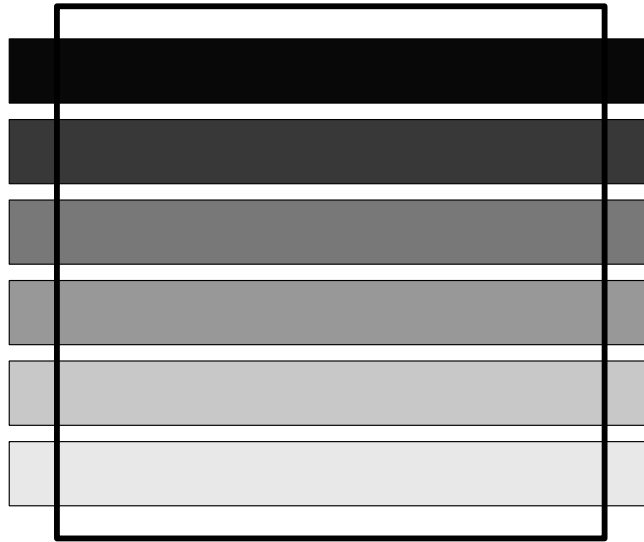


LES CAHIERS DU LANCI



*VERS UNE ÉPISTÉMOLOGIE DE LA SIMULATION :
LA VIE ARTIFICIELLE COMME MÉTHODE PHILOSOPHIQUE*

Jean-Frédéric de Pasquale et Pierre Poirier

N° 2004-01

UQÀM
Université du Québec à Montréal

Le Laboratoire d'ANalyse Cognitive de l'Information (LANCI) effectue des recherches sur le traitement cognitif de l'information. La recherche fondamentale porte sur les multiples conceptions de l'information. Elle s'intéresse plus particulièrement aux modèles cognitifs de la classification et de la catégorisation, tant dans une perspective symbolique que connexionniste. La recherche appliquée explore les technologies informatiques qui manipulent l'information. Le territoire privilégié est celui du texte. La recherche est de nature interdisciplinaire. Elle en appelle à la philosophie, à l'informatique, à la linguistique et à la psychologie.

Publication du Laboratoire d'ANalyse Cognitive de l'Information
Directeur : Jean-Guy Meunier
Université du Québec à Montréal

Volume 3, Numéro 2004-01 – Décembre 2004

Document disponible en ligne à l'adresse suivante : www.lanci.uqam.ca

Tirage : 15 exemplaires

Aucune partie de cette publication ne peut être conservée dans un système de recherche documentaire, traduite ou reproduite sous quelque forme que ce soit - imprimé, procédé photomécanique, microfilm, microfiche ou tout autre moyen - sans la permission écrite de l'éditeur. Tous droits réservés pour tous pays. / All rights reserved. No part of this publication covered by the copyrights hereon may be reproduced or used in any form or by any means - graphic, electronic or mechanical - without the prior written permission of the publisher.

Dépôt légal – Bibliothèque Nationale du Canada
Dépôt légal – Bibliothèque Nationale du Québec
ISBN : 2-922916-05-7

© 2004 Jean-Frédéric de Pasquale et Pierre Poirier

Mise en page : Gaétan Piché

VERS UNE ÉPISTÉMOLOGIE DE LA SIMULATION : LA VIE ARTIFICIELLE COMME MÉTHODE PHILOSOPHIQUE

Jean-Frédéric De Pasquale et Pierre Poirier
Université du Québec à Montréal

1. Introduction

La simulation informatique représente désormais un outil essentiel de la science et de l'ingénierie, et elle est en voie de devenir un élément commun de la vie courante ; qu'on pense par exemple aux « nuages modélisés » qu'on nous présente au bulletin de météo chaque soir. On ne saurait trop exagérer l'importance de la simulation en science, laquelle constitue selon nous une révolution méthodologique à placer avec celles qui ont amené en science l'observation (laquelle date au moins de l'antiquité grecque) et l'expérimentation, (autour des 16^e et 17^e siècle) – il ne fait pas de doute selon nous que la simulation est la moins importante de ces trois révolutions, mais nous croyons néanmoins qu'elle est de même nature que celles-ci. Ayant construit un modèle mathématique adéquat d'un ou d'une classe de phénomènes¹, tel que décrit par une théorie, il est possible, au moyen d'une simulation informatique du phénomène modélisé, (1) d'étudier l'évolution du phénomène dans des conditions inobservées, voire inobservables (pour des raisons pratiques ou théoriques), et (2) de tester systématiquement, mais virtuellement, toutes les hypothèses impliquées par la théorie².

Les progrès en philosophie découlent souvent de révolutions méthodologiques, qu'on pense au rôle de la méthode dans l'avènement de la philosophie moderne, à celui de la logique et de l'analyse du langage en philosophie dite analytique ou de la réduction en phénoménologie. Récemment, certains philosophes ont proposé d'ajouter la modélisation mathématique et la simulation informatique à l'arsenal des méthodes dont dispose le philosophe pour accomplir sa tâche. Ainsi, Peter Danielson, dans son ouvrage *Artificial Morality* (Danielson 1992, voir

1. Nous ne pouvons pas discuter ici des conditions d'une modélisation mathématique adéquate d'un phénomène. Voir Ziegler et al. (2000).

2. Plus précisément, toutes les implications de la théorie telle que modélisée. Il ne faut pas confondre un test virtuel avec un test empirique, ce dernier étant la source ultime de toute justification en sciences. Cependant, on peut imaginer que, pour tester virtuellement une théorie donnée, nous définissions un modèle de l'environnement pertinent qui soit empiriquement validé et aussi complet que nos connaissances actuelles le permettent. Dans ce cas, on peut imaginer que le test virtuel a presque autant de valeur, sur le plan épistémique, qu'un test empirique.

également *Modeling Rationality, Morality, and Evolution*, 1998) utilise des agents (programmés en Prolog) pour montrer qu'il peut être rationnel d'être moral, c'est-à-dire de coopérer, et ce même dans un contexte darwinien où, de prime abord, il semblerait beaucoup plus avantageux de ne pas coopérer, c'est-à-dire de prendre ce que donnent les coopérants sans réciproquer le moment venu. Pour Danielson, il peut être avantageux, ne serait-ce que pour ouvrir la réflexion, d'étudier des agents plus simples, réalisant des tâches simples mais pertinentes (croit-il), en l'occurrence des jeux stratégiques comme le dilemme du prisonnier. Pour des raisons semblables, Paul Thagard (1992) utilise des réseaux compétitifs (*interactive-competitive networks*, Rumelhart et McClelland 1986) pour étudier la découverte et la justification des théories scientifiques de même que leur usage en contexte de construction d'explications scientifiques. Pour ces philosophes, et bien d'autres (Sloman 1978, Dennett, 1979, Bedau 1996, Bynum et Moor 1998, Grim, Mar et St-Denis 1998) la précision des modèles informatiques ajoutera une mesure de rigueur aux théories philosophiques et contraindra les expériences de pensée de manière à les rendre moins dépendantes de l'intuition.

Nous ne pouvons pas ici discuter de chacun de ces philosophes ni de toutes les questions soulevées par l'usage de la simulation en philosophie. Nous proposons plutôt quelque chose qui, du moins en première analyse, peut s'avérer efficace pour étudier cet usage : nous verrons comment le philosophe Daniel Dennett a utilisé les simulations dans sa pratique de la philosophie. Nous verrons alors que Dennett utilise la simulation à diverses fins, que nous chercherons à comprendre, cataloguer et enfin évaluer. Nous verrons également que Dennett conçoit un lien étroit entre la nouvelle pratique de la simulation et la pratique traditionnelle en philosophie de l'expérience de pensée. Il marie en fait expérience de pensée et simulation : il utilise des simulations comme support au travail cognitif d'imagination requis pour réfléchir à une situation complexe lors d'une expérience de pensée. Bref, il utilise la simulation entre autres comme prothèse cognitive lors d'expériences de pensée.

Il convient de souligner, avant de passer au développement, que Daniel Dennett est ici un choix stratégique à plusieurs égards. Tout d'abord, il convient de rappeler que Dennett est un des premiers et plus importants philosophes de l'esprit et des sciences cognitives. Son ouvrage de 1969, *Content and Consciousness*, est peut-être le premier ouvrage de philosophie à appliquer systématiquement la perspective naturaliste en philosophie de l'esprit. Comme le remarque Dennett lui-même dans une rétrospective intellectuelle de sa carrière (Dennett 1996), toutes les dimensions importantes de sa philosophie sont présentes dans *Content and Consciousness*, et seront développées à travers deux cycles d'études sur le contenu et la conscience, le premier se

terminant avec la parution en 1979 de *Brainstorms*, le second étudiant d'abord l'intentionnalité, et culminant en 1987 avec la publication de *La Stratégie de l'interprète*, puis étudiant la conscience, avec la parution en 1992 de *La conscience expliquée*. Depuis, Dennett cherche à intégrer le contenu et la conscience dans le cadre d'une philosophie darwinienne, ce que l'on pourrait concevoir comme le troisième cycle de réflexion sur la conscience et le contenu, débutant en 1995 avec *Darwin's Dangerous Idea* et se poursuivant récemment avec *Freedom Evolves* (Dennett 2003).

En second lieu, Dennett a lui-même réfléchi explicitement à la question de l'usage des simulations en philosophie et il a utilisé (sans toutefois, pour autant que nous le sachions, en avoir conçu) des simulations. Sur ce plan, Dennett est un ami de longue date de l'intelligence artificielle et sa collaboration avec Rodney Brooks (directeur du laboratoire d'intelligence artificielle du MIT) au projet COG n'est que la plus récente manifestation de cet intérêt de longue date. Ses réflexions sur le sujet ne sont pas celles d'un novice qui s'attarderait momentanément à la question.

