Une sémantique générale des croyances justifiées

Fabien Schang[1⊞]
Alexandre Costa-Leite[2⊠]

[1] Université de Recherche Nationale des Hautes Études Économiques de Moscou - (HSE-Russie)
[2] Université de Brasília (UnB) - Brésil

01 novembre 2015

Résumé

Nous proposons une logique épistémique quadrivalente $\boldsymbol{AR}_{\mathbf{4}_{\blacksquare}}$, sur la base d'une sémantique qui définit les valeurs des énoncés en termes de réponses à des questions posées. Les propriétés de ce système sont exposées et interprétées dans le cadre d'une épistémologie sociale ; celle-ci comporte une variété d'agents, ainsi qu'un ensemble de relations entre eux. Les notions de justification (forte et faible) et de quasi-vérité sont également revues à la lumière de ce système logique.

1 Introduction

Cet article a pour but d'expliquer les concepts de justification et de quasi-vérité par le biais d'une catégorie d'opérateurs qui ne sont pas modaux, étant donné qu'ils n'ont pas une sémantique des mondes possible à la Kripke et que ce sont des fonctions de vérité, contrairement aux opérateurs modaux. Pour ce faire, l'article se compose de la manière suivante :

1. une réflexion sur les deux types de justification, faible et forte, ainsi qu'une réflexion sur la quasi-vérité ;
2. une introduction au niveau technique de plusieurs opérateurs de croyance non-modaux ;

⊞Ce travail est le résultat d'un projet de recherche intégré à un *Programme de Recherche Fondamental* au sein de *l'Université de Recherche Nationale des Hautes Études Économiques* de Moscou.

⊠Nos remerciements s'adressent également au *Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)*, pour le financement du projet de recherche 486635/2013-9 qui a permis la visite du premier auteur à Brasília et le développement des premières idées de cet article.

3. une étude du rapport entre justification, quasi-vérité, et croyance.

On propose dans ce qui suit une sémantique générale des croyances justifiées. “Générale”, parce qu'elle contient plusieurs opérateurs de croyance particuliers. “Justifiées”, parce que les opérateurs principaux définissent la croyance sur la base de relations logiques entre des informations données.

2 Justification, contingence et quasi-vérité

Soit φ une métavariable d'énoncé. Etant donné une logique bivalente sous-jacente, φ reçoit la valeur de vérité T (vrai) ou F (faux). Toutefois, pour savoir quelle est le valeur de φ on a besoin d'une justification. On va suivre l'article [6] dans cette section, où il y a une exposition des rapports entre justification et quasi-vérite. On a, d'un côté, des justifications qui ne sont pas une bonne garantie de vérité, c'est-à-dire des justifications faibles qui ne sont pas capables de garantir pleinement la vérité de φ. Les justifications faibles sont les plus courantes dans la vie quotidienne : quelqu'un qui affirme la vérité de φ, une expertise sur un crime, une photo, une vidéo filmée, ou n'importe quel autre type d'information pouvant être considéré comme une justification partielle. La propriété fondamentale de cette justification faible, c'est précisément le fait qu'elle ne donne pas une garantie absolue de la vérité d'un énoncé φ. D'un autre côté, il existe des justifications qui sont fortes : elles apportent à l'agent une information certaine, et l'agent croit absolument à la vérité de l'énoncé correspondant. Cette deuxième sorte de justification est celle que l'on trouve surtout en mathématiques, ainsi qu'en logique. Par exemple, la preuve d'Euclide selon laquelle “il existe une infinité de nombres premiers” ne donne pas seulement une justification à l'énoncé, mais aussi une raison décisive en vertu de laquelle cette justification ne peut pas être fausse. En ce sens, on peut dire que l'agent croit absolument à la vérité de l'énoncé en question. La même situation se vérifie partout dans les domaines des sciences formelles. En logique, par exemple, une preuve effectuée au sein d'un système de déduction naturelle donne une justification très forte de la vérité d'un énoncé. En conséquence, l'agent ne croit pas seulement mais sait qu'il s'agit d'une proposition vraie.

L'idée de *justification* se présente sous un format mathématique en logique contemporaine : les logiques de Newton da Costa (voir [7]), mais aussi les logiques de Sergei Artemov (voir [1]). Les logiques de justification les plus récentes ont été étudiées par Sergei Artemov, la différence provenant surtout du fait que, chez Artemov, la justification est un terme associé aux formules. Dans cet article, l'attention sera portée sur les premières. Dans son livre [7], Newton da Costa propose deux systèmes logiques pour traiter de la justification. La suite de l'article va présenter cette logique au sein d'un cadre plus général et uniforme. Appelons simplement $\boldsymbol{J}$

le système de Newton da Costa consacré à la justification. Le langage de la logique $\boldsymbol{J}$ possède deux opérateurs de justification : J (justification forte) et J' (justification faible), avec implication ($\rightarrow$), conjonction ($\wedge$) et négation ($\neg$). Ce système bimodal est une extension de la logique classique, il se compose des axiomes suivants :

A. $J(\varphi \rightarrow \psi) \rightarrow (J\varphi \rightarrow J\psi)$

B. $J\varphi \rightarrow \neg J \neg \varphi$

C. $J'(\varphi \wedge \psi) \rightarrow (J'\varphi \wedge J'\psi)$

D. $J'\varphi \leftrightarrow J'J'\varphi$

E. $J\varphi \rightarrow J'\varphi$.

ainsi que d'un ensemble de règles d'inférence :

(RI-1) Si $\vdash \varphi$, alors $\vdash J\varphi$.

(RI-2) Si $\vdash \varphi \leftrightarrow \psi$, alors $\vdash J'\varphi \leftrightarrow J'\psi$.

(RI-3) Si $\vdash \varphi$, alors $\vdash J'\varphi$.

On conjecture que les axiomes de la logique $\boldsymbol{J}$ sont corrects et complets par rapport aux classe de structures de Kripke F telles que $F = \langle W, R_j, R'_j \rangle$ et $R_j \subseteq R'_j$, où la dernière restriction permet de valider l'axiome d'interaction (E) lorsque R_j est réflexive. Les conditions de vérité de $J\varphi$ et $J'\varphi$ dans une structure de Kripke sont les mêmes pour l'opérateur fort $\Box$ de la logique modale normale, avec R_j réflexive.

