

Connaissance et information : une étude comparative critique

Jean Robillard

Téluq-UQÀM

Résumé

Cet article a pour objet une étude comparative de deux concepts, ceux de connaissance et d'information, sur lesquels se fonde la plupart des théories de la connaissance apparues depuis la seconde moitié du vingtième siècle, et la philosophie des sciences qui s'en sert à titre de modèles d'analyse ou de corpus de thèses à vocation explicative ou justificative. Son but est d'analyser certaines thèses, celles de Dretske en particulier, unissant ces concepts au sein d'un même domaine théorique dans le but ultime de fonder une théorie de la connaissance sur une théorie sémantisée de l'information. Nous nous limiterons à quelques thèses choisies en raison de leur exemplarité et, surtout, de l'influence qu'elles ont exercée au cours de l'histoire récente de la philosophie de la connaissance et des sciences. La thèse que nous défendrons est la suivante : à l'encontre de l'usage courant des concepts de connaissance et d'information, il appert que ces deux concepts ne sont pas réductibles l'un à l'autre, tant sur le plan logique que sur celui de leur contribution à diverses théories qui les intègrent parmi les concepts opératoires des explications qu'elles promettent de fournir des objets ou des événements sur lesquels elles portent.

1. Introduction

The ultimate reality, according to Bohr, is not the thing in itself, but the sum total of our information, quantified in terms of probabilities, about the thing. (...) In fancy words, Bohr believed that physics is not about ontology, the science of essence, but about epistemology, the study of how we know what we know, and of the limitations to our knowledge. Epistemology, in turn, has always been concerned with the flow of information (...): *Science is about information*.¹

Les concepts de connaissance et d'information sont associés depuis au moins un siècle. Or, cela n'a pas résulté en une union des plus heureuses. Outre les effets de mode ou de contexte social, économique et politique, encouragés par des investissements massifs durant et après la Deuxième Guerre², plusieurs chercheurs de tous les horizons des sciences (physiques,

mathématiques, sociales et humaines) sont en quelque sorte montés dans le train, filant à vive allure, de la recherche de solutions à des problèmes qui semblaient insolubles ou, du moins, d'une très grande complexité. Les problèmes en question ont eu trait, entre autres choses, à la valeur des connaissances dans un univers au sein duquel connaître devint rapidement assimilé à un acte défini formellement comme une inférence – qu'elle soit *a*) déductive ou *b*) inductive, cela dépend de l'accent mis sur *a*') la dimension justificatrice de l'analyse épistémologique ou sur *b*') la dimension opposée, en principe du moins, de la découverte de nouvelles connaissances, laquelle aura motivé un déplacement focal de l'analyse épistémologique vers les aspects psycho-cognitifs et individuels, voire personnels, de la structuration des connaissances. À chaque dimension correspondent des problèmes, des programmes de recherche et des théories et philosophies de la connaissance. Mais depuis plus de vingt ans on cherche à les réunir. Ce travail vise à interpréter les théories de la connaissance sous l'angle de la théorie de l'information. Est-ce possible de le faire?

Cet article a pour objet une analyse comparative des concepts de connaissance et d'information. Son but est d'analyser certaines thèses, celles de Dretske en particulier, unissant ces concepts au sein d'un même domaine théorique. S'il est juste qu'une étude détaillée des diverses théories de la connaissance contribueraient à l'éclosion d'une analyse plus solide par cela même, le contexte d'un article ne s'y prête guère. Nous devons donc nous limiter à quelques thèses choisies en raison de leur exemplarité et, surtout, de l'influence qu'elles ont exercée au cours de l'histoire récente de la philosophie de la connaissance et des sciences. La thèse que nous défendrons est la suivante : à l'encontre de l'usage courant des concepts de connaissance et d'information, il appert que ces deux concepts ne sont pas réductibles l'un à l'autre, tant sur le plan logique que sur celui de leur contribution à diverses

théories qui les intègrent parmi les concepts opératoires des explications qu'elles promettent de fournir des objets ou des événements sur lesquels elles portent.

2. Le problème de l'union

La question classique de la science, soit celle de savoir quels sont les objets du monde observable et comment en expliquer, et les propriétés, et les relations causales que celles-ci entretiennent entre elles, identifient méta-scientifiquement l'objet de la science. Elle présuppose en outre que la science est appuyée d'une ontologie propre ou, du moins, et selon la tradition positiviste logique, que de la science découleront quelques principes avérés favorisant l'émergence d'une ontologie débarrassée de toute coloration métaphysique. L'idéal positiviste de la science classique n'aura pas vu que cet objectif était en lui-même appuyé sur un principe proprement métaphysique, soit celui de la Vérité en tant que principe représentatif d'un ordre naturel et objet premier de toute connaissance possible et exprimable sous la forme de « Tables de Lois de la Nature ». Mais il aura fallu que cette science si sûre d'elle-même introduise dans ses méthodes et ses principes analytiques un principe d'incertitude pour que cette métaphysique s'écroule lourdement et qu'une nouvelle la remplace aussitôt. L'incertitude, celle autant de la connaissance acquise par l'application de la méthode statistique à l'étude des phénomènes de thermodynamique (cf. Maxwell), que celle de la mesure effectuée par les astronomes (cf. Laplace) en raison de la qualité relative de leurs instruments, cette incertitude dut être néanmoins solidement appuyée par quelque chose d'autre, déjà présent dans l'analyse statistique, c'est-à-dire la rationalité probabiliste. Si le monde est stochastique, alors son explication sera rien de moins que probabilitaire³. Cette dichotomie est présente dans l'œuvre de Hempel et fait aussi l'objet d'une vive attention chez tous les épistémologues et philosophes des sciences qui ont œuvré grosso modo durant la

période allant de 1930 à 1980, dont les travaux de Popper qui, à cet égard, reviennent aisément à la mémoire⁴.