2. Expérience de pensée et simulation

Avant de décrire et d'analyser les usages de la simulation chez Dennett, il nous faut préciser deux concepts dont nous parlerons tout au long du texte, soit les notions de simulation et d'expérience de pensée. Nous donnerons de chaque notion une caractérisation générale accompagnée d'exemples, sans prétendre en fournir une définition stricte, du moins dans un premier temps, puisque nous reviendrons sur la définition à la fin de l'article. Commençons par la notion d'expérience de pensée.

2.1 Qu'est-ce qu'une expérience de pensée ?

Comme son nom l'indique, une expérience de pensée est une expérience menée avec la pensée. Plus précisément, réaliser une expérience de pensée, c'est, selon Gendler (2002), « raisonner au sujet d'un scénario imaginaire dans l'intention de confirmer ou d'infirmer une hypothèse ou une théorie. » (p. 388). On retrouve surtout des expériences de pensée en physique et en philosophie, domaines où il faut parfois réfléchir à des hypothèses ou théories abstraites, qui demandent une évaluation, mais pour lesquelles on ne peut, pour des raisons théoriques ou pratiques, concevoir une véritable expérience. Prenons le cas d'une célèbre expérience de pensée en physique. En 1871, le physicien James Clerk Maxwell nous propose d'imaginer le système

suisant : une chambre remplie d'un gaz est séparée par un mur dans lequel est percée une ouverture. À l'embouchure de l'ouverture du trou se trouve un petit démon qui, au moyen d'une porte, peut ou bien ouvrir ou refermer l'ouverture. Le démon regarde les molécules se dirigeant vers l'ouverture et ouvre ou referme la porte selon leur vitesse (disons qu'il ouvre la porte pour laisser entrer les molécules rapides et referme la porte pour les molécules lentes). Après un certain temps, le démon aura réussi à redistribuer l'énergie dans le système et aura ainsi créé une source d'énergie que l'on pourrait, par exemple, utiliser pour générer de l'électricité. Puisqu'un tel démon est impossible, l'expérience de pensée de Maxwell (communément nommée « le démon de Maxwell ») entend soulever certains problèmes avec la thermodynamique statistique.

Comme le souligne Gendler (2002), plusieurs typologies des expériences de pensée ont été proposées (philosophique vs scientifique (Horowitz et Massey 1991), factuelles vs conceptuelle/valuationnelle (Gendler 2000), constructive vs destructive (Brown 1991) mais aucune n'a su rallier l'intérêt général. Nous proposerons donc notre propre typologie et distinguerons ici les expériences de pensée conceptuelles, visant à tester nos intuitions quant au sens de divers concepts, et les expériences de pensée théoriques, visant quant à elles à établir ou invalider des thèses ou hypothèses théoriques (cette distinction est similaire à celle de Gendler 2000).

Le premier type d'expérience consiste à imaginer des situations diverses permettant de tester des hypothèses à propos des conditions d'application d'un concept. En faisant varier les éléments des cas d'application, on peut déterminer ceux qui sont essentiels et ceux qui sont accessoires. C'est le principe même de la variation eidétique des phénoménologues, qui permet d'isoler ce qui est essentiel à un concept. On peut ranger sous ce premier type d'expérience de pensée le cas célèbre de la « Terre-Jumelle » de Putnam (1985). Putnam nous demande d'imaginer un monde qui est une réplique parfaite, molécule pour molécule, de notre monde, à l'exception du fait que dans ce monde, le liquide incolore et inodore qu'on nomme « eau », que l'on boit, qui constitue à 98 % les cellules, dans lequel on se lave, etc. (bref un liquide qui a les mêmes propriétés macroscopiques que notre eau) n'est pas de l' H_2O mais plutôt une molécule complexe dont on simplifie le nom par XYZ. Lorsque la contrepartie de Putnam sur Terre-Jumelle prononce le mot « eau », à quoi réfère-t-elle ? Et lorsque Putnam lui-même est transporté sur Terre-Jumelle et qu'il utilise le mot « eau » pour décrire le XYZ, fait-il une erreur ? Cette expérience est conçue pour tester nos intuitions à propos de concepts sémantiques comme la « référence » et la « signification » (et vise à démontrer qu'il existe une composante

irréremédiablement externe à la signification des termes d'une langue, bref que la signification n'est pas uniquement « dans la tête »).

Le second type d'expérience de pensée, plus proche quant à lui de la déduction logique, vise à faire un travail théorique. Il s'agit dans ce cas de montrer qu'une thèse est compatible avec une théorie donnée ou qu'un mécanisme peut rendre compte d'un phénomène observé. Par exemple, Griffith et Sterelny (1999) nous demande d'imaginer un gène ayant deux effets : le porteur du gène en question a une barbe verte et est poussé à sacrifier son potentiel reproducteur si ce sacrifice permet à un autre porteur du gène d'augmenter suffisamment son propre potentiel reproducteur. Dans ce cas, la fréquence du gène dans la population augmentera, et ce au dépend des effets potentiellement néfastes du gène sur ses porteurs. Cet argument démontre la possibilité théorique de l'existence de tels gènes dits « hors-la-loi ». L'expérience du démon de Maxwell est de même nature.

Plusieurs doutes ont été soulevés à l'égard de l'usage d'expériences de pensée. Il est impossible de les présenter toutes ici (voir Gendler 2000, 2002) mais nous discutons un problème pertinent dans le présent contexte. Plusieurs recherches en psychologie ont montré que nous sommes de très mauvais raisonneurs, étant sujets à une foule de biais cognitifs (voir par exemple Tversky et Kahneman 1983). Que peut-on penser alors de la valeur d'une méthode fondée exclusivement sur le raisonnement et l'intuition ? À tout le moins, il faudrait s'assurer que nos inférences sont valides et, pour cette raison, nous applaudissons l'apprentissage obligatoire de la logique chez les philosophes et, lorsque possible, la formalisation de nos expériences de pensée. Mais les raisonnements sur les processus nécessitent d'autres outils, sinon théoriquement, du moins dans la pratique. C'est ici qu'entrent en jeu la modélisation mathématique et la simulation qui peuvent servir de prothèse cognitive aux expériences de pensée. Nous décrivons d'abord la simulation en général puis nous nous concentrons sur la simulation pratiquée comme prothèse cognitive aux expériences de pensée.

2.2 Qu'est-ce qu'une simulation ?

Selon le Grand Dictionnaire Terminologique, la simulation est l'« ensemble des techniques permettant d'imiter, aussi fidèlement que possible, des phénomènes réels grâce à des programmes intégrant des modèles mathématiques fondés sur des paramètres qui interviennent dans la réalité. » (*Le grand dictionnaire terminologique*, 1999) Cette définition est insuffisante pour notre propos, mais elle peut suffire ici (il faudrait cependant préciser tout de suite que les phénomènes que modélise une simulation sont toujours des *processus*).

Un cas prototypique de simulation est le cas où l'on étudie un système que l'on peut isoler de son environnement, du moins dans une certaine mesure, et au sujet duquel on dispose d'un modèle théorique précis qui permet de décrire l'évolution du système pour un ensemble de conditions initiales. Ainsi, étant donnée une position de départ, un pendule aura une trajectoire particulière et son évolution sera décrite par une équation différentielle.³ Puisqu'on peut dès lors calculer les états successifs du système, chaque état étant calculé à partir du précédent, on peut simuler l'évolution du pendule. Dans certains cas, la seule manière de connaître l'état futur du système est de calculer ses états successifs, ce qui rend le recours à la simulation nécessaire. Le premier type de simulation repose sur une théorie d'arrière plan, et est déterministe. Ce ne sont pas là des conditions nécessaires au développement d'un modèle. Un autre cas classique de simulation repose uniquement sur des contraintes venant d'une connaissance suffisante d'un phénomène particulier ; et parfois, les règles de transitions entre états sont probabilistes. C'est le cas des modélisations de phénomènes de files d'attente : en connaissant la distribution des temps de service et de l'arrivée des clients, on peut estimer, grâce à des simulations, quel sera le temps d'attente pour tel nombre de guichets.

Il va de soi que l'on ne rencontre que rarement ce cas prototypique dans les faits. La plupart des simulations ont un rapport indirect au modèle théorique sous-jacent, ou encore se font en l'absence d'un modèle théorique précis ou complet ou de faits déterminant les détails de la simulation. Dans certains cas, de nouvelles hypothèses devront être introduites pour rendre le phénomène simulable, hypothèses qui pourront être éventuellement testées : la simulation joue alors un rôle heuristique. Dans d'autres cas, il faut, pour rendre la simulation possible, utiliser des stratagèmes d'implémentations et des simplifications plus ou moins radicales de la théorie. Enfin, certaines simulations visent moins à étudier le comportement d'un système qu'à faire une preuve de possibilité, pour démontrer qu'un phénomène donné peut s'expliquer de manière mécaniste.

Les sciences cognitives et la vie artificielle nous ont habitués à des simulations qui ne dérivent pas directement de modèles théoriques précis. Par exemple, des modèles symboliques comme ceux de Winograd ou Schank⁴ ou les réseaux de neurones artificiels utilisent des simplifications et font des hypothèses nouvelles à propos des phénomènes étudiés. Et les simulations utilisées par Elman (1995) pour modéliser l'apprentissage de la grammaire par un réseau connexionniste et

3. Il faut noter que nous simplifions délibérément : les équations du pendule ne déterminent une évolution pour une condition initiale donnée que si on ignore un ensemble de facteurs (friction, etc.).