Newton da Costa a proposé également un concept très utile pour la notion de vérité dans le domaine des sciences empiriques : le concept de *quasi-vérité* (voir [8] et [7]). L'idée de Newton da Costa a pour but de formaliser des propositions au sujet desquelles nous ne disposons d'aucune connaissance, des propositions auxquelles manque une justification forte. Afin de rendre compte du statut partiel ou provisoire des vérités scientifiques, la fonction principale de la quasi-vérité est de décrire la pratique scientifique de façon plus réaliste, au sein d'un cadre formel. Pour définir la quasi-vérité, da Costa fait d'abord usage de la définition de vérité de Tarski. De ce fait, l'idée de quasi-vérité exprime l'état d'ignorance dans lequel nous sommes par rapport à certaines propositions factuelles. De plus, la notion de quasi-vérité est appropriée pour parler du monde (par le biais d'énoncés empiriques) et de la connaissance que nous en avons. Les énoncés empiriques ne sont pas nécessaires mais contingents, de sorte que la valeur de vérité d'un énoncé empirique sera le faux demain si c'est le vrai aujourd'hui (et vice versa). L'essence d'un énoncé contingent réside précisément dans le changement, et c'est pour cette raison que Newton da Costa a proposé de parler de quasi-vérité lorsque le débat se situe dans le domaine des sciences empiriques. A noter qu'un énoncé quasi-vrai pourrait être qualifié tout aussi bien de "quasi-faux", compte tenu de son caractère contingent. Mais la notion de quasi-vérité reste cépendant prioritaire, dans la mesure où l'objectif essentiel d'une théorie scientifique est d'accéder à la vérité (et non à la fausseté).

Sélon l'article [6], il existe évidemment une connexion naturelle entre le fait d'être une quasi-vérité et celui d'avoir une justification faible. En effet, si une proposition est contingente alors elle ne peut pas être justifiée au sens fort ; elle ne peut l'être que faiblement et, donc, constitue une quasi-vérité. Bien que l'idée originale de Newton da Costa développe le concept de quasi-vérité conformément à la définition de la vérité de Tarski, nous allons proposer dans ce qui suit une version assez différente de la sienne, mais avec le pouvoir explicatif et les propriétés réclamées par da Costa (voir Bueno et De Souza dans [4], p.188) : *ouverture*, *partialité*, et *incomplétude de la connaissance*. De ce fait, nous procédons comme dans [6] et utilisons la quasi-vérité comme un opérateur défini dans la logique $\boldsymbol{J}$:

(Quasi-vérité) φ est une quasi-vérité ssi :
- il existe une justification faible pour φ ;
- il existe une justification faible pour la négation de φ.

On peut dès lors concevoir un opérateur de quasi-vérité au sein du langage de $\boldsymbol{J}$: $QV\varphi := J'\varphi \wedge J'\neg\varphi$. On peut aussi essayer d'axiomatiser une logique de la quasi-vérité à partir de la logique de justification, ou traiter la quasi-vérité comme un concept primitif. [1]

Notre objectif est de trouver une famille d'opérateurs de croyance non-modaux capables de formaliser les idées de justification (surtout la faible) et le concept de quasi-vérité.

3 Information, justification, croyance

Si les logiques multivalentes sont moins utilisées de nos jours, c'est avant tout pour une raison d'ordre technique : non seulement la sémantique des mondes pos-

[1] L'idée d'un opérateur de quasi-vérité est proposée par Otávio Bueno dans son article [3], mais sans faire usage des logiques de justification. Il existe par ailleurs d'autres approches du concept de quasi-vérité, parmi lesquelles celle développée par Coniglio et Silvestrini dans [5] sous la forme d'une logique non-modale de la quasi-vérité.

sibles (ou relationnelle) de Kripke est devenue un modèle majeur pour traiter d'informations plus complexes que les propositions de la logique classique ; mais de plus, les systèmes multivalents semblent incapables de produire une sémantique assez riche pour traiter des informations de type modal : en vertu du théorème de Dugundji (voir [10]), aucune matrice finie ne peut caractériser l'ensemble des systèmes modaux de Lewis situés entre $\boldsymbol{S1} - \boldsymbol{S5}$. Etant donné que le concept de croyance est principalement rendu en logique sous la forme d'un opérateur modal, on voit mal ce qui pourrait motiver la construction d'une nouvelle logique multivalente de la justification. Malgré cela, les croyances sont traitées dans ce qui suit comme des opérateurs unaires mais pas comme des modalités : ils sont définis au sein d'une sémantique algébrique, ou mutivalente, et ce ne sont pas des opérateurs modaux au sens où ils ne sont pas considérés dans une sémantique de type relationnel. Mais il reste un second obstacle à la réalisation de ce projet, cette fois-ci d'ordre conceptuel.

3.1 Information et croyance

Par opposition à la sémantique quadrivalente de Belnap (voir [2]), mieux connue sous l'acronyme $\boldsymbol{FDE}$ (First-Degree Entailment), Dubois a déplacé la polémique dans [9] sur le terrain de l'épistémologie et des concepts liés au domaine de la croyance : selon lui, la logique de Belnap crée une confusion intolérable entre valeurs de vérité et croyances. Que sont ces valeurs de vérité, et à quoi servent-elles ?

Pour donner une idée plus intuitive de son modèle, Belnap a présenté celui-ci comme un réseau d'informations fournies par des ordinateurs. Ainsi, on dit d'un énoncé qu'il est *vrai* s'il existe parmi le réseau au moins un ordinateur donnant une information favorable à l'énoncé ; à l'inverse, on dit qu'il est *faux* s'il existe un ordinateur donnant une information défavorable à ce sujet. On reconnaît derrière ces descriptions les valeurs de vérité classiques que sont le vrai et le faux, respectivement. Quant aux deux autres cas de figure, ce sont les valeurs de vérité *non-classiques* : *ni vrai, ni faux*, lorsqu'il existe un ordinateur ne disposant d'aucune information au sujet de l'énoncé ; *vrai et faux*, lorsqu'un ordinateur dispose des deux types d'information à la fois. Soient T, F, B et N les symboles respectifs de ces quatre nouvelles valeur de vérité ; Dubois reproche aux deux dernières de ne pas en être, dans la mesure où B correspond à un état épistémique de contradiction (il y a un excès d'information au sujet de l'énoncé) et N à l'état épistémique opposé d'ignorance (il y a une pénurie d'information au sujet de l'énoncé). Si Dubois dit vrai, le projet d'une logique de justification multivalente repose entièrement sur une confusion conceptuelle. De plus, la sémantique de Belnap contient des connecteurs dont les matrices aboutissent à des résultats contre-intuitifs : si $v(\varphi) = \mathrm{B}$ et $v(\psi) = \mathrm{N}$, alors (i) $v(\varphi \wedge \psi) = \mathrm{F}$ et (ii) $v(\varphi \vee \psi) = \mathrm{T}$. Comment expliquer qu'une formule φ

composée de deux valeurs distinctes en produise une troisième, lorsque φ est une conjonction ou une disjonction ? De plus, Dubois voit un défaut supplémentaire majeur dans l'invalidité des formules du tiers exclu et de la non-contradiction : (iii) $v(\neg(\varphi \wedge \neg\varphi)) = \mathrm{N}$ et (iv) $v(\varphi \vee \neg\varphi) = \mathrm{N}$ lorsque $v(\varphi) = \mathrm{N}$, sachant que N n'est pas une valeur désignée dans $\boldsymbol{FDE}$.