Mais cette histoire serait bien incomplète s'il fallait ne s'en tenir qu'à ce seul axe d'un raccourci utile aux seules fins de notre démonstration. Car en même temps se tramait le développement d'une théorie de l'information dont on estima rapidement qu'elle était susceptible d'applications diverses, insoupçonnées par ses concepteurs mêmes. Entre 1920 et 1950, l'information comme concept et comme objet devint de plus en plus légitime au fur et à mesure que se développaient en parallèle des technologies scientifiques et des résultats de recherche qui, en physique, en sociologie, ou en économie, contribuèrent à rendre du monde dans sa totalité une « image » dont la taille de l'échelle déconcertait. De l'infiniment petit à l'infiniment grand, le regard scientifique ne pouvait plus désormais fournir de ce monde une description à partir de règles et de lois uniformément valides. À une nouvelle science devait obligatoirement correspondre une nouvelle philosophie de la science.

Mais il n'était pas nécessairement question que cette philosophie fût, dans son essence, une ontologie. On retenait malgré tout les leçons viennoises, parfois sans même en dégager les excès normatifs. Mais peu importe. Cette nouvelle philosophie dégagerait de la science les principes nécessaires à expliquer, enfin!, ce que connaître veut dire et en quoi il suffirait d'en décrire les mécanismes intimes pour que la science fût pleinement justifiée dans ses principes et dans ses objectifs; puisque, paradoxalement, le principe de la prépondérance de la description des mécanismes du développement des connaissances accédait au statut de principe de justification ultime de la science.

Entra en scène la théorie de l'information. Shannon, en 1948-49, réussit à formaliser ce que d'autres avant lui ne surent faire. L'idée – géniale – fut de s'inspirer de la thermodynamique de Maxwell, et de mettre en relief⁵ le concept d'incertitude imbriqué dans le calcul de l'entropie, afin d'illustrer comment un phénomène de communication non seulement lui ressemblait, mais trouvait là les fondements nécessaires à toute interprétation quelle qu'elle soit de ce qui se passe lorsqu'un signal quelconque, émis à partir d'une source déterminée, transite via un canal dont on sait seulement avec assurance qu'il est un lieu de transit stochastique, et aboutit à l'autre extrémité passablement réduit en force et en clarté, parce que réduit en nombre; ce phénomène n'est explicable qu'en vertu d'une hypothèse ergodique stipulant que ce processus aléatoire ne produit d'événement connaissable qu'en vertu de la considération de la stabilité de la source à partir de laquelle a lieu le transit⁶. La probabilité que la réception des signaux soit quantitativement équivalente à l'entrée et à la sortie du canal devient ainsi le résultat d'un calcul, le résultat étant que l'information y est ensuite définie comme la réduction de cette incertitude. Le niveau de grande généralité de cette théorie, mathématique rappelons-le, permit ensuite plusieurs adaptations et appropriations.

Or, appliquée à d'autres domaines que celui des réseaux de télécommunication d'alors, que pouvait donc bien signifier une telle théorie? En particulier, si, du point de vue du développement scientifique, la théorie shannonienne de l'information pouvait servir de modèle interprétatif aux phénomènes observés, comme par exemple en balisant la recherche de modèles computationnels des phénomènes physiques ou économiques⁷, que pouvait-elle dire de la connaissance scientifique ou de la connaissance tout court? Ce à quoi l'on assista, dans un premier temps, c'est à une réinterprétation des processus d'apprentissage, et plus généralement cognitifs, à partir d'une incorporation, de la part de la psychologie d'inspiration behavioriste, de la théorie de la communication. La question devint alors de comprendre ce

qui se passait entre les stimuli et les réponses, comment l'assimilation de ce qui devint ainsi une « information » pouvait être représentée au moyen des catégories proprement informationnelles.

Mais qu'en fut-il de la philosophie et de l'épistémologie? Cette incorporation de nouveaux concepts au corpus de la science et de la philosophie de la science exigea de revoir les fondements mêmes des thèses empiristes et sensualistes, à l'aune de théories rapidement baptisées « expérimentales », telle la psychologie cognitive, en acceptant l'idée, issue de la physique des particules, que la simulation par ordinateur pouvait tenir lieu et concurrencer toute autre forme d'expérimentation, préservant du coup à la science son autorité historique sur « la » méthode, soit les dispositifs conceptuels et techniques permettant d'enquêter sur la structure du monde. « Observer » un comportement ou un événement obtenu par simulation informatique fut posé comme l'équivalent formel de l'observation d'un comportement ou d'un événement mondain, c'est-à-dire non simulé informatiquement⁸; surtout si la structure du programme développé à cet effet respecte les lois scientifiques connues et les règles d'inférence usuelles (surtout déductives, car les règles de l'inférence inductive n'existent pas en soi, se limitant fondamentalement aux axiomes et aux règles du calcul des probabilités). Mais plus fondamentalement encore, la philosophie et l'épistémologie intégrèrent à leur tour l'idée *non démontrée* que la connaissance était le produit d'un processus informationnel⁹. Les vieilles thèses empiristes et sensualistes, revues et corrigées, posèrent dès lors que toute connaissance est ou bien une information organisée d'une manière quelconque, ou bien une prise en charge par le sujet – c'est-à-dire par son cerveau et les processus cognitifs inobservables qui sont cependant regroupés sous la catégorie des propriétés neurocognitives – , de signaux lui provenant de son environnement. La science pût dès lors être rassurée : il existe en vertu de ces thèses un monde extérieur à l'observateur (i.e. la thèse externaliste de la

connaissance¹⁰) mais ce monde n'est perceptible et connaissable qu'en raison de sa structure interne propre, laquelle est *ontologiquement* constituée d' « information »¹¹. Les signaux perceptuels sont les véhicules de cette information que le cerveau traite comme s'il était bel et bien une machine computationnelle.

