionnement d'un modèle ou d'un simulateur qui est une représentation du système ou de l'organisme. (...) Le fonctionnement du modèle peut être étudié et, à partir de celui-ci, des propriétés concernant le comportement du système réel ou de son sous-système peuvent être déduites. (Shubik 1960, 909)

Shubik met en évidence deux caractéristiques principales qui sont au cœur de cette vision. C'est-à-dire qu'une simulation est une représentation ou une description du comportement d'un système cible, et que les propriétés d'un tel système cible peuvent être déduites. Le premier trait est central à ce point de vue, dans la mesure où il lui donne le nom. Mettre l'accent sur la capacité de représentation des simulations, par opposition aux modèles mathématiques, suggère qu'elles sont quelque peu différentes. Comme nous le verrons plus loin, cette différence réside dans le nombre de transformations par lesquelles passe un modèle mathématique – ou plutôt une série de modèles mathématiques – pour aboutir à une simulation informatique. La deuxième caractéristique, en revanche, met en évidence l'utilisation de simulations informatiques comme proxy pour comprendre quelque chose sur le système cible. C'est-à-dire que les chercheurs sont capables de déduire les propriétés du système cible sur la base des résultats de la simulation.

Ces deux caractéristiques, notons-le, sont absentes du point de vue de la résolution de problèmes. Le contraire, cependant, n'est pas vrai. Comme mentionné précédemment, comprendre les simulations informatiques de cette manière ne renie pas certaines prétentions du point de vue des techniques de résolution de problèmes. En particulier, la capacité de calculer des modèles complexes est une caractéristique des simulations informatiques généralement présente dans toutes les définitions. Par exemple, Shubik dit : « Le modèle se prête à des manipulations qui seraient impossibles, trop coûteuses ou irréalisables à effectuer sur l'entité qu'il dépeint » (909). Il est intéressant de noter qu'à mesure que l'on avance dans le temps, les préoccupations concernant la puissance de calcul tendent à disparaître.

Près de deux décennies plus tard, en 1979, G. Birtwistle a formulé la définition suivante pour les simulations informatiques :

> La simulation est une technique de représentation d'un système dynamique par un modèle afin d'obtenir des informations sur le système sous-jacent. Si le comportement du modèle correspond correctement

---

calcul. Maintenant, étant donné que les chercheurs veulent simuler un système cible réel, ils doivent le décrire le plus précisément possible. Si ce système cible est un système naturel, comme une sphère d'uranium, alors la précision dicte que la simulation est limitée à la masse de la sphère.

> les caractéristiques de comportement pertinentes du système sous-jacent, nous pouvons tirer des conclusions sur le système à partir d'expériences avec le modèle et ainsi nous épargner tout désastre.
> (Birtwistle 1979, 1)

De la même manière que Teichroew et Lubin ont présenté le point de vue de la résolution de problèmes, Birtwistle explique également clairement les principales caractéristiques de la description des modèles de point de vue comportemental. D'après la définition ci-dessus, il est clair que la représentation d'un système cible est au cœur des simulations informatiques; à condition d'avoir la bonne représentation, les chercheurs peuvent tirer des conclusions sur ce système cible. Notons aussi que, contrairement à Teichroew et Lubin qui considèrent les simulations informatiques comme une dernière ressource, pour Birtwistle c'est un élément crucial de la recherche scientifique qui aide à prévenir les catastrophes. Les attitudes opposées envers les simulations informatiques ne peuvent pas trouver deux meilleurs représentants.

Une autre définition digne d'être mentionnée vient de Robert E. Shannon, un ingénieur industriel qui a beaucoup travaillé sur la clarification de la nature des simulations informatiques (voir son travail de (Shannon 1975) et (Shannon 1978)).

> Nous définirons la simulation comme le processus de conception d'un modèle d'un système réel et de réalisation d'expériences avec ce modèle dans le but de comprendre le comportement du système et/ou d'évaluer diverses stratégies de fonctionnement du système. Il est donc essentiel que le modèle soit conçu de manière à ce que le comportement du modèle imite le comportement de réponse du système réel aux événements qui se produisent au fil du temps. (Shannon 1998, 7)

Encore une fois, nous pouvons voir comment Shannon souligne l'importance de représenter un système cible, ainsi que la capacité de déduire – et d'évaluer – nos connaissances à partir de simulations informatiques. Ce qui est peut-être l'aspect le plus remarquable de la définition de Shannon est l'accent marqué mis sur la méthodologie des simulations informatiques. Pour lui, il est essentiel que le modèle de la simulation imite le comportement du système cible.
Il ne suffit pas, comme on le trouve chez d'autres auteurs, que le modèle décrive correctement le comportement pertinent du système cible. Il faut prêter attention à la manière dont la simulation est conçue, car c'est là que l'on trouvera des motifs – et des problèmes – pour tirer des inférences sur le système cible.

Ces dernières idées se poursuivent, avec plus ou moins de succès, dans la littérature ultérieure liée à ce point de vue. Un bon exemple est le livre de 2004 de Paul Humphreys, où il présente un compte rendu détaillé de la méthodologie des simulations informatiques. Eric Winsberg, quelques années plus tard, a également fait un effort intéressant pour montrer comment les décisions de conception affectent les évaluations épistémologiques. Selon Winsberg, les décisions de conception présentes et passées fondent notre confiance dans les résultats des simulations informatiques. Examinons maintenant leurs positions plus en détail.[9] Plus tôt, j'ai mentionné

qu'en 1990, Humphreys a élaboré une définition de travail pour les simulations informatiques. Bien qu'il ne l'ait présentée que comme une définition de travail, il a reçu de vigoureuses objections qui l'ont pratiquement forcé à changer sa version originale.

---

[9] Il existe de nombreux autres auteurs contemporains qui méritent notre attention. Le travail le plus important est celui de Claus Beisbart, qui prend des simulations informatiques comme arguments (Beisbart 2012). C'est-à-dire une structure inférentielle englobant une prémisse et une conclusion. Un autre cas intéressant est Rawad El Skaf et Cyrille Imbert (El Skaf et Imbert 2013), qui conceptualisent les simulations informatiques comme des « scénarios de déploiement ». Malheureusement, l'espace ne me permet pas de discuter plus en détail de ces auteurs.
étendue.

point de vue sur les simulations informatiques. L'un des principaux critiques était Stephan Hartmann, qui a souligné que sa définition de travail manquait la nature dynamique de l'ordinateur simulations. Après les objections initiales de Hartmann, Humphreys forgea une nouvelle définition, basée cette fois sur la notion de modèle computationnel.

Selon sa nouvelle caractérisation, les simulations informatiques reposent sur un modèle informatique sous-jacent qui implique des représentations d'un système cible. D'abord coup d'œil, cette définition ressemble beaucoup aux définitions standard discutées jusqu'à présent. Cependant, le diable est dans les détails. Afin d'apprécier pleinement le tour d'Humphreys, il faut décortiquer sa définition de modèle computationnel, compris comme le sextuple :

<modèle de calcul, hypothèses de construction, ensemble de corrections, interprétation, justification initiale, représentation de sortie>10

Un modèle de calcul est, en fait, le résultat d'un traitement calculable modèle théorique. Un modèle théorique, à son tour, est le genre de modèle très général description mathématique que l'on peut trouver dans un ouvrage scientifique. Cela inclut les équations aux dérivées partielles, telles que les équations elliptiques (par exemple, l'équation de Laplace), paraboliques (par exemple, l'équation de diffusion), hyperbolique (par exemple, l'équation d'onde) et les équations différentielles ordinaires, entre autres. Un exemple éclairant de modèle théorique est la deuxième loi de Newton, car elle décrit une contrainte très générale sur la relation entre les forces, la masse et l'accélération. La principale caractéristique de la théorie modèles est que les chercheurs pourraient les spécifier de différentes manières. Pour Par exemple, la fonction de force dans la deuxième loi de Newton pourrait être soit une force gravitationnelle force, une force électrostatique, une force magnétique ou tout autre type de force.

Maintenant, un modèle de calcul ne peut pas simplement être choisi à partir de la théorie modèle. C'est le genre de fonctionnalité qui guide le point de vue de la résolution de problèmes, mais pas la description des modèles de point de vue comportement. De ce dernier point de vue, il y a est une méthodologie complète qui sert d'intermédiaire entre le modèle informatique et le modèle théorique à explorer. Concrètement, le processus de construction d'un modèle implique un certain nombre d'idéalisations, d'abstractions, de contraintes et approximations du système cible dont les chercheurs doivent tenir compte. De plus, à un moment donné, le modèle de calcul doit être validé par rapport aux données. Que se passe-t-il lorsqu'il ne correspond pas à ces données ? Eh bien, la réponse est que les chercheurs avoir une série de méthodes bien établies pour corriger le modèle de calcul afin de garantir des résultats précis. Selon Humphreys, la construction en tant qu'ensemble de sommations et de correction - composants deux et trois dans le sextuple - remplit précisément ces rôles. Sans eux, le modèle de calcul pourrait même ne pas être calculable.

Maintenant, afin d'avoir une représentation précise du système cible, les variables, les fonctions, etc. dans le modèle de calcul doivent recevoir une interprétation. Par exemple, dans la première dérivation d'une équation de diffusion, l'interprétation de la fonction représentant le gradient de température dans un conducteur cylindrique parfaitement isolé est essentielle pour décider si l'équation de diffusion

---

[dix] Pour plus de détails, voir (Humphreys 2004, 102-103)

tion représente correctement le flux de chaleur dans une barre métallique donnée (Humphreys 2004, 80). L'interprétation par le chercheur du modèle de calcul constitue une partie de la justification de l'adoption de certaines équations, valeurs et fonctions. Les modèles de calcul, dit Humphreys, ne sont « pas de simples conjectures mais des objets pour lesquels une justification distincte pour chaque idéalisation, approximation et principe physique est souvent disponible, et ces justifications sont transférées à l'utilisation du modèle ». (81).

Enfin, la représentation de sortie, c'est-à-dire la visualisation du modèle de calcul, se décline en différentes saveurs. Il peut s'agir d'un tableau de données, de fonctions, d'une matrice et, plus important en termes de compréhension, de représentations dynamiques telles que des vidéos ou des visualisations interactives. Comme nous le verrons en détail dans la section 5.2.1, les visualisations jouent un rôle fondamental dans notre gain épistémique à l'aide de simulations informatiques, et donc dans leur succès général en tant que nouvelles méthodes de recherche scientifique et technique.

Eric Winsberg est le deuxième philosophe de notre liste. Selon lui, il existe deux caractéristiques fondamentales qui distinguent de manière significative les simulations informatiques des autres formes de calcul. Tout d'abord, beaucoup d'efforts sont déployés pour mettre en place le modèle qui sert de base aux simulations informatiques, ainsi que pour décider quels résultats de simulation sont fiables et lesquels ne le sont pas. Deuxièmement, les simulations informatiques utilisent une variété de techniques et de méthodes qui facilitent l'inférence à partir des résultats (Winsberg 2010). Comme discuté précédemment dans cette section, ces deux caractéristiques sont typiques de la description des modèles de point de vue comportemental.

En outre, Winsberg souligne à juste titre que la construction de simulations informatiques est guidée, mais non déterminée, par la théorie. Cela signifie que, bien que les simulations informatiques s'appuient sur un contexte théorique, elles englobent généralement des éléments qui ne sont pas directement liés aux théories ni ne font partie de celles-ci. Un exemple de ceci sont les « fictionnalisations », c'est-à-dire les principes contraires aux faits qui sont inclus dans le modèle de simulation dans le but d'augmenter la fiabilité et la fiabilité de ses résultats. Comme nous l'avons vu précédemment, Humphreys a fait un point similaire avec les hypothèses de construction et l'ensemble de corrections. Winsberg illustre ensuite les fictionnalisations avec deux exemples, la « viscosité artificielle » et le « confinement de la vorticité ». Dans les simulations de dynamique des fluides, ces techniques sont utilisées avec succès bien qu'elles n'offrent pas de comptes rendus réalistes de la nature des fluides. Pourquoi sont-ils alors utilisés ? Il y a plusieurs raisons, y compris bien sûr qu'ils font largement partie de la pratique des techniques de construction de modèles sur la dynamique des fluides. D'autres raisons incluent le fait que ces fictionnalisations facilitent le calcul d'effets cruciaux qui seraient autrement perdus, et que sans ces fictionnalisations, les résultats des simulations sur la dynamique des fluides ne pourraient être ni précis ni justifiés.

Les discussions précédentes montrent qu'il n'est tout simplement pas possible d'intégrer le concept de simulations informatiques dans un corset conceptuel. Ainsi, notre question initiale : 'que sont les simulations informatiques ?' ne peut pas être répondu de manière unique. Il semble qu'en définitive, cela dépendra des engagements des praticiens. Alors que le point de vue de la résolution de problèmes s'intéresse davantage à la recherche de solutions à des modèles complexes, le point de vue de la description des modèles de comportement s'intéresse à la représentation précise d'un système cible. Les deux offrent de bonnes conceptualisations des simulations informatiques, et les deux

ont plusieurs problèmes à affronter. Discutons ensuite de trois types différents de simulations informatiques que l'on trouve dans la pratique scientifique et technique.

1.2 Types de simulations informatiques

Avant d'aborder les différentes classes de simulations informatiques, discutons brièvement d'une brève classification des systèmes cibles généralement associés aux simulations informatiques. Cette classification, en plus d'être non exhaustive - ou précisément à cause de cela - n'a aucune attente d'être unique. D'autres façons de caractériser les systèmes cibles - ainsi que les modèles qui représentent ces systèmes cibles - peuvent conduire à une taxonomie nouvelle et améliorée.

Après avoir mentionné tous les avertissements habituels, commençons par le plus familier de tous les systèmes cibles, c'est-à-dire les systèmes cibles empiriques. Ce sont des phénomènes empiriques - ou des phénomènes du monde réel - sous toutes les formes et saveurs. Les exemples incluent le rayonnement de fond des micro-ondes et le mouvement brownien en astronomie et en physique, la ségrégation sociale en sociologie, la concurrence entre les fournisseurs en économie et les scramjets en ingénierie, parmi de nombreux autres exemples.

Naturellement, les systèmes cibles empiriques sont le système cible le plus répandu dans la simulation informatique. C'est principalement parce que les chercheurs sont sérieusement engagés dans la compréhension du monde empirique, et les simulations informatiques fournissent une méthode nouvelle et efficace pour atteindre ces objectifs. Désormais, afin de représenter des systèmes cibles empiriques, les simulations informatiques implémentent des modèles qui sous-tendent théoriquement les phénomènes du monde réel à l'aide de lois, de principes et de théories acceptés par la communauté scientifique. Le modèle newtonien du mouvement planétaire, par exemple, décrit le comportement de deux corps quelconques interagissant l'un avec l'autre par une poignée de lois et de principes. Malheureusement, tous les systèmes cibles empiriques ne peuvent pas être représentés de manière aussi simple et précise.

Plus communément, les simulations informatiques représentent des phénomènes du monde réel en incluant une pléthore d'éléments provenant de sources différentes – et parfois incompatibles. Prenez par exemple les scramjets, les statoréacteurs à combustion dans lesquels la combustion a lieu dans un flux d'air supersonique. L'utilisation des équations de Navier-Stoke est typiquement à la base des simulations de dynamique des fluides. Cependant, l'admission d'un scramjet comprime l'air entrant via une série d'ondes de choc générées par la forme spécifique de l'admission ainsi que la vitesse de vol élevée, contrairement à d'autres véhicules à respiration aérienne qui compriment l'air entrant par des compresseurs - ou d'autres pièces mobiles. Les simulations de scramjets ne peuvent donc pas être entièrement caractérisées par les équations de Navier-Stoke.
Au lieu de cela, les couches limites laminaires et turbulentes, ainsi que l'interaction avec les ondes de choc, produisent un modèle d'écoulement complexe instable en trois dimensions. Une simulation fiable de ce qui se passe dans l'admission est alors réalisée au moyen de simulations numériques directes haute fidélité et de simulations de grandes turbulences. C'est le

conception du modèle, programmé et construit par les ingénieurs, et pas seulement les équations de Navier Stoke, qui permet une simulation fiable (Barnstorff 2010).[11]

Un autre système cible important est ce que l'on appelle le système cible hypothétique. Ce sont des systèmes cibles où aucun phénomène empirique n'est décrit. Elles sont plutôt théoriques ou imaginaires. Un système cible théorique décrit des systèmes ou des processus au sein de l'univers fournis par une théorie, qu'elle soit mathématique (par exemple, un tore), physique (par exemple, une résistance de l'air égale à zéro) ou biologique (par exemple, une population infinie). tions). Prenons comme exemple le célèbre problème des Sept Ponts de Königsberg ¨ ou le[12] problème du voyageur de commerce[13]. Ainsi compris, le système cible n'est pas empirique, mais au contraire il a les propriétés d'un système mathématique ou logique. Une simulation informatique mettant en œuvre ces modèles est essentiellement théorique et est généralement conçue pour explorer les propriétés sous-jacentes du modèle.

Les systèmes cibles imaginaires, en revanche, représentent des scénarios imaginaires inexistants. Par exemple, une épidémie de grippe en Europe compte comme un tel système cible. En effet, un tel scénario est susceptible de ne jamais exister, même si cela ne signifie pas qu'il ne se produira jamais. Une simulation d'un tel scénario fournit aux chercheurs la compréhension nécessaire de la dynamique d'une flambée épidémique pour la planification des mesures de prévention et des protocoles de confinement, ainsi que pour la formation du personnel. Les systèmes imaginaires peuvent être, à leur tour, divisibles en deux autres types, à savoir contingents et impossibles (Weisberg 2013). Le premier représente un scénario qui, en tant que fait contingent, n'existe pas. Ce dernier représente un scénario nomologiquement impossible. La simulation d'une épidémie est un exemple de la première, tandis que l'exécution d'une simulation qui viole les lois connues de la nature est un exemple de la seconde.

La première chose à noter à propos de cette classification est que les simulations informatiques peuvent représenter un système cible mais rendre les résultats d'un autre système cible. Il s'agit d'un mécanisme de «saut» courant qui pourrait être inoffensif ou qui pourrait jeter une ombre de doute sur les résultats. Un exemple simple montrera comment cela est possible. Envisagez de simuler le mouvement planétaire en mettant en œuvre un modèle newtonien ; maintenant instantanément tie $G = 2m^{3}\,kg^{-1}\,s^{-2}$ comme condition initiale. L'exemple montre une simulation qui implémente initialement un système cible empirique, mais rend les résultats d'un système cible imaginaire nomologiquement impossible. Au meilleur de notre connaissance actuelle de l'univers, il n'y a pas une telle constante gravitationnelle. Par conséquent, les résultats d'une simulation qui en principe auraient dû être sanctionnés empiriquement (par exemple, en validant par rapport à des données empiriques) ne peuvent être confirmés que théoriquement.[14]

[11] Je devrais également mentionner qu'il existe plusieurs autres artifices également impliqués dans la conception et la programmation de simulations informatiques. À cet égard, la section 4.2 présente et discute certaines d'entre elles, telles que les procédures d'étalonnage et les méthodes de vérification et de validation.

[12] Le problème peut être mieux décrit comme trouver un moyen de traverser chacun des sept ponts de la ville de Königsberg une seule fois. Le problème, résolu par Euler en 1735, a jeté les bases de la théorie des graphes.

[13] Le problème du voyageur de commerce décrit un vendeur qui doit voyager entre un nombre N de villes et maintenir les frais de déplacement aussi bas que possible. Le problème consiste à trouver la meilleure optimisation du parcours du vendeur.

[14] L'exécution d'une deuxième simulation informatique qui pourrait confirmer ces résultats devient une pratique courante (Ajelli et al. 2010).

D'une certaine manière, ces questions font partie du charme général et de la malléabilité de l'utilisation des simulations informatiques, mais elles doivent être prises au sérieux par les philosophes et donc les ciologues de la science. Cela dit, et en examinant de près la pratique de la simulation informatique, on peut voir comment les chercheurs disposent de quelques « astuces » qui aident à faire face à des situations comme le « saut ». Par exemple, une solution à la simulation du satellite sous contrainte de marée serait de définir G comme une constante globale de valeur 6,67384x10−11m 3 kg−1 s La valeur de la masse variable du $^{-2}$. Malheureusement, il ne s'agit là que d'une solution palliative puisque la satellite pourrait être définie sur n'importe quelle valeur irréaliste.
Encore une fois, les chercheurs pourraient établir des limites inférieures et supérieures sur la taille du satellite et la masse de la planète, mais cette solution ne fait que poser la question s'il n'y a pas un autre moyen de "tromper la simulation".

Soit par modélisation, soit par instanciation, les simulations informatiques peuvent créer plusieurs scénarios hors de l'esprit des chercheurs. Comment les chercheurs transforment-ils cette situation apparemment désastreuse en quelque chose d'avantageux ? La réponse est, je crois, dans la façon dont les scientifiques et les ingénieurs considèrent les simulations informatiques comme des processus fiables. C'est-à-dire en fournissant des raisons de croire que les simulations informatiques sont un processus fiable qui donne, la plupart du temps, des résultats corrects. J'aborderai ces questions plus en détail dans le chapitre 4.

Parallèlement à une classification des systèmes cibles, je propose maintenant une classification des simulations informatiques. Dans le même esprit, ce classement ne se veut ni exhaustif, ni concluant, ni unique. Ici, je divise les simulations en trois classes, sur la base du traitement standard que les simulations informatiques ont reçu de la littérature spécialisée (Winsberg 2015)). Il s'agit d'automates cellulaires, de simulations basées sur des agents et de simulations basées sur des équations.

1.2.1 Automates cellulaires

Les automates cellulaires sont le premier de nos exemples de simulations informatiques. Ils ont été conçus dans les années 1940 par Stanislaw Ulam et John von Neumann alors qu'Ulam étudiait la croissance des cristaux en utilisant un simple réseau de treillis comme modèle, et von Neumann travaillait sur le problème des systèmes auto-réplicants. L'histoire raconte qu'Ulam a suggéré à von Neumann d'utiliser le même type de réseau en treillis que le sien, créant ainsi un algorithme auto-réplicateur bidimensionnel.

Les automates cellulaires sont des formes simples de simulations informatiques. Une telle simplicité découle à la fois de leur programmation et de la conceptualisation sous-jacente. Un automate cellulaire standard est un système mathématique abstrait où l'espace et le temps sont considérés comme discrets ; il consiste en une grille régulière de cellules, dont chacune peut être dans n'importe quel état à un instant donné. Typiquement, toutes les cellules sont régies par la même règle, qui décrit comment l'état d'une cellule à un instant donné est déterminé par les états d'elle-même et de ses voisines à l'instant précédent. Stephen Wolfram définit les automates cellulaires de la manière suivante :

> [...] mathématiques pour des systèmes naturels complexes contenant un grand nombre de composants simples identiques avec des interactions locales. Ils consistent en un réseau de sites, chacun avec

> un ensemble fini de valeurs possibles. La valeur des sites évolue de manière synchrone en temps discret étapes selon des règles identiques. La valeur d'un site particulier est déterminée par le précédent valeurs d'un voisinage de sites qui l'entourent. (Wolfram 1984b, 1)

Bien qu'il s'agisse d'une caractérisation plutôt générale de cette classe de simulation informatique, la définition ci-dessus donne déjà les premières idées quant à leur domaine d'application. Les automates cellulaires ont été utilisés avec succès pour modéliser de nombreux domaines dans dynamique sociale (par exemple, dynamique comportementale pour les activités coopératives), biologie (par exemple, motifs de certains coquillages) et les types chimiques (par exemple, la réaction de Belousov-Zhabotinsky).

L'un des exemples les plus simples et les plus canoniques d'automates cellulaires est le jeu de la vie de Conway. La simulation est remarquable car elle fournit un cas de émergence de patterns et dynamiques d'auto-organisation de certains systèmes. Dans cette simulation, une cellule ne peut survivre que s'il y a deux ou trois autres cellules vivantes dans sa cellule. voisinage immédiat. Sans ces compagnons, la règle indique que la cellule meurt soit de surpeuplement, s'il a trop de voisins vivants, soit de solitude, s'il en a trop peu. Une cellule morte peut reprendre vie à condition que il y a exactement trois voisins vivants. En vérité, il y a peu d'interaction - comme un attendrait d'un jeu - en plus de créer une configuration initiale et d'observer comment ça évolue. Néanmoins, d'un point de vue théorique, le Jeu de la Vie peut calcule n'importe quel algorithme calculable, ce qui en fait un exemple remarquable de machine de Turing universelle. En 1970, Game of Life de Conway ouvre un nouveau champ de la recherche mathématique : le domaine des automates cellulaires (Gardner 1970).

Les automates cellulaires élémentaires fournissent des cas fascinants à la science contemporaine. L'idée de ces automates est qu'ils sont basés sur un réseau unidimensionnel infini de cellules avec seulement deux états. A des intervalles de temps discrets, chaque cellule change d'état en fonction de son état actuel et de l'état de ses deux voisins. Règle 30 en est un exemple qui produit des modèles complexes, apparemment aléatoires, à partir de règles simples et bien définies (voir figure 1.4). Par exemple, un modèle ressemblant à la règle 30 apparaît sur la coquille des espèces d'escargots coniques Conus textile (voir figure 1.5). D'autres exemples sont basés sur ses propriétés mathématiques, comme l'utilisation de la règle 30 comme un générateur de nombres pour les langages de programmation, et en tant que chiffrement de flux possible pour utilisation en cryptographie. L'ensemble de règles qui régit l'état suivant de la règle 30 est affiché dans la figure 1.4.

modèle actuel	111	110	101	100	011	010	001	000
nouvel état pour la cellule centrale	0	0		0	1	1	1	1	0

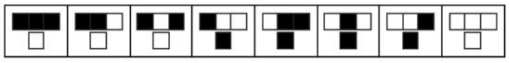


Fig. 1.4 Règle 30. http://mathworld.wolfram.com/CellularAutomaton.html

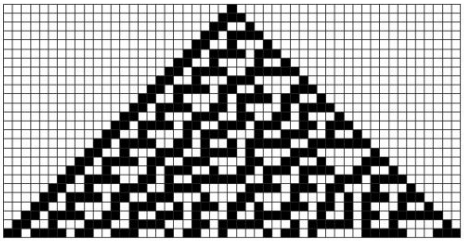


Fig. 1.5 Modèle créé par la règle 30. http://mathworld.wolfram.com/CellularAutomaton.html

Les automates cellulaires consacrent un ensemble de vertus méthodologiques et épistémologiques uniques. Pour n'en citer que quelques-uns, ils s'accommodent mieux de l'erreur car ils rendent des résultats exacts du modèle qu'ils implémentent. Comme les approximations avec le système cible sont quasi inexistantes, tout désaccord entre le modèle et les données empiriques peut être attribué directement au modèle qui a réalisé l'ensemble de règles. Une autre vertu épistémologique est soulignée par Evelyn Fox-Keller, qui explique que les automates cellulaires manquent de fondement théorique au sens familier du terme. Autrement dit, "ce qui doit être simulé n'est ni un ensemble bien établi d'équations différentielles [...] ni les constituants physiques fondamentaux (ou particules) du système [...] mais plutôt le phénomène lui-même" (Fox Keller 2003, 208). Les approximations, les idéalisations, les abstractions et autres sont des concepts qui préoccupent très peu le praticien des automates cellulaires.

Maintenant, tout n'est pas parfait pour les automates cellulaires. Ils ont été critiqués pour plusieurs raisons. L'une de ces critiques porte sur les hypothèses métaphysiques derrière cette classe de simulation. Il n'est pas clair, par exemple, que le monde naturel soit en fait un lieu discret, comme le supposent les automates cellulaires. De nombreux travaux scientifiques et techniques d'aujourd'hui sont basés sur la description d'un monde continu.
Pour des raisons moins spéculatives, c'est un fait que les automates cellulaires sont peu présents dans les domaines scientifiques et techniques. La raison en est, je crois, en partie culturelle. Les sciences physiques sont toujours le point de vue accepté pour décrire le monde naturel, et elles sont écrites dans le langage des équations différentielles partielles et des équations différentielles ordinaires (PDE et ODE respectivement).

Naturellement, les partisans des automates cellulaires concentrent leurs efforts pour montrer leur pertinence. En toute honnêteté, de nombreux automates cellulaires sont plus adaptables et structurellement similaires aux phénomènes empiriques que les PDE et les ODE (Wolfram 1984a, vii). Il a été souligné par Annick Lesne, physicienne théoricienne de renom, que des comportements discrets et continus coexistent dans de nombreux phénomènes naturels, selon l'échelle d'observation. Pour elle, c'est un indicateur non seulement de la base métaphysique de nombreux phénomènes naturels, mais aussi de l'adéquation des automates cellulaires à la recherche scientifique et technique (Lesne 2007). Dans le même ordre d'idées, Gérard Vichniac pense que les automates cellulaires ne cherchent pas seulement à s'accorder numériquement avec un système physique, mais aussi à faire correspondre la structure propre du système simulé, sa topologie, ses symétries et ses propriétés « profondes » (Vichniac 1984 : 113). . Tommaso Toffoli a une position similaire à ces auteurs, au point qu'il a intitulé un article :

"Les automates cellulaires comme alternative (plutôt qu'une approximation) différentielle
équations en physique de modélisation » (Toffoli 1984), mettant en évidence les automates cellulaires comme
remplacement naturel des équations différentielles en physique.

Malgré ces efforts et de nombreux autres auteurs pour montrer que le monde pourrait être
plus adéquatement décrites par les automates cellulaires, la plupart des disciplines scientifiques et
d'ingénierie n'ont pas encore complètement changé. La plupart des travaux effectués en
ces disciplines sont principalement basées sur des simulations basées sur des agents et sur des
équations. Comme mentionné précédemment, dans les sciences naturelles et l'ingénierie, la plupart
des théories physiques et chimiques utilisées en astrophysique, géologie, changement climatique et
comme implémenter PDE et ODE, deux systèmes d'équations qui sont à la base de
simulations basées sur des équations. Les systèmes sociaux et économiques, en revanche, sont
mieux décrites et comprises au moyen de simulations multi-agents.

1.2.2 Simulations basées sur des agents

Bien qu'il n'y ait pas d'accord général sur la définition de la nature d'un "agent",
le terme fait généralement référence à des programmes autonomes qui contrôlent leurs propres actions
en fonction des perceptions de l'environnement d'exploitation. En d'autres termes, basée sur l'agent
les simulations interagissent «intelligemment» avec leurs pairs ainsi qu'avec leur environnement.

La caractéristique pertinente de ces simulations est qu'elles montrent comment le total
Le comportement d'un système émerge de l'interaction collective de ses parties. Déconstruire ces
simulations en leurs éléments constitutifs supprimerait l'ajout
valeur qui a été fournie en premier lieu par le calcul des agents. Il
est une caractéristique fondamentale de ces simulations, alors, que l'interaction des
divers agents et l'environnement induit un comportement unique de l'ensemble
système.

De bons exemples de simulations à base d'agents proviennent des sciences sociales et
comportementales, où elles sont fortement présentes. L'exemple le plus connu de simulation basée
sur des agents est peut-être le modèle de ségrégation sociale de Schelling15 .
description très simple du modèle de Schelling consiste en deux groupes d'agents vivant
dans un 2-D16 , nxm matrix 'damier' où les agents sont placés au hasard. Chaque

---

15 Bien que de nos jours le modèle de Schelling soit exécuté par des ordinateurs, Schelling lui-même a averti
contre son utilisation pour la compréhension du modèle. Au lieu de cela, il a utilisé des pièces de monnaie ou d'autres éléments pour montrer comment
ségrégation s'est produite. À cet égard, Schelling déclare : « Je ne saurais trop vous exhorter à obtenir le
nickels et pennies et faites-le vous-même. Je peux vous montrer un résultat ou deux. Un ordinateur peut faire
cent fois pour vous, en testant les variations des demandes du quartier, les ratios globaux, la taille des
quartiers, etc. Mais il n'y a rien de tel que de le parcourir par vous-même et de voir
le processus se déroule tout seul. Cela prend environ cinq minutes - pas plus de temps qu'il ne me faut pour décrire
le résultat que vous obtiendriez » (Schelling 1971, 85). L'avertissement de Schelling contre l'utilisation des ordinateurs
est une anecdote amusante qui illustre comment les scientifiques pouvaient parfois échouer à prédire le rôle de
ordinateurs dans leurs domaines respectifs.

16 Schelling a également introduit une version 1-D, avec une population de 70 agents, avec les quatre plus proches
voisins de part et d'autre, la préférence consiste à ne pas être minoritaire, et la règle de migration est
que celle qui est mécontente se déplace vers le point le plus proche qui réponde à ses exigences (149).

l'agent individuel a un voisinage 3x3, qui est évalué par une fonction d'utilité
qui indique les critères de migration. C'est-à-dire l'ensemble de règles qui indique comment
déplacer – si possible – en cas de mécontentement d'un agent (voir Fig. 1.6)

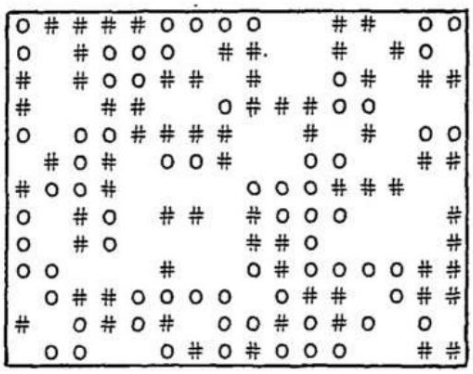


Fig. 1.6 Répartition aléatoire initiale des agents dans un damier de 13 lignes x 16 colonnes, avec un total de 208 cases. (Schelling 1971, 155)

Le modèle de ségrégation de Schelling est un exemple canonique de simulations à base d'agents17. Mais des simulations à base d'agents plus complexes peuvent être trouvées dans la littérature. Il est maintenant standard que les chercheurs modélisent différents attributs, préférences et comportement dans les agents. Nigel Gilbert et Klaus Troitzsch énumèrent l'ensemble des attributs qui sont typiquement modélisés par des agents (Gilbert et Troitzsch 2005) :

1. Connaissance et croyance : Puisque les agents fondent leurs actions sur leur interaction avec leur environnement ainsi que d'autres agents, il est crucial de pouvoir modéliser leur système de croyances. La distinction traditionnelle entre la connaissance, en tant que véritable croyance justifiée, et la simple croyance peut alors être modélisée. Ceux dont les informations peut être défectueux ou purement et simplement faux doivent être modélisés pour agir dans leur environnement d'une manière manière assez différente de celle des agents dont les informations sont correctes, en conséquence de savoir.
2. Inférence : la connaissance et la croyance sont possibles parce que les agents sont capables de déduire informations de leur propre ensemble de croyances. De telles inférences sont modélisées de différentes manières, parfois sur des hypothèses très intuitives. Par exemple, l'agent A pourrait en déduire qu'une source de «nourriture» est proche de l'agent B en sachant - ou simplement en croyant – que l'agent B a « mangé de la nourriture ».
3. Objectifs : Puisque les agents sont, pour la plupart, programmés comme des entités autonomes, ils sont généralement motivés par une sorte d'objectifs. La survie est un bon exemple de cela, car cela pourrait nécessiter la satisfaction d'objectifs subsidiaires, tels que l'acquisition

[17] En vérité, selon la façon dont le modèle de Schelling est conçu et programmé, il pourrait également se qualifier en tant qu'automate cellulaire. Merci Andres Il ci˘ c de me l'avoir signalé.

sources d'énergie, de nourriture et d'eau, tout en évitant tout danger. Modéliser ces objectifs subsidiaires n'est pas facile, car les chercheurs doivent décider comment évaluer l'importance et la pertinence de plusieurs objectifs. Différentes décisions de conception conduisent à différents agents guidés par un but, et donc à une approche globale différente. comportement du système. Notons que les buts d'un agent sont différents attribut de savoir et de déduire. Alors que les premiers visent à guider le comportement général des agents dans leur environnement, celui-ci dépend de ou de vraies informations pour gouverner leur comportement.

4. Planification : Afin de satisfaire ses objectifs, un agent doit avoir un moyen de déterminer quelles décisions sont les meilleures. Généralement, un ensemble de règles de condition-action sont programmés comme constituant de l'agent. Par exemple, une fonction d'utilité est programmé pour la satisfaction d'une source d'énergie et pour ce qui compte comme 'danger.' La planification consiste à déduire quelles actions conduisent à un objectif souhaité, quelles l'état est requis avant que cette action ait lieu, et quelles actions sont nécessaires pour arriver à l'état désiré. A cet égard, la planification est très sophistiquée car l'agent doit peser plusieurs options, y compris avoir une règle de « paiement » sur leurs décisions, déterminer où il doit être à un moment donné dans le futur, etc. Il a été objecté que la planification des agents n'est pas réaliste planification parce que la plupart des actions humaines sont motivées par des décisions de routine, une tendance à discerner, voire des jugements instinctifs qui ne peuvent être modélisés par un plan calculé.
5. Langage : la transmission d'informations entre les agents est essentielle pour toute simulation basée sur les agents. Un exemple intéressant est le modèle multi-agents des colonies de fourmis d'Alexis Drogoul et Jacques Ferber (Drogoul, Ferber et Cambier 1992). Selon les auteurs, les agents peuvent communiquer en propageant un « stimuli » au environnement. Ces stimuli peuvent être reçus et transmis de différentes manières. Lorsque les fourmis reçoivent ces stimuli, elles activent un comportement amical. Cependant, lorsque un prédateur reçoit les stimuli, il déclenche un comportement agressif. Dans cette simulation à base d'agents particulière, un tel mécanisme de communication est très simple et ne vise en aucun cas à transmettre une quelconque signification. D'autres exemples de langage de modélisation dans les simulations multi-agents sont la négociation de contrats (Smith et Davis 1981), la communication des décisions, et même un agent menaçant un autre avec la « mort » (Gilbert et Troitzsch 2005).
6. Modèles sociaux : certaines simulations basées sur des agents, comme le modèle de ségrégation de Schelling discutés précédemment, visent à modéliser l'interrelation entre les agents dans un environnement plus vaste. De ce fait, les agents sont capables de créer leur propre topologie compte tenu de l'ensemble de règles, de leur interaction avec d'autres agents et de la configuration, entre autres paramètres. Par exemple, dans le modèle de ségrégation de Schelling, les agents créent des topologies de ségrégation différentes selon le côté de la grille, la fonction d'utilité, et la position initiale des agents.

1.2.3 Simulations basées sur des équations

La classe de simulations informatiques la plus couramment utilisée en sciences et en ingénierie est peut-être celle que l'on appelle les simulations basées sur des équations. A leur base, ce sont les implémentation d'un modèle mathématique sur l'ordinateur physique qui décrit un système cible. Parce que la majeure partie de notre compréhension du monde passe par l'utilisation de descriptions mathématiques, ces simulations sont de loin les plus populaires dans les domaines scientifiques et techniques. Naturellement, les exemples ne manquent pas. La dynamique des fluides, la mécanique des solides, la dynamique des structures, la physique des ondes de choc et la chimie moléculaire ne sont qu'une poignée de cas. William Oberkampf, Timothy G. Trucano et Charles Hirsch a développé une longue liste qu'ils jugent appeler « ingénierie informatique » et « physique informatique » (Oberkampf, Trucano et Hirsch 2003). Notons que leur étiquetage ne met l'accent sur les simulations qu'en physique et domaines de l'ingénierie. Bien qu'il soit exact de dire que l'écrasante majorité des des simulations basées sur des équations peuvent être trouvées dans ces domaines, cela ne rend pas justice à la myriade de simulations que l'on trouve également dans d'autres domaines scientifiques. Pour étendre les exemples, le modèle Solow-Swan de croissance économique est un cas en économie, et le Le modèle Lotka-Volterra du prédateur-proie fonctionne aussi bien pour la sociologie que pour la biologie.

Comme mentionné, cette classe de simulations implémente des modèles mathématiques sur le ordinateur. Mais, est-ce tout simplement ainsi ? De la section 1.1, nous avons appris que cette façon de penser correspond au point de vue de la résolution de problèmes de la simulation informatique. Selon ses partisans, il y a peu de médiateur entre le modèle mathématique et sa mise en œuvre sur ordinateur sous forme de simulation informatique. L'opposé vue, la description du modèle de point de vue du comportement, suppose qu'il y a en fait un méthodologie qui facilite la mise en œuvre d'une multiplicité de modèles mathématiques dans l'ordinateur. Pour avoir une meilleure compréhension d'une équation typique basée sur simulation – et pour déterminer quel point de vue est le plus proche des pratiques réelles –, Analysons brièvement un exemple de simulation récente sur l'âge du désert du Sahara.

Zhongshi Zhang et al. croyait que le désert du Sahara a émergé au cours de l'étape tortonienne - il y a environ 7 à 11 millions d'années - de l'époque du Miocène supérieur après une période d'aridité dans la région nord-africaine (Zhang et al. 2014) . Prouver leur hypothèse, l'équipe de Zhang a décidé de simuler le changement climatique dans ces régions à des échelles de temps géologiques et sur les 30 derniers millions d'années. L'âge du Sahara, selon la simulation est, en effet, d'environ 7 à 11 millions d'années. Avec ça résultat en main, Zhang et al. ont pu s'opposer à la plupart des estimations classiques de l'âge du Sahara, qui le situe à environ 2 à 3 millions d'années au début de les glaciations du Quaternaire. Comment ont-ils réellement simulé une expérience empirique aussi complexe ? système?

Premièrement, les auteurs ne mettent pas en œuvre un grand modèle de changement climatique sur le ordinateur et calculez-le jusqu'à obtenir les résultats. C'est la résolution de problèmes point de vue qui comprennent les simulations informatiques uniquement comme des moyens de calcul. Les simulations informatiques ressemblent beaucoup à un établi de laboratoire, où les scientifiques combinent subtilement des éléments de théorie, des fragments de données et de nombreuses connaissances et expertises. instinct. Le processus est en fait complexe, désordonné et, dans de nombreux cas, non standardisé. L'équipe de Zhang a utilisé une famille de modèles, chacun effectuant des tâches différentes

et représentant un aspect différent du système cible. Ils ont utilisé des versions basse et haute résolution du modèle norvégien du système terrestre (NorESM-L) pour tenir compte de la série d'époques géologiques, et du modèle communautaire de l'atmosphère version 4 (CAM4) comme composant atmosphérique de NorESM-L. En fait, le modèle NorESM-L est en soi une hiérarchie de petits modèles - ou de simples composants d'un modèle plus large - représentant la terre, la banquise, l'océan, etc.

Il n'y a, en vérité, aucun grand modèle qui pourrait nous renseigner sur l'âge du Sahara. Au lieu de cela, un patchwork de modèles – certains connus et bien établis, d'autres spéculatifs – des lois, des principes, des données et des éléments de théorie est ce qui conforme la simulation de l'équipe de Zhang. Cela ne devrait pas surprendre, car on suppose généralement qu'il n'y a pas de théorie générale qui sous-tend ou guide les simulations informatiques. De plus, les simulations incluent généralement des informations non linguistiques, telles que les décisions de conception, les biais de modèle possibles, les incertitudes identifiées et les clauses de non-responsabilité « non inclus dans ce modèle ». Bentsen et al., lors de la description de CAM4, fournissent un bon exemple d'une telle clause de non-responsabilité : "[a] les effets indirects des aérosols sur les nuages à phase mixte et de glace (par exemple (Hoose et al. 2010) ne sont pas inclus dans la version actuelle de CAM4 -Oslo » (Bentsen et al. 2013, 689)).

Malgré leur manque de fondement théorique complet, ces simulations sont toujours très fiables car elles représentent un système cible spécifique et sont généralement validées par des méthodes de vérification et de validation standard (voir section 4.2.2). À cet égard, Zhang et al. nous rappellent constamment que le modèle fonctionne bien pour simuler le climat préindustriel, que CAM4 simule raisonnablement bien les modèles de précipitations africaines modernes et d'autres positions de confirmation similaires des modèles. De tels rappels, bien sûr, ne peuvent arrêter les objections contre les résultats de la simulation. En particulier, les critiques des travaux de Zhang soulignent le manque de preuves pour valider leurs résultats. Stefan Kropelin est l'un des principaux détracteurs de l'utilisation de simulations informatiques pour ce type de système cible. Il admet que, bien que le modèle soit intéressant, il s'agit principalement de "spéculations numériques basées sur des preuves géologiques quasi inexistantes (...) Rien de ce que l'on peut trouver au Sahara n'a plus de 500 000 ans, et en termes de climat saharien même notre connaissance des 10 000 derniers est pleine de lacunes » (Kroepelin 2006).
La réponse de Zhang et al. est que la preuve de l'apparition précoce de l'aridité saharienne est très controversée. Mathieu Schuster est également en désaccord avec l'interprétation des données de Kropelin. Selon lui, « s'il est vrai que l'on sait trop peu de choses sur la géologie ancienne de la région [...] l'étude tchadienne de 2006 [...] ainsi que celles qui ont fait état d'augmentations de la poussière et du pollen provenant des sédiments, contenait « des éléments de preuve solides pour étayer nos nouvelles découvertes » » (Schuster 2006). En fait, la simulation de Zhang et de son équipe vient étayer certaines affirmations déjà dans la littérature. Anil Gupta et son équipe affirment une augmentation de l'upwelling dans l'océan Indien il y a environ 7 à 8 millions d'années (Gupta et al. 2004) ; et Gilles Ramstein et son équipe ont utilisé des expériences de modélisation pour montrer que les températures estivales eurasiennes augmentent en réponse au rétrécissement de la Téthys, ce qui améliorerait également la circulation de la mousson (Ramstein et al. 1997).

Affirmer que les simulations informatiques ne sont pas fiables, ou que leurs résultats ne correspondent pas à la façon dont le monde est, nécessite plus qu'une simple affirmation qu'il n'y a pas

«preuve» qui soutient le manque de fiabilité de la simulation.[18] D'autres indicateurs de la fiabilité joue également un rôle central. Par exemple, la capacité de la simulation à expliquer et prédire des phénomènes directs ou connexes. Selon Zhang et al., leur simulation montre que la mousson d'été africaine a été considérablement affaiblie par la mer de Téthys se rétrécissant au cours de l'étage tortonien, permettant l'altération de la climat moyen de la région. Un tel changement climatique, spéculent les chercheurs, « a probablement causé les changements dans la flore et la faune asiatiques et africaines observés au cours de la même période, avec des liens possibles avec l'émergence des premiers hominis en Afrique du Nord » (Zhang et al. 2014, 401). Chose intéressante, les chercheurs n'ont pu parvenir à une telle conclusion qu'au moyen de simulations informatiques.

Permettez-moi de terminer cette section par une brève description des méthodes de calcul générales pour résoudre les simulations basées sur des équations. Selon le problème et la disponibilité des ressources, une ou plusieurs des méthodes suivantes s'appliquent : méthodes analytiques, analyse numérique et techniques stochastiques.

- Solutions exactes : c'est la méthode la plus simple de toutes. Il consiste à réaliser le opérations spécifiées dans la simulation de la même manière qu'un mathématicien ferait avec un stylo et du papier. Autrement dit, si la simulation consiste à additionner 2+2, alors le résultat doit être 4 - par opposition à une solution approximative. Les ordinateurs ont la même capacité à trouver les solutions exactes à certaines opérations que tout autre mécanisme de calcul, y compris notre cerveau. L'efficacité de ce méthode dépend cependant de la taille du "mot" dans un ordinateur assez grand pour effectuer l'opération mathématique.[19] Si l'opération dépasse sa taille, alors des mécanismes d'arrondi et de troncature interviennent pour opération possible, mais avec une perte de précision.
- Méthodes numériques mises en œuvre par ordinateur : cette méthode fait référence aux méthodes mises en œuvre par ordinateur pour calculer le modèle de simulation par approximation. Bien que les études mathématiques sur l'analyse numérique soient antérieures à l'utilisation des ordinateurs, elles gagnent en importance avec l'introduction des ordinateurs à des fins scientifiques et techniques. Ces méthodes sont utilisées pour résoudre les PDE et équations ODE, et comprennent l'interpolation linéaire, la méthode Runge-Kutta, la Méthode d'Adams-Moulton, polynôme d'interpolation de Lagrange, élimination gaussienne et méthode d'Euler, entre autres. Notons que chaque méthode est utilisé pour résoudre un type spécifique de PDE et ODE, selon le nombre les dérivées impliquent la fonction inconnue de n variables.
- Techniques stochastiques : pour les dimensions d'ordre supérieur, les solutions exactes et les méthodes numériques mises en œuvre par ordinateur deviennent prohibitives dans termes de temps de calcul et de ressources. Les techniques stochastiques reposent sur des méthodes qui utilisent des nombres pseudo-aléatoires ; c'est-à-dire des nombres générés par un numeri

---

[18] .Cela est particulièrement vrai pour le sens traditionnel de « preuve » (c'est-à-dire fondé sur des données empiriques) selon lequel Kropelin fait référence. D'autres formes de preuves incluent également les résultats de simulations bien établies, la vérification et la validation, la convergence des solutions, etc.

[19] Un « mot » représente l'unité minimale de données utilisée par une architecture informatique particulière. C'est un groupe de bits de taille fixe qui sont traités comme une unité par le processeur.

cal.20 La méthode stochastique la plus connue est la méthode de Monte Carlo,
ce qui est particulièrement utile pour simuler des systèmes avec de nombreux degrés couplés
de liberté tels que les fluides, les matériaux désordonnés, les solides fortement couplés, et
structures cellulaires, pour n'en citer que quelques-unes.

1.3 Remarques finales

Ce chapitre avait pour seul objectif de répondre à la question « Qu'est-ce qu'un ordinateur ?
simulation?' Il s'agit bien sûr d'une question importante, car elle pose les bases d'une
une grande partie de ce qui est discuté sur les simulations informatiques plus loin dans ce livre. Pour ça
raison, la première partie du chapitre traite de quelques remarques historiques sur la
de nombreuses tentatives de définition des simulations informatiques, qu'elles soient proposées par des ingénieurs,
des scientifiques ou des philosophes. Dans ce contexte, j'ai distingué deux types de définitions. Ceux
qui mettent l'accent sur la puissance de calcul des simulations informatiques - appelées le point de vue de la
technique de résolution de problèmes - et celles qui nécessitent des simulations informatiques,
comme caractéristique principale, la capacité à représenter un système cible donné – appelée description de
schémas de comportement du point de vue. Bien qu'il existe quelques définitions
où les deux points de vue sont combinés, et sans doute celui qui ne correspond pas à notre
distinction, en général, les chercheurs de tous les domaines s'accordent à conceptualiser l'informatique
simulations selon l'un ou l'autre point de vue.

La deuxième partie de ce chapitre a traité de trois types différents de simulations informatiques, comme on en
trouve couramment dans la littérature. Ce sont des automates cellulaires, des simulations basées sur des agents
et des simulations basées sur des équations. Comme averti, ce n'est ni un
taxonomie exhaustive ni offre une classification unique. Cela pourrait être relativement simple
montrer comment une simulation basée sur des agents pourrait être interprétée comme des automates cellulaires
(par exemple, lorsqu'ils se concentrent sur leur nature en tant qu'agents/cellules), ou en tant que simulation basée
sur des équations (par exemple, si la structure interne d'un agent est constituée d'équations). L'essentiel est de voir
quelle caractéristique de la simulation informatique est mise en évidence. Ici, je vous propose quelques
critères de caractérisation sonore de chaque type. Une dernière mise en garde est cependant émise quant à la
méthodologie et l'épistémologie propres à chaque type. Ce n'est pas
difficile de montrer que chaque type de simulation informatique implique des
préoccupations méthodologiques et épistémologiques, et nécessitent donc une approche différente.
traitement à leur manière. Dans le rappel de ce livre, je concentre mon attention uniquement
sur les simulations dites basées sur des équations.

---

[20] Le préfixe "pseudo" reflète le fait que ces méthodes sont basées sur un algorithme qui produit
nombres sur une base récursive, répétant éventuellement la série de nombres produits. Le pur hasard
dans les ordinateurs ne peut jamais être atteint.

Les références

Ajelli, Marco, Bruno Gonc¸alves, Duygu Balcan, Vittoria Colizza, Hao Hu, Jose´ J. Ramasco, Stefano Merler et Alessandro Vespignani. 2010. "Comparaison des approches informatiques à grande échelle à la modélisation épidémique : modèles de métapopulation basés sur les agents et modèles structurés." BMC Infectious Diseases 10 (190): 1–13.

Ardourel, Vincent et Julie Jebeile. 2017. "Sur la supériorité présumée des solutions analytiques sur les méthodes numériques." Journal européen de philosophie des sciences 7 (2): 201–220.

Banks, Jerry, John Carson, Barry L. Nelson et David Nicol. 2010. Simulation de système à événements discrets. Upper Saddle River, New Jersey : Prentice Hall.

Barnstorff, Kathy. 2010. Le X-51A effectue le vol Scramjet le plus long.

Beisbart, Claus. 2012. « Comment les simulations informatiques peuvent-elles produire de nouvelles connaissances ? » Journal européen de philosophie des sciences 2: 395–434.

Bentsen, M., I. Bethke, JB Debernard, T. Iversen, A. Kirkevag, ?. Seland, H. Drange, et al. 2013. "Le modèle norvégien du système terrestre, NorESM1-M - Partie 1 : Description et évaluation de base du climat physique." Développement de modèles géoscientifiques 6, no. 3 (mai): 687–720.

Birtwistle, GM 1979. DEMOS Un système de modélisation d'événements discrets sur Simula. (Réimpression 2003). La presse MacMillan.

Ceruzzi, Paul E. 1998. Une histoire de l'informatique moderne. Presse du MIT. ISBN : 0-262-53203-4.

De Mol, Liesbeth et Giuseppe Primiero. 2014. "Affronter l'informatique comme technique: vers une histoire et une philosophie de l'informatique." Philosophie et technologie 27 (3): 321–326.

————. 2015. "Quand la logique rencontre l'ingénierie: introduction aux problèmes logiques dans l'histoire et la philosophie de l'informatique." Histoire et philosophie de la logique 36 (3): 195–204.

Drogoul, Alexis, Jacques Ferber et Christophe Cambier. 1992. "La simulation multi-agents comme outil de modélisation des sociétés : application à la différenciation sociale dans les colonies de fourmis." Dans Simulating Societies, édité par G. Nigel Gilbert, 49–62. Guilford : Université du Surrey.

El Skaf, Rawad et Cyrille Imbert. 2013. "Déploiement dans les sciences empiriques : expériences, expériences de pensée et simulations informatiques." Synthèse 190 (16): 3451–3474.

Fox Keller, Evelyne. 2003. "Modèles, simulations et" expériences informatiques "." Dans The Philosophy of Scientific Experimentation, édité par Hans Radder, 198–215. University of Pittsburgh Press.

Frigg, Roman et Julian Reiss. 2009. « La philosophie de la simulation : Hot New
Problèmes ou même vieux ragoût ? » Synthèse 169 (3): 593–613.

Garner, Martin. 1970. "Les combinaisons fantastiques du nouveau solitaire de John Conway
jeu "la vie". Scientifique américain 223 (4): 120–123.

Gilbert, G. Nigel et Klaus G. Troitzsch. 2005. Simulation pour le spécialiste des sciences sociales.
2e éd. LCCB : HM51 .G54 2005. Maidenhead, Angleterre ; New York, NY :
Presse universitaire ouverte. ISBN : 978-0-335-21600-0.

Guala, Francesco. 2002. "Modèles, simulations et expériences". En mode basé sur un modèle
Reasoning: Science, Technology, Values, édité par L. Magnani et NJ Nersessian, 59–74.
Académique Kluwer.

Gupta, Anil K., Raj K. Singh, Sudheer Joseph et Ellen Thomas. 2004. "Indien
Evénement océanique de haute productivité (10–8 Ma) : Lié au Refroidissement Global ou au
Initiation des moussons indiennes ? Géologie 32 (9): 753.

Hartmann, Stéphane. 1996. « Le monde en tant que processus : simulations dans le naturel et
Sciences sociales." Dans Modélisation et simulation en sciences sociales de la
Philosophy of Science Point of View, édité par R. Hegselmann, Ulrich Mueller,
et Klaus G. Troitzsch, 77–100. Springer.

Himeno, Ryutaro. 2013. "La plus grande simulation de réseau neuronal réalisée à l'aide de K
Ordinateur."

Hoose, Corinna, Jon Egill Kristj ansson, Jen-Ping Chen et Anupam Hazra. 2010.
"Une paramétrisation basée sur la théorie classique de la nucléation hétérogène de la glace
par la poussière minérale, la suie et les particules biologiques dans un modèle climatique mondial.
Journal des sciences de l'atmosphère 67, no. 8 (août): 2483-2503.

Humphreys, Paul W. 1990. "Simulations informatiques". PSA: Actes de la réunion biennale de
l'Association de philosophie des sciences 2: 497–506.

————. 2004. Extension de nous-mêmes : science computationnelle, empirisme et méthode
scientifique. Presse universitaire d'Oxford.

Kroepelin, S. 2006. "Revisiter l'âge du désert du Sahara." 312, non. 5777 (mai):
1138b–1139b.

Laboratoire, Los Alamos National. 2015. La plus grande simulation de biologie computationnelle
Imite la nanomachine la plus essentielle de la vie.

Lesne, Annick. 2007. "La controverse discrète versus continue en physique."
Structures mathématiques en informatique 17 (2): 185–223.

McMillan, Claude et Richard F. Gonzalez. 1965. Analyse des systèmes : un ordinateur
Approche des modèles de décision. Homewood/Ill : Irwin.

Morrisson, Margaret. 2009. "Modèles, mesure et simulation informatique :
Changer le visage de l'expérimentation. Études philosophiques 143 (1): 33–57.

Morrisson, Margaret. 2015. Reconstruire la réalité. Modèles, mathématiques et simulation
ulations. Presse universitaire d'Oxford.

Naylor, Thomas H., JM Finger, James L. McKenney, Williams E. Schrank et
Charles C. Holt. 1967. "Vérification des modèles de simulation informatique." Management
Science 14 (2): 92–106.

Oberkampf, William L, Timothy G Trucano et Charles Hirsch. 2003. Verifica tion, Validation, and
Predictive Capability in Computational Engineering and
La physique. Laboratoires nationaux de Sandia.

Oren, Tucer. 1984. "Activités basées sur un modèle : un changement de paradigme. En Simulation et
Méthodologies basées sur des modèles : une vue intégrative, édité par Tuncer Oren, B. P
Zeigler et MS Elzas, 3–40. Springer.

————. 2011a. « Un examen critique des définitions et environ 400 types de modélisation
et simulation. SCS M&S Magazine 2 (3): 142–151.

————. 2011b. "Les multiples facettes de la simulation à travers une collection d'environ 100
Définitions. SCS M&S Magazine 2 (2): 82–92.

Oren, Tuncer, BP Zeigler et MS Elzas, éd. 1982. Actes de l'OTAN
Institut d'études avancées sur la simulation et les méthodologies basées sur les modes : un
Vue intégrative. Série ASI de l'OTAN. Springer Verlag.

Parker, Wendy S. 2009. « Est-ce que ça compte vraiment ? Simulations informatiques,
expériences et matérialité. Synthèse 169 (3): 483–496.

Ramstein, Gilles, Fred´eric Fluteau, Jean Besse, and Sylvie Joussaume. 1997. "Effet de
l'orogenèse, du mouvement des plaques et de la distribution terre-mer sur le climat eurasien
Changement au cours des 30 derniers millions d'années. 386, non. 6627 (avril): 788–795.

Saam, Nicole J. 2016. « Qu'est-ce qu'une simulation informatique ? L'avis d'un passionné
Débat." Journal de philosophie générale des sciences : 1–17. ISSN : 15728587.
doi : 10.1007/s10838-016-9354-8.

Schelling, Thomas C. 1971. "Sur l'écologie des micromotifs." Affaires nationales 25
(Automne).

Schuster, M. 2006. "L'âge du désert du Sahara." Sciences 311, non. 5762 (février) : 821–821.
ISSN : 0036-8075, 1095-9203. doi:10 . 1126 / sciences.
1120161. http://science.sciencemag.org/content/311/
5762/821.full.

Shannon, Robert E. 1975. Simulation de systèmes : l'art et la science. Englewood
Falaises, NJ : Prentice Hall.

————. 1978. "Conception et analyse d'expériences de simulation". Dans les actes
de la 10e Conférence sur la simulation hivernale - Volume I, 53–61. Presse IEEE.

———. 1998. "Introduction à l'art et à la science de la simulation." Dans Actes de la 30e conférence sur la simulation hivernale, 7–14. Los Alamitos, Californie, États-Unis : IEEE Computer Society Press.

Shubik, Martin. 1960. "Simulation de l'industrie et de l'entreprise." La revue économique américaine 50 (5): 908–919.

Smith, Reid G. et Randall Davis. 1981. "Frameworks for Cooperation in Distributed Problem Solving". Transactions IEEE sur les systèmes, l'homme et la cybernétique 11 (1): 61–70. ISSN : 0018-9472. doi : 10.1109/TSMC.1981.4308579.

Teichroew, Daniel et John Francis Lubin. 1966. "Simulation informatique - Discussion de la technique et comparaison des langages." Communications de l'ACM 9, no. 10 (octobre): 723–741. ISSN : 0001-0782.

Toffoli, Tommaso. 1984. "CAM : Une machine à automate cellulaire à haute performance." Physica D: Phénomènes non linéaires 10 (1-2): 195–204. ISSN : 01672789. doi : 10.1016/0167-2789(84)90261-6.

Vallverdu, Jordi. 2014. « Que sont les simulations ? Une approche épistémologique. Pro cédia Technology 13:6–15.

Vichniac, Gérard Y. 1984. "Simuler la physique avec des automates cellulaires." Physica D: Phénomènes non linéaires 10 (1-2): 96–116. ISSN : 01672789. doi : 10.1016/0167-2789(84)90253-7.

Weisberg, Michel. 2013. Simulation et similarité. Presse universitaire d'Oxford.

Winsberg, Éric. 2010. La science à l'ère de la simulation informatique. Presse de l'Université de Chicago.

———. 2015. « Simulations informatiques en sciences ». Dans The Stanford Encyclopedia of Philosophy, édité par Edward N. Zalta.

Wolfram, Stephen. 1984a. "Préface." Physique 10D : vii–xii.

———. 1984b. "Universalité et complexité dans les automates cellulaires." Physique D : Phénomènes non linéaires : 1–35.

Woolfson, Michael M., et Geoffrey J. Pert. 1999a. Une introduction aux simulations informatiques . Presse universitaire d'Oxford.

———. 1999b. SATELLIT.POUR.

Zhang, Zhongshi, Gilles Ramstein, Mathieu Schuster, Camille Li, Camille Contoux et Qing Yan. 2014. "Aridification du désert du Sahara causée par le rétrécissement de la mer de Téthys au cours du Miocène supérieur." Nature 513, non. 7518 (septembre) : 401–404.

Chapitre 2
Unités d'analyse I : modèles et simulations informatiques

Théories, modèles, montages expérimentaux, prototypes : ce sont quelques-unes des unités d'analyse typiques que l'on trouve dans les travaux scientifiques et d'ingénierie standard. La science et l'ingénierie sont bien sûr peuplées d'autres unités d'analyse tout aussi décisives qui facilitent notre description et notre connaissance du monde. Ceux-ci incluent des hypothèses, des conjectures, des postulats et une foule de mécanismes théoriques. Les simulations informatiques sont la nouvelle acquisition dans le domaine scientifique et technique qui compte comme de nouvelles unités d'analyse.1 Quels sont les composants des simulations informatiques qui composent une telle nouvelle unité ? Qu'est-ce qui les différencie des autres unités d'analyse ? Les réponses à ces questions sont proposées ici.

Dans le chapitre précédent, j'ai tourné notre intérêt vers les simulations informatiques qui implémentent des équations telles qu'elles sont régulièrement trouvées et utilisées dans les sciences et l'ingénierie. Ce chapitre vise à préciser ces remarques générales. À cette fin, la première section clarifie la notion de modèle scientifique et d'ingénierie, car elle est à la base des simulations basées sur des équations. J'ai également mentionné que leur mise en œuvre n'est pas directement sur l'ordinateur physique – rappelons que c'était une hypothèse fondamentale de la description des modèles de comportement point de vue – mais plutôt médiatisée par une méthodologie appropriée pour les simulations informatiques. La deuxième partie de ce chapitre présente en détail comment les modèles sont mis en œuvre en tant que « modèles de simulation ». À cette fin, je présente et discute trois principaux éléments constitutifs des simulations informatiques, à savoir les spécifications, les algorithmes et les processus informatiques. Je présenterai ensuite d'importants problèmes sociaux, techniques et philosophiques qui émergent de cette caractérisation. En faisant cela, j'espère, nous enracinerons le statut des simulations informatiques en tant qu'unités d'analyse à part entière.

[1] Le Big Data - dont je parle au chapitre 6 - et l'apprentissage automatique devraient également être inclus en tant que nouvelles unités d'analyse en science et en ingénierie.

2.1 Modèles scientifiques et techniques

La recherche scientifique et technique d'aujourd'hui dépend fortement des modèles. Mais comment ça un « modèle scientifique » ? et qu'est-ce qu'un « modèle d'ingénierie » ? sont-ils différents ? et comment les étudier ? À première vue, ces questions peuvent trouver une réponse dans de nombreuses manières différentes. Tibor Muller et Harmund M uller ont fourni dix-neuf exemples éclairants des différentes façons dont la notion de modèle se trouve dans la littérature (Muller et M¨ uller 2003, 1-31). Certaines notions mettent l'accent sur les usages et les finalités des modèles. Un exemple est Canadarm 2 - formellement la Station spatiale Remote Manipulator System (SSRMS) - qui a d'abord été conçu comme un modèle pour l'assistance et la maintenance à bord de la Station Spatiale Internationale puis est devenu le bras robotique que l'on connaît aujourd'hui. D'autres notions sont plus intéressées dans la compréhension de l'apport épistémologique d'un modèle. Par exemple, un newtonien modèle donne un aperçu du mouvement planétaire, alors qu'un modèle ptolémaïque ne parvient pas à le faire. Par ailleurs, certaines définitions mettent en évidence la valeur propédeutique de modèles avant tout. Dans ce dernier cas, le modèle ptolémaïque est aussi précieux que le modèle newtonien puisqu'ils illustrent tous deux des normes scientifiques différentes.

Une manière standard de concevoir des modèles dépend de leurs propriétés matérielles - ou de leur absence de celle-ci. Appelez modèles de matériaux ces modèles qui sont physiques dans un simple sens, tels que des modèles en bois, en acier ou tout autre type de matériau (par exemple, Canadarm); et appellent modèles conceptuels ou abstraits les modèles qui sont les produit de l'abstraction, des idéalisations ou des fictionnalisations d'un système cible, tel que modèles théoriques, modèles phénoménologiques et modèles de données, entre autres types (par exemple, le modèle newtonien et ptolémaïque du système planétaire).2

A première vue, les modèles matériels semblent plus proches de la notion d'expérimentation en laboratoire que les modèles abstraits, eux-mêmes plus proches des simulations informatiques. Comme nous discuterons dans la section 3, certains philosophes ont utilisé cette distinction pour fonder leurs affirmations sur le pouvoir épistémologique des simulations informatiques. Spécifiquement, il a été avancé que le caractère abstrait des simulations informatiques réduit leur capacité à inférer au monde, et qu'elles sont donc épistémiquement inférieures aux simulations de laboratoire. expérimentation. Il s'avère cependant qu'il y a autant d'abstraction en jeu dans l'expérimentation en laboratoire qu'il y en a dans les simulations informatiques - bien que, naturellement, pas dans la même quantité. L'inverse, en revanche, n'est pas vrai. Ordinateur les simulations n'ont aucune utilité pour les modèles de matériaux, car tout ce qu'ils peuvent mettre en œuvre sont modèles conceptuels. Discutons maintenant brièvement de ces idées.

Les modèles matériels sont parfois destinés à être un « morceau » du monde et, en tant que tels, une réplique plus ou moins précise d'un système cible empirique. Prenons comme exemple le utilisation d'un faisceau de lumière pour comprendre la nature de la lumière. Dans un tel cas, le le modèle matériel et le système cible partagent des propriétés et des relations évidentes : le modèle et le système cible sont constitués des mêmes matériaux. Dans ce cas particulier, la distinction entre un modèle de faisceau lumineux et un faisceau lumineux en lui-même est uniquement programmatique : il n'y a pas de réelles différences si ce n'est qu'il est plus facile

---

2 Pour un excellent traitement sur les modèles et les simulations informatiques, voir les travaux de Margaret Mor prison (Morrison 2015).

pour le chercheur de manipuler le modèle qu'un véritable faisceau de lumière. Ce fait est particulièrement vrai lorsqu'il existe des différences d'échelle entre le modèle et le système cible. Dans certains cas, il est beaucoup plus facile d'allumer une lampe de poche que d'essayer de capter la lumière du soleil qui passe par la fenêtre.

Bien sûr, les modèles de matériaux n'ont pas toujours besoin d'être constitués des mêmes matériaux que le système cible. Considérez par exemple l'utilisation d'un réservoir d'ondulation pour comprendre la nature de la lumière. Le réservoir d'ondulation est un modèle de matériau au sens simple, car il est composé de métal, d'eau, de capteurs, etc. Cependant, sa configuration matérielle de base diffère considérablement de son système cible, qui est léger. Qu'est-ce qui pourrait amener les scientifiques à croire que le réservoir d'ondulation pourrait les aider à comprendre la nature de la lumière ? La réponse est que les ondes sont facilement reproduites et manipulées dans un réservoir d'ondulation, et par conséquent, il est très utile pour comprendre la nature ondulatoire de la lumière.

Nous avons ici un exemple de modèle matériel – c'est-à-dire le réservoir d'ondulation – qui implique certains niveaux d'abstraction et d'idéalisation par rapport à son système cible. L'argument est simple. Étant donné que le milieu dans lequel les ondes se déplacent dans le réservoir d'ondulation (c'est-à-dire l'eau) est différent du système cible (c'est-à-dire la lumière), il doit y avoir une représentation abstraite de haut niveau qui relie ces deux ensemble. Cela vient avec les équations de Maxwell, l'équation d'onde de D'Alambert et la loi de Hook. C'est parce que les deux systèmes (c'est-à-dire les ondes d'eau et les ondes lumineuses) obéissent aux mêmes lois et peuvent être représentés par le même ensemble d'équations que les chercheurs peuvent utiliser un modèle matériel d'un type pour comprendre un type matériellement différent.

Maintenant, quelles que soient les nombreuses façons dont les modèles de matériaux peuvent être interprétés, ils ne peuvent jamais être mis en œuvre sur un ordinateur. Comme on pouvait s'y attendre, la raison est purement basée sur leur matérialité : seuls les modèles conceptuels conviennent comme simulations informatiques. Les modèles conceptuels sont une représentation abstraite – et parfois formelle – d'un système cible et, en tant que tels, ils se situent à l'opposé ontologique des modèles matériels. Cela signifie qu'ils sont intangibles, qu'ils ne sont pas voués à la détérioration – même s'ils sont oubliables – et qu'ils existent en quelque sorte hors du temps et de l'espace, tout comme les mathématiques et l'imagination.

On s'accorde généralement à considérer ces types de modèles comme des structures interprétées qui facilitent la réflexion sur le monde[3] . De plus , les philosophes s'accordent également à dire que divers niveaux d'abstractions, d'idéalisations et d'approximations sont attachés à ces modèles dans le cadre de leurs structures inhérentes[5]. Un exemple

---

[3] Pour deux excellents livres sur les modèles scientifiques, voir (Morgan et Morrison 1999 ; Gelfert 2016).

[4] Les modèles et théories scientifiques incluent parfois des termes sans interprétation spécifique, comme « trou noir », « mécanisme », etc. Celles-ci sont connues sous le nom de métaphores et sont généralement utilisées pour combler le vide introduit par ces termes. Leur utilisation est d'inspirer une sorte de réponse créative chez les utilisateurs du modèle qui ne peut être égalée par un langage littéral. De plus, il existe une relation particulière entre le langage métaphorique et les modèles, une relation qui inclut des subtilités sur les modèles en tant que métaphores (Bailer Jones 2009, 114). Ici, je n'ai aucun intérêt à explorer ce côté de la modélisation.

[5] Le traitement philosophique de l'abstraction, des idéalisations et des approximations est assez similaire dans la littérature. L'abstraction vise à ignorer les caractéristiques concrètes que possède le système cible afin de se concentrer sur leur configuration formelle. Les idéalisations, en revanche, se présentent sous deux formes : Les idéalisations aristotéliciennes, qui consistent à « dépouiller » des propriétés que nous croyons ne pas être

éclairera certains des termes utilisés jusqu'ici. Prenons un modèle mathématique de le mouvement planétaire. Les physiciens commencent par rassembler certaines théories (par exemple, un modèle newtonien) qui abstrait et idéalise déjà le mouvement planétaire, des bits de preuves empiriques (par exemple, des données recueillies à partir d'observations) qui sont présélectionnés et post-traités, et comme tels soumis à une série de décisions méthodologiques, épistémologiques et éthiques, et enfin un récit qui conceptualise et structure le modèle dans son ensemble.

En raison de sa simplicité et de son élégance, cette manière d'appréhender les modèles est privilégiée par un grand nombre de chercheurs de toutes disciplines. Notons la forte présence d'un lien représentationnel avec le système cible. Modèles abstraits, idéaliser, et approximer parce que leur fonction est de représenter un système cible comme aussi précisément que possible, en omettant les détails inutiles. Les philosophes ont systématiquement attiré l'attention sur le besoin impérieux de modèles à représenter. Les raisons, on le croit, découlent de notre notion de progrès de la science. Autrement dit, la science serait capable de mieux faire progresser la compréhension du monde avec des modèles qui réalisent certaines fonction épistémologique (par exemple, explication, prédiction, confirmation). De telles fonctions épistémologiques, à leur tour, sont comprises comme détenant des liens représentationnels avec le monde.

Les modèles non représentationnels, c'est-à-dire les modèles qui remplissent des rôles autres que celui de représenter un système cible - tels que des rôles propédeutiques, pragmatiques et esthétiques - sont, cependant, acquérant rapidement plus d'importance dans la philosophie des modèles.[6] Ici, cependant, nous nous intéressons surtout aux modèles qui représentent, plus ou moins précisément, le système cible prévu. La raison en est en partie parce que la plupart des pratiques et de la théorie des simulations informatiques sont encore effectuées avec des modèles qui représentent un système cible, et en partie parce qu'il existe peu d'études sur le côté non représentationnel de simulations informatiques. Cela dit, trois types de modèles basés sur leur capacité de représentation émergent, à savoir les modèles phénoménologiques, les modèles de données, et des modèles théoriques.[7]

Un exemple standard de modèles phénoménologiques est le modèle de goutte liquide de la noyau atomique. Ce modèle décrit plusieurs propriétés du noyau atomique, telles que comme tension et charge superficielles, sans réellement postuler de mécanisme sous-jacent. Voici la principale caractéristique des modèles phénoménologiques : ils miment des propriétés observables plutôt que d'avancer sur une structure sous-tendant le système cible. Une telle caractéristique ne devrait cependant pas suggérer que certains aspects des modèles phénoménologiques ne pourraient pas être déduits de la théorie. De nombreux modèles incorporent des principes et des lois associés à la théorie, tout en restant phénoménologiques.

---

pertinent pour nos besoins ; et les idéalisations galiléennes, qui impliquent des distorsions délibérées. Pour ce qui est de approximations, elles sont une représentation inexacte d'une caractéristique du système cible, pour des raisons telles que les limitations pratiques dans l'approche d'un système ou la traçabilité.

[6] Les travaux d'Ashley Kennedy sur le rôle explicatif des modèles non représentationnels (Kennedy 2012) et de Tarja Knuuttila sur la dimension matérielle des modèles qui en fait des objets de connaissances et leur permet de faire la médiation entre différentes personnes et diverses pratiques (Knuuttila 2005).

[7] Il va sans dire que ces trois modèles n'épuisent pas non plus tous les modèles trouvés en science et l'ingénierie, ou les nombreuses dimensions que les philosophes utilisent pour analyser la notion. Ils sont, cependant, une assez bonne caractérisation du type de modèles convenant à une simulation informatique.

Le modèle de la goutte de liquide est à nouveau l'exemple. Alors que l'hydrodynamique explique
tension superficielle, l'électrodynamique compte pour la charge. Le modèle de la goutte liquide reste néanmoins un
modèle phénoménologique.

Pourquoi les scientifiques et les ingénieurs seraient-ils intéressés par de tels modèles ? Une raison est
qu'imiter les fonctionnalités d'un système cible est parfois plus facile à gérer que
la théorie d'un tel système cible. Les théories fondamentales telles que l'électrodynamique quantique pour la matière
condensée ou la chromodynamique quantique (QCD) pour le nucléaire
la physique sont plus accessibles lorsqu'un modèle phénoménologique est utilisé à la place du
la théorie elle-même.8

Un fait intéressant à propos des modèles phénoménologiques est que les chercheurs ont
ont parfois minimisé leur place dans les disciplines scientifiques et techniques. À
beaucoup, une raison fondamentale d'utiliser des modèles est qu'ils postulent des
mécanismes des phénomènes, facilitant ainsi la formulation d'affirmations significatives
concernant le système cible. Les modèles qui se contentent de décrire ce que nous observons, comme les modèles
phénoménologiques, ne possèdent pas une propriété aussi fondamentale. Fritz Londres,
l'un des frères qui a développé le modèle phénoménologique Londres-Londres de
supraconductivité, ont insisté sur le fait que leur modèle ne devait être considéré que comme un
remplacement jusqu'à ce qu'une approximation théorique puisse être élaborée.

Les modèles de données sont similaires aux modèles phénoménologiques. Ils partagent tous les deux le manque
d'un fondement théorique, et ne capturent donc que les caractéristiques observables et mesurables d'un
phénomène.9 Malgré cette similitude, il existe également des différences qui
un modèle de données qui vaut la peine d'être étudié en lui-même. D'abord phénoménologique
les modèles sont basés sur l'évaluation du comportement du système cible, alors que
les modèles de données sont basés sur les données reconstruites recueillies à partir de l'observation et
mesurer les propriétés d'intérêt sur le système cible. Une autre différence importante réside dans leur méthodologie.
Les modèles de données sont caractérisés par une collection de
données bien organisées, et donc la conception et la construction diffèrent des modèles phénoménologiques. En
particulier, les modèles de données nécessitent plus de machinerie statistique et mathématique que tout autre
modèle car les données collectées doivent être filtrées
pour le bruit, les artefacts et autres sources d'erreur.