Une réponse à ces objections est donnée notamment par Wansing et Belnap (voir [15]) : il n'y a pas de confusion dans $\boldsymbol{FDE}$ entre valeurs de vérité et croyance, dans la mesure où les valeurs de vérité (classiques et non-classiques) sont l'expression d'états d'information qui ne sont pas des états épistémiques. Autrement dit, les valeurs T, F, B et N ne symbolisent pas des attitudes de croyance mais des situations impersonnelles en vertu desquelles il existe ou non un argument favorable ou défavorable à un énoncé donné. Une interprétation informelle de ces valeurs donne ceci :

$v(\varphi) = \mathrm{T}$ ssi il existe une information en faveur de (pour) φ
il n'existe pas d'information en défaveur de (contre) φ ;
$v(\varphi) = \mathrm{F}$ ssi il n'existe pas d'information en faveur de (pour) φ
il existe une information en défaveur de (contre) φ ;
$v(\varphi) = \mathrm{B}$ ssi il existe une information en faveur de (pour) φ
il existe une information en défaveur de (contre) φ ;
$v(\varphi) = \mathrm{N}$ ssi il n'existe pas d'information en faveur de (pour) φ, et
il n'existe pas d'information en défaveur de (contre) φ.

Les objections de Dubois sont sans effet contre cette interprétation des valeurs de vérité en termes d'information impersonnelle. D'une part, (i) signifie qu'il n'existe pas d'information en faveur de $\varphi \wedge \psi$ mais qu'il existe une information en faveur de $\varphi \wedge \psi$; (ii) dit le contraire : il existe une information en faveur de $\varphi \vee \psi$, aucune en défaveur de $\varphi \vee \psi$. Quant à (iii) et (iv), ces résultats ne posent aucun problème dès lors qu'ils ne sont plus interprétés en termes d'états épistémiques ou de croyances : (iii) et (iv) résultent du fait qu'il n'existe aucune information au sujet de φ, que ce soit en sa faveur (avec (iv)) ou en sa défaveur (avec (iii)). D'autre part, la vérité et la fausseté ainsi définies deviennent des valeurs logiques indépendantes : un énoncé peut être dit vrai et faux à la fois sans produire aucune incohérence au niveau des états épistémiques des agents : dans le cas où $v(\varphi) = B$, on constate simplement un excès d'information tel qu'il existe des arguments à la fois favorables et défavorables à φ.

Une fois cette objection conceptuelle surmontée, l'essentiel est encore à accomplir puisqu'il s'agit pour nous de construire une logique des justifications au pluriel. Quelle différence logique y a-t-il entre une information et une croyance ? Pour établir une connexion entre les concepts d'information, de justification et de croyance, nous proposons la définition qui suit :

(Just) Un énoncé φ est *justifié* si et seulement l'agent est autorisé à croire que φ (est le cas).

Le caractère normatif de cette définition repose sur le fait qu'un agent croit ou non un énoncé en vertu des *normes* épistémiques auxquelles il souscrit ; ces normes seront décrites plus loin (section 3.3.), en termes de conditions de vérité et de fausseté variables pour un même énoncé.

On voit ainsi que la justification sert de concept intermédiaire entre l'attitude d'un agent vis-à-vis d'un énoncé φ quelconque et sa valeur de vérité présumée, sur la base d'une information préalable. De plus, cette définition conduit à un pluralisme logique dans la mesure où la diversité des modes de justification pour un agent produit une variété de logiques de croyance distinctes. Avant de voir en quoi ces différents systèmes consistent, examinons d'abord la logique qui leur apporte les mêmes informations de base.

3.2 Logique d'information : la logique AR_4

La logique qui suit est un langage basé sur un modèle de questions-réponses nommé $\boldsymbol{AR_{m^n}}$: à un ensemble ordonné de n questions portant sur un énoncé correspond un ensemble ordonné de m types de réponses données par un agent-locuteur. Le nombre des questions et réponses dépend de l'énoncé examiné, selon le degré de complexité de l'information qu'il transmet. Dans le cas qui nous occupe à présent, deux questions portent sur un énoncé φ donné : est-il vrai ? est-il faux ? Rappelons que le modèle proposé est conforme à celui de Belnap, où la vérité et la fausseté ne signifient pas l'existence de faits mais d'informations favorables ou défavorables. Deux réponses sont susceptibles d'être données pour chacune des deux questions : oui, ou non.

On peut facilement adapter le système $\boldsymbol{FDE}$ de Belnap (1977) à partir de cette modélisation dialogique de l'information. Soit $A(\varphi) = \langle a_1(\varphi), a_2(\varphi)\rangle$ la valeur logique d'un énoncé φ du langage $INFO$ de la logique en question, où A est une fonction de valuation telle que $A : INFO \rightarrow \{0,1\}\times\{0,1\}$ (0 pour la réponse négative *non*, 1 pour la réponse positive *oui*). INFO correspond au langage de la logique propositionnelle, où chaque formule est l'expression d'une information sans justification.

L'ensemble des quatre valeurs de vérité simples de Belnap donne les quatre valeurs structurées qui suivent : T = $\langle 1,0\rangle$, F = $\langle 0,1\rangle$, B = $\langle 1,1\rangle$, et N = $\langle 0,0\rangle$, justifiant ainsi l'appellation $\boldsymbol{AR_{m^n}} = \boldsymbol{AR_{2^2}} = \boldsymbol{AR_4}$ de notre logique d'information multivalente.

Une différence majeure entre $\boldsymbol{FDE}$ et $\boldsymbol{AR_4}$ concerne la relation de conséquence, et plus particulièrement la relation sémantique établie entre les quatre valeurs de vérité. Dans $\boldsymbol{FDE}$, il existe une relation d'ordre partiel entre T, F, B et N : cet ordre est représenté notamment dans Ginsberg (1988) par un bi-treillis composé de deux niveaux distincts : le niveau de vérité t, et le niveau d'information i. Dans ce treillis à deux étages, on distingue les deux types de relation suivants :

T $>_t$ B $>_t$ F
T $>_t$ N $>_t$ F,

où B et N sont incomparables relativement à $>_t$ (l'une n'est pas *plus vraie* que l'autre).

B $>_i$ T $>_i$ N
B $>_i$ F $>_i$ N,

où T et F sont incomparables relativement à $>_i$ (l'une ne contient pas plus d'information que l'autre).

La théorie des treillis joue un rôle majeur dans la définition algébrique de la conséquence tarskienne, puisque la relation $\varphi \models_{t/i} \psi$ est satisfaite si et seulement si $v(\varphi) <_{t/i} v(\psi)$. En revanche, il n'y a pas de relation d'ordre dans $\boldsymbol{AR_4}$: chaque valeur logique est logiquement distincte des trois autres, et la seule notion qui importe pour définir la relation de conséquence est celle de désignation : $\varphi \models_{t/i} \psi$ est satisfaite si et seulement si $A(\psi) \in D$ chaque fois que $A(\varphi) \in D$, D symbolisant l'ensemble des valeurs logiques dites *désignées* telles que $D = \{\langle 1, 0\rangle, \langle 1, 1\rangle\}$.