Le principal défaut de toutes ces thèses, toutefois, consiste à faire de l'information le point de départ *et* le point d'arrivée du « canal » par lequel transite le signal. Elles considèrent toutes également que l'information est une « chose »; que celle-ci soit considérée comme une caractéristique événementielle ou phénoménale, ou non, il n'importe. Et elles établissent toutes avec assurance et autorité que la connaissance détenue par un sujet est le résultat du *traitement* d'une information lui provenant de l'extérieur ou à l'opposé de ses sensations proprioceptives.

En un mot : elles font de l'information et de la connaissance une seule et même chose, des synonymes – ou alors ce ne sont que des termes dont le sens reste indifférencié. Premier constat : il faudrait bien cependant convenir d'un critère de démarcation, s'il fallait à tout prix préserver intacte l'épistémologie informationnelle. Ces deux objectifs sont-ils toutefois atteignables?

3. Le critère de démarcation

Démarquer le concept de connaissance du concept d'information peut sembler a priori une question résolue par au moins un argument de taille : si, par le concept d'information, l'on réfère à la théorie mathématique de l'information de Shannon, et si, par le concept de connaissance l'on réfère au concept classique défini à partir de théories sensualistes, empiristes ou intentionnalistes, etc.¹², alors ce critère s'impose de lui-même. Il s'agit du

critère du sens ou de la théorie du sens que présuppose le concept de connaissance au contraire ou à la différence du concept d'information.

En effet, dans un premier temps, il est bien connu que la théorie de Shannon ne fait aucune place à la dimension sémantique de l'information; l'auteur l'a lui-même écrit en introduction de son opuscule¹³. Le concept shannonien d'information n'est pas un concept utile d'un point de sémantique dans la mesure où il réfère uniquement à une quantité (mesurée en bits) d'information transmise dans un signal par le truchement d'un canal; et pour pouvoir effectuer cette évaluation de la quantité d'information transmise, l'observateur doit faire entrer en scène la théorie des probabilités, et donc effectuer une analyse statistique de la transmission de signaux, considérant chaque occurrence d'un signal transmis comme comportant toujours la même structure *a priori*, c'est-à-dire que chaque occurrence d'une information équivaut au logarithme négatif (de base 2, par simple souci de commodité) de la probabilité d'une quantité de bits. Ce concept est en fait celui de la mesure de l'incertitude de la transmission d'un message dont la codification est assimilée à une moyenne de probabilités de transmission de bits à partir d'une source.

Le savoir de l'observateur, savoir spécialisé s'il en est, est préalable à toute discrimination d'une information, d'une part; et l'information dont il est question, d'autre part, est celle qui est *théoriquement* contenue dans le signal qui lui sert de véhicule, indépendamment du savoir de l'observateur quant au moyen statistique d'en évaluer la quantité. Or, lorsque cette évaluation est faite et que le résultat est vérifié, c'est-à-dire lorsque l'observateur en a calculé le résultat probable, on ne peut plus en parler comme une *information au sens de Shannon*, mais d'une information à propos de quelque chose d'autre, soit à propos d'un événement quelconque. Il n'y a qu'un calcul qui porte sur une quantité théorique de bits. À ce moment

précis, cette « information » à propos de quelque chose n'est plus redevable à la théorie de Shannon : ce terme a le sens que lui octroie le dictionnaire.

La différence entre le concept shannonien d'information et le concept usuel de connaissance se situe quelque part dans le fait que la connaissance est, semblablement en cela au concept *ordinaire* d'information, toujours à *propos de quelque chose d'autre*. Cette relation est souvent analysée comme une fonction de représentation. Or, si cela est le cas, c'est que l'on considère qu'une connaissance possède un sens précis (conventionnel ou non, socialement déterminé ou non, peu importe), qu'elle signifie quelque chose à propos d'un état du monde¹⁴ auquel elle réfère. Et, autre particularité de la connaissance, elle peut être vraie ou fausse, précisément en raison de cette relation de signification, ce que ne peut être une information au sens de Shannon, laquelle n'est que probable à divers degrés et à diverses conditions, sur la base du calcul de l'évaluation statistique de la transmission de bits par un signal; rappelons que le signal est le véhicule des bits, il n'en est pas le « référent ».

Un second constat s'impose donc : s'il doit exister un critère de démarcation entre le concept de connaissance et celui d'information, compte tenu des brèves définitions que j'en ai données plus haut, celui du sens semble approprié.