En conséquence, l'ensemble des problèmes entourant les modèles de données est significativement différent
des autres types de modèles. Par exemple, un problème standard est de décider
quelles données doivent être supprimées et selon quels critères. Un problème connexe consiste à décider quelle
fonction de courbe représente toutes les données nettoyées. Est-ce que ça va être une courbe
ou plusieurs courbes? Et quels points de données devraient être laissés de côté lorsqu'aucun
la courbe les convient-elles toutes ? Les problèmes concernant l'ajustement des données dans une courbe sont généralement traités
avec inférence statistique et analyse de régression.

De nombreux exemples de modèles de données proviennent de l'astronomie, où il est courant
pour trouver des collections de grandes quantités de données obtenues à partir de l'observation et de la mesure
événements astronomiques. Ces données sont classées par paramètres spécifiques, tels que la luminosité, le
spectre, la position céleste, le temps, l'énergie, etc. Un problème habituel pour le
l'astronome est de savoir quel modèle peut être construit à partir d'un tas de données donné.

---

8 Pour les raisons pour lesquelles c'est le cas, voir (Hartmann 1999, 327).

9 Comme nous le verrons dans la section 6.2, les chercheurs font des efforts importants pour trouver une structure derrière
grandes quantités de données.

À titre d'exemple, prenons le modèle de données de l'observatoire virtuel, un projet mondial où les métadonnées sont utilisées pour la classification des données d'observatoire, la construction de modèles de données et simulation de nouvelles données.10

Considérons maintenant un autre type de modèle, celui qui postule un mécanisme théorique sous-jacent du système cible, et appelons-le modèle théorique. Traditionnel les représentants de ce type de modèle sont les modèles Navier-Stoke de la dynamique des fluides, les équations de Newton du mouvement planétaire et le modèle de ségrégation de Schelling. Ainsi compris, les modèles théoriques incarnent des connaissances issues de théories bien établies et sous-tendent les structures profondes du système cible. Pour plusieurs raisons, ces types de modèles ont été les favoris des chercheurs. Cependant, à mesure que les modèles de données et les modèles phénoménologiques gagnent du terrain avec les nouvelles technologies, elles se positionnent épistémiquement au même niveau que les théories des modèles.

Enfin, les modèles en ingénierie reçoivent parfois un traitement différent des modèles en sciences. La raison en est qu'ils sont conçus comme des modèles pour 'faire' plutôt que pour 'représenter'. Malheureusement, cette dichotomie occulte fait que toutes sortes de modèles sont utilisés dans un but précis et, à cet égard, ils servent autant à « faire » qu'à « représenter ». De plus, penser présuppose ainsi une distinction entre science et ingénierie, qui dans le domaine des simulations informatiques semble être très dissous. Où est-ce que la sociologie se termine et l'ingénierie commence dans une simulation mettant en œuvre le Schelling modèle de ségrégation ? Bien sûr, il y a les sciences, et il y a les ingénieurs, et dans de nombreux cas, ces deux grandes disciplines peuvent être bien différenciées11 . sont aussi beaucoup d'intersections. Les simulations informatiques, je crois, résident dans l'un de ces intersection. De ce fait, et pour la suite de ce livre, je traite des modèles scientifiques et les modèles d'ingénierie d'une manière similaire.12

2.2 Simulations informatiques

2.2.1 Composants des simulations informatiques

Dans le chapitre précédent, j'ai réduit la classe des simulations informatiques d'intérêt aux simulations basées sur des équations. William Oberkampf, Timothy G. Trucano et

---

dix Parmi les groupes internationaux travaillant sur des modèles d'observatoire de données, il y a l'International Alliance Observatoire Virtuel (Alliance 2018) ; et l'analyse du milieu interstellaire des galaxies isolées (AMIGA) (milieu interstellaire des galaxies isolées 2018)

11 Cela est particulièrement vrai pour les modèles de matériaux. Par exemple, le modèle conceptuel du Canadarm comprenaient des notions distinctes de l'ingénierie et distinctes des sciences. Mais une grande partie était un mélange des deux. Cependant, c'est avec le bras mécanique qu'il devient incontestablement un artefact d'ingénierie. C'est sur ce point que se situe une grande partie du lourd travail de la philosophie de la technologie est en cours.

12 De bons travaux sur la compréhension de la modélisation en ingénierie peuvent être trouvés dans (Meijers 2009), en particulier Partie IV : Modélisation en sciences de l'ingénieur.

Charles Hirsch attire l'attention sur ce type de simulations informatiques en « ingénierie et physique computationnelles », qui comprend la dynamique des fluides computationnelle, la mécanique des solides computationnelle, la dynamique des structures, la physique des ondes de choc et la chimie computationnelle, entre autres.13 simulations informatiques c'est ainsi souligner leur double dépendance. D'une part, il y a une dépendance vis-à-vis des sciences naturelles, des mathématiques et de l'ingénierie. D'autre part, ils s'appuient sur les ordinateurs, l'informatique et l'architecture informatique. Bien entendu, les simulations basées sur des équations s'étendent également à d'autres domaines de recherche. Dans le sciences de la vie, par exemple, les simulations basées sur des équations sont importantes en biologie synthétique, et en médecine, il y a eu des progrès incroyables avec la clinique in silico essais. On pourrait également mentionner les nombreuses simulations réalisées dans le domaine de l'économie et de la sociologie.

Les philosophes ont largement reconnu l'importance d'étudier la méthodologie des simulations informatiques pour comprendre leur place dans la recherche scientifique et technique. Eric Winsberg est d'avis que la crédibilité de l'informatique simulations ne vient pas uniquement des références que lui fournit la théorie, mais aussi et peut-être dans une large mesure à partir de références établies dans la construction de modèles techniques employées dans la construction de la simulation.14 De même, Margaret Morrison considère que les inexactitudes de représentation entre les modèles de simulation et les modèles mathématiques peuvent être résolus à différentes étapes de la modélisation, telles que l'étalonnage et les tests.15 En outre, Johannes Lenhard a soutenu avec force que la Le processus de modélisation par simulation prend la forme d'une coopération exploratoire entre l'expérimentation et la modélisation (Lenhard 2007). Je suis convaincu que Winsberg, Morrison et Lenhard ont raison dans leurs interprétations. Winsberg a raison de pointer que la fiabilité des simulations informatiques a diverses sources, et donc il ne peut se limiter au modèle théorique dont sont issues les simulations. Morrison a raison de nous persuader que la capacité de représentation de la simulation modèle peut également être abordé lors des étapes de modélisation. Et Lenhard a raison pointant la relative autonomie des simulations informatiques par rapport aux modèles, données et phénomènes.

S'il est vrai que certains chercheurs utilisent pour leurs simulations des logiciels informatiques, il n'est pas rare de trouver de nombreux autres chercheurs concevant et programmer leurs propres simulations. La raison en est qu'il y a plus qu'une satisfaction intellectuelle à programmer ses propres simulations, il y a en fait de bonnes raisons épistémiques de le faire. En s'impliquant dans la conception et la programmation des simulations informatiques, les chercheurs apprennent à savoir ce qui est conçu et mis en œuvre, et par conséquent, ils ont une meilleure idée de ce à quoi on peut s'attendre dans termes d'erreurs, d'incertitudes, etc. En fait, plus les chercheurs sont impliqués dans la conception, la programmation et la mise en œuvre de simulations informatiques, mieux c'est préparés qu'ils sont pour comprendre leurs simulations.

---

13 Cela peut être trouvé dans (Oberkampf, Trucano et Hirsch 2003).

14 Voir, par exemple, son travail dans (Winsberg 2010).

15 Pour plus d'informations sur ces idées, voir (Morrison 2009).

Il existe deux sources générales qui alimentent la méthodologie des simulations informatiques. D'une part, il existe des règles formelles et des méthodes systématiques qui permettent équivalence formelle entre modèles mathématiques et modèles de simulation. Ceci est illustré par certaines méthodes de discrétisation qui permettent la transformation de fonctions dans ses homologues discrets. Des exemples de méthodes de discrétisation sont les Méthode Runge-Kutta et méthode Euler.16 Il existe également de nombreux langages de développement formels destinés à l'analyse des conceptions et à l'identification des caractéristiques clés. de modélisation, y compris les erreurs et les fausses déclarations. Ces langages, comme la notation Z, VDM ou LePUS3, utilisent généralement une sémantique de langage de programmation formelle. (c'est-à-dire la sémantique dénotationnelle, la sémantique opérationnelle et la sémantique axiomatique) pour le développement de simulations informatiques et de logiciels informatiques en général.

La seconde source d'une méthodologie de simulations informatiques a un côté plus pratique. La conception et la programmation de simulations informatiques ne dépendent pas uniquement sur les machines formelles, mais aussi sur les savoirs dits experts (Collins et Evans 2007). Ces connaissances comprennent les « trucs », le « savoir-faire », les « expériences passées », et une foule de mécanismes non formels pour concevoir et programmer leurs simulations. Il n'est pas facile de cerner la forme et les usages de ces connaissances car cela dépend sur de nombreux facteurs, tels que les communautés, les institutions et les antécédents scolaires personnels. Shari Lawrence Pfleeger et Joanne M. Atlee rapportent que les connaissances d'experts, aussi présent et aussi nécessaire qu'il soit dans la vie des logiciels informatiques, n'en est pas moins subjectif et dépendant des informations actuellement disponibles, comme les données accessibles et l'état de développement réel de l'unité logicielle.17 Dans ce contexte, il n'est pas surprenant que de nombreuses décisions concernant les logiciels informatiques qui étaient traditionnellement prises par le chercheur expert sont désormais réalisées par des algorithmes automatisés.

Une synthèse de ces deux sources consiste à s'appuyer sur des logiciels bien documentés paquets. D'une part, nombre de ces colis ont fait l'objet d'un contrôle formel ; d'autre part, les équipes impliquées dans leur développement sont les experts en la matière. Un exemple en sont les générateurs de nombres aléatoires, qui sont au cœur de stochastique simulations telles que les simulations de Monte Carlo. Pour ces progiciels, il est important de montrer formellement qu'ils se comportent de la manière dont ils ont été spécifiés, car dans de cette façon, les chercheurs connaissent les limites de sa fonctionnalité - nombre pseudo-aléatoire inférieur et supérieur généré - erreurs possibles, répétitions, etc. Choisir le le bon moteur de génération de nombres aléatoires est une décision de conception importante, car la précision et la précision des résultats en dépendent.18

Les deux sources mentionnées ci-dessus sont présentes lors des nombreuses étapes de conception et de programmation des simulations informatiques. Négliger les routines non formelles, par exemple, conduit à l'idée erronée qu'il existe une méthode axiomatique pour concevoir simulations informatiques. De même, ne pas reconnaître la présence de et les méthodes formelles conduisent à une vision des simulations informatiques comme un outil non structuré, discipline sans fondement. Le défi est donc de comprendre le rôle que chacun joue dans l'activité complexe qu'est la méthodologie des simulations informatiques. C'est

---

16 Pour une discussion sur les techniques de discrétisation, voir (Atkinson, Han et Stewart 2009), (Gould, Tobochnik, et Christian 2007), et (Butcher 2008) entre autres dans un très large corpus de littérature.

17 Pour plus d'informations sur le cycle de vie des logiciels informatiques, voir (Pfleeger et Atlee 2009).

18 Les loci classici sur ce sujet sont (Knuth 1973) et (Press et al. 2007)

il est également vrai que, malgré les nombreuses façons dont les simulations informatiques sont conçues et programmées, il existe une méthodologie normalisée qui fournit les bases de nombreuses simulations informatiques. Il est difficile d'imaginer une équipe de scientifiques et d'ingénieurs modifiant leurs méthodologies générales chaque fois qu'ils conçoivent et programment une nouvelle simulation informatique. La méthodologie des simulations informatiques, pourrions-nous dire, repose sur des idéaux de stabilité de comportement, de fiabilité, de robustesse et de continuité dans la conception et la programmation.19

Dans ce qui suit, je discute d'une méthodologie pour les logiciels informatiques en général, et pour les simulations informatiques en particulier.20 L'idée est d'avoir un aperçu général de ce qui comprend un modèle de simulation typique et de ce que cela signifie d'être une simulation informatique. J'aborderai également quelques questions qui ont retenu l'attention des philosophes. J'espère qu'à la fin de cette section, nous serons convaincus que les simulations informatiques représentent une nouvelle unité d'analyse pour la science et l'ingénierie.

2.2.1.1 Spécifications

Chaque instrument scientifique nécessite l'élucidation de sa fonctionnalité et de son opérabilité, de sa conception et de ses contraintes. Le type d'information présenté ici constitue la spécification de cet instrument scientifique. Prenez par exemple le thermomètre à mercure dans le verre avec les spécifications suivantes :

1. Insérez du mercure dans une ampoule de verre fixée à un tube de verre de diamètre étroit ; le volume du mercure dans le tube doit être bien moindre que le volume dans l'ampoule ; marques de calibrage sur le tube qui varient en fonction de la chaleur donnée ; remplissez l'espace au-dessus du mercure avec de l'azote.
2. Ce thermomètre ne peut être utilisé que pour mesurer les liquides, la température corporelle et les conditions météorologiques. il ne peut pas mesurer au-dessus de 50◦ C, ou en dessous de -10◦ C ; il a un intervalle de 0,5◦ C entre les valeurs et une précision de ±0,01◦ C ; le mercure dans le thermomètre se solidifie à -38,83◦

C ; 3. Instructions pour l'utilisation correcte du thermomètre : insérer l'ampoule de mercure située à l'extrémité du thermomètre dans le liquide (sous le bras, ou à l'extérieur mais pas en plein soleil) à mesurer, repérer la valeur indiquée par la hauteur de la barre de mercure, et ainsi de suite.

---

[19] Un bon exemple de cela, mais au niveau du matériel informatique, est l'architecture de von Neumann, qui est un standard dans la conception des ordinateurs depuis son rapport de 1945 (von Neumann 1945). Naturellement, il faut aussi prendre en considération les modifications apportées à l'architecture informatique amenées par le changement technologique.

[20] Permettez-moi de noter que même si je ne présente que trois composants de simulations informatiques, la possibilité d'en avoir jusqu'à six a été évoquée. Giuseppe Primiero reconnaît jusqu'à six « niveaux d'abstraction » auxquels les systèmes de calcul sont examinés, à savoir l'intention, la spécification, l'algorithme, les instructions du langage de programmation de haut niveau, les opérations d'assemblage/de code machine, une exécution (Primiero 2016). Il existe de nombreuses approches philosophiques de la nature des spécifications, des algorithmes et des processus informatiques que je ne pourrai pas discuter. Une courte liste de références inclurait certainement (Copeland 1996 ; Piccinini 2007, 2008 ; Primiero 2014 ; Zenil 2014).

En plus de ces spécifications, nous connaissons généralement des informations pertinentes sur le système cible. Dans le cas du thermomètre, il peut être pertinent de savoir que l'eau change d'état de liquide à solide à 0◦ C ; que si le liquide n'est pas correctement isolé, alors la mesure peut être biaisée par une autre source de température ; que la loi Zéro de la thermodynamique justifie la mesure de la propriété physique « température », et ainsi de suite.

Outre la présentation des informations nécessaires à la construction d'un instrument, la spécification est également fondamentale pour établir la fiabilité de l'instrument et l'exactitude de ses résultats. Toute mauvaise utilisation du thermomètre, c'est-à-dire toute utilisation qui enfreint explicitement ses spécifications, peut entraîner des mesures inexactes. Inversement, dire qu'un thermomètre effectue la mesure requise, que la mesure est précise et exacte et que les valeurs obtenues sont des mesures fiables, c'est aussi dire que la mesure a été effectuée dans les limites des spécifications données par le fabricant.

Un cas anecdotique d'utilisation abusive d'un instrument a été évoqué par Richard Feynman alors qu'il faisait partie du comité de recherche enquêtant sur la catastrophe du Challenger. Il se souvient avoir eu la conversation suivante avec le fabricant d'un pistolet à balayage infrarouge :

> Monsieur, votre pistolet scanner n'a rien à voir avec l'accident. Il a été utilisé par les gens ici d'une manière contraire aux procédures de votre manuel d'instructions, et j'essaie de comprendre si nous pouvons reproduire l'erreur et déterminer quelles étaient réellement les températures ce matin-là. Pour ce faire, j'ai besoin d'en savoir plus sur votre instrument. (Feynman 2001, 165-166)

La situation était probablement très effrayante pour le fabricant, car il aurait pu penser que le pistolet de balayage ne fonctionnait pas comme prévu. Ce n'était en fait pas le cas. Au lieu de cela, c'est le matériau utilisé dans les joints toriques de la navette qui est devenu moins résistant par temps froid, et donc pas correctement étanche lors d'une journée inhabituellement froide à Cap Canaveral. Le point ici est que pour Feynman, ainsi que pour tout autre chercheur, la spécification est un élément d'information clé sur la conception et l'utilisation appropriées d'un instrument.

Ainsi comprises, les spécifications remplissent à la fois une fonction méthodologique et une fonctionnalité épistémologique. D'un point de vue méthodologique, cela fonctionne comme les « modèles » pour la conception, la construction et l'utilisation d'un instrument. D'un point de vue épistémique, il fonctionne comme le dépositaire et le référentiel de nos connaissances sur cet instrument, ses résultats potentiels, ses erreurs, etc. Dans ce contexte, la spécification a un double objectif. Il fournit des informations pertinentes pour la construction d'un instrument ainsi qu'un aperçu de sa fonctionnalité. À partir de l'exemple du thermomètre ci-dessus, le point 1 illustre comment construire un thermomètre, y compris comment le calibrer ; le point 2 illustre les limites supérieure et inférieure dans lesquelles le thermomètre mesure et peut être utilisé comme un instrument fiable ; le point 3 illustre l'utilisation correcte du thermomètre pour des mesures réussies.

Dans le contexte des simulations informatiques21, les spécifications jouent un rôle similaire à celui des instruments : ce sont des descriptions du comportement, des composants et des capacités.

[21] Notons que les différences entre la spécification d'une simulation informatique et tout autre logiciel informatique sont minimes. Cela est dû au fait que les simulations informatiques sont une sorte de logiciel informatique. À cet égard, la différence la plus concrète réside dans le type de modèle mis en œuvre.

liens de la simulation informatique en fonction du système cible. Brian Canwell
Smith, un philosophe intéressé par les fondements de l'informatique, la définit comme une "description
formelle dans un langage formel standard, spécifiée en termes de modèle,
dans lequel le comportement souhaité est décrit. (Cantwell Smith 1985, 20). L'exemple utilisé est celui d'un
système de distribution de lait, qui peut être spécifié au minimum comme un système de distribution de lait.
voiture de livraison qui visite chaque magasin en parcourant la distance la plus courte possible au total.

Arrêtons-nous ici un instant et analysons la définition de Cantwell Smith. Bien que
éclairante, cette définition ne rend pas compte de ce que les chercheurs appellent aujourd'hui une «
spécification ». Il y a deux raisons à cela. Elle est d'abord réductrice. Deuxièmement, il n'est pas suffisamment
inclusif. Elle est réductrice car la notion de spécification est identifiée
avec une description formelle du comportement du système cible. Cela signifie que
le comportement prévu de la simulation est décrit en termes de machinerie formelle.
L'ingénierie logicielle, cependant, a clairement montré que les spécifications ne peuvent pas être entièrement
formalisé. Elles doivent plutôt être conçues comme une forme de description « semi-formelle » du
comportement d'un logiciel informatique. Dans ce dernier sens, formel aussi bien que
des descriptions non formelles coexistent dans la spécification. Autrement dit, mathématique
et des formules logiques coexistent avec des décisions de conception en langage naturel clair, ad hoc
des solutions à la traçabilité informatique, etc. Vient ensuite la documentation
pour le code informatique, les commentaires des chercheurs à son sujet, etc. Bien qu'il y ait
accord général que la formalisation complète de la spécification est un objectif souhaitable, il est
pas toujours réalisable, surtout pour le type de spécifications très complexes
que supposent les simulations informatiques.

La définition n'est pas suffisamment inclusive en ce sens que la spécification de la bonne
Le comportement d'une simulation n'est pas indépendant de la manière dont elle est mise en œuvre. Une
spécification ne décrit pas seulement le comportement prévu de la simulation conformément à
système cible, mais intègre également les contraintes pratiques et théoriques du
ordinateur physique. Cela signifie que les préoccupations concernant la calculabilité, les performances,
l'efficacité et la robustesse de l'ordinateur physique font généralement également partie de la spécification
de la simulation. Dans l'exemple ci-dessus de la simulation de livraison de lait,
la spécification ne capture que les aspects généraux du système cible, ceux qui
sont plus importants pour le chercheur, quels que soient les détails sur la façon dont la livraison
de lait est effectivement effectuée. C'est tout à fait exact. A cela s'ajoutent les contraintes
adaptés à la physicalité de l'ordinateur doivent être ajoutés. En conséquence, la spécification sera peuplée
de nouvelles entités et relations, de solutions alternatives à la
problèmes à résoudre, et une approche globalement différente d'une solution purement formelle
fournirait.

Une idée plus précise de ce qui constitue les spécifications est donnée par Shari Pfleeger
et Joanne Atlee.22 Ces mathématiciens et informaticiens pensent que tout
la spécification du programme revient à décrire les propriétés visibles externes d'une unité logicielle, ainsi
que les fonctions d'accès du système, les paramètres, les valeurs de retour et
des exceptions. En d'autres termes, la spécification prend en charge à la fois la cible modélisée
système ainsi que l'unité logicielle.

à savoir un modèle qui décrit un système cible destiné à des fins scientifiques et d'ingénierie générales.

[22] Voir (Pfleeger et Atlee 2009).

Il existe une série d'attributs et de fonctionnalités généraux qui sont généralement inclus en tant que partie du cahier des charges. Voici une liste restreinte :

Objectif : il documente la fonctionnalité de chaque fonction d'accès, la modification de variables, accès aux E/S, etc. Cela doit être fait avec suffisamment de détails pour que d'autres développeurs puissent identifier les fonctionnalités qui correspondent à leurs besoins ;

Préconditions : ce sont des hypothèses que le modèle inclut et qui doivent être disponible pour que d'autres développeurs sachent dans quelles conditions l'unité fonctionne correctement. Les conditions préalables incluent les valeurs des paramètres d'entrée, les états des ressources globales, autres unités logicielles, etc. ;

Protocoles : cela inclut des informations sur l'ordre dans lequel les fonctions d'accès doit être invoqué, les mécanismes dans lesquels les modules échangent des messages, etc. exemple, un module accédant à une base de données externe doit être correctement autorisé ;

Postconditions : tous les effets visibles possibles de l'accès aux fonctions sont documentés, y compris les valeurs de retour, les exceptions, les fichiers de sortie, etc. Les postconditions sont important car ils indiquent au code appelant comment réagir de manière appropriée à un sortie de la fonction ;

Attributs de qualité : ce sont les performances et la fiabilité du modèle visibles aux développeurs et aux utilisateurs. Dans l'exemple du satellite en orbite autour d'une planète en section 1.1, l'utilisateur doit spécifier la variable TOLERANCE, c'est-à-dire erreur absolue maximale qui peut être tolérée dans n'importe quelle coordonnée de position. Si c'est réglé trop bas, le programme peut devenir très lent ;

Décisions de conception : le cahier des charges est aussi le lieu où se déroulent les décisions politiques, éthiques et les décisions de conception sont mises en œuvre dans le cadre du modèle de simulation.

Traitement des erreurs : il est important de préciser le flux de travail de l'unité logicielle et comment se comporter dans des situations anormales, telles qu'une entrée non valide, des erreurs lors de la erreurs de calcul, de manipulation, etc. Les exigences fonctionnelles du cahier des charges doivent indiquer clairement ce que la simulation doit faire dans ces situations. Un bien La spécification doit suivre des principes spécifiques de robustesse, d'exactitude, d'exhaustivité, de stabilité et des desiderata similaires.

Documentation : toute information supplémentaire doit également être documentée dans la spécification, telle que les spécificités des langages de programmation utilisés, les bibliothèques, les relations détenues par des structures, des fonctions supplémentaires mises en œuvre, etc. ;

Dans cette veine, et à l'instar des instruments scientifiques, les spécifications jouent deux rôles centraux rôles : ils jouent un rôle méthodologique en tant que schéma directeur pour la conception, la programmation, et la mise en œuvre de la simulation. Ce rôle consiste notamment à lier la représentation et la connaissance du système cible à la connaissance du

système informatique (c'est-à-dire l'architecture de l'ordinateur, le système d'exploitation, les langages de programmation, etc.). Cela signifie que les spécifications aident à « relier » ces deux types de connaissances. Il ne s'agit bien sûr pas d'un lien fantasmagorique ou mystérieux, mais plutôt de la base de l'activité professionnelle des informaticiens et ingénieurs : une simulation doit être spécifiée le plus précisément possible avant d'être programmée dans un algorithme, car cela fait gagner du temps. , de l'argent, des ressources et, plus important encore, réduit la présence d'erreurs, de fausses déclarations et les risques d'erreurs de calcul.

Elle joue également un rôle épistémologique dans le sens où les spécifications sont dépositaires et dépositaires de nos connaissances sur le système cible et, comme on vient de le mentionner, de l'unité logicielle également. En ce sens, les spécifications sont une unité cognitivement transparente dans le sens où les chercheurs peuvent toujours comprendre ce qui y est décrit. En fait, on peut dire qu'il s'agit de l'unité la plus transparente dans une simulation informatique. Ceci est immédiatement clair lorsqu'on le compare à l'algorithme et au processus informatique : le premier, bien que toujours accessible cognitivement, est écrit dans un langage de programmation inadapté à un humain pour le suivre d'un bout à l'autre ; ce dernier, en revanche, occulte tout accès à la simulation informatique et à son évolution dans le temps.

Illustrons ces points par un exemple. Considérons la spécification d'une simulation simple, telle que la simulation du satellite en orbite sous contrainte de marée discutée à la section 1.1. Une spécification possible comprend des informations sur le comportement du satellite, telles que le fait qu'il s'étend le long du rayon vecteur. Il inclut également les éventuelles contraintes à la simulation. Woolfson et Pert indiquent que si l'orbite n'est pas circulaire, alors la contrainte sur le satellite est variable et donc il se dilate et se contracte le long du rayon vecteur de manière périodique. De ce fait, et du fait que le satellite n'est pas conçu pour être parfaitement élastique, il y aura des effets d'hystérésis et une partie de l'énergie mécanique sera convertie en chaleur qui sera rayonnée. Une solution ad-hoc intéressante à l'élasticité du satellite consiste à le représenter par trois masses, chacune de même valeur reliées par des ressorts de même longueur libre. Naturellement, un certain nombre d'équations sont également incluses dans la spécification.[23]

Ces éléments et d'autres doivent ensuite être inclus dans la spécification avec des informations sur l'ordinateur physique. Un exemple simple de ce dernier point est que la masse de Jupiter ne peut pas être représentée sur un ordinateur dont l'architecture est basée sur un système 32 bits. La raison en est que la masse de Jupiter est de 1,898x1027 kg, et ce nombre ne peut être représenté qu'en 128 bits ou plus.

Continuons maintenant avec l'étude de l'algorithme, c'est-à-dire la structure logique chargé d'interpréter la spécification dans un langage de programmation approprié.

2.2.1.2 Algorithmes

La plupart de nos activités quotidiennes peuvent être décrites comme de simples ensembles de règles que nous répétons systématiquement. On se réveille à une certaine heure le matin, on se brosse les dents,

---

[23] Pour plus de questions philosophiques sur les spécifications, voir (Turner 2011).

prendre une douche et prendre le bus pour aller travailler. On modifie notre routine, bien sûr, mais juste assez pour la rendre plus avantageuse d'une certaine façon : ça nous donne plus de temps au lit, ça minimise la distance entre les arrêts, ça satisfait tout le monde dans la maison.

D'une certaine manière, cette routine saisit le sens de ce que nous appelons un algorithme, car elle répète systématiquement un ensemble de règles bien définies encore et encore. Jean-Luc Chabert, un historien qui a consacré beaucoup de temps à la notion d'algorithme, la définit comme "un ensemble d'instructions pas à pas, à exécuter assez mécaniquement, de manière à atteindre un résultat souhaité" (Chabert 1994, 1) . Une routine est une sorte d'algorithme ou, pour être plus précis, elle pourrait être transformée en algorithme. Mais en soi ce n'est pas un algorithme. Il nous faut donc préciser cette notion.

Reconnaissons d'abord que la notion d'algorithme existait bien avant qu'un mot ne soit inventé pour le décrire. En fait, les algorithmes ont une histoire omniprésente qui remonte aux Babyloniens et à leur utilisation pour décoder les points de droit, aux professeurs de latin qui utilisaient des algorithmes pour élaborer sur la grammaire et aux clairvoyants pour prédire l'avenir. Leur popularité commence avec les mathématiques – les méthodes d'Euler et de Runge-Kutta, et la série de Fourier ne sont que quelques exemples – et s'étend à l'informatique et à l'ingénierie. C'est à ce dernier arrêt que réside notre intérêt pour les algorithmes. Bref, pour nous, algorithme est synonyme d'algorithme informatique. [24]

Désormais, les algorithmes informatiques reposent sur l'idée qu'ils font partie d'une procédure systématique, formelle et finie - mathématique et logique - pour la mise en œuvre d'ensembles d'instructions spécifiques. Chabert, empruntant à l'Ecyclopaedia Britannica, les définit comme « une procédure mathématique systématique qui produit – en un nombre fini d'étapes – la réponse à une question ou la solution d'un problème » (2). Selon cette interprétation, le système de particules dans la cellule sans collision suivant, tel que présenté par Michael Woolfson et Geoffrey Pert, est considéré comme un algorithme :

1. Construisez une grille pratique dans l'espace à une, deux ou trois dimensions dans lequel le système peut être défini. [...]

2. Décidez du nombre de superparticules, électrons et ions, et attribuez-leur des positions. Pour obtenir le moins de fluctuation aléatoire possible dans les champs, il est nécessaire d'avoir autant de particules que possible par cellule. [...]

3. En utilisant les densités aux points de grille, l'équation de Poisson : $2\varphi = -\rho/\varepsilon 0$

...

25 N. Si le temps total de simulation n'est pas dépassé alors retour à 3.

---

[24] Il existe une origine étymologique intéressante pour le mot « algorithme ». Les archives montrent que le mot dérive en partie d'al-Khwārizmı, un mathématicien persan du IXe siècle, auteur du plus ancien ouvrage d'algèbre connu. Le monde vient aussi du latin algorismus et du grec ancien αριθμος, qui signifie « nombre ».

[25] Une discussion plus détaillée des étapes impliquées dans le processus pour les champs purement électrostatiques peut se trouvent ici (Woolfson et Pert 1999, 115).

Un examen attentif révèle que l'exemple comprend des machines mathématiques ainsi que des déclarations en langage clair. D'une certaine manière, cela ressemble plus à une spécification pour un système de particules dans la cellule sans collision qu'à un algorithme informatique approprié. Et pourtant, il se qualifie comme un algorithme selon la définition de Chabert. Ce dont nous avons besoin, c'est d'une définition plus précise.

Dans les années 1930, le concept d'algorithme informatique a été popularisé dans le cadre de la programmation informatique. Dans ce nouveau contexte, la notion a subi quelques altérations par rapport à sa formulation mathématique d'origine, principalement dans le langage de base : des mathématiques à une multitude de constructions syntaxiques et sémantiques. Cependant, ce n'était pas tout. Une liste restreinte avec les nouvelles fonctionnalités comprend :

1. Un algorithme est défini comme un ensemble fini et organisé d'instructions, destiné à apporter la solution à un problème, et qui doit satisfaire certains ensembles de conditions ;

2. L'algorithme doit pouvoir être écrit dans une certaine langue ; 3. L'algorithme est une procédure qui s'effectue pas à pas ; 4. L'action à chaque étape est strictement déterminée par l'algorithme, les données d'entrée et les résultats obtenus aux étapes précédentes ;
5. Quelles que soient les données d'entrée, l'exécution de l'algorithme se terminera après un nombre fini d'étapes ; 6. Le comportement de l'algorithme est physiquement instancié lors de l'implémentation sur la machine informatique.[26]

Remarquons que de nombreuses structures en informatique et en ingénierie réussissent également à se qualifier d'algorithmes avec ces caractéristiques. Les pseudo-codes (par exemple, l'algorithme 1) sont notre premier exemple. Ce sont des descriptions qui remplissent la plupart des conditions ci-dessus, sauf qu'elles ne sont pas implémentables sur la machine physique. La raison en est qu'ils sont principalement destinés à une lecture humaine, avec un soupçon de syntaxe formelle. De ce fait, les pseudo-codes se retrouvent assez fréquemment dans les spécifications d'un système cible, car ils facilitent la transition vers un algorithme[27].

---

[26] Ce ne sont là que quelques-unes des caractéristiques attribuées aux algorithmes par Chabert. Un plus l'histoire détaillée des algorithmes informatiques peut être trouvée dans (Chabert 1994, 455).

[27] Pour une discussion approfondie sur les différentes notions d'algorithme, voir le débat entre Robin Hill dans (Hill 2013, 2015) et Andreas Blass, Nachum Dershowitz et Yuri Gurevich dans (Blass et Gurevich 2003 ; Blass, Dershowitz et Gurevich 2009). Ici, nous devons nous plonger profondément dans des discussions philosophiques subtiles.

Algorithme 1 Pseudo-code

Le pseudo-code est une description non formelle de haut niveau de la spécification. Il est destiné à se concentrer sur le comportement opérationnel de l'algorithme plutôt que sur une syntaxe particulière. En ce sens, il utilise un langage similaire à un langage de programmation mais dans un sens très vague, omettant généralement des détails qui ne sont pas essentiels à la compréhension de l'algorithme.

```
Nécessite : n ≥ 0_x 6= 0
Assurez-vous que : y = x^n
        1
    y si n < 0 alors
        X 1/x
        N -n sinon


        X X
        N n fin
    si tant
    que N 6= 0 faire
        si N est pair alors
            XX   X _
            NN/2 sinon
        {N est impair} y
            y   X
            NN −1
        fin si
    fin tant que
```

Le plus souvent, la notion d'algorithme est adaptée à un langage de programmation. Maintenant, puisque l'univers des langages de programmation est considérablement vaste, un exemple de chaque type suffirait pour nos besoins. Je prends alors Fortran, Java, Python et Haskell comme quatre représentants des langages de programmation. Le premier est un exemple de langage de programmation impératif (voir algorithme 2) ; le second est un langage de programmation orienté objet (voir algorithme 3) ; Python est un bon exemple de langage interprété (voir algorithme 4) ; et Haskell est l'exemple de la programmation fonctionnelle (voir algorithme 5).

Algorithme 2 Fortran

Fortran est un langage de programmation impératif particulièrement adapté au calcul numérique et au calcul scientifique. Il est très populaire parmi les chercheurs travaillant avec des simulations informatiques, car les formules sont facilement implémentées alors que les performances de calcul restent élevées. En raison de sa polyvalence et de son efficacité, Fortran s'est imposé dans des domaines de calcul intensif tels que la prévision numérique du temps, l'analyse par éléments finis et la dynamique des fluides computationnelle, entre autres. C'est aussi un langage très populaire dans le calcul haute performance.

Le programme GDC calcule le plus grand diviseur commun entre deux nombres entiers saisis par l'utilisateur :

```
programme
        PGCD implicite
        aucun entier :: a, b, c

        write(*,*) 'Donnez-moi deux entiers positifs : '
        read(*,*) a, b if
                (a < b) then
                c=un
                un=b
                b=c
        fin si

        faire
                c = MOD(a, b) if
                (c == 0) exit


        a=bb=c
        end do write(*,*) 'Le PGCD
est ', b end program GCD
```

Algorithme 3 JAVA

Java est un langage de programmation informatique à usage général. Certaines de ses principales caractéristiques sont qu'il est concurrent, basé sur les classes, orienté objet et spécifiquement conçu pour avoir le moins de dépendances d'implémentation possible. Ce dernier point a fait de Java un langage de programmation populaire puisque ses applications sont généralement compilées une fois et exécutées sur n'importe quelle machine virtuelle Java, quelle que soit l'architecture de l'ordinateur.

SNP calcule la somme de deux entiers saisis par l'utilisateur :

```
package SNP ;
importer java.util.Scanner ;

public class addTwoNumbers private
        static Scanner sc; public static
        void main(String[] args) int Number1, Number2,
                Sum ; sc = nouveau Scanner
                (System.in);

                System.out.println("Entrez le premier chiffre : "); a =
                sc.nextInt();

                System.out.println("Entrez le deuxième chiffre :  »); b =
                sc.nextInt();

                somme = a +
                                                        ”
                b; System.out.println("La somme est =          + somme);
```

Algorithme 4 Python

Python est un langage de programmation de haut niveau largement utilisé pour la programmation à usage général. C'est l'une des nombreuses langues interprétées disponibles aujourd'hui. L'une des philosophies de conception de Python est la lisibilité du code, renforcée en utilisant l'indentation des espaces blancs pour délimiter les blocs de code. Il utilise également une syntaxe remarquablement simple qui permet aux programmeurs d'exprimer des concepts complexes en moins de lignes de code.

La fonction pgcd calcule le plus grand diviseur commun de deux entiers saisis par l'utilisateur :

```
def pgcd(a, b):
         a = abs(a)
        b = abs(b)
                while b:
                        a, b = b, a%b
                retourner un
```

Algorithme 5 Haskell

Haskell est un langage de programmation standardisé, à usage général, purement fonctionnel, avec une sémantique non stricte et un typage statique fort. Être fonctionnel signifie que Haskell traite le calcul comme l'évaluation de mathématiques. les fonctions. A cet égard, la programmation en Haskell se fait avec des expressions ou déclarations au lieu d'instructions - par opposition à la programmation impérative. Haskell propose une évaluation paresseuse, une correspondance de modèles, une compréhension de liste, des classes de types et polymorphisme de type. Une caractéristique principale des fonctions dans Haskell est qu'elles n'ont pas d'effets secondaires, c'est-à-dire que le résultat d'une fonction est déterminé par son entrée et uniquement par son apport. Les fonctions peuvent donc être évaluées dans n'importe quel ordre et seront toujours renvoie le même résultat - à condition que la même entrée soit transmise. C'est un grand avantage de la programmation fonctionnelle qui rend le comportement d'un programme beaucoup plus facile à comprendre et à prévoir.

La fonction plus ajoute 1 à 2 et affiche le résultat :

```
plus :: Int − > Int − > Int
plus = (+)

principal = faire
        soit res = plus 1 2
        putStrLn $ ”1+2 = ” ++ afficher la résolution
```

Les algorithmes présentent plusieurs caractéristiques auxquelles il convient de prêter attention. D'un point de vue onto-logique, un algorithme est une structure syntaxique encodant l'information défini dans le cahier des charges. Comme je le discuterai plus loin dans cette section, pour plusieurs raisons techniques et pratiques, les algorithmes ne peuvent pas coder toutes les informations présentées dans le spécification. Au contraire, certaines informations seront ajoutées, d'autres seront perdues et d'autres sera simplement modifié.

Notons également que les études en informatique utilisent la notion de syntaxe et de sémantique d'une manière assez différente de la linguistique. Pour l'informatique, la syntaxe est l'étude des symboles et de leurs relations au sein d'un système formel ; il comprend généralement une grammaire (c'est-à-dire une séquence de symboles sous forme de formules bien formées), et la théorie de la preuve (c'est-à-dire une séquence de formules bien formées qui sont considérées comme les orèmes). D'autre part, la sémantique est l'étude de la relation entre un système formel, qui est syntaxiquement spécifié, et un domaine sémantique, qui est spécifié par un domaine assurant l'interprétation des symboles du domaine syntaxique. Dans le cas de l'implémentation de l'algorithme sur le calculateur numérique, le domaine sémantique correspond aux états physiques de l'ordinateur lors de l'exécution des instructions programmées dans l'algorithme. J'appelle ces états physiques l'ordinateur processus, et il sera discuté dans la section suivante.

Deux autres propriétés ontologiques intéressantes sont que l'algorithme est abstrait et officiel. Il est abstrait parce qu'il se compose d'une chaîne de symboles sans aucun élément physique. relations causales agissant sur eux : tout comme une structure logico-mathématique, un al-

gorithm est causalement inerte et déconnecté de l'espace-temps. Il est formel parce qu'il suit les lois de la logique qui indiquent comment manipuler systématiquement les symboles28. Ces traits ontologiques offrent plusieurs avantages épistémiques intéressants, dont deux ont particulièrement occupé les philosophes et les informaticiens. Ce sont la corrélation de syntaxe et le transfert de syntaxe. La corrélation syntaxique est la possibilité de maintenir l'équivalence entre deux structures algorithmiques. Le transfert de syntaxe, quant à lui, consiste à modifier l'algorithme pour lui faire exécuter une fonctionnalité différente. Permettez-moi d'en dire plus sur chacun.

La corrélation syntaxique peut être clarifiée par un exemple mathématique. Considérons un système de coordonnées cartésien et un système de coordonnées polaire. Il existe des transformations mathématiques bien établies qui aident à établir leur équivalence. Étant donné les coordonnées cartésiennes (x, y), leurs coordonnées polaires équivalentes sont fixées par $(r,\theta)=(\sqrt{x^2+y^2}, 2\tan^{-1}\frac{y}{x})$. De même, étant donné un ensemble de coordonnées polaires (r,θ), on peut trouver les coordonnées cartésiennes correspondantes sans trop d'effort en utilisant (x, y) = r cosθ,r sinθ).[29]

Une idée similaire peut être utilisée dans les algorithmes. Considérons l'extraction suivante de l'algorithme 6 et de son équivalent dans l'algorithme 7. Nous savons que les deux sont logiquement équivalents car il existe une preuve formelle de cela (donnée dans la table de vérité 2.1).[30]

Algorithme 6 si

··· (α) alors {a} sinon {b} ...

[28] Pour plus de détails, voir les travaux de Donald E. Knuth et Edsger W. Dijkstra dans (Knuth 1974), (Knuth 1973) et (Dijkstra 1974). Les deux auteurs sont reconnus comme des contributeurs majeurs au développement de l'informatique en tant que discipline rigoureuse. Knuth est également l'auteur de The Art of Computer Programming, un chef-d'œuvre en quatre volumes - selon le décompte d'aujourd'hui - qui couvre de nombreux types d'algorithmes de programmation et leur analyse. Dijkstra, à son tour, est l'un des premiers pionniers et promoteurs de l'informatique en tant que discipline universitaire. Il a contribué à jeter les bases de la naissance et du développement du génie logiciel, et ses écrits ont jeté les bases de nombreux domaines de recherche en informatique, en particulier dans la programmation structurée et l'informatique concurrente.

[29] Comme il ressort de l'exemple, la corrélation syntaxique soulève la question « dans quelle mesure deux systèmes d'équations sont-ils équivalents ? Ce n'est pas une question facile à répondre. Dans le cas des systèmes de coordonnées cartésiens et polaires, on pourrait objecter que des restrictions doivent être imposées aux systèmes de coordonnées polaires - c'est-à-dire, pour la fonction tan, le domaine est tous les cept $\pm\frac{\pi}{2}, \pm\frac{3\pi}{2}, \pm\frac{5\pi}{2}, \ldots,$ nombres réels ex où la fonction est indéfinie - et donc ils ne le sont pas. isomorphe. Quelque chose de similaire se produit avec le Lagrangien et l'Hamiltonien, comme je le dis ci-dessous pour les algorithmes.

[30] Certes, les algorithmes sont des structures beaucoup plus complexes que les exemples que j'utilise ici. Il n'est donc pas surprenant que l'équivalence entre algorithmes ne puisse pas être simplement établie par une table de vérité. Des machines mathématiques et informatiques plus complexes sont nécessaires et, en fait, utilisées. Une approche standard consiste à construire une classe d'équivalence d'algorithmes (Blass, Dershowitz et Gurevich 2009). Prenons par exemple un algorithme de tri, dont la formulation est qu'il renvoie une permutation ordonnée d'une liste d'entrée, pour une certaine définition de l'ordre. En montrant qu'une fonction donnée a la propriété de retourner une permutation ordonnée – pour une même définition de l'ordre – alors on pourrait affirmer que les deux algorithmes appartiennent à la même classe. Trouver une équivalence d'algorithme est central pour de nombreuses procédures de vérification logicielles et matérielles.

Algorithme

··· 7 si (non-α) alors {b} sinon {a} ...

Tableau 2.1 Tables de vérité d'équivalence de l'Algorithme 6 et de

α	{a}	{b}
T	T	ø
F	ø	T

l'Algorithme 7. non-α {a} {b}

T	ø	T
F	T	ø

Un avantage épistémique de la corrélation syntaxique est qu'elle augmente le nombre d'implémentations possibles et équivalentes pour une simulation informatique donnée. Considérons par exemple un ensemble d'équations lagrangiennes et hamiltoniennes, car elles sont corrélées dans les systèmes dynamiques. Les chercheurs ont le choix entre l'une ou l'autre formulation en fonction des besoins qui ne sont pas strictement liés à la représentation du système cible (par exemple, la compréhension d'un système d'équations est plus simple que l'autre pour un système cible donné, la performance de la simulation est amélioré avec l'un ou l'autre ensemble d'équations, etc.). De cette façon, les chercheurs ne sont plus coincés avec un ensemble d'équations pour déterminer comment les mettre en œuvre, mais concentrent plutôt leurs efforts et leurs préoccupations sur d'autres aspects de la simulation tels que les performances et la simplicité. Prenons par exemple le calcul rapide suivant. Pour un système avec un espace de configuration de dimension n, les équations hamiltoniennes sont un ensemble de 2n ODE couplées du premier ordre. Les équations lagrangiennes, quant à elles, sont un ensemble de n ODE non couplés du second ordre. Ainsi, implémenter des hamiltoniens sur des langrangiens pourrait donner un réel avantage en termes de performances, d'utilisation de la mémoire et de vitesse de calcul. Quelque chose de très similaire se produit lorsque l'on calcule à la main un système de coordonnées cartésien et polaire.

La deuxième caractéristique épistémique des algorithmes est le transfert de syntaxe. Cela fait référence à l'idée simple qu'en ajoutant – ou en soustrayant – seulement quelques lignes de code dans l'algorithme, les chercheurs sont capables de réutiliser le même code pour différents contextes de représentation. Dans de tels cas, le transfert de syntaxe présuppose des changements minimes dans l'algorithme. Un cas très simple peut être illustré en ajoutant à l'algorithme 3 quelques lignes pour calculer la sommation de 3 nombres, comme le montre l'algorithme 8. Notons que le transfert de syntaxe est une idée fondamentale derrière les modules et les bibliothèques : un code très similaire peut être utilisés dans des contextes différents mais liés.

Le transfert de syntaxe permet alors aux chercheurs de réutiliser leur code existant afin de l'adapter à différents contextes, ainsi que de généraliser le code afin d'inclure plus de résultats, élargissant ainsi - ou rétrécissant - la portée de l'algorithme.

Algorithme 8 JAVA paquet étendu

```
SNP ; importer
java.util.Scanner ;

public class addThreeNumbers private
        static Scanner sc; public static
        void main(String[] args) int Number1, Number2,
                Number3, Sum ; sc = nouveau Scanner
                (System.in);

                System.out.println("Entrez le premier chiffre : "); a =
                sc.nextInt();

                System.out.println("Entrez le deuxième chiffre :  »); b =
                sc.nextInt();

                System.out.println("Entrez le troisième chiffre : "); c =
                sc.nextInt();

                somme = a + b +
                                                        ”
                c ; System.out.println("La somme est =          + somme);
```

La corrélation syntaxique et le transfert syntaxique sont des pratiques courantes chez les chercheurs qui programment leurs propres simulations. Il n'est pas rare de voir comment la même simulation grandit et rétrécit en ajoutant et en éliminant certains modules ainsi qu'en en modifiant d'autres. Cela fait partie de la maintenance et de l'amélioration standard du code.
Or, il est également possible que le transfert de syntaxe rende le code trop volumineux, et donc impossible à maintenir sans un coût élevé. Lorsqu'une telle situation se produit, il est probablement temps pour un nouveau code. Dans des cas comme celui-ci, la corrélation syntaxique joue un rôle important, car de nombreuses fonctions de l'ancien code seront conservées – et correctement modifiées – dans le nouveau code.

Attachés à la corrélation syntaxique et au transfert syntaxique viennent quelques problèmes philosophiques. La plus importante est la question "quand deux algorithmes sont-ils identiques ?" La question émerge dans le contexte où la corrélation syntaxique et le transfert syntaxique présupposent des modifications d'un algorithme original, conduisant à un nouvel algorithme. Dans le cas de la corrélation syntaxique, celle-ci se présente sous la forme d'un nouvel algorithme réalisant les mêmes fonctionnalités que l'ancien algorithme. Dans le cas du transfert de syntaxe, cela se présente sous la forme d'un algorithme modifié – et donc, à proprement parler, nouveau.

Il y a deux réponses généralement acceptées à cette question. Soit deux algorithmes sont logiquement équivalents, c'est-à-dire que les deux algorithmes sont structurellement et formellement

similaires31 ou qu'ils sont équivalents sur le plan comportemental, c'est-à-dire que les deux algorithmes se comportent une mode similaire. Permettez-moi de discuter brièvement de ces deux approches.

L'équivalence logique est l'idée que deux algorithmes sont structurellement similaires, et qu'une telle équivalence peut être démontrée par des moyens formels. J'illustre un très simple équivalence logique en utilisant l'algorithme 6 et l'algorithme 7, puisque les deux sont formellement isomorphes les uns aux autres - la preuve, encore une fois, est dans le tableau 2.1. Procédures formelles de tout type – comme une table de vérité – sont de bons garants de la similarité structurelle.32 Ainsi, les si ... alors le conditionnel dans l'algorithme 6 est structurellement équivalent au conditionnel dans algorithme 7.

Malheureusement, l'équivalence logique n'est pas toujours réalisable en raison des contraintes pratiques ainsi que des contraintes théoriques. Des exemples de contraintes pratiques incluent des cas d'algorithmes qui sont humainement impossibles à vérifier formellement. Un autre exemple est formel procédures trop consommatrices de temps et de ressources. Exemples de théorie les contraintes incluent les cas où le langage de programmation codé dans la simulation fait référence à des entités, des relations, des opérations, etc., dont la procédure de vérification de la similarité structurelle est incapable de tenir compte.

Pour faire face à ces contraintes, le milieu universitaire et l'industrie ont uni leurs forces et a créé une multitude d'outils qui automatisent le processus de vérification et de vérification. Bien que il existe plusieurs vérificateurs de modèles et sémantiques disponibles, un exemple puissant utilisé dans la vérification du logiciel est ACL2. Une logique de calcul pour l'applicatif Common Lisp (ACL2) a été explicitement conçu pour prendre en charge le raisonnement automatisé et ainsi aider à la reconstruction des classes d'équivalence des algorithmes.

L'équivalence comportementale, quant à elle, consiste à s'assurer que les deux les algorithmes se comportent de manière similaire (par exemple, en produisant les mêmes résultats33). Maintenant, bien que l'équivalence comportementale semble plus simple à réaliser que l'équivalence structurelle, elle comporte ses propres problèmes. Par exemple, on craint que l'équivalence comportementale est fondée sur des principes inductifs. Cela signifie qu'un ne pourrait garantir l'équivalence que jusqu'au temps t, lorsque les deux algorithmes se comportent de la même manière. Mais rien ne garantit qu'au temps t +1 le comportement des algorithmes sera toujours le même. L'équivalence comportementale ne peut être garantie que jusqu'au moment où les deux algorithmes sont comparés. En fait, on pourrait faire le cas que deux algorithmes sont

---

31 L'isomorphisme serait la meilleure option ici, puisque c'est le seul -morphisme qui pourrait justifier équivalence totale entre algorithmes. Cependant, les morphismes alternatifs sont également discutés dans le littérature, par exemple, dans les travaux de (Blass, Dershowitz et Gurevich 2009) et (Blass et Gourevitch 2003).

32 Pour les lecteurs intéressés par des discussions philosophiques approfondies, il est indispensable de clarifier la notion de « similarité structurelle ». Ici, je suppose que l'on peut objectivement décider quand un algorithme est structurellement semblable à un autre. La littérature sur la similarité et, plus généralement, sur la représentation théorique est assez vaste. Une suggestion est pour commencer (Humphreys et Imbert 2012).

33 Ceci sous une interprétation donnée de "les mêmes résultats", sinon nous posons la question lorsque deux ensembles de résultats sont identiques. À cette fin, nous pourrions sélectionner des procédures mathématiques et algorithmiques externes aux deux algorithmes comparés qui établissent une similitude acceptable parmi les résultats.

comportementalement équivalent uniquement pour cette exécution des algorithmes. D'autres exécutions pourrait montrer une divergence de comportement.34

Un dernier problème découlant de l'équivalence comportementale est qu'elle pourrait masquer une équivalence logique. Cela signifie que deux algorithmes ont des comportements divergents bien qu'ils sont structurellement équivalents. Un exemple est un algorithme qui implémente un cartésien ensemble de coordonnées alors qu'un autre implémente des coordonnées polaires. Les deux algorithmes sont structurellement équivalents, mais dissemblables sur le plan comportemental.35 Dans ce genre de cas, il reste la question du type d'équivalence qui doit prévaloir.

Ce sont quelques-unes des discussions standard trouvées dans la philosophie de l'informatique science. Ici, je n'ai fait qu'effleurer la surface, et beaucoup plus de possibilités et de besoins à dire. La leçon que je voudrais tirer, cependant, est que la corrélation syntaxique et le transfert syntaxique ont un prix. Si les chercheurs sont prêts à payer ce prix - et à quel point ce prix est-il réellement élevé - est une question qui dépend de plusieurs variables, telles que les intérêts du chercheur, les ressources disponibles et les ressources réelles. urgence de trouver une solution. Pour certaines situations, des solutions prêtes à l'emploi exister; pour d'autres, l'expérience du chercheur reste la monnaie la plus précieuse.

Jusqu'à présent, nous avons discuté des algorithmes avec leurs conséquences philosophiques. Nous devons encore dire quelque chose sur le lien entre la spécification et l'algorithme.

Idéalement, la spécification et l'algorithme devraient être étroitement liés, c'est-à-dire que La spécification doit être complètement interprétée comme une structure algorithmique. En réalité, c'est rarement le cas, principalement parce que la spécification fait un usage intensif de langage naturel alors que dans l'algorithme, chaque terme doit être interprété littéralement. Un exemple typique est l'utilisation de métaphores et d'analogies. De nombreux modèles scientifiques et techniques incluent des termes qui n'ont pas d'interprétation spécifique, comme "trou noir" ou « mécanisme » (Bailer-Jones 2009).36 Des métaphores et des analogies sont ensuite utilisées pour l'écart introduit par ces termes qui n'ont pas d'interprétation littérale. En faisant cela, les métaphores et les analogies inspirent une sorte de réponse créative chez les utilisateurs du modèle auquel ne peut rivaliser le langage littéral. Cependant, si la même métaphore termes sont implémentés dans un algorithme, ils nécessitent une interprétation littérale sinon ils ne peuvent pas être calculés.

Bien sûr, pour faciliter l'interprétation de la spécification en algorithme, les chercheurs se sont à nouveau appuyés sur l'automatisation apportée par les ordinateurs. Pour cela, il existe une multitude de langages spécialisés qui formalisent la spécification,

---

34 De nombreux chercheurs hésitent à accepter comme une véritable préoccupation que le code informatique puisse être différent selon le moment de l'exécution. Le philosophe James Fetzer (Fetzer 1988) soulevé cette question une fois dans le contexte de la validation du programme. En réponse, de nombreux informaticiens et ingénieurs ont objecté qu'il savait peu de choses sur la façon dont les logiciels et le matériel informatique travail. Bien sûr, les objections n'étaient pas simplement ad hominem, mais contenaient de bonnes raisons de rejeter Position de Fetzer. En tout cas, Fetzer a soulevé une véritable question philosophique et il convient de la traiter en tant que tel.

35 Notons que ce raisonnement repose sur la notion de « comportement ». Si par là on prend simplement "exactement les mêmes résultats", les deux algorithmes donnent clairement des résultats différents.

36 Nous devons être prudents ici car, dans des cas comme les neurosciences, des termes tels que "mécanisme" avoir une définition à part entière (par exemple, (Machamer, Darden et Craver 2000) et (Craver 2001))

facilitant ainsi sa programmation dans un algorithme. Le langage de spécification algébrique commun (CASL), la méthode de développement de Vienne (VDM), la spécification comportementale abstraite (ABS) et la notation Z ne sont que quelques exemples.37 Modèle la vérification est également utile car elle teste automatiquement si un algorithme répond aux spécification requise, et donc il aide également à son interprétation. En bref, l'interprétation d'une spécification dans un algorithme a une longue tradition en mathématiques, en logique et en informatique, et elle ne représente pas vraiment une approche conceptuelle. problème ici.

J'ai également mentionné que la spécification comprend des éléments non formels, tels que des connaissances d'experts et des décisions de conception qui ne peuvent pas être formellement interprétées. Cependant, ces éléments non formels doivent également être inclus – et inclus correctement – dans le algorithme, sinon ils ne feront pas partie du calcul du modèle. Imaginer par exemple la spécification d'une simulation d'un système de vote. Pour cette simulation pour réussir, les modules statistiques sont mis en œuvre de manière à donner une répartition raisonnable de la population électorale. Lors de la phase de spécification, les chercheurs décident de donner plus de pertinence statistique à des variables telles que le sexe, le genre, et la santé par rapport à d'autres variables, comme l'éducation et le revenu. Si cette décision de conception n'est pas correctement programmé dans le module statistique, la simulation ne reflètent jamais la valeur de ces variables, bien qu'elles figurent dans la spécification.

Cet exemple vise à montrer qu'un algorithme doit être capable d'interpréter des éléments formels et non formels inclus dans la spécification, en particulier parce qu'il n'y a pas de méthodes formelles pour interpréter les connaissances d'expertise, les expériences passées, etc.

Où nous mène cette discussion dans la carte méthodologique des simulations informatiques ? D'une part, nous comprenons mieux la nature des simulations informatiques en tant qu'unités d'analyse. D'autre part, nous avons une meilleure compréhension de la méthodologie des spécifications, des algorithmes et de leur relation.

Avant de continuer, permettez-moi de clarifier une partie de la terminologie que j'ai utilisée. Appelez modèle informatique l'ensemble du processus de spécification et de programmation d'un système informatique donné. Si le but d'un tel modèle est de simuler un système cible, alors appelons-le le modèle de simulation. Les seules différences visibles entre ces deux est que le premier est une version généralisée du second. Différences plus importantes émergeront lors de la discussion des questions épistémologiques et pratiques dans ce qui suit chapitres. De plus, en implémentant un modèle informatique sur l'ordinateur physique nous obtenons un processus informatique (à discuter ensuite). Suivant la même structure comme précédemment, si le modèle mis en oeuvre est un modèle de simulation, alors on dispose d'un calculateur simulation.

---

[37] Pour CASL, voir par exemple (Bidoit et Mosses 2004). Pour VDM, voir par exemple (Bjorner et Henson 2007). Pour l'ABS, voir http://abs-models.org/concept/. Et pour la notation Z, voir pour exemple (Spivey 2001).

2.2.1.3 Processus informatiques

Dans les sections précédentes, j'ai utilisé la notion de spécification de Cantwell Smith comme point de départ de nos études. Il est maintenant temps de compléter son idée par l'analyse des processus informatiques.38

Selon l'auteur, un programme informatique est l'ensemble des instructions qui s'exécutent sur l'ordinateur physique. Une telle caractérisation diffère grandement de la notion de spécification, telle que discutée précédemment par Cantwell Smith. Alors qu'un programme informatique est un processus causal qui se déroule sur l'ordinateur physique, la spécification est une entité abstraite et, dans une certaine mesure, formelle. De plus, Cantwell Smith souligne que « le programme doit dire comment le comportement doit être atteint, généralement étape par étape (et souvent avec des détails atroces). La spécification, cependant, est moins contrainte : tout ce qu'elle a à faire est de spécifier ce que serait un comportement approprié, indépendamment de la manière dont il est accompli » (Cantwell Smith 1985, 22. Emphasis original.). Ainsi comprises, les spécifications sont déclaratives au sens où elles énoncent dans un langage donné – naturel et formel – comment le système évolue dans le temps. A cet égard, les spécifications désignent des langages de haut niveau pour résoudre des problèmes sans exiger explicitement la procédure exacte à suivre. Les programmes informatiques, quant à eux, sont procéduraux, car ils déterminent par étapes comment l'ordinateur physique doit se comporter.

Pour illustrer la différence entre la spécification et le programme informatique, reprenons la description du système de distribution de lait. Cette simulation précise qu'un camion de livraison de lait doit effectuer une livraison dans chaque magasin en parcourant la distance la plus courte possible au total. Selon Cantwell Smith, il s'agit d'une description de ce qui doit se passer, mais pas de la manière dont cela se passera. Il est de la responsabilité du programme informatique de montrer comment la livraison de lait se déroule réellement : "conduisez quatre pâtés de maisons vers le nord, tournez à droite, arrêtez-vous à l'épicerie Gregory au coin, déposez le lait, puis conduisez 17 pâtés de maisons vers le nord-est, [. ..] » (22).

Bien qu'elle soit correcte à bien des égards, la définition de « programme informatique » de Cantwell Smith n'est pas convaincante. La principale préoccupation ici est qu'elle ne parvient pas à saisir la différence entre une procédure par étapes comprise comme des formules syntaxiques (c'est-à-dire l'algorithme) et une procédure par étapes qui place la machine physique dans les états causaux appropriés (c'est-à-dire processus informatique).39 Comme nous le verrons bientôt, il ne s'agit pas d'une différence anodine, mais elle est au cœur de nombreuses discussions sur la vérification des logiciels. En particulier, ne pas tenir compte de cette distinction est à la base de la confusion entre la description par le chercheur du comportement du système de distribution de lait et les étapes réelles suivies par l'ordinateur.

---

38 Je n'aborde pas la question de l'architecture de calcul sur laquelle le logiciel informatique est exécuté. Cependant, il devrait être clair d'après le contexte que nous nous intéressons aux ordinateurs à base de silicium - par opposition aux ordinateurs quantiques ou aux ordinateurs biologiques. Il faut aussi dire que d'autres architectures représentent un ensemble de problèmes différent (Berekovic, Simopoulos et Wong 2008)
(Rojas et Hashagen 2000).

39 Pour compléter les étapes allant de la spécification au programme informatique, il convient également d'ajouter une foule d'étapes intermédiaires, telles que la compilation de l'algorithme, l'allocation de la mémoire, le stockage de masse, etc. Comme ces étapes intermédiaires ne constituent pas des unités d'analyse pour les simulations informatiques, il n'est pas nécessaire de les détailler.

La notion de programme informatique40 doit donc être subdivisée en deux : la
l'algorithme et le processus informatique.41 Puisque nous avons discuté des algorithmes dans certains
détail dans la section précédente, il est temps d'aborder la notion de processus informatique
et sa relation avec les algorithmes.

Permettez-moi de commencer par cette dernière question, car elle aide à donner un sens à la notion
du processus informatique lui-même. Comme tout chercheur qui a programmé au moins une fois dans
leur vie le sait, pour implémenter un algorithme sur l'ordinateur, il faut
compilez-le d'abord. La compilation consiste essentiellement en une mise en correspondance d'un domaine
interprété (c'est-à-dire l'algorithme) dans un domaine d'interprétation (c'est-à-dire le processus informatique)42.
En d'autres termes, un algorithme est implémenté comme un processus physique car l'ordinateur est capable
d'interpréter et d'exécuter l'algorithme de la bonne manière. Maintenant, comment est-ce
possible?

À la base, chaque matériel informatique est constitué de microélectronique construite à partir de milliards
de portes logiques. Ces portes logiques sont l'implémentation physique de la logique
opérateurs 'et', 'ou' et 'non' qui, lorsqu'ils sont combinés, suffisent à interpréter toute logique
et les opérations arithmétiques – et par conséquent, tout le reste.43 À ce niveau de description, tous les
ordinateurs sont fondamentalement les mêmes, peut-être à l'exception du
nombre de portes logiques utilisées. Cependant, c'est jusqu'à quel point l'identité entre
les ordinateurs vont, car aux niveaux supérieurs, tous les ordinateurs ne partagent pas la même architecture.

Un schéma ascendant relie ces portes logiques au processus informatique en
impliquant une cascade d'instructions machine complexes et de langages qui "parlent" à
l'un l'autre. Cela commence par Microcode,44 utilisé comme ensemble d'instructions au niveau matériel
capable d'implémenter des instructions de code machine de niveau supérieur, au compilateur, chargé de
convertir les instructions de l'algorithme en un code machine, de sorte que
ils peuvent être lus et exécutés comme un processus informatique. Ainsi compris, un ordinateur
le processus s'exécute sur un ordinateur physique car il existe plusieurs couches d'interpréteurs
qui traduisent un ensemble d'instructions en langage machine approprié.

Illustrons ces points par un cas simplifié de l'opération mathématique
2+2 écrit en langage C. Considérons l'algorithme 9.

---

40 Bien que je soutienne cette distinction, je garde la notion de « programme informatique » comme un moyen compact
pour faire référence à l'algorithme et au processus informatique à la fois.

41 Le concept de « processus informatique » est, à son tour, très ambigu, car il peut être utilisé pour désigner
à (i) des encodages d'algorithmes, (ii) des encodages d'algorithmes compilables, (iii) des encodages
d'algorithmes pouvant être compilés et exécutés par une machine. James H. Moor a suggéré
à cinq interprétations différentes (Moor 1988). Voir aussi (Fetzer 1988, 1058). Mon interprétation est
semblable au troisième.

42 Sur ce point on pourrait consulter les travaux de William Rapaport dans (William J. Rapaport 1999)
et (William J Rapaport 2005).

43 Il existe des langages spécifiques – appelés « langages de description matérielle » – tels que le VHDL
et Verilog qui facilitent la construction de ces portes logiques dans les microcircuits physiques (Cohn 1989)
(Ciletti 2010). Ces langages de description de matériel sont essentiellement des langages de programmation pour
l'architecture matérielle.

44 Le microcode a été développé à l'origine pour remplacer le câblage des instructions du processeur.
De cette manière, le comportement et la programmation du processeur sont remplacés par des routines microprogrammées
plutôt que par des circuits dédiés.

Algorithme 9 Un algorithme simple pour l'opération 2+2 écrit en langage C

```
void main()
{
retour(2+2)
}
```

En code binaire, le nombre 2 est représenté par '00000010, alors que l'opération plus est effectuée, par exemple, par un additionneur Ripple-carry. Le compilateur convertit alors ces instructions en code machine prêt à être exécuté sur un ordinateur physique. Une fois le processus informatique se termine, la solution est affichée sur le moniteur d'écran, dans ce cas 4, qui est 00000100 en code binaire.

À partir de ces considérations, nous avons maintenant suffisamment d'éléments pour présenter quelques questions philosophiques clés. Considérez la question suivante : si l'ordinateur les processus sont la réalisation physique d'algorithmes, qui sont abstraits – et parfois formels – comment concevoir la relation entre l'un et l'autre ?

Pour beaucoup, le seul but des algorithmes est de prescrire les règles que l'ordinateur les processus doivent suivre sur l'ordinateur physique. Notons que comprendre le relation entre les algorithmes et les processus informatiques de cette manière n'implique pas une revenons à la notion de « programme informatique » de Cantwell Smith, car les algorithmes et les processus informatiques sont encore deux entités distinctes. Il y a plusieurs raisons qui jusqu'à cette réclamation. Premièrement, les algorithmes et les processus informatiques sont ontologiquement différents. Alors que les algorithmes sont des entités abstraites, les processus informatiques sont causals dans un sens physique simple. De plus, alors que les algorithmes sont cognitivement accessibles (c'est-à-dire, les chercheurs peuvent comprendre et, dans une certaine mesure, suivre les instructions données dans l'algorithme), les processus informatiques sont cognitivement opaques. Une troisième raison est que une méthodologie du logiciel informatique nécessite de reconnaître l'existence d'algorithmes comme intermédiaires entre les spécifications et les processus informatiques.[45] Ainsi compris, les algorithmes et les processus informatiques sont ontologiquement dissemblables, mais ils sont épistémiquement équivalents. Cela signifie que les informations codées dans l'algorithme est physiquement instancié par le processus informatique, et considéré épistémiquement à égalité. L'exemple est l'addition de 2 + 2 programmée dans l'algorithme 9. En supposant une fonctionnalité appropriée du compilateur ainsi que de l'ordinateur physique, le résultat du calcul de cet algorithme est l'addition réelle, c'est-à-dire 4.

Accepter l'équivalence épistémique entre algorithmes et processus informatiques a une certaine parenté avec le débat sur la vérification dans les logiciels informatiques. c'est-à-dire étant donné la vérification formelle d'un algorithme, le chercheur pourrait-il avoir confiance que le processus informatique se comporte également de la manière prévue ? Répondre positivement à cette question signifie qu'il existe une méthode - formelle - qui garantit que les processus informatiques se comportent comme prévu dans les spécifications et tel que programmé dans les algorithmes. Répondre par la négative, en revanche, soulève la question de savoir sur quelles bases les chercheurs ont-ils

[45] Il y a eu des discussions sur la question de savoir si les spécifications - et les algorithmes - sont des exécutables, c'est-à-dire si un processus informatique peut les calculer fidèlement. Voir (Fuchs 1992).

faire confiance aux résultats des processus de calcul. Permettez-moi maintenant de présenter plus en détail le débat sur la

vérification46. Les informaticiens et philosophes CAR Hoare (Hoare 1999) et Edsger J. Dijkstra (Dijkstra 1974) pense que les logiciels informatiques sont de nature mathématique. Dans ce contexte, les programmes informatiques peuvent être formellement vérifiés, c'est-à-dire que l'exactitude des algorithmes peut être prouvée - ou réfutée - par rapport à certaines propriétés formelles de la spécification, un peu comme une preuve mathématique. La vraie différence avec les mathématiques est que la vérification des logiciels informatiques requiert sa propre syntaxe. À cette fin, Hoare a créé ses triplets, qui consistent en un système formel - états initial, intermédiaire et final - qui obéissent strictement à un ensemble de règles logiques. Le triplet Hoare a la forme : {P} C {Q} où P et Q sont des assertions – respectivement précondition et postcondition – et C est une commande. Lorsque la précondition est remplie, la commande établit la postcondition. Les assertions sont des formules dans la logique des prédicats avec un ensemble spécifique de règles. L'une de ces règles est le schéma d'axiome d'instruction vide : ; un autre est le schéma de l'axiome d'affectation : et {P} Skip{P} $\frac{}{\{P[E/x]\}x := E\{P\}}$, ainsi de suite (Hoare 1971).

Pour Hoare et Dijkstra, les programmes informatiques sont "fiables et obéissants" (Dijkstra 1974, 608) car ils se comportent exactement comme indiqué par la spécification. La charge de travail repose donc sur l'algorithme et sur la manière dont il interprète correctement la spécification.47 Une fois qu'un algorithme est formellement vérifié, les résultats du calcul seront tels que prévus dans la spécification.

Le point de vue opposé part du fait que les algorithmes sont ontologiquement différents des processus informatiques et conclut qu'ils doivent également être épistémiquement différents. Alors que les premières sont en effet des expressions mathématiques adaptées à la vérification mathématique – ou logique –, l'exactitude de la seconde ne peut être abordée qu'en utilisant des méthodes empiriques. C'est l'affirmation du philosophe James Fetzer (Fetzer 1988), qui estime que les processus informatiques pourraient contenir une interprétation de nature causale et donc sortir du champ des méthodes logiques et mathématiques.
Cela signifie que, lorsque l'algorithme est implémenté sur la machine physique où les facteurs causaux sont en jeu, l'ensemble du programme informatique devient en quelque sorte « causal ». Son argument se termine par l'affirmation que la vérification formelle est donc impossible en informatique, et que les méthodes de validation comme les tests doivent se voir accorder une place plus importante.48

---

46 Il existe un important complexe industriel et académique dédié aux méthodes de vérification et de validation. Ici, je m'intéresse à quelques promoteurs et détracteurs originaux. Une discussion complète peut être trouvée dans (Colburn, Fetzer et Rankin 2012).

47 Dijkstra est très préoccupé par les programmeurs et leur éducation. A l'époque, le cursus sur la vérification formelle était presque inexistante.

48 Je simplifie bien sûr le débat. L'argument de Fetzer est certainement plus élaboré et mérite d'être étudié en lui-même. Pour commencer, il fait une série de différences concernant la machine sur laquelle un algorithme et un processus sont exécutés (par exemple, des machines abstraites, des machines physiques), différences que j'ai ignorées ici. De plus, en plaidant pour les possibilités de vérification formelle, Fetzer fait une distinction essentielle entre les mathématiques pures et les mathématiques appliquées, ces dernières nécessitant l'interprétation d'un système physique (Fetzer 1988, 1059). Comme il le dit si la fonction d'un programme est de satisfaire les contraintes imposées par une machine abstraite pour laquelle il existe une interprétation prévue par rapport à un système physique, alors le comportement de ce système ne peut pas

De nombreux informaticiens et philosophes se sont opposés avec véhémence à l'argument de Fetzer, l'accusant de ne pas comprendre les bases de la théorie de l'informatique[49]. moyen - notre cerveau, une calculatrice, un boulier - est toujours nécessaire pour le calcul et la preuve. L'argument de Fetzer ne tient donc que si nous reconnaissons qu'il existe une différence qualitative entre le support physique utilisé pour implémenter un algorithme (c'est-à-dire l'ordinateur physique) et le support physique utilisé par un mathématicien faisant une preuve (c'est-à-dire notre cerveau) (Blanco et Garcia 2011).

Malgré ces objections, je crois que Fetzer a raison sur deux points. Premièrement, il a raison de dire que les algorithmes et les processus informatiques ne peuvent être conceptualisés comme la même entité mathématique, mais nécessitent plutôt un traitement différent. J'ai déjà mentionné ce point plus tôt, lorsque j'ai séparé ontologiquement les algorithmes en tant qu'entités abstraites et formelles des processus informatiques en tant que causalement liés. Il a cependant tort de penser que ce sont là des motifs de rejet de la vérification formelle et de l'équivalence épistémique. De nos jours, de nombreux algorithmes sont formellement vérifiés (par exemple, les protocoles cryptographiques et les protocoles de sécurité) et, lorsqu'ils sont exécutés sur l'ordinateur, les processus informatiques sont traités comme épistémiquement équivalents à ces algorithmes.

Deuxièmement, Fetzer souligne le rôle de la validation - ou des tests - comme étant plus pertinent qu'on ne le pensait à l'origine. Je suis fondamentalement d'accord avec Fetzer sur ce point. Les méthodes de validation ne peuvent être absorbées et remplacées par une vérification formelle, même lorsque celle-ci est possible. En fait, dans le contexte des simulations informatiques, une combinaison de méthodes de vérification et de validation est largement utilisée pour asseoir l'exactitude des résultats. Ces questions font l'objet du chapitre 4. Comme nous le verrons là-bas, les méthodes de vérification se focalisent sur la relation modèle-spécification, et vont donc au-delà de la vérification « formelle » ; les méthodes de validation, quant à elles, se concentrent sur la relation modèle-monde, et sont donc fondamentales pour s'assurer que notre modèle représente fidèlement le système cible.

Enfin, il y a une autre hypothèse que nous devons mentionner. Pour nos besoins actuels, je suppose qu'il n'y a pas d'erreurs de calcul ou d'artefacts mathématiques de quelque nature que ce soit que le processus informatique introduit dans les résultats. Ce présupposé est philosophiquement inoffensif et techniquement réalisable. Il s'ensuit donc que les équations programmées dans l'algorithme sont résolues de manière fiable par le processus informatique et que les résultats concernent la spécification et l'algorithme.

La figure 2.1 résume les trois unités de logiciel informatique et leurs connexions. Au niveau supérieur, il y a la spécification, où les décisions pour le logiciel informatique sont prises et intégrées complètement. L'algorithme est l'ensemble d'instructions qui interprète la spécification et prescrit le comportement de la machine. Comme mentionné précédemment, j'appelle la paire <spécification, algorithme> le modèle de simulation.

être soumis à une vérification absolue concluante, mais nécessite plutôt une enquête inductive empirique pour étayer des vérifications relatives non concluantes. (Fetzer 1988, 1059-1060)

[49] Une série de lettres d'indignation ont été envoyées au rédacteur en chef des communications de l'ACM, Robert L. Aslzedzurst, immédiatement après la publication de l'article de Fetzer. Heureusement, les lettres ont été publiées avec une réponse de l'auteur, ce qui montre non seulement l'esprit démocratique de l'éditeur, mais la réaction fascinante de la communauté face à cette question.

Un modèle de simulation englobe donc toutes les informations pertinentes sur le système cible et, à cet égard, constitue l'unité la plus transparente de la simulation informatique. Enfin, il y a le processus informatique en tant qu'implémentation sémantique de l'algorithme sur l'ordinateur physique.