L'ensemble des constantes logiques figurant dans $\boldsymbol{AR_4}$ est caractérisée de façon similaire à celles de $\boldsymbol{FDE}$, à l'exception notoire de l'implication. L'ensemble $INFO$ des formules de $\boldsymbol{AR_4}$ comprend des énoncés d'information, sans aucune connotation épistémique ; ces formules sont construites conformément à la forme de Backus-Naur :

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi$$

Les constantes logiques de $INFO$ sont définies à partir des opérateurs booléens $\sqcap$ et $\sqcup$ appliqués à des valeurs de réponse 1 et 0, telles que $1 \sqcap 0 = 0 \sqcap 1 = 0 \sqcap 0 = 0$ et $1 \sqcup 0 = 0 \sqcup 1 = 1 \sqcup 1 = 1$. Les définitions sont les suivantes, où les symboles des valeurs logiques sont simplifiés sous la forme générale $\langle x, y\rangle = xy$. Pour tout énoncé φ dont la valeur logique est $A(\varphi) = \langle a_1(\varphi), a_2(\varphi)\rangle$:

Atomes : $A(p) = 11$ ou $A(p) = 10$ ou $A(p) = 01$ ou $A(p) = 00$

Négation : $A(\neg\varphi) = \langle a_2(\varphi), a_1(\varphi)\rangle$

Conjonction : $A(\varphi \wedge \psi) = \langle a_1(\varphi) \sqcap a_1(\psi), a_2(\varphi) \sqcup a_2(\psi)\rangle$

Disjonction : $A(\varphi \vee \psi) = \langle a_1(\varphi) \sqcup a_1(\psi), a_2(\varphi) \sqcap a_2(\psi)\rangle$

Implication : $A(\varphi \rightarrow \psi) = \langle a_1(\varphi) \sqcap a_1(\psi), a_1(\varphi) \sqcap a_2(\psi)\rangle$

Les matrices de ces constantes logiques sont comparables à celles de ***FDE***, moyennant une transcription des lettres symboliques de [2] en valeurs booléennes structurées de $\boldsymbol{AR_4}$:

	$\neg$
11	11
10	01
01	10
00	00

$\wedge$	11 10 01 00
11	11 11 01 01
10	11 10 01 00
01	01 01 01 01
00	01 00 01 00

$\vee$	11 10 01 00
11	11 10 11 10
10	10 10 10 10
01	11 10 01 00
00	10 10 00 00

$\rightarrow$	11 10 01 00
11	11 10 01 00
10	10 10 00 00
01	01 00 00 00
00	00 00 00 00

De façon similaire, $\boldsymbol{AR_4}$ dispose d'un système d'axiomes (et aussi des tableaux) comparable à celui de ***FDE***, à l'exception des règles d'inférence pour l'implication : le comportement spécifique de cette constante dans notre logique de l'information est tel que les paradoxes de l'implication matérielle y sont évités ; cela n'empêche toutefois pas ses deux propriétés principales d'être préservées, à savoir le Modus Ponens (MP) et le Modus Tollens (MT). Ainsi, une implication forte est vraie dans le seul cas où l'antécédent et le conséquent sont vrais tous les deux ; elle est fausse uniquement lorsque l'antécédent est vrai et le conséquent est faux. Dans tous les autres cas de figure, l'implication n'est ni vraie ni fausse : l'absence d'argument favorable ou défavorable entraîne une absence de décision globale et, par conséquent, une indétermination de l'énoncé dans son ensemble. Cette caractérisation non-standard du conditionnel est permise par l'indépendance logique des concepts de vérité et de fausseté dans $\boldsymbol{AR_4}$: la non-vérité n'est pas synonyme de fausseté, et inversement. Le caractère symétrique ou asymétrique des valeurs de vérité sera examiné plus loin, relativement au comportement des agents doxastiques dans la logique de justification $AR4_{\blacksquare}$. [2]

Soit t^* une traduction des formules de ***FDE*** en formules de $\boldsymbol{AR_4}$. Les propriétés des connecteurs de ***FDE*** sont préservées par t^* dans $\boldsymbol{AR_4}$, à l'exception

[2]Pour plus de détails à ce sujet, voir [12].

notoire du cas de l'implication qui n'existe pas dans le système de Belnap. Pour cette raison, t^* est une traduction conservatrice de $\boldsymbol{FDE}$ dans $\boldsymbol{AR_4}$.[3]

3.3 Logiques de justification : $AR_{4_\blacksquare}$

S'il n'y a qu'une seule logique de l'information, il existe en revanche plusieurs logiques de justification ; la différence entre ces dernières réside dans les diverses contraintes imposées par les locuteurs sur les conditions d'acceptation et de rejet des énoncés. Pour mieux comprendre cette diversité, prenons l'exemple d'un énoncé peu scientifique tel que *Dieu existe*. Observons tout d'abord le cas du croyant : celui-ci voit dans les beautés de la Nature un signe manifeste en faveur de l'existence du Créateur et, malgré l'objection de la présence du mal dans les actions humaines, le croyant privilégie le premier argument en faveur de l'existence de Dieu et rejette l'argument en sa défaveur. Bien que des arguments défavorables soient disponibles, ce type d'agent ne les prend pas en considération dès lors que des arguments favorables s'y opposent. Dans le cas de l'athée, c'est le contraire : celui-ci privilégie l'argument défavorable à l'existence de Dieu, malgré la disponibilité d'arguments favorables. On peut encore relever deux autres types d'attitude en matière de croyances religieuses : l'agnostique, qui refuse de prendre position tant qu'il n'y a pas d'argument décisif en faveur de l'une ou l'autre position ; quant à l'éclectiste, la pluralité des arguments opposés l'incite à prendre en compte les deux points de vue sans en privilégier aucun.

L'exemple mentionné ici peut paraître incongru, dans le cadre d'une analyse logique de la justification ; mais il a pour but de montrer que les locuteurs obéissent à des types de conduite réguliers et donc rationnels, quel que soit le contenu de l'énoncé à accepter ou rejeter. Dans la majeure partie des cas, l'agent opte pour un seul type d'argument et détermine sa préférence en fonction de critères variables tels que le *poids* de l'argument. Mais en l'absence de critères suffisamment fiables pour délibérer avec certitude, l'agent peut choisir une issue *par défaut* : on le voit par exemple dans le *principe de précaution*, en vertu duquel les agents pensent qu'il y a plus à perdre qu'à gagner dans la poursuite d'un projet où le risque d'incidence n'est pas nul. En d'autres termes, le moindre argument défavorable à une prise de décision implique le rejet du projet et conduit à l'inaction des membres de la communauté. Un tel comportement est comparable à celui de l'athée en matière religieuse, par opposition au pari de Pascal qui voit bien plus à gagner qu'à perdre dans la croyance à l'existence du Créateur.