4. Dretske sur le sens

F. Dretske est l'un parmi les rares philosophes qui ont abordé de front cette question. Mais c'est sans doute lui qui a le mieux cerné la distinction entre les deux concepts. Pourtant, dans son *Knowledge and the Flow of Information*¹⁵, Dretske fait du sens non pas ce qui servirait à cliver respectivement les domaines d'appartenance des concepts de connaissance et

d'information, il en fait au contraire ce qui permet de *réunir* ces deux domaines. Pour Dretske, en effet :

The information embodied in a signal (linguistic or otherwise) is only incidentally related to the meaning (if any) of that signal. Typically, of course, we communicate with one another, exchange information, by exploiting the conventional meaning of signs. We convey information by using signs that have a meaning corresponding to the information we wish to convey. But this practice should not lead one to confuse the meaning of a symbol with the information, or amount of information, carried by the symbol.

According to this usage, then, signals *may* have a meaning but they *carry* information. What information a signal carries is what it is capable of “telling” us, telling us *truly*, about another state of affairs. Roughly speaking, information is that commodity capable of yielding knowledge, and what information a signal carries is what we can learn from it. If everything I say to you is false, then I have given you no information.¹⁶

(...)

Information is what is capable of yielding knowledge, and since knowledge requires truth, information requires it also.¹⁷

Le sens dont il est question, on le voit, est ce concept usuel en sémantique des langages formels. Fort bien. Il nous apparaît que cela soit le seul concept qui puisse effectivement mener à une thèse valable sur la valeur de la connaissance – ce qui n'est pas une exigence supérieure à la moyenne des théories de la connaissance. Or, le principal problème de la théorie dretskeenne se situe ailleurs, dans le fait que l'auteur doive impérativement, afin de parvenir à unir les concepts d'information et de connaissance par le biais de la catégorie de sens, ramener le concept d'information, qu'il emprunte dans un premier temps à Shannon, sur le terrain d'une théorie de l'information fondée sur les usages linguistiques ordinaires du *mot* « information ». Le passage par l'analyse du langage constitue sans aucun doute un génial stratagème mais les bénéfices, sur le plan de la théorie de la connaissance, sont plutôt minces. Car, en effet, ce passage entraîne Dretske à rééditer les thèses typiques à la philosophie analytique du langage, sur le statut intentionnel des propositions véhiculées par les phrases formées selon les règles de la syntaxe des langues naturelles; et ainsi à comparer les contenus

propositionnels, vérifiables par rapport à leur valeur de vérité, aux contenus informationnels véhiculés par un signal (linguistique : p. ex. un mot ou une phrase simple) dans un canal (p. ex. : le son [ou plus exactement, l'air] ou, selon les principaux exemples fournis par l'auteur, l'écrit).

Pour Dretske, une connaissance est donc produite (il utilise abondamment le verbe « *to yield* ») lorsque le contenu intensionnel d'une expression linguistique garantit une transmission parfaite (il en fait une norme établie sur la base d'une probabilité conditionnelle égale à 1, donc équivalent à l'absence de « bruit ») d'une information à propos d'un état du monde dont par ailleurs la description extensionnelle, exprimée sous une forme prédicative, peut être équivalente à une autre description du même état sans que pour autant cette différence sur le plan de la description extensionnelle n'affecte la valeur de l'information transmise lorsqu'elle est comme ici définie en tant que tributaire de l'intensionnalité de l'expression qui la véhicule. Deux prédicats peuvent décrire extensionnellement le même état du monde, mais ce n'est pas cela qui leur confère leur valeur informationnelle et cognitive. Ce qui le leur confère est le fait que l'intentionnalité des phénomènes décrits par les prédicats est déterminée par l'intensionnalité de ceux-ci¹⁸. L'intentionnalité des formes linguistiques confère ainsi aux descriptions leur statut de représentation cognitive.

If a signal carries the information that *s* is *F*, it does not necessarily carry the information that *s* is *G* despite the *extensional* equivalence of “*F*” and “*G*”. Even though “*F*” and “*G*” are true of exactly the same things (have the same *extension*), the information that *s* is *F* is different from the information that *s* is *G*.¹⁹

Il y a connaissance si l'information transmise *représente adéquatement* un état du monde ($P(i) = 1$). Et une telle représentation est un *état* d'un système cognitif quelconque; un état intentionnel, bien sûr, formé sous l'égide de deux conditions préalables : 1) le système cognitif doit posséder un savoir préalable à propos de la relation d'information établie entre

une description et les propriétés de l'état du monde qu'elle décrit sous une forme prédicative;

2) l'information transmise doit être causalement liée à l'état représentationnel et intentionnel du système cognitif en tant que cet état est causalement déterminé par des croyances soutenues par l'information adéquatement transmises.

Le représentationalisme dretskéen établit de la sorte que la relation entre une information et une connaissance est causale et qu'elle est, de plus, dépendante de la forme dans laquelle elle a lieu. Ceci est d'une grande importance pour Dretske dans la mesure où sa théorie de la connaissance tient compte, voire intègre, la conceptualité informationnelle classique. Mais ce n'est pas qu'une question de vocabulaire, évidemment. Car le fait de tenir compte de la théorie de l'information, concernant une théorie de la connaissance qui se veut réaliste, impose que l'on accorde à la première la fonction d'expliquer comment se construisent non pas le sens et la connaissance, mais les voies permettant aux signaux d'être interprétés comme autant de véhicules de la valeur de la croyance en un état donné du monde. Or cela requiert que l'on fasse de la connaissance et du processus informationnel de sa construction, le résultat d'un mécanisme dont l'existence en soi n'a plus rien à voir avec les propriétés propres à un système cognitif, celui-ci devant sa structure (connaissance, sens et croyances) à une seule et même cause, soit à une information certaine. Or, une information certaine, dans le langage informationnel, est certes une possibilité, mais il n'en reste pas moins qu'il s'agit bel et bien d'une possibilité hautement théorique²⁰. De telle sorte que le représentationalisme dretskéen a beau vouloir offrir de l'esprit une théorie qui en fait le miroir du monde, la probabilité que cela soit le cas est une question qui malheureusement reste ouverte²¹.