4. Ces modèles sont décrits par exemple dans (Crevier 1997). Il s'agit des modèles SHRDLU de Winograd (pp. 121-126) et des divers modèles inspirés de la notion de script développés sous la direction de Schank et Abelson (pp. 195-205).

par Axelrod (1984) pour modéliser l'évolution de la coopération visent davantage à donner une preuve de la possibilité pour certains types de mécanismes d'expliquer certains phénomènes.

2.3 Dennett : les simulations comme expériences de pensée

Dennett a lui-même réfléchi sur la simulation et l'on retrouve chez lui les différentes fonctions que nous venons d'énumérer. Le concept unificateur qu'il propose pour comprendre le rôle cognitif des simulations, et particulièrement le rôle qu'ils pourraient jouer en philosophie, est celui d'expérience de pensée, dont nous avons discuté à l'instant.

Dans son article "*Artificial Intelligence as Philosophy and Psychology*", Dennett (1978) prétend ainsi que les simulations *sont* des expériences de pensées : "*AI programs are not empirical experiments but thought-experiments prosthetically regulated by computers*" (p. 117). Ce qu'il entend par expérience de pensée dans ce texte est clair par sa définition de la « Question de l'IA » : « Comment un système (présentant les propriétés A, B, C...) pourrait réaliser le comportement X ? » (p. 111) Dans ce premier contexte, l'expérience de pensée permet la découverte en imagination d'un mécanisme possible qui présente certaines propriétés que l'on croit essentielles et qui produit un phénomène auquel on s'intéresse. Si nous sommes capable d'imaginer un tel mécanisme, alors nous avons prouvé, à tout le moins, qu'il est possible que ce mécanisme soit présent dans le système que nous étudions. Tout comme certains psychologues contemporains comparent l'imagination aux simulations, et décrivent l'usage de l'imagination comme une « simulation » mentale de ce qui est imaginé, il semble que Dennett voit les simulateurs comme des « prothèses pour l'imagination ».

Dans d'autres contextes, sa compréhension des expériences de pensée est plus traditionnellement philosophique – il s'agit des expériences de pensées que nous avons appelées conceptuelles (la citation qui fait référence au cas de la Terre Jumelle, dont nous avons déjà parlé) :

“Twin Earth sets internal similarity to maximum, so that external context can be demonstrated to be responsible for whatever our intuitions tell us. Thus these thought experiments mimic empirical experiments in their design, attempting to isolate a crucial interaction between variables by holding other variables constant. (Dennett 1994)”

Y'a-t-il un lien entre la compréhension qu'a Dennett des simulations comme expériences de pensée et cette conception plus traditionnelle des expériences de pensée ? Nous nous poserons cette question lors de notre analyse de ses usages de la simulation. Ainsi, une seconde fonction de la simulation chez Dennett, conformément à cet emploi des expériences de pensée en

philosophie, serait donc de servir d'outil prosthétique à des expériences de pensée conceptuelles – bien qu'il nous faudra valider cette intuition.

Une troisième conception de la simulation, qui n'est pas incompatible avec les deux précédentes la voit comme une méthode pour déduire certaines conséquences logiques de nos théories qui sont difficilement accessibles à l'esprit humain :

“[...] A simulation program is ultimately a high speed generator of consequences that some theory assigns to various antecedent conditions [...]” (Dennett, 1979, p. 192).

Enfin, Dennett présente à deux reprises au moins les simulations comme des outils d'assistance à la pensée, outils qui reposent sur l'usage de « micromondes », de simplifications utiles, évitant que des détails inutiles obscurcissent les caractéristiques importantes du problème. Ainsi, dans *Freedom Evolves*, il justifie son usage de la simulation par le fait que les erreurs qu'il dénonce résultent d'illusions centrales, dont il est difficile de se départir, au cœur de la façon dont nous nous représentons la liberté. Nous aurions donc besoin d'outils d'assistance à la pensée qui éviteront les pièges que recèlent nos conceptions intuitives :

“Getting clear about this aspect of the complex relationship between fundamental physics and biology sounds terrifying, but fortunately, there is a *toy* version of that relationship that is just what we need. The difference between a toy and a tool can evaporate if the toy can help us understand things that are otherwise too complex for us to keep track of.” (Dennett 2003, 25-26)

En somme, trois fonctions distinctes seraient assignées à la simulation : un outil de construction d'explications possibles, un générateur de conséquences théoriques, et un micromonde simplifiant un phénomène d'intérêt d'une manière utile. Sur la base de sa théorie de la simulation comme expérience de pensée, nous faisons de plus l'hypothèse qu'une quatrième fonction de la simulation devrait figurer chez Dennett, bien que nous ne connaissions pas de preuve directe que Dennett la propose explicitement: la simulation serait alors un outil d'analyse conceptuelle.

3. Le jeu de la Vie⁵

Munis de ces deux notions, simulation et expérience de pensée, et ayant pris connaissance de leur relation étroite dans la pensée de Dennett, nous pouvons maintenant nous attaquer aux textes où il utilise la simulation dans un but philosophique, et, ensuite, essayer de voir quelle est la nature de ces usages. Pour ne pas allonger inutilement le texte en discutant d'une multitude de formalismes de modélisation mathématique⁶, nous avons choisi de nous restreindre à un seul formalisme, les automates cellulaires, et, pour être plus précis, un automate cellulaire en particulier : le Jeu de la Vie (*The Game of Life*) développée par le mathématicien John Conway.⁷ Nous verrons cependant que Dennett les utilise à une multitude de fins, et nous faisons l'hypothèse que les *types* d'usages qu'il en fait épuisent les types d'usages possibles de la simulation en philosophie. Nous y reviendrons.

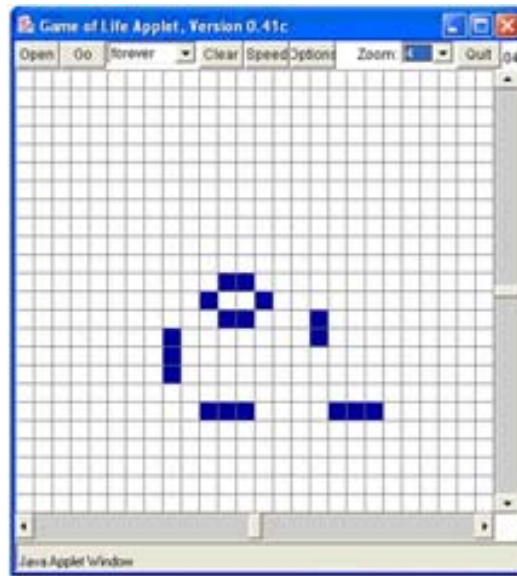


Figure 1. Le jeu de la vie.

5. Cette section est adaptée de Fissette et Poirier (2000). On ne comprend vraiment l'illustration de Dennett que lorsque l'on a observé l'univers du Jeu de la vie. Nous invitons le lecteur à « jouer » lui-même au jeu de la vie. Le lecteur pourra aisément trouver une simulation (par applet java) sur Internet en composant « Game of Life » ou « Conway's Game of life » sur son moteur de recherche favori.

6. Pour ne parler que des grands formalismes de modélisation mathématique que l'on retrouve, on peut mentionner les modélisations logiques, par automates formels, par systèmes de Lyndermayer, par réseaux de neurones, par algorithmes génétiques, par automate cellulaire, par réseaux bayésiens, par systèmes dynamiques. De plus en plus, également, certains chercheurs utilisent des modèles qui sont des hybrides entre des modèles mathématiques et des modèles physiques réels (des robots par exemple). Qui plus est, tous ces formalismes ne sont pas mutuellement exclusifs, de sorte que l'on retrouve des modélisations mathématiques tablant sur deux ou trois de ces formalismes.

7. Le Jeu de la Vie a été introduit au grand public par Martin Gardner (1970) dans sa chronique pour la revue *Scientific American*. Voir aussi Gardner (1983) et Poundstone (1985) pour des traitements plus élaborés.

Le Jeu de la Vie est un automate cellulaire composé d'une grille sur laquelle certaines cellules sont noires, d'autres blanches. La grille se veut infinie pour chacune de ses deux dimensions. Ceci signifie en outre que chacune des cellules de la grille est entourée de huit cellules adjacentes.

Étant donnée une configuration initiale, la distribution des cellules noires et blanches sur la grille « évolue » selon une règle simple, que nous nommerons la « Règle de la Vie » :

Pour chaque cellule C à un temps t , si, parmi ses huit voisines,

- moins de deux sont noires, alors C sera blanche à $t + 1$;
- exactement deux sont noires, alors C demeurera à $t + 1$ dans l'état où elle est à t ;
- exactement trois sont noires, alors C sera noire à $t + 1$;
- plus de trois sont noires, alors C sera blanche à $t + 1$.

Pour déterminer l'état de la grille à $t + 1$, il suffit de calculer l'état de chacune de ses cellules à $t + 1$ selon la règle précédente. Pour déterminer l'état de la grille à $t + 2$, il s'agit de répéter le processus à partir de l'état de la grille à $t + 1$, et ainsi de suite. L'évolution de la grille est donc entièrement déterminée par la règle. Cet automate cellulaire est un système purement déterministe : étant données la règle et une configuration initiale, il est possible de déterminer avec une précision absolue l'état de la grille à tout moment dans le futur. L'univers du Jeu de la Vie est entièrement laplacien.