[3]On suppose pour l'instant que $\boldsymbol{AR_4}$ dispose d'un théorème de complétude, sur la base de la traduction ci-dessus. Mais cela reste encore à démontrer.

Essayons de dégager les caractéristiques de ces agents, sous la forme d'opérateurs de croyance unaires qui reposent sur des critères de justification distincts.

Positivisme

L'agent positiviste :

- croit que φ si et seulement si il existe un argument en faveur de φ, ou il n'existe pas d'argument en défaveur de φ ;

- croit que $\neg\varphi$ si et seulement si il n'existe pas d'argument en faveur de φ, et il existe un argument en défaveur de φ ;

Négativisme

L'agent négativiste :

- croit que φ si et seulement si il existe un argument en faveur de φ, et il n'existe pas d'argument en défaveur de φ ;

- croit que $\neg\varphi$ si et seulement si il n'existe pas d'argument en faveur de φ, ou il existe un argument en défaveur de φ ;

Scepticisme

L'agent sceptique :

- croit que φ si et seulement si il existe un argument en faveur de φ, et il n'existe pas d'argument en défaveur de φ

- croit que $\neg\varphi$ si et seulement si il n'existe pas d'argument en faveur de φ, et il existe un argument en défaveur de φ ;

Ecléctisme

L'agent éclectiste :

- croit que φ si et seulement si il existe un argument en faveur de φ, ou il n'existe pas d'argument en défaveur de φ

- croit que $\neg\varphi$ si et seulement si il n'existe pas d'argument en faveur de φ, ou il existe un argument en défaveur de φ.

Soit $\blacksquare$ un ensemble d'opérateurs de croyance justifiée. On appellera $\boldsymbol{AR}_{\mathbf{4}_\blacksquare}$ la logique de justification générale, comprenant les quatre cas particuliers $\{\boxbar, \boxminus, \boxtimes, \boxplus\}$. Le langage de $\boldsymbol{AR}_{\mathbf{4}_\blacksquare}$ se compose d'un ensemble d'énoncés de croyance justifié $JUST$ qui contient l'ensemble des énoncés d'information $INFO$, conformément à la forme de Backus-Naur ci-dessous :

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid \blacksquare\varphi$$

On peut décrire les quatre attitudes épistémiques par des opérations unaires $\blacksquare$ sur des énoncés d'information φ avec une valuation $A : JUST \rightarrow \{0, 1\} \times \{0, 1\}$. Pour ce faire, nous ajoutons à $\sqcap$ et $\sqcup$ un opérateur booléen de complémentation $-$ tel que $-(1) = 0$:

Positivisme : $A(\boxbar\varphi) = \langle a_1(\varphi) \sqcup -(a_2(\varphi)), -(a_1(\varphi)) \sqcap a_2\varphi \rangle$

Négativisme : $A(\boxminus\varphi) = \langle a_1(\varphi) \sqcap -(a_2(\varphi)), -(a_1(\varphi)) \sqcup a_2\varphi \rangle$

Scepticisme : $A(\boxtimes\varphi) = \langle a_1(\varphi) \sqcap -(a_2(\varphi)), -(a_1(\varphi)) \sqcap a_2\varphi \rangle$

Eclectisme : $A(\boxplus\varphi) = \langle a_1(\varphi) \sqcup -(a_2(\varphi)), -(a_1(\varphi)) \sqcup a_2\varphi \rangle$

D'une part, la distinction entre "positivistes" et "négativistes" réside dans l'attitude de ces agents à l'égard des énoncés affirmatifs et négatifs : les premiers privilégient les arguments positifs ou favorables à un énoncé affirmatif de type φ, tandis que les seconds privilégient les arguments négatifs ou favorables à un énoncé négatif de type $\neg\varphi$. La notion d'"affirmativisme" eut été préférable afin d'éviter toute ambiguïté avec le positivisme d'Auguste Comte, ou le positivisme logique ; mais la lourdeur de la formule a fait pencher ici la décision taxinomique en faveur du concept de positivisme.

D'autre part, la distinction entre “sceptiques” et “éclectiques” réside dans l'attitude minimalement ou maximalement tolérante des agents correspondants à l'égard d'un énoncé donné : le sceptique réclame un argument décisif ou exclusif en sa faveur ou défaveur, tandis que l'éclectique accepte et inclut des arguments contradictoires issus de théories distinctes.
La distinction principale entre ces quatre types d'agent repose sur le rôle décisionnel accordé à la *charge de la preuve* dans la formulation de leurs normes épistémiques. Ainsi, il n'y a pas de charge particulière pour les sceptiques et les éclectiques, dans la mesure où ces agents accordent une valeur égale aux arguments favorables et défavorables à un énoncé donné. En revanche, le poids accordé aux arguments favorables et défavorables n'est pas le même pour les positivistes et les négativistes : un “positiviste” tel que le croyant impute la charge de la preuve aux arguments défavorables de l'athée relatifs à l'inexistence de Dieu, tandis qu'un “négativiste” tel que l'athée impute la charge de la preuve aux arguments favorables du croyant relatifs à l'existence Dieu.

Chaque attitude peut être caractérisée par une matrice quadrivalente des opérateurs de croyance, partant d'un ensemble d'informations identiques sur l'énoncé φ et de critères de justification variables.

φ	$\boxbar\varphi$	$\boxminus\varphi$	$\boxtimes\varphi$	$\boxplus\varphi$
11	10	01	00	11
10	10	10	10	10
01	01	01	01	01
00	10	01	00	11

Au sens large du terme, ces opérateurs unaires décrivent des modalités de croyance dans la mesure où un mode porte avant tout sur un contenu propositionnel. Au sens strict ou logique, ce ne sont pas des modalités parce que ces opérateurs ne sont pas définis au sein d'une sémantique relationnelle (des mondes possibles) et, surtout, parce qu'ils sont, contrairement aux opérateurs modaux, des fonctions de vérité.

Les agents positivistes et négativistes ont un comportement *normal*, à la fois consistant et complet : non seulement aucun d'entre eux ne peut accepter d'argument favorable à la fois à un énoncé φ et son contradictoire $\neg\varphi$; mais de plus, chacun d'eux est susceptible de croire un énoncé "par défaut". Ainsi, le positiviste finit par croire φ en l'absence d'arguments favorables à son contradictoire $\neg\varphi$, tandis que le négativiste finit par croire $\neg\varphi$ en l'absence d'arguments favorables à son contradictoire φ.