5. Ontologie informationnelle et causalité probabiliste

Quelques mots, avant de conclure, sur deux aspects des plus intéressants de la relation entre nos deux concepts, celui de l'ontologie qui se dégage des thèses dont celles de Dretske nous a servi d'exemple; et sur la notion – fort complexe – de causalité que ces thèses et théories informationnelles impliquent sans que leurs auteurs toujours ne le remarquent.

Rappelons pour commencer que la théorie de Shannon a pour référence, au sens logique du terme, non pas des objets du monde mais des concepts mathématiques. Tout le problème alors consiste en la détermination des critères qui permettent usuellement d'associer un concept mathématique, soit ici le concept d'information, à un événement du monde concret. Vieille question s'il en est. Or, il est possible d'au moins prendre le problème à partir d'une thèse sur le dualisme des propriétés des concepts et des objets concrets, comme par exemple celle que Mario Bunge soutient²².

La traitant comme d'une hypothèse philosophique, Bunge situe sa thèse dualiste sur le plan méthodologique, non sur celui de la métaphysique, et la caractérise « (...), comme l'affirmation selon laquelle alors que les objets concrets existent (tout court), les objets conceptuels sont fiction, c'est-à-dire existent (sur le plan conceptuel) par convention. »²³ Il en tire les deux conclusions suivantes :

Cela implique que construits et choses (concrètes) ont des propriétés fondamentalement différentes. Cela n'empêche pas qu'ils puissent partager certaines propriétés. (...).

Tout construit qui ne sera pas conforme à notre postulat de division des objets en choses et en construits sera dit *ontologiquement mal constitué*. L'attribution des propriétés conceptuelles à des objets physiques, et de propriétés physiques (ou chimiques, biologiques ou sociales) à des objets conceptuels, fait partie de la catégorie des objets mal constitués.²⁴

Le dualisme des propriétés bungeën subdivise donc la classe des « objets » à partir d'un critère ontologique, celui de la constitution propre de ces objets selon qu'ils sont concrets ou

conceptuellement construits. L'objectif d'une telle division est de clarifier ainsi la classe de référence propre aux sciences qu'il appelle « factuelles » et aux sciences dites « formelles ». Ainsi, aux premières correspond la classe des objets concrets et à la seconde celle des construits. Un tel dualisme, toutefois, ne peut être interprété comme étant lui-même constitutif d'une ontologie, au sens métaphysique du terme, attribuant à l'ensemble des objets des propriétés existentielles indépendantes de la convention que Bunge lui-même établit. Il s'agit certes d'un dualisme rationaliste – ou « conceptualiste » selon le terme utilisé par Bunge – mais participant d'une philosophie matérialiste et en ce sens opposée à la métaphysique cartésienne²⁵.

À partir de la thèse précédente, il est aisé de voir que les théories qui, à l'instar de celle de Dretske, associent les concepts d'information à celui de connaissance, qui plutôt les lient à partir du concept de sens qu'ignore la théorie mathématique de l'information, font un mésusage de ce dernier et contribuent ainsi à la formation d'un concept de connaissance ontologiquement mal constitué au sens de Bunge. L'ontologie informationnelle, qui postule l'existence d'une information indépendante de l'activité de construction épistémologique du sujet, est donc incapable de rendre compte adéquatement des phénomènes de connaissance qu'elle tente pourtant de cerner.

Par ailleurs, la *représentation* (théorique, ou conceptualiste) des processus cognitifs effectuée grâce aux catégories de la théorie de l'information entraîne nécessairement l'usage d'un formalisme qui peut certes rendre de grands services à une science qui a pour objet de tels processus. Mais le problème de la référence des modèles informationnels formels ainsi développés se pose plus que jamais dans la mesure où ce qu'il s'agit maintenant de préciser c'est fondamentalement de savoir si ces modèles représentent des phénomènes

informationnels physiquement observables (auquel cas, les concepts risquent de conserver leur caractère d'être ontologiquement mal constitués), ou bien s'ils représentent des mécanismes inférentiels particuliers susceptibles d'instruire sur le type de rationalité qui est la mieux adaptée à la tâche : le cas échéant, la classe de référence de ces théories ne peut être rien d'autre que constituée de concepts, non de phénomènes physiques sinon ceux que la simulation par ordinateur produit en raison du fonctionnement physique de ce dernier, traitement de l'information compris – ce qui entraîne que la simulation, comme *procédé méthodologique*, est bel et bien un processus informationnel, non une modélisation de phénomènes physiques²⁶.