Lorsque nous observons une seule cellule, nous ne la verrons que s'allumer (devenir blanche) ou s'éteindre (devenir noire). Rien d'autre ne peut se produire, étant donnée la règle déterminant le comportement des cellules. Mais lorsque nous observons d'une certaine distance une grille suffisamment grande, alors nous ne percevons plus des cellules qui passent du noir au blanc et vice-versa, conformément à la règle. Nous voyons plutôt des configurations dynamiques, des *patterns*, de cellules noires et blanches. À ce niveau, la grille n'est plus peuplée de cellules mais d'une foule d'objets exotiques : des clignotants, des ruches, des voiliers, des vaisseaux spatiaux, des gloutons, des canons, des galaxies, des horloges, des pulsars, etc. La figure 1 (voir ci-dessus) montre d'ailleurs une ruche, trois clignotants et un « objet » qui disparaîtra au temps suivant (la prochaine application de la règle)⁸. Et ces objets exotiques participent à des événements variés :

8. Remarquez que, puisqu'elle est un objet non dynamique, vous voyez la ruche sur l'image. Mais les clignotants sont des figures dynamiques (avec une périodicité de 2) qu'il n'est pas possible de voir sur un médium fixe comme une feuille de papier. L'état suivant du clignotant, non représenté ici, est une ligne de trois cellules ayant le même centre que la ligne verticale de trois cellules que vous voyez ici. Bref, un clignotant est une ligne verticale de trois cellules ayant la cellule n pour centre, suivi d'une ligne horizontale de trois cellules ayant aussi n pour centre, suivi d'une ligne verticale..., et ainsi de suite, à l'infini, à moins qu'un objet externe ne vienne perturber le clignotant.

les canons font feu, les prédateurs chassent, les voiliers glissent sur l'écran, les gloutons mangent des voiliers, etc.

Sur le plan « physique », c'est-à-dire mathématique dans le cas de cette simulation, tout est déterminé et prévisible, il n'existe que des cellules noires et blanches et le temps n'amène que des changements de couleur. Mais sur le plan « biologique », où *The Game of Life* porte à proprement parler le titre de « jeu » et manifeste virtuellement une certaine forme de « vie », tout est aussi déterminé mais l'imprévisible abonde (ce qui rend le jeu attrayant à regarder), il existe une foule d'objets exotiques et le temps amène une variété d'événements.

4. Un philosophe s'amuse au Jeu de la Vie

Dennett s'est plusieurs fois référé au Jeu de la Vie dans le cadre de son travail de philosophe. Nous en présentons trois ici, que nous discuterons à la section 5.

4.1 La « réalité » des états mentaux

Dennett est bien connu en philosophie de l'esprit pour s'opposer au Réalisme Intentionnel de philosophes comme Jerry Fodor (1975). Pour un réaliste intentionnel, les croyances et autres états mentaux (désirs, intentions, etc.) existent vraiment et sont autant d'états du cerveau. Dennett rejette cette idée mais ne veut pas pour autant rejeter tout usage du vocabulaire intentionnel en psychologie (s'il le faisait, il opterait pour une position semblable à celle des behavioristes). Pour Dennett, l'adoption de ce qu'il nomme « la stratégie intentionnelle », caractérisée, justement, par l'usage du vocabulaire intentionnel, est utile en psychologie parce qu'elle permet de prédire des événements qu'on ne pourrait prédire autrement. Mais est-ce que cela signifie que, pour Dennett, les croyances et désirs n'existent pas vraiment ? Qu'elles ne sont que des fictions utiles pour prédire le comportement (position qu'on nomme « instrumentalisme ») ? Dennett répondra, après avoir été tenté par l'idée de l'instrumentalisme, que les croyances existent bel et bien mais ne sont observables que d'un certain point de vue, le point de vue de celui qui adopte, justement, la stratégie de l'interprète. Mais comment l'existence peut-elle être liée ainsi à l'adoption d'un point de vue ? C'est pour répondre à cette question que Dennett, pour la première fois, utilise le jeu de la Vie.

On pourrait en effet se demander quel est le mode d'existence des clignotants, par exemple, dans le jeu de la vie. Existents-ils au même titre que les cellules noires et blanches ? Oui, croit

Dennett, car ce ne sont que des configurations dynamiques, des *patterns* de cellules noires et blanches. Poser leur existence n'implique rien d'autre que poser l'existence des cellules qui s'allument et s'éteignent. Mais pourtant, nous n'observons ces objets et événements que d'un certain point de vue : lorsque plusieurs dizaines de cellules sont observées simultanément et que nous laissons notre appareil visuel faire son travail normal de synthèse des configurations. Certains objets et événements qui existent donc bien réellement ne sont toutefois observables que d'un point de vue particulier. Tel est le mode d'existence des *patterns*. Pareillement, nous observons les cellules noires et blanches s'allumer et s'éteindre lorsque nous observons les cellules individuellement. Il existerait donc deux points de vue pour observer le Jeu de la vie. Nommons le premier « point de vue cellulaire » et le second « Le point de vue de la Vie ». Du point de vue cellulaire, il n'y a qu'un type d'objet (la cellule) et seulement deux types d'événements peuvent se produire (la cellule s'allume, la cellule s'éteint). Tout ce qui se produit à ce niveau est entièrement déterminé par la Règle de la Vie. Du point de vue de la Vie, où tout est autant déterminé par la Règle de la Vie, il existe cependant un nombre infini de type d'objets et d'événements.

Pour Dennett, les croyances, désirs et autres éléments mentionnés lorsqu'on adopte la stratégie de l'interprète existent au même titre que les configurations cellulaires du Jeu de la Vie, à cette différence près que ce n'est pas le travail ordinaire de notre système visuel qui perçoit les configurations, mais le travail ordinaire de notre appareil conceptuel, de nos hypothèses analytiques, pour utiliser l'expression quinienne. Équipés des concepts propres à la stratégie de l'interprète, nous « percevons » des configurations dynamiques dans le comportement humain et ces configurations existent tout autant que les comportements individuels qui les composent, tout comme les configurations de cellules existent tout autant que les cellules individuelles.

C'est ainsi que Dennett négocie un moyen terme entre le réalisme fort des réalistes intentionnels, dont Dennett dira qu'il est de « force industrielle » (*Industrial Strength Realism*), et un antiréalisme tout aussi fort, niant résolument l'existence des états mentaux.

4.2 Ordre, Design et Principe Anthropique

Dans son ouvrage *Darwin's Dangerous Idea*, Dennett (1995) revient au Jeu dans le cadre de deux arguments philosophiques (voir les pages 163 à 181, et 173 à 181). Chacun de ces arguments repose sur une analogie entre les lois de la physique et la Règle de la Vie ainsi qu'une autre analogie entre les êtres vivants (de notre monde) et les *patterns* évoluant dans le Jeu de la Vie (notamment les machines auto-reproductrices et la machine de Turing – Conway et ses

collaborateurs ayant montré qu'une telle machine peut exister dans le monde bidimensionnel du Jeu de la Vie).

Le premier argument philosophique concerne l'Ordre, qui est à la base du type de Design propre à l'évolution. Selon Dennett, Darwin peut être interprété comme s'il disait : Donnez-moi l'Ordre, et je vous montrerai comment le Design peut en émerger sans intervention divine. Cet argument entend soutenir la thèse voulant que l'Ordre et le Design diffèrent quantitativement et non qualitativement. Les voiliers du Jeu de la Vie, par exemple, ne peuvent être considérés comme ayant été conçus (*designed*) par les chercheurs : ce sont des *patterns* subsistants qui, étant donnée la Règle de la Vie, émergent spontanément de certains états initiaux. Et, évidemment, ceci est vrai de *tout* ce qui existe dans le jeu de la vie : tout ce qui existe dans une exécution particulière du Jeu de la Vie particulier est déductible, *logiquement* déductible, des conditions initiales et de Règle de la Vie (la « physique » du Jeu). Imaginons maintenant, et c'est là l'expérience de pensée, qu'on ait placé dans une exécution du Jeu de la Vie des êtres capables de se reproduire (et il y en a⁹) et qu'on les ait laissés entrer en compétition pour les ressources et se reproduire de manière différentielle selon leur capacité à s'approprier les ressources, il est permis de croire que nous aurions alors tendance à considérer leurs descendants comme ayant été davantage « conçus » (*designed*) que leurs ancêtres. Mais il ne s'agit pas d'une distinction de principe (*principled*). Toute configuration du Jeu de la Vie est purement le résultat de l'Ordre (et cela fait écho au cri de Marx : Darwin a réduit la téléologie à l'ordre matériel : p. 65). Ce qui est produit par Design – s'il y a une telle chose – est donc aussi produit par l'Ordre. Et d'autre part même les éléments les plus simples sont considérés par Dennett comme ayant une fonction : il mentionne les molécules de glucoses comme ayant une fonction de stockage de l'énergie, et dans le monde de la vie on pourrait penser aux gliders qui ont une fonction dans les Machines de Turing Universelles dont Conway a prouvé la possibilité. Bref, il n'y a vraiment pas de distinction de principe entre le Design et l'Ordre, entre ce qui est un produit du Design et ce qui est simplement ordonné.

Le second argument philosophique concerne le fameux principe anthropique¹⁰. Imaginons, c'est là l'expérience de pensée, que les êtres dont nous parlions dans le paragraphe précédent

9. Rappelons que les automates cellulaires ont été inventés par John von Neumann justement dans l'intention de réfléchir sur les conditions minimales nécessaires pour qu'un être soit capable de se reproduire.

10. « Selon ce principe, les conditions de l'évolution aboutissant à l'homme sont tellement improbables, elles supposent de telles coïncidences pour qu'apparaisse la vie terrestre, que l'évolution depuis le Big Bang devait nécessairement avoir une orientation, peut-être une intention, voire un projet : l'homme. » (Propos de Jacques Languirand, Guide Ressources, Vol. 08, N° 01, septembre 1992; <http://radio-canada.ca/par4/gr/gr0801.htm>). Languirand expose dans ce passage l'interprétation forte (ou modale) du principe, qui nous concerne ici, il y a également une interprétation plus faible (non modale), dont nous ne parlons pas.