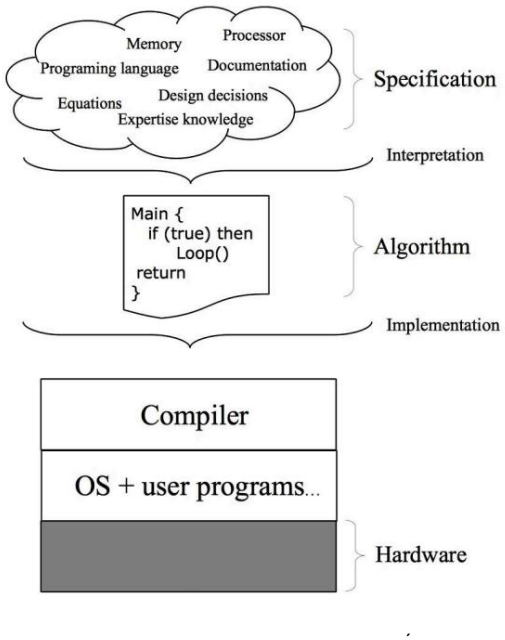


Fig. 2.1 Un schéma général des logiciels informatiques. Imprimé en (Duran 2014)

2.3 Remarques finales

Ce chapitre s'est concentré sur les simulations informatiques en tant que nouveaux types d'unités d'analyse. En identifiant et en reconstruisant la spécification, l'algorithme et le processus informatique, nous sommes en mesure de mieux comprendre la nature des logiciels informatiques en général et des simulations informatiques en particulier.

La leçon à retenir est que les simulations informatiques sont des unités d'analyse très complexes, celles qui associent les étapes de conception et de décision dans la spécification et l'algorithme, avec la mise en œuvre sur l'ordinateur physique. Chacune de ces étapes fait émerger une série de préoccupations, qu'elles soient liées aux aspects sociaux des simulations informatiques, à leurs possibilités techniques ou à des questions philosophiques profondes. Bien que ce chapitre ait abordé bon nombre de ces préoccupations, il est possible et nécessaire d'en dire beaucoup plus, en particulier en mettant l'accent sur la nature des simulations informatiques.

Les résultats obtenus ici nous suivront tout au long de l'ouvrage. Pour exemple, au chapitre 5.1.1, je défends les possibilités des simulations informatiques de donner une véritable compréhension du monde en l'expliquant. Comme nous le verrons là-bas, les spécifications et les algorithmes qui constituent le modèle de simulation permettent la force explicative des simulations informatiques. De même, j'utiliserai ces unités d'analyse explicitement lors de notre discussion dans la section 4.3.1, lorsque je discute des erreurs dans le logiciel et le matériel, et implicitement au chapitre 6 lorsque je fais référence aux simulations informatiques comme un nouveau paradigme de la science.

Les références

Alliance, Observatoire Virtuel International. 2018. Consulté le 26 février 2018. ht tp://www.ivoa.net.

Atkinson, Kendall E., Weimin Han et David E. Stewart. 2009. Solution numérique des équations différentielles ordinaires. John Wiley et fils.

Bailer-Jones, Daniela. 2009. Modèles scientifiques en philosophie des sciences. Université de Pittsburgh Press.

Berekovic, Mladen, Nikitas Simopoulos et Stephan Wong, éds. 2008. Intégré Systèmes informatiques : architectures, modélisation et simulation. Springer.

Bidoit, Michel, et Peter D. Mosses. 2004. Manuel de l'utilisateur CASL : Introduction à Utilisation du langage de spécification algébrique commun. Springer.

Bjorner, Dines et Martin C. Henson. 2007. Logiques des langages de spécification. Springer.

Blanco, Javier et Pío García. 2011. « Une erreur catégoriale dans le débat sur la vérification formelle. Dans The computational turn : Past, presents, futures ?, édité par C. Ess et R. Hagengruber. Mv-Wissenschaft, Munster, ¨ Université d'Arhus.

Blass, Andreas, Nachum Dershowitz et Yuri Gurevich. 2009. "Quand sont deux Algorithmes les mêmes ? » Le Bulletin de Logique Symbolique, no. 250 : 145–168.

Blass, Andreas et Youri Gourevitch. 2003. "Algorithmes: Une quête de définitions absolues." Dans Bulletin of European Associationfor Theoretical Computer Science, 195–225. Octobre. https : / / www . Microsoft. com / fr - us / recherche/publication/164-algorithmes-quête- définitions absolues/.

Boucher, James. 2008. "Modélisation des circuits neuronaux : l'avenir est prometteur." Le Lancet Neurology 7 (5): 382–383.

Cantwell Smith, Brian. 1985. "Les limites de l'exactitude." ACM SIGCAS Computers and Society 14 (1): 18–26.

Chabert, Jean-Claude, éd. 1994. Une histoire des algorithmes. Du caillou à la micropuce. Springer.

Ciletti, Michael D. 2010. Conception numérique avancée avec Verilog HDL. Prentice Hall.

Cohn, Avra. 1989. "La notion de preuve dans la vérification du matériel." Journal d'Au tomate Raisonnement 5, no. 2 (juin): 127–139.

Colburn, Timothy, James H Fetzer et RL Rankin. 2012. Vérification de programme : problèmes fondamentaux en informatique. Vol. 14. Springer Science & Business Media.

Collins, Harry et Robert Evans. 2007. Repenser l'expertise. Université de Chicago Presse.

Copeland, B Jack. 1996. "Qu'est-ce que le calcul?" Synthèse 108 (3): 335–359.

Craver, CF F. 2001. "Fonctions de rôle, mécanismes et hiérarchie." Philosophie des sciences 68 (1): 53–74.

Dijkstra, Edsger W. 1974. "La programmation en tant que discipline de nature mathématique." Mensuel mathématique américain 81 (6): 608–612.

Duran, Juan M. 2014. « Expliquer les phénomènes simulés : une défense du pouvoir épistémique des simulations informatiques ». Thèse de doctorat, Universitat Stuttgart.

Fetzer, James H. 1988. "Vérification de programme : l'idée même." Communication de l'ACM 37 (9): 1048-1063.

Feynman, Richard P. 2001. Que vous importe ce que les autres pensent ? WW Norton & Entreprise.

Fuchs, Norbert E. 1992. "Les spécifications sont (de préférence) exécutables." Journal de génie logiciel 7 (5): 323–334. http : // portail . acm. org/ citation.cfm?id=146587.

Gelfert, Axel. 2016. Comment faire de la science avec des modèles. Mémoires Springer en philosophie . Springer. ISBN : 978-3-319-27952-7 978-3-319-27954-1, consulté le 23 août 2016.

Gould, Harvey, Jan Tobochnik et Wolfgang Christian. 2007. Une introduction aux méthodes de simulation par ordinateur. Applications aux systèmes physiques. Pearson Ad dison Wesley.

Hartmann, Stéphane. 1999. "Modèles et histoires en physique des hadrons." Dans Models as Mediators: Perspectives on Natural and Social Science, édité par Mary S. Morgan et Margaret Morrison. La presse de l'Universite de Cambridge.

Hill, Robin K. 2013. "Qu'est-ce qu'un algorithme est et n'est pas." Communication de la ACM 56 (6): 8–9.

————. 2015. « Qu'est-ce qu'un algorithme ? » Philosophie et technologie 29 (1): 35–59.

Hoare, CAR 1971. Informatique.

————. 1999. Une théorie de la programmation: dénotationnelle, algébrique et opérationnelle sémantique. Rapport technique. Recherche Microsoft.

Humphreys, Paul et Cyrille Imbert, éd. 2012. Modèles, simulations et représentations. Routledge Studies en philosophie des sciences. Routledge. ISBN : 978-0-415-89196-7 978-0-203-80841-2.

Milieu interstellaire des galaxies isolées, AMIGA. 2018. Accédé le 26 février 2018. http://amiga.iaa.es/p/1-homepage.htm.

Kennedy, Ashley Graham. 2012. "Une vision non représentationaliste de l'explication du modèle." Études en histoire et philosophie des sciences Partie A 43 (2): 326–332.

Knuth, Donald E. 1973. L'art de la programmation informatique. Addison-Wesley.

————. 1974. "L'informatique et sa relation avec les mathématiques." L'Américain Mathematical Monthly 81 (4): 323–343.

Knuuttila, Tarja. 2005. Modèles comme artefacts épistémiques : vers un non-représentationnaliste Compte de représentation scientifique. Département de philosophie, Université de Helsinki. ISBN : 952-10-2797-5.

Lenhard, Johannes. 2007. « Simulation informatique : la coopération entre l'expérimentation et la modélisation ». Philosophie des sciences 74: 176–194.

Machamer, Peter, Lindley Darden et Carl F. Craver. 2000. "Penser à Mech anismes. Philosophie des sciences 67 (1): 1–25.

Meijers, Anthony, éd. 2009. Philosophie de la technologie et des sciences de l'ingénieur. Elsevier.

Moor, James H. 1988. "Le sophisme de la pseudoréalisation et l'argument de la salle chinoise." Dans Aspects of Artificial Intelligence, édité par James Fetzer. Springer, 1er janvier. ISBN : 978-1-55608-038-8. doi:10.1007/978-94-009-2699-8_2. http://dx.doi.org/10.1007/978-94-009-2699-8_2.

Morgan, Mary S. et Margaret Morrison, éd. 1999. Modèles comme médiateurs : Perspectives sur les sciences naturelles et sociales. La presse de l'Universite de Cambridge.

Morrisson, Margaret. 2009. "Modèles, mesure et simulation informatique : Changer le visage de l'expérimentation. Études philosophiques 143 (1): 33–57.

————. 2015. Reconstruire la réalité. Modèles, mathématiques et simulations. Ox ford University Press.

Muller, Tibor et Harmund Muller. 2003. Modélisation en sciences naturelles. Springer.

Oberkampf, William L, Timothy G Trucano et Charles Hirsch. 2003. Verifica tion, Validation, and Predictive Capability in Computational Engineering and La physique. Laboratoires nationaux de Sandia.

Pfleeger, Shari Lawrence et Joanne M. Atlee. 2009. Génie logiciel : Théorie et pratique. Prentice Hall.

Piccinini, Gualtiero. 2007. "Mécanismes informatiques". Philosophie des sciences 74: 501–526.

————. 2008. "Calcul sans représentation." Études philosophiques 137 (2): 205–241. ISSN : 00318116. doi : 10.1007/s11098-005-5385-4.

Press, William H., Saul A. Teukolsky, William T. Vetterling et Brian P. Flannery. 2007. Recettes numériques. L'art du calcul scientifique. La presse de l'Universite de Cambridge.

Primiero, Giuseppe. 2014. "Sur l'ontologie du processus informatique et l'épistémologie du calculé." Philosophie et technologie 27 (3): 485–489.

————. 2016. "L'information dans la philosophie de l'informatique." Dans The Rout ledge Handbook of Philosophy of Information, édité par Luciano Floridi, 90–106.

Rapaport, William J. 2005. "La mise en œuvre est une interprétation sémantique : pensées." Journal of Experimental & Theoretical Artificial Intelligence 17 (4): 385–417.

Rapaport, William J. 1999. « La mise en œuvre est une interprétation sémantique » [en en]. Le Moniste 82 (1): 109-130.

Rojas, Rauul et Ulf Hashagen, éd. 2000. Les premiers ordinateurs. Histoire et architectures. Presse du MIT.

Spivey, JM 2001. La Notation Z : Un Manuel de Référence. Prentice Hall.

Tourner, Raymond. 2011. "Spécification". Esprits et Machines 21 (2): 135–152. ISSN : 09246495. doi : 10.1007/s11023-011-9239-x.

Von Neumann, Jean. 1945. Première ébauche d'un rapport sur l'EDAVAC. États-Unis Département des munitions de l'armée de l'Université de Pennsylvanie.

Winsberg, Éric. 2010. La science à l'ère de la simulation informatique. Université de Presse de Chicago.

Woolfson, Michael M., et Geoffrey J. Pert. 1999. SATELLIT.FOR.

Zénil, Hector. 2014. "Qu'est-ce que le calcul semblable à la nature ? Une approche comportementale et une notion de programmabilité. Philosophie & Technologie 27, no. 3 (septembre): 399. doi:10.1007/s13347- 012- 0095- 2. http://dx.doi. org/10.1007/s13347-012-0095-2.

chapitre 3
Unités d'analyse II : Expérimentation en laboratoire et simulations informatiques

Lorsque les philosophes ont fixé leur attention sur les simulations informatiques, trois de grandes lignes d'étude ont émergé (Duran 2013a). La première ligne d'étude se concentre sur la recherche d'une définition appropriée pour les simulations informatiques. Une étape fondamentale vers la compréhension des simulations informatiques est précisément de mieux appréhender leur nature en s'approchant d'une définition. C'était le sujet de notre premier chapitre, où nous avons suivi les définitions remontent au début des années 1960.

Le deuxième axe d'étude relie les simulations informatiques à d'autres unités d'analyse plus familières aux chercheurs, telles que les modèles scientifiques et l'expérimentation en laboratoire. Cette relation étant établie sur une base comparative, le type de questions soulevées dans ce contexte sont, entre autres, « les simulations informatiques sont-elles une forme de modèles scientifiques, ou sont-ils des formes d'expérimentation ? quel genre de connaissances auquel le chercheur devrait s'attendre en utilisant une simulation informatique, en comparaison avec sorte de connaissances obtenues en utilisant des modèles et des expériences scientifiques ? » La comparaison de simulations informatiques avec des modèles scientifiques a fait l'objet du chapitre 2, où J'ai discuté de leurs principaux constituants (c'est-à-dire les spécifications, les algorithmes et les processus). Ce chapitre traite de la comparaison des simulations informatiques avec l'expérimentation en laboratoire. À cet égard, les principales questions ici sont "l'ordinateur peut-il les simulations produisent-elles le type de connaissances sur le monde que produisent les expériences en laboratoire ? à quels égards les simulations informatiques sont-elles plus – ou moins – adaptées à rendre une connaissance fiable du monde empirique ? Ces questions ont été au cœur de nombreux débats philosophiques sur et autour des simulations informatiques. Ici, je m'intéresse à reconstruire une partie de ce débat, ses hypothèses et conséquences.

Enfin, le troisième axe aborde les simulations informatiques au pied de la lettre, en posant la question de leur pouvoir épistémologique à fournir la connaissance et la compréhension d'un système cible donné. Cette troisième ligne est indépendante des deux autres dans la mesure où il ne s'intéresse ni à la définition de simulations informatiques ni à l'établissement des comparaisons avec des manières plus familières d'enquêter sur le monde. Le reste chapitres sont consacrés à étoffer bon nombre des problèmes soulevés par l'ordinateur simulations.

3.1 Expérimentation en laboratoire et simulations informatiques

Les simulations informatiques ont été régulièrement comparées à l'expérimentation en laboratoire, car dans de nombreux cas, elles sont utilisées de manière similaire et à des fins similaires.1 L'expérimentation est généralement conçue comme une activité multidimensionnelle issue non seulement de la complexité entourant les phénomènes empiriques étudiés, mais également de la pratique de l'expérimentation en elle-même, qui est complexe dans la conception et structure complexe. C'est pourquoi, quand les philosophes parlent d'expériences, ils font référence à une foule de sujets, de méthodologies et de façons de pratiquer entrelacés science. Les expériences, par exemple, sont utilisées pour observer des processus, détecter de nouveaux entités, mesurant des variables, et même dans certains cas pour « tester » la validité d'un théorie. Les questions qui guident cette section sont alors de savoir dans quelle mesure peut-on dire que les simulations informatiques sont-elles épistémiquement proches – voire supérieures – de l'expérimentation ? et quel ensemble de caractéristiques les rend deux distincts - ou similaires - les pratiques? Commençons par mieux comprendre ce qu'est l'expérimentation.

L'idée d'expérimenter avec la nature remonte aux premiers temps de civilisation. Aristote a consigné son observation de l'embryologie du poussin dans son Historia Animalium (Aristote 1965), facilitant notre compréhension précoce du poulet et le développement humain. En fait, les études d'Aristote ont correctement déduit le rôle de le placenta et le cordon ombilical chez l'homme. Bien que sa méthodologie soit imparfaite à plusieurs égards, elle ressemble néanmoins beaucoup à la méthode scientifique moderne : observer, mesurer et documenter chaque étape de la croissance – trois aspects pratique scientifique encore en usage jusqu'à aujourd'hui2. Or, malgré son indéniable centralité dans notre compréhension moderne du monde empirique, l'expérimentation en laboratoire a pas toujours reçu l'appréciation qu'il mérite.

Ce n'est qu'avec l'arrivée de l'empirisme logique dans les années 1920 et 1930 que les expériences ont commencé à recevoir une certaine attention dans la philosophie générale des sciences. À Pour l'empiriste logique, cependant, l'expérimentation représentait moins un problème philosophique en soi qu'une méthodologie subsidiaire pour comprendre la théorie. Dans fait, l'utilisation la plus importante des expériences était pour la confirmation et la réfutation d'une théorie, la question philosophique la plus répandue à l'époque.

Quelques décennies plus tard, l'empirisme logique a commencé à faire l'objet d'un certain nombre d'objections et d'attaques de différents bords. Une objection particulière a joué un rôle fondamental dans leur disparition, connue plus tard sous le nom de sous-détermination. de la théorie par la preuve. En son cœur, cette objection stipule que les éléments de preuve recueillis de l'expérimentation pourrait être insuffisant pour la confirmation ou la réfutation d'une théorie donnée à un moment donné. Les empiristes logiques n'avaient alors d'autre choix que d'embrasser l'expérimentation comme partie intégrante de la recherche scientifique et philosophique.

---

1 Permettez-moi d'apporter une précision terminologique et une délimitation de sujet. La clarification c'est que j'emploie indifféremment et sans autre discussion les notions d'expérimentation en laboratoire et d'expérimentation. Les subtilités de la distinction n'ont aucun intérêt pour notre propos. Pour ce qui est de la délimitation, je laisse de côté les expériences sur le terrain car elles nécessitent généralement une approche différente approche philosophique.

2 La méthode expérimentale moderne, cependant, doit être attribuée à Galileo Galilei.

Robert Ackerman et Deborah Mayo, deux grands noms de la philosophie de
expérimentation, se réfèrent à l'ère où l'expérimentation est au centre de la recherche philosophique en tant que
nouvel expérimentalisme. [3] Le nouvel expérimentalisme, tel que présenté,
complète la vision traditionnelle, basée sur la théorie, de l'empirisme logique avec une vision plus
vision expérimentale de la pratique scientifique.

Bien que les tenants du nouvel expérimentalisme s'intéressent à différentes
sortes de problèmes émergeant des expérimentations et de leurs pratiques, ils partagent tous
affirmation selon laquelle l'expérimentation scientifique est au cœur d'une grande partie de notre compréhension
du monde empirique. Le philosophe de l'expérimentation Marcel Weber propose
cinq tendances générales qui caractérisent le nouvel expérimentalisme. Premièrement, l'expérimentation est
exploratoire, c'est-à-dire qu'elle vise à découvrir de nouveaux phénomènes et des régularités empiriques.
Deuxièmement, les nouveaux expérimentateurs rejettent l'idée que l'observation et l'expérimentation
est guidé par la théorie. Ils soutiennent que dans un nombre important de cas, sans théorie
des expériences sont possibles et se produisent dans la pratique scientifique. Troisièmement, le nouvel
expérimentalisme a donné un nouveau souffle à la distinction entre observation et expérimentation.
Quatrièmement, les partisans du nouvel expérimentalisme ont contesté l'idée positiviste
que les théories se rapportent d'une manière ou d'une autre à la nature sur la base de résultats expérimentaux. Et
Cinquièmement, il a été souligné qu'une plus grande attention doit être accordée à la pratique expérimentale afin
de répondre aux questions concernant l'inférence scientifique et les tests théoriques.
(Weber 2005).

Le passage d'un schéma traditionnel « descendant » (c'est-à-dire de la théorie au monde empirique) à une
conceptualisation « ascendante » est la marque distinctive des nouvelles
expérimentalisme. Même des notions comme celle de phénomène naturel ont subi des transformations. Selon le
nouveau point de vue expérimentaliste, un phénomène peut être
des voitures directement observables qui s'écrasent devant nos maisons, à l'invisible
microbes, événements astronomiques et monde quantique.

Les chercheurs donnent le label « expérimentation » à un large éventail d'activités. L'observation du poussin
par Aris totle est peut-être l'utilisation la plus directe du terme.
Il existe une autre utilisation plus large du terme qui implique une intervention ou une manipulation
de nature. L'idée est très simple et séduisante : les scientifiques manipulent un montage expérimental comme
s'ils manipulaient le phénomène empirique lui-même. Quel que soit le gain épistémique du premier, il peut être
extrapolé au second. Sous
cette notion plusieurs activités peuvent être identifiées. L'un d'entre eux est la découverte de nouvelles entités,
une occupation très précieuse dans la recherche scientifique. Par exemple, le ¨ de Wilhelm Rontgen
la découverte des rayons X est un bon exemple de la découverte d'un nouveau type de rayonnement
en manipulant la nature.

Certains types de mesures de quantités sont un autre exemple de manipulation de la nature.
Par exemple, mesurer la vitesse de la lumière au milieu des années 1800 nécessitait un faisceau de
lumière pour se refléter sur un miroir à quelques kilomètres de distance. L'expérimentation a été mise en place

---

[3] Les travaux d'Ackerman se trouvent dans (Ackermann 1989), et de Mayo dans (Mayo 1994). Laisser
que le lecteur soit avisé que je saute plusieurs années de bonne philosophie de l'expérimentation.
Norwood R. Hanson, philosophe des sciences et farouche opposant à
l'empirisme logique, qui a apporté des contributions fondamentales au passage des expériences
comme méthodologie subsidiaire de la théorie, aux expériences comme unités d'étude à part entière. Pour
références, voir (Hanson 1958).

de telle manière que le faisceau devrait passer à travers les espaces entre les dents d'une roue en rotation rapide. La vitesse de la roue a donc été augmentée jusqu'à ce que la la lumière de retour passait à travers l'espace suivant et pouvait être vue. Une solution très astucieuse utilisé par Hippolyte Fizeau pour améliorer la précision des mesures passées.

Comprendre les expériences et la pratique expérimentale de cette manière soulève des questions sur la relation entre les expériences et les simulations informatiques. Sont-ils épistématiquement équivalents ? ou la capacité de manipuler le monde réel donne-t-elle aux expériences un avantage épistémique sur les simulations informatiques ? Le critère peut-être le plus célèbre pour l'analyse des simulations informatiques et des expériences de laboratoire est le argument dit de la matérialité. À sa base, l'argument de la matérialité dit que, dans véritables expériences, les mêmes causes matérielles sont à l'œuvre dans le dispositif expérimental comme dans le système cible ; dans les simulations informatiques, au contraire, il y a une correspondance formelle entre le modèle de simulation et le système cible.

Ainsi compris, l'argument de la matérialité offre différentes formes de et engagements épistémologiques. Une telle forme nécessite des expériences à faire des mêmes causes matérielles que le système cible, tandis que les simulations informatiques partager une correspondance formelle avec un tel système cible. Selon cette interprétation, les inférences sur le système cible sont plus justifiées dans une expérience que dans une simulation informatique.4 Alternativement, un argument peut être avancé où les expériences sont similaires aux simulations informatiques, et donc les inférences par les deux sont également justifié.

Dans ce qui suit, je reproduis en partie un de mes articles paru en 2013 où je discuter en détail des différentes façons de comprendre l'argument de la matérialité et son impact dans l'évaluation épistémologique des simulations informatiques. Cet article prévoit, je espoir, un niveau de détail technique et philosophique similaire à celui du livre. Laisse-moi enfin dire que beaucoup plus de travaux philosophiques sur la relation entre les simulations informatiques et l'expérimentation ont été publiés après cet article. Les exemples sont les excellent travail d'Emily Parke (Parke 2014), Michela Massimi et Wahid Bhimji (Massimi et Bhimji 2015), et plus récemment Claus Beisbart (Beisbart 2017).

3.2 L'argument de la matérialité5

Une grande partie de l'intérêt philosophique actuel pour les simulations informatiques découle de leur présence prolongée dans la pratique scientifique. Cet intérêt s'est centré sur les études des caractère expérimental des simulations informatiques et, à ce titre, sur les différences – et les similitudes – entre les simulations informatiques et les expériences en laboratoire. Le l'effort philosophique s'est donc principalement concentré sur l'établissement de la base de ce contraste; spécifiquement en comparant la puissance épistémique d'une simulation informatique avec celle d'une expérience de laboratoire. L'intuition de base a été

4 La reconstruction la plus compréhensible de l'argument de la matérialité est donnée par Wendy Parker dans (Parker 2009).

5 Le texte suivant a été partiellement publié dans (Duran 2013b). Publié avec la permission de Cambridge Scholars Publishing.

que si les simulations informatiques ressemblent à des expériences de laboratoire dans des domaines épistémiques pertinents
respects, alors ils peuvent aussi être sanctionnés comme un moyen de fournir une compréhension de
le monde.

La littérature standard sur le sujet distingue les simulations informatiques des expériences de laboratoire pour des raisons à la fois ontologiques et représentationnelles. Le fait que une simulation informatique est une entité abstraite, et n'a donc qu'une relation formelle au système étudié, contraste avec une expérience de laboratoire, qui a généralement un lien de causalité avec le système cible. Ces différences ontologiques et représentationnelles ont suggéré à certains philosophes que l'établissement de
la validité est une tâche beaucoup plus difficile pour les simulations informatiques que pour les simulations de laboratoire.
expériences. Pour d'autres, cependant, cela a été une motivation pour reconsidérer la pratique expérimentale et la considérer comme une activité plus large qui inclut également des simulations en tant que nouvel outil scientifique. Ces deux approches, selon moi, partagent une logique commune qui impose des restrictions à l'analyse épistémologique des simulations informatiques.

Le critère le plus connu pour distinguer les simulations informatiques et les expériences de laboratoire sont données par l'argument dit de la matérialité. Parker a fourni un compte rendu utile de cet argument:

> Dans les expériences authentiques, les mêmes causes « matérielles » sont à l'œuvre dans systèmes cibles, alors que dans les simulations il n'y a qu'une correspondance formelle entre les simulation et systèmes cibles [...] les inférences sur les systèmes cibles sont plus justifiées lorsque les systèmes expérimentaux et cibles sont faits de la « même matière » que lorsqu'ils sont faits de différents matériaux (comme c'est le cas dans les expériences sur ordinateur). (Parker 2009, 484)

Deux réclamations sont faites ici. La première est que les simulations informatiques sont des entités abstraites, alors que les expériences partagent le même substrat matériel que la cible. La seconde , essentiellement épistémique, est que les inférences sur les systèmes cibles empiriques sont davantage justifiées par des expériences que par des simulations informatiques. en raison des relations matérielles que le premier entretient avec le monde.

La littérature actuelle a combiné ces deux affirmations en deux propositions différentes : soit on accepte les deux affirmations et on encourage l'idée qu'il vaut mieux être matériel justifie les inférences sur le système cible que d'être abstrait et formel (Guala 2002 ; Morgan 2005); ou on rejette les deux affirmations et on encourage l'idée que les simulations informatiques sont de véritables formes d'expérimentation et, en tant que telles, épistémiquement au même titre que les pratiques expérimentales (Morrison 2009 ; Winsberg 2009 ; Parker 2009). Je prétends que ces deux groupes de philosophes, qui semblent superficiellement en désaccord, partagent en fait une logique commune dans leur argumentation. Concrètement, ils se disputent tous pour des engagements ontologiques qui fondent leurs évaluations épistémiques sur ordinateur simulations. J'appellerai cette justification le principe de matérialité.

Afin de montrer que le principe de matérialité est à l'œuvre dans la majeure partie de la littérature philosophique sur les simulations informatiques, j'aborde trois points de vue distincts, à savoir:

---

6 Une partie de la terminologie dans la littérature reste non spécifiée, comme les causes « matérielles » ou les « choses » (Guala 2002). Je les prends ici pour signifier des relations causales physiques, telles que décrites, par exemple, par (Dowe 2000). Dans le même ordre d'idées, lorsque je parle de causes, de causalité et de termes similaires, ils doivent être interprété de la manière indiquée ici.

a) Les simulations informatiques et les expériences sont ontologiquement similaires (les deux partagent la même matérialité avec le système cible) ; par conséquent, ils sont épistémiquement à égalité (Parker 2009);
b) Les simulations informatiques et les expériences sont ontologiquement différentes. Alors que le premier est de nature abstraite, le second partage la même matérialité avec le phénomène étudié ; par conséquent, ils sont épistémiquement différents (Guala 2002 ; Giere 2009 ; Morgan 2003, 2005) ; c) Les simulations informatiques et les expériences sont ontologiquement similaires (les deux sont « en forme de modèle ») ; par conséquent, ils sont épistémiquement sur un pied d'égalité (Morrison 2009 ; Winsberg 2009).

Avec ces trois points de vue à l'esprit, le principe de matérialité peut être recadré d'un autre point de vue : c'est en raison de l'attachement des philosophes à l'abstraction - ou à la matérialité - des simulations informatiques que les inférences sur le système cible en sont plus - ou moins - justifiée que les expériences de laboratoire.

L'objectif principal ici est de montrer que les philosophes de la simulation informatique adhèrent, d'une manière ou d'une autre, au principe de matérialité. Je suis également intéressé à décrire certaines des conséquences de l'adoption de ce raisonnement. En particulier, je suis convaincu que fonder l'analyse philosophique sur le principe de matérialité, comme semble le faire la majeure partie de la littérature actuelle, met un corset conceptuel à l'étude de la puissance épistémologique des simulations informatiques. L'étude philosophique des simulations informatiques ne doit pas être restreinte, ni limitée par des engagements ontologiques a priori. En analysant des thèmes dans la littérature, je montre donc que le principe de matérialité n'engendre pas une conceptualisation utile du pouvoir épistémique des simulations informatiques.

Les sections suivantes sont divisées d'une manière qui correspond aux trois utilisations de l'argument de la matérialité énumérées ci-dessus. La section intitulée l'identité de l'algorithme traite de l'option a); la section intitulée material stuff comme critère aborde l'option b), qui se décline en deux versions, la version forte et la version faible ; et enfin l'option c) est abordée dans la section intitulée modèles en tant que médiateurs (total).

3.2.1 L'identité de l'algorithme

La formulation par Wendy Parker de l'argument de la matérialité occupe une place de choix dans la littérature récente sur la simulation informatique. Suivant (Hartmann 1996), Parker définit une simulation informatique comme une séquence d'états ordonnée dans le temps qui représente de manière abstraite un ensemble de propriétés souhaitées du système cible. L'expérimentation, d'autre part, est l'activité consistant à mettre la configuration expérimentale dans un état particulier en y intervenant et à étudier comment certaines propriétés d'intérêt dans la configuration changent en conséquence de cette intervention (Parker 2009, 486). 7

7 L'« intervention » est conçue comme la manipulation des relations causales physiques dans le cadre expérimental.

L'objectif de Parker est de montrer que les simulations informatiques et les expériences partagent le même même base ontologique, et d'utiliser cette base pour justifier l'affirmation selon laquelle les simulations informatiques et les expériences sont épistémiquement sur un pied d'égalité. À son avis, le problème central est que les définitions actuelles de la simulation par ordinateur ne sont pas considérées comme un expérimenter parce qu'ils manquent des mécanismes d'intervention cruciaux. En effet, c'est le caractère abstrait du modèle qui empêche les simulations informatiques de servir comme systèmes intermédiaires. La solution à ce problème consiste à interpréter la notion d'études de simulation informatique en tant que simulation informatique où une intervention est fait dans l'ordinateur physique lui-même. Ainsi définie, une étude de simulation informatique est considéré comme une expérience.

> Une étude de simulation informatique [...] consiste en une activité plus large qui comprend la définition des état du calculateur numérique à partir duquel une simulation va évoluer, déclenchant cette évolution en démarrant le programme informatique qui génère la simulation, puis en collectant des informations sur la façon dont diverses propriétés du système informatique, telles que les valeurs stockées à divers endroits de sa mémoire ou les couleurs affichées sur son moniteur, évoluent à la lumière de l'intervention antérieure. (488)

La notion d'intervention est désormais redéfinie comme l'activité de mise en place état du système informatique et déclenchant son évolution ultérieure. Ainsi comprise, une étude par simulation informatique est une expérience au sens propre, car maintenant le système intervenu est le calculateur numérique programmé (488). Sur cette base, Parker affirme qu'il existe une équivalence ontologique entre les simulations informatiques et des expériences, ce qui lui permet à son tour de revendiquer une équivalence dans leur pouvoir épistémologique.

Notamment, elle n'explique pas ce que cela signifie pour une étude de simulation informatique être épistémiquement puissant. Au lieu de cela, elle limite l'argument à affirmer qu'un l'épistémologie des simulations informatiques devrait refléter le fait qu'il s'agit comportement du système informatique qui leur fait des expériences sur un matériau réel système – et donc épistémiquement puissant.

L'influence du principe de matérialité peut maintenant être explicitée. Tout d'abord Parker semble exiger que l'ordinateur numérique soit le "substrat" du système simulé, puisque cela lui permet de revendiquer une équivalence ontologique entre ordinateur études de simulation et expériences. De plus, étant donné que la simulation informatique l'étude est l'activité consistant à mettre l'ordinateur physique dans un état initial, déclenchant l'évolution de la simulation, et la collecte de données physiques telles qu'indiquées par des impressions, des affichages à l'écran, etc. (489), puis la valeur épistémique de la simulation informatique études correspond également à celle des expériences. En ce sens, l'évolution du comportement de l'ordinateur programmé représente les caractéristiques matérielles du phénomène étant simulé. Enfin, la compréhension qu'a le chercheur d'un tel phénomène est justifiée par son évolution sur l'ordinateur physique. Etudes de simulation informatique et les expériences sont donc ontologiquement à égalité, de même que leur pouvoir épistémologique.

Ici, j'ai brièvement décrit les principales affirmations de Parker. Le problème avec son compte, Je crois, c'est qu'on ne sait toujours pas quelles sont les raisons de considérer la matérialité de l'ordinateur numérique comme l'acteur pertinent dans l'épistémologie de l'informatique. simulations. Permettez-moi de formuler cette préoccupation en d'autres termes. À mon avis, les motivations de Parker sont de renverser l'argument de la matérialité en montrant que les simulations informatiques

et les expériences sont ontologiquement sur un pied d'égalité – tout comme leur pouvoir épistémique. Ce Le mouvement, comme je l'ai soutenu, est fondé sur une logique derrière le même argument de matérialité qu'elle essaie de renverser. La question est alors de savoir quel rôle joue le matérialité du jeu de l'ordinateur numérique dans l'évaluation du pouvoir épistémique de études de simulation informatique? Permettez-moi maintenant d'offrir trois interprétations possibles à ce question.

Tout d'abord, Parker considère la matérialité de l'ordinateur numérique comme jouant le rôle fondamental de « provoquer » le système cible (c'est-à-dire de créer une existence causale le phénomène simulé). En d'autres termes, les changements de comportement que les le scientifique observe dans l'ordinateur physique sont des instanciations des représentations intégrée à la simulation informatique. De telles représentations sont naturellement des représentations d'un système cible. De cette façon, l'ordinateur physique se comporte comme s'il était le phénomène empirique étant simulé dans l'ordinateur programmé. À cet égard, Parker dit que "[l]e système expérimental dans une expérience informatique est l'ordinateur numérique programmé (un système physique fait de fil, de plastique, etc.)" (Parker 2009, 488-489). Je ne sais pas si Parker utilise une métaphore ou, à la place, elle nous exhorte à prendre cette citation au pied de la lettre. Dans (Duran 2013b, 82), j'appelle cette interprétation le « phénomène dans la machine » et je montre comment il est techniquement impossible à obtenir.

Une deuxième interprétation possible est que le système d'intérêt est la physique ordinateur lui-même, quel que soit le système empirique représenté. Dans ce scénario, le la chercheuse exécute ses simulations comme d'habitude, en ne prêtant attention qu'aux changements dans le comportement de l'ordinateur physique. Ces changements de comportement deviennent la substance de l'enquête du scientifique, alors que le système cible n'est considéré que comme le référence pour la construction du modèle de simulation. Dans ce contexte, la chercheur apprend d'abord et avant tout en recueillant des informations sur les propriétés de l'ordinateur physique (c'est-à-dire les valeurs dans sa mémoire et les couleurs sur le moniteur (Parker 2009, 488)). Si c'est la bonne interprétation, alors Parker doit montrer que le scientifique peut accéder cognitivement aux différents états physiques de l'ordinateur, quelque chose qu'elle ne fait pas. Les philosophes se sont demandé s'il était possible pour accéder à différents emplacements à l'intérieur d'un ordinateur (par exemple, la mémoire, le processeur, le bus informatique, etc.) et l'accord général est que ces endroits sont cognitivement inaccessibles pour l'homme sans aide. Il existe un principe directeur épistémique opacité attribuée aux processus de calcul qui exclut toute possibilité d'accéder cognitivement aux états internes de l'ordinateur physique (voir ma discussion dans la rubrique 4.3). De plus, même si les scientifiques pouvaient réellement accéder à ces emplacements - disons, s'ils étaient aidés par un autre ordinateur - on ne sait toujours pas pourquoi accéder à ces emplacements seraient de toute pertinence pour comprendre les résultats d'un ordinateur simulation.

Une troisième interprétation est que Parker prend la matérialité de l'ordinateur physique jouer un rôle pertinent dans l'interprétation des résultats (490). Selon cette interprétation, les pannes matérielles, les erreurs d'arrondi et les sources analogues d'erreurs de calcul affectent les résultats de la simulation de différentes manières. Comme c'est le cas des ordinateurs et du calcul, alors l'affirmation de Parker doit être que l'ordinateur physique affecte les résultats finaux d'une simulation informatique et, par conséquent, leur assimilation épistémologique.

session. Si c'est la bonne interprétation, alors je crois qu'elle a raison. Dans le chapitre 4, je présente et discute de la manière dont les chercheurs peuvent savoir que les résultats des simulations informatiques sont corrects malgré les nombreuses sources d'erreurs impliquées dans le calcul.

3.2.2 Matériel comme critère

Les partisans de la « substance matérielle comme critère » sont peut-être les meilleurs interprètes de l'argument de la matérialité. Selon ce point de vue, il existe des différences ontologiques fondamentales et irréconciliables entre les simulations informatiques et les expériences, ces dernières étant épistémiquement supérieures. Il existe deux versions de ce compte : une version forte et une version faible.

La version forte soutient que les relations causales responsables de l'apparition du phénomène doivent également être présentes dans le dispositif expérimental. Cela signifie que l'expérience doit reproduire les relations causales présentes dans le système empirique. Selon la version forte, l'expérience est donc un « morceau » du monde.

Prenons comme exemple un faisceau lumineux utilisé pour comprendre la nature de la propagation de la lumière. Dans ce cas, le montage expérimental est identique au système cible ; c'est-à-dire qu'il s'agit simplement du système empirique à l'étude. Il s'ensuit que toute manipulation de la configuration expérimentale traite les mêmes causes que le phénomène, et qu'un aperçu de la nature de la lumière peut être fourni par notre compréhension de l'expérience contrôlée (c'est-à-dire le faisceau de lumière (Guala 2002)).

Appliquée aux simulations informatiques, la version forte considère que la simple correspondance formelle entre l'ordinateur et le système cible fournit une base suffisante pour minimiser leur statut de dispositifs épistémiques. S'il n'y a pas de relations causales présentes, alors le pouvoir épistémique des inférences ainsi faites sur le monde est dégradé.

La version faible, en revanche, assouplit certaines des conditions imposées par la version forte à l'expérimentation. Selon ce point de vue, une expérience contrôlée ne nécessite que l'ensemble des relations causales pertinentes qui provoquent le phénomène. Dans cette veine, les tenants de la version faible ne s'engagent pas sur une reproduction complète du phénomène étudié, comme le fait la version forte, mais plutôt sur l'ensemble des causes pertinentes qui caractérisent le comportement du phénomène.

Illustrons la version faible avec un exemple simple : un réservoir d'ondulations peut être utilisé comme représentation matérielle de la lumière, donnant ainsi un aperçu de sa nature d'onde. Pour le partisan de la version faible, il suffit d'avoir une collection représentative de correspondances causales entre la configuration expérimentale et le système cible pour que le premier donne un aperçu du second. La relation entre l'expérience et le phénomène du monde réel fait donc partie d'un sous-ensemble de toutes les relations causales. Une chambre à brouillard détecte les particules alpha et bêta, tout comme un compteur Geiger peut les mesurer, mais aucun des deux instruments n'est un «morceau» du phénomène étudié et n'interagit pleinement avec toutes sortes de particules. Il s'ensuit que

la pratique expérimentale, telle qu'illustrée par la détection et la mesure des particules, dépend d'un ensemble complexe mais partiel de toutes les relations causales existant entre le montage expérimental et le système cible.

Appliquée à l'évaluation générale des simulations informatiques, la version faible présente une image plus complexe et plus riche, qui offre des degrés de matérialité attribuée à des simulations informatiques.

Malgré ces différences, cependant, les deux versions partagent le même point de vue concernant les simulations informatiques ; à savoir qu'ils sont épistémiquement inférieurs aux expériences. Cette affirmation découle de la conceptualisation ontologique décrite précédemment et découle du même raisonnement qui sous-tend le principe de matérialité.

3.2.2.1 La version forte

Je crois que Francesco Guala défend la défense de la version forte quand il suppose qu'une expérience reproduit les relations causales présentes dans le phénomène. A cet égard, il suppose d'emblée l'existence de principes fondamentaux différences entre les simulations informatiques et les expériences fondées sur la causalité.

> La différence réside dans le type de relation qui existe entre, d'une part, un système expérimental et son système cible, et, d'autre part, un système simulant et son système cible. Dans le premier cas, la correspondance tient à un niveau "profond", "matériel", alors que dans le second la similitude n'est certes que 'abstraite' et 'formelle' [...] Dans une véritable expérience, le même des causes « matérielles » comme celles du système cible sont à l'œuvre ; dans une simulation, ils ne le sont pas, et la relation de correspondance (de similarité ou d'analogie) est de caractère purement formel (Guala 2002, 66-67. C'est moi qui souligne).

Pour Guala, les changements de matérialité et leur pouvoir épistémologique peuvent être compris en termes de partage de choses « identiques » et « différentes ». L'affaire de le ripple-tank est paradigmatique à cet égard. Selon Guala, les médias de traversés par les ondes sont constitués de matières « différentes » : tandis qu'un milieu est l'eau, le l'autre est léger. Le réservoir d'ondulation est donc une représentation de la nature ondulatoire de la lumière uniquement parce qu'il existe des similitudes dans le comportement à un niveau très abstrait (c'est-à-dire à niveau des équations de Maxwell, de l'équation d'onde de D'Alambert et de la loi de Hook). Les deux systèmes obéissent aux « mêmes » lois et peuvent être représentés par le « même » ensemble de équations, bien qu'elles soient constituées d'éléments "différents". Cependant, les vagues d'eau sont pas des ondes lumineuses, et une différence de matérialité suppose une différence de vision épistémique de la nature (66).

L'exemple du ripple-tank est extrapolé aux études sur les simulations informatiques, car il permet à Guala d'affirmer que la différence ontologique entre expérimentations et simulations fonde aussi des différences épistémologiques (63). Sa fidélité à le principe de matérialité est donc indiscutable : il y a une nette distinction entre ce que nous pouvons apprendre et comprendre par l'expérimentation directe, et ce que nous pouvons apprendre par une simulation informatique. Le gain épistémique de ce dernier est inférieur à celui du premier et c'est parce que, de ce point de vue, il y a un engagement ontologique envers la causalité en tant que épistémiquement supérieur qui détermine l'épistémologie des simulations informatiques.

Considérons maintenant quelques objections au point de vue de Guala. Parker a objecté que sa position était trop restrictive pour les expériences, ainsi que pour l'informatique.
simulations (Parker 2009, 485). Je suis d'accord avec elle sur ce point. La conceptualisation par Guala des expériences et des simulations informatiques impose des restrictions artificielles sur les deux sont difficiles à étayer par des exemples dans la pratique scientifique. De plus, et complémentaire à l'objection de Parker, je crois que Guala adopte une perspective qui considère les deux activités comme chronologiquement mutuellement exclusives : c'est-à-dire que la simulation informatique devient un outil pertinent lorsque l'expérimentation ne peut être mis en œuvre. STRATAGEM, une simulation informatique de la stratigraphie, nous fournit avec un exemple ici : quand les géologues sont confrontés à des difficultés pour réaliser expériences contrôlées sur la formation des strates, elles font appel à des simulations informatiques comme le remplacement le plus efficace (2002, 68).8 Une telle tendance à une évaluation disjonctive des deux activités est une conséquence naturelle de la prise en compte les simulations sont épistémiquement inférieures à l'expérimentation. En d'autres termes, il s'agit d'un conséquence naturelle de l'adoption du principe de matérialité.

3.2.2.2 La version faible

Pour un partisan de la version faible, je me tourne vers le travail de Mary Morgan. Elle a a présenté l'analyse la plus riche et la plus exhaustive que l'on trouve actuellement dans la littérature concernant les différences entre les expériences et les simulations informatiques.

Morgan s'intéresse principalement aux expériences dites vicariantes, c'est-à-dire :

> Les expériences qui impliquent des éléments de non matérialité soit dans leurs objets soit dans leurs interventions et qui résultent de la combinaison de l'utilisation de modèles et d'expériences, une combinaison qui a créé un certain nombre de formes hybrides intéressantes (Morgan 2003, 217).

Ayant ainsi exposé les caractéristiques des expériences vicariantes, elle se tourne ensuite vers les question de savoir comment ils fournissent une base épistémique pour l'inférence empirique. Brièvement, le plus de « choses » sont impliquées dans l'expérience vicariante, plus elle est épistémiquement fiable devient. En termes simples, les degrés de matérialité déterminent les degrés de fiabilité. Comme le commente Morgan : « pour des raisons d'inférence, l'expérience reste la solution préférable mode d'enquête parce que l'équivalence ontologique fournit un pouvoir épistémologique » (Morgan 2005, 326).

Morgan adhère donc à la version faible, car une expérience vicariante est caractérisée par différents degrés de matérialité, par opposition à la version forte qui estime que les expériences doivent être un « morceau » du monde. Du point de vue de la matérialité principe, cependant, il n'y a pas de différences fondamentales entre les deux versions : elle considère également l'ontologie pour déterminer la valeur épistémologique des simulations informatiques. La différence réside, encore une fois, dans l'analyse détaillée des différents types de expériences impliquées dans la pratique scientifique. Permettez-moi maintenant d'aborder brièvement son récit.

Comme indiqué ci-dessus, les expériences vicariantes peuvent être classées en fonction de leur degré de matérialité ; c'est-à-dire les différents degrés auxquels la matérialité d'un objet est

---

8 Guala admet que les expériences et les simulations informatiques sont des outils de recherche appropriés, producteurs de connaissances comme il les appelle, mais seulement dans des contextes différents (2002 : 70).

présent dans le montage expérimental. Le tableau 3.1 résume quatre classes d'expériences : expérience de laboratoire idéale (également appelée expérience matérielle), deux types d'expériences hybrides et enfin expérience de modèle mathématique. Comme l'indique le tableau, la classification se fait en fonction du type de contrôle exercé sur la classe d'expérience, des modalités de démonstration de la fiabilité des résultats obtenus, du degré de matérialité et de la représentativité de chaque classe.

La première et la dernière classe nous sont déjà bien connues : un exemple d'expérience de laboratoire idéale est le faisceau de lumière, car elle demande un effort de la part du scientifique pour isoler le système, une attention rigoureuse au contrôle des circonstances perturbatrices et une intervention sous ces conditions de contrôle. Un exemple d'expérience de modèle mathématique, en revanche, serait le fameux problème mathématique des sept ponts de Königsberg ; c'est-à-dire une classe d'expériences dont les exigences de contrôle sont satisfaites par des hypothèses simplificatrices, dont la méthode de démonstration est via une méthode mathématique/logique déductive et dont la matérialité est, comme prévu, inexistante (Morgan 2003 : 218).

Parmi les nombreuses différences entre ces deux classes d'expériences, Morgan met l'accent sur les contraintes imposées naturellement par la causalité physique et sur celles imposées artificiellement par des hypothèses :

> L'action de la nature crée des limites et des contraintes pour l'expérimentateur. Il y a aussi des contraintes dans les mathématiques du modèle, bien sûr, mais le point critique est de savoir si les hypothèses qui y sont faites sont les mêmes que celles de la situation représentée et il n'y a rien dans les mathématiques elles-mêmes pour garantir que ils sont (220).

Les expériences hybrides, quant à elles, peuvent être conçues comme des expériences intermédiaires entre les deux autres : elles ne sont ni entièrement matérielles ni entièrement mathématiques9.) objets semi-matériels », tandis que les expériences virtuelles sont celles « dans lesquelles nous avons des expériences non matérielles mais qui peuvent impliquer une sorte d'imitation d'objets matériels » (216). Le tableau 3.1 résume à nouveau les propriétés des quatre types d'expériences vicariantes en montrant leurs relations de représentation et d'inférence.

Les différences entre les expériences virtuelles et virtuelles peuvent être illustrées par l'exemple d'un os coxal de vache utilisé comme substitut de la structure interne des os humains. Pour mener à bien une telle expérience, il existe généralement deux alternatives : on peut utiliser une image 3D de haute qualité de l'os iliaque qui crée une carte détaillée de la structure osseuse, ou, alternativement, une image 3D informatisée de l'os stylisé os; c'est-à-dire une grille 3D informatisée représentant la structure de l'os stylisé. Selon Morgan, l'image 3D a un degré de vraisemblance plus élevé pour la structure de l'os iliaque réel car elle en est une représentation plus fidèle, par opposition à la mathématisation représentée par la grille 3D informatisée (230).

---

9 Morgan dit à ce sujet : "[b]y analysant le fonctionnement de ces différents types d'expériences hybrides, nous pouvons suggérer une taxonomie de choses hybrides entre les deux qui incluent des expériences virtuelles (entièrement immatérielles dans l'objet d'étude et dans l'intervention mais qui peuvent impliquer l'imitation d'observations) et virtuellement des expériences (presque une expérience matérielle en vertu de l'objet virtuellement matériel d'entrée) » (Morgan 2003, 232).

Tableau 3.1 Types d'expériences : laboratoire idéal, hybrides et modèles mathématiques avec relations de représentation (Morgan 2003, 231).

	Laboratoire idéal	Expériences hybrides		Mathématique
	expérience			expérience modèle
		Virtuellement	Virtuel	
Contrôles sur :				
Entrées expérimentales	expérimentales sur supposées			assumé
Apports expérimentaux	d'intervention ; supposé supposé			assumé
Environnement expérimental	sur intervention supposée et environnement			assumé
Simulation expérimentale de démonstration :	expérimentale/méthode			déductif
	en laboratoire	mathématique à l'aide d'un objet modèle		dans un modèle
Degré de matérialité de :				
Matériel d'entrée	semi-matériel		immatériel	mathématique
Matériel d'intervention	non matériel		immatériel	mathématique
Matériel de sortie	non matériel		non ou pseudo-matériel	mathématique
Représentant représentant de ... et relations d'inférence	... à la même chose dans le monde retour à d'autres types de		la représentation de...	
	représentant pour...			les choses du monde
	... à similaire dans le monde...			

La première est qualifiée d'expérience virtuelle, tandis que les secondes sont appelées expériences virtuelles.

Maintenant, quelles sont les différences entre tous ces différents types d'expériences ? Comme le montre le tableau 3.1, alors qu'une expérience virtuelle est semi- ou immatérielle, une l'expérience de laboratoire idéale est strictement matérielle. Aussi les méthodes de démonstration sont aussi sensiblement différents. La distinction entre une expérience virtuelle et une modèle mathématique, en revanche, semble se situer uniquement dans la méthode de démonstration, expérimentale pour la première et déductive pour la seconde. Morgan montre également comment les modèles de cours boursiers, bien qu'ils soient mathématiques modèles simulés sur ordinateur, peut être qualifiée d'expérience virtuelle en raison des données d'entrée et de l'observation des résultats (225). Les frontières entre cependant, les quatre classes d'expériences sont non fixées et dépendent de facteurs extérieurs à l'expérience en question. Par exemple, si une grille 3D de l'os de la vache fait l'utilisation de mesures réelles de l'os de vache comme données d'entrée, alors ce qui était à l'origine un l'expérience virtuelle devient virtuellement une expérience.

L'analyse épistémologique est fonction du degré de matérialité de la classe de l'expérience : « l'équivalence ontologique fournit un pouvoir épistémologique » (Morgan 2005, 326), comme l'indique Morgan. La rétro-inférence au monde à partir d'un système expérimental peut être mieux justifiée lorsque l'expérience et le système cible sont le même matériel. Comme l'explique Morgan : « l'ontologie est importante parce qu'elle affecte le pouvoir d'inférence » (324). Une simulation informatique, par exemple, ne peut pas tester les hypothèses théoriques du système représenté parce qu'il a été conçu pour

fournir des résultats cohérents avec les hypothèses intégrées. Une expérience en laboratoire, sur d'autre part, a été explicitement conçu pour laisser les faits sur la cible système "parler" par eux-mêmes. Selon Morgan, c'est donc le substrat matériel sous-jacent à une expérience qui est responsable de son pouvoir épistémique. D'où le l'expérience de laboratoire idéale est épistémiquement plus puissante qu'une expérience virtuelle ; à son tour, une expérience virtuelle est plus puissante qu'une expérience virtuelle, et ainsi de suite. Puisque les simulations informatiques ne peuvent être conçues que comme des expériences hybrides ou en tant qu'expériences mathématiques, il s'ensuit qu'elles sont toujours moins épistémiquement puissantes que les expériences de laboratoire idéales. Dans l'esprit de Morgan, il y a donc degrés de matérialité qui déterminent les degrés de pouvoir épistémique.

Dans ce contexte, Morgan utilise les termes surprise et confusion pour décrire les états épistémologiques du scientifique face aux résultats d'une simulation informatique et de une expérience matérielle, respectivement. Les résultats d'une simulation informatique ne peuvent que surprendre le scientifique car son comportement peut être retracé et réexpliqué dans termes du modèle sous-jacent. Une expérience matérielle, d'autre part, peut aussi bien surprendre que déconcerter le scientifique, car elle peut apporter des éléments nouveaux et inattendus. comportements inexplicables du point de vue de la théorie actuelle (Morgan 2005, 325), (Morgan 2003, 219). La matérialité de l'expérience fonctionne alors comme la garantie épistémique que les résultats peuvent être nouveaux, contrairement à la simulation, qui considère les résultats comme pouvant être expliqués en termes de modèle sous-jacent. Cela montre comment les idées de Morgan concernant les expériences et les simulations informatiques portent l'empreinte du principe de matérialité. Il présente le même raisonnement, mettant la matérialité comme caractéristique prédominante pour l'évaluation épistémique.

Malgré l'accent mis par Morgan sur la place de la matérialité dans la découverte de nouveaux phénomènes, il existe des exemples d'expériences virtuelles dont puissance est nettement supérieure à toute expérience de laboratoire idéale. Prenons comme exemple simple la dynamique de la micro fracture des matériaux. Il est impossible de savoir quoi que ce soit sur les micro-fractures sans l'aide d'ordinateurs. En effet, seul le calcul l'efficacité des méthodes d'éléments finis et la forte discontinuité multi-échelles peuvent nous dire quelque chose sur les micro fractures des matériaux (Linder 2012). La leçon ici est que comprendre quelque chose du monde ne vient pas nécessairement de expériences matérielles, ou à quelque degré de matérialité que ce soit. Ni un champ expérience ni une image 3D haute définition ne fournirait la compréhension de la dynamique des micro-fractures qui peut être fournie par une analyse mathématique précise modèle.

3.2.3 Les modèles comme médiateurs (total)

Le dernier compte de ma liste est celui que j'ai appelé "les modèles en tant que médiateurs (total)". Comme le titre l'indique, ce récit est directement influencé par les modèles de Morgan et Morrison en tant que médiateurs (Morrison 2009). Le livre est une défense du rôle médiateur de modèles dans la pratique scientifique, car elle considère que la pratique scientifique n'est ni motivée par des théories, ni purement sur la manipulation directe de phénomènes du monde réel. Dans-

Au lieu de cela, la pratique scientifique a besoin de la médiation de modèles pour réussir à
réalisation de ses objectifs. Une théorie ne peut donc pas être directement appliquée au phénomène,
mais seulement par la médiation d'un modèle ; de même, en pratique expérimentale,
les modèles restituent les données des mesures et des observations sous une forme disponible
à usage scientifique. Dans ce qui suit, je me concentre sur le rôle médiateur des modèles dans la pratique
expérimentale, puisque le partisan de l'approche des modèles comme médiateurs (total) est
plus intéressé par l'analyse de simulations informatiques à la lumière d'expériences. Je vais
laisser ainsi inanalysé le rôle médiateur des modèles dans le contexte de la théorie (voir mon
discussion dans le chapitre précédent).

Selon le promoteur des modèles comme compte médiateur (total), la pratique expérimentale consiste
à obtenir, par la manipulation de phénomènes, des données
qui nous renseigne sur certaines propriétés d'intérêt. Ces données, cependant, sont dans un tel
état brut qu'il est impossible de le considérer comme fiable ou représentatif des propriétés mesurées ou
observées. Au contraire, pour que ces données brutes aient une quelconque utilité scientifique,
il est nécessaire de le traiter davantage en filtrant le bruit, en corrigeant les valeurs, en mettant en œuvre
des techniques de correction des erreurs, etc. Ces techniques de correction sont
conduites par des modèles théoriques et, à ce titre, sont chargées de fiabiliser
données.

La pratique scientifique est alors conçue comme fortement médiatisée par des modèles ; et la
connaissance scientifique ne s'obtient plus uniquement par notre intervention dans le monde,
mais aussi par la médiation conceptuelle que représente le rapport modèle/monde. Dans ce
veine, l'analyse épistémique s'intéresse désormais aux données filtrées, corrigées,
et affiné par des modèles, plutôt que les données brutes collectées en manipulant directement
le vrai monde.

Les simulations informatiques devraient facilement s'intégrer dans cette nouvelle image de la pratique
scientifique. On pourrait penser que puisqu'ils sont conçus comme des modèles mis en œuvre sur le
ordinateur numérique, alors leurs résultats doivent être des données produites par un modèle fiable dans un
sens direct. Malheureusement, ce n'est pas ce que le partisan des modèles comme
(total) médiateurs a à l'esprit. Pour elle, il est juste de dire que les simulations informatiques
sont des modèles fonctionnant sur un ordinateur numérique, et il est également correct de dire qu'il existe
aucune intervention dans le monde au sens empiriste. Néanmoins, les données obtenues en exécutant
une simulation sont «brutes» dans le même sens que les données collectées par un
instrument scientifique.10 La raison en est qu'il existe des caractéristiques matérielles
système cible qui sont modélisés dans la simulation, et donc représentés dans le
données simulées finales (53). Les données simulées doivent donc être post-traitées par un autre modèle
théorique, exactement de la même manière que les données brutes. Autrement dit, simulé
les données doivent également être filtrées, corrigées et affinées par un autre ensemble de modèles afin de
produire des données qui peuvent être utilisées de manière fiable dans la pratique scientifique. Ontologiquement parlant,
alors, il n'y a pas de différences entre les données produites par un instrument scientifique et
données produites par une simulation informatique. Ces résultats donnent raison à l'affirmation
qu'il n'y a pas non plus de différences épistémiques entre ces deux types de données.

---

dix Afin de garder ces deux notions de données distinctes, je continuerai à me référer aux données collectées
par l'instrument scientifique comme des "données brutes", tandis que je ferai référence aux données obtenues en exécutant le
simulation informatique en tant que « données simulées ».

Permettez-moi maintenant d'élaborer un peu plus sur ces points. En 2009, Margaret Morrison publié une contribution fondamentale au débat sur la mesure dans le contexte de simulations informatiques. Dans cet ouvrage, elle affirmait que certains types d'ordinateurs les simulations ont le même statut épistémique que les mesures expérimentales précisément car les deux types de données sont ontologiquement et épistémiquement comparables.

Pour illustrer ce point, considérons brièvement son exemple de mesure du g force.11 Dans une mesure expérimentale, soutient Morrison, un instrument scientifique mesure une propriété physique jusqu'à un certain degré de précision, bien qu'une telle la mesure ne reflétera pas nécessairement une valeur exacte de cette propriété. La différence entre précision et exactitude est d'une importance primordiale pour Morrison ici : alors que le premier est lié à la pratique expérimentale d'intervention dans nature – ou le calcul du modèle dans la simulation – ce dernier est lié à la médiation des modèles en tant que rendu fiable des données. Dans ce contexte, une mesure précise consiste en un ensemble de résultats dans lequel le degré d'incertitude de la valeur estimée est relativement petit (Morrison 2009, 49); d'autre part, une mesure précise consiste en un ensemble de résultats proches de la vraie valeur de la physique mesurée propriété.12

La distinction entre ces deux concepts constitue la pierre angulaire de la stratégie de Morrison : les données recueillies à partir d'instruments expérimentaux ne fournissent mesures de g, alors que des mesures fiables doivent avant tout être des représentations exactes de la valeur mesurée. C'est dans ce contexte que Morrison considère que les données brutes doivent être post-traitées dans la recherche de l'exactitude (pour le cas de la mesure de g, Morrison propose le pendule ponctuel idéal comme modèle).

Du point de vue de Morrison, la fiabilité des données mesurées est donc fonction du niveau de précision, qui dépend d'un modèle théorique plutôt que de l'instrument scientifique - ou sur la simulation informatique.

> Le niveau de sophistication de l'appareil expérimental détermine la précision de la mesure, mais c'est l'analyse des facteurs de correction qui détermine la précision. Dans en d'autres termes, la manière dont les hypothèses de modélisation sont appliquées détermine la précision réelle de la mesure de g. Cette distinction entre précision et exactitude est très importante –un jeu de mesures précis donne une estimation proche de la vraie valeur de la quantité étant mesurée et une mesure précise est celle où l'incertitude dans l'estimation la valeur est petite. Afin de nous assurer que notre mesure de g est précise, nous devons nous fier largement sur les informations fournies par nos techniques/hypothèses de modélisation (49).

Les simulations informatiques, tout comme les instruments scientifiques, partagent le même sort d'être précises mais pas exactes. Pour ces derniers, la précision est adaptée aux contraintes physiques de mesure du monde réel ; pour le premier, la précision se présente sous la forme de contraintes physiques et logiques dans le calcul (par exemple, erreurs d'arrondi, troncature

---

11 Morrison discute également de l'exemple plus sophistiqué de la mesure du spin (Morrison 2009, 51).

12 La différence entre précision et exactitude est encadrée par Franklin dans l'exemple suivant : « une mesure de la vitesse de la lumière, c = (2,000000000 ± 0,000000001)x1010 cm/s est précise mais imprécis, alors qu'une mesure c = (3.0±0.1)x1010 cm/s est plus précise mais a une plus faible précision » (Franklin 1981, 367). Pour plus de détails sur l'exactitude et la précision, voir le chapitre 4.

erreurs, etc.). La dichotomie précision/exactitude s'applique donc à l'informatique
les simulations comme à la mesure expérimentale, rendant les deux pratiques à la fois tologiquement égales
au niveau des données précises, et épistémiquement égales au niveau des
des données exactes. Ainsi compris, l'argument de la matérialité est également présent ici : égal
l'ontologie détermine l'épistémologie égale. Et c'était précisément l'intention qui sous-tendait l'analyse de
Morrison : « le lien entre les modèles et la mesure est ce qui
fournit la base pour traiter certains types de résultats de simulation comme épistémiquement
au même titre que les mesures expérimentales, voire comme mesures elles-mêmes »
(36).

Au final, Morrison applique une philosophie de modélisation et d'expérimentation
sur une philosophie des simulations informatiques. C'est aussi une conséquence du suivi
le principe de matérialité ; c'est-à-dire qu'il n'y a pas d'analyse des simulations informatiques en elles-mêmes,
mais seulement à la lumière d'une philosophie plus familière. En faisant
données brutes et données simulées ontologiquement égales, et le post-traitement un autre
étape épistémique, Morrison applique des techniques de modèles à des simulations informatiques, quelles
que soient les particularités de ces dernières. Avec ce mouvement à l'esprit, Morrison a également
réduit la classe des simulations informatiques à celles qui sont utilisées comme mesure
dispositifs; et ce faisant, elle restreint l'analyse épistémique à ces simulations. La question qui reste ouverte
est de savoir si la stratégie de Morrison fonctionne également
pour toutes sortes de simulations informatiques (c'est-à-dire pour celles utilisées à d'autres fins que
mesure).

3.3 Remarques finales

De nombreux chercheurs utilisent des simulations informatiques comme s'il s'agissait de dispositifs
expérimentaux fiables. Une telle pratique suppose que les simulations soient épistémiquement sur un
à égalité avec l'expérimentation en laboratoire. En d'autres termes, les simulations informatiques rendent à
au moins autant et aussi qualitativement bonne connaissance de l'environnement empirique
monde comme expérimentation de laboratoire standard. Mais c'est une présupposition injustifiée à moins
que des raisons ne soient données qui fondent le pouvoir épistémologique de l'ordinateur.
simulations.

Suite aux discussions de ce chapitre, nous remarquons que la confiance des
chercheur utilisant des simulations informatiques pourrait être affecté par « l'argument de la matérialité ». Cet
argument dit que notre ensemble de connaissances scientifiques est adapté à
et dépend de l'identification des relations causales physiques interagissant dans
monde. Les simulations informatiques sont, pour beaucoup, des systèmes abstraits qui ne représentent que
des phénomènes du monde réel. Il s'ensuit alors que l'expérimentation en laboratoire, la méthode traditionnelle
source qui alimente le corps des croyances scientifiques, fournit toujours la plus fiable
voie pour connaître et comprendre le monde. La conclusion est alors simple : les chercheurs doivent préférer
mener des expériences plutôt que des simulations informatiques, d'autres
choses étant égales.

Mais nous devons nous demander si c'est vraiment le cas. Il y a beaucoup de
exemples où les simulations informatiques sont en fait des sources de connaissances plus fiables.

à l'avant-garde du monde que l'expérimentation traditionnelle. pourquoi est-ce le cas? pourquoi les chercheurs sont-ils si confiants dans l'utilisation des simulations informatiques pour fournir une vue sur le monde ? Ces questions exigent le genre de traitement philosophique sur l'expérimentation et les simulations informatiques que ce chapitre fournit.

Le chapitre a présenté trois points de vue différents sur la façon dont les philosophes comprennent actuellement l'étude épistémologique des simulations informatiques. J'ai montré que les trois utilisent le même raisonnement que le guide pour leur argumentation. j'ai appelé ça rationnel le principe de matérialité, et je le conceptualise comme les philosophes l'engagement dans un récit ontologique des simulations informatiques – et de l'expérimentation – qui détermine l'évaluation de leur pouvoir épistémique.

La conclusion générale est que les philosophes qui acceptent le principe de matérialité sont moins susceptibles de reconnaître ce qui distingue l'épistémologie de l'informatique simulations que ceux qui ne le font pas. La conclusion est modeste et vise à encourager certains changements dans le traitement philosophique des simulations informatiques. Pour exemple, Anouk Barberousse, Sara Franceschelli et Cyrille Imbert (Barberousse, Franceschelli, et Imbert 2009) ont apporté une contribution importante à la notion de données simulées par ordinateur, et Paul Humphreys a poursuivi leurs travaux en analysant plus en détail la notion de données (Humphreys 2013). Un autre excellent exemple est fourni par le rôle de la simulation informatique dans la modélisation climatique menée par Wendy Parker en elle (Parker 2014) et Johannes Lenhard et Eric Winsberg dans (Lenhard et Winsberg 2010).

Malgré ces excellents travaux, il reste encore beaucoup à faire. Selon moi, un domaine de recherche potentiellement fructueux consiste à reconsidérer certains thèmes classiques de la philosophie des sciences à travers le prisme des simulations informatiques. En ce sens, une revue de les notions traditionnelles d'explication, de prédiction, d'exploration, etc. pourraient fonctionner comme point de départ.[13]

De toute évidence, il y a une façon de faire de la philosophie des sciences qui est fortement ancrée sur une enquête empirique illustrée par l'expérimentation. Le principe épistémique directeur est que la source ultime de connaissance est donnée par une interaction avec, et manipulation du monde. Cependant, le succès continu des simulations informatiques remet en question ces principes : premièrement, il y a une tendance croissante à ce que les services représentent plutôt qu'ils n'interviennent dans le monde ; deuxième, informatique les méthodes éloignent les humains du centre de l'entreprise épistémologique (Humphreys 2009, 616). La seule conclusion définitive est que la philosophie l'enquête sur le pouvoir épistémologique des simulations informatiques est une tâche ardue devant.

Les références

Ackermann, Robert. 1989. "Le nouvel expérimentalisme." Le British Journal pour la Philosophie des sciences 40 (2): 185–190.

[13] J'aborde ces entreprises au chapitre 5.

Aristote. 1965. Histoire animale. Presse universitaire de Harvard.

Barberousse, Anouk, Sara Franceschelli et Cyrille Imbert. 2009. "Simulateur informatique ulations comme expériences. 169 (3): 557–574.

Beisbart, Claus. 2017. « Les simulations informatiques sont-elles des expériences ? Et sinon, comment sont ils étaient liés les uns aux autres ? Revue européenne de philosophie des sciences : 1– 34.

Dowe, Phil. 2000. Causalité physique. La presse de l'Universite de Cambridge.

Duran, Juan M. 2013a. "Un bref aperçu de l'étude philosophique des simulations informatiques." Bulletin APA sur la philosophie et les ordinateurs 13 (1): 38–46.

————. 2013b. "L'utilisation de 'l'argument de la matérialité' dans la littérature sur les simulations informatiques." Dans Computer Simulations and the Changing Face of Sci entific Experimentation, édité par Juan M. Duran et Eckhart Arnold, 76–98. Édition des boursiers de Cambridge.

Franklin, Allan. 1981. "Qu'est-ce qui fait une 'bonne' expérience?" Le journal britannique pour la philosophie des sciences 32 (4): 367–374.

Giere, Ronald N. 2009. "La simulation informatique change-t-elle le visage de l'expérimentation?" Études philosophiques 143 (1): 59–62.

Guala, Francesco. 2002. "Modèles, simulations et expériences". En mode basé sur un modèle Reasoning: Science, Technology, Values, édité par L. Magnani et NJ Nersessian, 59–74. Académique Kluwer.

Hanson, Norwood Russel. 1958. Modèles de découverte : une enquête sur les fondements conceptuels de la science. La presse de l'Universite de Cambridge.

Hartmann, Stéphane. 1996. « Le monde en tant que processus : simulations dans le naturel et Sciences sociales." Dans Modélisation et simulation en sciences sociales de la Philosophy of Science Point of View, édité par R. Hegselmann, Ulrich Mueller, et Klaus G. Troitzsch, 77–100. Springer.

Humphreys, Paul W. 2009. "La nouveauté philosophique de la simulation informatique Méthodes. Synthèse 169 (3): 615–626.

————. 2013. « Sur quoi portent les données ? Dans Computer Simulations and the Chang ing Face of Scientific Experimentation, édité par Juan M. Duran et Eckhart Arnold. Édition des boursiers de Cambridge.

Lenhard, Johannes et Eric Winsberg. 2010. "Holisme, retranchement et l'avenir du pluralisme des modèles climatiques." Études en histoire et philosophie des sciences Partie B - Études d'histoire et de philosophie de la physique moderne 41 (3): 253–262. ISSN : 13552198. http://dx.doi.org/10.1016/j.shpsb.2010. 07.001.

Linder, C. 2012. "Une enquête numérique sur le modèle de saturation de déplacement électrique exponentiel dans la fracturation des céramiques piézoélectriques." Technique Mécanique 32:53–69.

Massimi, Michela et Wahid Bhimji. 2015. « Simulations et expériences informatiques : le cas du boson de Higgs ». Etudes d'Histoire et de Philosophie de Science Partie B: Études en histoire et philosophie de la physique moderne 51: 71–81. Consulté le 7 juin 2016.

Mayo, Deborah G. 1994. "Le nouvel expérimentalisme, les hypothèses d'actualité et Apprendre de l'erreur. PSA: Actes de la réunion biennale de l'Association de philosophie des sciences 1: 270–279.

Morgan, Mary S. 2003. "Expériences sans intervention matérielle." Dans The Philosophy of Scientific Experimentation, édité par Hans Radder, 216–235. University of Pittsburgh Press.

————. 2005. "Expériences versus modèles : nouveaux phénomènes, inférence et surprise". Journal of Economic Methodology 12 (2): 317–329.

Morrisson, Margaret. 2009. "Modèles, mesure et simulation informatique : Changer le visage de l'expérimentation. Études philosophiques 143 (1): 33–57.

Parke, Emily C. 2014. "Expériences, simulations et privilège épistémique." Philosophie of Science 81 (4): 516–536.

Parker, Wendy S. 2009. « Est-ce que ça compte vraiment ? Simulations informatiques, expériences et matérialité. Synthèse 169 (3): 483–496.

————. 2014. "Simulation et compréhension dans l'étude du temps et du climat." Perspectives sur la science 22 (3): 336–356.

Weber, Marcel. 2005. Philosophie de la biologie expérimentale. L'université de Cambridge Presse. ISBN : 978-0-511-49859-6.

Winsberg, Éric. 2009. "Un conte de deux méthodes." 169 (3): 575–592.

Chapitre 4

Faire confiance aux simulations informatiques

S'appuyer sur des simulations informatiques et se fier à leurs résultats est essentiel pour l'avenir épistémique de cette nouvelle méthodologie de recherche. Les questions qui nous intéressent dans ce chapitre sont de savoir comment les chercheurs construisent généralement la fiabilité des simulations informatiques ? et qu'est-ce que cela signifierait exactement de faire confiance aux résultats de simulations informatiques ? Lorsque nous tentons de répondre à ces questions, un dilemme se pose. D'une part, il semble qu'une machine ne puisse pas être entièrement fiable dans le sens où elle n'est pas capable de rendre des résultats absolument corrects. Plusieurs choses peuvent mal tourner et se produisent généralement : d'une erreur de calcul systématique à un chercheur imprudent qui trébuche sur un cordon d'alimentation. Il est vrai que les chercheurs développent des méthodes et construisent des infrastructures destinées à accroître la fiabilité des simulations informatiques et de leurs résultats. Cependant, l'exactitude et la précision absolues sont intrinsèquement une chimère dans la science et la technologie.

D'un autre côté, les chercheurs font confiance aux résultats des simulations informatiques – et il est important que nous continuions à renforcer cette confiance – parce qu'ils fournissent des informations correctes (ou à peu près correctes) sur le monde. À l'aide de simulations informatiques, les chercheurs sont en mesure de prédire les états futurs d'un système du monde réel, d'expliquer pourquoi un phénomène donné se produit et d'effectuer une multitude d'activités scientifiques standard et nouvelles.

Ce dilemme met en avant une distinction qui n'est pas toujours explicite dans le langage scientifique et technologique, mais qui est néanmoins centrale pour évaluer la confiance dans les simulations informatiques. La distinction est entre savoir que les résultats sont corrects et comprendre les résultats. Dans le premier cas, les chercheurs savent que les résultats sont corrects parce que la simulation par ordinateur est un processus de calcul fiable.1 Mais cela ne signifie pas que les chercheurs ont également compris les résultats de la simulation par ordinateur. Pour les comprendre, cela signifie être capable de relier les résultats à un corpus plus large de croyances scientifiques, telles que des théories scientifiques, des lois et des principes valables qui régissent la nature. En connaissant et en comprenant les résultats des simulations informatiques, les chercheurs connaissent et comprennent également quelque chose sur le système cible simulé.

---

[1] Puisque j'ai également fait référence aux simulations informatiques en tant que méthodes, nous pourrions également dire qu'il s'agit de méthodes de calcul fiables. Je les utilise indifféremment.

Le plus souvent, les chercheurs se réfèrent simplement aux résultats de leurs simulations en disant « nous faisons confiance à nos résultats » ou « nous faisons confiance à nos simulations informatiques », ce qui signifie qu'ils savent que les résultats sont corrects (ou approximativement corrects) d'un système cible, que ils comprennent pourquoi ils sont corrects (ou approximativement corrects) d'un système cible, ou les deux. L'un des principaux objectifs de ce chapitre est d'éclairer cette distinction et son rôle

dans les simulations informatiques2. Les philosophes soutiennent depuis longtemps que la connaissance et la compréhension sont deux concepts épistémiques distincts et qu'ils doivent donc être traités séparément. Dans le premier sens ci-dessus, les chercheurs font confiance aux simulations informatiques lorsqu'ils savent que les résultats sont une bonne approximation des données réelles mesurées et observées. Dans le second sens, les chercheurs font confiance aux simulations informatiques lorsqu'ils comprennent les résultats et comment ils se rapportent au corpus de croyances scientifiques. La différence peut être éclairée par un exemple simple. On pourrait savoir que 2 + 2 = 4 sans vraiment comprendre l'arithmétique. Une analogie avec des simulations informatiques peut bien sûr également être établie. Les chercheurs peuvent savoir que la trajectoire simulée d'un satellite donné sous contrainte de marée est correcte par rapport à la trajectoire réelle sans comprendre pourquoi les pics de la simulation se produisent (voir figure 1.3).

Compte tenu de cette distinction, deux questions différentes sont soulevées, à savoir, « comment les chercheurs savent-ils que les résultats de la simulation sont corrects pour le système cible » et « quel type de compréhension pourrait être obtenu ? » Pour répondre à la première question, nous devons discuter des méthodes disponibles pour augmenter la fiabilité des simulations informatiques ainsi que des sources d'erreurs et d'opacités qui diminuent une telle fiabilité. Pour répondre à la deuxième question, nous devons aborder certaines des nombreuses fonctions épistémiques offertes par les simulations informatiques. La première question fait donc l'objet de ce chapitre tandis que la deuxième question fait l'objet du chapitre suivant. Commençons par clarifier la distinction entre connaissance et compréhension3.

---

2 Permettez-moi de préciser que l'analyse suivante a des engagements forts à la représentation d'un système cible. La raison d'emprunter cette voie est que la plupart des chercheurs s'intéressent davantage aux simulations informatiques qui implémentent des modèles représentant un système cible. Cependant, un point de vue non représentationnaliste est également possible et souhaitable, c'est-à-dire un point de vue qui admet que des affirmations telles que « les résultats suggèrent une augmentation de la température dans l'Arctique telle que prédite par la théorie » et « les résultats sont cohérents avec les résultats expérimentaux », sont des affirmations solides, au lieu de simplement "les résultats sont corrects du système cible". Ce changement signifie que les simulations informatiques sont des processus fiables bien qu'elles ne représentent pas un système cible.