Or, si tel est le cas, alors la représentation informationnelle de la connaissance, comme de ses processus, est nécessairement vouée à s'inscrire dans le cadre d'une approche indéterministe de la science, et plus précisément dans le cadre d'une théorie probabiliste de la causalité au sein de laquelle les mécanismes qu'elle étudie ne peuvent pas être qualifiés de mécanismes causaux au sens mécaniste et causaliste classique du terme. Et ce, pour la simple raison que la théorie de l'information n'est pas une théorie causale au sens classique non plus. (Ce qui fait que le principe de cause dreiskéen ne peut être admis comme valable même s'il est nécessaire à l'argument du sens de la théorie dreiskéenne.) Or, s'il est possible de se représenter un concept de causalité auquel est associé celui d'indéterminisme²⁷, c'est-à-dire de concevoir un principe de relation causale comme tributaire des lois de la théorie des probabilités, alors il est tout aussi possible de concevoir un concept causal de connaissance qui intégrerait ces mêmes lois; soit, par exemple à partir d'un principe qui irait comme suit : « la validité d'une connaissance est une mesure de probabilité non nulle d'une relation causale quelconque entre deux ou plusieurs événements, actions, phénomènes, etc. » Or, même dans ce cas, l'on ne saurait en déduire le caractère informationnel du processus causal lui-même, non plus que la

constitution informationnelle de la connaissance en tant que telle. Il faut distinguer le modèle, les principes théoriques que l'on soutient, la caractérisation des phénomènes modélisés, et la méthode de modélisation elle-même.

6. Conclusion : quand connaître est *être formé de* plutôt que *donner forme à*

Si les concepts de connaissance et d'information ne sont pas logiquement réductibles l'un à l'autre, s'ils ne sont pas non plus synonymes, doit-on pour autant en conclure que leur usage doit être proscrit? Cela nous semble être une conclusion exagérée. Car l'usage de ces concepts est fort répandu. Mais cela ne veut dire pour autant que tous ces usages sont épistémologiquement prudents et légitimes. Car le fait d'associer ces deux concepts sans avoir effectué un retour critique sur la valeur de cette association, cela revient dans les faits à accepter que le principe de l'association prévaut sur le sens qu'elle sert à produire sur le plan scientifique et épistémologique.

Or, si on retient comme valable le critère du sens comme celui devant servir à cliver les domaines propres à la théorie de l'information et à la théorie de la connaissance, il va de soi que, même modélisé informationnellement, le concept de connaissance ne saurait pour autant être défini comme l'est celui d'information. Le critère du sens permet alors d'entrevoir la possibilité de formaliser un concept de connaissance qui soit sémantiquement équivalent au concept usuel d'information, soit qui veuille dire « être formé de » ou « avoir la forme de ». Ainsi défini, le concept de connaissance retrouve sa pleine autonomie vis-à-vis la théorie mathématique de l'information, autonomie qui, en fait, revient à avoir éliminé de son usage et de sa définition le recours à des analogies trop nombreuses. Le processus de transit informationnel que décrit la théorie de Shannon, les règles et les axiomes sur lesquels elle s'appuie, ne pourront plus, dès lors servir à décrire un processus d'acquisition ou de

développement de connaissance comme étant justiciable d'un acte donnant forme, dans une sorte d'absolu, à un contenu (une « information »).

Bibliographie

BUNGE, Mario (1983), *Épistémologie*, Paris : Maloine S.A. Éditeur.

CARNAP, Rudolf, *Logical Foundations of Probabilities*, Chicago: University of Chicago Press & Routledge & Kegan Paul, 1962 (1950).

CHISHOLM, Roderick M. (1999), « The problem of the criterion », in POJMAN, Louis P., *The Theory of Knowledge. Classical and Contemporary Readings*, New York : Wadsworth Publishing Company, pp. 26-35

COHEN, L. Jonathan (1991), *The Probable and the Provable*, Hampshire (Ang.) : Gregg Revivals (Oxford University Press, 1977).

COHEN, L. Jonathan (1989), *An Introduction to the Philosophy of Induction and Probability*, Oxford: Clarendon Press.

CREVIER, Daniel (1997), *À la recherche de l'intelligence artificielle*, Paris : Flammarion, coll. Champs.

DRETSKE, F. (1981), *Knowledge and the Flow of Information*, Cambridge, (Mass.) : M.I.T. Press.

EELLS, Ellery (1996 [1991]), *Probabilistic causality*, Cambridge: Cambridge University Press.

FETZER, James H. ed. (1988), *Probability and Causality. Essays in Honor of Wesley C. Salmon*, Dordrecht: D. Reidel.

FLORIDI, Luciano (Ed.), *Philosophy of Computing and Information*, Oxford (UK) : Blackwell Publishing, 2004, 371 p.

GALISON, Peter (1997), *Image and Logic. A Material Culture of Microphysics*, Chicago and London : University of Chicago Press.

GALISON, Peter, STUMP, David J. (1996), *The Disunity of Science. Boundaries, Contexts and Power*, Stanford : Stanford University Press.

GALISON, Peter (1987), *How Experiments End*, Chicago and London: University of Chicago Press.

GEORGE, Frank H., *Philosophical Foundations of Cybernetics*, Chatham : Abacus Press, 1979, 157 p.

GIGERENZER, Gerd, *et.al.*, *The Empire of Chance. How Probability Changed Science and Everyday Life*, Cambridge : Cambridge University Press, 1997 (1989), 340 p.

GRANGER, Gilles-Gaston (1998), *L'irrationnel*, Paris : Éditions Odile Jacob, coll. Philosophie.

GRANGER, Gilles-Gaston (1995), *Le probable, le possible et le virtuel*, Paris : Éditions Odile Jacob.

GRANGER, Gilles-Gaston (1994), *Formes, opérations, objets*, Paris : Vrin, coll. Mathesis.

GRANGER, Gilles-Gaston (1992), *La vérification*, Paris : Éditions Odile Jacob, coll. Philosophie.

GRANGER, Gilles-Gaston (1988), *Essai d'une philosophie du style*, Paris : Éditions Odile Jacob.