(ceux qui étaient capables de se reproduire dans les confins du Jeu de la Vie) ont évolués à partir d'une certaine configuration initiale et qu'ils ont suffisamment évolué pour découvrir les lois de la « physique » de leur monde (c'est-à-dire, ils ont découvert que tout dans leur monde obéit la Règle de la Vie). Supposons également qu'ils se soient rendus compte que – comme c'est le cas – seule la Règle de la Vie peut donner lieu à la Vie.¹¹ Et supposons enfin qu'ils en déduisent l'intervention d'un créateur bienveillant, ayant, pour que la vie soit possible, créé un monde ayant précisément ces lois-là. Ou bien supposons, pour adapter la citation de la note 10, qu'ils se disent : « les conditions de l'évolution aboutissant à nous sont tellement improbables, elles supposent de telles coïncidences pour qu'apparaisse la Vie, que l'évolution depuis notre Big Bang devait nécessairement avoir une orientation, peut-être une intention, voire un projet : nous-mêmes. ». Et ils auraient raison ! Le mathématicien John Conway a *créé* le Jeu de la Vie précisément dans ce but. Dans leur monde, Dieu a un nom : John Conway ! Mais leur inférence serait néanmoins invalide (rappelons que la validité d'une inférence dépend uniquement de la méthode utilisée pour obtenir la conclusion peu importe que cette dernière soit vraie ou fausse – nous avons donc ici le cas d'un raisonnement invalide appuyant une thèse vraie). Du fait qu'aucun jeu de lois fondamentales de la physique autre que celui qui nous gouverne ne permet la vie, certains cosmologues (physiciens et autres) ont conclu qu'un Créateur a dû choisir le jeu actuel de lois spécifiquement pour qu'il aboutisse à la vie, voire à nous. Mais, nous rappelle Dennett, cette inférence est aussi invalide que dans le cas précédent (et comme dans le cas précédent, ceci ne signifie pas que la conclusion est fausse, simplement que l'argument offert ne la justifie pas). Dennett le montre en mettant en parallèle le jeu de la vie, un argument philosophique traditionnel, et un argument cosmologique, tous montrant une alternative possible au principe anthropique :

- 1) Un processus de recherche par essais et erreurs aurait également pu découvrir la Règle de la Vie, ou encore un processus d'essai de toutes les possibilités. Même si Dennett ne mentionne pas explicitement ces algorithmes, il semble que sa présentation du jeu de la vie et la terminologie employée par la suite (“a purely algorithmic Darwinian process of world-trying”, p. 177) les rend implicites dans la discussion.¹²

11. C'est un fait empirique que, bien que certaines règles génèrent des phénomènes intéressants, aucune autre règle peut ne générer des phénomènes complexes où l'on pourrait voir une forme de vie : ces règles mènent rapidement à des attracteurs fixes ou périodiques simples (elles « gèlent ») ou à expansions désordonnées (elles « explosent »).

12. Nous sommes face à un problème d'interprétation. D'une part Dennett ne fait pas référence explicitement à une éventuelle implémentation de ces processus dans des algorithmes appliqués au Jeu de la Vie. D'autre part considérer que la référence à de telles implémentations n'est pas implicite rend sans objet la mention par Dennett du Jeu de la Vie dans ce contexte. Nous avons donc choisi de faire l'hypothèse (audacieuse ?) que Dennett met en parallèle de tels algorithmes et les processus qu'il décrit.

-
- 2) Hume a imaginé deux façons dont l'ordre aurait pu surgir dans le monde : une « mécanique stupide » qui, par essais et erreurs, recherche un monde capable d'abriter la vie ; ou un essai de tous les mondes possibles : si les possibilités sont finies, alors un nombre fini de monde est possible, et si le temps est infini, un autre « mécanisme aveugle » peut essayer au hasard toutes les possibilités.
 - 3) Selon certains physiciens, seuls certains univers sont capables de donner naissance à des trous noirs, lesquels sont eux-mêmes des univers, dont seuls certains sont capables de donner naissance à des trous noirs, et ainsi de suite, indéfiniment. Un univers incapable de donner naissance à au moins un trou noir ne laissera pas de « descendants ». On peut également penser, avec d'autres physiciens, que notre univers n'est que l'un des univers ayant existé, et qu'il s'insère dans une séquence infinie d'univers ainsi générés. Si la capacité de créer des trous noirs est liée à la capacité d'abriter la vie (le carbone est à la base de la vie (telle que nous la connaissons) mais aussi nécessaire au processus d'effondrement stellaire conduisant aux trous noirs), alors il est nécessaire que notre univers possède les lois fondamentales à la vie puisqu'il contient des trous noirs.

4.3 Déterminisme et liberté

Dans *Freedom Evolves* (2003), Dennett cherche à développer une théorie dite « compatibiliste » de la liberté, c'est-à-dire qu'il cherche à montrer que la liberté humaine est compatible avec une nature physique déterministe. Mais, selon lui, le phénomène de la liberté humaine est trop complexe pour penser directement sa relation au déterminisme éventuel des lois de la physique¹³. Il recourt donc à une simplification, inspirée de Quine (1969), qui permet d'éliminer toutes les caractéristiques qui ne semblent pas essentielles au problème du déterminisme. En simplifiant beaucoup, on peut imaginer un monde sur la base d'un espace (un jeu de coordonnées cartésiennes) tridimensionnel où, à chaque instant, chacun des points de l'espace est, ou bien occupé par un atome d'un certain type, ou bien inoccupé. Un monde serait donc une séquence de telles configurations d'atomes. Pour simplifier davantage, on peut imaginer qu'il n'existe qu'un seul type d'atome, lequel est, à un instant donné, ou bien présent, ou bien absent à un point particulier de l'espace. On peut également discrétiser l'espace, de façon à pouvoir représenter un point de l'espace par un triplet d'entiers. Dennett nomme de tels mondes, des *mondes de Démocrite*. Prenons maintenant l'ensemble des mondes de Démocrite possibles. La vaste majorité de ceux-ci sont totalement incompréhensibles: les atomes y surgissent de nulle part pour disparaître tout aussi soudainement, dans un chaos complet (Dennett appelle ces mondes des *mondes nihilistes*). Un petit nombre de mondes de Démocrite présentent des régularités, ce sont les mondes

13. Les lois de la physique quantique ne sont pas, de fait, déterministes, du moins dans leur formulation traditionnelle. Cependant, le but de Dennett est de montrer que, même si des modèles déterministes – les modèles dit à variables cachées par exemple – venaient à être trouvés, cela ne changerait rien à notre liberté.

réguliers. Pour ces mondes, il est possible de spécifier les règles déterminant le passage d'un état à un autre, et ce soit avec certitude (ce monde régulier est alors dit « déterministe ») soit avec une probabilité plus élevée, mais non totale ($p = 1$), de passer à certains états plutôt qu'à d'autres (ce monde régulier est alors dit « indéterministe »)¹⁴. Tels sont donc les mondes de Démocrite.

Dennett nous présente ensuite le Jeu de la Vie (pp. 36-51 et 56-62) en rappelant qu'une exécution du Jeu de la Vie est un Monde de Démocrite. Bien que l'espace du Jeu de la Vie soit bidimensionnel, chaque point à chaque instant y est dans l'un ou l'autre de deux états, soit occupé (par un atome) ou non, selon l'interprétation des Mondes de Démocrite, soit vivant ou mort, selon l'interprétation traditionnelle de l'état des cellules du Jeu de la Vie. Comme le Jeu de la Vie présente des régularités que celles-ci sont telles que chaque état détermine totalement celui qui le suit: le monde de la Vie est donc un *Monde de Démocrite régulier déterministe*. Bref, en discutant du « Monde de la Vie », on élimine une autre caractéristique inessentielle pour penser la question de la compatibilité entre la liberté et le déterminisme : la troisième dimension.

On sait par ailleurs grâce aux travaux de Conway et de ses étudiants que le Jeu de la Vie permet l'existence d'une machine de Turing Universelle, de configurations persistantes, de systèmes autoreproducteurs, ainsi que le déroulement d'un processus d'évolution limité, bref tous les ingrédients permettant l'évolution d'une classe particulière de configurations (patterns) qui sont capables d'éviter d'être détruites par un débris perdu. Dennett nomme ces configurations particulières « les éviteurs »¹⁵ (*avoiders*)¹⁶. Une telle constatation forme la prémisse de l'argument suivant :

- Dans certains mondes déterministes, il existe des créatures qui évitent d'être détruites (nos « éviteurs » dans le Jeu de la Vie).
- Donc, dans certains mondes déterministes, certains événements sont évités.
- Tout ce qui est évité est évitable.
- Par conséquent, dans certains mondes déterministes tout n'est pas inévitable.
- Par conséquent le déterminisme n'implique pas l'inévitabilité. (Dennett 2003, p. 56)

14. Dennett pourrait utiliser le concept de complexité de Kolmogorov pour distinguer les mondes réguliers (déterministes ou indéterministes) des mondes nihilistes, pour autant que le temps et l'espace soient finis : un monde est nihiliste si sa description est incompressible. Dennett connaît ce concept mais ne l'applique pas dans ce contexte.

15. Soulignons pour éviter la confusion que les « éviteurs » ne sont pas des types particuliers de configurations du Jeu mais un type de types de configurations : toutes celles qui sont capables d'éviter d'être détruites.

16. "...so we can say with some confidence that our toy deterministic world is one in which all the necessary ingredients exist for the evolution of... avoiders !" (p. 51).