3 Connaître et comprendre sont des concepts qui expriment nos états épistémologiques et, en un sens, ils peuvent être considérés comme « mentaux ». Si tel est le cas, les neurosciences et la psychologie sont des disciplines mieux préparées à rendre compte de ces concepts. Une autre façon de les analyser consiste à étudier les concepts en eux-mêmes, à montrer leurs hypothèses et leurs conséquences, et à étudier leur structure logique. C'est dans ce dernier sens que les philosophes discutent généralement des concepts de connaissance et de compréhension.

4.1 Connaissance et compréhension

Selon les théories classiques, la connaissance consiste à avoir des raisons de croire à un fait – également appelé « connaissance descriptive » ou « savoir que ». Dans un jargon plus philosophique, connaître quelque chose, c'est avoir une croyance vraie sur ce quelque chose, et être justifié d'avoir une telle croyance4. Les épistémologues, c'est-à-dire les philosophes spécialisés en théorie de la connaissance, énoncent trois conditions générales de la connaissance. Suivant la littérature standard, les schémas suivants s'imposent : un sujet S connaît une proposition p si et seulement si :

(i) p est vrai,
(ii) S croit que p, (iii) S
est fondé à croire que p

Les schémas ci-dessus sont connus sous le nom de 'Justified True Belief' - ou JTB en abrégé –, où la première prémisse représente la vérité, la seconde la croyance et la troisième la justification. Les épistémologues y voient les conditions minimales pour qu'un sujet revendique un savoir.

Reconstruisons maintenant JTB dans le cadre de simulations informatiques. Appelons p la proposition générale « les résultats d'une simulation informatique sont corrects (ou approximativement corrects) du système cible », et S les chercheurs utilisant des simulations informatiques. Il s'ensuit alors que S connaît p si :

(i) il est vrai que les résultats d'une simulation informatique sont corrects (ou approximativement corrects) du système cible, (ii)
le chercheur pense qu'il est vrai que les résultats sont corrects (ou approximativement corrects) du système cible, (iii) le chercheur est
fondé à croire qu'il est vrai que les résultats sont corrects (ou approximativement corrects) du système cible

La condition (i), la condition de vérité, est en grande partie non controversée. La plupart des épistémologues s'accordent à dire que ce qui est faux ne peut pas être connu, et il n'y a donc pas grand-chose à débattre autour de cette condition. Par exemple, il est faux de croire que Jorge L. Borges a écrit Principia Mathematica, ou qu'il est né en Allemagne. C'est un exemple du genre de chose que personne ne revendiquerait – ou ne serait en mesure de revendiquer – comme étant du savoir. De même, aucun chercheur ne prétendrait connaître les résultats de simulations informatiques qui dépendent d'opérations arithmétiques de base telles que a+b = (b+a) +1.

La condition (ii), la condition de croyance, est plus controversée que la condition de vérité, mais encore largement acceptée par les épistémologues. Il stipule essentiellement que pour connaître p, S doit croire en p. Bien qu'il s'agisse d'une affirmation apparemment évidente, elle a reçu plusieurs objections de philosophes qui considèrent que la connaissance sans croyance est également possible (Ichikawa et Steup 2012). Considérons par exemple un quiz où l'étudiant est invité à répondre à plusieurs questions concernant la littérature argentine. Un

---

[4] Il existe de nombreux bons ouvrages philosophiques sur la notion de connaissance. La littérature spécialisée comprend (Steup et Sosa 2005 ; Haddock, Millar et Pritchard 2009 ; Pritchard 2013).

telle question est « où est né Jorge L. Borges ? ». L'étudiante ne fait pas confiance à sa réponse parce qu'elle la considère comme une simple supposition. Pourtant, elle parvient à bien répondre à de nombreuses questions, notamment en disant "Buenos Aires, Argentine". Cet étudiant a-t-il des connaissances sur la littérature argentine ? Selon JTB, elle le fait. Ceci est un exemple évoqué par Colin Radford dans (Radford 1966), et compte comme une belle argumentation philosophique contre JTB.

Or, ni la condition de croyance telle que présentée par les tenants du JTB ni les critiques à son encontre ne nous intéressent ici. Il en est ainsi non seulement en raison de la complexité inhérente du sujet qui nous éloignerait trop de notre sujet principal, mais principalement parce qu'il y a de bonnes raisons de penser qu'il est très peu probable que les chercheurs s'en tirent avec de simples suppositions sur les résultats de simulations informatiques. Premièrement, il serait franchement assez étonnant que quelqu'un puisse deviner les résultats d'une simulation informatique - en fait, dans la section 4.3.2, je m'oppose à cette possibilité. Deuxièmement, il existe des méthodes fiables qui réduisent les possibilités et le besoin de toute chance épistémique quant à l'exactitude des résultats. Je considère alors que la condition de croyance ne nous concerne pas vraiment, et je passe au vrai problème de la simulation informatique, c'est-à-dire la condition (iii), la condition de justification.

L'importance de la condition (iii) est qu'une croyance doit être correctement formée pour être une connaissance. Une croyance peut être vraie et pourtant être une simple supposition chanceuse, ou pire encore, induite. Si je lance une pièce et que je crois sans raison particulière qu'elle tombera sur pile, et si par hasard la pièce tombe sur pile, alors il n'y a aucune base - autre que le hasard - pour dire que ma croyance était vraie. Personne ne peut revendiquer la connaissance sur la base du simple hasard. Considérons maintenant le cas d'un avocat qui emploie des sophismes pour induire un jury dans une croyance donnée au sujet d'un accusé. Le jury peut considérer cette croyance comme vraie, mais si la croyance n'est pas suffisamment fondée, elle ne constitue pas une connaissance et n'est donc pas fondée pour juger une personne (Ichikawa et Steup 2012).

Comment pourrions-nous accomplir la justification dans des simulations informatiques ? Il existe plusieurs théories de la justification trouvées dans la littérature spécialisée qui viennent à notre aide. Ici, je m'intéresse particulièrement à la soi-disant théorie du reliabilisme de la justification. Le reliabilisme, dans sa forme la plus simple, considère qu'une croyance est justifiée dans le cas où elle est produite par un processus fiable, c'est-à-dire un processus qui tend à produire une forte proportion de croyances vraies par rapport aux fausses. Une façon d'interpréter cela dans le contexte des simulations informatiques est de dire que les chercheurs sont fondés à croire que les résultats de leurs simulations sont corrects ou valides par rapport à un système cible parce qu'il existe un processus fiable (c'est-à-dire la simulation informatique) qui, la plupart du temps, produit des résultats exacts et précis plutôt que des résultats inexacts et imprécis5. Le défi consiste maintenant à montrer comment les simulations informatiques peuvent être considérées comme un processus fiable.

Alvin Goldman est le plus éminent défenseur du reliabilisme. Il l'explique de la manière suivante : « la fiabilité consiste en la tendance d'un processus à produire des croyances qui sont vraies plutôt que fausses » (Goldman 1979, 9-10. Emphase dans l'original). Son

---

5 Strictement parlant, p devrait se lire : « les résultats de leurs simulations sont corrects », et donc les chercheurs sont fondés à croire que p est vrai. Pour simplifier, je dirai simplement que les chercheurs ont raison de croire que les résultats de leurs simulations sont corrects. Cette dernière phrase, bien sûr, est considérée comme vraie.

La proposition met en évidence la place d'un processus de formation de croyance dans les étapes vers la connaissance. Considérons, par exemple, les connaissances acquises par un processus de raisonnement, comme l'exécution d'opérations arithmétiques de base. Les processus de raisonnement sont, dans des circonstances normales et dans un ensemble limité d'opérations, hautement fiables. Il n'y a rien d'accidentel dans la vérité d'une croyance que 2 + 2 = 4, ou que l'arbre devant ma fenêtre était là hier et, à moins que quelque chose d'extraordinaire ne se produise, il sera au même endroit demain6. Ainsi, selon le reliabiliste, une croyance produite par un processus de raisonnement se qualifie, la plupart du temps, comme une instance de connaissance.

La question se tourne maintenant vers ce que cela signifie pour un processus d'être fiable et, plus spécifiquement à nos intérêts, ce que cela signifie pour l'analyse des simulations informatiques. Illustrons la première réponse avec un exemple de Goldman :

> Si une bonne tasse d'expresso est produite par une machine à expresso fiable et que cette machine reste à disposition, alors la probabilité que la prochaine tasse d'expresso soit bonne est supérieure à la probabilité que la prochaine tasse d'expresso soit bonne étant donné que la première bonne tasse a été heureusement produite par une machine peu fiable. Si une machine à café fiable produit un bon expresso pour vous aujourd'hui et reste à votre disposition, elle peut normalement vous produire un bon expresso demain. La production fiable d'une bonne tasse d'espresso peut ou non être dans la relation de causalité singulière avec toute bonne tasse d'espresso ultérieure.
>
> Mais la production fiable d'une bonne tasse d'espresso augmente ou améliore la probabilité d'une bonne tasse d'espresso par la suite. Cette amélioration de la probabilité est une propriété précieuse à avoir. (28. Je souligne)

La probabilité est ici interprétée objectivement, c'est-à-dire comme la tendance d'un processus à produire des croyances qui sont vraies plutôt que fausses. L'idée centrale est que si un processus donné est fiable dans une situation, il est très probable que, toutes choses étant égales par ailleurs, le même processus sera fiable dans une situation similaire. Notons que Goldman est très prudent lorsqu'il s'agit d'exiger l'infaillibilité ou la certitude absolue du compte fiabiliste. Au lieu de cela, un compte rendu à long terme de la fréquence ou de la propension à la probabilité fournit l'idée d'une production fiable de café qui augmente la probabilité d'une bonne tasse d'espresso par la suite.

Empruntant à ces idées, nous pouvons maintenant dire que nous sommes fondés à croire que les simulations informatiques sont des processus fiables si les deux conditions suivantes sont remplies :

> (a) Le modèle de simulation est une bonne représentation du système cible empirique ;7 et (b) Le processus de calcul n'introduit pas de distorsions pertinentes, d'erreurs de calcul ou une sorte d'artefact mathématique.

A minima ces deux conditions doivent être réunies pour avoir une simulation informatique fiable, c'est-à-dire une simulation dont les résultats représentent la plupart du temps correctement les phénomènes empiriques. Permettez-moi d'illustrer ce qui se passerait si l'une des conditions ci-dessus n'était pas remplie. Supposons d'abord que la condition (a) ne soit pas remplie, comme c'est le cas de l'utilisation du modèle ptolémaïque pour représenter le mouvement planétaire. Dans ce cas,

---

6 Notons que ces exemples montrent qu'un processus fiable peut être purement cognitif, comme dans un processus de raisonnement ; ou externe à notre esprit, comme le montre l'exemple d'un arbre devant ma fenêtre.

7 Comme mentionné dans la première note de bas de page, nous n'avons pas strictement besoin de représentation. Les simulations informatiques pourraient être fiables pour les cas où elles ne représentent pas, par exemple lorsque le modèle mis en œuvre est bien fondé et qu'il a été correctement mis en œuvre. Je ne discuterai pas de tels cas.

bien que la simulation puisse rendre des résultats corrects, ils ne représentent aucun réel
système planétaire et par conséquent les résultats ne peuvent être considérés comme des connaissances
du mouvement planétaire. Le cas est similaire si la condition (b) n'est pas remplie. Cela signifie
que pendant les étapes de calcul, il y a eu un artefact quelconque conduisant le
simulation pour rendre des résultats incorrects. Dans ce cas, les résultats de la simulation sont
devrait échouer à représenter le mouvement planétaire. La raison en est que les erreurs de calcul
affectent directement et minimisent le degré de précision des résultats.

Dans la section 2.2.1, j'ai décrit avec certains détails les trois niveaux de logiciel informatique ; à
savoir, la spécification, l'algorithme et le processus informatique. moi aussi
affirment que les trois niveaux utilisent des techniques de construction, de langage et de
des méthodes formelles qui rendent les relations entre eux dignes de confiance : il existe des techniques
de construction bien établies basées sur des langages communs et des
des méthodes qui relient la spécification à l'algorithme, et permettent l'implémentation de ce dernier sur
le calculateur numérique. C'est l'ensemble de ces relations qui
faire de la simulation par ordinateur un processus fiable. En d'autres termes, ces trois niveaux de logiciels
sont intimement liés aux deux conditions ci-dessus : la conception du
spécification et l'algorithme remplissent la condition (a), tandis que le calculateur
le processus remplit la condition (b). Il s'ensuit qu'une simulation par ordinateur est un processus fiable
parce que ses constituants (c'est-à-dire la spécification, l'algorithme et l'ordinateur
processus) et le processus de construction et d'exécution d'une simulation reposent, individuellement et
conjointement, sur des méthodes fiables. Enfin, à partir de l'établissement de la fiabilité
d'une simulation informatique, il s'ensuit que nous sommes fondés à croire (c'est-à-dire que nous savons)
que les résultats de la simulation représentent correctement le système cible.

Nous pouvons maintenant assimiler le réalisme de Goldman à notre question sur la connaissance
dans les simulations informatiques : les chercheurs ont raison de croire que les résultats d'un
la simulation par ordinateur sont correctes d'un système cible parce qu'il existe un processus fiable - la
simulation par ordinateur - dont la probabilité que le prochain ensemble de résultats soit
correct est supérieur à la probabilité que le prochain ensemble de résultats soit correct étant donné
que les premiers résultats ont été heureusement produits par un processus peu fiable. En d'autre
mots, les résultats sont dignes de confiance car les simulations informatiques sont des processus fiables
qui produisent, la plupart du temps, des résultats corrects (ou approximativement corrects). Le problème
consiste maintenant à préciser comment fiabiliser les processus des simulations informatiques.
Arrêtons-nous maintenant ici et reprenons ce problème dans la section 4.2 où je discute de certains
des conditions de fiabilité des simulations informatiques. Il est maintenant temps de discuter
compréhension.

Au début de ce chapitre, j'ai mentionné que savoir que 2 + 2 est une opération fiable qui conduit à 4
n'implique pas une compréhension de l'arithmétique. Comprendre, contrairement à savoir, semble
impliquer quelque chose de plus profond et peut-être même
plus précieux qui est de comprendre que quelque chose est le cas.

Pourquoi une analyse sur la compréhension est-elle importante ? La réponse courte est que la
compréhension scientifique est essentiellement une notion épistémique qui implique des activités
scientifiques telles que l'explication, la prédiction et la visualisation de notre monde environnant. Il y a,
cependant, s'accordent généralement à dire que la notion de compréhension est difficile à définir. Nous
disons que nous "comprenons" pourquoi la Terre tourne autour du Soleil, ou que la vitesse

d'une voiture pourrait être mesurée en dérivant la position du corps par rapport au temps.
Mais trouver les conditions dans lesquelles nous comprenons quelque chose est étonnamment plus difficile que de savoir.

Une première caractérisation considère la compréhension comme le processus de constitution d'un corpus cohérent de croyances scientifiques vraies (ou proches des vraies croyances) sur le monde réel. De telles croyances sont vraies (ou proches de la vérité) dans le sens où nos modèles, théories et déclarations sur le monde fournissent des raisons de croire que le monde réel n'est pas susceptible d'être significativement différent (Kitcher 1989, 453).

Naturellement, toutes les croyances scientifiques ne sont pas strictement vraies. Parfois, nous n'avons même pas une compréhension parfaite du fonctionnement de nos théories et modèles scientifiques, et encore moins une compréhension complète de la raison pour laquelle le monde est tel qu'il est. Pour ces raisons, la notion de compréhension doit aussi admettre certaines faussetés. La philosophe Catherine Elgin a inventé un terme adéquat pour ces cas; elle les appelle « faussetés heureuses » comme un moyen de montrer le côté positif d'une théorie de ne pas être strictement vraie. Ces faussetés heureuses sont les idéalisations et les abstractions que les théories et les modèles prétendent. Par exemple, les scientifiques sont très bien conscients qu'aucun gaz réel ne se comporte de la manière décrite par la théorie cinétique des gaz. Cependant, la loi des gaz parfaits rend compte du comportement des gaz en prédisant leur mouvement et en expliquant les propriétés et les relations. Un tel gaz n'existe pas, mais les scientifiques prétendent comprendre le comportement des gaz réels en se référant à la loi des gaz parfaits (c'est-à-dire pour référencer un corpus cohérent de croyances scientifiques) (Catherine Elgin 2007, 39).

Or, bien que tout corpus scientifique de croyances soit truffé de fausses fausses heureuses, cela ne signifie pas que la totalité de notre corpus de croyances est fausse. Un corps cohérent de croyances majoritairement fausses et infondées, comme l'alchimie ou le créationnisme, ne constitue toujours pas une compréhension de la chimie ou de l'origine des êtres, et il ne constitue certainement pas un corpus cohérent de croyances scientifiques. Dans cette veine, la première exigence pour avoir une compréhension du monde est que notre corpus soit majoritairement peuplé de croyances vraies (ou proches de la vérité).

Pris sous cet angle, il est primordial de rendre compte des mécanismes par lesquels les nouvelles croyances sont incorporées dans le corpus général des croyances vraies, c'est-à-dire de la manière dont il est peuplé. Gerhard Schurz et Karel Lambert affirment que « comprendre un phénomène P, c'est savoir comment P s'inscrit dans ses connaissances de base » (Schurz et Lambert 1994 : 66). Elgin fait écho à ces idées lorsqu'elle dit que "la compréhension est avant tout une relation cognitive à un ensemble d'informations assez complet et cohérent".
(Catherine Elgin 2007, 35).

Il existe plusieurs opérations qui permettent aux scientifiques de peupler notre corpus scientifique de croyances. Par exemple, une dérivation mathématique ou logique d'un ensemble d'axiomes incorpore de nouvelles croyances bien fondées dans le corpus de l'arithmétique ou de la logique, les rendant plus cohérents et intégrés. Il y a aussi une dimension pragmatique qui considère que nous incorporons de nouvelles croyances lorsque nous sommes capables d'utiliser notre corpus scientifique de croyances pour une activité épistémique spécifique, comme le raisonnement, le travail sur des hypothèses, etc. Elgin, par exemple, attire l'attention sur le fait que comprendre la géométrie implique que l'on doit être capable de raisonner géométriquement sur de nouveaux problèmes, d'appliquer la perspicacité géométrique dans différents domaines, d'évaluer les limites du raisonnement géométrique pour la tâche à accomplir, etc. suite (C. Elgin 2009, 324).

Ici, je m'intéresse à décrire quatre manières particulières d'incorporer de nouveaux croyances dans le corpus de connaissances scientifiques. Ce sont, à titre d'explication, par la prédiction, par l'exploration d'un modèle et par la visualisation. Pour cela, je montre comment chacune de ces fonctions épistémiques fonctionne comme un processus de mise en cohérence capable d'incorporer de nouvelles croyances dans notre corpus de croyances. Dans certains cas, le processus de population est assez simple. Les philosophes travaillant sur l'explication scientifique, par exemple, ont largement admis que le but de l'explication est précisément de fournir une compréhension de ce qui est étant expliqué. Le philosophe Jaegwon Kim dit que « l'idée d'expliquer quelque chose est inséparable de l'idée de le rendre plus intelligible ; chercher un expliquer quelque chose, c'est chercher à le comprendre, à le rendre intelligible » (Kim 1994, 54). Stephen Grimm, un autre philosophe de l'explication, fait la même chose point avec moins de mots : « comprendre est le but de l'explication » (Grimm 2010). L'explication est donc une force motrice importante pour la compréhension scientifique. Nous pouvons mieux comprendre le monde parce que nous pouvons expliquer pourquoi il fonctionne ainsi fait, et ainsi peupler notre corpus scientifique de croyances. Un compte réussi de l'explication des simulations informatiques doit alors montrer comment rendre la compréhension en simulant un morceau du monde. Un argument similaire est utilisé pour les autres fonctions épistémologiques des simulations informatiques. C'est pourtant le sujet du prochain chapitre.

4.2 Bâtir la confiance

Plus haut, j'ai affirmé que le défi pour affirmer que les résultats d'une simulation informatique sont fiables consiste à montrer qu'ils sont produits par un processus fiable, à savoir la simulation informatique. La question qui nous préoccupe maintenant est donc par quels moyens les chercheurs pourraient-ils renforcer la fiabilité des simulations informatiques ? Au fil des ans, les chercheurs ont mis au point différentes méthodes qui facilitent une telle but. Dans les sections suivantes, je consacre du temps à discuter de ces méthodes et comment ils affectent la confiance du chercheur dans les résultats des simulations informatiques.

4.2.1 Exactitude, précision et étalonnage

Précisons tout d'abord trois termes qui sont essentiels pour évaluer la fiabilité des simulations informatiques. Ce sont l'exactitude, la précision et l'étalonnage. Contemporain Les études ont généré deux principaux corps de recherche. D'une part, il y a théories mathématiques de la mesure dont la principale préoccupation est la mathématique représentation des grandeurs, normalisation, unités et systèmes, et méthodes de détermination des rapports et des grandeurs. D'autre part, il existe des théories philosophiques concernés par les hypothèses méthodologiques, épistémologiques et métaphysiques de mesure. Ici, nous nous intéressons bien sûr à ce dernier.

En sciences, les chercheurs effectuent des mesures d'un certain nombre de quantités d'intérêt qui peuvent être plus ou moins précises de cette quantité. Qu'est-ce que cela signifie? Les comptes rendus traditionnels de la théorie de la mesure supposent que la précision fait référence à l'ensemble de mesures qui fournissent une valeur estimée proche de la valeur réelle de la quantité mesurée. À cet égard, l'exactitude fait référence au fait que la quantité a été correctement déterminée par comparaison. Par exemple, la mesure de la vitesse de la lumière c = (3,0±0,1)x108 m/s est précise dans la mesure où elle est proche de la vraie valeur de la vitesse de la lumière dans le vide, soit 299 792 458 m/s.

Une telle conceptualisation semble juste, sauf qu'elle présuppose la connaissance de la vraie valeur de la vitesse de la lumière. C'est-à-dire que les chercheurs ont accès à la valeur réelle et fixe d'une quantité dans la nature, et que le moyen par lequel nous accédons à une telle valeur est accordé (c'est-à-dire impartial). Malheureusement, ce n'est pas une image réaliste de la mesure standard en science et en ingénierie. Au contraire, il est plus courant de voir des chercheurs s'efforcer d'obtenir une valeur au moyen de techniques et d'instruments de mesure qui introduiront inévitablement un certain biais. Pour ces raisons, les chercheurs supposent qu'ils ne peuvent, au mieux, mesurer que des valeurs approximatives d'une quantité. Par exemple, en mesurant la vitesse de la lumière, les chercheurs idéalisent le milieu dans lequel la lumière se déplace, en l'occurrence le vide. D'autres idéalisations doivent également être en place, telles que la température et la pression, et la stabilité des unités de mesure. La suppression de ces idéalisations impliquerait un scénario plus complexe où, très probablement, les chercheurs ne connaîtront jamais la véritable valeur de la vitesse de la lumière (Teller 2013).

C'est à cause d'arguments philosophiques comme celui-ci que notre monde devient moins certain. Bien sûr, le fait que dans de nombreux contextes les chercheurs ne puissent pas mesurer la vraie valeur d'une quantité ne signifie pas que, pour toutes les disciplines scientifiques et techniques, mesurer la vraie valeur d'une quantité est impossible. En trigonométrie, par exemple, on peut théoriquement connaître la vraie valeur de chaque angle d'un rectangle. Cependant, pour de nombreuses disciplines scientifiques et d'ingénierie, il est vraiment difficile, voire impossible, de parler de mesurer la vraie valeur de quantités (empiriques). Même dans les cas où la mesure d'une valeur est obtenue par des moyens théoriques, il y a des raisons de croire qu'il existe un ensemble d'hypothèses qui imposent des contraintes sur la mesure. Prenons l'exemple de la mesure d'une seconde. « Depuis 1967, précise Eran Tal, la seconde est définie comme la durée d'exactement 9 192 631 770 périodes de rayonnement correspondant à une transition hyperfine du césium 133 à l'état fondamental (BIPM 2006). Cette définition se rapporte à un atome de césium non perturbé à une température de zéro absolu. Étant une description idéalisée d'un type de système atomique, aucun atome de césium réel ne satisfait jamais cette définition » (Tal 2011, 1086). La leçon à retenir est que la mesure d'une vraie quantité - en particulier dans la nature - n'est pas possible, et cela affecte la définition de la précision par comparaison.

Une autre façon d'analyser la précision consiste à s'intéresser de plus près aux pratiques scientifiques et techniques, en particulier à la conception et à l'utilisation que les chercheurs font des instruments. Par exemple, si la conception et l'utilisation d'un thermomètre justifient l'attribution de la température à un objet dans une plage étroite, alors nous pouvons dire que ce thermomètre est précis. Ainsi entendue, une mesure précise suppose un accès complet à la conception de l'instrument, aux lois qui le régissent - en l'occurrence

cas, thermodynamique - conditions externes qui pourraient affecter l'instrument (par exemple, une main qui réchauffe le thermomètre) et les utilisations de l'instrument.

Une troisième voie consiste à considérer qu'une mesure est exacte par rapport à un étalon de mesure donné. Le cas de la mesure de la vitesse de la lumière en est un exemple. S'il est établi en standard que la vitesse de la lumière est de 299 792 458 m/s, alors une mesure de c = (3,0±0,1)x108 m/s est précise dans la mesure où elle est proche de la valeur standard. Le grand avantage de faire dépendre la précision des normes de mesure est que les chercheurs n'ont pas besoin de présupposer une valeur fixe dans la nature, mais plutôt une valeur fixe dans le temps. Cela signifie, à son tour, que différentes périodes de l'histoire de la science et de la technologie donneront naissance à de nouvelles normes de mesure – et modifieront les normes existantes. Un exemple probant est fourni, encore une fois, par la vitesse de la lumière. En 1907, la valeur de c était d'environ 299 710 ± 22 km/s, mais en 1950, le résultat établi était de 299 792,5 ± 3,0 km/s. Chacune de ces mesures, et les nombreuses qui ont suivi, dépendent d'une méthode et d'un instrument différents qui sont généralement adaptés à un moment précis. De ce point de vue, il est absolument indispensable de combiner les méthodes, les instruments et la temporalité socio-technologique dans la conceptualisation de la notion de précision.

Pourtant, une quatrième manière alternative de comprendre la précision est sur de multiples bases comparatives avec d'autres instruments scientifiques et techniques. Ainsi, si le résultat d'un instrument concorde sur la même valeur avec d'autres instruments, de préférence d'un type différent (par exemple, thermomètre à mercure, thermomètre infrarouge, thermomètre numérique), alors il est dit précis (Tal 2011, 1087).

Ce qui compte comme des résultats précis dépend alors des méthodes de mesure, des normes, des instruments et, bien sûr, des communautés scientifiques et techniques. À cet égard, les chercheurs peuvent considérer comme exacts les résultats d'une simulation par ordinateur s'ils concordent étroitement avec la valeur d'une quantité obtenue par référence à d'autres valeurs, à condition qu'elles se situent dans une plage raisonnablement étroite ; par référence à un étalon de mesure, tel que la vitesse de la lumière ou le zéro absolu ; ou en comparant les résultats avec d'autres instruments scientifiques - y compris, dans de nombreux cas, d'autres simulations informatiques. Un bon exemple de ce dernier cas est fourni par Marco Ajelli et son équipe (Ajelli et al. 2010), qui comparent côte à côte les résultats de deux simulations différentes : un modèle stochastique basé sur un agent et un modèle stochastique structuré de métapopulation. Selon les auteurs, « les résultats obtenus montrent que les deux modèles fournissent des profils épidémiques en très bon accord aux niveaux de granularité accessibles par les deux approches » (1).

L'exactitude est souvent liée à la précision, bien qu'une analyse minutieuse montre qu'elles doivent être différenciées. L'analyse des facteurs de correction et des hypothèses du modèle détermine la précision d'une mesure, alors que c'est le niveau de sophistication de l'instrument scientifique qui détermine la précision d'une mesure.
L'exemple classique qui permet de distinguer ces deux éléments est celui du tir à l'arc sur cible. La précision de l'archer est déterminée par la proximité des flèches autour du centre de la cible. La précision de l'archer, quant à elle, correspond à la proximité (ou à la largeur) des flèches dispersées. Plus les flèches sont serrées, plus le

archer. Maintenant, si la distance du centre de la cible est considérée comme trop grande, alors l'archer c'est-à-dire avoir peu de précision.

La distinction entre précision et exactitude est très importante. Un précis la mesure donne une estimation proche de la vraie valeur de la grandeur mesurée. Afin de rendre une mesure précise, les chercheurs doivent s'appuyer largement sur les informations concernant leurs modèles et leur machinerie mathématique. Une mesure précise, en revanche, est une mesure où l'incertitude sur la valeur estimée est raisonnablement petit. Afin de rendre une mesure précise, les chercheurs s'appuient sur la sophistication de leurs instruments. Comme je le montrerai plus tard, dans les simulations informatiques, la précision dépend également des informations sur le modèle du chercheur et sur les mathématiques. machinerie.

En théorie de la mesure, la précision fait référence à l'ensemble de mesures répétées dans des conditions inchangées pour lesquelles l'incertitude des résultats estimés est relativement faible. Afin d'augmenter la précision, il existe un certain nombre de méthodes statistiques qui aident à fournir des résultats plus précis, tels que l'erreur standard et l'écart type. Un bon exemple est encore fourni par la mesure de la vitesse de la lumière; un ensemble de mesures de la vitesse de la lumière c dont l'écart type est environ ± 0,0000807347498766737 est moins précis que celui dont l'écart type est d'environ ± 0,0000008073474987667, car l'incertitude des résultats estimés est plus petit. Ainsi comprise, la sophistication de l'instrument scientifique (par exemple, un laser pour mesurer la vitesse de la lumière) détermine la précision de la mesure de c, alors que la précision de la mesure dépend de l'analyse de la correction facteurs, hypothèses du modèle, etc.

En simulation informatique, la précision est un élément important à prendre en compte. Dans des circonstances typiques, les chercheurs supposent, à juste titre selon moi, que l'ordinateur en tant que la machine physique est un instrument raisonnablement précis. C'est-à-dire que les résultats obtenus à partir du calcul sont répartis dans une plage acceptable (par exemple, dans un distribution normale). Étant donné que les études sur la précision se concentrent également sur les sources d'instruments imprécision, comme les variations incontrôlées de l'équipement, les dysfonctionnements et défaillance générale, il est important de signaler les mêmes sources qui affectent l'ordinateur simulations. L'homologie avec les simulations informatiques peut alors être établie par des facteurs tels que le contrôle de la surchauffe des ordinateurs, les erreurs de périphérique d'E/S et exceptions de mémoire, entre autres.

Outre la couche matérielle, les simulations informatiques impliquent une couche logicielle qui peut également introduire des imprécisions dans le calcul. Habituellement, la quantité limitée de les bits utilisés pour stocker un nombre conduisent à des troncatures et des arrondis connus - et inconnus hors erreurs. Un exemple simple consiste à stocker « sin(0.1) » en simple précision IEEE standard à virgule flottante.8 Si des calculs ultérieurs utilisent ce nombre, le l'erreur a tendance à être amplifiée, ce qui risque de compromettre la précision du calcul. Dans le but de faire la mesure de sin(0.1) plus précise, les chercheurs améliorent généralement leur base matérielle - par exemple, d'une architecture 64 bits à une architecture 128 bits. Bien sûr, cela signifie seulement que l'arrondi a tendance à être plus petit - ou à avoir moins d'impact dans les résultats finaux – mais ne sont pas nécessairement éliminés. C'est une raison pour laquelle, dans

[8] L'exemple fonctionne également pour montrer les inexactitudes introduites en calculant "sin(0.1) dans IEEE virgule flottante simple précision.

Le calcul haute performance, le matériel joue un rôle si important. Malheureusement, les modifications du matériel peuvent être assez coûteuses, de sorte que la précision doit également être traitée au moyen d'une analyse numérique et d'une bonne programmation.

Enfin, le calibrage – également connu sous le nom de « réglage » – fait référence à l'ensemble des méthodes qui permettent d'apporter de petits ajustements aux paramètres du modèle de sorte que le degré d'exactitude et de précision souhaité atteigne une norme de fidélité spécifique. L'étalonnage se produit généralement - mais pas toujours - lorsqu'un modèle comprend des paramètres sur lesquels il existe une grande incertitude, et donc la valeur du paramètre est déterminée en trouvant la meilleure façon d'adapter les résultats aux données disponibles. Le paramètre en question est alors considéré comme un paramètre libre qui peut être "réglé" selon les besoins. L'étalonnage est alors l'activité de recherche et d'ajustement des valeurs du paramètre libre qui sont les meilleures pour rendre compte des données disponibles.

En simulation informatique, Marc Kennedy et Anthony O'Hagan identifient deux formes d'étalonnage. Premièrement, les "entrées d'étalonnage" qui sont des entrées qui prennent des valeurs fixes mais inconnues pour toutes les mesures et observations utilisées pour l'étalonnage.
Ainsi comprises, les entrées d'étalonnage sont les entrées que nous souhaitons connaître à travers le processus d'étalonnage de la simulation. Deuxièmement, les « entrées variables » qui consistent en toutes les autres entrées dont la valeur peut changer pendant l'exécution d'une simulation informatique. Celles-ci décrivent généralement la géométrie et les conditions initiales et aux limites associées à des aspects spécifiques du système cible.9 Pour les deux formes d'étalonnage, il existe une multitude de méthodes disponibles, notamment des estimations de paramètres par régression non linéaire, des méthodes bayésiennes et une analyse de sensibilité pour évaluer la le contenu informatif des données et l'identification des mesures existantes qui dominent le développement du modèle.

Bien que faisant partie de la pratique standard des simulations informatiques, l'étalonnage présente des inconvénients importants. L'une des principales préoccupations liées à l'étalonnage est qu'il pourrait entraîner des données « comptées deux fois ». C'est-à-dire que les données utilisées pour l'étalonnage de la simulation informatique peuvent également être utilisées pour l'évaluation de la précision des résultats d'une simulation informatique. Cette préoccupation, particulièrement présente au sein de la communauté de la simulation climatique, interroge la circularité et les postures d'auto-confirmation :

> certains commentateurs estiment qu'il y a une circularité non scientifique dans certains des arguments fournis par les GCMers [modélisateurs de la circulation générale] ; par exemple, l'affirmation selon laquelle les MCG peuvent produire une bonne simulation s'accorde mal avec le fait que des aspects importants de la simulation reposent sur [. . .] réglage. (Shackley et al. 1998, 170)

Un autre problème avec l'étalonnage est les "incertitudes résiduelles" qui l'accompagnent. Comme nous l'avons vu, le calage consiste à rechercher un ensemble de valeurs des entrées inconnues telles que les données disponibles correspondent le plus possible aux sorties du modèle. Ces valeurs servent à plusieurs fins, et elles sont toutes des estimations des vraies valeurs de ces paramètres. La nature « estimative » de ces entrées entraîne une incertitude résiduelle sur ces entrées. En d'autres termes, l'étalonnage n'élimine pas l'incertitude, il la réduit simplement. Ce fait doit être reconnu dans une analyse ultérieure

9 Pour plus d'informations sur cette question, voir (McFarland et Mahadevan 2008), (Kennedy et O'Hagan 2001) et (Trucano et al. 2006)

du modèle.

Les efforts ci-dessus pour caractériser l'exactitude, la précision et l'étalonnage découlent de considérations sur la fiabilité des simulations informatiques et la confiance ultérieure que les chercheurs ont dans leurs résultats. Si les simulations informatiques sont des processus fiables qui donnent des résultats corrects, nous avons besoin d'un moyen de caractériser ce que serait un résultat "correct". Ainsi, la raison de notre récente discussion. Comme il est d'usage dans ce livre, nous présentons et discutons, même brièvement, des problèmes philosophiques liés aux concepts et aux idées. L'étape suivante consiste à donner des raisons pour dire que les simulations informatiques sont des processus fiables. À cette fin, je propose une brève discussion sur la littérature sur la vérification et la validation comme les deux méthodes les plus importantes pour accorder la fiabilité.

4.2.2 Vérification et validation

Vérification et validation10 sont les noms généraux donnés à une multitude de méthodes utilisées pour accroître la fiabilité des modèles scientifiques et des logiciels informatiques. Comprendre leur rôle s'avère alors essentiel pour l'évaluation, la crédibilité et le pouvoir d'élicitation des résultats des ordinateurs. Désormais, lorsque l'on s'intéresse aux simulations informatiques, les méthodes de vérification et de validation sont adaptées à des tâches spécifiques. Dans ce qui suit, je discuterai d'abord des généralités des méthodes de vérification et de validation et discuterai plus tard en quoi elles augmentent la fiabilité des simulations informatiques.

Afin de rendre compte de la fiabilité des simulations informatiques ainsi que de sanctionner l'exactitude des résultats, les chercheurs ont à leur disposition des procédures formelles (par exemple, pour confirmer la bonne implémentation de la spécification dans l'ordinateur) et des repères (c'est-à-dire, valeurs de référence précises) auxquelles comparer les résultats du calcul (par exemple, avec d'autres sources de données). En vérification, les méthodes formelles sont au centre de la fiabilité des logiciels informatiques, tandis qu'en validation, le benchmarking est responsable de la confirmation des résultats (Oberkampf et Roy 2010, Préface). En d'autres termes, dans les méthodes de vérification, la relation d'intérêt est entre la spécification - y compris le modèle - et le logiciel informatique, alors que dans les méthodes de validation, la relation d'intérêt est entre le calcul et le monde empirique (par exemple, les données expérimentales obtenues en mesurant et méthodes d'observation) (Oberkampf, Trucano et Hirsch 2003)11.

Voici deux définitions largement acceptées et utilisées par la communauté des chercheurs :

> Vérification : le processus consistant à déterminer qu'un modèle informatique représente avec précision le modèle mathématique sous-jacent et sa solution.

[dix] Également appelées « validité interne » et « validité externe », respectivement.

[11] Pour être plus juste avec la proposition générale d'Oberkampf, Truncano, Roy et Hirsch, je dois également mentionner l'analyse de l'incertitude et comment elle se propage tout au long du processus de conception, de programmation et d'exécution de simulations informatiques. Pour un traitement plus philosophique de la vérification et de la validation, ainsi que des exemples concrets, voir (Oreskes, Shrader-Frechette et Belitz 1994 ; Kuppers et Lenhard 2005 ; Hasse et Lenhard 2017).

> Validation : processus consistant à déterminer dans quelle mesure un modèle est une représentation exacte du monde réel du point de vue des utilisations prévues du modèle. (Oberkampf, Trucano et Hirsch 2003)

Maintenant, cette façon de décrire la vérification et la validation est largement acceptée et utilisé, notamment dans de nombreux traitements philosophiques de simulations informatiques. Éric Winsberg, par exemple, considère que "la vérification, [...] est le processus qui consiste à déterminer si le résultat de la simulation se rapproche ou non des vraies solutions à les équations différentielles du modèle original. La validation, quant à elle, est la processus pour déterminer si oui ou non le modèle choisi est une bonne représentation de le système du monde réel aux fins de la simulation » (Winsberg 2010, 19-20). Un autre exemple de philosophe discutant de la vérification et de la validation des simulations informatiques est Margaret Morrison. Bien qu'elle adopte une définition plus large de méthodes de vérification et de validation, et pense même que ces deux méthodes ne sont pas toujours clairement divisible, elle minimise néanmoins la nécessité de méthodes de vérification affirmant que la validation est plus cruciale qu'une méthode d'évaluation de la fiabilité de la simulation informatique (Morrison 2009, 43).

Les communautés scientifiques et techniques, d'autre part, ont un champ d'action plus large et ensemble de définitions plus diversifié à proposer, toutes adaptées aux spécificités des systèmes à l'étude.[12] Discutons-les maintenant séparément et soulignons ce qui est si spécifique sur les simulations informatiques.

4.2.2.1 Vérification

La vérification dans les simulations informatiques consiste à s'assurer que la spécification pour une simulation donnée est correctement mis en œuvre en tant que modèle de simulation. La littérature propose plusieurs méthodes de vérification adaptées aux logiciels informatiques en général, mais il existe deux méthodes particulièrement importantes pour les simulations informatiques, à savoir, vérification du code et vérification du calcul[13]. Leur importance réside dans le fait que les deux les méthodes se concentrent sur l'exactitude de la discrétisation, un élément clé pour la mise en œuvre de modèles mathématiques sous forme de simulations informatiques.

La vérification du code est définie comme le processus consistant à déterminer si les algorithmes numériques sont correctement implémentés dans le code informatique, ainsi qu'à identifier erreurs potentielles dans le logiciel (Oberkampf, Trucano et Hirsch 2003, 32). Dans ce À cet égard, la vérification de code fournit un cadre pour développer et maintenir un code de simulation informatique fiable.

William Oberkampf et Timothy Trucano ont fait valoir qu'il est utile de poursuivre séparer la vérification de code en deux activités, à savoir la vérification d'algorithmes numériques et l'ingénierie de la qualité logicielle. Le but de la vérification des algorithmes numériques est d'examiner l'exactitude mathématique de la mise en œuvre de tous les

---

[12] Voir (Oberkampf et Roy 2010, 21-29) pour une analyse sur la diversité des concepts. Regarde aussi (Salari et Kambiz 2003 ; Sargent 2007 ; Naylor et al. 1967 ; Naylor, Wallace et Sasser 1967).

[13] Également appelée vérification de solution dans (Oberkampf et Roy 2010, 26), et numérique estimation de l'erreur dans (Oberkampf, Trucano et Hirsch 2003, 26).

algorithmes numériques affectant la précision numérique des résultats de la simulation. Le but de cette méthode de vérification est de démontrer que la
les algorithmes mis en œuvre dans le cadre du modèle de simulation sont correctement mis en œuvre et fonctionnant comme prévu (William L. Oberkampf et Timothy G. Trucano 2002, 720).

L'ingénierie de la qualité logicielle, au contraire, met l'accent sur la détermination si le modèle de simulation produit les résultats corrects – ou à peu près corrects. Le Le but de l'ingénierie de la qualité logicielle est de vérifier le modèle de simulation et les résultats de la simulation sur un matériel informatique spécifique, dans un environnement logiciel spécifié - y compris les compilateurs, les bibliothèques, les E/S, etc. Ces procédures de vérification sont principalement utilisé pendant le développement, les tests et la maintenance de la simulation modèle (721).

D'autre part, la vérification des calculs est définie comme la méthode qui empêche trois types d'erreurs : erreur humaine dans la préparation du code, erreur humaine dans l'analyse des résultats, et les erreurs numériques résultant du calcul de la solution discrétisée du modèle de simulation. Une définition pour la vérification des calculs est ce qui suit:

> Vérification du calcul : le processus de détermination de l'exactitude des données d'entrée, la précision numérique de la solution obtenue, et l'exactitude des données de sortie pour un simulation particulière. (Oberkampf, Trucano et Hirsch 2003, 34)

Ainsi comprise, la vérification par calcul est le côté empirique de la vérification. Il est basée sur la comparaison entre les résultats de la simulation et des solutions très précises du modèle scientifique. Dans un sens, la vérification des calculs est similaire à l'évaluation de la validation dans la mesure où les deux comparent les résultats estimés avec les résultats corrects. Il contrôle le plus souvent les taux de convergence spatiale et temporelle, itératifs convergence, indépendance des solutions pour coordonner les transformations, etc. autres procédés (26).

4.2.2.2 Validation

Le processus de validation (aussi appelé test) consiste à montrer que les résultats de la simulation correspondent, plus ou moins fidèlement et précisément, à ceux
obtenu par mesure et observation du système cible. Oberkampf et Tru cano mettent en évidence trois aspects clés de la validation :

> i) quantification de la précision du modèle informatique en comparant ses réponses avec des réponses mesurées expérimentalement,
>
> ii) interpolation ou extrapolation du modèle de calcul aux conditions correspondant à l'utilisation prévue du modèle, et
>
> iii) déterminer si la précision estimée du modèle de calcul, pour les conditions de l'utilisation prévue, satisfait aux exigences de précision spécifiées. (WL Oberkampf et TG Trucano 2008, 724)

Bien que les méthodes de validation viennent naturellement à de nombreux expérimentateurs, puisqu'ils s'attendent à reproduire – par opposition à représenter ou imiter – un morceau du monde14, ils sont plutôt une question complexe dans le contexte des simulations informatiques. Voici quand certaines inquiétudes apparaissent quant à la fiabilité de la validation.

Un problème majeur vient du fait que la plupart des méthodes de validation sont inductives, et on doit donc s'attendre à rencontrer des problèmes typiques d'induction. Le problème ici est que la méthode ne permet de valider un modèle que jusqu'à un certain point, et donc la validation complète est absolument impossible en raison du grand nombre de comparaisons nécessaire - sans parler de l'improbabilité d'avoir tous les résultats possibles à portée de main. C'est la raison pour laquelle la validation est connue, principalement parmi les informaticiens, comme un méthode de détection de la présence d'erreurs, mais non conçue pour établir leur absence.15

Un autre problème est que la validation dépend de la capacité à comparer des résultats de simulation informatique avec des données empiriques. Une telle relation duelle nécessite évidemment la présence des deux, les résultats de la simulation et les données recueillies à partir d'une source empirique. Cela conduit à exclure les nombreuses simulations informatiques pour pour lesquels il n'y a pas de données empiriques correspondantes. En ce sens, la validation n'est qu'un concept approprié pour les cas où une simulation informatique représente un système réel, et non un système possible ou concevable (par exemple, une simulation qui viole une constante de la nature, telle qu'une simulation avec une force gravitationnelle égale à $G = 1mkg^{-1} s^{-2}$).

Cela dit, il est important de mentionner qu'avec l'introduction de simulations informatiques dans des contextes expérimentaux, la validation ne dépend pas exclusivement sur des résultats contrastés avec des données empiriques. Ajelli et son équipe montrent comment il est possible d'exécuter différentes simulations informatiques et d'utiliser leurs résultats pour affirmer la fiabilité pour chacun – dans ce cas, il n'y a pas seulement convergence de résultats, mais aussi de clés (Ajelli et al. 2010).16

La figure 4.1 montre dans un organigramme comment la vérification (vérification du code et vérification des calculs) et les méthodes de validation sont mises en œuvre dans la pratique scientifique courante. Le modèle conceptuel ici est le produit de l'analyse et de l'observation le système physique d'intérêt (c'est-à-dire ce que nous avons appelé le modèle scientifique). En clé applications de la physique computationnelle (telles que la dynamique des fluides computationnelle, la mécanique des solides computationnelle, la dynamique des structures, la physique des ondes de choc et la chimie computationnelle), le modèle conceptuel est dominé par l'ensemble des EDP utilisées pour représentant des grandeurs physiques.

Deux autres types de modèles peuvent être identifiés : un modèle mathématique, à partir duquel le modèle de calcul ou de simulation est créé, et un modèle physique qui, pour simplicité, nous nous identifierons à une expérience (rappelons notre traitement de l'expérimentation dans la section 3). Le modèle de calcul, dans notre propre terminologie, est un programme informatique opérationnel qui implémente le modèle de simulation en tant qu'ordinateur simulation.

---

14 Alors que c'est une affirmation valable pour certaines formes d'expérimentations scientifiques, ce n'est pas le cas pour d'autres comme l'économie et la psychologie.

15 Cette affirmation est largement attribuée à Edsger Dijkstra.

16 A proprement parler, Ajelli et al. effectuent une analyse de robustesse (Weisberg 2013).

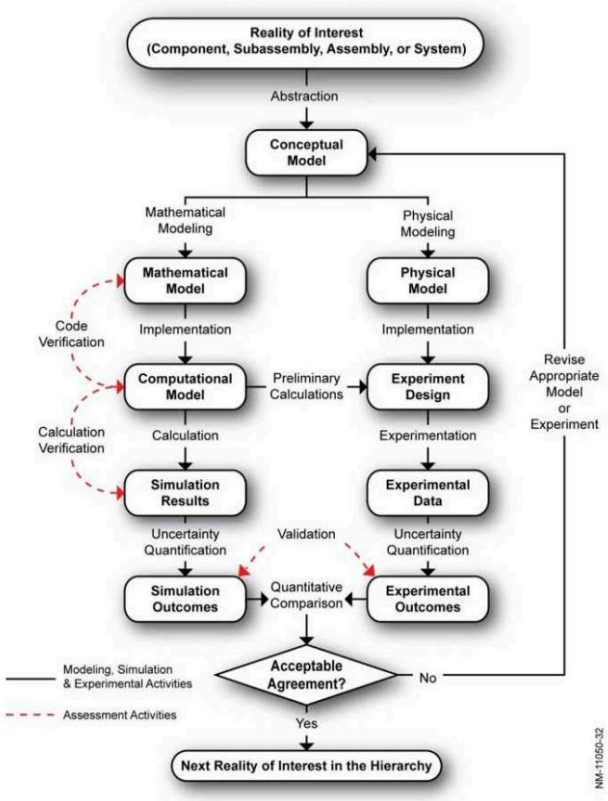


Fig. 4.1 Phases de modélisation et de simulation, et rôle de la vérification et de la validation (ASME 2006, 5).

Notons que la figure montre également que la vérification du code porte sur la fidélité entre le modèle conceptuel et le modèle mathématique, alors que la vérification du calcul porte sur l'accord entre les résultats de la simulation et le résultats attendus du modèle de calcul. La validation, en revanche, est une adéquation quantitative entre les solutions de la simulation et les mesures ou observations expérimentales. Cette adéquation, encore une fois, peut être déterminée par un comparaison qui fournit un accord acceptable entre les solutions des deux modèles impliqués.[17]

Les vérifications et les validations sont deux piliers fondamentaux pour affirmer la fiabilité des simulations informatiques (Duran et Formanek 2018). Cela signifie que les chercheurs ont de bonnes raisons d'affirmer que leurs simulations informatiques sont des processus fiables, et donc qu'ils sont fondés à croire que les résultats des simulation sont correctes.

[17] Pour une discussion plus détaillée sur la figure 4.1, voir (Oberkampf et Roy 2010, 30).

Ce résultat est important car, comme nous le verrons dans le chapitre suivant, la fiabilité des simulations informatiques étaye l'affirmation selon laquelle les chercheurs peuvent prétendre pour l'explication, la prédiction et d'autres activités épistémiques. En d'autres termes, fiable les simulations informatiques permettent également aux chercheurs de revendiquer la compréhension des résultats.

Un dernier et très court commentaire sur la vérification et la validation avant d'introduire le sujet suivant. Bien que ni le logiciel ni le matériel ne puissent être entièrement vérifiés ou validés, les chercheurs développent encore des méthodes qui réduisent la survenue d'erreurs et augmente la crédibilité de la simulation. Cet intérêt sert de preuve pour l'importance des deux méthodes pour la fiabilité générale des simulations informatiques.

4.3 Erreurs et opacité

L'utilisation de certains termes, comme la confiance, la certitude, l'exactitude et la fiabilité ne doit pas donner la fausse impression que l'histoire des simulations informatiques est une histoire de succès. Comme nous le verrons dans les sections suivantes, les simulations informatiques pas un accès totalement transparent et sécurisé au monde, mais plutôt un accès qui est peuplé d'erreurs et d'incertitudes. De manière très générale, chaque discipline dans l'histoire humaine a traité du manque et de la perte de connaissances. En ce sens, les simulations informatiques font également partie de la grande histoire de la science et de la technologie, et il n'y a donc rien de particulièrement nouveau à leur sujet. À un niveau plus local, cependant, plusieurs problèmes émergent strictement de l'utilisation des nouvelles technologies.
– en particulier, les ordinateurs – et les nouvelles méthodes basées sur ces technologies – en notamment les simulations informatiques.

Jusqu'à présent, mes efforts se sont concentrés sur la recherche de simulations informatiques comme sources pour la connaissance du monde. Ce point de vue est bien sûr correct, car les simulations informatiques constituent des méthodes puissantes pour fournir des informations précises sur le monde qui nous entoure. En fait, nous consacrons le chapitre suivant à discuter de la façon dont les simulations informatiques expliquent, prédisent et exécutent plusieurs fonctions épistémologiques. Mais comme mentionné précédemment, les simulations informatiques sont aussi source d'opacités, erreurs et incertitudes qui pourraient affecter leur fiabilité et miner la confiance les chercheurs s'appuient sur les résultats d'une simulation. Pour cette raison, il est important de discuter, même brièvement, des sources d'erreurs et d'opacités qui imprègnent la pratique de simulations informatiques, ainsi que les enjeux philosophiques qui leur sont adaptés. Également important sera de discuter de certains des mécanismes standard utilisés par les chercheurs qui aident à atténuer ces erreurs et à contourner les opacités. Permettez-moi de dire que mon traitement ici sera irrémédiablement injuste compte tenu de la complexité de ces questions. J'espère, cependant, être en mesure d'esquisser les problèmes de base et de suggérer des solutions possibles.

4.3.1 Erreurs

Les erreurs sont un problème séculaire dans les disciplines scientifiques et technologiques. Alors que certains indiquent que quelque chose s'est mal passé, d'autres peuvent en fait offrir un aperçu de ce qui s'est mal passé (Deborah G Mayo 1996). Dans tous les cas, l'importance d'étudier les erreurs est qu'à un moment donné, elles affecteront la fiabilité des simulations informatiques, et donc la confiance du chercheur dans leurs résultats. Pour cette raison, les chercheurs doivent fournir des mesures strictes et robustes pour détecter les erreurs ainsi que des moyens de les prévenir. Lorsqu'elles se produisent, cependant, il est important de savoir comment les traiter, réduire les effets négatifs et les inverser lorsque cela est possible.[18] Une première approche des erreurs les divise en erreurs arbitraires, telles que le produit d'une chute de tension pendant calcul ou un membre du laboratoire négligent qui trébuche sur le cordon d'alimentation de l'ordinateur ; et les erreurs systématiques, c'est-à-dire les erreurs inhérentes au processus de conception, de programmation et de calcul d'un modèle de simulation.

Les erreurs arbitraires ont peu de valeur pour nous. Leur probabilité d'occurrence est très faible, et ils n'apportent aucune contribution à la fréquence d'un processus de calcul pour produire des résultats corrects. Au contraire, ce sont les contributeurs silencieux et solitaires qui, une fois détectés, sont faciles à faire disparaître. Il n'est pas surprenant que les chercheurs mettent peu d'efforts pour essayer de comprendre leur probabilité de se produire. Au lieu de cela, les installations de recherche conçoivent des protocoles et des mesures de sécurité qui aident à y faire face, voire à les éradiquer complètement. Les câbles installés dans des canaux de câbles spéciaux à l'écart de la zone de travail principale sont une mesure de sécurité qui évite avec succès aux membres du laboratoire négligents de trébucher dessus. De même, les sources d'énergie sont très stables de nos jours, et certaines installations ont même des générateurs de secours. En cas de panne générale, les données sont généralement répliquées sur plusieurs serveurs de manière à ce que les simulations informatiques puissent reprendre à partir de la dernière exécution. Ainsi, les erreurs arbitraires ne doivent pas être considérées comme une menace pour la fiabilité générale des simulations informatiques.

Les erreurs systématiques, en revanche, jouent un rôle plus important. Je les divise en deux types, à savoir les erreurs matérielles et les erreurs logicielles. Comme son nom l'indique, le premier type d'erreur est lié au dysfonctionnement de l'ordinateur physique, tandis que le second découle d'erreurs de conception et de programmation de simulations informatiques - y compris des erreurs de packages prêts à l'emploi.[19]

4.3.1.1 Erreurs matérielles

Les erreurs physiques sont liées au dysfonctionnement temporel ou permanent du microprocesseur de l'ordinateur, de la mémoire, des erreurs de périphérique d'E/S et, en général, de tout

---

[18] Pour une excellente analyse des erreurs et de leur incidence sur la pratique scientifique en général, voir (Debo rah G. Mayo 2010 ; Mayo et Spanos 2010). Pour savoir comment les erreurs affectent l'informatique en particulier, voir (Jason 1989). Et pour le rôle des erreurs en informatique, voir (Parker 2008). Je prends ici que les erreurs affectent négativement le calcul.

[19] Pour un aperçu des erreurs dans le cycle de conception et de production des systèmes informatiques, voir (Seibel 2009 ; Fresco et Primiero 2013 ; Floridi, Fresco et Primiero 2015).

composant de l'ordinateur susceptible d'altérer le processus de calcul normal d'une simulation.

De toutes les sources imaginables d'erreurs matérielles, la plus excentrique vient peut-être de l'espace : les supernovae, les trous noirs et d'autres événements cosmiques peuvent en fait faire planter les ordinateurs. Ces erreurs matérielles sont appelées "erreurs logicielles" et sont produites par un phénomène astronomique simple : les rayons cosmiques frappant l'atmosphère terrestre. Lorsque les rayons cosmiques entrent en collision avec des molécules d'air, ils produisent une « pluie d'air » de protons, de neutrons et d'autres particules à haute énergie qui peuvent frapper les composants internes de l'ordinateur. S'ils s'approchent de la mauvaise partie d'une puce, les électrons qu'ils suivent créent un 1 ou un 0 numérique à partir de nulle part (Simonite 2008).

Ces types d'erreurs sont appelés "erreurs logicielles" car leur présence n'entraîne pas de dommages permanents à l'ordinateur, ni de modifications des caractéristiques physiques du matériel. Au contraire, les erreurs logicielles ne corrompent qu'un ou plusieurs bits dans un programme ou une valeur de données, altérant ainsi les données et donc le processus de calcul, sans produire de dommages visibles sur le matériel. Au milieu des années 90, IBM a testé près de 1 000 dispositifs de mémoire à différents niveaux de la mer - vallées, montagnes et grottes - et le résultat a montré que plus l'altitude était élevée, plus les erreurs logicielles se produisaient, alors que le taux d'erreurs logicielles sur les dispositifs testés dans les grottes était presque nulle.

Les erreurs logicielles affectent généralement le système de mémoire de l'ordinateur ainsi que certaines logiques combinatoires utilisées dans les circuits, telles que l'unité logique arithmétique. Dans le cas des circuits mémoire, la source des erreurs logicielles sont les particules énergétiques qui génèrent suffisamment de charge libre pour perturber l'état de la cellule mémoire. Dans le cas de la logique combinatoire, des pointes de tension ou des courants transitoires peuvent perturber la synchronisation de l'horloge provoquant des erreurs logicielles qui se propagent et qui sont finalement verrouillées à la sortie de la chaîne logique (Slayman 2010). Les résultats évidents sont des résultats inconnus et non fiables.

Voici un exemple simple qui illustre à quel point ce type d'erreur est nocif pourrait en fait être. Considérez les deux sous-routines suivantes dans le langage C :

Algorithme 10 Opérateur au niveau du

```
bit si (a & b) puis
    impression "Les démocrates ont gagné" sinon

    printout "Les Républicains ont gagné" end if
```

Algorithme 11 Opérateur logique

```
if (a && b) then
    impression "Les démocrates ont gagné" else

impression "Les républicains ont gagné" end if
```

Dans le langage C, les opérateurs comme & sont appelés au niveau du bit car ce sont des opérations effectuées au niveau du bit en changeant simplement un 1 en 0, et vice-versa. De plus, le langage C n'inclut pas le concept de variable booléenne. Au lieu de cela, "faux" est représenté par 0, et "vrai" peut être représenté par n'importe quelle valeur numérique non égale à 0. Ce fait, généralement bien caché, permet aux programmeurs d'écrire, dans certaines circonstances : a & b tout comme a && b.

Supposons maintenant que a = 4 (000001002) et b = 4 (000001002). Dans ce cas, les conditionnels des deux algorithmes sont évalués comme étant égaux et donc "Les démocrates ont gagné" est la sortie finale. C'est le cas parce que l'opérateur logique && prend des valeurs non nulles (algorithme 11), et donc la première instruction est toujours appelée.
De même, l'opération au niveau du bit ajoute à une valeur non nulle a&b = (000001002) dans (algorithme 10).

Maintenant, si une erreur logicielle se produit sur la 3ème bouchée rendant a = 0 (000000002), alors l'opérateur logique de l'algorithme 11 sera toujours évalué à "Les démocrates ont gagné" simplement parce que c'est une valeur non nulle, alors que l'opérateur au niveau du bit évaluerait à "Les républicains ont gagné" puisqu'il évalue à une valeur nulle a & b = (000000002) (algorithme 10).

L'exemple montre que si la bonne particule énergétique frappe le bon endroit dans la mémoire où la valeur de « a » est stockée, alors les résultats pourraient être très différents – dans ce cas, la démocratie étant la principale victime. On peut soutenir que l'exemple est hautement improbable, tout autant qu'il est malheureux. Les possibilités réelles qu'une telle erreur se produise réellement sont extrêmement faibles, voire pratiquement impossibles. Et ce non seulement parce que les probabilités sont contre, mais aussi parce que les langages de programmation sont devenus très stables et robustes. Malgré ces considérations, il s'agit d'une possibilité réelle que les constructeurs prennent très au sérieux. Au fur et à mesure que les composants matériels réduisent leurs dimensions et leur consommation d'énergie, et que les puces RAM deviennent plus denses, la sensibilité au rayonnement augmente considérablement, et donc la probabilité qu'une erreur logicielle se produise (Baumann 2005).

Pour contrer cet effet, les entreprises technologiques se concentrent sur la conception de meilleures puces et l'amélioration des technologies de vérification des erreurs. En fait, le géant informatique Intel a systématiquement travaillé à l'intégration d'un détecteur de rayons cosmiques intégré dans ses puces. Le détecteur repérerait soit les coups de rayons cosmiques sur les circuits à proximité, soit directement sur le détecteur lui-même. Lorsqu'il est déclenché, il active une série de circuits de vérification des erreurs qui rafraîchissent la mémoire, répètent les procédures les plus récentes et demandent le dernier message envoyé au circuit concerné (Simonite 2008). De cette manière, Intel vise à réduire les erreurs logicielles et donc à augmenter la fiabilité du composant matériel.

Il existe bien sûr d'autres types d'erreurs matérielles systématiques en plus des erreurs logicielles. Ceux-ci viennent généralement en combinaison avec le logiciel qui gère le matériel.
Peut-être que l'erreur matérielle et logicielle la plus célèbre - ou la plus tristement célèbre, devrais-je dire - de l'histoire est attribuée au bogue Pentium FDIV, du microprocesseur Intel. L'objectif d'Intel était de multiplier par 3 l'exécution d'un scalaire à virgule flottante et par 5 fois le code vectoriel, par rapport aux microprocesseurs précédents. L'algorithme utilisé aurait une table de correspondance pour calculer les quotients intermédiaires nécessaires à la division en virgule flottante. Cette table de recherche se composerait de 1066 entiers, 5

dont, en raison d'une erreur de programmation, n'ont pas été téléchargés dans le
tableau logique. Lorsque ces 5 cellules ont été accédées par l'unité à virgule flottante, il serait
récupérer un 0 au lieu du +2 attendu, censé être contenu dans le
cellules "manquantes". Cette erreur a perturbé le calcul et a donné lieu à un calcul moins précis
nombre que la bonne réponse (Halfhill 1995). Malgré le fait que les chances de
le bogue apparaissant au hasard a été calculé à environ 1 sur 360 milliards, le Pentium
Le bogue FDIV a coûté à Intel Co. une perte d'environ 500 millions de dollars de revenus avec le remplacement
des processeurs défectueux.

La leçon à retenir est que la technologie plus avancée n'est pas immédiatement
se traduit par des calculs plus fiables. Des erreurs logicielles apparaissent avec l'introduction
de la carte de circuit imprimé moderne et de la technologie à base de silicium. En vérité, cependant, les erreurs
matérielles sont les moins préoccupantes pour la plupart des chercheurs travaillant sur des simulations
informatiques. C'est principalement parce que, comme nous l'avons vu jusqu'ici, spécifier, programmer,
et l'exécution de simulations informatiques est une pratique pilotée par logiciel. Les chercheurs comptent sur
leur matériel, et ils ont de très bonnes raisons de le faire. De plus, lorsque la question de la fiabilité des
simulations informatiques se pose, la plupart des philosophes, moi-même
inclus, réfléchissent à des moyens de traiter les erreurs logicielles, et non les erreurs matérielles.
Sur cette note, passons maintenant à la discussion des erreurs logicielles.

4.3.1.2 Erreurs logicielles

Les erreurs logicielles sont sans doute la source d'erreurs la plus fréquente en informatique. Ils conduisent à
des instabilités dans le comportement général du logiciel informatique,
et compromettent sérieusement la fiabilité des simulations informatiques.

Les erreurs logicielles peuvent être trouvées dans une myriade de lieux et de pratiques. Par exemple,
la pratique de la programmation est une source principale d'erreurs logicielles, car la programmation peut
devenir extrêmement compliqué. Un compilateur défectueux et un langage informatique défectueux également
mettre en évidence les préoccupations concernant la fiabilité des logiciels informatiques. En outre,
les erreurs de discrétisation sont une source majeure d'erreurs dans les simulations informatiques car elles
proviennent du processus d'interpolation, de différenciation et d'intégration d'une série de
équations mathématiques.

Certaines de ces erreurs peuvent certainement être évitées, mais d'autres s'avèrent plus compliquées.
Une mauvaise programmation, par exemple, peut être imputée à un programmeur maladroit. C
Lawrence Wenham a énuméré plusieurs signes qui font d'un mauvais programmeur (Wenham
2012). Ceux-ci incluent l'incapacité de raisonner sur le code (par exemple, la présence de
'code vaudou', ou code qui n'a aucun effet sur le but du programme mais qui est de toute façon diligentement
maintenu), mauvaise compréhension de la programmation du langage
modèle (par exemple, créer plusieurs versions du même algorithme pour gérer différents
types ou opérateurs), une mauvaise connaissance chronique des fonctionnalités de la plate-forme (par
exemple, réinventer des classes et des fonctions intégrées aux langages de programmation), la
incapacité à comprendre les pointeurs (par exemple, allouer des tableaux arbitrairement grands pour des
collections de longueur variable) et difficulté à voir à travers la récursivité (par exemple, penser que le
nombre d'itérations va être passé en paramètre). La liste s'étend significativement

icément. La programmation, dans tous les cas, est une activité intellectuellement exigeante dans laquelle même le programmeur le plus qualifié et le plus talentueux est susceptible de faire des erreurs.

Notons au passage que ces erreurs logicielles ont en commun le fait qu'elles sont toutes lié à l'humain. Comme indiqué, les erreurs de programmation sont principalement commises par les programmeurs. Un exemple stupide – et pourtant désastreux – est le Mars Climate Orbiter avec lequel La NASA a perdu tout contact près d'un an après son lancement en décembre 1998. Le conseil chargé d'enquêter sur l'accident a conclu que huit facteurs contribuaient à la catastrophe, dont l'un était les modèles informatiques au sol responsables de la navigation de la sonde. Une erreur de programmation a empêché les modèles informatiques de traduire les unités non-SI de livres-secondes (c'est-à-dire les unités anglaises) en newtons secondes métriques SI (unités métriques).20

À ce stade, on pourrait conclure que les erreurs logicielles sont dues à l'homme et par conséquent, éradique avec la formation appropriée. On pourrait penser alors que c'est vrai même pour les cas d'un compilateur défectueux et d'un langage informatique défectueux, car ils sont basé sur de mauvaises spécifications et des procédures de mise en œuvre manquantes (par exemple, une erreur appel), et sont donc également adaptés à l'homme. De plus, même le processus de discrétisation des équations mathématiques est taillé sur mesure, car elles sont encore portées pour la plupart par des mathématiciens – ou des informaticiens, ou des ingénieurs. À En fin de compte, les erreurs logicielles sont des erreurs humaines.

Les choses sont en fait un peu plus complexes qu'ici décrites. Il y a plusieurs des sources d'erreurs qui ne dépendent pas d'habitudes de programmation difficiles à éradiquer – ou des erreurs produites à partir de notre capacité cognitive limitée – mais de la propagation des erreurs par le processus itératif de calcul ; c'est-à-dire le type d'erreurs logicielles qui les ordinateurs, plutôt que les humains, introduisent dans le processus de simulation. Un exemple peut illustrer ce point. Une façon de résoudre les fonctions non linéaires consiste à approximer les résultats avec des méthodes itératives. Si tout se passe bien, c'est-à-dire si la procédure de discrétisation et la programmation a posteriori dans un modèle de simulation sont correctes, alors la ensemble de solutions de la simulation convergent vers la bonne valeur avec une petite marge de erreur.21 Bien qu'il s'agisse d'une pratique courante, dans quelques cas, l'ensemble de solutions sont imprécis en raison d'une accumulation continue d'erreurs lors du calcul. Ces types d'erreurs sont connus sous le nom d'erreurs de convergence, et ils deviennent un sujet d'inquiétude lorsque leur présence passe inaperçue.

Il est bien connu que les deux erreurs de convergence les plus importantes sont l'arrondi erreurs et erreurs de troncature. Typiquement, les premiers sont introduits par la taille du mot de l'ordinateur, d'où une précision limitée des nombres réels. Erreurs de troncature, d'autre part, sont des erreurs faites en raccourcissant une somme infinie à une taille plus petite et en l'approximant par une somme finie.

Afin d'illustrer les effets négatifs potentiels des erreurs d'arrondi dans une simulation informatique, considérons à nouveau l'exemple présenté à la page 11 d'un satellite en orbite autour d'une planète. Là, si l'équation correspondant à la quantité d'énergie totale E (équation 1.1) doit décroître, alors le demi-grand axe a doit devenir

---

[20] Arthur Stephenson, président du comité d'enquête sur l'échec de la mission Mars Climate Orbiter, croyait en fait que c'était la principale cause de la perte de contact avec Mars Climate Orbiter sonde. Voir (Douglas et Savage 1999).

[21] À condition, bien sûr, qu'il n'y ait pas d'erreurs matérielles impliquées.

plus petit. Mais si le moment cinétique H (équation 1.2) doit être constant, alors l'excentricité e doit devenir plus petite. C'est-à-dire que l'orbite doit s'arrondir22. Ainsi expliquée, la tendance de l'excentricité orbitale est régulièrement à la baisse, comme le montre la figure 1.3.

Maintenant, il est bien connu que les chercheurs utilisent des simulations informatiques parce qu'elles sont moins chères, plus rapides et plus faciles à mettre en place que de construire un satellite et de le mettre en orbite. De nombreux philosophes ont même affirmé que, pour ces raisons, les simulations informatiques agissent comme si elles étaient le véritable satellite en orbite autour d'une planète sous l'effet des marées. Je crois que rien n'est plus éloigné de la vérité. Les chercheurs sont bien conscients des limites de leurs simulations et savent que même si les pics montrés sur la figure 1.3 peuvent être attribués à un satellite du monde réel en orbite autour d'une planète du monde réel et ainsi de suite, ils ne peuvent toujours pas attribuer la tendance à la baisse constante qui qu'ils voient dans la même visualisation. Pourquoi pas? Parce que ce n'est pas quelque chose qui se produit réellement dans le monde réel, mais un artefact de la simulation (c'est-à-dire une erreur d'arrondi). Si cet effet devait en fait être attribué au comportement du satellite du monde réel, alors nous verrions l'orbite du satellite devenir circulaire. Mais encore une fois, ce n'est que l'artefact du calcul d'une erreur d'arrondi dans le modèle de simulation (Duran 2017). Woolfson et Pert sont bien sûr bien conscients de ce fait - car ils sont également programmeurs - et peuvent donc prendre les mesures appropriées pour les éviter ou, si cela est inévitable, pour faire face à une telle erreur.

Comme exemple d'erreurs de troncature mesurées, on peut citer la méthode Runge-Kutta qui, soit dit en passant, est également utilisée dans la simulation du satellite en orbite autour de la planète. Les chercheurs estiment que cet algorithme a une erreur de troncature locale de l'ordre de $O(h\ p + 1)$, et une erreur cumulée totale de l'ordre de $nChp+1 = C(x^{-} - x0)h\ p$. Cependant, comme les deux sont des fonctions itératives, la dérivation de l'erreur locale et totale dépendra de chaque itération et il est donc difficile de déterminer l'erreur exacte. Woolfson et Pert incluent dans le sous-programme NBODY de leur simulation le sous-programme Runge-Kutta avec contrôle automatique des pas, et suggèrent aux chercheurs de « surveiller » tout résultat imprévu.

On peut désormais aisément anticiper que la connaissance préalable de l'existence d'une erreur ainsi que le fait de disposer de moyens pour la mesurer constituent des avantages significatifs pour l'évaluation globale des résultats des simulations informatiques. Dans cette simulation particulière, les auteurs soulignent que bien que la simulation d'un satellite en orbite autour d'une planète soit relativement simple à comprendre, certains effets n'auraient peut-être pas été évidents à première vue. Le fait que les pointes se produisent, par exemple, est l'un de ces effets. Selon les auteurs, une bonne simulation apportera toujours des caractéristiques nouvelles, inattendues et importantes du système cible étudié (Woolfson et Pert 1999, 22). A cette réflexion, il faut ajouter que la présence d'erreurs apporte aussi des résultats nouveaux et inattendus que les chercheurs doivent apprendre à gérer.

---

22 C'est l'interprétation de l'échange entre l'énergie et le moment cinétique (Woolfson et Pert 1999, 18). Notons que les auteurs ne parlent pas d'« erreurs », mais uniquement d'arrondir l'orbite. Cet exemple montre également que les erreurs d'arrondi peuvent être interprétées comme faisant partie intégrante de la programmation d'une simulation informatique. Ceci, bien sûr, ne les empêche pas d'être qualifiés d'"erreurs".

4.3.2 Opacité épistémique

La discussion précédente visait à montrer comment les erreurs peuvent contribuer à l'imprécision générale des résultats et ainsi compromettre la fiabilité des simulations informatiques. Tel que présenté, il existe autant de sources d'erreurs que de manières de traiter
avec eux. Tout bien considéré, on peut dire à juste titre que de nombreuses erreurs – mais pas tous, naturellement - sont corrigibles dans une certaine mesure, et donc ils ne sont pas si critiques à la fiabilité des simulations informatiques. Malheureusement, en informatique – et donc dans les simulations informatiques – il existe une source beaucoup plus inquiétante de méfiance que les erreurs, c'est-à-dire l'opacité épistémique.

L'histoire de ce concept remonte bien avant l'utilisation des ordinateurs à des fins scientifiques. Cependant, c'est Paul Humphreys qui a introduit le terme comme un
marque distinctive de l'informatique (Humphreys 2004). Selon lui, une caractéristique essentielle de l'opacité épistémique est que les chercheurs sont incapables,
êtres humains limités, de connaître tous les états pertinents d'un processus de calcul donné à un moment donné. L'argument est assez convaincant. Il dit qu'aucun humain -
ou un groupe d'humains - pourrait éventuellement examiner chaque élément du calcul processus pertinent pour l'évaluation et la justification des résultats. Encore une fois, la « justification des résultats » signifie ici simplement avoir des raisons de croire que les résultats sont
correct. L'opacité épistémique est alors conçue comme la perte irrécupérable de connaissances sur un processus de calcul donné suivie de l'incapacité à justifier la
résultats d'un tel processus (148). Pour caractériser plus formellement l'opacité épistémique, je reproduire la définition de Humphrey :

> un processus est essentiellement épistémiquement opaque à [un agent cognitif] X si et seulement s'il est impossible, étant donné la nature de X, pour X de connaître tous les éléments épistémiquement pertinents du processus (Humphreys 2009, 618).

Décomposons cette caractérisation en ses principales composantes. D'abord, le genre
de processus que Humphreys a à l'esprit est un processus de calcul - comme le calcul d'un modèle de simulation. On pourrait bien sûr se demander s'il existe des
processus également qualifiés d'épistémiquement opaques. Comme suggéré, le concept n'est pas réservé uniquement aux processus informatiques, mais il a plutôt une longue histoire en mathématiques et la sociologie. Plus loin dans cette section, je discute du point de vue de deux mathématiciens et un philosophe qui revendiquent trois formes d'opacité aux racines mathématiques et sociologiques qui affectent aussi les simulations informatiques.

Une autre composante importante de la définition ci-dessus est la notion d'éléments épistémiquement pertinents pour chaque processus. Pour autant que nous sachions, un élément épistémiquement pertinent dans un processus de calcul est toute fonction, variable, pointeur mémoire.
et, d'une manière générale, tout élément intervenant directement ou indirectement dans le calcul du modèle aux fins de rendu des résultats. Enfin, l'agent cognitif X
fait référence à un nombre quelconque de chercheurs impliqués dans un processus épistémiquement opaque. Le nombre de chercheurs n'est bien sûr pas pertinent.

Nous pouvons maintenant reconstruire positivement la caractérisation épistémique de Humphreys. opacité de la manière suivante. Les simulations informatiques sont épistémiquement opaques à tout nombre de chercheurs si et seulement s'il est impossible de connaître l'évolution dans le temps

des variables, des fonctions, etc., dans le processus de calcul. La conséquence de l'opacité épistémique est, encore une fois, que les chercheurs sont incapables de justifier les résultats de leurs simulations.

Ainsi comprise, l'opacité épistémique est un argument solide qui met la pression sur les simulations informatiques en tant que nouvelles méthodes dans les sciences et l'ingénierie. En effet, si les chercheurs ne peuvent pas justifier leurs résultats, quelles sortes de raisons ont-ils de leur faire confiance et, par conséquent, d'utiliser les résultats pour des prédictions et des explications ? Pour illustrer le problème, considérons à nouveau la simulation de la mise en orbite d'un satellite sous contrainte de marée comme discuté dans la section 1.1. Si la simulation est arrêtée à n'importe quelle étape aléatoire, il est au-delà de n'importe quel nombre de chercheurs de reconstruire l'état actuel de la simulation, de rétrodire les états précédents et de prédire les états futurs. Ainsi, les chercheurs sont incapables de justifier les pics illustrés à la figure 1.3 comme le comportement d'un satellite du monde réel sous le stress des marées. Pour autant qu'ils en sachent, les pointes pourraient simplement être du bruit ou un artefact du calcul. L'opacité épistémique donne donc à de nombreux philosophes de bonnes raisons de rejeter l'affirmation selon laquelle les simulations informatiques sont des sources fiables d'informations sur le monde (par exemple, (Guala 2002 ; Parker 2009)).