HAACK, Susan (1978), *Philosophy of Logics*, Cambridge (Mass.) : Cambridge University Press.

HACKING, Ian, *L'émergence de la probabilité*, Paris : Seuil, 2002, coll. Liber, 276 p.

HACKING, Ian, Michel Dufour (2004), *L'ouverture au probable. Éléments de logique inductive*, Paris : Armand Colin.

HEIMS, Steve J. (1982), *John von Neumann and Norbert Wiener*, Cambridge : M.I.T. Press.

HEMPEL, Carl (2000 [1996, 1972]), *Éléments d'épistémologie*, Paris : Armand Colin, coll. Cursus

HOUDÉ, Olivier *et al.*, *Vocabulaire de sciences cognitives*, Paris : PUF, coll. Psychologie et sciences de la pensée, 1998, 417 p.

JEFFREY, Richard (2002), *Epistemology Probabilized*, disponible à :

<http://www.princeton.edu/%7Ebayesway/IPMU.pdf>

JEFFREY, Richard (2004), *Subjective Probability. The Real Thing*, Cambridge: Cambridge University Press.

LAFONTAINE, Céline (2004), *L'empire de la cybernétique. Des machines à penser à la pensée machine*, Paris : Seuil.

MATTELART, Armand (1994), *L'invention de la communication*, Paris : La Découverte, coll. Sciences humaines et sociales.

MIROWSKI, Philip (2002), *Machine Dreams. Economics Becomes a Cyborg Science*, Cambridge (UK): Cambridge University Press.

PIERCE, John R. (1980), *An Introduction to Information Theory. Symbols, Signals and Noise*, New York : Dover Publications inc.

POJMAN, Louis P., *The Theory of Knowledge. Classical and Contemporary Readings*, New York : Wadsworth Publishing Company, 1999, 643 p.

POPPER, Karl R. (1973), *La logique de la découverte scientifique*, Paris : Payot.

QUINE, Willard Van Orman (1960), *Word and Object*, Cambridge : M.I.T. Press,

REICHENBACH, Hans (1949), *The theory of Probability*, Berkeley : University of California Press.

SALMON, Westley C. (1988), "Dynamic rationality: Propensitiy, Probability, and Credence", dans FETZER, James H., ed., (1998), *Probability and Causality. Essays in Honor of Wesley C. Salmon*, Dordrecht: D. Reidel.

SALMON, Westley C., (1971), *Statistical Explanation & Statistical Relevance*, Pittsburgh: University of Pittsburgh Press.

SALMON, Westley C. (1963), *Logic*, Englewood Cliffs (N.J.): Prentice hall, inc.

SAYRE, Kenneth M. (1996), “Cybernetics”, in SHANKER, Stuart (ed.), *Philosophy of Science, Logic and Mathematics in the Twentieth Century*, London: Routledge, 1996, coll. History of Philosophy, Vol.9, pp. 292-314.

SEGAL, Jérôme (2003), *Le Zéro et le Un. Histoire de la notion scientifique d'information au 20^e siècle*, Paris : Éditions Syllepse, coll. Matériologiques.

SHAFIR, E., TVERSKY, A., “Thinking Through Uncertainty: Nonsequential Reasoning and Choice”, *Cognitive Psychology*, 24, 1992, pp. 449-474.

SHANNON, Claude, WEAVER, Warren (1963 [1949]), *The Mathematical Theory of Communication*, Urbana and Chicago : University of Illinois Press.

VON BAEYER, Hans Christian (2004), *Information: The New Language of Science*, Cambridge (Mass.): Harvard University Press.

¹ VON BAEYER (2004), p. 65. Italiques originaux.

² Cet article ne porte pas sur le récit de ces événements historiques. Le lecteur intéressé pourra consulter : HEIMS (1982); LAFONTAINE, (2004); MIROWSKI, (2002); SEGAL (2003).

³ Rappelons que deux grandes tendances en théorie des probabilités coexistent : une tendance fréquentiste et objectivante et une tendance dite « bayésienne » ou subjectiviste. Voir la bibliographie. Ces deux postures épistémologiques supposent des ontologies distinctes et des théories de la connaissance à l'avenant. Si la tendance fréquentiste affirme reproduire idéalement le monde via la modélisation et le calcul, la tendance bayésienne quant à elle ne se prononce pas sur la relation monde-modèle probabiliste, seulement sur le *locus primus* des probabilités : la pensée humaine. Or, dans les deux cas, l'explication des phénomènes mondains retiennent l'idée et se fonde sur elle, selon laquelle l'estimation probabilitaire des causes des phénomènes n'est certes pas une garantie de la vérité de l'explication. C'est ce que notre commentaire souligne, sans doute lapidairement.

⁴ HEMPEL, (2000); POPPER, (1973). Il est important de situer ces débats au sein d'un contexte plus large, celui de la recherche des principes de l'unité des sciences, recherche inspirée du programme du Cercle de Vienne. Voir à ce sujet : GALISON, STUMP (1996).

⁵ Il est probable que cela lui fut suggéré par von Neuman. Mais les sources divergent à ce sujet.

⁶ L'ergodicité est essentielle à la démonstration en tant que l'état stationnaire et stable de la source (hypothèse première) autorise l'hypothèse (seconde) d'un nombre fini de signaux. À ce sujet : PIERCE (1980).

⁷ Cela nécessita la mise au point d'ordinateurs suffisamment puissants couplés à des langages de programmation à l'avenant. À ce sujet voir : SEGAL, *op. cit.*; GALISON, (1997); MATTELART (1994); MIROWSKI (2002); CREVIER (1997).