Cette affirmation est un point crucial dans la démonstration de la compatibilité de la liberté et du déterminisme. Si, dans un monde déterministe, on peut authentiquement éviter certains événements (la destruction par le contact avec un débris), alors on ne voit pas pourquoi la liberté y serait impossible. Mais cette interprétation de l'expérience de pensée (celle qui voudrait voir des « éviteurs » dans le monde du Jeu de la Vie) n'est pas aussi claire qu'il paraît. Dennett s'en rend compte et invente un dialogue avec un sceptique. Celui-ci objecte qu'en fait, aucune créature n'évite réellement quoi que ce soit: les créatures qui *semblent éviter* des événements ne font que cela : *sembler éviter*. Puisque le monde en question est déterministe et que les « éviteurs » exécutent nécessairement les comportements qui mènent à l'évitement, il s'ensuit que l'événement n'aurait pas eu lieu de toute façon. Est-ce que l'idée d'*éviter un événement* ne signifie pas que l'on *change l'avenir*? Dennett répond que l'idée même de changer l'avenir est incohérente (p. 59) : si, en faisant X, on change l'avenir A en l'avenir B, alors A n'a jamais été l'avenir. Revenons aux mondes de Démocrite. Le fait qu'il n'y ait qu'une seule succession d'états du monde – et donc un seul avenir – est indépendant du fait que le monde soit déterministe ou non : c'est simplement la définition de ce qu'est un monde. Le déterminisme est une propriété *de* cette série, qui dit qu'il existe des règles qui déterminent tous les états futurs du monde étant donné un état présent ou passé.

Une autre objection est alors possible : si le futur est déterminé par le passé, l'« éviteur » ne *pouvait* pas ne pas éviter ce qu'il évite. Donc il n'évite rien, puisque l'événement ne *pouvait* pas se produire. La réponse de Dennett ne ferait donc que déplacer le problème du factuel au modal. Dennett (au chapitre 3) répondra à cet argument en disant que déterminisme et possibilité sont compatibles, et utilisera de nouveau les Mondes de Démocrite¹⁷. La possibilité, selon Dennett, dépend du fait que, dans l'ensemble des mondes possibles suffisamment similaires au nôtre, un événement se produit. Dans le cas du Jeu de la Vie, pour déterminer si l'« éviteur » *pouvait* éviter un débris, il faudrait considérer l'ensemble des Mondes de Démocrite bidimensionnels respectant la Règle de la Vie, et où la configuration macroscopique étudiée est reproduite, avec éventuellement des modifications microscopiques légères (ce qui rendrait le monde considéré « suffisamment similaire »). Le Jeu de la Vie joue un rôle particulier dans ce débat, dans la mesure où nous *savons* que les mondes respectant la Règle de la Vie sont déterministes : “*Now that we have Conway's Life world under our belt, we can always check our intuitions by recasting the issue into Conway's Life world, where we do know the physics perfectly and know that it is deterministic.*” (p. 65)

17. Rappelons à cet effet que les Mondes de Démocrite justement ont été inventés par Quine justement pour traiter des problèmes liés aux modalités.

Comme cette partie de la discussion dépasse notre propos, nous terminerons ici notre présentation des usages de la simulation chez Dennett pour passer à une analyse critique de ces exemples et de leur rapport à sa théorie de la simulation, et pour finalement évaluer sa thèse selon laquelle la simulation peut apporter quelque chose à la philosophie.

5. La nature et la valeur des arguments par simulation

Nous avons vu que Dennett formule, et ce à trois reprises dans son œuvre, des arguments qui utilisent le Jeu de la vie et dont l'objectif est d'appuyer différentes thèses philosophiques :

1. un argument pour soutenir une certaine forme de réalisme des états mentaux,
2. un argument en faveur d'une certaine compréhension de la distinction entre les concepts d'Ordre et de Design,
3. un argument militant en faveur du rejet du principe anthropique et
4. un argument en faveur de la thèse compatibiliste au sujet de la liberté humaine.

Nous avons également vu que Dennett a réfléchi de manière générale sur l'usage de la simulation en philosophie et a dans ce contexte été amené à formuler quatre fonctions de la simulation :

1. La simulation comme un outil de construction d'explications possibles,
2. la simulation comme générateur de conséquences théoriques,
3. la simulation comme micromonde simplifiant un phénomène d'intérêt d'une manière utile.
4. La simulation comme outil d'analyse conceptuelle.

Nous aimerions maintenant tirer quelques leçons épistémologiques. En particulier, nous voudrions d'abord classer chacun des usages du Jeu de la Vie par Dennett comme une instance particulière de l'une des fonctions de la simulation que mentionne Dennett lorsqu'il en discute de manière générale. Ensuite, nous voudrions examiner plus précisément les arguments utilisant le Jeu de la Vie pour en déterminer l'interprétation la plus plausible et, surtout, pour évaluer leur validité ou, le cas échéant, leur invalidité. Ceci nous permettra, en conclusion (§6), d'apprécier la valeur de la simulation en philosophie. Si les arguments de Dennett sont valides, nous aurons là une preuve qu'il est bel et bien possible d'utiliser fructueusement la simulation en philosophie. Une autre façon de poser le même problème serait de se demander s'il y a une raison de principe

pour que la simulation soit préférée, dans certain cas, à l'argumentation traditionnelle. Est-ce que, sans perdre de leur force logique ou rhétorique, les arguments de Dennett auraient pu être formulés sans faire appel à des prémisses reposant sur les propriétés de cet automate cellulaire qu'est le Jeu de la Vie.

5.1. Classification des arguments

Nous avons vu que Dennett attribue quatre fonctions aux simulations en philosophie : (1) pour imaginer des mécanismes alternatifs de production d'un phénomène, (2) comme micromonde pour simplifier les phénomènes du monde réel, nous permettant toutefois de faire des inférences ayant une certaine validité quand au monde réel, (3) pour générer « à haute vitesse » les conséquences logiques de nos théories et, enfin, bien qu'ici c'est plus une hypothèse de notre part, qui reste à démontrer, (4) pour fournir des cas sur lesquels tester nos intuitions sémantiques. Lesquelles de ces fonctions sont utilisées par Dennett dans les exemples étudiés ?

L'exemple de *Real Patterns* ne semble pas exactement un usage de la *simulation*, du moins si, par « simulation », on entend une *représentation* dynamique d'un (type de) système(s) ou de processus. Dennett veut prouver que les patterns peuvent être réels même si leur réalité dépend du pouvoir prédictif de leur postulation en tant qu'entités réelles. À cette fin, il n'utilise pas la simulation pour montrer qu'une situation modélisée peut s'expliquer de telle et telle façon, ou pour montrer qu'une situation a telle et telle conséquences ; bref, il n'utilise pas le pouvoir du Jeu de la Vie comme *représentation*. Il l'utilise plutôt en tant que tel, comme objet informatique présentant certaines caractéristiques, dont celle d'avoir des *patterns* ayant des propriétés similaires à celles que Dennett attribue aux croyances, c'est-à-dire de n'être observable que d'un point de vue. Bien que nous ayons inclus cet usage du Jeu de la Vie dans notre étude de la place de la simulation informatique chez Dennett, il appert après examen que le Jeu de la Vie n'est pas vraiment utilisé dans ce texte comme *simulation*. Il s'agit là d'une leçon (négative) importante : dans certains contextes, certains programmes peuvent cesser d'être des simulations, alors qu'ils le sont dans d'autres. Tout ce qui brille n'est pas de l'or.

Lorsqu'il étudie la distinction Ordre/Design dans *Darwin's Dangerous Idea*, Dennett utilise la simulation pour tester nos intuitions sémantiques. Dans ce cas, nos intuitions au sujet des concepts d'Ordre et de Design sont la variable dépendante, alors que les variables indépendantes sont les types d'objets du Jeu de la Vie auxquels on applique les concepts. Lorsque le Jeu de la Vie est mis à profit pour explorer les débats autour du principe anthropique, la simulation sert à l'imagination d'un mécanisme alternatif : peut-on imaginer un mécanisme expliquant l'apparition

des lois du Jeu de la Vie sans faire appel à un créateur ? Dennett met en parallèle les mécanismes imaginables et les mécanismes similaires postulés par la cosmologie contemporaine. Ici, l'argument par simulation est un argument par analogie.

Dennett présente explicitement le Jeu de la Vie dans *Freedom Evolves* comme une simplification utile nous permettant de penser les relations entre physique et biologie, et entre déterminisme et liberté. Cette fonction « micromonde » de la simulation est effectivement celle que Dennett pratique dans *Freedom Evolves* : plutôt que de se donner une représentation complète du monde selon des principes déterministes et de montrer que l'inévitabilité ne s'ensuit pas – ce qui, vu la complexité des phénomènes en jeu, serait hors de la portée d'un être humain – Dennett utilise un modèle simplifié pour établir son argument. On peut également considérer ce cas comme une instance de la conception d'un mécanisme expliquant un phénomène : il s'agit de montrer qu'il existe un moyen de construire un comportement d'évitement à partir d'une physique déterministe. Et finalement, il y aura, comme nous le verrons, une part d'expérience de pensée conceptuelle. Cet exemple est donc hybride.