Pour mettre encore plus en perspective l'opacité épistémique, opposez-la à certaines formes d'erreur. Comme indiqué précédemment, certaines erreurs matérielles peuvent être annulées et entièrement neutralisées en ayant des redondances dans le système, par exemple. Les erreurs logicielles sont, dans de nombreux cas, anticipées grâce à de bonnes pratiques de programmation et à des méthodes de vérification et de validation. Lorsque les erreurs sont uniquement prises en compte, notre manque de connaissances est censé n'être que temporaire. Une fois détectée et modifiée, notre connaissance du processus informatique est restaurée, et avec elle la capacité des chercheurs à justifier les résultats des simulations informatiques. L'opacité épistémique, en revanche, suggère une perte profonde et permanente de connaissances, une incertitude irréversible sur un processus informatique que les chercheurs ne sont pas en mesure de contrôler ou d'inverser. En conséquence, les résultats sont au-delà de toute justification possible.

L'opacité épistémique n'est donc pas une forme d'erreur. C'est clair. On pourrait raisonnablement soutenir, cependant, que l'opacité épistémique nous est familière de la même manière que l'abstraction et les idéalisations nous sont familières. L'argument ici est que tous les trois sont des formes de donner des degrés de détail sur un processus donné (par exemple, un système cible, un processus informatique, etc.), et donc une façon de perdre des connaissances. Mais contrairement à l'opacité épistémologique, les notions d'abstraction et d'idéalisation renvoient à des manières de négliger certains aspects du processus afin d'en enrichir notre connaissance. Les chercheurs font abstraction de la couleur du sable dans le désert du Sahara car elle n'est absolument pas pertinente pour estimer son âge (Kroepelin 2006 ; Schuster 2006). De même, des idéalisations ont lieu dans la reconstruction des effets indirects des aérosols sur les nuages en phase mixte et de glace, car ils ne sont pas inclus dans la simulation utilisée par Benstsen et al.
(Bentsen et al. 2013, 689). Contrairement à l'opacité épistémique, les abstractions et les idéalisations ont donc pour objectif général d'améliorer notre connaissance, et non de la diminuer.
De plus, contrairement à l'abstraction et à l'idéalisation, la présence de l'opacité épistémique est imposée aux chercheurs, plutôt que choisie par eux.

La clé pour comprendre l'opacité épistémique est d'examiner les mathématiques et la façon dont elles gèrent la notion de « surveillabilité » des preuves et des calculs. Les vérités mathématiques telles que les théorèmes, les lemmes, les preuves et les calculs sont en principe contrôlables ; que

c'est-à-dire que les mathématiciens ont un accès cognitif aux équations et aux formules, ainsi qu'aux chaque étape d'une preuve et d'un calcul. Avec l'introduction des ordinateurs, l'arpentabilité en mathématiques devient un peu plus opaque. Un exemple historiquement intéressant qui illustre le type d'angoisse épistémique que produit une telle opacité est la preuve de le théorème des quatre couleurs de Kenneth Appel et Wolfgang Haken (Appel et Haken 1976a, 1976b). Donald MacKenzie rappelle que lorsque Haken a présenté la preuve, le public s'est scindé en deux groupes vers l'âge de quarante ans. Mathématiciens quarante ne pouvaient pas être convaincus qu'un ordinateur pouvait fournir une réponse mathématiquement correcte preuve; et les mathématiciens de moins de quarante ans ne pouvaient être convaincus qu'une preuve a pris 700 pages de calculs manuels pourrait être correct (MacKenzie 2001, 128). Le anecdote montre comment la capacité d'enquête est au centre de la confiance épistémologique dans une méthode mathématique et informatique. En fin de compte, Appel et Hanken avaient fournir des raisons indépendantes pour lesquelles leur programme était fiable et donc rendu des résultats fiables.

Sous ce titre, il est relativement simple de faire le lien entre survérabilité et opacité épistémique : la première empêche la seconde. A l'ère des ordinateurs, cependant, on pourrait à juste titre se demander s'il est nécessaire d'étudier une simulation informatique pour prétendre à la fiabilité et à la confiance. Le but de l'exécution de l'ordinateur simulations semble être, précisément, d'éviter des calculs compliqués en utilisant les machine pour faire le gros du travail. En fait, les implications qui suivent l'opacité épistémique contrastent grossièrement avec le succès des simulations informatiques dans la pratique scientifique et technique.

recherche?

La réponse à cette question a déjà été donnée au début de cette section. Le reliabilisme, comme nous en avons discuté précédemment, est le moyen le plus efficace de contourner opacité épistémique. À la fin de ce chapitre, je montrerai comment la fiabilité aide à cet effort. Mais avant, nous devons aborder toutes les formes imaginables d'opacité pour simulations informatiques.

Dans un article récent, Andreas Kaminski, responsable du département de philosophie à le High Performance Computing Center Stuttgart (HLRS), Michael Resch, directeur du HLRS, et Uwe Kuster, chef du département de méthodes numériques ont présenté trois formes différentes d'opacité qu'ils ont appelées : opacité sociale, opacité technologique et opacité mathématique , qui a un intérieur et un extérieur ..
interprétation (Kaminski, Resch et Kuster 2018). Toutes trois sont des formes d'opacité concernées par la fiabilité des simulations informatiques et la mesure dans laquelle les chercheurs peuvent se fier à leurs résultats. Examinons-les brièvement à tour de rôle24.

---

23 Pour un exemple de simulations informatiques épistémiquement opaques mais réussies, voir (Lenhard 2006).

24 Les idées d'un autre auteur sur l'opacité épistémique et la confiance épistémique qui méritent d'être prises en compte sont celles de Julian Homme nouveau. Pour Newman, l'opacité épistémique est un symptôme que les modélisateurs n'ont pas réussi à adopter pratiques du génie logiciel (Newman 2015). Au lieu de cela, en développant la bonne ingénierie et les bonnes pratiques sociales, soutient Newman, les modélisateurs pourraient éviter plusieurs formes d'opacité épistémique et, finalement, rejettent l'affirmation de Humphreys selon laquelle les ordinateurs sont un autorité épistémique. Comme il le dit explicitement : "[...] un logiciel bien architecturé n'est pas épistémiquement

L'opacité sociale est un autre nom de la division du travail, largement discuté parmi les épistémologues sociaux. Lorsque les projets sont trop complexes, prennent beaucoup de temps ou incluent un grand nombre de participants, la division du travail est la meilleure voie vers le succès. Prenons par exemple la mesure d'une quantité dans la nature. Les physiciens savent généralement comment détecter une telle quantité, quel instrument utiliser et comment analyser les données. Ils pourraient même savoir dans quelle fourchette une telle quantité devrait être attendue, et ce qu'elle signifie pour une théorie physique donnée. Les ingénieurs connaissent peu le travail de mesure et les préoccupations du physicien. Au lieu de cela, ils savent dans les moindres détails comment construire un instrument précis capable de détecter la quantité d'intérêt. Enfin, nous avons les mathématiciens, contributeurs silencieux qui élaborent les mathématiques pour l'instrument et, parfois, pour la théorie physique également. Il s'agit bien sûr d'un cas simplifié et plutôt idéalisé de division du travail. Il s'agit d'illustrer que différents chercheurs collaborent vers un même but, en l'occurrence détecter et mesurer une grandeur dans la nature au moyen d'un instrument précis. La division du travail est une stratégie très efficace qui implique des chercheurs de toutes les disciplines – ainsi que divers chercheurs au sein de la même discipline.

Kaminsky et al. affirment que, dans un contexte de division du travail, les chercheurs connaissent leur propre travail mais pas celui des autres, et qu'ils doivent donc s'appuyer sur des expertises, des solutions et des normes professionnelles qui ne sont pas les leurs (Kaminski, Resch et Kuster 2018, 267). L'exemple est une simulation informatique mettant en œuvre un module lié à une bibliothèque de logiciels. Généralement, ces modules et bibliothèques circulent entre les projets de recherche, les différentes communautés et les techniciens au point que personne, à lui seul, ne connaît tous les détails du module. Il existe une abondante littérature dans les études sociales et technologiques qui étayent leur affirmation : les instruments, les modules informatiques et les artefacts ne sont pas seulement un produit technologique, mais un produit social (Longino 1990). Ainsi comprise, l'opacité sociale est le manque de connaissances que les chercheurs d'une communauté ont sur un produit technologique – ou un changement technologique – d'une autre communauté.

L'opacité technologique, quant à elle, emprunte aux premières idées en mathématiques où les chercheurs utilisent des théorèmes, des lemmes et une multitude de machines mathématiques sans avoir une connaissance spécifique de la preuve formelle qui garantit leur vérité (Kaminski, Resch et Kuster 2018 , 267). L'idée des auteurs est que quelque chose de similaire peut être dit à propos des instruments technologiques. Les chercheurs utilisent une myriade d'instruments indépendamment de toute connaissance approfondie qu'ils peuvent ou non avoir de l'instrument. Par «compréhension profonde», Kaminski et al. désigne toute idée qui va au-delà de la simple connaissance de l'utilisation réussie de l'instrument.

Bien que l'opacité sociale et technologique soient reconnues comme étant des sources pertinentes qui affectent négativement l'évaluation des résultats, les auteurs accordent plus de poids à l'opacité mathématique comme forme centrale d'opacité épistémique pour les simulations informatiques. Dans ce contexte, ils revendiquent deux formes d'opacité mathématique, à savoir une forme internaliste et une forme externaliste (270). L'opacité mathématique interne est conçue comme l'agent cognitif incapable d'examiner le modèle de simulation en raison de

opaque : sa structure modulaire facilitera la réduction des erreurs initiales, la reconnaissance et la correction des erreurs commises, puis l'intégration systématique de nouveaux composants logiciels » (Newman 2015, 257).

sa complexité. En effet, il est extrêmement difficile, voire impossible, d'étudier un modèle de simulation qui inclut des propriétés mathématiques complexes (par exemple, commutatives, distributive, etc.) et les machines de calcul (par exemple, les conditions, la communication E/S, etc.). L'opacité mathématique externe, quant à elle, consiste en ce qu'un agent cognitif est incapable de résoudre le modèle mathématique par ses propres moyens, et donc avoir à l'implémenter sur l'ordinateur. Il s'ensuit que le processus informatique de résolution du modèle n'est plus cognitivement accessible par l'agent. Ainsi compris, l'approche externaliste est très proche des idées d'opacité épistémique présentée par Humphreys.

Dans ce contexte, deux questions viennent à l'esprit. Il faut d'abord se demander dans quelle mesure ces formes d'opacité représentent en effet un problème pour l'appréciation et la justification des résultats. Rappelons que justifier les résultats signifie avoir des raisons à croire que les résultats sont corrects. Dans la mesure où l'opacité épistémique est source de méfiance, il faut se poser la question. Deuxièmement, il y a la question de savoir s'il y a moyens de contourner toute forme d'opacité épistémique. Ma réponse est qu'il y en a. Dans En fait, j'ai déjà présenté une solution au début de ce chapitre. Faites-nous savoir répondre à chaque question à tour de rôle.

Kaminsky et al. ont raison de souligner que les aspects sociaux, technologiques et le point de vue internaliste des mathématiques sont des formes de préoccupation épistémique. Je suis sceptique, cependant, qu'ils représentent un problème pour l'évaluation des résultats de l'informatique. simulations. Mes raisons découlent du fait que Kaminski et al. ne pas rendre explicite ce qui constitue un élément épistémiquement pertinent pour chaque processus. Quand on fait ces éléments unis, il devient clair que ces formes d'opacité ne sont pas nécessairement compromettent la justification des résultats des simulations informatiques. Pour mettre le même point plus précisément, j'identifie deux caractéristiques de ces processus qui les rendent plus épistémiquement "transparent" - mais cette transparence peut être mesurée - et donc pas une réelle menace pour la justification des résultats d'une simulation informatique.

Premièrement, les trois formes d'opacité dépendent de la bonne quantité de description. Généralement, les chercheurs ne s'intéressent qu'à une quantité limitée d'informations qui comptent pour la justification des résultats. Lorsque le bon montant est obtenu, alors ils ont le niveau de transparence recherché. Par exemple, sachant qu'un moteur pseudo-aléatoire module produit les nombres suivants {0,763,0,452,0,345,0,235...} pourrait être moins pertinent épistémiquement pour justifier les résultats de la simulation que de savoir que les résultats sont compris entre $0 < i < 1$. La raison en est que les chercheurs pourraient préférer cette dernière formulation parce qu'elle est suffisamment précise, plus simple dans sa formulation, et mathématiquement plus gérable. Il n'y a pas de raisons intrinsèques qui font que connaissance de chaque nombre pseudo-aléatoire plus pertinente sur le plan épistémique que simplement fournissant une gamme.

La bonne quantité de description est donc un moyen de réduire la pression sociale, les processus technologiques et mathématiques étant épistémiquement opaques. Comme le montre l'exemple de l'opacité mathématique interne, fournir une plage plutôt que chaque la valeur individuelle contribue mieux à la justification des résultats.

Dans cette optique, on pourrait également développer des exemples d'opacité sociale et technologique. Par exemple, de nombreux chercheurs n'ont aucune idée de la façon dont les ordinateurs localisent variables et leurs valeurs dans la mémoire. Cependant, ce fait n'empêche pas pro-

les grammaires de préciser dans leur codage où dans la mémoire une variable doit être
situé. En sachant cela, les chercheurs sont en mesure de justifier pourquoi les résultats ont un certain
erreur de troncation - disons, parce qu'il s'agit d'une mémoire de 8 Mo et de la taille de la valeur stockée
est de 16 Mo. C'est un exemple de la façon dont l'opacité technologique ne signifie pas nécessairement
affecter la justification des résultats.25 En outre, les chercheurs pourraient justifier les résultats
de leurs simulations sans avoir aucune information sur la façon dont les procédures de stockage et de
récupération des valeurs en mémoire ont été conçues et programmées. Autrement dit,
l'opacité sociale n'implique pas non plus l'absence de justification.

Une deuxième caractéristique qui plaide en faveur de la transparence épistémique est le droit
niveau de description d'un processus. C'est l'idée que les éléments épistémiquement pertinents
sont adaptés à la description à différents niveaux du processus. Contrairement au précédent
caractéristiques qui mettent l'accent sur la quantité d'informations, ici l'accent est mis sur la
profondeur d'une quantité donnée d'informations. Ainsi, à de bas niveaux de description, certains processus
sont opaques, alors qu'à certains niveaux plus élevés, ils ne le sont pas. Dans un cadre technologique
processus, par exemple, les chercheurs ne connaissent généralement pas dans les moindres détails tous
les éléments épistémiquement pertinents adaptés à l'instrument, mais cela ne semble guère être un problème.
argument en faveur de l'opacité. Pour illustrer ce point, imaginez un cas fictif où
les chercheurs connaissent chaque détail du fonctionnement d'un ordinateur physique, de la partie
que chaque transistor joue dans l'ordinateur, aux lois physiques impliquées qui permettent
l'ordinateur fonctionne comme il le fait. Vient alors la question : est-ce qu'un chercheur
profiter de cet excès de connaissances ou, au contraire, serait-il un fardeau pour la justifi cation des
résultats ? Il semble assez évident qu'une connaissance approfondie d'un processus pourrait
s'avérer, en fait, contre-productif.

Dans le cas des processus sociaux, par exemple, les chercheurs échangent des idées et des
informations pertinentes avec des collègues sur les décisions de conception et de programmation sur le
fonctionnalité d'un module logiciel. Les processus sociaux ne sont pas des pratiques obscurantistes,
mais elles sont plutôt bien documentées (Latour et Woolgar 2013). Ce point de vue
s'applique également au point de vue internaliste des mathématiques, si nous avions cru l'affirmation
que les mathématiques sont, dans une certaine mesure, un processus social (De Millo, Lipton et Perlis
1979).

Loin d'établir la transparence épistémique des processus sociaux, technologiques et mathématiques
internes, les deux caractéristiques évoquées ci-dessus visent à
inquiétudes quant à la prétendue opacité de ces processus. Ce que Kaminski et al. appeler 'opacité'
est, en fait, une attitude négligente vis-à-vis de ces processus. La division du travail consiste à négliger la
connaissance détaillée du travail des autres chercheurs pour déplacer le nôtre.
avant. Les processus technologiques négligent les informations sur les instruments et les appareils afin de
pouvoir utiliser cette technologie plus efficacement. Et enfin, les processus mathématiques internes utilisent
un principe de négligence similaire, puisqu'ils négligent
des informations sur les étapes spécifiques d'une preuve afin de faciliter l'établissement de
autres vérités mathématiques.

À cet égard, les processus mathématiques sociaux, technologiques et internes négligent
informations afin d'améliorer notre compréhension épistémique. En d'autres termes, ils ne sont pas

25 Humphreys a utilisé un argument similaire pour souligner que les chercheurs n'ont pas besoin de connaître les détails d'un instrument afin de savoir que les résultats d'un tel instrument sont corrects (par exemple, que l'entité observée existe réellement) (Humphreys 2009, 618).

visait à saper la justification des résultats, mais plutôt à renforcer leur évaluation épistémologique. Les chercheurs connaissent ces formes de négligence comme ils les utilisent systématiquement dans des abstractions et des idéalisations. Philosophie standard de la science considère que l'abstraction vise à ignorer les caractéristiques concrètes que possède le système cible afin de se concentrer sur leur configuration formelle (Frigg et Hartmann 2006). Les idéalisations, en revanche, se présentent sous deux formes : tandis que les idéalisations aristotéliciennes consistent à « supprimer » des propriétés que nous estimons non pertinentes pour nos fins, les idéalisations galiléennes impliquent des distorsions délibérées (Weisberg 2013).

La similitude entre les trois formes d'opacité, d'abstraction et d'idéalisation découle, encore une fois, du fait que tous ces processus sont destinés à améliorer notre confiance dans les résultats d'une simulation plutôt que de la saper. Les processus sociaux sont exclusivement conçu pour le succès de la collaboration. On peut dire quelque chose de similaire sur les processus technologiques. La modularisation, par exemple, a été créée pour permettre aux chercheurs de se concentrer sur ce qui est le plus pertinent dans leur travail. Progrès dans la science dépend fortement de ces formes d'opacité, tout autant qu'elle dépend de abstraction et idéalisation.[26]

Pour les raisons données ci-dessus, il semble que nous ne pouvons pas classer le point de vue social, technologique ou internaliste des processus mathématiques comme épistémiquement opaque. au sens donné au début de cette section ; c'est-à-dire que notre perte de connaissances ne peut être annulé, neutralisé ou anticipé. Cela ne veut bien sûr pas dire qu'ils ne constituent pas une question épistémologique en soi. Ils soulèvent des questions importantes concernant la pratique scientifique et technique, mais en principe il y a rien à voir avec le problème de l'opacité épistémique qui nous intéresse ici.

L'opacité mathématique externe, ou simplement l'opacité épistémique, est une situation assez différente. animal. Alors que l'opacité sociale, technologique et mathématique interne place l'humain au centre de leur analyse, dans le contexte de l'opacité épistémique de Humphreys – ou l'opacité mathématique externe de Kaminski et al. – les humains n'ont pas une telle un rôle pertinent. Au lieu de cela, Humphreys se concentre sur le processus de calcul et dans comment il devient épistémiquement opaque. Ainsi compris, les processus informatiques, et pas les humains, sont essentiels pour comprendre l'opacité épistémique. C'est la raison pourquoi Humphreys affirme que les humains ont été déplacés par les ordinateurs du centre de production de connaissances. Les humains, pour reformuler Humphreys, font partie d'un ancien épistémologie.

Nous pourrions maintenant répondre à notre deuxième question, qui vise à considérer les moyens pour contourner l'opacité épistémique.[27]. Fait intéressant, la réponse à cette question peut être remonte au début de ce chapitre, où nous discutons des formes d'octroi fiabilité aux simulations informatiques.[28]

---

[26] Humphreys lui-même établit des parallèles entre les processus sociaux et l'épistémologie sociale, et conclut qu'il n'y a pas de réelle nouveauté dans l'un ou l'autre qui affecterait davantage les simulations informatiques. mesure qu'ils affectent toute autre discipline scientifique, artistique ou technique (619).

[27] En fait, le reliabilisme pourrait être utilisé pour contourner toutes les formes d'opacité v.gr., d'opacité sociale, d'opacité technologique et d'opacité mathématique interne.

[28] J'introduis et discute le reliabilisme dans le contexte des simulations informatiques pour la première fois en (Duran 2014).

Pour récapituler brièvement, rappelez-vous de la section 4 notre discussion sur la façon dont les chercheurs fondé à croire que les résultats d'une simulation informatique sont corrects d'une cible système. Là, nous avons dit qu'il existe un processus fiable - la simulation informatique - dont la probabilité que le prochain ensemble de résultats soit correct est supérieure à la probabilité que la prochaine série de résultats est correcte étant donné que les premiers résultats étaient juste heureusement produit par un processus peu fiable. Une simulation informatique est un processus fiable parce qu'il existe des méthodes de vérification et de validation bien établies qui confèrent confiance dans les résultats.[29] En d'autres termes, notre confiance dans les résultats d'une processus opaque tel qu'une simulation informatique est donné par des processus externes à la simulation elle-même, mais qui fondent leur fiabilité – et qui sont, dans et par eux-mêmes, non opaques.

4.4 Remarques finales

Instaurer la confiance dans les simulations informatiques et leurs résultats n'est pas une mince affaire. À certains, il existe une barrière épistémologique infranchissable imposée par la nature même de l'ordinateur et des processus informatiques qui ne nous permettra jamais, à nous humains, de savoir comment se déroule le processus de simulation. Un tel point de vue permet au affirment que les simulations informatiques ne sont pas aussi fiables que les expérimentations en laboratoire, et donc leur signification épistémologique doit être réduite - une partie de cela a déjà été discutée au chapitre 3. Pour d'autres, moi y compris, nous n'avons pas besoin d'avoir une transparence épistémique totale sur un processus computationnel afin de revendiquer une connaissance. Au contraire, les chercheurs peuvent véritablement savoir quelque chose sur le monde, peu importe de l'opacité impliquée dans la simulation. Bien sûr, il faut encore quelques critères minimaux de ce qui constitue une simulation informatique fiable afin d'avoir toute prétention à la connaissance. L'un de ces critères est garanti par l'enracinement de l'ordinateur simulations en tant que processus fiables.

Tout au long de ce chapitre, il s'agissait simplement de montrer les nombreuses discussions autour de la confiance épistémique dans les simulations informatiques, ainsi que les nombreuses réflexions philosophiques avenues que ces discussions doivent emprunter. Affiner certains concepts nous a aidés à mieux saisir les problèmes sous-jacents, mais malheureusement ce n'est jamais assez. Ici je a pris une position claire, celle qui conçoit que la confiance dans les simulations informatiques peut être accordé pour des raisons épistémiques, et que le reliabilisme est la voie à suivre. Le prochain chapitre suppose une grande partie de ce qui a été dit jusqu'à présent en montrant comment une telle confiance est affichée dans le domaine scientifique et technique. Je les ai appelées "fonctions épistémiques" comme un moyen mettre en évidence les multiples formes de compréhension offertes par les simulations informatiques dans pratique scientifique et technique.

[29] Dans (Duran et Formanek 2018), nous étendons les sources de fiabilité à un historique des simulations informatiques (in) réussies, à l'analyse de robustesse et au rôle d'expert dans la sanction des simulations informatiques.

Les références

Ajelli, Marco, Bruno Gonc¸alves, Duygu Balcan, Vittoria Colizza, Hao Hu, Jose´ J. Ramasco, Stefano Merler et Alessandro Vespignani. 2010. "Comparaison des approches informatiques à grande échelle à la modélisation épidémique : modèles de métapopulation basés sur les agents et modèles structurés." BMC Infectious Diseases 10 (190): 1–13.

Appel, Kenneth et Wolfgang Haken. 1976a. "Chaque carte planaire est quadri colorable." Bulletin de la Société mathématique américaine 82 (5): 711–712.

————. 1976b. "Preuve du théorème des 4 couleurs." Mathématiques discrètes 16 (2): 179–180.

COMME MOI. 2006. Guide pour la vérification et la validation en mécanique computationnelle des solides. L'American Society of Mechanical Engineers, ASME Standard V&V 10-2006.

Baumann, R. 2005. "Erreurs logicielles dans les systèmes informatiques avancés." Conception IEEE Test d'ordinateurs 22, non. 3 (mai): 258–266.

Bentsen, M., I. Bethke, JB Debernard, T. Iversen, A. Kirkevag, ?. Seland, H. Drange, et al. 2013. "Le modèle norvégien du système terrestre, NorESM1-M - Partie 1 : Description et évaluation de base du climat physique." Développement de modèles géoscientifiques 6, no. 3 (mai): 687–720.

De Millo, Richard A., Richard J. Lipton et Alan J. Perlis. 1979. « Processus sociaux et preuves de théorèmes et de programmes ». Communications de l'ACM 22 (5): 271–281.

Douglas, Isbell et Don Savage. 1999. "Le comité d'échec de Mars Climate Orbiter publie un rapport, de nombreuses actions de la NASA sont en cours en réponse." https : / / mars . jpl.nasa.gov/ msp98/news/mco991110.html.

Duran, Juan M. 2014. « Expliquer les phénomènes simulés : une défense du pouvoir épistémique des simulations informatiques ». Thèse de doctorat, Universitat Stuttgart.

————. 2017. "Variétés de simulations : de l'analogique au numérique." In Science and Art of Simulation 2015, édité par M. Resch, Kaminski A. et P. Gehring. Springer.

Duran, Juan M., et Nico Formanek. 2018. "Motifs de confiance: Essential Epistemic Opacité et fiabilité computationnelle. inédit.

Elgin, C. 2009. « La compréhension est-elle factice ? » Dans Epistemic Value, édité par A. Had dock, A. Millar et DH Pritchard, 322–330. Presse universitaire d'Oxford.

Elgin, Catherine. 2007. "La compréhension et les faits." Études philosophiques 132 (1) : 33–42.

Floridi, Luciano, Nir Fresco et Giuseppe Primiero. 2015. « Sur les dysfonctionnements logiciel." Synthèse 192 (4): 1199-1220.

Fresco, Nir et Giuseppe Primiero. 2013. "Erreur de calcul". Philosophie et technologie nology 26 (3): 253–272.

Frigg, Roman et Stephan Hartmann. 2006. "Modèles scientifiques". Dans La Philosophie des sciences. An Encyclopedia, édité par S. Sarkar et J. Pfeifer, 740–749. Routledge.

Goldman, Alvin I. 1979. « Justification et connaissance ». Dans Justification and Knowl edge: New Studies in Epistemology, édité par George Sotiros Pappas, 1–23. Dordrecht : Springer.

Grimm, Stephen R. 2010. "Le but de l'explication." Studies in History and Philosophy of Science 41:337–344.

Guala, Francesco. 2002. "Modèles, simulations et expériences". En mode basé sur un modèle Reasoning: Science, Technology, Values, édité par L. Magnani et NJ Nersessian, 59–74. Académique Kluwer.

Haddock, Adrian, Alan Millar et Duncan Pritchard. 2009. Valeur épistémique. Ox ford University Press.

Halfhill, Tom R. 1995. "La vérité derrière le bogue du Pentium : à quelle fréquence les cinq Cellules vides dans la table de recherche FPU du Pentium Erreur de calcul ? » OCTET (Mars).

Hasse, Hans et Johannes Lenhard. 2017. "Boon and Bane : sur le rôle des paramètres ajustables dans les modèles de simulation." Dans Mathematics as a Tool, édité par Johannes Lenhard et M. Carrier. Boston Studies in the History and Philosophy of the Sciences.

Humphreys, Paul W. 2004. Nous étendre: science computationnelle, Empiri cisme et méthode scientifique. Presse universitaire d'Oxford.

————. 2009. "La nouveauté philosophique des méthodes de simulation par ordinateur." Syn ces 169 (3): 615–626.

Ichikawa, Jonathan Jenkins et Matthias Steup. 2012. "L'analyse des connaissances." Dans The Stanford Encyclopedia of Philosophy, édité par Edward N. Zalta.

Jason, Gary. 1989. "Le rôle de l'erreur en informatique." Philosophie 19 (4): 403–416. ISSN : 0048-3893.

Kaminski, Andreas, Michael Resch et Uwe Kuster. 2018. "Mathematische Opazit at. Uber Rechfertigung und Reproduzierbarkeit in der Computersimulation. Dans Jahrbuch Technikphilosophie, édité par Alexander Friedrich, Petra Gehring, Christoph Hubig, Andreas Kaminski et Alfred Nordmann, 253–278. nomos Verlagsgesellschaft.

Kennedy, Marc C et Anthony O'Hagan. 2001. "Calibrage bayésien de l'ordinateur des modèles." Journal of the Royal Statistical Society: Series B (Statistical Method ology) 63 (3): 425–464.

Kim, Jaegwon. 1994. "Connaissance explicative et dépendance métaphysique."
Questions philosophiques 5 (1994): 51–69.

Kitcher, Philippe. 1989. "Unification explicative et la structure causale de la
Monde." Dans Scientific Explanation, édité par Philip Kitcher et Wesley C.
Saumon, 410–505. Presse de l'Université du Minnesota.

Kroepelin, S. 2006. "Revisiter l'âge du désert du Sahara." 312, non. 5777 (mai):
1138b–1139b.

Kuppers, Gunter et Johannes Lenhard. 2005. "Validation de la simulation : Modèles dans les sciences
sociales et naturelles." Revue des sociétés artificielles et sociales
simulation 8 (4).

Latour, Bruno et Steve Woolgar. 2013. La vie du laboratoire : La construction des sci
faits avérés. Presse universitaire de Princeton.

Lenhard, Johannes. 2006. "Surpris par un nanofil : simulation, contrôle et compréhension."
Philosophie des sciences 73 (5): 605–616. ISSN : 0031-8248.

Longino, Helen E. 1990. Science as social knowledge: Values and objectivity in
enquête scientifique. Presse universitaire de Princeton.

MacKenzie, Donald A. 2001. Mécanisation de la preuve : informatique, risque et confiance à l'intérieur
Technologie. Presse du MIT.

Mayo, Deborah G. 1996. Erreur et croissance des connaissances expérimentales. University of
Chicago Press.

Mayo, Deborah G. 2010. "Apprendre de l'erreur, des tests sévères et de la croissance de
Connaissance théorique." Dans Erreur et Inférence. Échanges récents sur le raisonnement
expérimental, la fiabilité, l'objectivité et la rationalité de la science,
édité par Deborah G. Mayo et Aris Spanos. Presse de l'Université de Chicago.

Mayo, Deborah G. et Aris Spanos, éd. 2010. Erreur et inférence. Changements récents sur le
raisonnement expérimental, la fiabilité, l'objectivité et la rationalité de la science. La presse de
l'Universite de Cambridge.

Mc Farland, John et Sankaran Mahadevan. 2008. "Test de signification multivariée et étalonnage de
modèle sous incertitude." Méthodes informatiques appliquées
Mécanique et ingénierie 197 (29-32): 2467–2479.

Morrisson, Margaret. 2009. "Modèles, mesure et simulation informatique :
Changer le visage de l'expérimentation. Études philosophiques 143 (1): 33–57.

Naylor, Thomas H., JM Finger, James L. McKenney, Williams E. Schrank et
Charles C. Holt. 1967. "Vérification des modèles de simulation informatique." Management
Science 14 (2): 92–106.

Naylor, Thomas H., WH Wallace et WE Sasser. 1967. "Un modèle de simulation informatique de
l'industrie textile." Journal of the American Statistical Association 62 (320): 1338–1364.

Newman, Julien. 2015. « Opacité épistémique, holisme de confirmation et dette technique : La simulation informatique à la lumière du génie logiciel empirique. Dans l'histoire et Philosophie de l'informatique – Troisième conférence internationale, HaPoC 2015, édité par F. Gadducci et M. Tavosanis, 256–272. Springer.

Oberkampf, WL et TG Trucano. 2008. "Repères de vérification et de validation." Ingénierie et conception nucléaires 238 (3): 716–743.

Oberkampf, William L., et Christopher J. Roy. 2010. Vérification et validation en calcul scientifique. La presse de l'Universite de Cambridge.

Oberkampf, William L., et Timothy G. Trucano. 2002. "Vérification et validation en dynamique des fluides computationnelle." Progrès des sciences aérospatiales 38 (3): 209-272.

Oberkampf, William L, Timothy G Trucano et Charles Hirsch. 2003. Verifica tion, Validation, and Predictive Capability in Computational Engineering and La physique. Laboratoires nationaux de Sandia.

Oreskes, Naomi, K Shrader-Frechette et Kenneth Belitz. 1994. « Vérification, validation et confirmation des modèles numériques dans les sciences de la Terre ». Science 263 (5147): 641.

Parker, Wendy S. 2008. "Simulations informatiques à travers une lentille statistique d'erreur." Synthèse 163 (3): 371–384. ISSN : 0039-7857.

————. 2009. « Est-ce que ça compte vraiment ? Simulations informatiques, expériences, et la matérialité. Synthèse 169 (3): 483–496.

Prichard, Duncan. 2013. Qu'est-ce que cette chose appelée connaissance ? Routledge.

Radford, Colin. 1966. "La connaissance par des exemples." Analyse 27 (1): 1–11.

Salari, Patrick et Knupp Kambiz. 2003. Vérification des codes informatiques dans Com science et ingénierie de réputation. Chapman & Hall.

Sargent, Robert G. 2007. « Vérification et validation des modèles de simulation ». Conférence de simulation d'hiver : 124–137.

Schurz, Gerhard et Karel Lambert. 1994. "Esquisse d'une théorie de la compréhension scientifique." Synthèse 101 : 65–120.

Schuster, M. 2006. "L'âge du désert du Sahara." Sciences 311, non. 5762 (février) : 821–821. ISSN : 0036-8075, 1095-9203. doi:10 . 1126 / sciences. 1120161. http://science.sciencemag.org/content/311/ 5762/821.full.

Sebel, Peter. 2009. Codeurs au travail. Apress.

Shackley, Simon, Peter Young, Stuart Parkinson et Brian Wynne. 1998. « Incertitude, complexité et concepts de bonne science dans la modélisation du changement climatique : Les GCM sont-ils les meilleurs outils ? » Changement climatique 38 (2): 159–205.

Simonite, Tom. 2008. "Chaque puce informatique devrait-elle avoir un détecteur de rayons cosmiques?" NouveauScientifique. https://www.newscientist.com/blog/technology/2008/03/do-we-need-cosmic-ray-alerts-for.html .

Slayman, C. 2010. "Erreurs légères - Histoire passée et découvertes récentes." Dans le rapport final de l'atelier international sur la fiabilité intégrée de l'IEEE 2010, 25–30. doi:1 0.1109/IIRW.2010.5706479.

Steup, M., et E. Sosa, éds. 2005. Débats contemporains en épistémologie. Noir Bien.

Tal, Eran. 2011. "Quelle est la précision de la seconde standard?" Philosophie des sciences 78 (5): 1082–1096. ISSN : 0031-8248. doi:10.1086/662268. http://www. jstor.org/stable/info/10.1086/662268%5C%5Cnpapers3 : //publication/doi/10.1086/662268.

Teller, Paul. 2013. "Le concept de mesure-précision." Synthèse 190 (2): 189–202. ISSN : 00397857. doi : 10.1007/s11229-012-0141-8.

Trucano, TG, LP Swiler, T. Igusa, WL Oberkampf et M. Pilch. 2006. « Étalonnage, validation et analyse de sensibilité : de quoi s'agit-il ». Ingénierie de la fiabilité et sécurité du système 91 (10-11): 1331–1357. ISSN : 09518320. doi :10. 1016/j.press.2005.11.031.

Weisberg, Michel. 2013. Simulation et similarité. Presse universitaire d'Oxford.

Wenham, C. Lawrence. 2012. "Signes que vous êtes un mauvais programmeur."

Winsberg, Éric. 2010. La science à l'ère de la simulation informatique. Presse de l'Université de Chicago.

Woolfson, Michael M., et Geoffrey J. Pert. 1999. Une introduction aux simulations informatiques. Presse universitaire d'Oxford.

Chapitre 5
Fonctions épistémiques des simulations informatiques

Le chapitre précédent a fait une distinction entre connaître et comprendre. Dans les simulations informatiques, cette distinction nous permet de distinguer quand les chercheurs font confiance aux résultats et quand ils les comprennent. Dans ce chapitre, nous explorons différentes formes de compréhension à l'aide de simulations informatiques. À cette fin, j'ai divisé le chapitre entre les fonctions épistémiques qui ont une forme linguistique, et celles qui se caractérisent pour avoir une forme non linguistique. Cette distinction vise à mieux aider à catégoriser les différentes manières dont les chercheurs obtiennent une compréhension du monde qui nous entoure au moyen de simulations informatiques. En effet, parfois les simulations informatiques nous ouvrent le monde sous forme de symboles (par exemple, par l'utilisation des mathématiques, du code informatique, de la logique, de la représentation numérique), alors que parfois le monde est accessible par des visualisations et des sons. Dans ce qui suit, j'analyse les études sur l'explication scientifique, les prédictions et les stratégies exploratoires en tant que formes linguistiques qui permettent de comprendre le monde, et la visualisation comme un cas pour les formes non linguistiques.

5.1 Formes linguistiques de compréhension

5.1.1 Force explicative

Toute théorie de l'explication scientifique vise à répondre à la question « pourquoi q ? »1 , où q pourrait être virtuellement n'importe quelle proposition. Considérez les questions pourquoi suivantes : 'pourquoi la fenêtre s'est-elle cassée ?', 'pourquoi le nombre d'élèves qui abandonnent l'école 1 augmente-t-il chaque année ?', 'pourquoi n'est pas défini pour x = 0' dans le contexte de l'infiniesimal classique. calcul? Les chercheurs répondent à ces questions de différentes manières. Prendre pour

[1] Dans certains cas, nous pouvons nous attendre à une explication en posant des questions « comment q ». Par exemple, 'comment le chat a-t-il grimpé à l'arbre ?' est une question qui demande une explication sur la façon dont le chat a réussi à monter sur un arbre. Bien que certaines théories de l'explication placent les questions de comment au centre de leurs préoccupations, nous ne nous intéresserons ici qu'aux questions de pourquoi.

exemple la première question. Un chercheur pourrait expliquer correctement une fenêtre brisée en soulignant qu'une pierre lancée dessus a provoqué la rupture de la fenêtre. Un autre chercheur pourrait avoir une explication faisant appel à la dureté des matériaux – minéraux et verre – et au fait que le premier provoque la rupture du second. Parce que les minéraux qui composent la roche sont plus durs que ceux qui composent le verre, la fenêtre se brisera à chaque fois qu'une pierre lui sera lancée. Un autre chercheur pourrait utiliser comme explication la structure moléculaire des matériaux et montrer comment les propriétés d'une structure provoquent la rupture de l'autre. Quel que soit le niveau de détail utilisé pour l'explication, ils indiquent tous que c'est le rocher qui provoque la rupture de la fenêtre.

Faire appel aux causes pour fournir une explication n'est pas toujours possible ni même approprié. Considérez notre troisième question du pourquoi $\frac{1}{x}$ indéfini pour x = 0 ? Pas de montant des causes peut en fait fournir une bonne explication à cette question du pourquoi. Au lieu de cela, nous devons dériver la réponse d'un ensemble de schémas en utilisant la théorie du calcul différentiel. Un telle explication est la suivante : considérer $\frac{1}{x}$ lorsque x tend vers 0. Considérons maintenant sa limite positive et négative. Ainsi, $\lim_{x \to 0+} \frac{1}{x} = +\infty$, alors que $\lim_{x \to 0-} \frac{1}{x} = -\infty$.

1 Il s'ensuit que $\lim_{x \to 0} \frac{1}{x}$ n'existe pas et n'est donc pas défini.

Les explications ci-dessus visent à illustrer deux éléments de base dans toute théorie de l'explication. Premièrement, les relations explicatives qui permettent de répondre pourquoi les questions peuvent prendre plusieurs formes. Parfois, les chercheurs peuvent expliquer en soulignant les causes qui provoquent q, alors qu'il est parfois préférable de dériver q d'un corpus de croyance scientifique - comme les théories scientifiques, les lois et les modèles.2

L'exigence de base pour l'approche causale est que nous identifions comment q s'inscrit dans le lien causal. C'est-à-dire quelles sont les causes qui provoquent q comme effet. Ainsi, le rocher (c'est-à-dire la cause) a fait casser la vitre (c'est-à-dire l'effet). Le principal défi de toute approche causale est d'énoncer la notion de cause, une question vraiment difficile.
C'est ici qu'un important travail philosophique est nécessaire.

L'alternative à une approche causale consiste à dériver q d'un ensemble de croyances bien établies. Nous avons déjà montré comment cela pouvait être fait dans le cas où 1 expliquant pourquoi d'explication prétendent $\frac{1}{x}$ est indéfini pour x = 0. Fait intéressant, les partisans de ce point de vue également que leur récit pourrait expliquer des cas tels que la fenêtre brisée. À cette fin, les chercheurs doivent reconstruire sous forme de phrases schématiques quelques théories bien connues pertinentes pour l'explication et en déduire le fait que la fenêtre s'est brisée. Certains candidats évidents sont la mécanique newtonienne pour la trajectoire de la roche, la chimie pour les caractéristiques de liaison chimique des verres et la théorie des matériaux pour spécifier le type de verre ainsi que ses propriétés physiques, entre autres.

---

2 Les deux théories majeures de l'explication sont les théories ontiques, où la causalité est au centre et les théories épistémiques, où la dérivation est au centre. Les principaux défenseurs de la première sont (Salmon 1984), (Woodward 2003) et (Craver 2007). Les principaux défenseurs de ce dernier sont (Hempel 1965), (Friedman 1974) et (Kitcher 1989). Le lecteur intéressé par les nombreuses autres théories de l'explication devrait approcher (Salmon 1989).

La deuxième composante de toute théorie de l'explication est la compréhension. Pourquoi les chercheurs devraient-ils poser des questions pourquoi ? Pour quelles raisons les scientifiques et les ingénieurs doivent-ils s'intéresser à expliquer telle ou telle chose ? La réponse est qu'en expliquant, les chercheurs peuvent mieux comprendre pourquoi quelque chose est le cas. Expliquer pourquoi la roche a brisé la fenêtre fait progresser notre compréhension de la physique (par exemple, la trajectoire des projectiles) et de la chimie (par exemple, la résistance des cristaux), ainsi que le simple phénomène d'une fenêtre brisée par une roche.

Pour mettre la question de la compréhension dans une certaine perspective, considérons à quel point l'explication est fondamentale pour rejeter les fausses théories. Un exemple historique intéressant est la théorie du phlogistique, l'opinion prédominante à la fin du XVIIIe siècle sur la combustion chimique. Selon cette théorie, les corps combustibles sont riches en une substance appelée phlogiston qui est libérée dans l'air lors de la combustion. Ainsi, lors de la combustion du bois, du phlogistique est émis dans l'air en laissant des cendres comme résidu. Maintenant, la théorie du phlogistique repose sur deux principes qui s'appliquent à la combustion, à savoir que le corps brûlé perd de la masse et que l'air est « rempli » de cette substance appelée « phlogistique

». La théorie du phlogistique a été mise sous pression lorsqu'elle n'a pas été en mesure d'expliquer comment, lors de la combustion, certains métaux gagnaient en fait de la masse au lieu d'en perdre, violant le premier principe. Dans une tentative de sauver la théorie, certains partisans ont suggéré que le phlogistique avait en fait une masse négative, et donc au lieu d'alléger la masse totale du corps, il le rendrait plus lourd en accord avec la plupart des mesures. Malheureusement, une telle suggestion soulève la question de ce que cela signifie pour le phlogistique d'avoir une masse négative, un concept qui n'était pas responsable par la physique de la fin du XVIIIe siècle. D'autres partisans ont suggéré que le phlogistique émis par ces métaux était, en fait, plus léger que l'air. Cependant, une analyse détaillée basée sur le principe d'Archimède a montré que les densités de magnésium ainsi que sa combustion ne pouvaient pas expliquer une augmentation totale de masse. Aujourd'hui, nous savons que la théorie du phlogistique n'a pas été en mesure d'expliquer l'augmentation de poids de certains métaux lors de la combustion, ce qui en fait une fausse théorie de la combustion.

L'importance de l'explication scientifique pour les simulations informatiques est double. D'une part, elle confère aux simulations une fonction épistémologique principale, à savoir fournir une compréhension de ce qui est simulé. D'autre part, il supprime le simple rôle des simulations en tant que recherche de l'ensemble des solutions à un modèle mathématique insoluble - l'approche standard du point de vue de la résolution de problèmes. Dans ce contexte, trois questions nous intéressent ici. Ce sont, dans l'ordre : « qu'expliquons-nous lorsque nous expliquons avec des simulations informatiques ? », « comment est-il possible d'expliquer des simulations informatiques ? » et enfin « à quel type de compréhension devrions-nous nous attendre en expliquant ? »

Répondre à la première question était plutôt simple : les chercheurs veulent expliquer les phénomènes du monde réel. Tel est le format standard que l'on retrouve dans la plupart des théories de l'explication. Soit une théorie, une hypothèse ou un modèle parmi de nombreuses autres unités d'explication ont la force explicative pour rendre compte d'un phénomène dans le monde.

C'est l'idée avancée par Ulrich Krohş (2008) et Paul Weirich (2011), et interrogé plus tard par moi-même (Duran 2017). Pour ces auteurs, le pouvoir explicatif des simulations informatiques découle du modèle mathématique sous-jacent qui est mis en œuvre sur l'ordinateur capable de rendre compte des phénomènes du monde réel.3 Weirich précise ce point lorsqu'il affirme que « [p]our que la simulation soit explicative, le modèle doit être explicatif » (Weirich 2011, Résumé). De même, Krohs soutient que « dans le triangle du processus du monde réel, du modèle théorique et de la simulation, l'explication du processus du monde réel par la simulation implique un détour par le modèle » (Krohs 2008, 284). Bien que ces auteurs soient en désaccord dans leur interprétation de la façon dont les modèles mathématiques sont mis en œuvre en tant que simulation informatique, et comment ils représentent le système cible, ils conviennent que les simulations ne sont qu'instrumentales dispositifs pour trouver l'ensemble des solutions aux modèles mathématiques. Ainsi compris, les modèles mathématiques implémentés dans la simulation ont la force explicative plutôt que la simulation informatique elle-même.

Mon point de vue diffère de Krohs et Weirich en ce que, pour moi, les chercheurs ont accéder en priorité aux résultats de la simulation, et donc leur intérêt pour l'explication réside dans la comptabilisation de tels résultats. Naturellement, les chercheurs seront même intéressés à comprendre également le monde réel représenté par leur résultats. Cependant, une telle compréhension du monde vient à un stade ultérieur. Pour illustrer ma position, prenons l'exemple des pointes de la figure 1.3. Les chercheurs ont accéder aux pointes de la visualisation, et c'est ce qu'ils veulent expliquer. Le la question qui leur est alors posée est "pourquoi les pointes se produisent-elles" et "pourquoi y a-t-il une tendance à la baisse ? L'importance d'expliquer les résultats de la simulation est que les chercheurs sont en mesure d'expliquer également les vrais pics, avaient un vrai satellite, planète, distance, force de marée, etc. comme spécifié dans le modèle de simulation dans l'espace. Ainsi, en exécutant une simulation informatique fiable et en expliquant ses résultats, les chercheurs peuvent expliquer pourquoi certains phénomènes dans le monde réel se produisent, comme le montre la simulation, sans s'engager réellement dans une quelconque interaction avec le monde lui-même. L'explication dans les simulations informatiques est une caractéristique cruciale qui met en évidence leur pouvoir épistémologique, indépendamment des comparaisons avec des modèles scientifiques ou de l'expérimentation. De plus, comme je le montrerai plus loin dans cette section, la seule unité capable de comptabiliser pour les résultats est le modèle de simulation, par opposition au modèle mathématique qui Krohs et Weirich prétendent.

En résumé, Krohs et Weirich considèrent que les modèles mathématiques mis en œuvre dans la simulation informatique a une force explicative, alors que je soutiens que c'est le modèle de simulation l'unité d'analyse qui devrait en fait tenir ce rôle. En outre, Krohs et Weirich pensent que l'explication est celle d'un phénomène du monde réel, alors que Je prétends que les chercheurs sont d'abord intéressés à expliquer les résultats de la simulation, et plus tard le phénomène du monde réel qu'ils représentent.

La question suivante est de savoir comment une explication scientifique des simulations informatiques est-elle possible ? Pour répondre à cette question, il faut se reporter au début de ce section. Là, j'ai mentionné deux approches principales de l'explication scientifique, à savoir,

---

3 Rappelez-vous notre discussion sur les simulations informatiques comme techniques de résolution de problèmes dans la section 1.1.1.

l'approche causale qui nécessite de situer q dans le lien causal, et l'explication inférentielle qui enracine l'explication par dérivation de q à partir d'un ensemble de croyances scientifiques.

Pour illustrer cette terminologie grossière, utilisons à nouveau l'exemple du satellite sous contrainte de marée et expliquons pourquoi les pics de la figure 1.3 se produisent. L'explication causale consiste à montrer que, comme condition initiale, la position du satellite est à sa distance la plus éloignée de la planète, donc les pointes ne se produisent que lorsqu'elles sont au plus près. Lorsque cela se produit, le satellite est étiré, ce qui est causé par la force de marée exercée par la planète. De même, l'inertie fait que le renflement de marée du satellite est en retard sur le rayon vecteur. Le décalage et l'avance dans le renflement de marée du satellite donnent un moment cinétique de spin à l'approche et le soustraient à la récession. Lorsqu'il s'éloigne du point proche, le renflement de la marée est en avance sur le rayon vecteur et l'effet est donc inversé. Les pointes sont alors causées par l'échange entre le spin et le moment cinétique orbital autour de l'approche la plus proche (voir (Woolfson et Pert 1999, 21)). Le récit inférentiel, au lieu de cela, reconstruirait d'abord le modèle de simulation sous forme de phrases schématiques - mettant en œuvre, entre autres, la mécanique newtonienne - et montrerait ensuite comment une dérivation des pointes, comme indiqué dans la visualisation, a lieu.

Bien que les deux explications semblent valables, il existe une différence fondamentale qui les distingue. Alors que dans l'approche causale la relation explicative dépend d'une relation externe objective (c'est-à-dire des relations causales), dans l'approche inférentielle l'explication est quantifiée sur l'ensemble des connaissances scientifiques actuelles et des croyances établies. Parce que les simulations informatiques sont des entités abstraites, tout comme les mathématiques et la logique, il est assez naturel de penser qu'une explication des pics dépend d'un corpus de croyances scientifiques (c'est-à-dire la simulation informatique) plutôt que de relations causales exogènes.4 Dans (Duran 2017) Je plaide en faveur

de la première position.5 À ce stade, on pourrait être tenté de supposer que, si le modèle de simulation explique les résultats et que les résultats sont le sous-produit du calcul du modèle de simulation, alors il doit y avoir une sorte de circularité argumentative entre ce que les chercheurs veulent expliquer (c'est-à-dire les résultats de la simulation informatique) et l'unité explicative (c'est-à-dire le modèle de simulation) ? Pour aborder cette question, rappelons la distinction entre connaître et comprendre introduite précédemment. Il est important de ne pas confondre ce que les chercheurs savent des résultats de la simulation, c'est-à-dire que le modèle de simulation tient compte du modèle et rend les résultats, avec ce que le chercheur comprend des résultats. Les chercheurs expliquent parce qu'ils veulent comprendre quelque chose. Expliquer les résultats d'une simulation les aide à comprendre pourquoi un résultat s'est produit, indépendamment de la connaissance du comment et du fait qu'il s'est produit. Prenons une fois de plus l'exemple de la raison pour laquelle les pics de la figure 1.3 se produisent. Le fait que

---

4 Dans la section 6.2, je discute brièvement des tentatives d'affirmation de relations causales dans les simulations informatiques, c'est-à-dire de la capacité des chercheurs à déduire des relations causales à partir de simulations informatiques. Ce problème ne doit pas être confondu avec la mise en œuvre d'un modèle causal, ce qui est parfaitement possible si la spécification est correcte.

5 Une autre question importante qui plaide en faveur de l'explication par dérivation est que les résultats des simulations informatiques comportent des erreurs que les théories causales sont incapables de prendre en compte (voir (Duran 2017)).

les chercheurs savent que les résultats sont corrects car le modèle ne dit rien de la raison pour laquelle les pics se produisent. À moins que les chercheurs ne donnent une explication explicite indiquant que les pics se produisent en raison d'un échange entre le spin et le moment angulaire orbital autour de l'approche la plus proche, ils n'ont aucune idée réelle de la raison pour laquelle ces pics sont là. Les explications fonctionnent, quand elles fonctionnent, non seulement en vertu de la bonne relation explicative, mais aussi parce qu'elles fournissent une véritable compréhension scientifique. Le modèle ptolémaïque, par exemple, ne pouvait pas expliquer la trajectoire des planètes d'une manière épistémiquement significative puisqu'il ne permet pas de comprendre la mécanique planétaire. D'autre part, les modèles newtoniens classiques expliquent précisément parce qu'ils décrivent de manière compréhensible la structure du mouvement planétaire. Le critère pour ancrer un type de modèle plutôt qu'un autre comme explicatif réside, en partie, dans leur capacité à fournir une compréhension du phénomène étudié.

Nous arrivons enfin à notre dernière question, c'est-à-dire « à quel type de compréhension devrions-nous nous attendre en expliquant avec des simulations informatiques ? » Comme nous l'avons déjà mentionné, les chercheurs veulent expliquer parce qu'ils espèrent mieux comprendre le phénomène examiné et, ce faisant, rendre le monde plus transparent et compréhensible.

C'est un fait bien connu en philosophie que notre compréhension peut prendre différentes formes (Lipton 2001). L'une de ces formes consiste à identifier la compréhension avec le fait d'avoir de bonnes raisons de croire que quelque chose est le cas. Selon cette interprétation, les explications fournissent de bonnes raisons de croire les résultats des simulations informatiques - ou de croire que le monde se comporte de la manière dont les simulations informatiques le décrivent. Bien que ce point de vue soit attrayant, il ne fait pas la différence entre savoir que quelque chose est le cas et comprendre pourquoi cela se produit. Le fait que les simulations montrent qu'il devrait y avoir beaucoup plus de petites galaxies autour de la Voie lactée qu'il n'en est observé à travers les télescopes fournit une excellente raison de croire que c'est effectivement le cas, mais pas la moindre idée de pourquoi (Bœhm et al. 2014).

Une autre façon dont l'explication permet de comprendre consiste à réduire l'inconnu à quelque chose de plus familier – et donc connu. Ce point de vue s'inspire d'exemples tels que la théorie cinétique des gaz, où des phénomènes inconnus sont comparés à des phénomènes plus familiers tels que le mouvement de minuscules boules de billard. Malheureusement, ce point de vue souffre de nombreux inconvénients, y compris des problèmes concernant la signification de « être familier ». Ce qui est « familier » à un physicien peut ne pas l'être à un ingénieur.
De plus, de nombreuses explications scientifiques relient des phénomènes familiers à une théorie inconnue. Il n'y a peut-être rien de plus familier pour nous qu'un embouteillage matinal sur le chemin du travail. Cependant, des événements comme celui-ci nécessitent une modélisation très complexe où l'explication est loin d'être familière.

Notons que le point de vue de la familiarité a également des difficultés avec le soi-disant « pourquoi régresser ». Cela signifie que seul ce qui est familier est compris, et que seul ce qui est familier peut expliquer. Il s'ensuit que ce point de vue ne permet pas que ce qui n'est pas lui-même compris puisse néanmoins s'expliquer. Mais les chercheurs ont intérêt à permettre cela : ils veulent pouvoir expliquer les phénomènes même dans les cas où ils ne comprennent pas les théories et les modèles impliqués dans l'explication.

Une manière peut-être plus sophistiquée de comprendre consiste à pointer les causes qui provoquent un phénomène donné. C'est la forme qui prend le plus de causalité

récits d'explication. Notre première question de savoir pourquoi la fenêtre s'est brisée peut être pleinement comprise lorsque nous prenons en compte toutes les causes qui conduisent à ce scénario ou, comme les philosophes des sciences aiment à le dire, le phénomène se situe dans le lien causal. Ainsi, lorsqu'une pierre est lancée contre la fenêtre, il y a alors une série de relations causales qui finissent par produire un effet : la pierre lancée par ma main se déplace dans l'air et finit par atteindre la fenêtre qui se brise finalement.

A mon sens, aucune de ces formes de compréhension n'est adaptée à l'informatique simulations. Dans certains cas, les simulations informatiques ne donnent aux chercheurs aucune sorte de raisons de croire leurs résultats. Dans d'autres, la réduction au familier est simplement impossible si l'on tient compte du grand nombre d'incertitudes spécifiées et non spécifié dans le modèle de simulation. Il est ridicule de penser qu'une sorte de réductionnisme de simulations plus complexes à des simulations moins complexes - et peut-être plus familier - est même possible. Enfin, la possibilité d'identifier des relations causales et leur signaler une simulation semble tiré par les cheveux. Comme je le dis plus tard au chapitre 6, la causalité dans les simulations informatiques est plutôt un programme de recherche ouvert qu'un hypothèse initiale.

Outre les interprétations précédentes de la compréhension, il existe encore une autre forme cela s'avère assez prometteur pour les simulations informatiques. Cette interprétation est connu sous le nom de point de vue «unificationniste» parce qu'il suppose que la compréhension consiste à voir comment, ce qui a été expliqué, s'inscrit dans un tout unifié.

Pour l'unificationniste, la compréhension vient du fait de voir des liens et des points communs. modèles dans ce qui semblait initialement être des faits brutaux ou indépendants.[6] 'Voir' ici est considéré comme la manœuvre cognitive de réduction des résultats expliqués - ou du monde réel phénomènes – à un cadre théorique plus large, tel que notre corpus de croyances. Plusieurs philosophes des sciences ont adhéré à ce point de vue de la compréhension – mais pas nécessairement à l'unificationnisme. Gerhard Schurz et Karel Lambert dit que « comprendre un phénomène P, c'est savoir comment P s'inscrit dans son connaissances de base » (Schurz et Lambert 1994, 66), et Catherine Elgin affirme que « la compréhension est avant tout une relation cognitive à un ensemble assez compréhensif, ensemble cohérent d'informations » (Catherine Elgin 2007, 35). La réduction proposée par l'unificationniste s'accompagne de plusieurs avantages épistémologiques, tels que des résultats devenir plus transparent pour les chercheurs, qui à leur tour obtiennent une image plus unifiée de la nature, ainsi que renforcer et systématiser notre corpus de croyances scientifiques. Dans l'ensemble, dit l'unificationniste, le monde devient un endroit plus simplifié (Friedman 1974 ; Kitcher 1981, 1989).

Dans (Duran 2017), je soutiens que lorsque les résultats d'une simulation informatique sont compris en les expliquant, une manœuvre cognitive similaire est effectuée. Les chercheurs peuvent intégrer les résultats de la simulation dans un cadre théorique plus large, réduisant ainsi le nombre de résultats indépendants à rechercher. une explication. Ainsi, en expliquant pourquoi les pics de la figure 1.3 se produisent, les chercheurs élargissent leur corpus de connaissances scientifiques en incorporant un cas dérivé de la mécanique newtonienne.

---

[6] Pour une analyse des faits « bruts » et « indépendants », voir (Barnes 1994) et (Fahrbach 2005)

Le cas des simulations informatiques est particulièrement intéressant car il se déroule en réalité en deux étapes. Premièrement, les résultats sont inclus dans le corps des croyances scientifiques liés au modèle de simulation ; et deuxièmement, ils sont inclus dans notre plus grand corps de croyances scientifiques.

Permettez-moi d'illustrer ce dernier point avec une explication de l'occurrence des pointes dans la figure 1.3. En expliquant pourquoi les pointes apparaissent dans la visualisation, les chercheurs expliquent les raisons de leur formation. Une telle explication est possible, je crois, parce qu'il existe une structure de modèle bien définie qui permet aux chercheurs de dériver un description des pointes du modèle de simulation (Duran 2017). La compréhension des pointes vient donc de leur incorporation dans le corpus plus large de connaissances scientifiques. croyances qui est le modèle de simulation. Autrement dit, les chercheurs saisissent comment les résultats correspondent dans, contribuent à et sont justifiés par référence au cadre théorique postulé par le modèle de simulation. C'est précisément la raison pour laquelle les chercheurs peuvent pour expliquer l'apparition des pics ainsi que leur tendance à la baisse : les deux peuvent être théoriquement unifié par le modèle de simulation. De plus, puisque la simulation modèle dépend de connaissances scientifiques bien établies - dans ce cas, la mécanique newtonienne - les chercheurs sont en mesure de voir les résultats de la simulation d'une manière qui leur est maintenant bien connu, c'est-à-dire unifié avec le corps général des croyances scientifiques concernant la mécanique à deux corps.

Jusqu'à présent, l'image standard de l'unificationniste s'applique aux simulations informatiques. Mais je crois que nous pouvons prolonger cette image en montrant comment la compréhension des résultats comportent également une dimension pratique. Du point de vue de la recherche en simulation, comprendre les résultats, c'est aussi appréhender les difficultés techniques pour programmer des simulations plus complexes, plus rapides et plus réalistes, interpréter les processus de vérification et de validation, et véhiculer des informations pertinentes pour le système interne. mécanisme de la simulation. En d'autres termes, la compréhension des résultats permet également de dans le modèle de simulation, contribuant à améliorer les simulations informatiques. Par exemple, en expliquant et en comprenant les raisons pour lesquelles les pics tendent vers le bas, les chercheurs sont conscients de l'existence et ont les moyens de résoudre les erreurs d'arrondi, erreur de discrétisation, résolution de grille, etc. Dans les études de simulation informatique, les chercheurs veulent expliquer parce qu'ils veulent aussi comprendre et améliorer leurs simulations, tout autant qu'ils veulent comprendre les phénomènes du monde réel (Duran 2017).

Le dernier point que nous devons aborder ici est de savoir comment comprendre les phénomènes du monde réel en expliquant les résultats des simulations informatiques. Comme je l'ai présenté au début de cette section, Weirich et Krohs avaient pour objectif principal la explication des phénomènes du monde réel à l'aide de simulations informatiques. Les deux auteurs sont raison de penser que l'utilisation de simulations informatiques se justifie, dans un grand nombre de cas, car elles permettent de comprendre certains aspects du monde.

La question est maintenant de savoir si nous pouvons comprendre les phénomènes réels en expliquant résultats de simulations informatiques? Je crois que nous pouvons répondre positivement à cette question. Nous savons que la visualisation des résultats de la simulation représente la comportement d'un satellite du monde réel sous le stress des marées. Cela signifie que les résultats de la la simulation liée aux pointes représente, et peut donc être attribuée au comportement d'un satellite du monde réel. En ce sens, et suivant Elgin sur ce point (2007 ; 2009), il existe un droit – par représentation et attribution – aux résultats de

la simulation au comportement d'un satellite du monde réel. C'est précisément grâce à ce droit que nous sommes en mesure de relier notre compréhension des résultats de la simulation à notre compréhension du comportement du satellite du monde réel. Ce point peut également être fait au moyen de la capacité pratique qui présuppose comprendre quelque chose. Comme Elgin le soutient avec force, celui qui comprend détient la capacité d'utiliser les informations à sa disposition à des fins pratiques (Catherine Elgin 2007, 35).
Dans notre cas, les chercheurs pourraient en fait construire le satellite spécifié dans la simulation et le placer dans l'espace.

5.1.2 Outils prédictifs

Lorsque les philosophes ont concentré leur attention sur l'explication scientifique, ils se sont également tournés vers la prédiction scientifique. En fait, Carl Hempel et Paul Oppenheim, les deux principaux philosophes qui ont systématisé et fixé l'ordre du jour des études philosophiques sur l'explication scientifique, pensaient que la prédiction était le revers de la médaille. Ces idées ont fleuri dans et autour de 1948 avec leur ouvrage fondateur Studies in the Logic of Explanation (Hempel et Oppenheim 1948), et se sont poursuivies jusqu'à la disparition de l'empirisme logique en 1969 lors d'un symposium sur la structure des théories à Urbana, Illinois (Suppe 1977) .

Malgré la naissance commune, l'explication scientifique et la prédiction ont à peu près pris des directions différentes depuis. Alors que les études d'explication scientifique se sont considérablement développées en établissant différentes écoles de pensée, les philosophes travaillant sur la prédiction sont plus difficiles à trouver. Il est intéressant de noter qu'en ce qui concerne les études philosophiques sur les simulations informatiques, beaucoup plus d'efforts ont été consacrés à l'étude de la prédiction et beaucoup moins à l'explication scientifique. Cette asymétrie peut s'expliquer par les pratiques de simulations informatiques dans des contextes scientifiques et d'ingénierie. Les chercheurs sont plus intéressés à prédire les états futurs d'un système, plutôt qu'à expliquer pourquoi de tels états sont obtenus. Au début de ce chapitre, nous avons discuté en détail de l'explication scientifique. Il est maintenant temps d'attirer l'attention sur certains des principes de base de la prédiction scientifique dans le contexte des simulations informatiques. Mais d'abord, un exemple qui illustre les prédictions en science et aide à introduire la terminologie de base.

Un beau cas dans l'histoire des sciences est celui des prédictions d'Edmond Halley sur les comètes. Les croyances générales sur les comètes à l'époque de Halley étaient qu'elles étaient de mystérieux intrus astronomiques se déplaçant de manière imprévisible dans le ciel.
Bien que Halley ait fait des rétrodictions précises - c'est-à-dire des prédictions dans le passé - établissant que les comètes apparues en 1531, 1607 et 1682 étaient toutes des manifestations humaines du même phénomène, ses postdictions - c'est-à-dire des prédictions dans le futur - de la La prochaine apparition de la comète n'a pas été aussi réussie, confirmant ainsi la croyance scientifique établie de l'époque. En fait, il a postulé que la comète réapparaîtrait dans le ciel en 1758, un an plus tôt que son apparition réelle.

C'était le travail d'Alexis Clairaut, un éminent newtonien, de postdicter correctement la prochaine apparition de la comète, qui devait atteindre son périhélie en 1759.
Clairaut a basé ses postdictions sur des calculs qui incluraient des forces inconnues à l'époque, mais qui avaient du sens dans la théorie newtonienne. Ces forces faisaient principalement référence aux actions et à l'influence de planètes lointaines - rappelons qu'Uranus a été découverte en 1781 et que Neptune est restée inconnue jusqu'en 1846. Il est intéressant de noter qu'outre une postdiction réussie, ce qui était également en jeu était la confirmation de théorie comme la manière la plus adéquate de décrire le monde naturel. La lutte entre les factions a conduit de nombreux physiciens à rejeter d'abord les calculs de Clairaut, et beaucoup d'autres à anticiper avec une certaine joie l'échec de la théorie newtonienne. L'histoire se termine avec les calculs de Clairaut postdictant correctement la prochaine apparition de la comète de Halley - malgré quelques troncatures dans les termes supérieurs de son équation - et avec la vision newtonienne du monde qui s'impose de manière écrasante sur les théories moins adéquates.

L'exemple de la comète de Halley montre clairement l'importance de la prédiction pour la recherche scientifique – et, dans ce cas particulier, également pour confirmer la théorie newtonienne. Les prédictions réussies sont précieuses car elles vont au-delà de ce que les chercheurs savent le plus directement et le plus évidemment, fournissant des informations « cachées » sur les phénomènes et les systèmes empiriques.7

Deux caractéristiques importantes de la prédiction sont la dimension temporelle et la précision de la prédiction. Une erreur courante consiste à présumer que les prédictions consistent à dire quelque chose de significatif sur l'avenir (c'est-à-dire les postdictions), mais pas sur le passé (c'est-à-dire les rétrodictions). C'est généralement le cas lorsque les rétrodictions sont confondues par erreur avec des preuves scientifiques de quelque chose qui se passe. Ainsi, on dit que Halley a trouvé des preuves que la comète apparue en 1531, 1607 et 1682 était la même, mais pas qu'il a fait des rétrodictions. Bien que la preuve et la prédiction puissent, dans certains cas, être liées, il s'agit toujours de deux notions distinctes. Les rétrodictions – comme les postdictions – font partie des implications d'une théorie, et ce quelles que soient les contraintes temporelles. Les preuves, d'autre part, servent à soutenir ou à contrer une théorie scientifique. En ce sens, rétrodictions et postdictions sont la manifestation d'un même phénomène, à savoir des prédictions. Ainsi compris, il est correct de dire que Halley a prédit les trois occurrences de la comète avant son observation de 1682, tout comme Clairaut a prédit sa prochaine apparition en 1759 parce qu'il a fait des calculs à l'aide d'une théorie - une forme d'implication théorique.

Le langage de la prédiction est utilisé pour décrire des affirmations déclaratives sur le passé ainsi que des événements futurs faites à la lumière d'une théorie, et nous l'utiliserons donc ici. En effet, la dimension temporelle comporte une composante épistémique : « prédire c'est

---

7 Incidemment, ces mêmes caractéristiques rendent la prédiction intrinsèquement risquée, puisque l'accès à l'information cachée dépend généralement de normes partagées dans une communauté donnée, et est donc potentiellement intersubjectif. Or, une façon d'éviter les problèmes d'intersubjectivité est d'exiger des prédictions convergentes. Autrement dit, en ayant des domaines de recherche disparates qui prédisent tous vers des résultats similaires, notre confiance dans la prédiction doit inévitablement augmenter. C'est ce que la philosophe Heather Douglas appelle l'objectivité convergente (Douglas 2009 : 120). Nous devons toutefois garder à l'esprit qu'il existe également d'autres formes d'objectivité. Ici, nous n'allons pas nous préoccuper des questions de subjectivité et d'objectivité, car elles méritent une étude à elles seules. Pour d'autres références, le lecteur est renvoyé à ((Daston et Galison 2007) (Lloyd 1995), et bien sûr (Douglas 2009)).

faire une déclaration sur des sujets qui ne sont pas encore connus, pas nécessairement sur des événements qui ne se sont pas encore produits » (Barrett et Stanford 2006, 586). En d'autres termes, la prédiction concerne la « diction » d'un événement, passé, présent ou futur. Ainsi comprise, la prédiction revient à se demander « que nous dit la théorie ? », « quelles connaissances sont nouvelles ? Dans notre cas, la réponse est plutôt évidente : la prochaine apparition de la comète de Halley.

Étroitement liée à cette dimension temporelle se trouve la question « dans quelle mesure la théorie prédit-elle le résultat réel observé ? Il s'agit d'une question principale qui nous donne un achat sur la deuxième caractéristique de la prédiction pertinente pour notre discussion, à savoir l'exactitude des prédictions.

Dans le chapitre précédent, nous avons discuté de la précision comme l'ensemble des mesures qui fournissent une valeur estimée proche de la valeur réelle de la quantité mesurée.
Par exemple, le noyau de la comète de Halley mesure environ 15 kilomètres de long, 8 kilomètres de large et environ 8 kilomètres d'épaisseur - un noyau plutôt petit pour la vaste taille de son coma. La masse de la comète est également relativement faible, estimée à 2,2x1014 kg. Comme les tronomers ont calculé la densité moyenne à 0,6 g/cm3, ce qui indique qu'il est constitué d'un grand nombre de petits morceaux maintenus lâchement ensemble. Avec ces informations à portée de main, ainsi que les calculs de la trajectoire, les astronomes peuvent prédire avec précision que la comète de Halley est visible à l'œil nu avec une période de 76 ans.

L'exactitude des prédictions scientifiques dépend d'une combinaison de la nature de l'événement, de l'adéquation de nos théories et de l'état actuel de notre technologie.
Bien qu'il soit vrai qu'en utilisant des ordinateurs, nous sommes capables de prédire avec des degrés de précision plus élevés que ceux de Clairaut la prochaine fois que la comète de Halley apparaîtra dans le ciel,8 des prédictions sont possibles parce que la trajectoire de la comète peut être décrite de manière adéquate en utilisant la mécanique newtonienne . .

Maintenant, pour de nombreuses occasions, il est difficile d'obtenir des prédictions exactes. Il en est ainsi en raison de la nature des phénomènes à prévoir. En effet, il existe de nombreux événements et phénomènes pour lesquels on ne peut prédire leur comportement que dans une fourchette donnée de probabilité d'occurrence. Un exemple simple est le suivant. Supposons que je vous demande de choisir une carte et de la placer face cachée sur la table. Ensuite, je dois prédire sa couleur (c'est-à-dire 'coeur', 'carreau', 'trèfle' ou 'pique'). La théorie dit que j'ai 1 chance sur 4 d'avoir raison. La théorie n'implique pas quelle couleur vous venez de placer face cachée, mais plutôt que la probabilité que j'aie raison est de 4 . Bien que cela ne compte pas strictement comme une prédiction, cela nous dit quelque chose sur l'exactitude de la prédiction.

Un autre bon exemple de prédictions inexactes sont les systèmes dits chaotiques.
Ce sont des systèmes très sensibles aux conditions initiales, où de petites erreurs de calcul peuvent rapidement se propager en grandes erreurs de prédiction. Cela signifie que certaines prédictions sont impossibles après un certain point du calcul. Le fait que les systèmes chaotiques passé un certain point rendent les prédictions difficiles et essentiellement im

---

[8] Clairaut a participé à plusieurs calculs de la période orbitale de la comète de Halley. Pour l'un des premiers calculs, il a divisé l'orbite cométaire en trois parties. Premièrement, de 0◦ à 90◦ d'anomalie excentrique, le premier quadrant de l'ellipse de périhélie, dont la comète a pris plus la moitié plus de 7 ans à parcourir. Deuxièmement, de 90◦ à , supérieure de l'orbite, que la comète le dernier 270◦ il a fallu plus de 60 ans pour traverser. Troisièmement, de , quadrant de l'ellipse, 270◦ à 360◦ qui a mis environ 7 ans à traverser. Les calculs de Clairaut ont été rectifiés à différents moments de l'histoire (Wilson 1993).

possible dans la pratique a des conséquences importantes pour la recherche scientifique. Un exemple standard de système chaotique est celui des conditions atmosphériques. Dans les prévisions météorologiques, il est très difficile de faire des prévisions à long terme en raison de la complexité inhérente du système qui ne permet des prévisions que jusqu'à un certain point. C'est la raison pour laquelle le comportement du temps est considéré comme chaotique, les plus petits changements dans les conditions initiales rendant les prévisions à long terme irréalisables, voire totalement impossibles.

Voyons maintenant comment la prédiction fonctionne dans les simulations informatiques en discutant de deux simulations d'une épidémie en Italie. La première simulation est un modèle stochastique à base d'agents, tandis que la seconde simulation est un modèle structuré de méta-population - connu sous le nom de GLobal Epidemic and Mobility - GLEaM. Le modèle basé sur les agents comprend une représentation explicite de la population italienne à travers des données très détaillées de la structure sociodémographique. De plus, et pour déterminer la probabilité de se déplacer d'une municipalité à l'autre, le modèle utilise une norme générale de modèle de gravité dans la théorie des transports. La dynamique de transmission épidémique repose sur une compartimentation des syndromes grippaux (SG) basée sur des modèles stochastiques qui intègrent les infections sensibles, latentes, asymptomatiques et les infections symptomatiques (Ajelli et al. 2010, 5). Marco Ajelli et son équipe, responsables de la conception et de la programmation de ces simulations, définissent le modèle à base d'agents comme « un modèle de simulation stochastique, spatialement explicite, à temps discret, où les agents représentent des individus humains [...] Les caractéristiques du modèle sont la caractérisation du réseau de contacts entre les individus sur la base d'un modèle réaliste de la structure sociodémographique de la population italienne » (4).

D'autre part, la simulation GLEaM intègre des bases de données de population à haute résolution - estimant la population avec une résolution donnée par des cellules de 15 x 15 minutes d'arc - avec l'infrastructure de transport aérien et les modèles de mobilité à courte distance. De nombreuses simulations GLEaM standard consistent en trois couches de données. Une première couche, où la population et la mobilité permettent le découpage du monde en régions géographiques. Cette partition définit une deuxième couche, le réseau de sous-population, où l'interconnexion représente les flux d'individus via les infrastructures de transport et les schémas de mobilité générale. Enfin, et superposée à cette deuxième couche, se trouve la couche épidémique, qui définit à l'intérieur de chaque sous-population la dynamique de la maladie (Balcan et al. 2009). Dans l'étude, le GLEaM représente également une partition en forme de grille où chaque cellule se voit attribuer l'aéroport le plus proche. Le réseau de sous-population utilise des données de recensement géographique et les couches de mobilité obtiennent des données de différentes bases de données, y compris la base de données de l'Association du transport aérien international consistant en une liste d'aéroports dans le monde reliés par des vols directs.

Comme prévu, il y a des avantages et des inconvénients à utiliser les simulations basées sur les agents et les simulations GLEaM. En ce qui concerne le modèle GLEaM, les réseaux de mobilité spatiale détaillés fournissent une description précise des voies de transport disponibles pour propager la maladie. Cependant, des estimations précises de l'impact de la maladie à un niveau plus local sont difficiles à obtenir en raison du faible niveau de détail contenu dans ce modèle. Quant à l'approche basée sur les agents, bien qu'elle soit très détaillée en ce qui concerne les structures des ménages, des écoles, des hôpitaux, etc., elle souffre de la collecte d'ensembles de données de haute confiance provenant de la plupart des régions du monde (Ajelli et

Al. 2010, 12). Bien que chaque simulation offre des fonctionnalités différentes, et affecte ainsi leur taux d'attaque prédictif respectif de la maladie, Ajelli et son équipe remarquent qu'il existe une convergence constante dans les résultats qui donne confiance dans chacune des prédictions des simulations. L'hétérogénéité du réseau de transport fourni par le modèle GLEaM, par exemple, permet des prédictions précises de la propagation spatio-temporelle de la maladie ILI à l'échelle mondiale. D'autre part, la représentation explicite des individus dans le modèle basé sur les agents facilite les prédictions précises de la propagation d'une épidémie à une échelle plus locale.