Les agents sceptiques et éclectiques ont un comportement *paranormal* au sens métalogique du terme, c'est-à-dire *inconsistant* ou *incomplet*. L'incomplétude est caractéristique du sceptique, qui admet les situations dans lesquelles il n'y a d'argument ni favorable ni défavorable à un énoncé donné ; dans un tel cas de figure, l'agent ne croit ni φ ni $\neg\varphi$, et cette situation est décrite par les première et dernière rangées de sa matrice caractéristique. L'inconsistance est caractéristique de l'éclectique, lorsqu'une situation apporte un argument favorable et défavorable à un énoncé. L'agent croit à la fois φ et $\neg\varphi$ dans ce cas de figure, comme le montrent également les première et dernière rangées de sa matrice caractéristique. Quoi qu'il en soit, on peut décrire les propriétés des agents rationnels par le biais d'une liste de formules utilisées couramment en logique modale :

(P1) $\blacksquare(\varphi \wedge \psi) \models \blacksquare\varphi \wedge \blacksquare\psi$;

(P2) $\blacksquare(\varphi \vee \psi) \models \blacksquare\varphi \vee \blacksquare\psi$;

(P3) $\blacksquare(\varphi \rightarrow \psi) \models \blacksquare\varphi \rightarrow \blacksquare\psi$;

(P4) $\blacksquare\varphi \models \varphi$;

(P5) $\blacksquare\varphi \models \blacksquare\blacksquare\varphi$;

(P1)-(P3) rappellent les propriétés modales de distribution sous conjonction, disjonction, et implication dans le système modal ***K*** ; (P4) correspond au principe de factivité du système ***T*** ; (P5) correspond au principe d'introspection positive du système ***S*4**.

On peut observer que les agents doxastiques ne s'accordent pas tous sur les mêmes propriétés exposées ci-dessus, comme le montre le tableau qui suit (où "PnC" symbolise la converse de la propriété numérotée *n*).

	$\boxbar$	$\boxminus$	$\boxtimes$	$\boxplus$
(P1)	✓	✓	×	✓
(P1C)	×	✓	✓	✓
(P2)	✓	✓	✓	✓
(P2C)	✓	×	×	✓
(P3)	×	×	×	×
(P3C)	✓	✓	✓	✓
(P4)	×	✓	✓	×
(P4C)	✓	×	×	✓
(P5)	✓	✓	✓	✓
(P5C)	✓	✓	✓	✓

Il va de soi que ces propriétés ne sont pas typiques de celles observées en logique épistémique modale. Par exemple, (P4C) est satisfaite par certains agents alors que la converse du principe modal de réflexivité ne l'est jamais dans un contexte modal. Cette dernière semble signifier qu'un agent de ce type croit tout ce qui est vrai ; mais une fois encore, il ne faut pas oublier que φ symbolise un énoncé d'information et ne décrit pas l'état du monde authentique. En d'autres termes, une formule telle que (P4C) exprime simplement la façon dont les agents se comportent vis-à-vis des informations positives dont ils disposent. Une autre raison plus sérieuse de rejeter le caractère modal des propriétés (P1)-(P5) est la validation systématique de (P5) et (P5C) : les quatre agents admettent aussi bien la formule initiale (P5) que sa converse, ce qui veut dire que l'itération des croyances est redondante dans $\boldsymbol{AR}_{\mathbf{4}_{\blacksquare}}$.

D'une certaine manière, on peut introduire un ordre entre les opérateurs de croyance en fonction de leur degré de tolérance, c'est-à-dire d'acceptation (de la vérité) des énoncés :[4]

[4]Les formules $\boxtimes\phi$-$\boxminus\phi$ et les formules $\boxbar\phi$-$\boxplus\phi$ sont logiquement *équivalentes*, au sens de la conséquence logique comme préservation de la vérité (relation tarskienne standard) : 1) tout ce qui est

$$\text{(TOL) } \boxtimes\varphi \dashv\models \boxminus\varphi \models \boxbar\varphi \dashv\models \boxplus\varphi$$

Il y a encore d'autres notions à développer dans ce domaine : citons par exemple les négations, semi-négations et oppositions qui concernent les relations de désaccord (partiel ou total) entre agents. Mais là n'est pas le but de cet article.[5] Revenons plutôt donc au sujet initial du texte, sur la base de notre nouveau cadre sémantique.

4 Retour à la justification et la quasi-vérité

Il est possible de retranscrire les idées de da Costa au sein de $\boldsymbol{AR}_{\mathbf{4}_{\blacksquare}}$. Pour ce faire, nous allons d'abord étudier la correspondance entre les deux opérateurs J et J' et les quatre opérateurs unaires étudiés jusqu'ici ; puis nous allons voir pourquoi le concept de quasi-vérité nécessite une extension des attitudes de croyance au sein de notre logique multivalente.

4.1 Justifications

Il est possible de prouver les deux résultats de correspondance qui suivent sur la base d'un calcul matriciel.

Théorème 4.1 *L'opérateur de justification J ne correspond pas a aucun opérateur de $\boldsymbol{AR}_{\mathbf{4}_{\blacksquare}}$.*

Théorème 4.2 *Les propriétés de l'opérateur de justification faible J' sont satisfaites par deux opérateurs de $\boldsymbol{AR}_{\mathbf{4}_{\blacksquare}}$: $\boxbar$, et $\boxplus$.*

Le premier résultat montre que les opérateurs de croyance ont un comportement ”non-normal”, par analogie avec les systèmes modaux de type qui ne sont satisfont pas le schéma de Kripke : $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$. Tous les opérateurs satisfont (P3C) mais aucun ne satisfait (P3), parce que les conditions dans lesquelles une implication est crue sont moins fortes que les conditions dans lesquelles une croyance en implique une autre. Cela dit, ce résultat repose moins sur le comportement des opérateurs unaires que sur celui du connecteur d'implication forte introduit dans $\boldsymbol{AR}_{\mathbf{4}}$: la propriété de distribution sous implication serait validée par le sceptique, si l'implication était définie au sens matériel qui produit les paradoxes

vrai pour les sceptiques est également vrai pour les négativistes, et réciproquement ; 2) tout ce qui est vrai pour les positivistes est également vrai pour les éclectiques, et réciproquement. En revanche, 3) tout ce qui est vrai pour les sceptiques et les négativistes est également vrai pour les positivistes et les éclectiques, mais 4) la réciproque n'est pas vraie.