5.2 Validité des arguments

Real Patterns. Nous avons souligné que le Jeu de la Vie dans *Real Patterns* n'a pas la fonction d'une simulation mais celle d'un cas concret servant au sein d'une preuve constructive (X est possible ; à preuve, en voici un). Bien que ce ne soit pas un cas de simulation, il est néanmoins possible d'évaluer la validité de cette preuve. À cette fin, reconstruisons son argument. Dennett croit que les croyances et les désirs (et autres attitudes propositionnelles) sont des patrons de comportement que l'on peut observer lorsqu'on adopte un certain point de vue. Il croit par ailleurs que ces patrons de comportements sont beaucoup trop complexes pour qu'on puisse les énoncer purement en termes comportementaux. S'il pouvait dire : « Regardez ici. Tel est le patron comportemental correspondant à ma croyance que Boston est à l'est de Chicago. », il pourrait prouver sa thèse par la simple présentation d'un cas concret. Mais c'est impossible. Dans le Jeu de la Vie, cependant, il y a aussi des patrons qui, à la fois existent ne sont observables que d'un certain point de vue, mais pour lesquels on peut aisément montrer comment chaque patron n'est au fond qu'un agencement d'événements de plus bas niveau. Le Jeu de la Vie est donc utilisé ici pour montrer qu'une possibilité théorique est concrètement possible. On pourrait explicitement formuler l'argument, tout à fait valide, de Dennett ainsi :

-
- Les attitudes propositionnelles sont des patterns de comportements qui existent vraiment (il s'agit de *patterns réels*, de *real patterns*) mais qu'on ne peut observer que d'un point de vue particulier. Ceci ne peut pas être démontré (du moins pour l'instant).
 - Une personne pourrait s'objecter à cette position en disant que l'idée même d'un « patron réel mais n'existant que d'un point de vue » est absurde.
 - Mais elle ne l'est pas. À preuve, les objets exotiques qui vivent dans les deux dimensions de l'automate cellulaire qu'est le Jeu de la Vie sont de patrons réels et mais n'existent que lorsque qu'on observe le jeu d'un certain point de vue (c'est-à-dire globalement en laissant les processeurs visuels effectuer leur travail).

Darwin's Dangerous Idea. Lorsque Dennett pose la question de la distinction entre Ordre et Design dans le monde de la vie, il se sert d'exemples (les voiliers, les êtres auto-reproducteurs, les êtres ayant évolués) pour montrer que la distinction entre l'Ordre et le Design est floue. Dans ce contexte, il semble que le Jeu de la Vie est sert de micromonde¹⁸, une simplification extrême mais utile des relations entre physique et biologie. Mais aussi, le Jeu fournit des contre-exemples (simulés) à une théorie considérant ces distinctions comme parfaitement nettes. Est-ce que l'argument tient ? Est-ce qu'on peut considérer que les concepts d'Ordre et de Design sont applicables ? Il semble que oui. La Règle de la Vie (qu'elle soit ou non une représentation des lois de la physique) semble tomber sous notre concept intuitif d'Ordre ; la production d'une configuration complexe, ayant évolué de surcroît, correspond assez à notre concept de Design. Dennett dit que la distinction n'en est pas une de principe. Comment faut-il l'entendre ? On pourrait penser que l'expérience sert à montrer que la distinction entre les concepts d'Ordre et de Design n'est pas « classique » (*crisp*) mais qu'il s'agit bien plus d'une paire de concepts *flous* (*fuzzy* ; selon la théorie des ensembles flous de Zadeh ; voir Bouchon-Meunier 1994). Que l'Ordre et le Design aient une frontière floue semble effectivement confirmé par le fait que la distinction entre un « objet trouvé » (p. 174) et une configuration ayant fait l'objet de design semble largement arbitraire. Mais les choses se compliquent, car Dennett semble vouloir prouver plus : comme nous l'avons dit, d'une part toute configuration du Jeu de la Vie est purement le résultat de l'Ordre, et d'autre part même les éléments les plus simples sont présentés par Dennett comme ayant une fonction (molécules de glucoses, et peut-être les voiliers). L'argument, ici, semble trop fort, pour qu'on y voie une simple analyse des concepts d'Ordre et de Design comme des concepts flous ; il semble que la distinction s'écroule et que ces choses soit à la fois Ordre et Design, selon le point de vue. Si c'est la conclusion désirée et que c'est ainsi que nous devons

18. Nous verrons que diverses interprétations du micromonde de la Vie sont possibles lors de notre analyse de l'argument à propos de la liberté et du déterminisme ; pour le moment nous nous contenterons d'une compréhension intuitive.

comprendre la phrase “...*there is no line to be drawn between merely ordered things and designed things*” (p. 174), il y a une contradiction apparente entre cet argument et l'importance du concept de Design chez Dennett. Une interprétation possible de cette apparente destruction de la distinction pourrait préserver la distinction tout en la rendant relative : la différence entre un voilier qui est le produit du Design et un voilier qui est le produit de l'Ordre est une propriété dépendant de leur rôle fonctionnel à l'intérieur d'entités plus complexes ; et ces entités elles-mêmes n'ont de propriétés téléologiques que dans la mesure où nous avons besoin de leur attribuer de telles propriétés pour prédire et expliquer leur comportement. Ce sont des *Vrais Patterns*, et toute la biologie en serait remplie. Cette interprétation est cohérente avec l'affirmation de Dennett selon laquelle son livre sur l'évolution vise à démontrer la place de la stratégie de l'interprète en biologie.¹⁹ Utiliser de nouveau le Jeu de la Vie pour faire cette affirmation serait intéressant, puisque Dennett l'a déjà utilisé pour préciser son concept de pattern objectif-subjectif. Cet argument aurait donc pu être valide (en acceptant comme prémisse l'existence de « Vrais Patterns ») ; le problème est que cette interprétation n'est pas directement justifiée par le texte de l'argument lui-même et demeure une conjecture.

Quant au second usage du Jeu de la Vie, il sert à imaginer des mécanismes alternatifs pour rendre compte de l'apparition des lois du Jeu de la Vie (de l'Ordre) sans faire appel à un créateur. En imaginant ce mécanisme, Dennett transfère les résultats à notre monde et recherche des mécanismes similaires dans la cosmologie contemporaine. L'argument peut être résumé comme suit :

- Dans certains Mondes de la Vie, des créatures croient que l'Ordre de leur monde est le fruit d'une intention bienveillante ; mais en fait, cet Ordre est (ou aurait pu être) le fruit d'un processus qui n'est ni intelligent ni bienveillant.
- Donc, dans certains mondes où il y a de la Vie, il est à la fois vrai que l'Ordre du monde est requis pour que la Vie puisse se manifester, et faux que cet Ordre est le fruit d'un processus bienveillant ou intelligent.
- Donc il est invalide de conclure, du fait qu'il y a de la Vie dans un monde et que l'Ordre du monde est requis pour que la Vie puisse s'y manifester, au fait que cet Ordre est le fruit d'un processus bienveillant ou intelligent.

Encore une fois, nous rencontrons le problème de l'interprétation des micromondes. Nous devons donc simplement dire que l'argument est valide si on peut passer du sens de « monde » utilisé dans la première prémisse au sens de monde utilisé dans la conclusion. Ceci peut se faire

19. “I recognize that I have yet to persuade many philosophers that they must take seriously my shocking line on how biology depends, in the end, on adopting the intentional stance towards the evolutionary process itself (that's one reason I am writing my next book on the idea of evolution)...” (Dennett 1992)

par divers dispositifs sémantiques, et la section suivante les analyses justement. Notons, avant de laisser *Darwin Dangerous Idea*, que notre interprétation de l'argument est conditionnelle : car nous avons fait l'hypothèse que la mise en parallèle des algorithmes appliqués au Jeu de la Vie était implicite dans le texte de Dennett ; or il est possible que ce ne soit pas le cas.

Freedom Evolves. Le cas le plus intéressant et le plus développé d'argument utilisant la simulation est certainement celui de *Freedom Evolves*. Nous devons donc lui porter une attention plus grande.

Supposons que Dennett ait réussi à nous convaincre dans ce texte que le déterminisme n'implique pas l'inévitabilité. Comment s'y est-il pris, et que vaut cette conviction ? Revoyons la première prémisse : dans certains mondes déterministes, des « éviteurs » évitent la destruction. Il s'agit à l'évidence d'un *contre-exemple* à la thèse selon laquelle le déterminisme implique l'inévitabilité, mais d'un étrange contre-exemple dont il n'est pas sûr qu'il puisse permettre la conclusion qui en est tirée. Comment devons-nous interpréter cet argument ? Doit-on considérer que le Jeu de la Vie est un *monde* ? Autrement dit, quel est le statut du Jeu de la Vie : est-ce un monde virtuel ayant sa réalité propre, ou est-ce simplement un système formel existant là où ce genre d'entité existe (dans notre tête, dans le ciel platonicien, peu importe), ou encore est-ce un programme fonctionnant dans un ordinateur particulier existant dans *notre* monde ? Ou doit-on, comme le suggère lui-même Dennett, considérer que le Jeu de la Vie est une représentation simplifiée, une « *toy version* » de la relation entre physique et biologie ? De ces interprétations dépendra notre conclusion à propos de la validité d'un tel argument.

Des philosophes et même certains scientifiques considèrent parfois que les simulations sont des *instances* de ce qu'elles simulent : par exemple les systèmes d'Intelligence Artificielle seraient *réellement* intelligents, les créatures de la vie artificielle seraient *réellement* vivantes, car c'est la forme qui détermine la vie, pas la matière. Dans ce cas, Dennett, s'il réussit, sous cette interprétation, à nous convaincre de sa première prémisse, aurait effectivement littéralement réfuté la thèse voulant que « si le déterminisme est vrai, alors tout est inévitable », et ce en nous en présentant un contre-exemple – un autre monde où le déterminisme est vrai et où certaines choses sont évitables. Mais cet argument reposerait alors sur une façon de pratiquer la science qui est pour le moins controversée.

Cette interprétation n'est heureusement pas la seule interprétation possible de la première prémisse. On peut aussi considérer que « monde » est entendu ici par analogie. On infère par analogie que dans notre monde, le déterminisme n'implique pas l'inévitabilité, puisque que dans

cet analogue de monde qu'est le jeu de la vie, on retrouve des analogues d'êtres vivants faisant quelque chose d'analogue à l'évitement.