En précisant les prédictions, la différence dans les amplitudes maximales varie en fonction de plusieurs facteurs, principalement basés sur les valeurs du nombre reproducteur.9 La taille moyenne de l'épidémie prédite par GLEaM est de 36 % pour un taux de reproduction de R0 = 1,5, alors que pour le modèle à base d'agents avec le même taux de reproduction, la taille moyenne de l'épidémie descend à 25 %. À l'autre extrémité de l'épidémie, la taille de l'épidémie prédite par GLEaM est de 56 % de la population pour un taux de reproduction de R0 = 2,3, alors qu'elle est de 40 % pour l'agent-basé. Les chercheurs ont observé une différence absolue d'environ 10 % pour R0 = 1,5 et d'environ 7 % pour R0 = 2,3. Des prédictions ont également été faites pour un taux de reproduction de R0 = 1,9, montrant un comportement similaire dans la taille moyenne de l'épidémie prédite par les deux simulations.

Pour Ajelli et al. ces prédictions semblent assez précises car elles voient une convergence dans la taille moyenne de l'épidémie. L'équipe, cependant, est incapable d'évaluer laquelle des deux prédictions est la meilleure. Le haut niveau de réalisme du modèle à base d'agents devrait, en principe, parler en faveur de la précision de la prédiction. Mais malheureusement, un modèle aussi réaliste n'est pas exempt d'hypothèses de modélisation. En effet, afin de déterminer la probabilité de se déplacer d'une commune à l'autre, Ajelli et al. mettre en œuvre un modèle de gravité général utilisé dans la théorie des transports et supposer une forme fonctionnelle de loi de puissance pour la distance – malgré le fait que d'autres formes fonctionnelles telles que la décroissance exponentielle peuvent également être envisagées (3). Une autre hypothèse découle de la prise d'un mélange homogène dans les ménages, les écoles et les lieux de travail, tandis que les contacts aléatoires dans la population générale sont supposés dépendre explicitement de la distance (3). Malgré le fait que les deux hypothèses sont parfaitement raisonnables, elles affectent inévitablement la prédiction et, par conséquent, les raisons de préférer la prédiction du modèle à base d'agents à celle du GLEaM. De même, la prédiction du modèle de métapopulation structurée est tout aussi difficile à évaluer précisément en raison de son manque de réalisme. "La valeur correcte", dit Ajelli et al., "devrait se situer entre les prédictions des modèles, comme en témoigne le fait que la différence entre les modèles diminue à mesure que R0 augmente, les modèles convergeant vers la même valeur pour l' attaque . taux » (8).

---

9 En épidémiologie, le taux de reproduction - ou nombre de reproduction - représente le nombre de cas émergents qu'un individu infecté génère en moyenne au cours d'une période infectieuse dans une population par ailleurs non infectée. Ainsi, pour R0 < 1, une épidémie infectieuse s'éteindra à terme, alors que pour un R0 > 1 une infection pourra se propager et infecter la population. Plus le nombre est grand, plus il sera difficile de contrôler l'épidémie. Ainsi, la métrique aide à déterminer la vitesse à laquelle une maladie infectieuse peut se propager dans une population en bonne santé. Dans le cas des deux simulations, les auteurs rapportent que pour de grands R0, les épidémies locales se généralisent à toutes les couches de la population, rendant la structure de la population de moins en moins pertinente.

Il faut ici souligner une différence cruciale entre les prédictions des simulations d'Ajelli et al., et celles de Clairaut : alors que les prédictions des plus tard peuvent être confirmées empiriquement, les prédictions des simulations informatiques ne pouvait pas. Sur quelles bases, alors, Ajelli et son équipe prétendent-ils que ces prédictions sont exactes ? Leur réponse est d'insister sur le fait que le bon accord conclu par les résultats des deux simulations constituent des raisons qui plaident en faveur d'une analyse précise prédiction.10 Assurément, le fait que les deux simulations sont, dans leurs propres termes, de bonnes représentations du système cible contribue également à la confiance du chercheur. Mais c'est la convergence des résultats qui fonde fermement la confiance dans des prévisions exactes.

Une telle position ne manque pas de bonnes raisons. La convergence des résultats parle en faveur des simulations ainsi que de la précision de la prédiction. Dans le chapitre précédent, nous avons vu comment les méthodes de vérification et de validation peuvent être utilisées pour accorder fiabilité sur la simulation et ainsi aider les chercheurs à se fier à leurs résultats. Le Une tournure intéressante introduite par Ajelli et son équipe est que la validation n'est pas contre des données empiriques, mais plutôt contre une autre simulation informatique.11 Cela signifie que la les résultats d'une simulation fonctionnent comme une instance de confirmation des résultats de l'autre simulation – et vice versa. C'est une pratique assez courante chez les chercheurs faisant un usage intensif de simulations informatiques et dont les systèmes cibles il n'y a pas de données empiriques disponibles pour la validation. En d'autres termes, la fiabilité des simulations informatiques est utilisée pour fournir des motifs de confiance dans leurs résultats qui, dans des conditions de convergence, servent à leur tour de fondement pour croire à l'exactitude de toutes sortes de prédictions.

Cela dit, les pratiques prédictives doivent être tempérées par quelques principes de précaution sains. Comme tout chercheur le sait, les modèles sont basés sur notre représentation du fonctionnement du monde réel. Bien que le corpus scientifique et technique croyances est extrêmement réussie, elle n'est en aucun cas infaillible. C'est pourquoi la validation contre des données empiriques est une pratique bonne et nécessaire dans la recherche scientifique et technique. Cela oblige pratiquement les chercheurs à "tester" leurs résultats contre le monde, et pour trouver une solution de contournement en cas de non-concordance. Dans ce cas particulier, Ajelli et son équipe considère la convergence des résultats comme un signe positif – bien qu'il ne s'agisse pas d'un exemple de confirmation – d'une prédiction précise. Donc beaucoup est, je crois, correct. Cependant, comme ils l'expliquent également dans leur article, la prédiction par convergence des résultats présuppose un socle commun dans les cadres de modélisation au sein desquels les simulations sont situés. Autrement dit, des informations telles que la paramétrisation, l'intégration de données,

---

[dix] Cette réponse soulève la question de savoir ce qu'est un "bon accord" pour les résultats de simulations informatiques. Ajelli et al. ne précisent pas comment cette notion doit être comprise. Nous pourrions supposer que nous avons un bon accord lorsque tous les résultats tombent dans une distribution donnée (par exemple, une distribution normale). Ceci, bien sûr, nécessite des spécifications. Comme nous l'ont appris les études épistémologiques de expérience, un accord sur les résultats de différentes techniques donne confiance non seulement dans les résultats, mais aussi dans la capacité des techniques à produire des résultats valides (Franklin 1986, chapitre 6). Le La question est alors de savoir dans quelle mesure peut-on considérer la simulation multi-agents et GLEaM deux techniques différentes.

[11] J'entends par là que les résultats de l'une ou l'autre des simulations informatiques ne peuvent, en principe, être empiriquement validé. Étant donné que chaque simulation informatique implémente des sous-modèles, il est possible que certaines d'entre elles ont été validées empiriquement.

et les conditions aux limites, et même les nombreuses approximations utilisées doivent être partagées parmi les chercheurs, sinon, Ajelli et al. affirment que les chercheurs seraient incapables de écarter les effets indésirables tels que ceux découlant des hypothèses de modélisation.

Plus inquiétant est la possibilité que les résultats des deux simulations convergent artificiellement. Pour éviter cela, les chercheurs essaient de comparer les résultats qui n'ont pas de base commune entre les simulations. Autrement dit, les simulations informatiques doivent être différenciés en termes d'hypothèses de modèle, de conditions initiales, de systèmes cibles, paramétrisation, calibration, etc. C'est l'approche choisie par M. Elizabeth Hallo ran et son équipe, comme le rapportent Ajelli et. Al. Halloran compare les résultats de trois modèles individuels d'une souche de grippe pandémique avec différentes hypothèses et des données initiales, une au niveau de la description d'une ville et deux au niveau de la description d'un pays (Halloran et al. 2008).

Malgré de nombreux efforts de précaution, comparer différents modèles et résultats est généralement une tâche difficile. Ajelli et al. soulignent que la comparaison de Halloran et al. est limitée aux hypothèses de chaque modèle ainsi qu'aux valeurs simulées disponibles. scénarios. À cet égard, ils ne définissent jamais explicitement un ensemble commun de paramètres, conditions initiales et approximations partagées par tous les modèles. La faible transmission scénario proposé par Halloran, par exemple, est comparé dans chaque modèle en utilisant des valeurs différentes pour le nombre reproducteur, avec le risque de ne pas pouvoir écarter l'effet de cette différence dans les résultats obtenus. En conséquence, la convergence des résultats est artificielle, de même que toute prédiction faite par les simulations.

5.1.3 Stratégies exploratoires

Pour nous, les stratégies exploratoires ont le caractère d'une activité de recherche dans le but de générer des découvertes significatives sur des phénomènes sans avoir à faire appel à, ni s'appuyer sur la théorie de tels phénomènes. Les stratégies exploratoires ont donc deux objectifs. D'une part, il entend provoquer les changements observables dans le monde ; d'autre part, il sert de banc d'essai pour de nouveaux concepts encore à stabiliser12.

La première question qui vient à l'esprit est alors pourquoi est-il important de mettre à part théorie de l'expérience? Une réponse évidente consiste à souligner que tout une théorie donnée pourrait ne pas être capable de fournir toutes les informations pertinentes sur un phénomène donné. Un bon exemple de ceci est le mouvement brownien. En 1827, le le botaniste Robert Brown a remarqué que des particules piégées dans des cavités à l'intérieur des grains de pollen dans l'eau se déplacerait. Ni Brown ni personne à l'époque n'a été en mesure de déterminer ni expliquer les mécanismes qui ont provoqué un tel mouvement. Ce n'est qu'en 1905 que

[12] Les stratégies exploratoires font partie de ces sujets qui attirent peu l'attention des philosophes, malgré leur centralité dans l'effort scientifique et technique. Heureusement, il existe quelques excellents travaux qui traitent de ces questions. Quelques exemples dont je vais parler ici viennent des travaux de Friedrich Steinle sur les expériences exploratoires (Steinle 1997), Axel Gelfert sur modèles exploratoires (Gelfert 2016), et sur les simulations informatiques exploratoires, nous avons Viola Schi affonati (Schiaffonati 2016) et P´ıo Garc´ıa et Marisa Velasco (Garc´ıa et Velasco 2013). Ici, Je vais adopter une approche assez différente de ces auteurs.

Albert Einstein a publié un article expliquant en détail comment le mouvement qui
Brown avait observé était le résultat du déplacement du pollen par l'eau individuelle
molécules.

Une autre raison importante pour dissocier la théorie de l'expérience est que, dans de nombreux
cas, des expériences sont utilisées pour démystifier une théorie donnée. Par exemple, August Weismann a
mené une expérience où il a enlevé la queue de 68 souris blanches à plusieurs reprises sur cinq générations
juste pour montrer qu'aucune souris n'est née sans queue ou
même avec une queue plus courte simplement parce qu'il les a coupés. L'intérêt de Weismann était de mettre
la fin du lamarckisme et de la théorie de l'hérédité des caractères acquis par
montrant comment il ne pouvait pas tenir compte d'une génération de souris à queue courte ou sans queue.

Après avoir proposé l'importance de l'expérimentation indépendante pour notre
connaissance du monde, peut-on aujourd'hui donner du sens à l'idée de stratégies exploratoires dans les
simulations en laboratoire et sur ordinateur ? Pour répondre à cette question, nous devons
revenir un peu en arrière, à la fois dans le temps et en nombre de pages, et revoir brièvement le contexte
où les idées sur les stratégies exploratoires ont fleuri. Au chapitre 3, j'ai mentionné
les empiristes logiques en tant que groupe de philosophes et de scientifiques intéressés à comprendre la
notion et les implications des théories scientifiques. Comme ils l'ont vite découvert,
c'est autant un terme familier qu'un concept évasif. Ils découvrent notamment
que la théorie et l'expérience sont plus imbriquées qu'on ne le pensait initialement.

L'une des principales positions de l'empiriste logique était de considérer les expériences non pas tant
comme un problème philosophique en soi, comme une méthodologie subsidiaire pour comprendre
théories. Les critiques ont profité de cette position pour attaquer les idéaux fondamentaux de la logique
empiristes. Pour beaucoup de ces critiques, les expériences et l'expérimentation étaient des problèmes
valeur philosophique authentique et devait être traité comme tel. Un particulièrement
point intéressant soulevé était l'objectivité des preuves d'observation, et comment
cette question relève de la théorie. Le problème peut être formulé comme suit : dans leurs expériences, les
chercheurs interagissent, observent et manipulent également des phénomènes du monde réel.
que de recueillir des preuves pour une analyse plus approfondie, tout comme Weismann, Lamarck et Brown l'ont fait
à leur manière il y a de nombreuses années. Le problème est alors de déterminer si les observations sont
objectives[13] ou dépendent du bagage théorique du chercheur.
connaissances – un problème connu sous le nom de « charge théorique ».

Les positions sur ce point étaient partagées. Dans de nombreux cas, les chercheurs ne pouvaient pas
garantir que l'expérimentation fournit des résultats significatifs sur les phénomènes sans
avoir à faire appel en quelque sorte à la théorie. La raison en est que chaque chercheur
aborde le monde avec quelques connaissances de base. Même Brown, lorsqu'il a observé les particules se
déplaçant dans l'eau, s'approchait du phénomène
avec l'esprit d'un scientifique. Plusieurs critiques des empiristes logiques, dont
Thomas Kuhn, Norwood Hanson et Paul Feyerabend, entre autres, étaient très
méfiant des idées d'objectivité dans les preuves d'observation. Pour eux, les chercheurs
ne peut pas vraiment observer, collecter et utiliser des preuves de laboratoire sans s'engager
eux-mêmes à une théorie donnée.

Pour illustrer l'ampleur du problème, prenons un cas précis de l'histoire
de l'astronomie. Les premiers astronomes d'observation avaient des instruments très simples pour ob

[13] Cette notion est utilisée dans le sens d'indépendant de tout chercheur, instrument, méthode ou théorie.

au service des étoiles. L'une des premières observations importantes faites par Galileo Galilei - outre le nombre de lunes en orbite autour de Jupiter – était de Saturne en l'an 1610. Retour puis, il a deviné à tort que Saturne était une grande planète avec une lune de chaque côté. Au cours des 50 années suivantes, les astronomes ont continué à dessiner Saturne avec les deux lunes, ou avec des "bras" sortant des poteaux. Ce n'est qu'en 1959 que la gens de Christiaan Huy a correctement déduit que les "lunes" et les "bras" étaient en fait le système d'anneaux. de Saturne. Cela était bien sûr possible grâce à l'amélioration de l'optique du télescope. Le point est que jusqu'à la découverte réelle du système d'anneaux, les astronomes cherchaient à Saturne de la même manière que Galilée l'a représenté.14

L'exemple montre plusieurs problèmes liés aux stratégies exploratoires et le problème de la charge théorique. Il montre que l'observation n'est pas toujours la plus fiable source de connaissances simplement parce que les instruments ne sont peut-être pas assez puissants – ou sont manipulés - pour réellement fournir des informations fiables sur le monde. Elle montre également que les attentes du chercheur sont une source majeure d'influence dans leurs rapports. Cela est particulièrement vrai dans les cas où une « autorité » a établi le terrain de travail sur une question donnée. Galileo est un exemple de la façon dont l'autorité est parfois incontestée.

Une combinaison de ces deux problèmes est fournie par l'histoire de la physique. Au début des années 1920, il y eut une grande controverse entre Ernest Rutherford et Hans Pettersson concernant l'émission de protons à partir d'éléments tels que le carbone et le silicium subissant un bombardement par des particules alpha. Les deux chercheurs ont mené des expériences similaires où ils ont pu observer un écran à scintillation pour flashs produits par des impacts de particules. Alors que le laboratoire de Pettersson a rapporté une observation positive, celui de Rutherford a rapporté n'avoir vu aucun des éclairs attendus de carbone ou silicium. James Chadwick, le collègue de Rutherford, a visité le laboratoire de Pettersson pour évaluer leurs données afin de travailler sur d'éventuelles erreurs dans leur propre approche. Pendant que les assistants de Pettersson lui montraient les résultats, Chadwick était manipuler l'équipement sans que personne ne s'en aperçoive. Les manipulations de Chadwick ont modifié les conditions normales de fonctionnement de l'instrument, assurant que aucune particule ne pouvait réellement toucher l'écran. Malgré cela, les assistants de Pettersson continuent ont signalé avoir vu des éclairs à un rythme très proche du taux signalé dans les conditions précédentes. Après ces événements, les données de Pettersson ont été incontestablement discréditées (Stuewer 1985, 284-288).

Plus à notre intérêt, l'exemple montre que les observations du chercheur sont façonnés par leur formation et par la théorie dans laquelle eux - et leur superviseur et associés – encadrent leurs expériences. Cela soulève la question suivante qui est à cœur de notre étude sur les stratégies exploratoires. Si l'observation – et d'autres formes d'expérimentation – est teinté d'attentes théoriques, dans quel sens tout le processus d'expérimentation - de la mise en place de l'expérience à l'évaluation de la données, y compris la manipulation d'instruments et de phénomènes - générer des résultats sur les phénomènes sans faire appel à la théorie ?

---

14 Voir (Brewer et Lambert 2001) et (Van Helden 1974).

Pour pouvoir répondre à cette question, nous devons préciser la notion de « charge théorique »15. Friedrich Steinle, un philosophe qui s'est consacré à l'étude de nombreux de ces questions, a fait valoir qu'une conception de la charge théorique comme celle exposés ici ne parviennent pas à rendre compte de la complexité et de la diversité des expérimentation (Steinle 1997, 2002). Selon lui, nous devons faire la distinction entre deux types d'expérimentation, à savoir les expériences « guidées par la théorie » et les expériences « exploratoires ». À son avis, les expériences axées sur la théorie présentent plus ou moins les mêmes caractéristiques décrites pour l'expérimentation d'observation au chapitre 3. les expériences sont mises en place et menées avec « une théorie bien formée à l'esprit, de la de la toute première idée, via la conception spécifique et l'exécution, à l'évaluation » (Steinle 1997, 69). Maintenant, dire qu'une expérience est "axée sur la théorie" suggère au moins trois différentes significations. Cela pourrait signifier que les attentes concernant ses résultats chutent dans le cadre fourni par une telle théorie ; cela pourrait signifier que la conception de l'expérience dépend plus ou moins de la théorie ; et cela pourrait signifier que le les instruments utilisés pour l'expérience dépendent fortement de la théorie. Ainsi compris, les expériences basées sur la théorie servent plusieurs objectifs spécifiques, tels que la détermination des paramètres et l'utilisation des théories comme outils heuristiques pour la recherche de nouveaux effets.

En revanche, les expériences exploratoires utilisent des stratégies caractérisées par le manque d'orientation théorique. Pour être plus précis, selon Steinle, aucun des significations susmentionnées attachées aux expériences axées sur la théorie s'appliquent à expériences exploratoires. Ainsi, une expérience exploratoire génère des résultats sur phénomènes qui ne font appel ni au cadre que la théorie fournit, ni à la théorie utilisée pour concevoir l'expérience, ni à la théorie construite dans les instruments qui sont utilisés. En d'autres termes, l'expérience et ses résultats fournissent des informations pertinentes sur les phénomènes pour leur propre compte.

Un problème sérieux avec ce point de vue est qu'il y a un manque de notion de théorie en place, ainsi qu'un manque de compréhension des niveaux de théorie impliqués dans la conception, l'exécution et l'analyse des résultats expérimentaux qui pourraient aider à caractériser ces stratégies en conséquence16. En fait, P´ıo Garc´ıa et Marisa Ve lasco soulignent que, pour défendre l'une quelconque des interprétations de Steinle, nous devons d'abord être capable de rendre compte des différents niveaux de théorie impliqués dans une expérience17, comme ainsi que pour déterminer dans lequel de ces niveaux les conseils théoriques sont les plus pertinents (García et Velasco 2013). En somme, il n'y a pas une théorie qui guide une expérience, et on ne sait pas quel ensemble de théories impliquées dans une expérience en fait une théorie conduit.

Dans ce contexte, l'idée que des expériences pourraient générer des découvertes sur des phénomènes strictement sans faire appel à la théorie commence à sembler difficile à ancrer. Nous pourrions, cependant, contentez-vous d'une interprétation plus générale, et peut-être plus faible, de ce que sont les stratégies exploratoires et du contexte dans lequel elles s'appliquent. Nous pourrions, alors, charac

---

15 Le lecteur doit être conscient qu'il existe de nombreuses subtilités dans la littérature sur la «théorie charge' que nous n'allons pas aborder ici. Pour une bonne source de discussion, voir (Hanson 1958) et (Kuhn 1962).

16 Ces idées peuvent être trouvées dans (Garc´ıa et Velasco 2013, 106).

17 Ian Hacking propose une première approche des différents types et niveaux de théorie impliqués dans une expérience dans (Hacking 1992).

caractériser les stratégies exploratoires par leur relative indépendance vis-à-vis
restrictions et leur capacité à générer des résultats significatifs qui ne peuvent être encadrés
– ou facilement encadrée – dans un cadre théorique bien établi. Un exemple paradigmatique peut être
trouvé dans l'histoire ancienne des phénomènes d'électricité statique,
interprété par Charles Dufay, André-Marie Ampère et Michael Faraday. Comme le souligne Koray Karaca,
ces expériences ont été menées dans un « nouveau domaine de recherche »
qui, à l'époque, n'avait ni cadre théorique bien défini ni bien établi (Karaca 2013). Les résultats, tels
qu'enregistrés, aident à faire progresser l'électromagnétisme comme
la discipline que nous connaissons aujourd'hui.

Ainsi comprises, les stratégies exploratoires sont censées remplir des fonctions épistémologiques très spécifiques. Ils sont particulièrement importants dans les cas où un domaine scientifique donné domaine est ouvert à révision en raison, par exemple, de son insuffisance empirique. Pour de tels cas, les stratégies exploratoires jouent un rôle fondamental dans la fortune des théories puisque leur les conclusions ne sont, par définition, pas encadrées par la théorie examinée. Ils sont également important lorsque, faiblement encadrées dans une théorie, les stratégies exploratoires des informations substantielles sur le monde qui ne sont pas impliquées par la théorie elle-même. Plus généralement, les résultats obtenus par les stratégies exploratoires sont significatifs à une variété d'objectifs, allant de questions plus pratiques telles qu'apprendre à manipuler les phénomènes, à des fins théoriques telles que le développement d'un cadre conceptuel alternatif18. Steinle souligne également comme fonction épistémique majeure des stratégies exploratoires le fait que leurs découvertes pourraient avoir des implications sur notre compréhension des concepts théoriques existants. C'est le cas lorsque, dans le tenter de formuler les régularités suggérées par les stratégies exploratoires, les chercheurs sont amenés à revoir les concepts et catégories existants, et sont contraints de formuler nouvelles afin d'assurer une formulation stable et générale de l'expérimentation résultats (Steinle 2002, 419).19 Face à cela, il n'y a pas une division totale entre expérience et théorie, mais plutôt une coexistence plus compliquée basée sur les diplômes d'indépendance, de capacité à produire des conclusions, etc.

Il est intéressant de noter que, même à l'heure actuelle où la technologie a évolué tant dans le laboratoire scientifique, les stratégies exploratoires appliquées dans des contextes expérimentaux sont toujours primordiales pour l'avancement général de la science et ingénierie. De plus, ce serait une erreur de déduire de ces exemples que ils sont liés à des périodes historiques de la science, à des domaines de recherche ou à des traditions scientifiques. Les travaux de Karaca sur la physique des particules à haute énergie, par exemple, sont une preuve de cela (Karaca 2013).

Nous devons maintenant nous demander s'il est possible de donner un sens aux stratégies exploratoires pour les simulations informatiques. La réponse à cette question est oui, et elle se présente sous la forme de l'une des utilisations les plus appréciées des simulations informatiques, à savoir leur capacité à montrez-nous un monde auquel nous ne pouvons pas facilement accéder. Une utilisation standard de l'ordinateur simulations consiste à étudier comment certains phénomènes du monde réel simulés dans l'ordinateur se comporterait dans certaines conditions spécifiques. Ce faisant, les chercheurs sont en mesure de favoriser leur compréhension de ce phénomène, indépendamment de

---

[18] Cette idée est discutée par Kenneth Waters dans (Waters 2007).

[19] Des fonctions plus exploratoires, telles que les points de départ de la recherche scientifique, les explications potentielles, et d'autres fonctions sont discutées dans (Gelfert 2016, 2018)

la théorie ou le modèle dans lequel le phénomène est encadré. En d'autres termes, la simulation fournit des informations sur des phénomènes qui vont au-delà du modèle mis en œuvre. Dans de tels cas, les simulations informatiques sont généralement utilisées pour produire des résultats pour le cas particulier en question, et non pour déduire ou dériver des solutions générales.
Illustrons cela par un exemple.

Prenez l'utilisation de simulations informatiques en médecine. Un cas important est d'étudier la résistance des os humains, pour laquelle il est essentiel de comprendre leur architecture interne. Dans les expériences matérielles réelles, la force est exercée mécaniquement, mesurée et les données sont collectées. Le problème avec cette approche est qu'elle ne permet pas aux chercheurs de faire la distinction entre la résistance du matériau et la résistance de sa structure. De plus, ce processus mécanique détruit l'os, ce qui rend difficile de voir et d'analyser comment la structure interne détaillée réagit à une force croissante. La meilleure façon d'obtenir des informations fiables sur la résistance des os humains est d'effectuer une simulation informatique.

Deux différents types de simulations ont été décrits au Laboratoire de biomécanique orthopédique de l'Université de Californie à Berkeley par Tony Keaveny et son équipe (Keaveny et al. 1994 ; Niebur et al. 2000). Le premier type, un véritable os de la hanche de vache, a été converti en une image informatisée, un processus qui impliquait de couper de très fines tranches de l'échantillon d'os et de les préparer de manière à permettre à la structure osseuse compliquée de se démarquer clairement des espaces non osseux. . Chaque tranche, ensuite, a été transformée en une image numérique (Beck et al. 1997). Ces images numérisées ont ensuite été réassemblées dans l'ordinateur pour créer une image 3D de haute qualité d'un vrai os de la hanche d'une vache. L'avantage de cette simulation est qu'elle conserve un degré élevé de vraisemblance dans la structure et l'apparence de chaque échantillon osseux particulier. À cet égard, peu de choses sont ajoutées, supprimées, filtrées ou remplacées dans le processus de préparation de l'os et dans le processus de transformation en modèle informatique.

Dans la deuxième simulation, un os stylisé est informatisé sous la forme d'une image de grille 3D. Chaque carré individuel dans la grille reçoit des largeurs assorties basées sur des mesures moyennes des largeurs internes des entretoises de vrais os de vache et inclinées les unes par rapport aux autres par un processus d'attribution aléatoire (Morgan 2003). Les avantages de l'os stylisé viennent en termes de familiarité avec le processus de modélisation. Les chercheurs commencent par émettre l'hypothèse d'une structure de grille simple à laquelle des détails et des fonctionnalités sont ajoutés au besoin. De cette façon, une structure abstraite idéalisée et simplifiée de l'os est créée dès le début.

Ainsi comprise, la première simulation ressemble à des procédures plus proches des montages expérimentaux alors que la seconde simulation ressemble à des méthodes rencontrées dans les pratiques de modélisation mathématique. A cet égard, toutes les caractéristiques de la deuxième simulation sont choisies par, et donc connues des chercheurs. Ce n'est pas nécessairement le cas pour la première simulation, où les chercheurs traitent d'un objet matériel qui a encore la capacité de surprendre et de confondre les chercheurs (223).

Dans les deux cas, néanmoins, la simulation consiste en la mise en œuvre d'un modèle mathématique utilisant les lois de la mécanique. L'ordinateur calcule ensuite l'effet de la force sur les éléments individuels de chaque grille et assemble les effets individuels en une mesure globale de la résistance donnée dans la structure osseuse. C'est

intéressant de noter que la simulation permet également un affichage visuel de la façon dont l'interne
la structure osseuse se comporte sous pression, ainsi que le point de fracture.

Les deux simulations sont exploratoires dans leur conception et leur objectif : elles étudient comment les
la structure des os se comporte dans des conditions spécifiques de stress et de pression, et
favorisent ainsi la compréhension du chercheur. De plus, les deux simulations permettent aux chercheurs
de comprendre comment l'architecture des os réagit lors d'accidents réels,
quelles sont les conditions d'une fracture osseuse et comment réparer au mieux les os,
toute information allant au-delà du modèle mathématique mis en œuvre.20

On pourrait bien sûr objecter que ces simulations informatiques, comme toutes les autres simulations
informatiques, sont guidées par la théorie en ce sens que le modèle mathématique avec
le calcul constitue le cadre théorique. Cependant, comme je viens de le mentionner, l'idée des simulations
informatiques en tant qu'expériences exploratoires est de générer
découvertes significatives sur les phénomènes sans entretenir de liens étroits avec la théorie. Laisser
Rappelons-nous qu'une expérience basée sur la théorie pourrait avoir trois significations différentes, aucune
dont, je crois, est applicable au type de simulation informatique ici illustré. Cela ne peut pas signifier que les
attentes concernant les résultats de la simulation
entrent dans le cadre fourni par une théorie donnée, puisque les résultats fournissent des informations qui
ne sont pas contenues dans le modèle mathématique. Cela ne pourrait pas non plus signifier
que la conception de l'expérience dépend, plus ou moins, de la théorie. L'exemple montre très clairement
que, outre la mise en œuvre d'une poignée de lois mécanistes,
il y a peu de théorie à l'appui de la simulation. Enfin, cela ne pouvait pas signifier que le
les instruments utilisés pour l'expérience dépendent fortement de la théorie. Bien qu'il soit vrai
que l'ordinateur est lié par la théorie et la technologie, en principe celles-ci jouent peu
rôle dans la production de résultats fiables.21

L'exemple montre ensuite deux simulations utilisées à des fins exploratoires. Semblable à ce que nous
avons dit à propos des expériences exploratoires, ces simulations génèrent une quantité importante de
découvertes qui ne peuvent pas (facilement) être encadrées dans un cadre théorique bien établi. Le fait que
les deux mettent en œuvre des modèles mathématiques utilisant les lois de la mécanique n'aide pas à tenir
compte des nouvelles preuves.
Le contraire, je crois, est vrai. Autrement dit, la nouvelle preuve obtenue à partir des résultats
contribuer à la consolidation, à la reformulation et à la révision des concepts, principes et hypothèses d'une
théorie médicale sur les os, d'une théorie physique sur les
la résistance des matériaux et les modèles intégrés dans la simulation. Ce sont, selon Steinle (Steinle 2002),
les principales caractéristiques des expériences exploratoires.
En ce sens, les résultats des simulations sont relativement indépendants des fortes
restrictions théoriques incluses dans les modèles de simulation.

---

20 Certains prétendent que l'information qu'une simulation informatique peut fournir est déjà contenue dans les modèles mis en œuvre. Je trouve cette affirmation particulièrement trompeuse pour deux raisons. D'abord,
car il existe plusieurs cas où les simulations informatiques produisent des phénomènes émergents qui ont été
pas strictement contenues dans les modèles mis en œuvre. Ainsi, l'allégation donne des informations erronées sur la portée des
modèles et dénature le rôle des simulations informatiques. Deuxièmement, parce que même si les modèles mis en œuvre contiennent
toutes les informations que la simulation est capable d'offrir, ce fait
ne dit rien sur les connaissances dont disposent les chercheurs. Il est pratiquement impossible et pragmatiquement insensé de
connaître l'ensemble de toutes les solutions d'un modèle de simulation. Précisément pour ces cas,
nous avons des simulations informatiques.

21 Rappelez-vous notre discussion au chapitre 4 sur la fiabilité des simulations informatiques.

De ce point de vue, les simulations informatiques remplissent un rôle de stratégies exploratoires au même titre que l'expérimentation. Naturellement, dans l'évaluation de leurs résultats, les chercheurs doivent encore prendre en compte le fait qu'il s'agit d'un modèle, comme opposé à interagir plus ou moins directement avec un monde non théorisé. Autre que cela, il semble que les activités exploratoires soient profondément impliquées dans l'utilisation et l'épistémique fonctions fournies par des simulations informatiques.22

5.2 Formes de compréhension non linguistiques

5.2.1 Visualisation

L'explication, la prédiction et les stratégies exploratoires se résument à fournir une compréhension du mot au moyen d'une certaine forme d'expression linguistique. Dans le cas de l'explication, c'est assez simple. Les chercheurs reconstruisent le modèle de simulation, une expression logique-mathématique, afin d'offrir une explication de leurs résultats. La prédiction, quant à elle, consiste à produire des résultats qui pourraient quantitativement dire aux chercheurs quelque chose de significatif sur un système cible. Enfin, exploratoire stratégies consistent à générer des découvertes significatives sur des phénomènes sous la forme de données (par exemple, nombres, matrices, vecteurs, etc.). En fin de compte, les trois fonctions épistémiques présentés ici sont liés à des formes de représentations linguistiques. Il y a aussi un mode de représentation alternatif, non linguistique, qui offre des moyens importants de comprendre un système cible : la visualisation des résultats de simulations informatiques.23

Les visualisations étant de véritables formes d'appréhension d'un système cible, elles ne peuvent être considérées comme véhiculant des informations redondantes déjà contenues dans le résultats d'une simulation informatique. L'étude philosophique sur les visualisations résiste toute interprétation qui les réduit à une simple contemplation esthétique ou à un support d'informations plus pertinentes. Au lieu de cela, les visualisations sont considérées comme épistémiquement précieux pour les simulations informatiques à part entière. A cet égard, les visualisations font partie intégrante des arguments du chercheur, et sont soumises à les mêmes principes de force et de solidité, d'acceptation et de rejet que les théories et modèles.

Maintenant, avant que quoi que ce soit puisse être dit sur les visualisations, nous devons d'abord remarquer que il y a plusieurs façons d'analyser le terme. Ici, nous ne sommes pas intéressés par le processus théoriquement alambiqué de post-traitement des résultats d'une simulation dans un visualisation. Cela signifie que nous ne nous intéressons pas aux transformations (par exemple, géométriques, topologiques, etc.), ni au type d'algorithmes qui fonctionnent selon

---

22 Certes, sous cette interprétation, de nombreuses simulations informatiques deviennent exploratoires. Non y voir un problème, car cette caractéristique cadre bien avec la nature et les usages des simulations informatiques. Cependant, je vois des philosophes des sciences s'opposer à mon interprétation, principalement parce que il supprime le statut spécial qui était à l'origine accordé à certains types d'expériences.

23 Étant donné que nous nous intéressons uniquement à la visualisation dans les simulations informatiques, les autres types de visualisation, tels que les graphiques, les photographies, les films vidéo, les radiographies et les images IRM, sont exclus de notre étude.

les données (par exemple, algorithmes scalaires, algorithmes vectoriels, algorithmes tensoriels, etc.). Nous sommes également pas intéressé par une position de post-traitement pour affiner les résultats. Pour nous, les visualisations en elles-mêmes sont le problème principal. Nous sommes intéressés par le résultat visuel d'une simulation informatique utilisée pour l'évaluation épistémologique. En bref, nous nous intéressons au type de compréhension que l'on obtient en visualisant une simulation, et ce que les chercheurs peuvent en faire.

Ici, nous restreignons également notre intérêt à quatre niveaux différents d'analyse de visualisation. Ce sont : la dimension spatiale (c'est-à-dire les visualisations 2D et 3D), la la dimension évolution-temps (visualisations statiques et dynamiques), la dimension manipulabilité (c'est-à-dire les cas où les chercheurs peuvent intervenir dans la visualisation et modifier en y ajoutant des informations, en changeant le point d'observation, etc.) et la dimension de codage (c'est-à-dire les standardisations utilisées pour coder les visualisations, comme la couleur, poste, etc). Chaque dimension individuellement, ainsi qu'en association avec d'autres, offrent différents types de visualisations.

La signification globale des visualisations est qu'elles sont un complexe de distributions spatiales et de relations d'objets, de formes, de temps, de couleurs et de dynamiques. Laura Perini explique que les visualisations sont en partie des interprétations conventionnelles dans le sens où la relation entre une image et le contenu n'est pas déterminée non plus par des caractéristiques de l'image ou par une relation de ressemblance entre l'image et sa référence. Plutôt, quelque chose d'extrinsèque à l'image et à sa référence est nécessaire pour déterminer leur relation de référence (Perini 2006). Un bon exemple de ces conventions est l'utilisation de certaines formes de codage, telles que la couleur – luminosité, contraste, etc. – glyphes, etc. Examinons de plus près la couleur. Le bleu représente toujours le froid dans une simulation incluant la température. Ce serait enfreindre les règles tacites de la pratique scientifique que de le changer en rouge ou en vert. De même pour certains symboles : une flèche, par exemple, représente aller - se déplacer, se déplacer - en avant - vers le haut, vers - quand son le sommet pointe vers le haut. Étant donné que l'interprétation et la compréhension des visualisations sont souvent trop naturel et automatique pour les chercheurs, des changements introduiraient la confusion et des retards inutiles dans la pratique scientifique.[24] En ce sens, l'utilisation et l'application des bonnes conventions est nécessaire pour comprendre les représentations visuelles (864).

Considérons maintenant un cas de modification de la distribution spatiale d'une image par ailleurs standard. En discutant avec un groupe de collègues de l'Université du Colorado, à Boulder, à propos de retourner une carte du monde à l'envers est un exemple tiré d'une expérience personnelle. Dans ce cas, l'Amérique du Sud serait « au-dessus » et l'Amérique du Nord serait 'ci-dessous.' « Nord » et « Sud » seraient également déplacés. Retourner le la carte de cette manière pourrait être une déclaration politique, puisque nous attribuons généralement "au-dessus" – dans la répartition spatiale sur la carte – être en quelque sorte meilleur ou plus important. Mes collègues ont trouvé l'idée assez séduisante, car elle provoque une "crampe mentale" voir le monde « à l'envers »[25]. Je veux simplement dire par là qu'il nous faudrait quelques

---

[24] Perini explore cette idée dans (Perini 2004, 2005).

[25] Deux points à souligner ici. Premièrement, l'idée d'une « crampe mentale » vient de Wittgenstein, qui dit que les problèmes philosophiques sont comparés à une crampe mentale à soulager ou à un nœud dans notre pensant être déliée (Wittgenstein 1976). Deuxièmement, il n'y a pas de monde « à l'envers », puisqu'il est simplement la façon dont nous - en tant qu'humains - avons décidé de le représenter. Tant que le nord et le sud sont conservés

secondes pour comprendre la nouvelle distribution spatiale du continent. Comme mentionné,
la normalisation des couleurs, des symboles et de la notation est fondamentale
interprétations qui facilitent la libre circulation de la pratique scientifique.

Dans ce contexte, les visualisations informatiques peuvent remplir plusieurs fonctions épistémiques.
Pensez à quel point un chercheur pourrait comprendre un système cible en regardant simplement
à la relation des objets répartis dans l'espace et dans le temps, leurs propriétés visuelles et leur comportement,
les couleurs, etc. Ils contribuent tous au sens global
de ce que le chercheur observe. De plus, les visualisations facilitent l'identification des problèmes dans l'ensemble
des équations qui constituent la simulation informatique.
En d'autres termes, une visualisation peut montrer où quelque chose s'est mal passé, inattendu, ou simplement
montrer une fausse hypothèse dans le modèle. Cela ne veut bien sûr pas dire
que les visualisations sont capables de fournir des solutions techniques. Comparez cette idée avec
méthodes de vérification et de validation. Dans ces derniers cas, ces méthodes visent à
localiser un ensemble de problèmes (par exemple, de mauvaises dérivations mathématiques dans la discrétisation
processus) et de fournir les outils théoriques qui conduisent à une solution. Dans le cas d
visualisation, l'identification des erreurs se fait par inspection visuelle, et cela dépend donc
sur l'œil exercé des chercheurs. En ce sens, les visualisations sont très utiles pour
aider à décider de différentes lignes de conduite et fournir des bases pour une
décisions, mais ils n'offrent pas d'outils formels pour résoudre les problèmes dans la simulation
modèle. À cet égard, les visualisations sont également importantes car elles étendent les utilisations
des simulations informatiques dans le domaine social et politique.

Permettez-moi maintenant d'illustrer cette discussion plutôt abstraite avec un exemple de tornade.
L'un des principaux problèmes de la science des tempêtes est que la quantité d'informations dont disposent les
chercheurs sur une véritable tornade est plutôt limitée. Même avec des
images satellites atmosphériques et chasseurs de tornades, l'état actuel des connaissances scientifiques
l'instrumentation ne peut pas fournir aux chercheurs une image complète de ce qui se passe.
Pour ces raisons et d'autres telles que la commodité et la sécurité de l'étude des tor nades à partir d'un ordinateur
de bureau, les chercheurs sont plus enclins à étudier ces phénomènes naturels à l'aide de simulations
informatiques. À cet égard, Lou Wicker du
Le laboratoire national des tempêtes violentes de la NOAA déclare que «[f] sur le terrain, nous ne pouvons pas comprendre
complètement ce qui se passe, mais nous pensons que le modèle informatique est un raisonnable
approximation de ce qui se passe, et avec le modèle, nous pouvons capturer l'ensemble
histoire » (Barker 2004).

La simulation que Lou Wicker, Robert Wilhelmson, Leigh Orf et d'autres recréent est la genèse d'un
supertwister similaire à celui vu à Manchester, Dakota du Sud.
en juin 2003. La simulation est alors basée sur un modèle de tempête qui comprend des équations de mouvement
pour les substances atmosphériques et aquatiques (par exemple, les gouttelettes, la pluie, la glace) et les tailles de grille
allant d'une résolution uniforme de cinq mètres à une résolution beaucoup plus élevée. Les données sont
utilisé pour semer les conditions pré-tornade, telles que la vitesse du vent, la pression atmosphérique et l'humidité
près de Manchester à l'époque. Étant donné que les données sont assez rares, il pourrait
représentent des points séparés par des distances de vingt mètres jusqu'à trois kilomètres.

---

fixe, nous sommes capables de représenter le globe comme nous l'aimons - un bon exemple en est le logo de la
Les Nations Unies. Pour une représentation artistique de cela, voir le travail de Jose Torres Garc ´ıa, America ´
Invertida, 1943.

Cela signifie que les chercheurs doivent tenir compte d'un tel éventail de distribution spatiale lors de l'analyse de la visualisation.

Naturellement, la visualisation de la tornade est réduite. L'échelle spatiale de la tornade va de quelques kilomètres de haut à quelques centaines de kilomètres de large et de profondeur. La durée totale de la simulation peut également varier en fonction de la résolution de la tornade et du nombre d'éléments qu'elle contient - par exemple, s'il y a plus d'une tornade ou si elle simule également la destruction d'une ville entière. En outre, l'échelle de temps d'une tornade dépend également de sa formation initiale, de son évolution et de sa disparition. Pour ces raisons, les simulations informatiques sont généralement mesurées en « heures de tempête ». La visualisation d'une tornade de type F3 au sein d'une super cellule dans (Wilhelmson et al. 2005) représente une heure d'évolution de la tempête, bien que la visualisation ne prenne qu'environ une minute et demie.

Notons également que l'intégration et la chorégraphie de la visualisation sont aussi importantes que la visualisation elle-même. Des choix doivent être faits afin de se concentrer sur les données et les événements les plus significatifs. Dans la visualisation de la tornade, des milliers de trajectoires calculées ont été éditées jusqu'à quelques-unes des plus significatives.
Il y a généralement deux raisons pour modifier les visualisations de cette manière : soit il y a des données qui ne sont pas pertinentes pour une visualisation, soit il y a des données qui ne sont pas pertinentes à certaines fins dans la visualisation. Par exemple, Trish Barker rapporte qu'une visualisation complète de (Wilhelmson et al. 2005) « ressemblerait à une assiette de cheveux d'ange » (Barker 2004) en raison de la quantité écrasante d'informations à afficher. Dans de tels cas, la visualisation échoue dans son objectif, car elle fournirait peu de compréhension du phénomène en question - un point similaire est soulevé avec la réduction de la visualisation dans le temps et l'espace.

Dans la plupart des cas, cependant, les chercheurs utilisent toutes les données disponibles, ils se contentent de mettre en évidence différents aspects de celles-ci. De cette manière, il est possible d'obtenir des informations distinctes à partir d'une même simulation. Et des informations distinctes conduisent à différents plans d'action, à des mesures de prévention et à l'identification de problèmes dans la simulation qui n'avaient pas été anticipés lors des étapes de conception et de programmation. Un exemple intéressant de ce dernier est donné par Orf, qui mentionne que dans la visualisation de la genèse de la tornade, la pluie ne centrifugeait pas hors de la tornade comme elle le devrait. Il conclut : « c'est quelque chose sur quoi nous devons travailler » (Orf et al. 2014).

Le but de la simulation n'est donc pas simplement de calculer des modèles mathématiques complexes, mais plutôt de visualiser la structure, la formation, l'évolution et la disparition de grandes tornades dommageables produites dans les supercellules. Pour y parvenir, un aspect crucial de la visualisation est son réalisme. C'est-à-dire que la visualisation doit avoir une résolution suffisante pour capturer l'afflux de tornade de bas niveau, les puits de précipitation minces qui forment des `` échos en crochet '' adjacents à la tornade principale, les nuages et, si possible, le niveau du sol au-dessus duquel la tornade passe et tous les débris ça se disperse.

Le réalisme visuel est plus qu'une caractéristique esthétique, il est essentiel à l'évaluation des simulations informatiques. Orf se rappelle avoir demandé aux chasseurs de tornade leur avis sur la visualisation de sa simulation (voir figure 5.1 et figure 5.2). Pour les experts de terrain, la visualisation semble assez raisonnable, c'est-à-dire que le réalisme visuel est convaincant malgré le manque de compréhension approfondie du modèle de simulation et des méthodes de visualisation de la tornade. Pour les chasseurs de tornades, il n'y a que l'aspect esthétique

expérience qui se rapproche de l'expérience réelle. Mais cela n'enlève rien à leur appréciation de la puissance épistémologique des simulations informatiques. Fiabilité de la modèle de simulation et la confiance dans les résultats sont fournis par Orf et son équipe, car ils en sont responsables dans la division du travail.

Un autre élément important pour le réalisme visuel est le codage des couleurs et des glyphes. Comme Robert Patterson raconte que les tubes de courant colorés représentent le mouvement et la vitesse de particules lorsqu'elles sont libérées dans le flux d'air, montrant la géométrie du flux d'air dans et autour la tornade.26 La variation de couleur des tubes de courant transmet des informations supplémentaires sur la température de l'air et le processus de refroidissement-réchauffement - tubes de courant sont orange en montant et bleu clair en descendant. Fait intéressant, la couleur joue un rôle informatif supplémentaire, comme la mise en évidence de la pression et du taux de rotation de la tornade. Les sphères dans le vortex à basse pression représentent la tornade en développement, qui s'élève dans le courant ascendant et sont colorées par la pression (voir figure 5.1 et figure 5.2). Inclinaison les cônes colorés en fonction de la température représentent la vitesse et la direction du vent au surface, et montrent l'interaction de l'air chaud et froid autour de la tor nade en développement (Wilhelmson et al. 2010).

Donna Cox, chef de la division des technologies expérimentales du NCSA, capture assez poétiquement l'effort d'équipe impliqué dans le développement de cette visualisation. Elle l'imagine collègues comme « une équipe de renaissance très travaillante et collaborative » (Barker 2004). En disant cela, Cox rend explicite le fait que cette visualisation particulière est un complexe processus, qui implique une consultation entre les chercheurs et les disciplines. Aboyeur explique en outre qu'à chaque étape du processus, les chercheurs sont tenus de consulter les uns aux autres afin de prendre des décisions éclairées sur la sélection des données et la meilleure façon d'en tirer des informations significatives.

À ce stade, il est intéressant de noter qu'une tornade satellite contrarotative apparaît du côté de la tornade principale (voir figure 5.2). Il s'agit d'un phénomène que le rapport d'experts comme rarement observé dans la nature, et il a été rarement enregistré par chasseurs de tempête. En fait, la deuxième tornade satellite contrarotative n'a pas été observée à Manchester, et les chercheurs ne s'attendaient pas non plus à son apparition. Néanmoins, les chasseurs de tor nado reconnaissent la possibilité d'une tornade satellite apparaissant compte tenu de la bonnes hypothèses intégrées dans la simulation informatique. Nous pourrions conclure que cela La deuxième tornade déroute les chercheurs dans le sens où elle est inattendue mais empiriquement possible. Un bref rappel ici s'impose. Au chapitre 3, nous avons parlé de Marie Les affirmations de Morgan selon lesquelles les simulations informatiques surprennent les chercheurs mais ne confondent pas parce que le comportement de la simulation peut être retracé et expliqué dans termes du modèle sous-jacent. Le lecteur est maintenant invité à revisiter ce chapitre dans lumière de cet exemple.

Revenant à l'analyse des visualisations dans les simulations informatiques, nous voyons que les chercheurs pourraient créer une simulation informatique très complexe d'une tornade en « temps réel » et la visualiser avec ses propriétés et son comportement. Grâce à ces visualisations, les chercheurs peuvent comprendre la formation, l'évolution et la disparition d'une tornade beaucoup plus efficacement que via toute autre forme linguistique (par exemple, des modèles mathématiques et de simulation, des matrices, des vecteurs ou toute forme numérique de représentation

26 Pour une vidéo complète montrant le développement de la tornade, voir http://avl.ncsa.illinois.edu/wp content/uploads/2010/09/NCSA F3 Tornade H264 864.mov

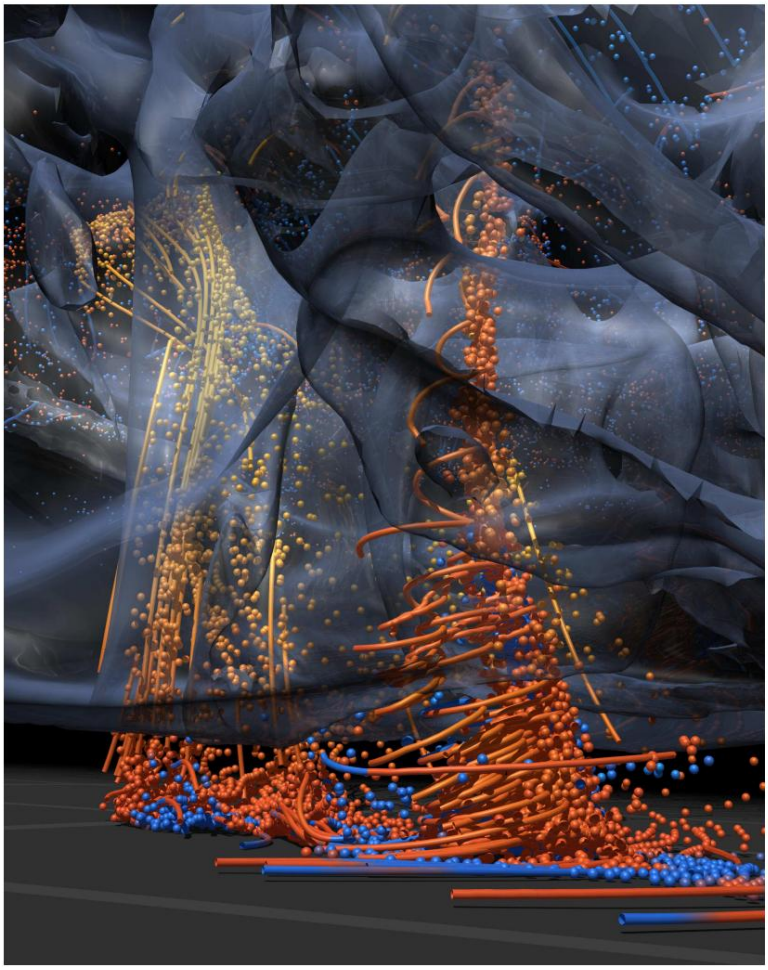


Fig. 5.1 Image de la visualisation d'une tornade de type F3 avec formation de nuages. Créé par le laboratoire de visualisation avancée du NCSA. Avec l'aimable autorisation du National Center for Supercomputing Applications (NCSA) et du conseil d'administration de l'Université de l'Illinois.

résultats de simulations informatiques). Les chercheurs sont également en mesure de rendre compte de la tornade inattendue du satellite à contre-rotation, de prédire son évolution, de décrire ses performances, d'étudier sa trajectoire et d'analyser les conditions initiales qui ont rendu possible cette tornade du satellite en premier lieu.

Ainsi comprise, cette visualisation est utile pour plusieurs efforts épistémiques tels que l'explication d'une série de questions liées à la tornade, la mesure de leurs valeurs internes et la prédiction des dommages potentiels de la tornade. Notons que le sens donné ici à l'explication et à la prédiction est d'être des fonctions épistémiques qui dépendent d'une base non linguistique. En ce sens, ces formes d'explication et de prédiction dépendent

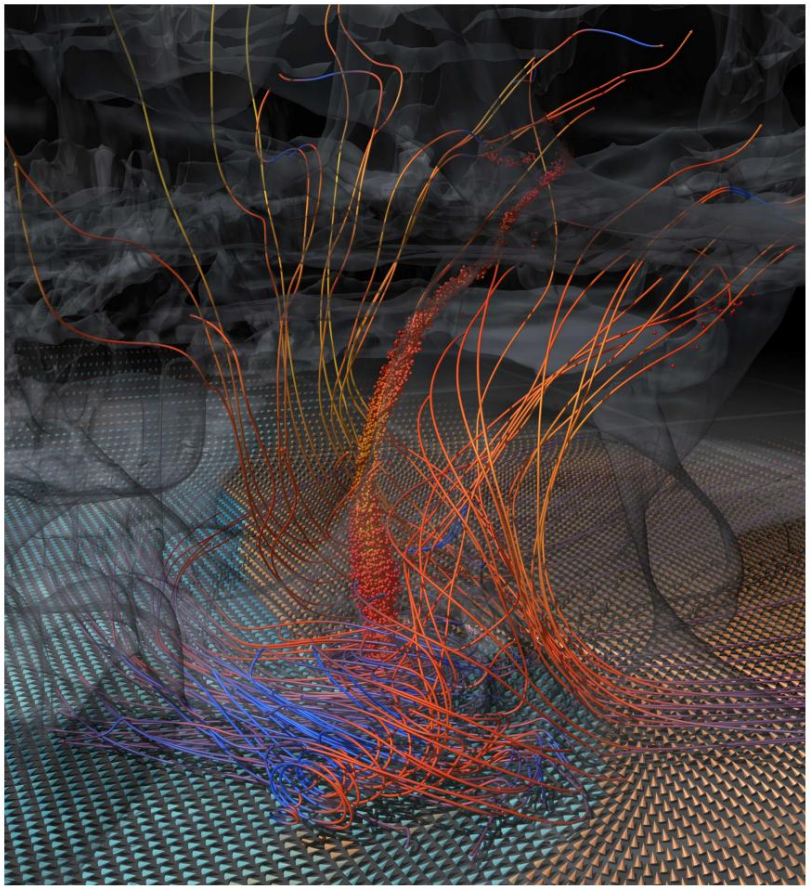


Fig. 5.2 Image de la visualisation d'une tornade de type F3 avec la formation d'une tornade satellite.
Créé par le laboratoire de visualisation avancée du NCSA. Avec l'aimable autorisation du National Center for Supercomputing Applications (NCSA) et du conseil d'administration de l'Université de l'Illinois.

sur une perception psychologique et esthétique des tornades, plutôt que sur une reconstruction linguistique de la visualisation – si cela est possible. Une autre utilisation importante de cette visualisation est la validation des données d'entrée initiales ainsi que la fiabilité du modèle de simulation sous-jacent. Si le comportement de la tornade s'écarte trop d'une vraie tornade - ou de tout ce que les experts anticipent - alors les chercheurs ont des raisons de croire que la simulation, les données d'entrée ou les deux sont incorrectes.

La visualisation de la tornade – représentée statiquement dans les figures 5.1 et 5.2 – peut être résumée comme une évolution dynamique 3D avec la possibilité de manipuler les conditions initiales et aux limites. Une caractéristique de ce type de visualisation, ainsi que de la plupart des visualisations dans les simulations informatiques, est qu'elles sont affichées sur un écran d'ordinateur. Bien que souligner quelque chose d'aussi évident puisse sembler arbitraire,

elle s'accompagne de limitations spécifiques quant à la capacité du chercheur à manipuler, visualiser et finalement comprendre les visualisations. Je reviendrai sur ce point à la fin de la rubrique.

Des formes plus sophistiquées de visualisations se trouvent dans les installations de recherche scientifique haut de gamme. Deux formes viennent à l'esprit, à savoir la «réalité virtuelle» (VR) et la «réalité augmentée» (AR). La première fait référence aux visualisations où le chercheur est capable d'interagir avec eux en utilisant des gadgets spéciaux, tels que des lunettes et une "baguette de souris". Au High Performance Computing Center Stuttgart (HLRS) de l'Université de Stuttgart, des chercheurs ont recréé une réplique entière de Forbach, une ville en Forêt-Noire, ainsi que la centrale électrique de Rudolf Fettweis. Projeter le visualisation sur une pièce spéciale - connue sous le nom d'environnement virtuel automatique de grotte (CAVE)27 – les chercheurs peuvent se promener dans Forbach, pénétrer à l'intérieur du barrage ainsi que turbines souterraines et inspecter le réservoir, le tout à l'aide d'un ensemble spécial de lentilles. De plus, les chercheurs pouvaient entrer dans n'importe quelle maison de la région et observer comment les projets de construction et de modernisation affectent les citoyens, la vie sauvage et l'environnement en général (Gedenk 2017). Le fait que la simulation soit, pour ainsi dire, en dehors de l'écran de l'ordinateur, introduit des avantages significatifs en termes de pratique. de simulations informatiques, ainsi que la compréhension des résultats et la communication eux.

Fig. 5.3 L'environnement virtuel automatique de la grotte (CAVE). Le CAVE offre aux chercheurs une pleine plonger dans un environnement de simulation 3D pour analyser et discuter de leurs calculs. Créé par le Département de visualisation au High-Performance Computing Center Stuttgart (HLRS), Université de Stuttgart. Avec l'aimable autorisation du HLRS, Université de Stuttgart.

27 La CAVE est une salle noire de trois mètres sur trois sur trois mètres avec cinq projecteurs DLP à puce unique avec une résolution de 1920 par 1200 pixels, chacun envoyant une image respective créant une image précise. rendu pour l'œil humain. Quatre caméras sont installées aux coins du plafond pour le suivi les entrées des chercheurs par les lunettes et la souris-baguette.

La deuxième forme de visualisation est connue sous le nom de «réalité augmentée» et, comme son nom l'indique, consiste à agréger des informations simulées sur le monde réel. Fait intéressant, il existe de nombreuses façons de le faire. Au HLRS, un cas typique de RA consiste à superposer des calculs pré-simulés sur une entité réelle taguée. Avec l'aide de technologies avancées, telles que les marqueurs de code et les caméras spéciales, la vision par ordinateur et la reconnaissance d'objets, la quantité de perspicacité et de compréhension d'un objet dans le monde réel peut être considérablement augmentée.

Un exemple simple mais illustratif de RA est le débit d'eau simulé de manière interactive et affiché autour d'une turbine Kaplan (voir figure 5.4). À cette fin, les chercheurs doivent d'abord pré-simuler le débit d'eau autour de la turbine à l'aide d'équations standard de dynamique d'écoulement - au HLRS, les chercheurs utilisent Fenfloss, un solveur rapide de Navier-Stokes qui calcule le débit d'eau. La turbine simulée, quant à elle, est un modèle de l'architecture de la turbine réelle, et doit donc être aussi précise que possible.

Une fois que les données pré-simulées sont disponibles, les chercheurs marquent avec des marqueurs de code des endroits spécifiques sur la turbine réelle pour visualiser le débit simulé. Avec ces informations, un générateur de maillage paramétrique crée la surface et le maillage informatique de la turbine hydraulique. Après quelques secondes de traitement, les résultats de la simulation sont affichés sur la turbine réelle imitant le flux réel.[28]

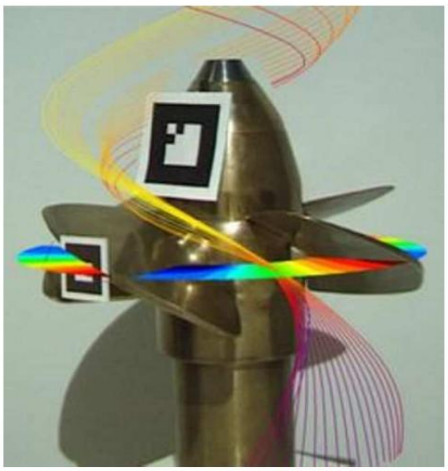


Fig. 5.4 Le débit d'eau est simulé de manière interactive et affiché au sommet d'une turbine Kaplan. Créé par le Département de visualisation du Centre de calcul haute performance de Stuttgart (HLRS), Université de Stuttgart. Avec l'aimable autorisation de Stellba Hydro GmbH & Co KG.

[28] Il existe quelques cas où la simulation est calculée en temps réel pendant la session AR. Le La principale préoccupation de ce type de technologie, cependant, est qu'elle est trop lente et prend trop de temps.

Demandons-nous maintenant quel type d'avantages épistémologiques la réalité virtuelle et la réalité augmentée ont-elles à offrir ?29 Comme on pouvait s'y attendre, chaque forme de visualisation a une valeur épistémique différente et fournit donc différentes formes de compréhension. Un dénominateur commun, cependant, est que la réalité virtuelle et la réalité augmentée « naturalisent » la simulation dans le sens où les chercheurs manipulent la visualisation comme s'ils manipulaient une chose réelle dans le monde. On pourrait dire que la simulation devient une extension naturelle du monde, un «morceau» de celui-ci dans un sens similaire attribué à l'expérimentation. Pour la réalité virtuelle, la naturalisation prend la forme de « marcher » à l'intérieur de la simulation, de « regarder » au-dessous et au-dessus, de « toucher » des objets, de « changer » leur emplacement, etc. Pour la RA, la naturalisation découle de l'intégration de la simulation dans le monde réel et le monde réel dans la simulation. L'écoulement simulé apparaît comme l'écoulement naturel et la vraie turbine fait partie de la simulation.

En conséquence, les résultats de la simulation deviennent épistémiquement plus accessibles qu'une simple visualisation sur l'écran de l'ordinateur, aussi réaliste et sophistiqué soit-il. AR et VR amènent la simulation dans le monde, et le monde dans la simulation. Ils s'emboîtent l'un dans l'autre, mélangeant le réel et le simulé en une nouvelle forme de réalité naturalisée. Le succès de la réalité augmentée et de la réalité virtuelle, en somme, est qu'ils exigent beaucoup moins d'efforts cognitifs de la part des chercheurs - ainsi que des décideurs politiques, des politiciens et du grand public - car ils font de la pratique de la simulation une expérience scientifique et technique plutôt naturelle.

Concentrons-nous d'abord sur la réalité virtuelle. Comme le montre la figure 5.3, les deux chercheurs se tiennent près d'une surface d'eau de manière très réaliste. L'image montre également les chercheurs tenant une "souris-baguette" spéciale qui les aide à naviguer dans la visualisation ainsi que dans le menu affiché sur le côté droit de la CAVE. L'utilisation de lunettes 3D est également importante, car elles aident à orienter la visualisation de telle sorte qu'elle semble naturelle à l'œil humain.

Grâce à ces gadgets, la navigation est simplifiée. La souris-baguette permet au chercheur de « marcher » dans la simulation, et les lunettes 3D permettent aux chercheurs de « regarder » derrière, au-dessus ou en dessous de différents objets simulés de la même manière qu'ils le feraient dans la vraie vie. Comme mentionné précédemment, l'avantage incontestable de l'utilisation de la RA est qu'elle amène la simulation dans le monde réel. Mais il existe également plusieurs autres avantages de la RA destinés à faciliter notre compréhension d'un système cible donné.

Lors d'une conversation au HLRS, Thomas Obst et Wolfgang Schotte me disent que la RA réussit très bien à intégrer les décideurs politiques, les politiciens et le grand public dans les résultats de la simulation. Reprenons l'exemple de la turbine (figure 5.4). Il est possible de montrer les résultats des simulations aux parties prenantes concernées sans avoir besoin d'être dans le CAVE. Au lieu de cela, les résultats d'une simulation peuvent être visualisés in situ avec la turbine, un ordinateur portable et une caméra. La portabilité est importante. En fait, la RA facilite l'explication de problèmes techniques complexes de manière simple et organique aux chercheurs qui n'ont pas été initialement impliqués dans la simulation. « Des chercheurs, mais aussi des politiques et le grand public » dit Obst,

---

[29] Je remercie Thomas Obst et Wolfgang Schotte du HLRS de m'avoir expliqué les détails de leur intéressant travail.

"peut relier et comprendre les résultats de la simulation d'une manière beaucoup plus facile avec AR que sur un écran d'ordinateur, ou même le CAVE."

Malheureusement, la RA a ses limites. Un problème central qui inquiète de nombreux chercheurs est que la RA dépend d'une étape pré-calculée, c'est-à-dire que les résultats visualisés ne sont pas calculés en temps réel mais plutôt pré-calculés. De ce fait, sa contribution à la simulation est limitée de plusieurs manières. Par exemple, aucune erreur en temps réel ne peut être détectée dans un environnement AR. De plus, si les conditions de conception changent (par exemple, des éléments ont été ajoutés ou soustraits à la simulation d'origine ou à la configuration du matériau), l'environnement AR peut être rendu totalement inutile.30 Un avantage important de la RV et de la RA par rapport aux

visualisations sur écran d'ordinateur est qu'ils n'obligent ni les chercheurs ni le public (hommes politiques, décideurs, etc.) à adopter une perspective fixe plutôt qu'une autre. Ce point a à voir avec la « naturalisation » d'une visualisation apportée par la réalité virtuelle et la réalité augmentée mentionnée précédemment.
Les chercheurs et le public peuvent manipuler la simulation, et ainsi concentrer leur attention sur ce qui leur importe le plus. Comparez cela avec une visualisation sur un écran d'ordinateur. L'utilisation d'une souris – ou peut-être d'un écran tactile – limite les points d'attention, fixe l'ordre d'importance et ménage une perspective donnée. Tout en utilisant des visualisations sur écran pour communiquer les résultats au public, les chercheurs ont une perspective pré-choisie de ce qu'il faut montrer (par exemple, en sélectionnant l'angle, en naviguant dans le menu des options, etc.). Au lieu de cela, lorsque la communication des résultats se fait via VR ou AR, le public est en mesure d'interagir avec la visualisation d'une manière différente, et donc de la comprendre selon ses propres termes, sans aucune perspective pré-choisie.

Dans un sens simple, avec la réalité virtuelle et la réalité augmentée, les chercheurs et le public n'ont pas besoin d'en savoir beaucoup sur les modèles de simulation ni sur leurs hypothèses pour comprendre les résultats des simulations informatiques. Alors que cela est également vrai pour les visualisations sur écran d'ordinateur, la réalité virtuelle et la réalité augmentée naturalisent l'expérience de visualisation.

5.3 Remarques finales

Tout au long de ce livre, j'ai soutenu que les simulations informatiques jouent un rôle central dans la recherche scientifique et technique contemporaine. Leur importance réside dans la puissance épistémologique qu'elles procurent en tant que méthodes de recherche. Dans le chapitre précédent, j'ai soutenu que les simulations informatiques fournissent des connaissances sur un système cible. Dans ce chapitre, je montre comment la compréhension est obtenue en utilisant des simulations informatiques dans la recherche scientifique et technique.

J'ai ensuite divisé le chapitre en deux formes de compréhension, à savoir les formes linguistiques et les formes non linguistiques de compréhension. La distinction vise à montrer qu'il existe des formes de compréhension qui dépendent d'une forme symbolique (par exemple, des formules, des spécifications, etc.) et d'autres formes de compréhension qui dépendent de

30 Ces problèmes pourraient être surmontés en calculant et en visualisant les résultats dans l'environnement AR en temps réel. Malheureusement, ce type de technologie a des coûts élevés en termes de processus de calcul, de temps et de stockage de mémoire.

formes non symboliques (par exemple, visualisations, sons, interactions). Les cas qui se qualifient comme les premiers sont la force explicative des simulations informatiques, leur utilisation prédictive et la possibilité de générer des découvertes significatives sur des phénomènes du monde réel. Chacun d'eux dépend, d'une manière ou d'une autre, du modèle de simulation en tant qu'ensemble de formules.

Quant aux formes de compréhension non linguistiques, j'ai abordé le cas des visualisations dans les simulations informatiques. Il est bien connu parmi les communautés scientifiques et d'ingénierie que la visualisation est un véhicule important pour comprendre les résultats des simulations informatiques. Dans la section consacrée à ce sujet, j'aborde les visualisations standard sur écran d'ordinateur, et les moins courantes mais centrales à des fins scientifiques et d'ingénierie, la réalité virtuelle et la réalité augmentée.

Les références

Ajelli, Marco, Bruno Gonc¸alves, Duygu Balcan, Vittoria Colizza, Hao Hu, Jose´ J. Ramasco, Stefano Merler et Alessandro Vespignani. 2010. "Comparaison des approches informatiques à grande échelle à la modélisation épidémique : modèles de métapopulation basés sur les agents et modèles structurés." BMC Infectious Diseases 10 (190): 1–13.

Balcan, Duygu, Vittoria Colizza, Bruno Gon\'calves, Hao Hu, Jose J Ramasco et Alessandro Vespignani. 2009. "Réseaux de mobilité multi-échelles et propagation spatiale des maladies infectieuses." Actes de l'Académie nationale des sciences des États-Unis d'Amérique 106 (51): 21484–21489.

Barker, Trish. 2004. "Chasse au supertwister." National Center for Supercomputing Applications - Université de l'Illinois à Urbana-Campaign. http://www. ncsa.illinois.edu/news/stories/supertwister.

Barnes, E. 1994. "Expliquer les faits brutaux." Philosophie des sciences 1: 61–68.

Barrett, Jeffrey A. et P. Kyle Stanford. 2006. « La philosophie des sciences. Une encyclopédie." Type. Prédiction, édité par S. Sarkar et J. Pfeifer, 585–599. Routledge.

Beck, JD, BL Canfield, SM Haddock, TJH Chen, M Kothari et TM Keaveny. 1997. "Imagerie tridimensionnelle de l'os trabéculaire à l'aide de la technique de fraisage à commande numérique par ordinateur." Os 21 (3): 281–287.

Bœhm, C., JA Schewtschenko, RJ Wilkinson, CM Baugh et S. Pascoli. 2014. « Utilisation des satellites de la Voie lactée pour étudier les interactions entre la matière noire froide et le rayonnement » [en en]. Avis mensuels de la Royal Astronomical Society : Lettres 445, no. 1 (novembre): L31–L35. Consulté le 18 juillet 2016.

Brewer, William F et Bruce L Lambert. 2001. "La charge théorique de l'observation et la charge théorique du reste du processus scientifique." Philosophie des sciences 68 (S3): S176–S186.

Craver, Carl F. 2007. Expliquer le cerveau : les mécanismes et l'unité de mosaïque de neurosciences. Presse universitaire d'Oxford.

Daston, Lorraine et Peter Galison. 2007. Objectivité. Livres de zone.

Douglas, Heather. 2009. Science, politique et idéal sans valeur. Presse de l'Université de Pittsburgh.

Duran, Juan M. 2017. "Variation de la durée explicative : explication scientifique pour les simulations informatiques." Études internationales en philosophie des sciences 31 (1): 27–45.

Elgin, C. 2009. « La compréhension est-elle factice ? » Dans Epistemic Value, édité par A. Had dock, A. Millar et DH Pritchard, 322–330. Presse universitaire d'Oxford.

Elgin, Catherine. 2007. "La compréhension et les faits." Études philosophiques 132 (1) : 33–42.

Fahrbach, Ludwig. 2005. "Comprendre les faits brutaux." Synthèse 145, no. 3 (juillet): 449–466.

Franklin, Allan. 1986. La négligence de l'expérience. La presse de l'Universite de Cambridge.

Friedman, Michel. 1974. « Explication et compréhension scientifique ». Le journal de Philosophie 71 (1): 5–19.

García, Pío et Marisa Velasco. 2013. "Stratégies exploratoires : expériences et simulations". Dans Computer Simulations and the Changing Face of Scientific Experimentation, édité par Juan M. Duran et Eckhart Arnold. Édition des boursiers de Cambridge.

Gedenk, Éric. 2017. "La salle de visualisation CAVE plonge les chercheurs dans les simulations." https : / / www . hlrs. de / fr / solutions - services / portefeuille-de-services/visualisation/ réalité-virtuelle/.

Gelfert, Axel. 2016. Comment faire de la science avec des modèles. Mémoires Springer en philosophie . Springer. ISBN : 978-3-319-27952-7 978-3-319-27954-1, consulté le 23 août 2016.

————. 2018. "Modèles à la recherche de cibles : modélisation exploratoire et cas des modèles de Turing." Dans Philosophy of Science, édité par A. Christian, D. Hommen, N. Retzlaff et G. Schurz, 245–271. Springer.

Piratage, Ian. 1992. "L'auto-justification des sciences de laboratoire." Dans Science as Practice and Culture, édité par Andrew Pickering, 29–64. Presse de l'Université de Chicago.

Halloran, M Elizabeth, Neil M Ferguson, Stephen Eubank, Ira M Longini, Derek AT Cummings, Bryan Lewis, Shufu Xu, Christophe Fraser, Anil Vullikanti, Timo thy C Germann, et al. 2008. "Modélisation du confinement en couches ciblé d'une pandémie de grippe aux États-Unis." Actes de l'Académie nationale des Sciences 105 (12): 4639–4644.

Hanson, Norwood Russel. 1958. Modèles de découverte : une enquête sur les fondements conceptuels de la science. La presse de l'Universite de Cambridge.

Hempel, C., et P. Oppenheim. 1948. « Études sur la logique de l'explication ». Philosophie des sciences 15 (2): 135–175.

Hempel, Carl G. 1965. Aspects de l'explication scientifique et autres essais dans le Philosophie des sciences. La presse libre.

Karaca, Koray. 2013. "Les sens forts et faibles de la charge théorique de l'expérimentation : expériences axées sur la théorie versus expériences exploratoires dans l'histoire de la physique des particules à haute énergie. Science en contexte 26 (1): 93–136.

Keaveny, TM, EF Wachtel, XE Guo et WC Hayes. 1994. "Mécanique Comportement de l'os trabéculaire endommagé » [en anglais]. Journal de biomécanique 27, Non. 11 (novembre): 1309–1318.

Kitcher, Philippe. 1981. "Unification explicative". Philosophie des sciences 48 (4): 507–531.

————. 1989. «Unification explicative et structure causale du monde». Dans Explication scientifique, édité par Philip Kitcher et Wesley C. Salmon, 410–505. Presse de l'Université du Minnesota.

Krohs, Ulrich. 2008. "Comment les simulations informatiques numériques expliquent les processus du monde réel." Études internationales en philosophie des sciences 22 (3): 277–292.

Kuhn, TS 1962. La structure des révolutions scientifiques. Université de Chicago Presse.

Lipton, Peter. 2001. "A quoi bon une explication?" Dans Explication, édité par G. Hon et S. Rackover, 43–59. Springer.

Lloyd, Elisabeth A. 1995. "Objectivité et double standard pour l'épisté féministe mologies. Synthèse 104 (3): 351–381.

Morgan, Mary S. 2003. "Expériences sans intervention matérielle." Dans The Philosophy of Scientific Experimentation, édité par Hans Radder, 216–235. University of Pittsburgh Press.

Niebur, Glen L, Michael J Feldstein, Jonathan C Yuen, Tony J Chen et Tony M Keaveny. 2000. "Modèles d'éléments finis à haute résolution avec résistance des tissus l'asymétrie prédit avec précision l'échec de l'os trabéculaire. Journal of biome chanics 33 (12): 1575–1583.

Orf, Leigh, Robert Wilhelmson, Lou Wicker, BD Lee et CA Finley. 2014. Parler
3B.3 lors de la Conférence sur les orages locaux violents. Madison, novembre. https:
//ams.confex.com/ams/27SLS/webprogram/Paper255451.
html%20https://www.youtube.com/watch?v=1JC79gzZykU%
5C&t=330s.

Perini, Laure. 2004. "Représentations visuelles et confirmation." Philosophy of Science 72 (5): 913–
916.

————. 2005. "La vérité en images." Philosophie des sciences 72 (1): 262–285.

————. 2006. "Représentation visuelle". Type. Représentation visuelle dans la philosophie des
sciences. Une encyclopédie, éditée par Sahotra Sarkar et Jessica Pfeifer,
863–870. Routledge.

Salmon, Wesley C. 1984. Explication scientifique et structure causale de la
Monde. Presse universitaire de Princeton.

————. 1989. Quatre décennies d'explication scientifique. Université de Pittsburgh
Presse.

Schiafonati, Viola. 2016. "Élargir la notion traditionnelle d'expérience en informatique : expériences
exploratoires." Éthique des sciences et de l'ingénierie 22, no. 3
(juin): 647–665. ISSN : 1471-5546. doi : 10.1007/s11948-015-9655-z.
https://doi.org/10.1007/s11948-015-9655-z.

Schurz, Gerhard et Karel Lambert. 1994. "Esquisse d'une théorie de la compréhension scientifique."
Synthèse 101 : 65–120.

Steinle, Friedrich. 1997. « Entrer dans de nouveaux domaines : utilisations exploratoires d'experimenta
tion. Philosophie des sciences 64: S65–S74.

————. 2002. "Expériences en histoire et philosophie des sciences." Perspectives sur
Sciences 10 (4): 408–432.

Stuewer, Roger H. 1985. "Désintégration artificielle et controverse Cambridge-Vienne." Observation,
expérience et hypothèse dans les sciences physiques modernes : 239–307.

Suppe, Frédéric, éd. 1977. La structure des théories scientifiques. Université d'Illi
nois Presse.

Van Helden, Albert. 1974. "Saturne et ses Anses." Revue d'histoire d'As
tronomy 5 (2): 105–121.