[5] Voir à ce sujet [13].

de l'implication. Pour cette raison, la traduction des propriétés de $\boldsymbol{J}$ dans $\boldsymbol{AR}_{4_\blacksquare}$ exige de symboliser l'implication de da Costa par la relation de conséquence : la formule $\varphi \to \psi$ du système de da Costa doit être traduite par la formule $\varphi \models \psi$ dans $\boldsymbol{AR}_{4_\blacksquare}$, sous peine d'empêcher toute validation des propriétés de J et J'. Tel est le prix à payer pour le renforcement de l'implication.[6]

4.2 Généralisations

Nous avons vu dans une section précédente le lien étroit qui relie les concepts de quasi-vérité et de contingence. En effet, un énoncé φ est dit quasi-vrai si et seulement si il existe toujours un argument en faveur de φ et en défaveur de φ. Or aucun des quatre opérateurs de croyance considérés jusqu'ici ne semble être un candidat valable à son interprétation dans $\boldsymbol{AR}_{4_\blacksquare}$. Malgré une certaine connexion entre l'attitude de l'éclectiste et l'idée de contingence, cet agent ne satisfait pas l'idée de quasi-vérité lorsque l'énoncé d'information reçoit une valeur bivalente, c'est-à-dire seulement vraie ou seulement fausse. Cela conduit à la création d'autres attitudes épistémiques, sur la base du domaine de valuation quadrivalent admis par $\boldsymbol{AR}_4$.

Si l'on applique l'algorithme combinatoire, il existe un nombre maximal de $(m^n)^n = (4^4)^1 = 256$ opérateurs unaires au sein d'une sémantique à $m = 4$ valeurs de vérité telle que la nôtre. La majorité de ces opérateurs est de type *dégénéré*, dans la mesure où leurs valeurs sont fixées sans tenir compte du processus vérifonctionnel d'évaluation d'un énoncé. En revanche, on peut dégager de cet ensemble maximal deux catégories d'opérateurs unaires : les opérateurs *unilatéraux*, et les opérateurs *bilatéraux*. Dans les premiers, un seul type d'information (favorable, ou défavorable) est exigé afin d'accepter ou rejeter un énoncé ; dans les seconds, les deux types d'information sont exigés à la fois. Les quatre opérateurs étudiés jusqu'ici sont de type bilatéral, puisque leurs conditions de croyance se définissent toujours à partir des deux types d'information disponibles.

Pour construire les opérateurs unilatéraux $\blacksquare^u$, un seule type de justification de forme $a_x(\varphi) = y$ est requis afin de spécifier l'attitude de l'agent épistémique :

$$\begin{gathered}\text{(UBeL)}\\ A(\blacksquare^u\varphi) = \langle a_1(\blacksquare^u\varphi), a_2(\blacksquare^u\varphi)\rangle\\ a_1(\blacksquare^u\varphi) = 1 \text{ ssi } (a_x\varphi) = y\\ a_2(\blacksquare^u\varphi) = 1 \text{ ssi } (a_x\varphi) = y\end{gathered}$$

Chaque information $a_x(\varphi)$ pouvant être soit favorable ($y = 1$) soit défavorable ($y = 0$) à l'énoncé φ, il en résulte par combinaison un total de 16 opérateurs unilatéraux basés sur un seul type d'information (favorable, ou défavorable à l'énoncé

[6]Pour une explication du fonctionnement du conditionnel fort de $\boldsymbol{AR}_4$, voir [12].

en question). On peut relever au moins deux attitudes unilatérales susceptibles d'intéresser l'épistémologie social : la credulité, et l'incredulité.

Crédulité
L'agent crédule :
- croit que φ si et seulement si il existe un argument en faveur de φ ;
- croit que $\neg\varphi$ si et seulement si il existe un argument en défaveur de φ.

Ce type d'agent croit *tout ce qu'on lui dit*, sur la base des informations dont il dispose : il ne tire aucune conséquence des possibles contradictions existant entre elles, sans aucune autre exigence logique que de les accepter au cas par cas. A l'inverse, l'agent incrédule prend toujours le contre-pied des informations disponibles et symbolise l'attitude contradictoire à celle de l'agent crédule.

Incrédulité
L'agent incrédule :
- croit que φ si et seulement si il n'existe pas d'argument en faveur de φ ;
- croit que $\neg\varphi$ si et seulement si il n'existe pas d'argument en défaveur de φ.

L'opérateur de crédulité constitue le pendant épistémique de l'opérateur unaire d'*affirmation* en logique bivalente : c'est une fonction triviale qui ne change pas la valeur de ses arguments. Par voie de conséquence, l'opérateur d'incrédulité est le pendant épistémique de la *négation* classique : c'est une fonction qui inverse toutes les valeurs de ses arguments. Les conditions d'acceptation et de rejet de ces deux types d'agent peuvent être décrites comme suit, puis caractérisées dans la liste ci-dessous des matrices d'opérateurs unilatéraux :

Crédulité : $A(\boxdot\varphi) = \langle a_1(\varphi), a_2(\varphi)\rangle$
Incrédulité : $A(\boxbox\varphi) = \langle -(a_1(\varphi)), -(a_2(\varphi))\rangle$

Dans les quatre premiers opérateurs $\blacksquare^u_1$-$\blacksquare^u_4$, $x = \langle 1, 1\rangle$, tandis que $x = \langle 1, 2\rangle$ avec $\blacksquare^u_5$-$\blacksquare^u_8$, $x = \langle 2, 1\rangle$ avec $\blacksquare^u_9$-$\blacksquare^u_{12}$, et $x = \langle 2, 2\rangle$ avec $\blacksquare^u_{13}$-$\blacksquare^u_{16}$.

φ	$\blacksquare^u_1\varphi$	$\blacksquare^u_2\varphi$	$\blacksquare^u_3\varphi$	$\blacksquare^u_4\varphi$	$\boxdot\varphi$	$\blacksquare^u_6\varphi$	$\blacksquare^u_7\varphi$	$\boxbox\varphi$
11	11	10	01	00	11	10	01	00
10	11	10	01	00	10	11	00	01
01	00	01	10	11	01	00	11	10
00	00	01	10	11	00	01	10	11

$\blacksquare^u_9\varphi$	$\blacksquare^u_{10}\varphi$	$\blacksquare^u_{11}\varphi$	$\blacksquare^u_{12}\varphi$	$\blacksquare^u_{13}\varphi$	$\blacksquare^u_{14}\varphi$	$\blacksquare^u_{15}\varphi$	$\blacksquare^u_{16}\varphi$
11	10	01	00	11	11	01	00
01	00	11	10	00	10	11	11
10	11	00	01	11	01	00	11
00	01	10	11	00	00	10	00

Dans le cas des opérateurs *bilatéraux* $\blacksquare^b$, les deux conditions $a_1(\varphi)$ et $a_2(\varphi)$ sont exigées pour déterminer la croyance des agents, la différence entre eux résidant dans la combinaison des réponses y obtenues.

(BBeL)

$$A(\blacksquare^b\varphi) = \langle a_1(\blacksquare^b\varphi), a_2(\blacksquare^b\varphi)\rangle\,;$$
$$a_1(\blacksquare^b\varphi) = 1 \text{ ssi } a_1(\varphi) = y \text{ et/ou } a_2(\varphi) = y\,;$$
$$a_2(\blacksquare^b\varphi) = 1 \text{ ssi } a_1(\varphi) = y \text{ et/ou } a_2(\varphi) = y.$$

Il existe 64 opérateurs bilatéraux distincts : quatre fois plus que les opérateurs unilatéraux, en raison des quatre combinaisons possibles de justification pour chacun d'eux (et-et, et-ou, ou-et, ou-ou) d'informations favorables et défavorables.