Ou bien, si on préfère l'idée que le Jeu de la Vie est une représentation simplifiée du rapport entre physique et biologie, on considère que, comme le Jeu de la Vie est une simplification extrême, mais pas totalement invalide sous les rapports qui nous intéressent (lois physiques déterministes, « objets » se déplaçant, se reproduisant, calculant, et évoluant), d'une représentation du monde réel, on peut, avec toutes les précautions nécessaires, inférer que les conséquences tirées de cette représentation tiennent dans le contexte d'une représentation plus réaliste du monde. C'est là ce que semble suggérer Dennett lorsqu'il introduit le Jeu de la Vie.

Une autre option serait de faire disparaître la notion de « monde » des prémisses en considérant son usage comme purement rhétorique. On infère la thèse qui nous intéresse en considérant que le fait qu'un système puisse être entièrement décrit de manière déterministe n'implique pas que rien dans ce système ne soit (analogue à) l'évitabilité. On aura alors éliminé un appel à l'analogie (celui entre le « monde » du Jeu de la Vie et notre monde) et on n'aura pas considéré que le Jeu de la Vie est une description simplifiée de notre monde (ce qui peut sembler un peu trop coûteux, quoique les automates cellulaires semblent être utilisés dans des simulations physiques) ou même un monde « réel » mais virtuel. Si on accepte d'appliquer le terme « éviteur » sans le considérer comme une métaphore, on pourrait dire que cet argument n'en est pas un par analogie, mais bien une expérience de pensée au sens strict : une preuve presque formelle de l'existence d'un système fermé déterministe où des patterns évitent des événements qui les détruiraient. Par conséquent, dire que l'univers est un système fermé déterministe n'impliquerait pas, de ce seul fait, que l'évitement est impossible. Dans ce dernier cas, on présente au lecteur (on lui demande d'imaginer) le phénomène d'évitement au sein du Jeu de la Vie (qui est un système déterministe) et on lui demande si le terme « évitement » est juste. Si les intuitions du lecteur penchent pour cette option, on aura fourni une expérience de pensée prouvant la compatibilité du déterminisme et de l'évitement. Notons cependant que le problème serait alors passé du problème (très complexe) de savoir si le déterminisme implique l'inévitabilité au problème (plus simple) de savoir si tel pattern du Jeu de la Vie peut être authentiquement dit éviter un autre pattern.

6. Conclusion : quelle est donc la valeur de la simulation en philosophie ?

Il semble donc que, dans certains cas, et moyennant certaines précautions, on peut dire qu'au moins un argument de Dennett réussit à utiliser la simulation de manière productive en philosophie, c'est-à-dire qu'il réussit à justifier une thèse philosophique sur la base d'une simulation (ou d'un type de simulation).

Notre dernier objectif dans cet article sera de déterminer la valeur des arguments par simulation en philosophie : ont-ils vraiment un avantage sur les arguments qui ne recourent pas à des simulations ? Que gagne au juste Dennett en utilisant le Jeu de la Vie dans l'argument qu'il développe dans *Freedom Evolves* ? Pourquoi ne pas avoir simplement demandé « Imaginez un monde déterministe... Imaginez que certains assemblages de particules soient détruits alors que d'autres réussissent à se reproduire... Certaines créatures, lorsqu'elles sont sur le chemin d'un assemblage de particules qui serait destructeur pour beaucoup d'autres créatures, vont se comporter de façon telle qu'elles ne seront pas détruites. Considèreriez-vous que ces créatures évitent la destruction ? »

Le gain, croyons-nous, est la clarté et la simplicité, de même que la facilité accrue de l'exercice de l'imagination. Le Jeu de la Vie est simple ; chacune des prémisses menant à l'existence des « éviteurs » est clairement démontrée (sinon par Dennett, du moins par Conway et ses étudiants). Aucun principe caché ne peut donc être invoqué : la physique est totalement explicite. Réfléchir au monde réel de manière abstraite ne permet pas cette clarté et cette simplicité. Il sera toujours possible à un sceptique d'objecter : « ça ne marchera pas dans notre monde ; aucune évolution n'aura lieu ; il est nécessaire qu'il y ait des effets quantiques mystérieux, et c'est pour cette raison que la vie a évolué dans notre monde. » En rendant la physique claire, les principes simples et totalement explicites, le Jeu de la Vie agit comme une prothèse pour l'imagination, laquelle peut désormais imaginer ce qui semble impossible lorsqu'on doit considérer un trop grand nombre de faits et/ou une situation difficile à dégager des présupposés au sujet de notre propre monde. C'est ce que dit Dennett avant de développer son argument : *“Getting clear about this aspect of the complex relationship between fundamental physics and biology sounds terrifying, but fortunately, there is a toy version of that relationship that is just what we need.”* (p. 26).

Ainsi Dennett a tenté d'introduire la simulation parmi les outils du philosophe, et l'a fait de diverses manières : en s'en servant pour produire des contre-exemples à des fins d'analyse conceptuelle (réalité des pattern²⁰ ; distinction Ordre/Design ; déterminisme et liberté) ; pour

20. Mais il faut garder à l'esprit la réserve que nous avons formulé sur la signification dans le cas du mot « simulation ».

imaginer des hypothèses explicatives alternatives (principe anthropique ; déterminisme et liberté) ; et comme micromonde permettant de fournir un cadre simplifié à nos raisonnements à propos de phénomènes complexes (déterminisme et liberté). Est-ce qu'il y a réussi ? Nous croyons qu'il est indéniable que certains usages sont valides (par exemple dans le cas du principe anthropiques). Dans le cas le plus intéressant (celui de la liberté et du déterminisme) la valeur heuristique du modèle est certaine, bien que les résultats soient (selon l'un d'entre nous) moins convaincants. Il reste qu'il s'agit certainement, avec les travaux de Thagard (1992) et de Danielson (1992) d'une des tentatives les plus élaborées d'introduire la simulation dans des recherches qui sont authentiquement philosophiques.

REMERCIEMENTS: Jean-Frédéric de Pasquale remercie le CRSH et le FCAR pour leur soutien financier, lequel a permis la rédaction de cet article.

Références

- Axelrod, R. (1984). *The Evolution of Cooperation*. New York : Basic Books.
- Bedau, M.A. (1996). "The nature of life", in M.A. Boden, éd., *The Philosophy of Artificial Life*, pages 332-357, Oxford University Press.
- Bouchon-Meunier, B. (1994). *La logique floue*. Paris : PUF (Que sais-je ?).
- Brown, J.R. (1991). *The Laboratory of the Mind : Thought Experiments in the Natural Sciences*. New York et Londres : Routledge.
- Bynum, T. et Moor, J.H., eds. (1998). *The Digital Phoenix : How Computers Are Changing Philosophy*. Oxford : Blackwell.
- Crevier, D. (1997). *À la recherche de l'Intelligence Artificielle*. Paris : Flammarion.
- Danielson P. (1992). *Artificial Morality : Virtuous Robots for Virtual Games*. Routledge, London.
- Dennett, D.C. (1979). "Artificial Intelligence as Philosophy and as Psychology." in *Brainstorms : Philosophical Essays on Mind and Psychology*. Harvester Studies in Philosophy. Hassocks, Sussex : Harvester Press.
- Dennett, D.C. (1994). Get Real, in *Philosophical Topics*, vol. 22, no. 1 & 2, pp. 505-568.
- Dennett D.C. (1995). *Darwin's Dangerous Idea : Evolution and the Meanings of Life*, New York : Simon & Schuster.
- Dennett, D. C. (2003). *Freedom Evolves*. New York : Viking Press.
- Elman, J. L. (1995). Language as a dynamical system. In R.F. Port & T. van Gelder (Eds.), *Mind as Motion : Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press. pp. 195-223.
- Fodor, J. (1975).. *The Language of Thought*, New York, Thomas Y. Crowell.
- Gardner, M.(1970). The fantastic combinations of John Conway's new solitaire game "life." *Scientific American*, 223, 120-123.
- Gardner, M. (1983). *Wheels, Life, and Other Mathematical Amusements*, W. H. Freeman and Company.
- Gendler, T.S. (2002). "Thought experiment", *Encyclopedia of Cognitive Science*, New York/Londres, Nature/Routledge.
- Le grand dictionnaire terminologique*, Gouvernement du Québec, Office de la langue française, Canada, CEDROM-SNI, 1999.
- Grim, P., Mar, G. et St. Denis, P. (1998). *The Philosophical Computer : Exploratory essays in philosophical computer modeling*. Cambridge, MA : The MIT Press.
- Horowitz, T. et Massey, G. éditeurs (1991). *Thought Experiments in Science and Philosophy*, Savage, MD : Rowman and Littlefield.
- Poundstone, W.(1985). *The Recursive Universe*. New York : Contemporary Publishing.

-
- Putnam, H. (1985). Signification, référence et stéréotypes. Trad. par Jean Khalifa. In : Philosophie, 1985 (2), n°5, p. 21-44.
- Quine, W. V. O. (1969). *Ontological Relativity and other Essays*, Columbia UP.
- Sloman, A. (1978). *The computer revolution in philosophy : philosophy science and models of mind*, Hassocks, Sussex, Harvester Press.
- Sterelny, K. et Griffiths, P.E. (1999). *Sex and Death. An Introduction to Philosophy of Biology*, Univ. of Chicago Press, Chicago.
- Thagard, P. R. (1992). *Conceptual Revolution*. Princeton, NJ : Princeton University Press.
- Tversky, A., et Kahneman, D. (1983). Extensional versus intensional reasoning : The conjunction fallacy in probability judgments. *Psychological Review*, 90, 293-315.
- Zeigler, B., Kim, T. et Praehofer, H. (2000). *Theory of Modeling and Simulation, 2nd edition*. New York : Academic Press.