Waters, C Kenneth. 2007. « La nature et le contexte de l'expérimentation exploratoire :
Une introduction à trois études de cas de recherche exploratoire. Histoire et philosophie des
sciences de la vie 29 (3) : 275–284.

Weirich, Paul. 2011. « Le pouvoir explicatif des modèles et des simulations : une exploration
philosophique ». Simulation et jeu 42 (2): 155–176. ISSN : 1046-
8781, 1552-826X.

Wilhelmson, Robert, Mathhew Gilmore, Louis Wicker, Glen Romine, Lee Cronce et Mark Straka. 2005. "Visualisation d'une tornade F3 dans un orage supercellulaire simulé." Procédure SIGGRAPH '05 ACM SIGGRAPH 2005 Catalogue d'art électronique et d'animation: 248–249. http : / / avl . ncsa. illinois.edu/what-we-do/services/media-downloads.

Wilhelmson, Robert, Lou Wicker, Matthew Gilmore, Glen Romine, Lee Cronce, Mark Straka, Donna Cox, et al. 2010. Visualisation d'un F3 Tornado : perspective d'un chasseur de tempête. Rapport technique. Laboratoire de visualisation avancée du NCSA oratoire. http : / / avl . ncsa. Illinois. edu / what - we - do / services/media-downloads%20https:// www.youtube.com/ watch?v=EgumU0Ns1Yl.

Wilson, Curtis. 1993. "Calcul de Clairaut sur le retour au XVIIIe siècle de la comète de Halley." Journal pour l'histoire de l'astronomie 24 (1-2): 1–15.

Wittgenstein, Ludwig. 1976. Zettel. Edité par GE Anscombe et Georg Henrik Von Wright. Presse de l'Université de Californie.

Woodward, James. 2003. Faire bouger les choses. Presse universitaire d'Oxford.

Woolfson, Michael M., et Geoffrey J. Pert. 1999. SATELLIT.FOR.

Chapitre 6
Paradigmes technologiques

Dans les chapitres précédents, nous avons discuté de la façon dont les philosophes, les scientifiques et les ingénieurs construisent l'idée que les simulations informatiques offrent une « nouvelle épistémologie » pour pratique scientifique. Ils entendaient par là que les simulations informatiques introduisent de nouvelles - et peut-être sans précédent - des formes de connaissance et de compréhension de l'environnement monde, des formulaires qui n'étaient pas disponibles auparavant. Alors que les scientifiques et les ingénieurs mettent l'accent sur la nouveauté scientifique des simulations informatiques, les philosophes tentent d'évaluer simulations informatiques pour leurs vertus philosophiques. La vérité de l'ancienne affirmation est hors de question, ce dernier, cependant, est plus controversé.

Les philosophes ont plaidé en faveur de la nouveauté des simulations informatiques à plusieurs reprises. Peter Galison, par exemple, défend l'idée que « les physiciens et les ingénieurs a rapidement élevé le Monte Carlo au-dessus du statut modeste d'un simple schéma de calcul numérique, [car] il en est venu à constituer une réalité alternative - dans certains cas un préféré – sur lequel une « expérimentation » pourrait être menée. Prouvé sur quoi à l'époque était le problème le plus complexe jamais entrepris dans l'histoire des sciences – la conception de la première bombe à hydrogène – le Monte Carlo a introduit la physique dans un lieu paradoxalement disloqué de la réalité traditionnelle qui emprunte à la fois domaines expérimentaux et théoriques, a lié ces emprunts entre eux et utilisé le bricolage qui en a résulté pour créer un Pays-Bas marginalisé qui n'était à la fois nulle part et partout sur la carte méthodologique habituelle. (Galison 1996, 119-120).

Naturellement, Galison n'est pas seul dans ses revendications. Beaucoup d'autres l'ont également rejoint plaidant en faveur de la valeur épistémologique, méthodologique et sémantique des simulations informatiques. Fritz Rohrlich, par exemple, est l'un des premiers philosophes à situer les simulations informatiques dans la carte méthodologique, se situant quelque part entre théorie et expérience. Il dit que « la simulation par ordinateur fournit [...] une analyse qualitative méthodologie nouvelle et différente pour les sciences naturelles, et cette méthodologie [...] se situe quelque part entre la science théorique traditionnelle et ses méthodes empiriques d'expérimentation et d'observation. Dans de nombreux cas, il s'agit d'un nouveau syntaxe qui remplace progressivement l'ancienne, et implique l'expérimentation de modèles théoriques d'une manière qualitativement nouvelle et intéressante. L'activité scientifique a ainsi atteint une nouvelle étape quelque peu comparable aux étapes qui ont lancé l'approche empirique (Galileo) et l'approche mathématique déterministe de la dynamique

(l'ancienne syntaxe de Newton et Laplace). La simulation informatique est par conséquent de intérêt philosophique considérable » (Rohrlich 1990, 507, original en italique). Auteurs comme Eric Winsberg ont également affirmé que «[c]omputer simulations have a distinct épistémologie [...] En d'autres termes, les techniques que les simulationnistes utilisent pour tenter pour justifier la simulation ne ressemblent à rien de ce qui passe habituellement pour de l'épistémologie dans le philosophie de la littérature scientifique. (Winsberg 2001, 447). Il est également connu pour avoir dit que « les simulations informatiques ne sont pas simplement des techniques de calcul numérique. Ils impliquent une chaîne complexe d'inférences qui servent à transformer les structures théoriques en connaissances concrètes spécifiques des systèmes physiques [...] ce processus de transformation [...] a sa propre épistémologie unique. C'est une épistémologie qui m'est inconnue à la plupart des philosophies des sciences » (Winsberg 1999, 275). De même, Paul Humphreys considère les simulations informatiques comme étant essentiellement différentes de la façon dont nous comprenons et évaluons les théories et modèles traditionnels (Humphreys 2004, 54). Le la liste des auteurs continue.

Notre traitement des fonctions épistémiques au chapitre 5 révèle une grande partie de la forme qui les études épistémologiques prennent pour des simulations informatiques. Tel que présenté et discuté, plusieurs fonctions épistémiques sont bien exécutées par des simulations informatiques, rendant la compréhension d'un système cible donné. Explication, prédiction, exploration, et les visualisations ne sont qu'une poignée de ces fonctions avec des spécificités épistémologiques. contribution à la pratique scientifique et technique. Cependant, de nombreuses autres fonctions épistémiques pourraient également être mentionnées. Penser à de nouvelles formes d'observation et de mesure le monde à travers des simulations informatiques, des modes d'évaluation des preuves et des moyens pour confirmer/infirmer des hypothèses scientifiques, entre autres. Les simulations dynamiques moléculaires en chimie, par exemple, ont désormais le potentiel de fournir des informations précieuses sur Résultats expérimentaux. Et les simulations d'états quantiques permettent aux scientifiques de sélectionner types d'atomes potentiels pour des cibles spécifiques, avant même de s'asseoir sur une paillasse de laboratoire (par exemple, la boîte à outils Atomistix). Toutes ces utilisations et fonctions des simulations informatiques sont, de nos jours, assez courant et dans une mesure raisonnable, critique pour le progrès de la science et de l'ingénierie modernes.

Nous avons également vu que, dans un nombre croissant d'occasions, les simulations précèdent expériences. Les simulations de dynamique moléculaire fournissent à nouveau des exemples utiles. Ces les simulations sont capables de fournir des images chimiques simples des intermédiaires ionisés et les mécanismes de réaction essentiels pour différents scénarios d'origine de la vie. La très les mêmes simulations aident à identifier les propriétés à l'échelle atomique qui déterminent cinétique macroscopique. Avec la présence de ces simulations, les expériences deviennent plus maniable et précis, car la simulation aide à réduire le nombre de matériaux utilisés, conditions réactives et configurations de densité. De plus, ces les simulations facilitent la manipulation des contraintes d'échelle de temps et de longueur qui limiterait autrement l'expérimentation moléculaire. À cet égard, les simulations informatiques complètent la pratique scientifique et technique en la rendant possible.

Dans ce contexte, certains scientifiques versés dans la philosophie ont soutenu que les simulations informatiques constituaient un nouveau paradigme de la recherche scientifique et technique. Le premier paradigme étant la théorie – et la modélisation – alors que le second paradigme étant l'expérimentation en laboratoire et sur le terrain. Dans ce chapitre, nous allons discuter de ce qu'il reviendrait à appeler les simulations informatiques un troisième paradigme de la recherche. De la même manière,

Le Big Data a été qualifié de « quatrième paradigme » de la recherche scientifique, ce qui rend impossible d'ignorer la symétrie avec les simulations informatiques. A cause de cela, dans ce chapitre, je vais tirer quelques similitudes et différences entre l'un et l'autre, dans l'espoir d'en comprendre la portée et les limites.

6.1 Les nouveaux paradigmes

La dénomination des simulations informatiques comme troisième paradigme de la recherche a son origine chez les partisans des études Big Data. En fait, les partisans des simulations informatiques n'en ont jamais parlé ainsi, malgré de fortes avancées en faveur de leur nouveauté scientifique et philosophique.

Concevoir les simulations informatiques comme un paradigme de la recherche – quelle que soit leur position ordinale – ajoute une certaine pression aux attentes que les chercheurs et les grand public ont sur eux. La physique est conçue comme un paradigme pour toutes les phénomènes qui se produisent parce qu'il donne un aperçu de ces phénomènes. La théorie électromagnétique, par exemple, est capable d'expliquer tous les phénomènes électromagnétiques et magnétiques. phénomènes, et la relativité générale généralise la relativité restreinte et la loi de Newton la gravitation universelle fournissant une description unifiée de la gravité en tant que propriété géométrique de l'espace et du temps. Pour appeler les simulations informatiques et le Big Data un paradigme, un pourrait penser, doit avoir une entrée épistémique similaire à celle de la physique en termes de perspicacité qu'ils fournissent ainsi que leur rôle d'autorité épistémique. Il est donc important discuter brièvement de l'étendue d'un tel titre (méritoire).

Jim Gray a suggéré d'appeler les simulations informatiques et le Big Data de nouveaux paradigmes de research1 dans une conférence donnée au National Research Council et au Computer Science and Telecommunications Board à Mountain View, CA en janvier 2007. Au cours de sa conférence, Gray a connecté les simulations informatiques et le Big Data avec une préoccupation principale existant dans la recherche contemporaine, c'est-à-dire que la pratique scientifique et technique sont touchés par un "déluge de données". La métaphore visait à attirer l'attention aux véritables préoccupations de nombreux chercheurs concernant la grande quantité de données stockés, rendus, rassemblés et manipulés par des praticiens de la science et de l'ingénierie. Pour donner un exemple simple, l'Australian Square Kilometre Array Pathfinder (ASKAP) se compose de 36 antennes de 12 mètres de diamètre chacune, réparties sur 4 000 mètres carrés, et travaillant ensemble comme un seul instrument rendant autant

---

1 Pour la plupart, je vais ignorer la position ordinale du paradigme. A cet égard, Je laisserai sans réponse la question de savoir s'il existe une hiérarchie présupposée entre les paradigmes. Comme nous l'avons vu au chapitre 3, le positiviste considérerait que l'expérimentation reste une position secondaire et confirmatoire de la théorie, où la théorie est plus importante. Après que les failles du positivisme ont été exposées, une nouvelle vague expérimentale a envahi la littérature montrant le vaste univers de l'expérimentation et leur importance philosophique. Les défenseurs des simulations informatiques devraient-ils et Big-Data affirment qu'il existe une nouvelle façon améliorée de faire de la science et de l'ingénierie, et que la théorie et l'expérimentation ne sont réservées qu'un rôle mineur, elles marcheraient dans le même sens route dangereuse comme le positiviste.

comme 700 To/seconde de données2. Ainsi compris, le déluge de données découle d'un excès dans la production et la collecte de données qu'aucune équipe de chercheurs ne peut traiter, sélectionner et comprendre sans l'aide supplémentaire d'un système informatique. C'est la raison principale pour laquelle le Big Data nécessite des algorithmes particuliers qui aident à trier ce qui est une donnée importante et ce qui ne l'est pas (par exemple, le bruit, les redondances, les données incomplètes, etc.).

Dans ce contexte, il est important d'élucider la notion de « paradigme » telle qu'utilisée par Gray et ses partisans car, jusqu'à présent, il n'existe pas de définitions disponibles. En particulier, c'est un terme qui ne peut être pris à la légère, surtout s'il a des implications dans le statut – culturel, épistémologique, social – d'une discipline et la manière dont le public l'acceptera. En philosophie, le « paradigme » est un terme théorique qui s'accompagne d'hypothèses spécifiques. Clarifier la signification d'un « paradigme » dans le contexte des simulations informatiques et du Big Data est donc notre première tâche.

Avant de commencer, une mise en garde doit être admise. Plus tôt, nous avons discuté de scénarios hybrides où les simulations informatiques s'intègrent à l'expérimentation en laboratoire. Dans les études Big Data, nous sommes confrontés à une situation similaire. Le Grand collisionneur de hadrons (LHC) est un bon exemple d'une telle intégration, car il combine de manière exquise la science et la technologie de pointe avec de grandes quantités de données, y compris, bien sûr, l'expérimentation, la théorie, le travail interdisciplinaire et les simulations informatiques. De nombreuses simulations au LHC visent à optimiser les ressources informatiques nécessaires pour modéliser la complexité des détecteurs et des capteurs, ainsi que la physique (Rimoldi 2011). D'autres, comme les simulations de pointe de Monte Carlo, calculent le signal du boson de Higgs du modèle standard et tout processus de fond pertinent. L'utilisation de ces simulations est d'optimiser les sélections d'événements, d'évaluer leur acceptation et d'évaluer les incertitudes systémiques (Chatrchyan et al. 2014). Ces simulations sont destinées à produire de grandes quantités de données qui doivent finalement être soigneusement conservées, sélectionnées et classées. Ainsi compris, les simulations informatiques et le Big Data sont profondément imbriqués dans le processus de rendu des données, de classification et d'utilisation, entre autres activités. Dans ce livre, j'ai intentionnellement évité de discuter de scénarios hybrides tels que le suggère le LHC. Cette décision est basée sur une raison très simple. Pour moi, avant de pouvoir appréhender les simulations informatiques en tant que systèmes hybrides, il est important de les comprendre d'abord individuellement. Alors que les scénarios hybrides offrent une vision plus riche, ils occultent également des aspects importants de l'analyse épistémologique et méthodologique.
À cet égard, lorsque j'aborderai les troisième et quatrième paradigmes de la recherche, j'aborderai les simulations informatiques et le Big-Data dans un scénario non hybride.

Thomas Kuhn est le premier philosophe à avoir analysé en profondeur la notion de « paradigme » dans la recherche scientifique. En discutant du fonctionnement des paradigmes en science, Kuhn note que «l'une des choses qu'une communauté scientifique acquiert avec un paradigme est un critère pour choisir des problèmes qui, bien que le paradigme soit pris pour acquis, peuvent être supposés avoir une solution. Dans une large mesure, ce sont les seuls problèmes que la communauté admettra comme scientifiques ou encouragera ses membres à entreprendre » (Kuhn 1962 : 37). Appeler les simulations informatiques et les mégadonnées un paradigme de la recherche scientifique

2 Bien qu'il existe plusieurs projets en science et en ingénierie reposant sur le Big Data, sa présence est beaucoup plus forte dans des domaines tels que les études sur les réseaux sociaux, l'économie et le grand gouvernement.

recherche a des implications spécifiques dans la manière dont les chercheurs mènent leur pratique, problèmes qui valent la peine d'être résolus et les bonnes méthodes pour rechercher de telles solutions.

Qu'est-ce donc qu'un « paradigme » ? Selon Kuhn, toute science mature (par exemple, physique, chimie, astronomie, etc.) connaît une alternance de phases de la science normale (par exemple, la mécanique newtonienne) et de révolutions (par exemple, le relativisme einsteinien). Pendant les périodes de science normale, une constellation d'engagements est fixée, y compris des théories, des instruments, des valeurs et des hypothèses. Celles-ci sont conformes à un « paradigme », c'est-à-dire le consensus sur ce qui constitue des cas exemplaires de bonne recherche scientifique. La fonction d'un paradigme est donc de fournir aux scientifiques des énigmes à résoudre. et de fournir les outils pour leur solution (Bird 2013). A titre d'exemples, Kuhn cite les équilibre chimique trouvé dans le Traité de l'élémentaire de chimie d'Antoine Lavoisier, la mathématisation du champ électromagnétique par James Clerk Maxwell, et la invention du calcul dans Principia Mathematica par Isaac Newton (Kuhn 1962, 23). Chacun de ces livres contient non seulement les théories clés, les lois et les principes de La nature, mais aussi – et c'est ce qui en fait des paradigmes – des guides d'application ces théories pour la résolution de problèmes importants (Bird 2013). De plus, ils fournissent également de nouvelles techniques expérimentales et mathématiques pour la résolution de tels problèmes. Des exemples de ce genre ont déjà été mentionnés : le bilan chimique pour le Traite` el´ ementaire de chimie et le calcul pour les Principia Mathematica.

Une crise scientifique survient lorsque la confiance des scientifiques dans un paradigme est perdue en raison de son incapacité ou de son incapacité à résoudre une énigme particulièrement inquiétante. Ceux-ci sont les « anomalies » qui émergent à l'époque de la science normale. Ces crises sont généralement suivie d'une révolution scientifique, ayant un paradigme existant remplacé par un rival. Lors d'une révolution scientifique, la matrice disciplinaire (c'est-à-dire la constellation des engagements partagés) subit des révisions, bousculant parfois jusqu'à la moelle corpus de croyances et de vision du monde. De telles révolutions émergent généralement de la nécessité de trouver de nouvelles solutions aux anomalies et de perturber de nouveaux phénomènes qui coexistaient dans les théories des périodes de science normale. L'exemple classique est le précession du périhélie de Mercure qui a fonctionné comme une position de confirmation pour relativité générale sur la mécanique newtonienne[3]. Les révolutions, cependant, n'affectent pas nécessairement le progrès scientifique, principalement parce que le nouveau paradigme doit moins certains aspects essentiels de son prédécesseur, en particulier le pouvoir de résoudre des problèmes quantitatifs problèmes (Kuhn 1962, 160ff.). Il est possible, cependant, que le nouveau paradigme perdre un certain pouvoir qualitatif et explicatif (Kuhn 1970 : 20). Dans tous les cas, nous pouvons dire qu'en période de révolution, il y a une augmentation globale de la résolution d'énigmes puissance, le nombre et la signification des énigmes, et les anomalies résolues par le paradigme révisé (Bird 2013).

Un paradigme renseigne donc les scientifiques sur la portée et les limites de leur recherche scientifique. domaine, au moment qui garantit que tous les problèmes légitimes peuvent être résolus dans ses propres termes. Ainsi compris, il semble que ni les simulations informatiques ni les Big Les données correspondent à cette description. Pour commencer, ce sont des méthodes qui implémentent des théories et des modèles, mais pas des théories en elles-mêmes, et donc ils ne sont pas adaptés à pro

[3] La précession du périhélie de Mercure a été expliquée par la relativité générale vers 1925 – avec des mesures successives et plus précises à partir de 1959 - bien qu'il s'agisse d'un "anormal" phénomène déjà connu en 1919.

mote une crise scientifique. Pourraient-ils conférer une théorie qui remet en question notre compréhension de base de, disons, la biologie ? Probablement, mais dans ce cas, ils n'auraient pas un statut différent de celui des expériences utilisées pour démystifier les théories sur la génération spontanée de vie complexe à partir de matière inanimée.4 On pourrait bien sûr spéculer que les simulations informatiques et le Big Data pourraient devenir, dans et par eux-mêmes, les ories d'une certaine sorte - ou des moyens pour une théorie. Il est vrai que certains philosophes ont déclaré la « fin de la théorie » provoquée par le Big Data, mais il y a peu de preuves que la pratique de la recherche se dirige réellement dans cette direction. De plus, les changements de paradigme viennent avec le nouveau paradigme retenant la même force explicative et prédictive que l'ancien, avec en plus la prise en charge des anomalies qui conduisent à la crise en premier lieu. Le Big Data, comme nous le verrons dans ce chapitre, non seulement se préoccupe peu d'« accumuler » des paradigmes précédents, mais il rejette ouvertement bon nombre de ses triomphes. Le plus évident est le rejet du besoin d'explications des phénomènes de toute sorte. Comme l'admettent de nombreux partisans du Big Data, il n'est pas possible d'expliquer pourquoi les phénomènes du monde réel se produisent, mais seulement de montrer qu'ils se produisent.

Qu'avait donc Gray à l'esprit lorsqu'il appelait les simulations informatiques et le Big Data le troisième et le quatrième paradigme, respectivement ? Commençons par remarquer sa division des paradigmes scientifiques en quatre moments historiques, à savoir,

1. Il y a des milliers d'années, la science était empirique décrivant les phénomènes naturels;
2. Branche théorique des dernières centaines d'années utilisant des modèles, des généralisations;
3. Depuis quelques années une branche informatique simulant des phénomènes complexes ;
4. Aujourd'hui : l'exploration de données eScience unifie théorie, expérimentation et simulation

> – des données captées par des instruments ou générées par simulateur,
>
> – traitées par le matériel, –
>
> informations/connaissances stockées dans l'ordinateur, –
>
> Les scientifiques analysent les bases de données/fichiers à l'aide de la gestion des données et des statistiques. Gris (2009, xviii)

Selon cette interprétation, un « paradigme » n'est pas tant un terme technique au sens donné par Kuhn, qu'un ensemble de pratiques de recherche cohérentes - comprenant des méthodes, des hypothèses et une terminologie - qu'une communauté de scientifiques et d'ingénieurs partagent entre eux. eux-mêmes. De telles pratiques de recherche n'exigent pas une révolution scientifique, ni n'en favorisent une. En fait, puisque les simulations informatiques et le Big Data utilisent les normes scientifiques et techniques actuelles, les théories, etc., ils semblent déjà être insérés dans un paradigme. La caractéristique des simulations informatiques et du Big Data, cependant, est la mécanisation et l'automatisation des données par les ordinateurs, ce qui manque clairement aux deux premiers paradigmes. Cela signifie que les méthodes utilisées et proposées dans les troisième et quatrième paradigmes sont très différentes de l'expérimentation de phénomènes et de la théorie du monde. Je me référerai aux simulations informatiques et au Big Data en tant que « paradigmes technologiques » pour tenter de mettre quelque distance par rapport à l'interprétation philosophique du « paradigme » présentée par Kuhn. Voyons maintenant si nous pouvons donner un sens à ces nouveaux paradigmes technologiques.

---

4 Louis Pasteur a montré que la génération spontanée apparente de micro-organismes était en fait due à l'air non filtré permettant la croissance bactérienne.

6.2 Big Data : comment faire de la science avec de grandes quantités de données ?

Tel que nous l'entendons aujourd'hui, Big Data5 fait référence à un grand ensemble de données dont la taille va bien au-delà de la capacité à capturer, gérer et traiter des données par un agent cognitif dans un délai raisonnable. Malheureusement, il n'existe pas de conceptualisation appropriée du Big Data dont la notion est généralement capturée en énumérant certaines de ses caractéristiques attribuées, telles que grandes, diverses, complexes, longitudinales, etc. Le problème principal d'avoir une liste de caractéristiques, de caractéristiques et attributs est qu'ils n'éclairent pas nécessairement le concept. Dire qu'il a quatre pattes, qu'il est poilu et qu'il aboie n'éclaire pas le concept de « chien ». En particulier, ces listes ont tendance à obscurcir plutôt qu'à éclairer la signification du Big Data, car nous n'avons pas avancé d'un pas en définissant « grand » en termes de « grand », « diversifié », etc. Une meilleure façon de caractériser le Big Data est nécessaire.

En 2001, Douglas Laney a proposé une définition du Big Data basée sur ce qu'il a appelé les « trois V » : volume, vélocité et variété (Laney 2001). Malheureusement, cette définition n'a pas très bien fonctionné. Aucun des « V » ne donne un aperçu réel de la notion, de la pratique et des utilisations du Big Data. Plus tard en 2012, la définition a été affinée par Mark Bayer et Laney lui-même comme "des actifs d'information à volume élevé, à grande vitesse et/ou à grande variété qui nécessitent de nouvelles formes de traitement pour permettre une prise de décision améliorée, la découverte d'informations et l'optimisation des processus". (Beyer et Laney 2012). Mieux adapté, mais toujours inadapté. Rob Kitchin a encore étendu la liste des caractéristiques et des fonctions qui constituent le Big Data, y compris une portée exhaustive, une résolution fine et uniquement indexicale dans l'identification, la relation, la flexibilité, etc. (Kitchin 2014). Malheureusement, aucune de ces définitions n'éclaire davantage la notion de Big Data qu'une simple liste. Des notions telles que « volume » et « variété » ne font pas avancer notre compréhension de « grand » et « grand ». Un sac de bonbons peut être varié et volumineux, et pourtant il ne dit rien sur le bonbon lui-même.

Mais il y a plus. Des termes comme « grand », « grand », « abondant », et d'autres peuvent être des constituants d'un prédicat relationnel, c'est-à-dire dans le cadre d'une relation de comparaison : cent mètres est un grand pâté de maisons ; et un livre de mille pages est un gros livre. Mais aucun de ces prédicats n'est absolu. Un pâté de maisons de cent mètres pourrait être grand pour un Allemand vivant dans une vieille ville de style médiéval. Mais il est de taille normale pour un citoyen argentin où la plupart des pâtés de maisons mesurent environ cent mètres. En d'autres termes, ce qui est important pour vous ne l'est peut-être pas pour moi (Floridi 2012). Bien sûr cela

---

5 Wolfgang Pietsch a suggéré une distinction entre le Big Data et la science à forte intensité de données. Alors que la première est définie en fonction de la quantité de données et des défis techniques qu'elle pose, la science à forte intensité de données fait référence « aux techniques avec lesquelles de grandes quantités de données sont traitées. Il convient en outre de distinguer les méthodes d'acquisition de données, de stockage de données et d'analyse de données. (Pietsch 2015). Il s'agit d'une distinction utile à des fins analytiques, car elle permet aux philosophes de tirer des conclusions sur les données, quelles que soient les méthodes impliquées ; en d'autres termes, distinguer le Big Data en tant que produit d'une discipline. Pour nous, cependant, cette distinction est inutile car nous sommes intéressés à étudier les techniques d'acquisition de données dans un contexte de composants techniques (par exemple, la vitesse, la mémoire, etc.). Une remarque similaire s'applique à la notion d'eScience de Jim Gray, comprise comme « là où l'informatique rencontre les scientifiques » (Hey, Tansley et Tolle 2009, xviii). Dans ce qui suit, bien que j'utilise indistinctement la notion de Big Data, de science intensive en données et d'eScience, les lecteurs doivent garder à l'esprit qu'il s'agit de domaines différents.

n'est pas un simple facteur culturel ou sociétal. Le changement technologique est particulièrement sensible aux prédicats relationnels. Ce qui aurait pu être considéré comme du « Big Data » au milieu 1950 est, à tous points de vue, insignifiant aujourd'hui. Comparez par exemple le montant de données pour l'élection présidentielle américaine de 1952, qui a utilisé l'UNIVAC avec un capacité de stockage d'environ 1,5 Mo par bande, avec les 700 To/s produits par l'australien Square Kilometre Array Pathfinder (ASKAP). Sous l'égide des deux étant "grandes", l'élection américaine et l'ASKAP peuvent être méthodologiquement et épistémiquement traitées comme égales, alors qu'elles ne devraient évidemment pas l'être. La leçon à tirer ici est que la seule liste des propriétés et des attributs des données ne fournit pas d'informations dans le concept, les pratiques et les usages du Big Data.

Avant de continuer, revenons un peu en arrière et demandons-nous : pourquoi est-il si important de avez-vous une définition du « Big Data » ? Une réponse possible est qu'en l'absence d'un concept et la spécification théorique de cette notion pourrait avoir de sérieuses implications dans régulation des pratiques technologiques, aux niveaux de dépendance individuelle, institutionnelle et gouvernementale. Tenez compte de la description suivante pour le financement de projets par la National Science Foundation (NSF): "L'expression" Big Data "dans cette sollicitation fait référence à des ensembles de données volumineux, divers, complexes, longitudinaux et/ou distribués générés à partir d'instruments, de capteurs, de transactions Internet, d'e-mails, de vidéos, de flux de clics et/ou toutes les autres sources numériques disponibles aujourd'hui et à l'avenir » (NSF 2012). Si quoi que ce soit, cette description est peu claire et vague, et elle n'aide pas à préciser ce qui constitue Le Big Data, ses finalités, ses limites, sa finalité, autant d'informations fondamentales pour demandeurs et exécuteurs de la subvention. De plus, l'utilisation excessive de synonymes ne signifie pas apporter par lui-même des éclaircissements sur ce à quoi renvoie le « Big Data », ni sur sa portée générale. Par rapport au document qui a remplacé la sollicitation deux ans plus tard, on relit avec beaucoup d'ambiguïté : « L'expression 'Big Data' fait référence à des données qui remettre en question les méthodes existantes en raison de leur taille, de leur complexité ou de leur taux de disponibilité. » (NSF 2014). Il est intéressant de noter que dans les deux sollicitations suivantes – 2015 et 2016 – la NSF a éliminé toutes les références aux définitions du Big Data. Au mieux, un peut trouver des énoncés indiquant « [w]hile les notions de volume, de vitesse et de variété sont couramment attribués aux problèmes de Big Data, d'autres problèmes clés incluent la qualité des données et provenance » (NSF 2016) en référence claire à la définition de Laney de 2001. Ces sollicitations de la NSF montrent combien il est difficile de définir, voire de caractériser clairement, la notion de Big Data.

A ma connaissance, il n'existe pas de définition non problématique et globale du Big Data. Ce fait, cependant, ne semble pas préoccuper la majorité de la littérature spécialisée. Certains auteurs affirment que la nouveauté du Big Data réside dans la simple quantité de données impliquées, en considérant cela comme intrinsèquement intuitif et suffisamment caractérisation valide. Dans le même sens, il a été affirmé que le Big Data se réfère à des ensembles de données dont la taille dépasse la capacité d'un logiciel de base de données commercial typique à capturer, stocker, gérer et analyser les données. Au lieu de cela, des périphériques de stockage et des logiciels spécifiques doivent être utilisés. À la base, la quantité de données traitées varie de quelques dizaines de pétaoctets à des milliers de zettaoctets, voire plus. Aussi naturellement qu'ils viennent, ces chiffres changent à mesure que de nouveaux ordinateurs entrent dans la technologie. scène en permettant plus de stockage et de vitesse de traitement informatique, à mesure que de nouveaux algorithmes sont développés pour trier et catégoriser les données, et que de nouvelles institutions

et les entreprises privées investissent de l'argent dans des installations et du personnel pour développer encore plus le Big Data. Pour traduire ces chiffres en valeurs et applications réelles, on estime qu'environ 2,5 exaoctets de données sont générés chaque jour, soit 2,5 x 1018 octets. Mark Liberman, professeur de linguistique à l'Université de Pennsylvanie, a estimé que le stockage requis pour conserver toute la parole humaine numérisée à 16 kHz, audio 16 bits, est d'environ 42 zettaoctets - ou 42x1021 octets - (Liberman 2003) .
De plus, étant donné que la production mondiale de données informatiques augmente à une vitesse sans précédent, les prévisions estiment qu'au moins 44 zettaoctets - c'est-à-dire 44x1021 octets seront produits d'ici 2020 (IDC 2014).6 Accepter

le fait qu'il n'existe pas de définition claire de ce qui constitue le Big Data ne doit pas suggérer que nous ne pouvons pas en identifier et mettre en évidence les caractéristiques spécifiques. Évidemment, la simple quantité de données est bien une caractéristique distinctive qui nous permet de discuter plus en détail de la méthodologie et de l'épistémologie du Big Data.
Par ailleurs, Sabina Leonelli, dans un article remarquable, a souligné à juste titre que la nouveauté du Big Data réside dans l'importance et le statut acquis par les données dans les sciences, ainsi que dans les méthodes, les infrastructures, les technologies et les compétences développées pour traiter ces données. (Leonelli 2014). Leonelli a raison de souligner que le Big Data élève les données au statut de « marchandise scientifique », dans le sens où elles assimilent d'autres unités d'analyse telles que les modèles et les théories dans leur pertinence scientifique. Elle a également raison de dire que le Big Data introduit des méthodologies, des infrastructures, des technologies et des compétences que nous n'avions pas auparavant. Le Big Data concerne donc la quantité de données, les nouvelles méthodes et infrastructures, ainsi que le développement de nouvelles technologies. Mais surtout, le Big Data est une façon entièrement nouvelle et différente de pratiquer la science et de comprendre le monde environnant, un peu comme les simulations informatiques. Je crois que c'est le sens que Gray avait à l'esprit lorsqu'il a qualifié les simulations informatiques et le Big Data de nouveaux paradigmes de la recherche scientifique.

Jusqu'à présent, nos efforts se sont concentrés sur la conceptualisation du Big Data. Voyons maintenant sa pratique concrète dans les domaines scientifiques et techniques. Une courte reconstruction du Big Data inclurait certaines pratiques au stade de la collecte, certaines au stade de la conservation et d'autres à l'utilisation des données. Commençons par la première étape, lorsque les données sont collectées par différents moyens de calcul.7 L'exemple

---

6 Certaines études établissent un lien entre la croissance de la mémoire système et la quantité de données stockées. Un rapport au Département de l'énergie des États-Unis montre qu'en moyenne, chaque 1 téraoctet de mémoire principale entraîne environ 35 téraoctets de nouvelles données stockées dans l'archive chaque année - plus de 35 téraoctets par 1 téraoctet sont effectivement stockés, mais 20 – 50 % sont effectivement supprimés en moyenne au cours de l'année (Hick, Watson et Cook 2010).

7 Notre traitement du Big Data sera axé sur les usages scientifiques. À cet égard, nous devons garder à l'esprit que, bien que la nature des données soit toujours computationnelle, elles sont également d'origine empirique. Permettez-moi ici de faire une digression rapide et de clarifier ce que j'entends par données computationnelles d'origine empirique. Dans l'expérimentation en laboratoire, les données recueillies pourraient provenir directement de la manipulation de l'expérience et donc au moyen de rapports de changements, de mesures, de réactions, etc., ainsi qu'au moyen de l'utilisation d'instruments de laboratoire. Un exemple du premier consiste à utiliser une boîte de Pétri pour observer le comportement des bactéries et la germination des plantes. Un exemple de ce dernier est la chambre à bulles qui détecte les particules chargées électriquement se déplaçant à travers un hydrogène liquide surchauffé. Le Big Data dans la recherche scientifique obtient une grande partie de ses données de sources similaires.

de l'ASKAP ci-dessus en est un bon exemple. Là, les données sont obtenues par un seul réseau de radiotélescopes et formatées de manière à les rendre compatibles avec différents ensembles de données et normes. Cependant, l'obtention de données à partir d'une seule constellation d'instruments n'est pas un cas typique pour le Big Data. Plus habituel est de trouver plusieurs sources différentes et, très probablement, incompatibles contribuant à agrandir les bases de données. On pourrait penser à combiner les données de l'ASKAP, qui est un radiotélescope pour explorer les origines des galaxies, avec les données obtenues de l'observatoire d'Atacama de l'université de Tokyo (TAO), un télescope optique-infrarouge pour comprendre le centre galactique (Yoshii et al. 2009).8 La collecte de données nécessite de s'assurer qu'elles sont formatées de manière à rendre les données compatibles avec les ensembles de données réels, ainsi que de s'assurer que les métadonnées sont normalisées de manière appropriée. Le formatage des données, ainsi que la standardisation des métadonnées, est une activité longue et coûteuse, bien que nécessaire pour assurer la disponibilité des données et qu'elles puissent éventuellement être analysées comme un ensemble unique d'informations.

C'est là qu'émerge le premier ensemble de problèmes pour le Big Data. La présence d'une constellation de sources augmente considérablement les problèmes de partage des données, de leur formatage et de leur normalisation, de leur conservation et de leur utilisation éventuelle. Étant donné qu'à ce stade initial, les principaux problèmes sont le partage, le formatage et la normalisation des données, concentrons-nous d'abord sur ceux-ci. Leonelli rapporte qu'il est fréquent que les chercheurs, qui souhaitent partager leurs données dans une base de données, doivent s'assurer que les données et les métadonnées qu'ils soumettent sont compatibles avec les normes existantes. Cela signifie trouver du temps dans leur agenda déjà chargé pour acquérir des connaissances actualisées sur ce que sont les normes et comment elles peuvent être mises en œuvre (Leonelli 2014, 4). Si les choses n'étaient pas assez difficiles, ni les universités ni les institutions publiques et privées n'apprécient toujours que leurs chercheurs consacrent du temps qualitatif à de telles entreprises. En conséquence, des efforts non reconnus incitent très peu les chercheurs à partager, formater et normaliser leurs données, avec la perte subséquente de collaboration professionnelle et d'efforts de recherche.

Aussi laborieuse et coûteuse que puisse être la collecte de données, si le Big Data doit devenir une réalité, le développement de bases de données, d'institutions, de législations, d'un financement de la recherche à long terme et d'un personnel formé est nécessaire. Ce point nous amène à la prochaine étape du processus de Big Data, impliquant nombre de ces acteurs dans la conservation des données (Buneman et al. 2008). Typiquement, curating consiste à rendre spécifiques

---

Comme suggéré, l'ASKAP obtient de grandes quantités de données en balayant le ciel et, à cet égard, les données ont une origine empirique. Même les données collectées sur les réseaux sociaux largement utilisées par les sociologues et les psychologues pourraient être considérées comme ayant une origine empirique. Pour opposer ces modes de collecte de données, prenons le cas des simulations informatiques. Avec ces dernières méthodes, les données sont produites par la simulation plutôt que collectées. Il ne s'agit pas d'une distinction fantaisiste, puisque les caractéristiques, l'évaluation épistémologique et la qualité des données varient considérablement d'une méthode à l'autre. Les philosophes se sont intéressés à la nature des données et à ce qui les différencie (par exemple (Barberousse et Marion 2013), (Humphreys 2013b) et (Humphreys 2013a)).

8 Un cas intéressant découle du Big Data biomédical où les données sont recueillies à partir d'une variété de sources incroyablement variées et complexes. Comme Charles Safran et al. l'indiquent, ces sources comprennent « les analyseurs automatiques de laboratoire, les systèmes de pharmacie et les systèmes d'imagerie clinique [...] complétés par des données provenant de systèmes soutenant les fonctions administratives de la santé telles que la démographie des patients, la couverture d'assurance, les données financières, etc. informations narratives, capturées électroniquement sous forme de données structurées ou transcrites en « texte libre » [...] dossiers de santé électroniques » (Safran et al. 2006, 2).

décisions concernant les données, allant des données à collecter à leur soin final et à leur documentation, y compris des activités telles que la structuration des données, l'évaluation de la qualité des données et l'offre de lignes directrices aux chercheurs intéressés par de grands ensembles de données. La conservation des données est une tâche difficile et souvent ingrate, car les conservateurs doivent s'assurer que les données sont facilement accessibles aux chercheurs. Elle est aussi très sensible, car le curating suppose une sélection, et donc une augmentation de l'idiocratie des données. Enfin, ce qui fait du processus de curation une activité complexe et controversée, c'est qu'il détermine en partie la manière dont les données seront utilisées à l'avenir. Malgré son importance, et tout comme la collecte de données, la curation n'est actuellement pas une activité valorisée au sein du système académique et institutionnel, et donc peu attractive pour les chercheurs.

Comme suggéré par la discussion précédente, les données qui sont rendues disponibles par le biais des bases de données ne sont pas des données « innocentes ». Il a été collecté, organisé et mis à la disposition de chercheurs ayant des objectifs spécifiques en tête. Cela ne veut bien sûr pas dire qu'il s'agit de données non fiables, mais plutôt qu'elles sont nécessairement biaisées, chargées de valeurs - épistémiques, cognitives, sociales et morales - qui doivent être affichées au grand jour avant leur utilisation. Malheureusement, cette dernière demande est plutôt difficile à satisfaire. Le chercheur, comme tout autre individu, participe à une société, une culture, un ensemble de valeurs prédéfinies qu'il projette – parfois à son insu – dans sa pratique. Fait intéressant, il a été avancé que le Big Data contrecarrait le risque de biais dans la collecte, la conservation et l'interprétation des données. En effet, l'accès à de grands ensembles de données augmente la probabilité que les biais et les erreurs soient automatiquement éliminés par une tendance «naturelle» des données à se regrouper - c'est ce que les sociologues et les philosophes appellent la triangulation. Par conséquent, plus les données sont collectées, plus il devient facile de les recouper entre elles et d'éliminer les données qui ressemblent à des valeurs aberrantes (voir (Leonelli 2014 ; Denzin 2006 ; Alison 2002)).

Quand la question sur le Big Data se résume aux usages de ces données, les réponses se dispersent dans tous les sens selon les disciplines : médecine (Costa 2014), études de durabilité (Mahajan et al. 2012), biologie (Callebaut 2012) et (Leonelli 2014), la génomique (Choudhury et al. 2014), l'astronomie (Edwards et Gaber 2014), pour ne citer que quelques utilisations du Big Data dans des contextes scientifiques (voir aussi (Critchlow et Dam 2013)). Dans la section suivante, je discute d'une telle utilisation en virologie. Au-delà des multiples applications du Big Data, la question de son utilisation est avant tout une question d'éthique. Les questions sur l'utilisation du Big Data en médecine renvoient à des questions sur le consentement et l'anonymisation du patient, ainsi que sur la confidentialité et la protection des données (Mittelstadt et Floridi 2016a). De même, les questions sur la propriété, la propriété intellectuelle et la manière de distinguer le Big Data académique du Big Data commercial sont des questions fondamentales d'ordre éthique. Les réponses à ces questions ont le pouvoir de nuire aux individus ou de contribuer à de nouvelles découvertes. Dans ce livre, je n'aborderai aucune de ces questions. La question des usages du Big Data est, pour moi, une question sur leur épistémologie. En ce qui concerne les questions d'éthique, je concentre mes efforts dans le domaine beaucoup moins développé de l'éthique des simulations informatiques (voir

article 7). Pour d'autres lectures sur l'éthique dans le Big Data, cependant, je recommande (Mit telstadt et Floridi 2016b), (Bunnik et al. 2016) et (Collmann et Matei 2016).9

6.2.1 Un exemple de Big Data

Dans un livre récent acclamé, Viktor Mayer-Schonberger et Kenneth Cukier (2013) en donnent un bel exemple qui dresse un portrait en pied du Big Data et de son utilisation potentielle à des fins scientifiques. L'histoire commence en 2009 avec l'épidémie mondiale d'une nouvelle souche de grippe A, le H1N1/09. Dans des situations comme celle-ci, la rapidité et l'exactitude des informations sont essentielles. Les retards et les imprécisions coûtent non seulement des vies, mais dans ce cas particulier menacent le déclenchement d'une pandémie. Autorités de santé publique aux États-Unis et en Europe étaient trop lents, prenant des jours voire des semaines pour obtenir les tenants et les aboutissants sorties de la situation. Bien sûr, plusieurs facteurs étaient en jeu qui ont contribué à la aggravation des retards. D'une part, les gens ne sont généralement pas traités pendant plusieurs jours avant d'aller chez le médecin. Les autorités sanitaires sont alors impuissantes leurs dossiers et leurs données dépendent de la collecte d'informations publiques (c'est-à-dire auprès des hôpitaux, cliniques et médecins privés), dont aucune n'implique une intrusion dans la vie privée des la maison. Les Centers for Disease Control and Prevention (CDC) aux États-Unis, et le système européen de surveillance de la grippe (EISS) s'appuyait sur les les données cliniques extraites de ces canaux, et donc les informations qu'ils détenaient était la plupart du temps incomplète pour une évaluation précise de la situation. De l'autre D'autre part, relayer les informations de ces sources vers les autorités sanitaires centrales pourrait prendre énormément de temps. Le CDC a publié des données nationales et régionales sur une base hebdomadaire, généralement avec un décalage de déclaration de 1 à 2 semaines. Récupérer et relayer l'information était un processus douloureux, imprécis et très lent. En conséquence, des informations sur l'endroit où il y avait eu une épidémie, ainsi qu'une carte précise des zones potentielles étaient, la plupart du temps, manquantes ou bien trop peu fiables.

En février de la même année, et quelques semaines seulement avant que le virus ne fasse la une des journaux, un groupe d'ingénieurs de Google a publié un article dans Nature décrivant une nouvelle façon d'obtenir des informations fiables qui pourraient potentiellement prédire le prochaine épidémie du virus, ainsi que pour fournir une carte précise de la propagation. Il était connu sous le nom de projet Google Flu Trends qui prône une approche simple mais assez idée nouvelle (Ginsberg et al. 2009). Selon les ingénieurs de Google, l'actuel le niveau d'activité grippale hebdomadaire aux États-Unis pourrait être estimé en mesurant fréquence relative des requêtes Internet spécifiques qui, selon les auteurs, était fortement corrélé avec le pourcentage de visites chez le médecin au cours desquelles un patient s'est présenté avec des symptômes pseudo-grippaux. En d'autres termes, ils ont utilisé des millions de gigaoctets des requêtes Internet et les a comparées avec les données du CDC. Tout cela pourrait être fait dans le temps incroyable d'une journée. Les choses s'accéléraient, et si ça continuait

9 Notons que les questions éthiques soulevées dans chacun de ces livres ne se limitent pas nécessairement à utilisations scientifiques et techniques du Big Data, mais elles s'étendent également aux entreprises, à la société et aux études sur le gouvernement et la loi.

de cette façon, les autorités sanitaires seraient en mesure de faire face aux épidémies dans un délai raisonnable laps de temps.

Selon ces ingénieurs de Google, "les requêtes de recherche sur le Web de Google peuvent être utilisées pour estimer avec précision les pourcentages de SG [syndrome grippal] dans chacun des neuf régions de santé publique des États-Unis » (1012). Le message était fort et attrayant, et bien sûr tout le monde en a pris note. Des auteurs comme Mayer-Schonberger et Cukier se sont aventurés sur le fait que « d'ici la prochaine pandémie, le monde disposera d'un meilleur outil pour prévoir et ainsi empêcher sa propagation » (Mayer Schonberger et Cukier 2013, 10). Les attentes sont élevées, tout comme les intérêts pour le succès du Big Data dans les disciplines scientifiques.

Malheureusement, le Big Data est tout sauf exact, et l'enthousiasme et les grands espoirs suscités par ces promoteurs doivent être mesurés. Pour commencer, il est loin d'être clair que Big Les données pourraient être utilisées pour « la prochaine pandémie », à moins que certaines caractéristiques ne soient partagées, tels que le type de pandémie et les moyens de récupérer des informations. Le Big Data est géographiquement, économiquement et socialement dépendant. Le choléra est actuellement classé comme une pandémie, bien qu'elle soit très rare dans les pays développés et industrialisés, et la façon dont il éclate et se propage est assez différente d'une maladie grippale. Fur thermore, les zones présentant un risque permanent de maladie, comme certains endroits en Afrique et l'Asie du Sud-Est, ne disposent pas de l'infrastructure technologique appropriée pour les citoyens pour accéder à Internet, sans parler d'interroger à l'aide de Google. Récupération des informations de ces endroits ne satisferaient probablement pas aux exigences minimales de corrélation.

On pourrait bien sûr se concentrer uniquement sur les cas réussis de Big Data. L'hypothèse examinée dans de tels cas est que le Big Data est un substitut plutôt qu'un complément à la collecte et à l'analyse de données traditionnelles. Ceci est connu dans la communauté sous le nom de « big data hubris »,10 et vise à souligner qu'une histoire de (relative) le succès n'est en aucun cas l'histoire d'une méthode. Même les ingénieurs de Google sont prudents pour nous avertir que le système n'a pas été conçu pour remplacer les diagnostics de laboratoire et la surveillance médicale. De plus, ils admettent que les requêtes de recherche dans le modèle utilisé par Google Flu Trends ne sont pas soumis uniquement par des utilisateurs confrontés symptômes pseudo-grippaux, et donc les corrélations observées sont susceptibles de fausses alertes causées par une augmentation soudaine des requêtes liées aux SG. « Un événement insolite, comme un rappel de médicament pour un remède populaire contre le rhume ou la grippe, pourrait provoquer une telle fausse alerte » (Ginsberg et al. 2009, 1014) déclare Jeremy Ginsberg, le scientifique principal qui a publié les résultats.

Bien que Google Flu Trends ne soit qu'un exemple de l'utilisation du Big Data, et certes pas le plus réussi, il décrit toujours bien les avantages du Big Data. Les données ainsi que leurs limites intrinsèques : plus de données ne signifie pas de meilleures informations, cela signifie simplement plus de données. De plus, toutes les données ne sont pas des données de « bonne qualité ». C'est le raison pour laquelle Ginsberg et al. préciser que l'utilisation des données de requête du moteur de recherche n'est pas conçu pour remplacer ni supplanter le besoin de diagnostics en laboratoire et de surveillance médicale. Des données médicales fiables sont encore obtenues par la surveillance traditionnelle

---

[dix] Le terme 'hubris' se trouve généralement dans les grandes tragédies grecques décrivant la personnalité du héros qualité comme étant d'une fierté extrême et stupide, ou détenant un excès de confiance dangereux souvent dans combinaison avec l'arrogance. Le comportement du héros est de défier les normes établies en défiant les dieux, avec pour résultat la propre chute du héros.

mécanismes, comme simplement aller chez le médecin. Un phénomène attendu lié
au modèle utilisé par Google Flu Trends est que, dans une situation de panique et d'inquiétude,
les individus en bonne santé provoqueront une augmentation des requêtes liées au SG, ce qui augmentera les estimations
des personnes réellement infectées et des scénarios de propagation potentiels. C'est parce que les requêtes
ne sont pas limités aux utilisateurs qui présentent effectivement des symptômes pseudo-grippaux, mais
plutôt toute personne qui publie la bonne requête. Les corrélations observées, donc,
ne sont pas significatives dans toutes les populations. Autrement dit, le système est susceptible
à un certain nombre de fausses alertes, pas toutes traçables et éliminables, rendant encore peu fiable
l'ensemble du programme d'anticipation d'une pandémie via le Big Data. Le
l'espoir est qu'avec les corrections et les utilisations appropriées de Google Flu Trends, il fournira un bon
aperçu du schéma de propagation des maladies liées au SG, ainsi
devenu un outil utile pour les responsables de la santé publique afin de leur donner une réponse précoce
et mieux informée en cas de pandémie réelle. À cette fin, Google Flu Trends a
désormais recentré sur la mise à disposition des données aux institutions spécialisées dans
recherche sur les maladies infectieuses d'utiliser les données pour construire leurs propres modèles.

Notons enfin que Google Flu Trends se nourrit de différentes sources de données
que l'exemple ASKAP. Dans le premier cas, les sources sont les données des requêtes de recherche,
tandis que dans ce dernier, la source est l'information astronomique fournie par les radiotélescopes. En ce
sens, les deux cas présentent de nombreuses différences, telles que la susceptibilité
de fausses alertes dans le cas de Google Flu Trends absent dans ASKAP. De plus, un
pourrait faire valoir que les données utilisées dans Google Flu Trends ne sont pas d'origine empirique,
comme ASKAP. Il y a bien sûr aussi des similitudes partagées par ces deux
systèmes. Leonelli soutient que Mayer-Schonberger et Cukier considèrent trois innovations fondamentales
apportées par le Big Data - également présentes dans le cas de Google Flu Trends
comme ASKAP. Le premier est l'exhaustivité et fait référence à l'affirmation selon laquelle l'accumulation de
grands ensembles de données permet à différents scientifiques, à différents moments, de fonder
leur analyse sur différents aspects d'un même phénomène. Vient ensuite le désordre.
Le Big Data pousse les chercheurs à accueillir la nature complexe et multiforme de
le monde, plutôt que de rechercher l'exactitude et la précision comme le fait la pratique scientifique standard.
En effet, selon elle, le point de vue de Mayer-Schonberger et Cukier considère qu'il est impossible d'assembler
le Big Data de manière garantie précise et homogène. Empruntant à la même idée, Mayer-Schonberger et
Cukier déclarent : "Le Big Data est désordonné, varie en qualité et est distribué sur un nombre incalculable
de serveurs dans le monde" (Mayer-Schonberger et Cukier 2013, 13). Troisièmement et finalement vient
« le triomphe des corrélations », c'est-à-dire un nouveau type de science (c'est-à-dire

la science apportée par le Big Data) guidé par des corrélations statistiques entre les données
valeurs plutôt que la causalité (Leonelli 2014).

C'est un bon endroit pour s'arrêter et poser la question suivante : qu'est-ce que Big
Des données à voir avec des simulations informatiques ? D'une part, et en principe, ils partagent tous deux
deux piliers communs. Ce sont le pilier informatique, qui
comprend la prise en compte de la vitesse de calcul, de la capacité de stockage des données et de la manipulation
grands volumes de données ; et le pilier cognitif, qui comprend des méthodes d'élucidation
pour représenter, valider, agréger et croiser des ensembles de données. Sur le
D'autre part, le Big Data et les simulations informatiques ont peu de choses en commun. Cela est
particulièrement vrai lorsqu'il s'agit d'affirmations épistémiques et méthodologiques. Nous avons déjà

a dit quelque chose sur les piliers informatiques et cognitifs, à la fois pour les Big Data et les simulations informatiques. La question qui nous intéresse maintenant est de savoir quelles sont les différences qui distinguent le Big Data et les simulations informatiques. Autrement dit, qu'est-ce qui en fait deux paradigmes technologiques distincts ?

6.3 La lutte pour la causalité : Big Data et simulations informatiques

Un objectif souhaitable pour toute méthode scientifique est de pouvoir fournir une connaissance et une compréhension du monde qui nous entoure. Si ces connaissances et cette compréhension sont fournies rapidement, comme dans le cas de Google Flue Trends, alors c'est certainement la plus préférable. Dans tous les cas, une telle méthode doit fournir des moyens de comprendre comment le monde est, sinon elle n'aura que peu d'intérêt à des fins scientifiques et techniques. Prenons comme exemple nos discussions précédentes sur la comparaison des simulations informatiques avec l'expérimentation en laboratoire au chapitre 3. Pourquoi les philosophes se donnent-ils la peine de discuter de ces questions ? L'une des principales raisons est que nous devons nous donner les moyens de garantir, dans des limites raisonnables, que les résultats des simulations informatiques sont au moins aussi fiables que nos meilleures méthodes actuelles d'enquête sur le monde, c'est-à-dire via l'expérimentation. Un raisonnement similaire s'applique à la théorie et aux modèles, comme discuté au chapitre 2.

Une caractéristique essentielle de l'expérimentation, et que reflète également une grande partie de la pratique de la modélisation, est qu'elle découvre des relations causales. Par exemple, l'antenne conçue par Arno Penzias et Robert Wilson à l'origine pour détecter les ondes radio a fini par détecter le rayonnement de fond cosmique des micro-ondes. Comment cela pourrait-il arriver? L'histoire est assez fascinante, mais le point important pour nous est que ces micro-ondes cosmiques interagissaient de manière causale avec l'antenne. Cela, bien sûr, ne s'est pas produit sans que les scientifiques aient tenté d'éliminer d'autres sources potentielles d'interaction causale, telles que les interférences terrestres - le bruit émanant de New York - ainsi que les excréments de chauves-souris et de colombes.

La raison pour laquelle les affirmations sur les relations causales sont à la base de la confiance des méthodes scientifiques et techniques est qu'elles fonctionnent comme une sorte de garantie pour notre connaissance et notre compréhension du monde. De nombreux philosophes ont dépeint les relations causales – ou la causalité – comme les pierres angulaires de toute entreprise scientifique. Le philosophe John L. Mackie a même qualifié la causalité de « ciment de l'univers » (Mackie 1980).

Comme le montre l'exemple de Penzias et Wilson, l'expérimentation est cruciale pour la science et l'ingénierie précisément parce qu'elle aide à découvrir les relations causales existant dans le monde. Rappelons du chapitre 3 la centralité de la causalité dans les discussions sur l'appartenance expérimentale des simulations informatiques. Nous avons vu que si nous pouvons montrer l'existence de relations causales dans des simulations informatiques, alors elles peuvent être traitées sur un pied d'égalité en tant qu'expériences. Si cela n'est pas possible, leur évaluation épistémologique reste ouverte. Dans la section 3, nous avons eu une discussion relativement longue sur les raisons qui poussent certains philosophes à croire que les simulations informatiques manquent de relations causales, et ce que cela signifie pour leur assimilation épistémologique.

session. Que la causalité provienne d'une interaction directe avec le monde - telle
comme dans l'expérimentation standard - ou la représentation des causes - comme dans les théories, les
modèles et les simulations informatiques - être capable d'identifier les relations causales est essentiel pour
progrès scientifique et technique. C'est vrai, en tout cas, si l'on exclut Big
Données. Dans le Big Data, a-t-on soutenu, la corrélation remplace la causalité dans la course
pour connaître le monde.

Mayer-Schonberger et Cukier sont deux auteurs influents qui préconisent fortement que la causalité soit
mise dans un tiroir et que les chercheurs adoptent la corrélation
plutôt comme le nouvel objectif de la recherche scientifique. Pour eux, il est tout à fait vain de diriger
efforts pour découvrir un mécanisme universel pour inférer des relations causales à partir de Big
Données. En fin de compte, tout ce dont nous avons besoin pour être satisfait épistémiquement dans le Big Data
sont des corrélations élevées (Mayer-Schonberger et Cukier 2013).

Une question à laquelle nous voulons répondre ici est de savoir si le Big Data, tel que Mayer
Schonberger et Cukier le décrivent, pourrait bien fonctionner avec une simple corrélation, ou si la causalité
doit être replacée dans le tableau général de la recherche scientifique et technique.
Ma conviction est que nous ne devrions pas abandonner si facilement la causalité, principalement parce que faire
vient donc avec un prix (élevé). Commençons par apporter quelques précisions.

La différence fondamentale entre la causalité et la corrélation est que la première est prise
comme une régularité dans un monde naturel et social stable et permanent à travers
temps et espace. La corrélation, quant à elle, est un concept statistique qui mesure
la relation entre deux valeurs de données données. Par exemple, il n'existe pas
relations de causalité entre jouer au football et casser des vitres, malgré la forte
corrélation que nous avions dans le quartier de mon enfance. Il y a cependant une causalité
lien entre un ballon de foot frappant la vitre de mon voisin et ce dernier pénétrant
pièces. La causalité, telle que nous l'entendons ici, nous indique les véritables raisons pour lesquelles
et comment quelque chose se passe. La corrélation, au contraire, ne nous dit que quelque chose sur un
relations statistiques entre deux points de données ou plus.

De l'exemple précédent, nous ne devons pas déduire que la corrélation n'est pas un
source de connaissance. Des études sur l'assurance automobile, par exemple, ont montré que les hommes
les conducteurs ont une plus forte corrélation avec les accidents de voiture que les conductrices, et par
conséquent les compagnies d'assurance ont tendance à facturer plus les hommes que les femmes pour leur assurance.
Il n'y a, bien sûr, aucun lien de causalité entre être un conducteur masculin et provoquer
un accident, ni être une conductrice et être une conductrice prudente. Une corrélation élevée est,
pour des cas comme celui-ci, assez bien...

Maintenant, est-ce ce que Mayer-Schonberger et Cukier ont à l'esprit lorsqu'ils plaident pour la
corrélation plutôt que la causalité dans le Big Data ? Pas assez. Selon ces auteurs,
la recherche de relations causales - une tâche très difficile et exigeante en soi - perd
tout son attrait avec l'accroissement exceptionnel des données disponibles sur un phénomène donné.
Rappelez-vous de la section précédente lorsque nous avons discuté de "l'exhaustivité"
et le « désordre » des données. Alors que le premier fait référence à la quantité de données, le second
met l'accent sur le rôle des données inexactes. Ensemble, affirment Mayer-Schonberger et Cukier, ils
fournissent le contexte épistémique nécessaire pour préférer une corrélation élevée à
causalité. Comment est-ce possible? Le raisonnement est le suivant : la corrélation consiste
dans la mesure de la relation statistique entre deux valeurs de données, et les données sont précisément

ce que les chercheurs ont en trop (c'est-à-dire la thèse de l'exhaustivité.11) Cela signifie
que les chercheurs sont capables de trouver des corrélations très élevées et stables dans un ensemble de données
un moyen relativement facile et rapide. De plus, les chercheurs savent que la corrélation pourrait
restent élevés même s'il y a des données inexactes, puisque la quantité est suffisante pour compenser toute
inexactitude trouvée dans les données (c'est-à-dire le désordre
thèse). La causalité devient alors une antiquité sous le nouveau soleil du Big Data, car
les chercheurs n'ont plus besoin de trouver les relations causales agissantes pour avoir des revendications
sur un phénomène donné. L'exemple qui éclaire ce raisonnement est Google
Tendances de la grippe à nouveau. Comme présenté, les ingénieurs de Google ont pu trouver des corrélations élevées
parmi les données malgré le fait que certaines requêtes provenaient de personnes
grippe régulière, et certains ne sont pas malades du tout. La grande quantité de données recueillies compense
les éventuelles inexactitudes qui se cachent dans ces quelques requêtes. Naturellement, le désordre
thèse qui permet de faire abstraction des quelques interrogations trompeuses suppose que
pourcentage de ces requêtes est négligeable. Sinon, la corrélation peut encore être élevée, mais
sur les mauvais résultats.

Comme je l'ai mentionné plus tôt, Mayer-Schonberger et Cukier ont raison sur les avantages de soutenir,
dans les bonnes circonstances, la corrélation plutôt que la causalité. Et
ils fournissent abondamment des exemples de cas réussis. Mais j'ai aussi averti que leur
les idées doivent être dans le contexte approprié pour être applicables. La plupart des auteurs
les exemples proviennent de l'utilisation du Big Data dans les grandes entreprises et les grands gouvernements.
Bien que l'avantage que représente le Big Data pour ces sphères de la
société est indéniable, ils ne représentent pas les usages et les besoins du Big Data dans les contextes
scientifiques et d'ingénierie. En effet, la corrélation semble être le bon concept conceptuel
outil pour trouver les liens pertinents entre nos préférences d'achat et nos choix futurs
– comme cela est décrit dans le cas d'Amazon.com (101) – et nos notes de crédit financier
avec notre comportement personnel - comme le montre The Fair Isaac Corporation avec l'utilisation de
Mégadonnées (56). Mais quand la question porte sur les usages du Big Data à des fins scientifiques et
à des fins d'ingénierie, la causalité semble retrouver sa valeur d'origine. Laisse-moi expliquer
pourquoi.

L'une des pertes majeures dans la corrélation commerciale sur la causalité est que l'ancien
ne peut pas dire précisément aux chercheurs pourquoi quelque chose se produit, mais plutôt quoi - ou que
- ça arrive. Mayer-Schonberger et Cukier l'admettent dès le début, mettant ces idées sous la forme suivante : «
si nous pouvons économiser de l'argent en connaissant le meilleur moment pour
acheter un billet d'avion sans comprendre la méthode derrière la folie des billets d'avion, c'est
assez bien. Les mégadonnées concernent ce qui n'est pas pourquoi » (14 - Souligné dans l'original).12

Selon Mayer-Schonberger et Cukier, le Big Data ne consiste donc pas à expliquer pourquoi quelque chose
est le cas, mais plutôt à offrir des connaissances prédictives et descriptives. Cela pourrait être considéré
comme une limitation sérieuse pour un
méthode promue comme le quatrième paradigme de la recherche scientifique. Physique sans

---

11 On prétend même que la quantité de données correspond au phénomène lui-même. C'est,
les données sont le phénomène (Mayer-Schonberger et Cukier 2013).

12 Les auteurs sont bien sûr conscients de l'importance de la causalité dans les sciences et l'ingénierie. Dans
cet égard, ils disent : « Nous aurons encore besoin d'études causales et d'expériences contrôlées avec soin.
des données conservées dans certains cas, comme la conception d'une pièce d'avion critique. Mais pour beaucoup tous les jours
besoins, savoir ce qui n'est pas pourquoi suffit » (191 - emphase dans l'original)

La capacité d'expliquer pourquoi les planètes tournent autour du soleil ne ferait pas progresser notre compréhension du système solaire. De même, ce qui sépare les évolutionnistes des créationnistes – et de leurs cousins, les défenseurs du dessein intelligent – est précisément leur capacité à expliquer une grande partie de notre monde biologique et géologique environnant. Un bon exemple d'évolution est la variation de couleur de la population de papillons poivrés à la suite de la révolution industrielle. À l'ère préindustrielle, la grande majorité de ces papillons avaient une coloration claire et marbrée qui leur servait de camouflage contre les prédateurs. On estime qu'avant la révolution industrielle, une variante uniformément sombre du papillon poivré représentait 2% de l'espèce. Après la révolution industrielle, la couleur de la population a connu un profond changement : jusqu'à 95 % présentaient une coloration foncée. La meilleure explication de la raison pour laquelle ce changement a eu lieu découle de la prise en compte de l'adaptation des papillons de nuit au nouvel environnement assombri par la pollution. L'exemple se présente comme un changement majeur causé par des mutations dans une espèce, fondant les idées de variation et de sélection naturelle. Le pouvoir épistémologique de la théorie de l'évolution réside dans le fait qu'elle peut expliquer pourquoi cela s'est produit, et pas seulement que cela s'est produit.

La perte de force explicative semble être un prix trop élevé à payer pour toute méthode qui se proclame scientifique. Nous avons discuté dans la section 5.1.1 comment l'explication pouvait compter comme une véritable fonction épistémologique des simulations informatiques, au point de les fonder comme des méthodes fiables pour connaître et comprendre le monde. Cela ne veut bien sûr pas dire qu'il faille abandonner le Big Data comme véritable méthode de recherche de la connaissance du monde. Il vise uniquement à attirer l'attention sur les limites du Big Data en tant que méthode de recherche scientifique et technique.13 Considérons

encore un autre cas, cette fois-ci venant de l'ingénierie. Le 28 janvier 1986, la navette spatiale Challenger s'est brisée 73 secondes après le début de son vol pour finalement se désintégrer au-dessus de l'océan Atlantique. En conséquence, la catastrophe a coûté la vie à sept membres d'équipage. Richard Feynman, lors de l'audition du Challenger, a montré que les caoutchoucs utilisés dans les joints toriques de la navette devenaient moins résistants dans les situations où l'anneau était exposé à de basses températures. Feynman a démontré qu'en comprimant un échantillon des joints toriques dans une pince et en l'immergeant dans de l'eau glacée, l'anneau ne retrouverait jamais sa forme d'origine. Avec cette preuve en place, la commission a pu déterminer que la catastrophe avait été causée par le joint torique primaire qui n'était pas correctement scellé par le temps exceptionnellement froid de janvier 1986.

Ce deuxième exemple montre que trouver les bons liens de causalité, et donc pouvoir fournir une véritable explication, n'est pas une affaire qui ne concerne que la science, mais aussi la technologie. Nous tenons en haute estime la science et la technologie précisément en raison de leur force de persuasion pour comprendre et changer le monde. Trouver les bonnes relations causales est l'un des principaux contributeurs à cette fin. Ainsi, lorsque Mayer-Schonberger et Cukier objectent que cerner les relations causales est une tâche difficile, cela ne doit pas être considéré comme une raison pour ne pas se lancer dans leur chasse. Et lorsque Jim Gray ordonne que le Big Data soit le quatrième paradigme de la recherche, il nous encourage à adopter sa nouveauté ainsi qu'à comprendre ses limites. Aucune méthode de calcul en science et

[13] On pourrait faire valoir que la causalité n'est pas nécessaire pour l'explication, et qu'étant donné le bon cadre explicatif, le Big Data pourrait être en mesure de fournir de véritables explications. Au meilleur de ma connaissance, un tel cadre explicatif fait toujours défaut.

la recherche en ingénierie est toute puissante qui pourrait dispenser les chercheurs de s'appuyer sur la théorisation et l'expérimentation traditionnelles.

En revanche, la notion de causalité – ou de relations causales – est en effet difficile à cerner. On pourrait être tenté de supposer que toutes les vraies régularités sont des relations causales. Malheureusement, ce n'est pas le cas. De nombreuses régularités qui nous sont familières n'ont aucune relation causale. La nuit suit le jour comme le jour suit la nuit, mais ni l'un ni l'autre ne cause l'autre. De plus, on pourrait supposer que toutes les lois de la nature sont en quelque sorte causales. Mais ici réside aussi une autre déception. La loi de Kepler sur le mouvement planétaire décrit l'orbite des planètes, mais elle n'offre aucune explication causale de ces mouvements. De même, la loi des gaz parfaits relie la pression au volume et à la température. Elle nous dit même comment ces quantités varient en fonction les unes des autres pour un échantillon de gaz donné, mais elle ne nous dit rien sur les relations causales existant entre elles.

Les philosophes des sciences sont intrigués par la causalité depuis des siècles. Aris totle est peut-être le philosophe le plus ancien et le plus célèbre à avoir discuté de la notion et de la nature de la causalité (Falcon 2015). David Hume, d'autre part, était célèbre pour avoir une vision sceptique de la causalité, une vision qui la considère comme une « coutume » ou une « habitude » qui produit une idée de connexion nécessaire (De Pierris et Friedman 2013). Il existe bien sûr une multitude d'interprétations philosophiques différentes de la causalité, allant de l'histoire ancienne à aujourd'hui.

La façon la plus intuitive d'interpréter la causalité est peut-être de la considérer comme ayant une sorte de caractéristique physique. Prenons l'exemple du ballon de football frappant à nouveau la fenêtre. Dans cet exemple, la balle interagit physiquement avec la fenêtre, la brisant. Pour les partisans de la causalité physique, deux objets sont causalement liés lorsqu'il y a un échange d'une quantité physique, comme une marque ou un moment de l'un sur l'autre.[14] La balle a eu un échange d'une quantité physique avec la fenêtre, modifiant ainsi il.

Malheureusement, toutes les relations causales en science et en ingénierie ne dépendent pas d'une manière ou d'une autre de l'échange d'une quantité physique. Prenons, par exemple, la phrase suivante : « L'absentéisme chez les enfants en âge scolaire est causé par des parents qui ont perdu leur emploi. Dans un tel cas, il est difficile, voire impossible, d'établir une relation causale physique entre l'absentéisme et la perte de son emploi. Dans de tels cas, les philosophes ont tendance à faire abstraction de la « physicalité » des relations causales et à mettre l'accent sur un niveau représentationnel (Woodward 2003). C'est cette dernière interprétation qui permet aux simulations informatiques d'embrasser la notion de causalité et au Big Data de la combattre.

Est-il alors possible de sauver la causalité dans le Big Data et ainsi de garder intacte notre vision générale de la pratique scientifique ? Pour répondre à cette question, il faut d'abord préciser comment la causalité se retrouve dans ces deux paradigmes technologiques. Les modèles de causalité existent dans la pratique scientifique et sont régulièrement mis en œuvre sous forme de systèmes informatiques. La caractéristique générale de ces modèles est qu'ils représentent fidèlement, dans des conditions précises, l'ensemble des relations causales jouant un rôle dans le système cible[15].

---

[14] Les principaux exemples sont (Salmon 1998), (Dowe 2000) et (Bunge 2017).

[15] Pour un compte rendu complet de cette interprétation de la causalité, voir (Pearl 2000).

Cependant, ce n'est pas la signification que nous donnons ici. Pour toute question sur la causalité pour avoir un sens à la fois pour les simulations informatiques et le Big Data, nous avons besoin en quelque sorte pour les mettre sur un pied d'égalité. Étant donné que le Big Data ne consiste pas à mettre en œuvre des modèles, mais plutôt à faire quelque chose avec de grandes quantités de données, alors le seul façon de parler de relations causales pertinentes pour les deux paradigmes est lorsque nous prenons le relations causales déduites des données. Pour un tel cas, un algorithme est chargé de reconstruire de telles relations causales à partir de grandes quantités de données. Dans ce qui suit, je me concentre uniquement sur les discussions sur les algorithmes pour le Big Data, mais ce qui est dit ici pourrait également s'appliquer aux simulations informatiques, à condition que la réduction de la quantité de données est prise en compte. Dans ce contexte, la seule différence entre les simulations informatiques et le Big Data est sur la quantité de données que chacun poignées.

Au début des années 1990, il y avait un vif intérêt à enraciner la découverte de relations causales dans les données. À la base, le problème consiste à avoir un algorithme qui identifie correctement les relations causales existant dans de grandes quantités de données. Pour illustrer comment les chercheurs connaissent généralement les relations causales, considèrent le cas du tabagisme comme un préjudice pour la santé humaine. Des études médicales montrent que sur plus de 7 000 produits chimiques dans la fumée de tabac, au moins 250 sont connus pour être nocifs. Et de ces 250, au moins 69 peuvent causer le cancer. Ces produits chimiques cancérigènes comprennent l'acétaldéhyde, les amines aromatiques, l'arsenic, etc. (US Department of Health and Human Prestations 2014). Comment ces informations sont-elles généralement obtenues ? Le National Cancer L'Institut rapporte que leurs résultats sont basés sur les preuves disponibles, obtenues par mener des essais médicaux standard.16 Dans un tel cas, les chercheurs ont à leur disposition un corpus bien fourni de connaissances scientifiques ainsi que des méthodes expérimentales bien établies à partir desquelles ils déduisent des informations fiables sur les produits chimiques qui causer des dommages à la santé humaine. Considérez à nouveau Google Flu Trends. Les chercheurs pourraient-ils, les médecins et les autorités de santé publique en déduisent des relations causales existantes qui lient souche du H1N1/09 avec un grand ensemble de données? Plus précisément, la question que L'enjeu est-il de savoir s'il existe un algorithme capable de recréer des relations de causalité existant entre la grippe A et les requêtes à partir des données ? Cette question garde en fait scientifiques et ingénieurs en haleine, car un grand triomphe du Big Data serait capable de déduire des relations causales à partir de données par des moyens non expérimentaux. Les réponses à cette question divise pourtant chercheurs et philosophes.

En 1993, Peter Spirtes, Clark Glymour et Richard Scheines (ci-après SGS) ont présenté leur point de vue sur la façon d'inférer des relations causales à partir de données dans un livre qu'ils ont appelé Causalité, prédiction et recherche. Là, ils prétendaient avoir trouvé une méthode pour découvrir des relations causales basées sur des données et qui ne nécessiteraient aucune connaissance du sujet. Cela signifie que les scientifiques et les ingénieurs pourraient déduire structures causales à partir des données collectées en exécutant simplement un algorithme spécial,

---

16 Il est intéressant de noter que le National Cancer Institute présente quatre catégories de relations causales, en fonction de la force des preuves disponibles. Il s'agit du "niveau 1 : les preuves sont suffisantes pour déduire une relation causale », « niveau 2 : les preuves sont suggestives mais pas suffisantes pour déduire une relation causale », « niveau 3 : les preuves sont insuffisantes pour déduire la présence ou l'absence d'une relation causale (qui englobe des preuves rares, de mauvaise qualité ou contradictoires) » et « niveau 4 : preuves suggère l'absence de relation causale » (US Department of Health and Human Services 2014).

moins de leurs connaissances antérieures sur les données ou leur structure causale. Comme prévu, une telle un algorithme s'est avéré assez sophistiqué, combinant théorie des graphes, statistiques, philosophie et informatique.

Bien sûr, l'algorithme de SGS n'était pas la première tentative d'inférer la causalité à partir des données. Les méthodes conventionnelles comprennent l'exploration de données, l'analyse de régression, les modèles de chemin et l'analyse factorielle, parmi d'autres méthodes causales existantes issues de l'économétrie et de la sociométrie contemporaines. Cependant, SGS a affirmé que la leur était une méthode supérieure. SGS critiquent en particulier l'analyse de régression sur la base que, en termes de sa structure causale - ses affirmations causales - n'est pas testable car elle n'implique aucune contrainte dans les données.

Plus généralement, SGS identifie trois problèmes inhérents aux méthodes conventionnelles. D'abord, ils identifient incorrectement les hypothèses causales, ils excluent les hypothèses causales et enfin ils incluent également de nombreuses hypothèses qui n'ont aucune signification causale. Deuxième, les spécifications de la distribution imposent généralement l'utilisation de procédures numériques limités par des raisons statistiques ou numériques. Et troisièmement, les restrictions sur le la recherche dans l'espace de données génère généralement une seule hypothèse, échouant de cette manière pour produire des hypothèses alternatives qui pourraient être valides étant donné le même espace de données.

La différence fondamentale entre l'algorithme de SGS et les tentatives conventionnelles est donc qu'aucune de ces dernières ne permet effectivement d'identifier la bonne relation causale. structures. Au contraire, et au mieux, ils sont capables de découvrir des schémas d'association, qui n'ont pas nécessairement une structure causale. Pour chacune de ces méthodes, la stratégie générale n'est pas satisfaisante puisque le but n'est pas seulement d'identifier la distribution estimée, mais aussi d'identifier la structure causale et de prédire l'avenir. résultats causaux des variables. L'algorithme SGS, prétend-on, est conforme à toutes de cela.

Dans ce contexte, SGS a développé un algorithme appelé TETRAD17. Maintenant, pour TETRAD pour réussir, les auteurs devaient prêter attention à trois éléments clés. Premièrement la idée d'un système causal avec une précision suffisante pour l'analyse mathématique nécessaire à rendre explicite. Deuxièmement, il était important de généraliser leur point de vue pour saisir une large éventail de pratiques scientifiques afin de comprendre les possibilités et les limites de la découverte des structures causales. Et troisièmement, ils devaient caractériser le probabilités prédites par une hypothèse causale, étant donné une intervention à une valeur (3).

En conséquence, SGS a pu proposer un algorithme qui relie les structures causales avec un graphe orienté avec les probabilités attribuées à chaque sommet de la graphique. De plus, les auteurs ont affirmé que l'algorithme concernait la découverte des structures causales dans les systèmes linéaires et non linéaires, les systèmes avec et sans rétroaction, les cas dans lesquels l'appartenance aux échantillons observés est influencée par les variables à l'étude, et une poignée d'autres cas (4). Quant aux détails de la Algorithme TETRAD, je suggère au lecteur intéressé d'aborder directement le livre de SGS et site web.18 Ici n'est pas le bon endroit pour la lourde machinerie probabiliste et mathématique qui s'y trouve. Un bref aperçu des objections que SGS a dû le visage, cependant, est en ordre.