⁸ À ce sujet, voir GALISON (1997) et GALISON (1987).

⁹ La philosophie analytique du langage n'échappe absolument pas à cette remarque critique, bien au contraire. Par exemple: « The less susceptible the stimulus meaning of an occasion sentence is to the influences of collateral information, the less absurdity there is in thinking of the stimulus meaning of the sentence as the meaning of the sentence. Occasion sentences whose stimulus meaning vary none under the influence of collateral information may naturally be called *observation sentences*, and their stimulus meaning may without fear of contradiction be said to do full justice to their meaning. » In QUINE (1960), p. 42. Italiques originaux. Par ailleurs, le professeur Daniel Lemire (Téluq-UQÀM), spécialiste et chercheur en extraction de données (« *data mining* ») nous a fait à ce sujet le commentaire suivant : « Dans le domaine de la découverte des connaissances, on distingue information et connaissance de la manière suivante : une règle, une association ou une classification est une connaissance... alors que les instances que l'on souhaite trier, associer ou classer sont des éléments d'information. Certes, les règles, associations et classifications peuvent devenir, à leur tour, de l'information, (...). » Courriel personnel, 19 juin 2006.

¹⁰ À laquelle s'oppose la thèse *internaliste* de la connaissance selon laquelle le sujet construit ses connaissances indépendamment de sa relation au monde. Voir Houdé (1998).

¹¹ Une précision : le monde n'est pas « donné » comme une structure d'information mais phénoménologiquement perçu comme tel et traduit, représenté, grâce à la méthode scientifique incorporant l'approche informationnelle. Nous y revenons plus loin.

¹² Nous ignorons volontairement de citer les théories sceptiques de la connaissance parce que ces théories ont en général refusé au concept de connaissance toute définition qui fût fondée sur autre chose que sur le paradoxe classique de la certitude de connaître à partir de la distinction entre perceptions vraies et perceptions fausses de la réalité connue ou crue. Pour une discussion de ce paradoxe et une proposition de sa solution, voir : CHISHOLM, Roderick M. (1999).

¹³ Voir : SHANNON (1963 [1949]), p.31.

¹⁴ Nous employons l'expression « état du monde » dans un sens très général qui n'exclue pas le fait que des connaissances sont parfois purement théoriques ou formelles : or, la présence même de celles-ci dans le monde, sur quelque support qu'elles apparaissent, contribue à constituer un état du monde particulier.

¹⁵ DRETSKE (1981).

¹⁶ *Op. cit.*, p. 44. Italiques originaux.

¹⁷ *Ibid.*, p. 45. Ce qui est manifestement faux eu égard à la théorie shannonienne de l'information dont Dretske fait par ailleurs une excellente présentation au début de son ouvrage.

¹⁸ Voir pp. 75-77, *op. cit.* Il faudrait ajouter que la relation d'information établie entre les descriptions prédicatives est pour Dretske une relation nomique qui garantit à son tour leur statut intentionnel aux Lois qui gouvernent la transmission de l'information.

¹⁹ *Op. cit.*, p.75. Italiques originaux.

²⁰ « Because of the variety of symbols events at the input, there is uncertainty in advance (a priori probability less than 1) about what event (E) will actually occur there at a given moment of operation. This uncertainty is removed when E actually occurs (with a posteriori probability of 1). The removal of this uncertainty is designated "information"»: SAYRE (1996), p. 297. Cette théorie n'a aucune prétention à décrire des phénomènes quels qu'ils soient, sauf évidemment le mécanisme grâce auquel la réduction de l'incertitude lors de la transmission d'un signal donnant alors lieu à l'avènement d'une « information ». Or, cette théorie n'est pas prédictive et en déduire l'existence d'une loi de l'avènement du sens lorsque la probabilité d'une information vraie s'avère réalisée (mais nous ne savons pas ce que cela veut vraiment dire pour une probabilité que d'être réalisée), comme le proposait Dretske, surajoute à la probabilité 1 de l'avènement de l'information l'application d'une autre fonction probabilitaire conditionnant ultimement l'avènement du sens : $P(P(i)) = 1 = \text{Sens}$.

²¹ Les textes récents de Dretske approfondissent cette thèse représentationnaliste sans toutefois reprendre l'argument du lien informationnel causal. Voir : DRETSKE, Fred, « The Mind's Awareness of Itself », <http://www.nyu.edu/gsas/dept/philo/courses/consciousness97/papers/dretske.html>. Une critique détaillée se trouve chez BACH. Kent, « Engineering the Mind », <http://userwww.sfsu.edu/~kbach/Dretske.html>.

²² BUNGE (1983). Ce penseur est choisi ici en raison de la simplicité de sa thèse et non parce que nous pensons qu'elle ne pose aucun problème. Les thèses de Granger (1995, 1994, 1992) sur les relations entre les modèles et les phénomènes conviendraient tout autant mais auraient nécessité de bien plus longues incursions explicatives.

²³ Bunge, *op. cit.*, p. 57.

²⁴ *Ibidem*.

²⁵ *Idem*, p. 63.

²⁶ Et il ne suffit pas qu'une définition de variable fasse *intensionnellement* référence à un phénomène physique non informationnel pour que le modèle informationnel développé corresponde *réellement* au phénomène visé par une telle variable, ou plusieurs.

²⁷ Voir : EELLS (1996 [1991]); FETZER (1988).