---

17 Pour un version à jour http:// de l'algorithme TETRAD, visite www.phil.cmu.edu/tetrad/current.html

18 Voir http://www.phil.cmu.edu/projects/tetrad/

Quelques années plus tard après la publication du livre de SGS, Paul Humphreys et David Freedman ont lancé une série d'objections visant à montrer que SGS n'a pas accompli ce qu'ils prétendaient ((Humphreys 1997) et (Freedman et Humphreys 1999)). Le cœur des objections de Humphreys et Freedman a deux sources. D'une part, l'analyse de SGS est adaptée à des aspects techniques qui ne reflètent pas le sens du lien de causalité. En fait, selon Humphreys et Freedman, la causalité est supposée dans les algorithmes mais jamais fournie : « les causes directes peuvent être représentées par des flèches lorsque les données sont fidèles au véritable graphe causal qui génère les données » (116). En d'autres termes, la causalité est définie en termes de causalité, avec peu de valeur ajoutée. La deuxième objection est fondée sur l'absence d'un traitement satisfaisant des problèmes réels d'inférence statistique issus de données imparfaites.

La très brève discussion présentée ci-dessus sur la reconstruction des relations causales à partir des données met en évidence un point important pour nous, à savoir l'existence de problèmes sérieux pour identifier les structures causales en informatique. Comme indiqué, ni les données produites par une simulation informatique et dont les chercheurs ne détiennent aucune information sur leurs structures causales, ni les données du Big Data ne peuvent être facilement utilisées pour déduire des relations causales.

Peut-être qu'une solution plus sur mesure est nécessaire. À cet égard, l'un des rares défenseurs de la causalité dans le Big Data est Wolfgang Pietsch, qui dans un article récent a plaidé pour le caractère causal de la modélisation dans le Big Data (Pietsch 2015). Pietsch présente et analyse plusieurs algorithmes, concluant qu'ils sont capables d'identifier la pertinence causale sur la base de l'induction éliminatoire et d'un compte de causalité faisant la différence. Cela signifie qu'un phénomène est examiné sous l'angle de la variation systématique des conditions potentiellement pertinentes afin d'établir la pertinence causale - ou la non-pertinence causale - de ces conditions. Selon Pietsch, celles-ci doivent être effectuées par rapport à un certain contexte ou arrière-plan déterminé par d'autres conditions (Pietsch 2015). L'auteur part donc du principe que la pertinence causale d'une condition pourrait être déterminée par la méthode de la différence. À la base, la méthode compare deux instances de relations causales qui ne diffèrent que sur la condition α, et s'accordent dans toutes les autres circonstances. Si, dans un cas, la condition α et le phénomène Θ sont présents, et dans un autre cas, les deux sont absents, alors α est causalement pertinent pour Θ. Une façon de déterminer cela consiste à retenir le contrefactuel suivant : si α ne s'était pas produit, Θ ne se serait pas produit. Suivant notre exemple en médecine, on pourrait dire que s'il n'y avait pas eu de tabagisme, il n'y aurait pas eu de cancer.

Bien sûr, je simplifie le compte rendu différentiel de la causalité. Au fur et à mesure que l'on entre dans les détails, il est possible de découvrir ce que ce compte a à offrir ainsi que ses limites. Les problèmes évoqués ci-dessus ne constituent qu'un exemple du type de difficulté que partagent ces deux paradigmes de la recherche scientifique. Fait intéressant, on pourrait en fait trouver plus de différences qui distinguent les simulations informatiques et le Big Data. L'une de ces différences largement acceptée est que les simulations examinent les implications d'un modèle mathématique, tandis que le Big Data cherche à trouver des structures dans de grands ensembles de données. Ceci, à son tour, nous aide à délimiter comme diamétralement opposées les sources à partir desquelles les chercheurs recueillent des données. Autrement dit, alors que les simulations informatiques produisent de grandes quantités de données à partir d'un modèle donné, le Big Data recrée une structure sur un phénomène donné à partir de grands ensembles de données. Une autre différence est que la capacité

La capacité de chacun à prédire, confirmer et observer un phénomène donné est différente. Alors que les simulations informatiques ancrent généralement ces fonctions épistémiques dans les hypothèses du modèle, le Big Data repose sur la collecte de méthodes et de processus de conservation sous-jacents à la structure des données. Leur approche méthodologique est donc différente.

Le point de départ des deux est donc différent. Alors que les simulations informatiques implémentent et résolvent un modèle de simulation, le Big Data examine une collection de données. Cela conduit à une dernière différence; à savoir, la nature de nos inférences sur le système cible. Dans les simulations informatiques, on tire généralement les conséquences du modèle de simulation, alors que le Big Data est principalement piloté par des inférences inductives.

6.4 Remarques finales

Lorsque Jim Gray a fait référence aux simulations informatiques et au Big Data comme troisième et quatrième paradigme respectivement, il prévoyait, je crois, l'avenir de la recherche scientifique et technique. Dans son dernier discours avant sa disparition en mer en 2007, il a déclaré : « [l]e monde de la science a changé, et cela ne fait aucun doute. Le nouveau modèle prévoit que les données soient saisies par des instruments ou générées par des simulations avant d'être traitées par des logiciels et que les informations ou connaissances qui en résultent soient stockées dans des ordinateurs. Les scientifiques ne peuvent examiner leurs données qu'assez tard dans ce pipeline. (Gray 2009, xix).

Les simulations informatiques se sont avérées sans aucun doute un outil fondamental pour l'avancement et le développement de la recherche scientifique et technique. Dans ce chapitre, j'ai tenté d'aborder de manière critique les simulations informatiques et le Big Data en montrant leur programme de recherche et en quoi ils diffèrent sur des questions spécifiques. Je suis conscient que je n'ai fait qu'effleurer la surface de deux méthodes profondément enracinées dans l'état actuel – et futur – de la recherche scientifique et technique. Néanmoins, laissez ce chapitre être une petite contribution pour une discussion beaucoup plus large.

Les références

Alison, Wylie. 2002. Penser à partir des choses: Essais sur la philosophie de l'archéologie. Presse de l'Université de Californie.

Barberousse, Anouk et Vorms Marion. 2013. « Simulations informatiques et données empiriques ». Dans Computer Simulations and the Changing Face of Scientific Experimentation, édité par Juan M. Duran et Eckhart Arnold. Édition des boursiers de Cambridge.

Beyer, Mark et Douglas Laney. 2012. "L'importance du 'Big Data' : une définition tion. Gartner.

Oiseau, Alexandre. 2013. "Thomas Kuhn". Dans The Stanford Encyclopedia of Philoso phy, automne 2013, édité par Edward N. Zalta.

Buneman, Peter, James Cheney, Wang-Chiew Tan et Stijn Vansummeren. 2008. "Bases de données organisées." Dans Actes du vingt-septième symposium ACM SIGMOD SIGACT-SIGART sur les principes des systèmes de bases de données, 1–12. GOUSSES '08. New York, NY, États-Unis : ACM.

Bunge, Mario. 2017. Causalité et science moderne. Routledge.

Bunnik, Anno, Anthony Cawley, Michael Mulqueen et Andrej Zwitter. 2016. Grand Défis des données : société, sécurité, innovation et éthique. Springer.

Callebaut, Werner. 2012. "Perspectivisme scientifique : la réponse d'un philosophe des sciences au défi de la biologie des mégadonnées." Études d'histoire et de philosophie des sciences Partie C : Études d'histoire et de philosophie des sciences biologiques et Sciences biomédicales 43 (1): 69–80. http : / / www . sciencedirect. com/science/article/pii/S1369848611000835.

Chatrchyan, Serguei, V. Khachatryan, AM Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Ero, C. Fabjan, M. Friedl, et al. 2014. "Measurement of the Properties of a Higgs Boson in the Four-Lepton Final State." Examen physique D 89 (9): 1–75.

Choudhury, Suparna, Jennifer R. Fishman, Michelle L. McGowan et Eric T. Juengst. 2014. « Big data, science ouverte et cerveau : enseignements tirés génomique. Accès 13. Déc. 2017, Frontiers in Human Neuroscience 8 (239). https://www.frontiersin.org/articles/10.3389/fnhum. 2014.00239/complet.

Collmann, Jeff et Sorin Adam Matei, éd. 2016. Raisonnement éthique dans le Big Data : Une analyse exploratoire. Springer.

Costa, FF 2014. "Mégadonnées en biomédecine." Découverte de médicaments aujourd'hui 19 (4): 433–440.

Critchlow, Terence et Kerstin Kleese van Dam, éd. 2013. Science intensive en données ence. Chapman/Hall/CRC.

De Pierris, Graciela et Michael Friedman. 2013. "Kant et Hume sur la causalité." Dans The Stanford Encyclopedia of Philosophy, édité par Edward N. Zalta. (Édition Hiver 2013). https://plato.stanford.edu/archives/ win2013/entries/kant-hume-causality/.

Denzin, Norman K. 2006. Méthodes sociologiques : Un livre source. Aldine Transac tion.

Dowe, Phil. 2000. Causalité physique. La presse de l'Universite de Cambridge.

Edwards, Kieran et Mohamed Medhat Gaber. 2014. Astronomie et mégadonnées.
Une approche de regroupement de données pour identifier la morphologie incertaine des galaxies.
Springer.

Faucon, Andréa. 2015. "Aristote sur la causalité." Dans l'Encyclopédie de Stanford de
Philosophie, édité par Edward N. Zalta. (édition printemps 2015). https://pl
ato.stanford.edu/archives/spr2015/entries/aristotle causality/.

Floridi, Luciano. 2012. "Big Data et leur défi épistémologique." Philoso phy and Technology 25 (4):
435–437.

Freedman, David, et Paul W. Humphreys. 1999. « Y a-t-il des algorithmes qui découvrent une
structure causale ? » Synthèse 121 (1): 29–54.

Galisson, Pierre. 1996. "Simulations informatiques et zone commerciale". La désunion
of Science: Frontières, contextes et pouvoir: 119–157.

Ginsberg, Jeremy, Matthew H Mohebbi, Rajan S Patel, Lynnette Brammer, Mark S
Smolinski et Larry Brillant. 2009. "Détection des épidémies de grippe à l'aide
Données de requête de moteur de recherche. » Nature 457 (7232): 1012–4.

Gris, Jim. 2009. "Jim Gray sur eScience : une méthode scientifique transformée." Dans
The Fourth Paradigm, édité par Tony Hey, Stewart Tansley et Kristin Tolle,
xvii–xxxi. Redmond, Washington : Microsoft Research.

Salut, Tony, Stewart Tansley et Kristin Tolle, éds. 2009. Le quatrième paradigme :
Découverte scientifique à forte intensité de données. Microsoft Corporation.

Hick, Jason, Dick Watson et Danny Cook. 2010. HPSS à l'ère de l'échelle extrême :
Rapport au DOE Office of Science sur HPSS en 2018-2022. Département d'En
ergie.

Humphreys, Paul W. 1997. "Une évaluation critique des algorithmes de découverte causale."
Dans Causality in Crisis?: Statistical Methods and the Search for Causal Knowl edge in the
Social Sciences, édité par Vaughn R. McKim et Stephen P.
Turner, 249-263.

————. 2004. Extension de nous-mêmes : science computationnelle, empirisme et méthode
scientifique. Presse universitaire d'Oxford.

————. 2013a. « Analyse des données : modèles ou techniques ? » Fondements des sciences 18
(3): 579–581.

————. 2013b. « A quoi correspondent les données ? Dans Computer Simulations and the
Chang ing Face of Scientific Experimentation, édité par Juan M. Duran et Eckhart Arnold.
Édition des boursiers de Cambridge.

IDC. 2014. "L'univers numérique des opportunités : des données riches et l'augmentation Valeur de l'Internet des objets. Consulté le 11 décembre 2017. https:// www.emc.com/leadership/digital-univers/2014iview/ executive-summary.htm.

Kitchin, Rob. 2014. « Mégadonnées, nouvelles épistémologies et changements de paradigme ». Big Data & Société 1 (1): 2053951714528481.

Kuhn, TS 1962. La structure des révolutions scientifiques. Université de Chicago Presse.

————. 1970. « Logique de la découverte ou psychologie de la recherche ? » Dans Critique et the Growth of Knowledge, édité par Imre Lakatos et Alan Musgrave, 4:1–24. La presse de l'Universite de Cambridge.

Laney, Doug. 2001. "META Delta". Stratégies de livraison d'applications 949 (février 2001): 4.

Léonelli, Sabine. 2014. « Quelle différence fait la quantité ? Sur l'épistémologie du Big Data en biologie. Big data & société, non. 1 : 1–11.

Liberman, Marc. 2003. "Linguistique Zettascale."

Mackie, JL 1980. Le Ciment de l'Univers. Presse universitaire d'Oxford.

Mahajan, RL, R. Mueller et CB Williams, J. Reed, TA Campbell et N. Ramakrishnan. 2012. "Cultiver les technologies émergentes et Black Swan." Congrès et exposition internationaux de génie mécanique ASME 6: 549–557.

Mayer-Schonberger, Viktor et Kenneth Cukier. 2013. Big Data : une révolution Cela transformera notre façon de vivre, de travailler et de penser. Cour Houghton Mifflin Har, mars.

Mittelstadt, Brent Daniel et Luciano Floridi. 2016a. "L'éthique des mégadonnées : enjeux actuels et prévisibles dans les contextes biomédicaux." Sciences et ingénierie Éthique 22 (2): 303–341.

————. 2016b. L'éthique des mégadonnées biomédicales. Springer.

NSF. 2012. "Techniques et technologies de base pour faire progresser la science des mégadonnées \& Ingénierie (BIGDATA). NSF 12-499. https://www.nsf.gov/ pubs/2012/nsf12499/nsf12499.htm.

————. 2014. "Techniques et technologies critiques pour l'avancement du Big Data Sci . nsf. ence \& Engineering (BIGDATA). NSF 14-543. https : / / www gov/pubs/2014/ nsf14543/nsf14543.htm.

————. 2016. "Techniques, technologies et méthodologies critiques pour l'avancement des fondements et des applications des sciences et de l'ingénierie du Big Data (BIG DATA)". NSF 16-512. https : / / www . nsf. gouv / pubs / 2016 / nsf16512/nsf16512.htm.

Perle, Judée. 2000. Causalité. Modèles, raisonnement et inférence. Université de Cambridge
Versité Presse.

Pietsch, Wolfgang. 2015. "La nature causale de la modélisation avec le Big Data." Philosophy &
Technology 29 (2): 137–171.

Rimoldi, Adèle. 2011. "Stratégies de simulation pour l' {vphantom}ATLASvphantom{}
Expérimentez au {vphantom}LHCvphantom{}. » Dans Journal of Physics: Conference Series,
vol. 331. Édition IOP.

Rohrlich, Fritz. 1990. "Simulation par ordinateur dans les sciences physiques." Philosophie
of Science Association 2:507–518.

Safran, Charles, Meryl Bloomrosen, W Edward Hammond, Steven Labkoff, Suzanne
Markel-Fox, Paul C Tang et Don E Detmer. 2006. « Vers un cadre national pour l'utilisation
secondaire des données de santé : une étude américaine sur l'informatique médicale
Livre blanc de l'association. Journal of the American Medical Informatics Association 14 (1): 1–9.

Salmon, Wesley C. 1998. Causalité et explication. Presse universitaire d'Oxford.
ISBN : 978-0-19-510864-4, consulté le 4 août 2016.

Spirtes, Peter, Clark Glymour et Richard Scheines. 1993. Causalité, prédiction,
et Rechercher. Presse du MIT.

Département américain de la santé et des services sociaux. 2014. Les conséquences sanitaires
du tabagisme - 50 ans de progrès : un rapport du Surgeon General - US
Département de la santé et des services sociaux, Centers for Disease Control and
Prévention, Centre national de prévention des maladies chroniques et de promotion de la santé,
Office on Smoking and Health. Atlanta, Géorgie.

Winsberg, Éric. 1999. "Modèles de sanction : l'épistémologie de la simulation". SCI
rence dans le contexte 12:275–292.

————. 2001. "Simulations, modèles et théories : systèmes physiques complexes
et leurs représentations. Philosophie des sciences 68 (S1): S442. ISSN : 0031-
8248.

Woodward, James. 2003. Faire bouger les choses. Presse universitaire d'Oxford.

Yoshii, Yuzuru, Kentaro Motohara, Takashi Miyata et Natsuko Mitani. 2009. "Le
Le télescope de 1 m de l'observatoire d'Atacama a commencé son fonctionnement scientifique,
détectant la ligne d'émission d'hydrogène du centre galactique dans l'infrarouge
Lumière." R. http://www.su-tokyo.ac.jp/en/press/2009/15.
html.

Chapitre 7
Éthique et simulations informatiques

Ce chapitre a pour seul but de poser la question suivante : y a-t-il une éthique qui émerge dans le cadre des simulations informatiques ? Afin de répondre correctement à cette question, nous devons enquêter sur la littérature spécialisée pour voir comment les problèmes ont été abordés. Le premier problème que nous rencontrons est la question de savoir si une telle éthique existe réellement, ou plutôt si les préoccupations morales dans les simulations informatiques peuvent être abordées par un cadre éthique plus familier. Les auteurs concernés par l'utilisation générale des ordinateurs ont déjà répondu par la négative à cette question, invitant à une évaluation appropriée de l'éthique informatique. Il est intéressant de noter ici la symétrie avec nos études antérieures sur l'épistémologie et la méthodologie des simulations informatiques, notamment avec la discussion présentée en introduction quant à leur nouveauté philosophique. On pourrait conjecturer que cela fait partie du processus d'introduction de nouvelles technologies dans la recherche scientifique et technique qui soulève les sourcils de l'esprit philosophiquement sceptique.

Malheureusement, très peu de travail a été consacré à l'élaboration complète d'une éthique pour s'adapter aux simulations informatiques. Seule une poignée d'auteurs ont abordé la question ci-dessus à leur valeur nominale. Dans ce chapitre, je les présente et les discute ainsi que leurs points de vue. À cet égard, j'aborderai les préoccupations concernant la fiabilité des simulations informatiques – un sujet qui nous est familier à ce stade – la représentation et la pratique professionnelle. De ces trois approches, seule la dernière ne dépend pas directement de l'épistémologie et de la méthodologie de la simulation informatique, et c'est pourquoi je consacrerai plus de temps à discuter plus en détail de la pratique professionnelle et d'un code d'éthique pour les simulations informatiques. Je reproduirai et discuterai également le code de déontologie officiel des chercheurs en simulations informatiques.

7.1 Éthique informatique, éthique en ingénierie et éthique en science

Au cours des trente dernières années, il y a eu un intérêt croissant pour la compréhension de l'éthique dans le contexte des ordinateurs. Un tel intérêt découle de la place omniprésente que les ordinateurs ont dans notre vie quotidienne, ainsi que de leur présence dans les domaines scientifiques et

pratique de l'ingénierie. L'« éthique informatique », comme on l'appelle aujourd'hui est la branche de l'éthique appliquée centrée sur les questions soulevées par les ordinateurs, a connu des débuts difficiles et a eu du mal à s'affirmer comme une discipline à part entière.

L'un des articles fondamentaux de l'éthique informatique a été écrit par James H. Moor et paru en 1985 sous le titre « Qu'est-ce que l'éthique informatique ? (Moor 1985). À À l'époque, les chercheurs travaillant sur l'éthique appliquée avaient la ferme conviction que l'introduction de nouvelles technologies, comme l'ordinateur, ne suggérait pas de nouveaux problèmes moraux. vaut la peine d'être étudié; des problèmes éthiques plutôt anciens et plus familiers pourraient être appliqués à ces technologies aussi bien1. Certains ont même tourné en dérision l'idée d'une éthique des ordinateurs en suggérant une éthique des machines à laver, et une éthique des voitures. C'est bientôt est devenu clair qu'il n'y a aucune raison réelle de comparer les appareils et les moyens de transport avec ordinateurs.

Mais revenons à l'éthique informatique. Le défi de Moor était double. Sur le d'une part, il devait montrer en quoi précisément les études sur l'éthique menées à son époque n'était pas en mesure de comptabiliser les ordinateurs. D'autre part, il devait montrent comment les ordinateurs ont apporté de nouvelles et authentiques questions éthiques sur la table. Maure stratégie était d'analyser la nature et l'impact social de la technologie informatique, et la formulation et la justification correspondantes des politiques d'utilisation éthique de ces technologie. Son inquiétude portait principalement sur le vide politique que les ordinateurs semblaient laisser entendre, comme c'était souvent le cas à l'époque, qu'il n'y avait ni réglementation ni politiques de conception, de programmation et d'utilisation des logiciels et du matériel.

Pour relativiser les idées de Moor, il existe en effet des cas où l'ordinateur n'est qu'un accessoire du problème moral, et il pourrait donc être résolu du point de vue des théories éthiques classiques sur l'individu et la société. Pour exemple, qu'il est mal de voler un ordinateur. A cet égard, l'analyse ne serait pas ont été très différents du vol d'une machine à laver ou d'une voiture. Mais Moor avait autre chose en tête. Pour lui, les problèmes d'éthique informatique survenaient parce qu'il n'y avait pas de cadres éthiques qui indiqueraient comment la technologie informatique devrait être utilisé, ni comment les utilisateurs, les fabricants, les programmeurs et, fondamentalement, tous ceux qui sont impliqués dans le développement de logiciels et de matériel devraient se comporter. Comme il a raison dit, "une tâche centrale de l'éthique informatique est de déterminer ce que nous devrions faire dans un tel cas, c'est-à-dire formuler des politiques pour guider nos actions. Bien sûr, certaines situations éthiques nous confrontent en tant qu'individus et d'autres en tant que société. L'éthique informatique comprend prise en compte des politiques personnelles et sociales pour l'utilisation éthique de l'ordinateur technologie » (266). Il s'avère qu'il a prévu bon nombre des questions éthiques qui se posent au centre d'une éthique des simulations informatiques (voir section 7.3.2).2

Ainsi furent les modestes débuts de l'éthique informatique. Moor principalement discuté la présence révolutionnaire des ordinateurs dans notre vie quotidienne, en tant qu'individus, institutions, gouvernements et société. Mais ce n'est qu'à la publication de Deborah Johnson's "Computer Ethics" (Johnson 1985) que le sujet a attiré plus de visibilité

---

1 La différence entre l'éthique et la morale est souvent formulée en termes que l'éthique est la science de la morale, alors que la morale est la pratique de l'éthique. Alors que la morale est l'ensemble des valeurs et normes, l'éthique est l'étude formelle et l'encodage de ces normes. je ne ferai pas ça différenciation ici. Au lieu de cela, je me référerai aux deux concepts indistinctement.

2 Un autre auteur qui a anticipé une grande partie de la discussion contemporaine est (McLeod 1986).

ité. Maintenant considéré comme un livre classique, Johnson a également contribué à la justification de
l'existence et l'unicité de l'éthique informatique. Contrairement à Moor, à qui elle a appris à jeter les bases de la
question, les principaux intérêts de Johnson se concentraient
sur les questions liées à Internet, à la confidentialité et aux droits de propriété. Prenons par exemple la compréhension
de la vie privée à l'ère d'Internet. L'exemple de Johnson sont les entreprises
comme Amazon, qui a affirmé qu'ils avaient besoin de stocker des informations sur leur
clients pour mieux les servir. La question ici est de savoir si l'envoi d'informations sur les nouvelles sorties de livres
fournit réellement un service, car l'entreprise
revendiqué au milieu des années 1990, ou c'était en quelque sorte une atteinte à la vie privée, puisque,
très probablement, les clients n'ont pas explicitement demandé ces informations. Ainsi compris,
Johnson affirme que "la charge de la preuve incombe aux défenseurs de la vie privée pour montrer soit que
il y a quelque chose de nuisible dans la collecte et l'échange d'informations ou qu'il y a
y a-t-il un avantage à tirer d'une collecte d'informations contraignante » (119). Les problèmes liés à la collecte de
grandes quantités de données à des fins autres que la «fourniture d'un
service » n'était évidemment pas sur la table dans les années 1990. Aujourd'hui, les questions éthiques que soulèvent
des entreprises telles qu'Amazon, Google et d'autres sont en fait de savoir si
ils façonnent notre intérêt pour les livres, les films, la musique et, finalement, tout
qui concerne notre personnalité, y compris nos opinions politiques.3

Les études sur l'éthique informatique mettent donc l'accent sur des questions sur la façon dont un
l'action utilisant un ordinateur peut être moralement bonne ou mauvaise selon ses motifs, ses conséquences, son
universalité et sa nature vertueuse. Ainsi comprises, ces études ont une
large éventail de sujets, de l'utilisation des ordinateurs dans et pour le gouvernement, à leur
utiliser dans notre vie privée. Cela signifie qu'ils ne se concentrent pas nécessairement exclusivement sur
les utilisations scientifiques et techniques des ordinateurs, sans parler spécifiquement des simulations informatiques.
Laissez-moi vous expliquer ceci. Des études sur la vie privée et les droits de propriété, par exemple,
ont un impact sur l'utilisation des ordinateurs dans des contextes scientifiques et techniques.
La propriété est généralement considérée comme un mécanisme de contrôle des données, qui pourrait facilement
inclure des cas de données provenant de tests médicaux et de dépistage de drogues. Un problème qui émerge dans ce
le contexte est la distinction entre données académiques et données publiques, d'une part, et
données commerciales, d'autre part. Cette distinction permet aux chercheurs, aux individus ou aux institutions de
conserver des attentes réalistes quant aux utilisations et implications potentielles de
leurs données (Lupton 2014). Ainsi comprises, les préoccupations éthiques concernant les droits de propriété
ont un impact sur l'utilisation des ordinateurs dans la pratique scientifique et technique.
Nos préoccupations ici, cependant, portent sur les questions morales émergeant exclusivement de la
utilisation des ordinateurs dans des contextes scientifiques et techniques, plutôt que des extrapolations
d'autres champs ou régions. Un exemple de ceci est les questions morales adaptées à la représentation des
simulations informatiques (7.2.2). Les préoccupations qui émergent dans ce contexte sont,
sans doute, exclusivement sur les utilisations de simulations informatiques.

---

3 Les études sur le Big-Data constituent un nouveau locus classicus sur les questions concernant les implications éthiques
des entreprises traitant de grandes quantités de données. Pour des cas d'utilisations réussies du Big-Data, voir
les travaux de (Mayer-Schonberger et Cukier 2013) et (Marr 2016) ; et pour le traitement philosophique, voir entre
autres (Zwitter 2014), (Beranger 2016) et (Mittelstadt et Floridi 2016).
Nous pourrions également mentionner les nombreuses institutions - gouvernementales, éducatives et du secteur privé
– qui sont créés pour comprendre l'éthique informatique. Un excellent exemple en est l'Oxford
Internet Institute, qui fait partie de l'Université d'Oxford.

À ce stade, nous devons nous demander s'il existe une branche de l'éthique informatique qui aborde les problèmes moraux émergeant dans le contexte des simulations informatiques ? Étonnamment, le nombre d'articles publiés sur l'éthique impliquant des simulations informatiques est très faible. Comme nous le verrons dans la section suivante, je présente quels sont sans doute les trois principaux arguments sur l'éthique des simulations informatiques disponibles dans la littérature spécialisée. Le premier aborde le travail de TJ Williamson, ingénieur et architecte qui présente les enjeux éthiques des simulations informatiques en examinant leurs problèmes épistémologiques. À cet égard, il relie joliment nos discussions passées sur l'épistémologie et la méthodologie des simulations informatiques avec des préoccupations morales. La deuxième étude est celle du philosophe et membre de la Société internationale d'éthique et des technologies de l'information Philip Brey. Brey fait le cas intéressant qu'un aspect significatif de l'éthique des simulations informatiques vient de leur capacité à (mal)représenter un système cible. Étant donné qu'une telle (fausse) représentation peut prendre plusieurs formes, il s'ensuit qu'il existe des responsabilités professionnelles spécifiques attachées à chacune d'entre elles. Notre troisième discussion provient d'un plus grand corpus de littérature écrite par Tuncer Oren, Andreas Tolk et d'autres. Oren est un ingénieur formé qui a beaucoup travaillé sur l'établissement de problèmes éthiques dans les simulations informatiques. Il a également été consultant principal dans l'élaboration du code de déontologie des chercheurs et des praticiens de la simulation informatique (voir section 7.3).

Nous savons que les simulations informatiques combinent deux composantes principales, à savoir l'informatique et les sciences générales - ou l'ingénierie, selon ce qui est simulé. Cela suggère que toute étude sur l'éthique des simulations informatiques doit inclure ces deux – ou trois – éléments. C'est-à-dire que les questions découlant de l'éthique des ordinateurs ainsi que de l'éthique des sciences et de l'ingénierie alimentent les études éthiques sur les simulations informatiques. Nous avons discuté de l'éthique informatique, en présentant très brièvement quelques-unes des principales préoccupations qui émergent dans ce domaine. Avant de discuter en détail des approches actuelles de l'éthique des simulations informatiques, parlons brièvement de l'éthique des sciences et de l'ingénierie.

Les études sur l'éthique des sciences et de l'ingénierie s'intéressent à différentes questions liées au comportement professionnel des chercheurs, à la manière dont ils mènent leurs travaux et expérimentations, et aux conséquences que ces produits et expérimentations ont sur la société. Maintenant, alors qu'une grande partie du travail quotidien du chercheur peut être effectuée sans aucune implication éthique visible, il existe plusieurs dispositions offertes par les études éthiques qui renforcent les bonnes pratiques scientifiques et d'ingénierie, ainsi qu'un cadre éthique dans lequel placer les implications de leurs pratiques. Les thèmes clés incluent la fabrication de données, qui se produit généralement lorsque les chercheurs veulent proposer plus vigoureusement et avec plus de conviction des hypothèses particulières. À titre d'exemple, nous pouvons citer les fausses greffes de peau de Summerlin en 1974 et l'origine des expériences sur le cancer de Spectorin en 1980 et 1981. De même, la falsification des données se produit lorsque les résultats sont modifiés pour correspondre aux résultats escomptés par le chercheur. Des cas de falsification de résultats par une jeune chercheuse ont été signalés parce qu'elle était sous pression pour dupliquer les résultats de son conseiller. Un exemple de cela est la falsification des résultats sur les récepteurs de l'insuline par Soman de 1978 à 1980.

analyse minative de ces questions, voir (Spier 2012). Les travaux d'Adam Briggle et
Cart Mitcham sont également essentiels pour les études sur l'éthique de la recherche scientifique (Briggle
et Mitcham 2012). Dans ce qui suit, je présente et discute certaines de ces questions dans
le contexte de l'éthique des simulations informatiques.

7.2 Un aperçu de l'éthique dans les simulations informatiques

7.2.1 Williamson

Commençons par (Williamson 2010) dont les travaux, bien que chronologiquement plus
plus récent que celui d'Oren et de Brey, a l'avantage d'être conceptuellement plus proche de nos discussions
les plus récentes sur la fiabilité des simulations informatiques - voir chapitre 4.

La principale motivation qui guide les préoccupations éthiques de Williamson est que l'ordinateur
les simulations pourraient contribuer à améliorer la vie humaine ainsi qu'à contribuer à un environnement
durable, aujourd'hui et à l'avenir. Malgré ces idéaux optimistes, Williamson
choisit de discuter des problèmes éthiques d'un point de vue négatif. Plus précisement,
Williamson soutient que les affirmations épistémiques sur la fiabilité des simulations informatiques peuvent
conduire à de fausses impressions de précision - et donc de légitimité - conduisant
finalement à des utilisations moralement inappropriées de simulations informatiques, telles que de mauvaises
prise de décision et affectation erronée des ressources. En d'autres termes, l'épistémologie de la simulation
informatique permet de questionner les valeurs et l'éthique. Dans ce
Dans ce contexte, Williamson montre qu'une combinaison de préoccupations épistémiques concernant
l'exactitude/la validité et un critère donné de fiabilité sont constitutives des problèmes moraux des simulations
informatiques. Analysons-les d'abord tour à tour et voyons plus tard comment
ils se combinent pour donner lieu aux préoccupations morales de Williamson.

Selon Williamson, la précision ou la validité d'un modèle de simulation est conçue
comme le degré auquel le modèle correspond aux phénomènes du monde réel sous
examen minutieux. Afin de fonder la précision du modèle de simulation, Williamson s'appuie
sur les méthodologies standards de validation empirique, « qui comparent
résultats avec des données mesurées dans le monde réel », la vérification analytique, « qui compare la sortie
de simulation d'un programme, d'un sous-programme, d'un algorithme ou d'un objet logiciel
avec des résultats issus d'une solution analytique connue ou d'un ensemble de solutions quasi-analytiques »,
et la comparaison intermodale, "qui compare le résultat d'un programme avec le
résultats d'autres programmes similaires » (404). Une grande partie de la discussion sur cette question
ont déjà été abordés dans la section 4.

Contrairement à l'exactitude ou à la validation, qui sont essentiellement des questions adaptées au
modèle de simulation, les critères de fiabilité sont étroitement liés aux objectifs d'un
problème donné. Autrement dit, cela dépend de poser les bonnes questions pour l'utilisation correcte
de la simulation. L'exemple qui illustre les critères de fiabilité vient
à partir de simulations informatiques de la performance environnementale des bâtiments. Dans de telles
simulations, les chercheurs doivent toujours poser des questions spécifiques liées à la
le niveau de confort thermique des occupants et les valeurs minimales des dépenses énergétiques (405).

Toutes ces questions supposent une structure plus ou moins cohérente de croyances qui constitutive de la fiabilité de la simulation. Plus précisément, Williamson suggère que la fiabilité des simulations informatiques peut être identifiée en posant quatre questions sur la pertinence.

- Absence : ce type de savoir doit-il être absent (ou présent) ?
- Confusion : y a-t-il une distorsion dans la définition de la connaissance ?
- Incertitude : quel degré de certitude est pertinent ?
- Inexactitude : quelle doit être l'exactitude des connaissances ? Est-ce sans importance parce que n'est pas assez précis ou est-ce inutilement précis ? (Williamson 2010, 405)

Avec ces idées en main, Williamson propose le cadre suivant pour la compréhension appropriée des préoccupations éthiques qui émergent dans le contexte des simulations informatiques.

Crédibilité (et absence)

Selon Williamson, la crédibilité de l'utilisation d'une simulation découle de deux sources. D'une part, la crédibilité de la simulation dépend de sa précision ou de sa validité. En effet, une simulation imprécise ne peut fournir de connaissance (voir notre discussion dans la section 4.1), et par conséquent, il devient une source de toutes sortes de problèmes éthiques. préoccupations. En fait, selon Williamson, il existe une proportion directe entre la précision et la crédibilité : plus une simulation informatique est précise, plus elle est crédible. c'est.

D'autre part, la crédibilité dépend d'une autorité capable de sanctionner la corrélation entre le modèle de simulation – et ses résultats – et le monde réel (406). À cet égard, selon Williamson, «la crédibilité fera défaut si la responsabilité morale (sinon la responsabilité juridique) dans l'application des résultats des simulations n'est pas acceptée à tous les niveaux» (406). Par « tous niveaux », il entend le rôle de l'expert sanctionnant la fiabilité de la simulation. Il est intéressant de noter comment Williamson fait une distinction claire entre la précision fournie par des méthodes dépendantes (telles que les méthodes de vérification et de validation), de l'expert avis. Alors que la plupart des auteurs dans la littérature ne considèrent pas les experts comme fiables sources pour sanctionner les simulations informatiques, Williamson est mal à l'aise avec limiter la crédibilité aux seules méthodes mathématiques. Il insiste pour que l'expert joue un rôle fondamental dans la simulation au point que « des juges réfléchis [...] s'accordent [ce qui] devrait être inclus » (406).

La crédibilité est cependant fragilisée dans deux cas précis, à savoir soit lorsque des les éléments d'un problème sont absents lorsque les experts conviennent qu'ils doivent être inclus dans la simulation (406), et lorsque la simulation est utilisée de manière inappropriée (par exemple, pour représentent des systèmes que la simulation n'est pas capable de représenter).

Illustrons ces points par l'utilisation d'une simulation de performances thermiques dans des immeubles à Adélaïde, Australie (407). Selon l'exemple, les simulations des émissions de gaz à effet de serre produites par le système de chauffage et de refroidissement ne correspondent pas aux performances réelles dérivées de plusieurs années d'énergie d'utilisation finale enregistrée

données. La raison en est que la simulation ne tient pas compte des appareils de chauffage et/ou de refroidissement trouvés dans le National House Energy Rating Scheme - NatHERS. Si ce n'était de l'expert qui détermine que les résultats sont erronés, la simulation aurait pu donner une fausse impression d'exactitude, et donc de légitimité. Les conséquences auraient été de graves dommages à l'environnement et à la santé humaine. De plus, le gouvernement australien n'aurait pas été en mesure de parvenir à une politique durable puisqu'une bonne estimation de la consommation potentielle d'énergie et des émissions de gaz à effet de serre ne fait pas partie de la simulation.

Transférabilité (et confusion)

La transférabilité s'entend comme la possibilité d'utiliser les résultats d'une simulation informatique au-delà de la portée prévue du modèle de simulation. Pour Williamson, les questions concernant la transférabilité doivent inclure la mesure dans laquelle les connaissances faisant autorité devraient être simplement scientifiques – adaptées à la simulation informatique – ou d'autres formes de connaissances doivent également être prises en compte.

Des problèmes moraux dans le contexte de la transférabilité surgissent lorsque des constructeurs et des architectes responsables veulent atteindre les normes imposées par l'industrie (par exemple, Star Rat ing) avec la conviction que, ce faisant, cela réduira davantage la consommation d'énergie et les émissions de gaz à effet de serre. Étant donné que le nombre d'étoiles est calculé à partir d'une simulation des charges globales de chauffage et de refroidissement en tenant compte de certaines hypothèses sur le bâtiment, les résultats ne peuvent pas être facilement transférables à toute autre construction, mais uniquement à celles qui sont conformes aux hypothèses initialement intégrées dans le modèle. À cet égard, une mauvaise compréhension des hypothèses intégrées dans une simulation donnée, ainsi que l'espoir de transférer les résultats d'une simulation dans un contexte différent, pourraient sérieusement fausser le système cible, avec pour conséquence une allocation erronée des ressources sans atteindre l'objectif souhaité (407) .

Fiabilité (et incertitude)

Pour Williamson, plusieurs simulations informatiques ne sont plus fiables car leurs résultats, après une période de temps spécifique, sont incertains par rapport à un système cible donné. Une telle incertitude, il faut le dire, n'est pas le produit de l'ignorance du chercheur, ni d'un manque de précision initiale des résultats de la simulation - ou de ses résultats - mais plutôt de la nature changeante des phénomènes du monde réel simulés. Prenons l'exemple suivant. C'est un fait bien connu que l'isolation thermique en vrac se détériore avec le temps. On s'attend à ce que l'isolation en fibre de verre dans un grenier se comprime considérablement et perde jusqu'à 30 % de son efficacité d'isolation en une dizaine d'années (407). Les implications sur la consommation d'énergie et l'impact environnemental seront bien sûr importantes, entraînant de sérieuses inquiétudes sur la santé humaine, l'utilisation rationnelle des ressources énergétiques, etc.

(Jamieson 2008). Dans l'esprit de Williamson, les études actuelles sur l'éthique de l'informatique

les simulations « ne tiennent pas compte de cette inconnue ou d'autres inconnues connues » (Williamson 2010, 407). Je crois qu'il a raison sur ce compte.

Confirmabilité (et inexactitude)

Les modèles de simulation comprennent une série d'hypothèses de construction, de conjectures et même de données qui ne correspondent pas nécessairement au statut épistémologique des formulations scientifiques et techniques (par exemple, la résistance de surface, l'émissivité du ciel et le coefficient de décharge sont inclus dans une simulation de performance thermique). Inévitablement, ces suppositions et conjectures conduisent à des incertitudes ou, comme Williamson aime à les appeler « des « inexactitudes » dans l'application de la simulation » (407). Un bon exemple de ceux-ci sont les modèles climatiques, où il n'est pas rare que différents modèles de circulation globale rapportent des résultats différents. Cela est principalement dû aux hypothèses intégrées à chaque modèle ainsi qu'aux données utilisées pour les instancier. Williamson le montre avec l'exemple de 17 modèles climatiques de circulation globale simulant une série chronologique de réchauffement de surface moyen mondial, comme indiqué dans le quatrième rapport d'évaluation (AR4) de 2007 du Groupe d'experts intergouvernemental sur l'évolution du climat (voir figure 7.1) (Meehl et al. 2007, 763). Selon lui, toutes les prédictions faites par ces modèles ne peuvent pas être justes ou probablement également fausses. Pour contrer cet effet, la moyenne des résultats est utilisée comme valeur la plus probable de l'état futur.

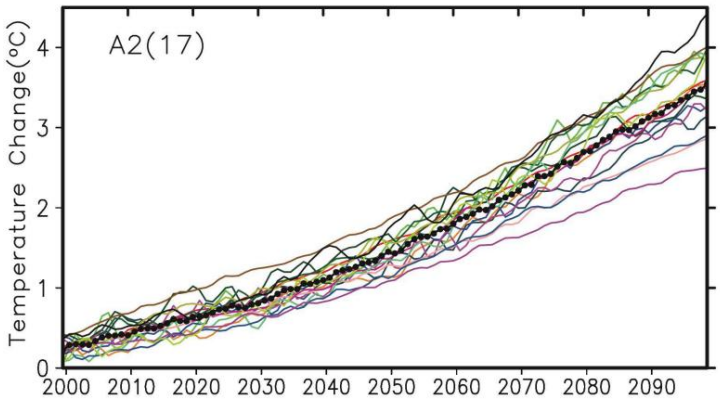


Fig. 7.1 Série chronologique du réchauffement de surface moyen mondial tel que rapporté dans les 17 modèles climatiques de circulation mondiale, Scénario A2 du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC). Les séries moyennes multi-modèles sont marquées de points noirs (Meehl et al. 2007, 763), (Williamson 2010, 408).

Naturellement, les inexactitudes trouvées dans les résultats des simulations informatiques soulèvent des inquiétudes, voire des objections directes, à l'utilisation des résultats pour des questions éthiquement sensibles. Mais les chercheurs sont bien conscients qu'une certitude totale est pratiquement

impossible. Le défi consiste donc à trouver des moyens d'équilibrer le manque d'informations - ou la désinformation inévitable - qui accompagne certains types d'ordinateurs.
simulations et les préoccupations éthiques qui se font pressantes.

Comme nous l'avons mentionné au début, Williamson conteste l'épistémique
la fiabilité des simulations informatiques, leur échec en tant que substituts de la réalité et les fausses impressions de légitimité. On pourrait ne pas être d'accord avec lui sur la base que tous
formes de connaissances (c'est-à-dire scientifiques, techniques, sociales, culturelles, etc.) peuvent être présentées être défectueux, incomplet ou imparfait. Cela signifie-t-il que notre utilisation de celui-ci a des effets négatifs conséquences morales ? La réponse est un fort "ça dépend". Dans certains cas, il est clair
vrai que des connaissances incomplètes entraînent des conséquences moralement inacceptables. Le la fabrication de données par manque de connaissances sur la manière de monter correctement une expérience scientifique a des conséquences morales – et juridiques – directes pour les chercheurs. Et l'utilisation de données fabriquées par des politiciens pour informer le public est une affaire sérieuse.

L'intéressant travail de Williamson consiste à montrer comment l'épistémologie des simulations informatiques étend son propre domaine et touche à la valeur
et éthiques, permettent de légitimer certaines pratiques et comptes
pour des décisions de conception et des réglementations éclairées par des simulations moralement acceptables. Parce que son domaine de recherche privilégié est l'architecture des bâtiments, son des exemples sont liés à l'éthique environnementale. Un bon et dernier exemple est le développement de simulations de performances thermiques des bâtiments qui permettent, selon
Williamson, une croissance impressionnante de la vente de climatisation et de la commercialisation de conditions de confort, deux pratiques écologiquement irresponsables
(Shove 2004).

7.2.2 Brey

Une source connexe de préoccupations éthiques provient de la capacité de représentation des simulations informatiques ainsi que leur utilisation professionnelle. Brey a fait valoir qu'il
sont de bonnes raisons de croire que les représentations ne sont pas moralement neutres. Il utilise comme exemples de fausses déclarations et de représentations biaisées dans les simulations informatiques. De même, Brey soutient que la fabrication et l'utilisation de simulations informatiques impliquent des choix éthiques, et soulève donc des questions sur la pratique professionnelle.
(Brey 2008, 369).4 Examinons ces questions dans l'ordre.

Selon Brey, les simulations informatiques peuvent causer du tort, ainsi qu'induire en erreur chercheurs au point de causer du tort, si la simulation ne respecte pas certaines
normes d'exactitude de la représentation. Quand une telle situation se produit-elle ? Dans de nombreux Dans certains cas, cela fait partie de l'objectif et de la fonctionnalité d'une simulation informatique de représenter de manière réaliste des aspects du monde réel. Par exemple, dans une simulation biomécanique des systèmes d'implants osseux, il est primordial de représenter de manière réaliste et précise
l'os humain, sinon toute information obtenue à partir de la simulation pourrait être
inutile (Schneider et Resch 2014).

---

4 Qu'on se le dise, le travail de Brey porte aussi sur les enjeux éthiques de la Réalité Virtuelle telle qu'utilisée en vidéo des jeux ou des visualisations non scientifiques. Ici, je me concentre uniquement sur sa vision des simulations informatiques.

Une question centrale à discuter maintenant est la question de savoir ce que cela signifie pour une simulation informatique d'être réaliste ? – ou, si vous préférez, que signifie le respect de normes acceptables d'exactitude de représentation ? Nous avons appris plus tôt de Williamson que l'exactitude des résultats d'une simulation informatique est primordiale pour les évaluations morales.
Maintenant, la façon dont Williamson comprend la notion de précision équivaut à la précision numérique. C'est-à-dire que l'exactitude des résultats d'une simulation informatique est comprise comme le degré auquel ils correspondent aux valeurs mesurées et observées dans un système cible réel. Brey, au contraire, a une interprétation différente de la précision à l'esprit. Selon ses propres termes, "[les normes de précision] sont des normes qui définissent le degré de liberté qui existe dans la représentation d'un phénomène, et qui spécifient quels types de caractéristiques doivent être inclus dans une représentation pour qu'elle soit précise, quel niveau de détail est nécessaire, et quels types d'idéalisations sont permises » (Brey 2008, 369).
En d'autres termes, pour Brey, la précision est synonyme de « réalisme ». Un modèle précis est un modèle qui représente de manière réaliste le système cible, c'est-à-dire qui réduit le nombre d'idéalisations, d'abstractions et de fictionnalisation du modèle, tout en augmentant le niveau de détail.

Prenant la précision dans ce sens, nous devons maintenant nous demander quelles sortes de problèmes éthiques émergent avec les fausses déclarations ? Selon Brey, les fausses représentations de la réalité sont moralement problématiques dans la mesure où elles peuvent entraîner une certaine forme de préjudice. Plus ces dommages sont importants et plus ils risquent de se produire, plus grande est la responsabilité morale des chercheurs. À ce stade, il est intéressant de noter comment Brey combine la fausse représentation des simulations informatiques - en principe moralement neutre en soi - avec la responsabilité des concepteurs et des fabricants pour assurer l'exactitude des représentations (370). Par exemple, des imprécisions dans la reconstruction d'un os simulé pour un futur implant osseux peuvent entraîner de graves conséquences pour les concepteurs et les fabricants, mais aussi pour le patient. Un autre bon exemple présenté par Brey est le fonctionnement d'un moteur dans un logiciel éducatif, où de fausses déclarations peuvent amener les élèves à avoir de fausses croyances et, plus tard, causer un certain préjudice (370).

Les représentations biaisées, en revanche, constituent pour Brey une deuxième source de représentations moralement problématiques dans les simulations informatiques (Brey 1999). Une représentation est biaisée lorsqu'elle représente sélectivement des phénomènes – ou, pourrions-nous ajouter, déforme sélectivement des phénomènes. Par exemple, il a été détecté que dans de nombreuses simulations climatiques mondiales, la quantité de dioxyde de carbone absorbée par les plantes a été sous-estimée d'environ un sixième. Cela explique pourquoi l'accumulation de CO2 dans l'atmosphère n'est pas aussi rapide que les modèles climatiques l'ont prédit (Sun et al. 2014). Une telle représentation biaisée ignore de manière injustifiée la contribution au réchauffement climatique apportée par certains types d'industries et de pays. En général, le problème des représentations biaisées est qu'elles entraînent généralement des désavantages injustes pour des individus ou des groupes spécifiques, ou promeuvent de manière injustifiée certaines valeurs ou certains intérêts par rapport à d'autres (Brey 2008, 370).

De plus, les représentations peuvent être biaisées en contenant des hypothèses implicites sur le système en question. Nous avons déjà présenté un exemple d'hypothèses implicites qui conduisent à des représentations biaisées dans notre discussion sur les algorithmes dans la section 2.2.1.2 – bien que dans un contexte différent. L'exemple consistait à imaginer la spéci-

fication pour une simulation d'un système de vote. Pour que cette simulation soit réussie,
les modules statistiques sont mis en œuvre de manière à donner une répartition raisonnable de la
population électorale. Lors de la phase de spécification, les chercheurs décident
donner plus de pertinence statistique à des variables telles que le sexe, le genre et la santé
d'autres variables, comme l'éducation et le revenu. Si cette décision de conception n'est pas programmée
dans le module statistique de manière appropriée, alors la simulation ne reflétera jamais la
valeur de ces variables, entraînant ainsi un biais dans la représentation des électeurs.

L'exemple du mode de scrutin fait ressortir certaines inquiétudes partagées par de nombreux
auteurs, dont Brey, Williamson et, comme nous le verrons plus tard, Oren. C'est-à-dire le rôle des
concepteurs dans la spécification et la programmation des simulations informatiques. Brey soutient que
les concepteurs de simulations informatiques ont la responsabilité de
réfléchir sur leurs valeurs et leurs idéaux, les préjugés potentiels et les fausses déclarations inclus dans
leurs simulations, et de s'assurer qu'ils ne violent pas d'importants principes·éthiques
(Brey 1999). À cette fin, Brey, ainsi que Williamson et Oren, ont recours aux principes de pratique
professionnelle et aux codes d'éthique. Comme nous le verrons plus loin, Oren présente non seulement
ses préoccupations concernant l'utilisation de simulations informatiques, mais il propose également une
solution par la mise en œuvre d'un code d'éthique strict exclusivement
pour les chercheurs travaillant avec des simulations informatiques.

7.2.3 Oren

Tuncer Oren a une description plus élaborée et approfondie de l'éthique dans les simulations informatiques.
Pour lui, le problème fondamental est de savoir si l'utilisation de simulations informatiques entraîne des
implications graves pour les êtres humains. Sa principale préoccupation vient de
le fait que des simulations informatiques sont utilisées pour soutenir les politiques et décisions cruciales
qui pourraient modifier notre vie actuelle et contraindre notre avenir. Dans les systèmes de gestion des déchets
de combustible nucléaire, par exemple, des simulations informatiques sont utilisées pour étudier le long terme
comportement, l'impact environnemental et social, ainsi que les moyens de contenir les déchets de
combustible nucléaire. De même, des simulations informatiques de fuites de déchets nucléaires dans le
Les rivières terrestres et sous-marines sont cruciales pour étayer les décisions sur la question de savoir si de telles
les déchets nucléaires devraient être enterrés au lieu de construire une installation spéciale pour le stockage.

Oren comprend bien sûr que les questions éthiques concernant l'utilisation des résultats des simulations
informatiques ont une relation directe avec la fiabilité des simulations informatiques. À cet égard, il partage
avec Williamson et Brey l'idée que l'épistémologie des simulations informatiques est à la base des
préoccupations morales. À propos de ça,
il dit : « l'existence de plusieurs techniques et outils de validation, de vérification et d'accréditation atteste
également de l'importance des enjeux de la simulation »
(Oran 2000). Comme pour Williamson et Brey, ce qui est important pour Oren, c'est de créer la base de la
crédibilité des simulations informatiques de manière à ce que les décideurs et les responsables politiques
puissent considérer les simulations comme un outil fiable. Cependant, je crois
qu'il va plus loin que Williamson et Brey dans son analyse lorsqu'il affirme que
une étude appropriée sur l'éthique des simulations informatiques nécessite un code bien défini
d'éthique qui complète les méthodes de vérification, de validation et d'accréditation. À

Dans son esprit, un code d'éthique permettrait d'établir plus facilement la crédibilité des chercheurs qui conçoivent et programment la simulation informatique - à la fois en tant qu'individus et en tant que groupes – et ne dépendent donc pas entièrement de la simulation informatique elle-même.[5]

Une question intéressante émergeant de la position d'Oren est qu'elle dépend fortement de l'honnêteté, du dévouement et de la bonne conduite des chercheurs . une hypothèse juste. La pratique scientifique et technique s'accompagne de l'adhésion explicite aux règles de bonne pratique scientifique, qui comprend le maintien de normes professionnelles, la documentation des résultats, la remise en question rigoureuse de toutes les découvertes et attribuant honnêtement toute contribution par et aux partenaires, concurrents et prédécesseurs (Deutsche Forschungsgemeinschaft 2013). Recherche nationale et internationale fondations, universités et fonds privés, tous comprennent que l'inconduite scientifique est une offense grave à la communauté scientifique et technique, ainsi que des poses menace majeure pour le prestige de toute institution scientifique. La recherche allemande Foundation (DFG) définit l'inconduite scientifique comme « l'intentionnelle et déclaration par négligence grave de mensonges dans un contexte scientifique, violation des droits de propriété intellectuelle ou entrave aux travaux de recherche d'autrui » (Deutsche Forschungsgemeinschaft 2013). Toute faute avérée est sévèrement sanctionnée. je discutera des idées d'Oren sur la pratique professionnelle et le code d'éthique dans les sections suivantes. Avant d'entrer dans les détails, permettez-moi de vous présenter une petite préoccupation concernant sa position.

Le problème qui, selon moi, imprègne la position d'Oren est que, si les chercheurs adoptent et respectent les règles d'une bonne pratique scientifique, alors il n'y a pas d'éthique particulière problèmes pour les simulations informatiques. La raison en est que, dans le cadre d'Oren, les questions éthiques sont adaptées aux chercheurs et peu de responsabilité est mise en place. sur la simulation informatique elle-même. Par cela, je ne veux bien sûr pas dire que le la simulation informatique est moralement responsable, mais plutôt que les préoccupations éthiques découlent de la simulation elle-même, comme dans le cas de Brey, où la fausse déclaration est une source d'inquiétude.

Oren a cependant raison de souligner à quel point les chercheurs pourraient bénéficier d'un code d'éthique. Les organismes professionnels ont une longue tradition d'attribution à ces codes pour les avantages évidents qu'ils apportent à leur pratique. Permettez-moi maintenant de discuter plus en détail la pratique professionnelle et le véritable code de déontologie des chercheurs travailler avec des simulations informatiques.

7.3 Pratique professionnelle et code de déontologie

Les codes de déontologie trouvent leur origine dans des sociétés professionnelles particulières, comme l'American Society of Civil Engineers, l'American Society of Mechanical Engineers, l'In

[5] Rappelons que pour Williamson, « l'expert » n'était pas nécessairement le chercheur qui conçoit et programmer une simulation informatique, mais plutôt le spécialiste du sujet qui est simulé.

[6] Andreas Tolk a un grand nombre de travaux sur l'éthique des simulations informatiques qui suivent un ligne de recherche similaire à celle d'Oren. Voir (Tolk 2017b, 2017a).

stitute of Electronic and Electrical Engineers, et autres.7 Comme le suggèrent leurs titres, ces sociétés reflètent la profession de leurs membres – dans ces cas, tous les ingénieurs – et leur spécialisation.

Or, il est typique de ces sociétés, groupes et institutions de souscrire à un code de déontologie, c'est-à-dire un ensemble d'aspirations, de règlements et de lignes directrices qui représente les valeurs de la profession à laquelle il s'applique.8 De tels codes de déontologie sont généralement comprises comme des normes normatives sur la façon dont les ingénieurs et les scientifiques doivent se conduire dans des circonstances spécifiques afin de continuer à faire partie de la profession. En ce sens, un code d'éthique contient des normes morales de ce qui est compris, par une communauté donnée, comme un comportement professionnel bon et acceptable.

Si l'éthique dans les simulations informatiques dépend d'un code d'éthique, comme le suggère Oren, alors il sera primordial pour nous de pouvoir comprendre les fonctions d'un code d'éthique, son contenu et l'applicabilité dans les études sur les simulations informatiques.

Les codes de déontologie peuvent être considérés comme la marque d'une profession et, à ce titre, les dispositions du code de déontologie sont traitées comme des lignes directrices qui doivent être suivies par les membres d'une société. L'une des principales fonctions du code de déontologie est de contenir une disposition selon laquelle un membre ne doit rien faire qui puisse jeter le discrédit sur la profession ou déshonorer le professionnel. Il s'agit d'un élément central sur lequel repose le code de déontologie d'Oren, car il exige que la conception et la programmation des chercheurs respectent les règles de bonnes pratiques professionnelles.

Une question importante ici est de savoir si un code d'éthique présuppose une obligation morale – et même légale – pour ceux qui y adhèrent. Certains philosophes considèrent qu'un code de déontologie n'est rien de plus qu'un accord entre les membres d'une profession donnée pour s'engager à respecter un ensemble commun de normes. Ces normes servent à établir les bases et les principes d'une profession partagée. Heinz Luegen biehl, un philosophe qui s'intéresse aux codes et à la formation des ingénieurs considère tout code d'éthique comme « un ensemble de règles éthiques qui doivent régir les ingénieurs dans leur vie professionnelle » (Luegenbiehl 1991). Ainsi compris, un code d'éthique ne sert qu'à conseiller les ingénieurs sur la façon dont ils doivent agir dans des circonstances données, mais il n'impose aucune obligation morale ou légale, et on ne s'attend pas non plus à ce que les ingénieurs se conforment strictement à un code d'éthique. En fait, selon cette interprétation, un code d'éthique n'est rien de plus qu'un engagement ou un devoir envers ses collègues professionnels.

Pour d'autres philosophes, un code de déontologie est un ensemble d'énoncés incarnant la sagesse collective des membres d'une profession donnée qui tend à tomber en désuétude.
Les ingénieurs en exercice consultent rarement ces codes, et encore moins les suivent. Il y a plusieurs raisons à cela. L'un des principaux est que les chercheurs constatent que certains des codes d'éthique conçus pour leurs disciplines contiennent des principes et des idéaux qui sont en conflit. Pourtant, d'autres philosophes trouvent le code d'éthique coercitif dans son intention et limitent ainsi les libertés du chercheur en tant que professionnels. En fait, une base pour rejeter tous les codes de

---

7 Une collection de plus de 50 codes d'éthique officiels publiés par 45 associations dans les domaines des affaires, de la santé et du droit se trouve dans (Gorlin 1994).

8 Les codes se présentent sous de nombreuses formes différentes. Ils peuvent être formels (écrits) ou informels (oraux). Ils portent une variété de noms, chacun avec des objectifs éducatifs et réglementaires légèrement différents. Les formes les plus courantes sont les codes d'éthique, les codes de conduite professionnelle et les codes de pratique. Ici, je vais discuter d'une forme générale de codes. Pour plus de détails, voir (Pritchard 1998).

l'éthique procède de l'affirmation qu'elles supposent une remise en cause de l'autonomie des agents moraux. La question « pourquoi avons-nous besoin d'un code d'éthique ? » est répondu par dire que c'est peut-être parce que les chercheurs ne sont pas des agents moraux.

Au-delà des complexités philosophiques derrière les codes d'éthique et des raisons pour lesquelles les chercheurs doivent les adopter ou carrément les rejeter, il y a plusieurs objectifs que tout code d'éthique s'attend à atteindre. Un objectif important est d'être inspirant, c'est-à-dire qu'il est censé inspirer les membres à être plus « éthiques » dans leur travail. conduire. Une objection évidente à cet objectif est qu'il présuppose que les membres d'un communauté professionnelle sont contraires à l'éthique, amoraux ou submoraux, et il est donc nécessaire pour les exhorter – même par des moyens inspirés – à être moraux. Un autre objectif important des codes de déontologie est qu'ils servent à sensibiliser les membres d'un ordre professionnel à faire part de leurs inquiétudes sur des questions morales liées à leur propre discipline.

Un code pourrait également offrir des conseils en cas de perplexité morale sur ce qu'il faut faire. Par exemple, quand est-il juste pour un membre de dénoncer un collègue pour un acte répréhensible ? Naturellement, un tel cas est différent d'un chercheur violant une règle morale ou même une loi. Signalement d'un collègue pour avoir vendu des médicaments sans l'autorisation les licences et les habitations ne devraient pas constituer un problème pour un code de déontologie.

Dans la plupart des cas, un code d'éthique fonctionne comme un code disciplinaire pour faire respecter certaines règles d'une profession sur ses membres afin de défendre l'intégrité de la profession et protéger ses normes professionnelles. Il est douteux, cependant, qu'aucune mesure disciplinaire l'action peut être gérée et appliquée par un code d'éthique, mais c'est, pour la plupart, un bonne source lorsqu'elle est utilisée comme critère pour déterminer la faute professionnelle.[9]

7.3.1 Un code de déontologie pour les chercheurs en simulations informatiques

Si à ce stade nous sommes convaincus de la valeur de l'éthique pour concevoir, programmer, et éventuellement en utilisant des simulations informatiques, comme le propose Oren, alors la question suivante est de savoir comment développer un code d'éthique professionnelle pour les chercheurs travaillant avec des simulations informatiques. Il sera tout aussi important de rendre explicites les responsabilités qui se trouvent au-dessus et au-delà de ce code d'éthique, que je prends également d'Oren.

Ce qui suit est le code de déontologie des simulationnistes de la Society for Modeling & Simulation International (SCS). Le SCS se consacre à faire progresser l'utilisation et compréhension de la modélisation et des simulations informatiques dans le but de résoudre des problèmes du monde réel. Il est intéressant de noter que, dans leur déclaration, la SCS déclare son engagement non seulement à l'avancement des simulations informatiques dans domaines de la science et de l'ingénierie, mais aussi dans le domaine des arts. Un autre objectif important de la SCS est de promouvoir la communication et la collaboration entre les professionnels du domaine des simulations informatiques. L'une des principales raisons d'avoir un code d'éthique est précisément de remplir ces objectifs constitutifs en tant que société.

---

[9] Pour en savoir plus sur l'éthique de la pratique professionnelle et les codes de déontologie, voir (Harris Jr et al. 2013).

Le code de déontologie des chercheurs travaillant dans les simulations informatiques, tel que publié sur le site Web du SCS (http://scs.org) est le suivant :

Les simulationnistes sont des professionnels impliqués dans un ou plusieurs des domaines suivants domaines :

• Activités de modélisation et de simulation. • Fournir des produits de modélisation et de simulation. • Fournir des services de modélisation et de simulation.

1. Développement personnel et profession

En tant que simulationniste, je vais :

1.1. Acquérir et maintenir une compétence et une attitude professionnelles.
1.2. Traiter équitablement les employés, les clients, les utilisateurs, les collègues et les employeurs.
1.3. Encourager et soutenir les nouveaux entrants dans la profession.
1.4. Soutenir les collègues praticiens et les membres d'autres professions qui engagé dans la modélisation et la simulation.
1.5. Aider les collègues à obtenir des résultats fiables.
1.6. Promouvoir l'utilisation fiable et crédible de la modélisation et de la simulation.
1.7. Promouvoir le métier de la modélisation et de la simulation ; par exemple, faire progresser la connaissance et l'appréciation du public de la modélisation et de la simulation et clarifier et contrer les déclarations fausses ou trompeuses.

2. Compétence professionnelle

En tant que simulationniste, je vais :

2.1. Assurer la qualité du produit et/ou du service en utilisant une méthodologie appropriée gies et technologies.
2.2. Rechercher, utiliser et fournir une critique professionnelle critique.
2.3. Recommander et stipuler des objectifs appropriés et réalisables pour tout projet.
2.4. Documenter les études de simulation et/ou les systèmes de manière compréhensible et précise aux personnes autorisées.
2.5. Fournir une divulgation complète des hypothèses de conception du système et des limites et problèmes connus aux parties autorisées.
2.6. Être explicite et sans équivoque sur les conditions d'applicabilité des modèles spécifiques et des résultats de simulation associés.
2.7. Mise en garde contre l'acceptation des résultats de modélisation et de simulation lorsqu'il n'y a pas suffisamment de preuves d'une validation et d'une vérification approfondies.
2.8. Assurer des interprétations et des évaluations approfondies et impartiales des résultats des études de modélisation et de simulation.

3. Fiabilité

En tant que simulationniste, je vais :

3.1. Soyez honnête à propos de toutes les circonstances qui pourraient conduire à un conflit d'intérêts HNE.

3.2. Honorer les contrats, les accords et les responsabilités et obligations assignées.

3.3. Aider à développer un environnement organisationnel qui soutient le comportement éthique.

3.4. Soutenir les études qui ne nuiront pas aux êtres humains (générations actuelles et futures) ni à l'environnement.

4. Droits de propriété et crédit dû

En tant que simulationniste, je vais :

4.1. Reconnaître pleinement les contributions des autres.
4.2. Attribuez un crédit approprié à la propriété intellectuelle.
4.3. Respectez les droits de propriété, y compris les droits d'auteur et les brevets.
4.4. Respecter les droits à la vie privée des individus et des organisations ainsi que confi dentialité des données et connaissances pertinentes.

5. Conformité au Code

En tant que simulationniste, je vais :

5.1. Adhérez à ce code et encouragez les autres simulationnistes à y adhérer.
5.2. Traitez les violations de ce code comme incompatibles avec le fait d'être un simulationniste.
5.3. Demander conseil à des collègues professionnels lorsqu'ils sont confrontés à un problème d'éthique. dilemme dans les activités de modélisation et de simulation.
5.4. Aviser toute société professionnelle qui soutient ce code des pratiques souhaitables mises à jour.

La justification de ce code d'éthique est fournie par Oren dans (Oren 2002). Selon l'auteur, il y a au moins deux raisons d'adopter un code d'éthique pour simulationnistes.10 Premièrement, parce qu'il existe quelques sociétés de simulation émergentes qui exigeront de leurs membres qu'ils adoptent un code d'éthique. De cette façon, et suivant la principes d'un code d'éthique, les membres peuvent montrer qu'ils acceptent leur responsabilité et leur imputabilité dans le développement, la programmation et l'utilisation de simulations informatiques. Deuxièmement, parce que les simulations informatiques sont une forme d'expérimentation de modèles, et peuvent donc affecter les humains et l'environnement de différentes manières. Dans

[dix] Oren parle en fait de trois raisons, la troisième étant la célébration du 50 [ème] anniversaire de la fondation de la Society for Modeling and Simulation International (SCS).

Dans ce contexte, les éthiciens doivent fournir une analyse des bonnes et des mauvaises actions, des bonnes et des mauvaises conséquences, et des utilisations justes et injustes des simulations informatiques.

Alors que les premières raisons données par Oren visent à établir des critères de responsabilités professionnelles, la seconde fournit un contexte pour traiter des conséquences de l'utilisation de simulations informatiques.

7.3.2 Responsabilités professionnelles

Le code d'éthique présenté ci-dessus définit ce qui constitue un bon comportement pour les scientifiques et ingénieurs professionnels travaillant avec des simulations informatiques. À cet égard, il concentre son attention sur la conception, la programmation et l'utilisation de simulations informatiques. Dans ce qui suit, je présente la discussion d'Oren sur le type de responsabilité attribuée à la pratique professionnelle des simulations informatiques. Suivant et élargissant (Oren 2000), les scientifiques et les ingénieurs ont un devoir envers le grand public, leurs clients, les employeurs, leurs collègues, leur profession et eux-mêmes. La liste de responsabilités suivante est tirée de (169).

- Responsabilité vis-à-vis du public :
  - Un chercheur doit agir en cohérence avec l'intérêt public.
  - Un chercheur doit promouvoir l'utilisation de la simulation pour améliorer l'existence humaine.
- Responsabilité vis-à-vis du client (un client de simulation est une personne, une entreprise ou un agent qui achète, loue ou loue un produit de simulation, un service de simulation ou un conseil basé sur la simulation) :
  - Un chercheur doit agir d'une manière qui est dans le meilleur intérêt du client. Cette responsabilité doit être conforme à l'intérêt public.
  - Un chercheur doit fournir/maintenir un produit et/ou des services de simulation pour résoudre le problème de la manière la plus fiable. Cela comprend l'utilité, c'est-à-dire l'aptitude à l'usage ainsi que le respect des normes professionnelles les plus élevées.
- Responsabilité vis-à-vis de l'employeur :
  - Un chercheur doit agir d'une manière qui est dans le meilleur intérêt de l'employeur, pourvu que les activités soient en alliance avec le meilleur intérêt du public et du client.
  - Un chercheur doit respecter les droits de propriété intellectuelle de ses employeurs actuels et/ou passés.
- Responsabilité vis-à-vis des collègues :
  - Un chercheur doit être juste et solidaire envers ses collègues.
- Responsabilité vis-à-vis de la profession :

– Un chercheur doit faire progresser l'intégrité et la réputation de la simulation profession conforme à l'intérêt général.
- Un chercheur doit appliquer la technologie de simulation de la manière la plus appropriée et ne doit pas forcer la simulation ou tout type de celle-ci comme un lit de Procuste.
– Un chercheur doit partager son expérience et ses connaissances pour faire progresser le métier de la simulation ; et ce, de concert avec les intérêts de son employeur et de son client.

• Responsabilité vis-à-vis de soi :

- Un chercheur doit continuer à améliorer ses capacités à avoir une vision et des connaissances appropriées pour concevoir des problèmes dans une perspective large et les appliquer pour la solution de problèmes de simulation.

Certes, il n'y a guère plus à ajouter à ce code d'éthique. Sous peine de répéter certains des problèmes évoqués ici, ce code de déontologie pourrait être étendu sur quelques points supplémentaires. En ce qui concerne la responsabilité envers un client, les chercheurs doivent être très attentifs à toutes les règles de confidentialité qui relient leur travail aux intérêts du client, ainsi qu'à sa propriété sur les conceptions, les résultats, etc. Cela pourrait être le cas de simulations sensibles comme en médecine. Dans de tels cas, le client fait confiance au chercheur pour garder les informations fournies en secret, un vote de confiance qui ne doit pas être brisé.

Dans le cas de la responsabilité d'un chercheur à l'égard du public, on pourrait ajouter que toute divulgation publique doit être informée de la manière la plus objective et la plus impartiale. Comme nous l'avons vu dans la section 5.2.1, la visualisation d'une simulation n'est pas directement claire pour le non-expert (par exemple, politicien, communicateur public), et donc la communication de la visualisation ne doit pas être médiatisée par des interprétations subjectives - dans la mesure où car c'est effectivement possible. Bien que cette responsabilité s'étende également aux clients et aux employeurs, c'est peut-être le public l'agent le plus sensible à protéger des visualisations biaisées.

Enfin, et comme règle générale à observer, les chercheurs doivent suivre des codes généraux de bonne conduite professionnelle (par exemple, les règles de bonne pratique scientifique fournies par la Fondation allemande pour la recherche (Deutsche Forschungsgemeinschaft 2013)). À cet égard, les scientifiques et ingénieurs travaillant avec des simulations informatiques n'ont pas un statut particulier en tant que chercheurs, mais sont plutôt soumis à des responsabilités spécifiques adaptées à leur profession.

7.4 Remarques finales

Les études sur l'éthique des simulations informatiques n'en sont qu'à leurs balbutiements. Dans ce chapitre, je me suis concentré sur la présentation des principales idées concernant l'éthique des simulations informatiques disponibles dans la littérature actuelle. À cet égard, nous avons vu trois points de vue différents. Williamson, qui met l'accent sur l'épistémologie et la méthodologie des simulations informatiques, précisant que leur fiabilité est primordiale pour

l'évaluation éthique. Brey, qui attire correctement l'attention sur les dangers des représentations erronées et des simulations biaisées. Enfin, Oren qui discute longuement de la forme que devrait prendre la responsabilité professionnelle dans le cadre des simulations informatiques.

Dans un contexte où les simulations informatiques sont omniprésentes dans les domaines scientifiques et d'ingénierie, où notre connaissance et notre compréhension des rouages le monde dépend d'eux, il est frappant qu'on ait si peu parlé de la morale conséquences qui découlent de la conception, de la programmation et de l'utilisation de simulations informatiques. J'espère que la discussion que nous venons d'avoir servira de point de départ à de nombreuses prochaines discussions fructueuses.

Les références

Béranger, Jérôme. 2016. Big data et éthique : la datasphère médicale. Elsevier.

Brey, Philippe. 1999. "L'éthique de la représentation et de l'action dans la réalité virtuelle." Éthique et Technologies de l'information 1 (1): 5–14.

————. 2008. "Réalité virtuelle et simulation informatique." Dans The Handbook of Information and Computer Ethics, édité par Kenneth Einar Himma et Herman T. Tavani, 361–384.

Briggle, Adam et Carl Mitcham. 2012. Éthique et science : une introduction. Came pont University Press.

Deutsche Forschungsgemeinschaft, éd. 2013. Sicherung guter wissenschaftlicher Praxis – Sauvegarder les bonnes pratiques scientifiques. WILY.

Gorlin, Rena A. 1994. Codes de responsabilité professionnelle. Livres BNA.

Harris Jr, Charles E, Michael S Pritchard, Michael J Rabins, Ray James et Elaine Engelhardt. 2013. Éthique de l'ingénieur : Concepts et cas. Cengage Apprentissage.

Jamieson, Dale. 2008. Éthique et environnement : une introduction. Université de Cambridge Versité Presse.

Johnson, Deborah G. 1985. «Éthique informatique». Falaises d'Englewood (NJ).

Lügenbiehl, Heinz. 1991. "Codes d'éthique et éducation morale des ingénieurs." Dans Ethical Issues in Engineering, édité par Deborah Johnson, 137–138. 4. Prentice Hall.

Lupton, Deborah. 2014. « La marchandisation de l'opinion des patients : le patient numérique faites l'expérience de l'économie à l'ère des mégadonnées. Sociologie de la santé et de la maladie 36 (6): 856–869. ISSN : 1467-9566. doi : 10.1111/1467-9566.12109. http://dx.doi.org/10.1111/1467-9566.12109.

Marr, Bernard. 2016. Big Data en pratique : Comment 45 entreprises prospères ont utilisé Big Data Analytics pour fournir des résultats extraordinaires. John Wiley et fils.

Mayer-Schonberger, Viktor et Kenneth Cukier. 2013. Big Data : une révolution Cela transformera notre façon de vivre, de travailler et de penser. Cour Houghton Mifflin Har, mars.

Mc Leod, John. 1986. "Mais, Monsieur le Président - est-ce éthique?" Actes du 1986 Conférence de simulation d'hiver.

Meehl, Gérard A., Thomas F. Stocker, William D. Collins, AT Friedlingstein, T. Gaye Amadou, M. Gregory Jonathan, Akio Kitoh, et al. 2007. "Climat Changement 2007 : La base des sciences physiques. Contribution du groupe de travail I au quatrième rapport d'évaluation du Groupe d'experts intergouvernemental sur le climat Changement." Type. Projections climatiques mondiales, éditées par S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, KB Averyt, M. Tignor et HL Miller, 747–845. Presse universitaire de Cambridge.

Mittelstadt, Brent Daniel et Luciano Floridi. 2016. « L'éthique du Big Data : enjeux actuels et prévisibles dans les contextes biomédicaux. Sciences et ingénierie Éthique 22 (2): 303–341.

Moor, James H. 1985. "Qu'est-ce que l'éthique informatique?" Métaphilosophie 16, no. 4 (octobre) : 266–275.

Oren, Tuncer I. 2000. "Responsabilité, éthique et simulation." Transactions 17 (4).

————. 2002. "Justification d'un code d'éthique professionnelle pour les simulationnistes." Dans Conférence d'été sur la simulation informatique, 428–433. Société pour l'informatique Simulation Internationale ; 1998.

Pritchard, J. 1998. « Codes de déontologie ». Dans Encyclopedia of Applied Ethics (Second Edition), deuxième édition, éditée par Ruth Chadwick, 494–499. San Diego : Aca demic Press.

Schneider, Ralf et Michael M Resch. 2014. "Calcul de l'efficacité discrète Raideur de l'os spongieux par simulations mécaniques directes. In Computational Surgery and Dual Training, édité par Garbey M., Bass B., Berceli S., Collet C. et Cerveri P., 351–361. Springer.

Poussez, Elizabeth. 2004. Confort, propreté et commodité : L'organisation sociale de la normalité. Editeurs Berg.

Spier, Raymond E. 2012. "Encyclopédie d'éthique appliquée." Type. Science et Engineering Ethics, Overview, édité par Ruth Chadwick, 14–31. Elsevier.

Sun, Ying, Lianhong Gu, Robert E Dickinson, Richard J Norby, Stephen G Palardy et Forrest M Hoffman. 2014. "Impact de la diffusion du mésophylle sur l'estimation de la fertilisation mondiale par le CO2 des terres." Actes de l'Académie nationale des Sciences 111 (44): 15774–15779.

Tolk, Andreas. 2017a. "Code d'éthique." In Le métier de la modélisation et de la simulation : discipline, éthique, éducation, vocation, sociétés et économie, édité par Andreas Tolk et Tuncer Oren, 35–51. Wiley & Fils.

————. 2017b. "Les sociétés de modélisation et de simulation façonnent la profession." Dans le métier de la modélisation et de la simulation : discipline, éthique, formation; vocation, sociétés et économie, édité par Andreas Tolk et Tuncer Oren, 131-150. Wiley & Fils.

Williamson, TJ. 2010. « Prédire la performance des bâtiments : l'éthique de l'informatique simulation." Recherche et information sur le bâtiment 38 (4): 401–410.

Zwitter, Andrej. 2014. "Éthique des mégadonnées". Big Data & Société 1 (2) : 2053951714559253. ISSN : 2053-9517. doi:10.1177/2053951714559253.