Parmi l'ensemble de ces opérateurs unilatéraux et bilatéraux, certains peuvent être qualifiés d'irrationnels parce qu'ils ne suivent pas les normes admises de l'épistémologie, telles que celle consistant à baser ses croyances sur les informations dont on dispose. On peut exprimer cette idée sous l'appellation d'un critère de confiance par défaut, en cas d'absence de témoignage adverse :

(RAT) Un opérateur $\blacksquare$ de $\boldsymbol{AR_4}$ est *rationnel* seulement si l'agent fait confiance à l'information minimale dont il dispose, que celle-ci soit favorable ou défavorable à l'énoncé :

$$a_i(\varphi) = 1 \text{ si } a_i(\varphi) = 1 \text{ ou } a_j(\varphi) = 0$$

L'information disponible est une information ni nécessaire, ni suffisante par définition à la croyance d'un agent. En revanche, elle peut le devenir en fonction de ses critères de justification épistémique. Le sceptique est un agent bilatéral rationnel, dont la rigueur des critères épistémiques oblige à tenir compte des deux types d'information possibles. L'agent incrédule est, quant à lui, un agent irrationnel, dans la mesure où il se dédit toujours des informations dont il dispose.

Une dernière catégorie d'agents peut être mentionnée, qui figure dans la liste maximale des 256 opérateurs unaires quadrivalents : celle des agents inintelligibles. Situés hors de la liste des opérateurs unilatéraux et bilatéraux, chacun d'eux incarne un opérateur *dégénéré* tel que l'attitude n'est pas toujours (sinon jamais) fonction du type d'information en présence. L'attitude intelligible renvoie, au contraire, au critère de prédictibilité :

(INT) Un opérateur $\blacksquare$ de $\boldsymbol{AR_4}$ est *intelligible* ssi l'agent détermine son attitude en fonction de l'information minimale dont il dispose :

$$a_i(\blacksquare\varphi) = 1 \text{ seulement si } f(a_i(\varphi))$$

Les agents rationnels et irrationnels sont tous deux des cas particuliers d'agents intelligibles, puisque leurs attitudes dépendent toujours des informations en présence. Et bien que l'agent irrationnel ne se conforme pas aux critères de rationalité épistémique, son attitude est cependant prévisible et, en ce sens, intelligible.

$$a_i(\blacksquare\varphi) = 1 \text{ seulement si } a_i(\varphi) = 0 \text{ et } a_j(\varphi) = 1$$

Les deux cas extrêmes d'attitude inintelligible sont ceux où l'agent accepte tous les énoncés ou rejette tous les énoncés, sans exception et sans condition. Appelons-les *relativistes* et *nihilistes*, respectivement.

Relativiste : $\blacksquare_1$
L'agent relativiste :
- croit que φ s'il existe un argument en faveur de φ ou s'il n'existe pas d'argument en faveur de φ ;
- croit que $\neg\varphi$ s'il existe un argument en défaveur de φ ou s'il n'existe pas d'argument en faveur de φ.

Ce type d'agent croit aussi tout ce qu'on lui dit, tout comme le crédule. Mais ses exigences épistémiques sont encore plus faibles, puisque la simple absence d'argument favorable (ou défavorable) ne l'empêche pas de prendre parti pour (ou contre) l'énoncé. A l'opposé de lui, l'agent nihiliste ne croit à rien parce que ses critères de croyance violent le critère logique de cohérence.

Nihiliste : $\blacksquare_{256}$
L'agent nihiliste :
- croit que φ s'il existe un argument en faveur de φ et s'il n'existe pas d'argument en faveur de φ ;
- croit que $\neg\varphi$ s'il existe un argument en défaveur de φ et s'il n'existe pas d'argument en faveur de φ.

A noter que la cohérence n'est pas synonyme de *consistance* : un agent est consistant s'il n'accepte pas un énoncé φ et sa négation $\neg\varphi$ en même temps. L'agent éclectiste est donc un agent inconsistant, mais cohérent vis-à-vis de chaque énoncé particulier φ : il n'accepte pas ce qu'il rejette tout à la fois, et réciproquement.[7]

Ces deux attitudes inintelligibles sont les contreparties épistémiques des opérateurs unaires de la *tautologie* $\top$ et de l'*antilogie* $\bot$, en logique bivalente.

φ	$\blacksquare^u_1\varphi$	$\blacksquare^u_{256}\varphi$
11	11	00
10	11	00
01	11	00
00	11	00

4.3 Quasi-vérité

Plutôt que de dresser la liste exhaustive des matrices caractéristiques des croyances bilatérales, notre attention va porter sur l'un de ces opérateurs parce qu'il semble correspondre au concept de quasi-vérité. Si la quasi-vérité est une vérité contingente toujours réfutable par définition, cela veut dire qu'il doit exister dans chaque cas un argument et un contre-argument en faveur de l'énoncé φ.

Quasi-vérité

Un agent épistémique :

- croit que φ est *quasi-vrai* si et seulement si il existe un argument en faveur de φ et un argument en défaveur de φ ;

- ne croit pas que φ est *quasi-vrai* si et seulement si il n'existe pas d'argument en faveur de φ ou il n'existe pas d'argument en défaveur de φ.

La matrice caractéristique de la quasi-vérité peut être symbolisée et définie comme suit :

$A(\boxslash\varphi) = \langle a_1(\varphi) \sqcap a_2(\varphi), -(a_1(\varphi) \sqcup -(a_2(\varphi))\rangle$

Le concept de *quasi-vérité* satisfait ou non les propriétés suivantes :

[7] Sur la différence entre consistance et cohérence, voir [14].

φ	$\boxslash\varphi$
11	10
10	01
01	01
00	01

	$\boxslash$
(P1)	$\times$
(P1C)	$\checkmark$
(P2)	$\checkmark$
(P2C)	$\times$
(P3)	$\times$
(P3C)	$\times$
(P4)	$\checkmark$
(P4C)	$\times$
(P5)	$\times$
(P5C)	$\times$

Deux propriétés supplémentaires intéressent la quasi-vérité, en raison de son comportement vis-à-vis de la vérité :

(P6) $\boxslash\varphi \models \boxslash\neg\varphi$;

(P6C) $\boxslash\neg\varphi \models \boxslash\varphi$.

La quasi-vérité est un concept ambivalent et, pour cette raison, il est mieux adapté à une logique à plusieurs valeurs de vérité. Notre sémantique est en mesure de répondre également à deux questions posées par Otávio Bueno dans [3], au sujet de l'opérateur de quasi-verité (ou 'QV' symbolise cet opérateur) :

(Q1) $QV\varphi \wedge QV\neg\varphi \models QV(\varphi \wedge \neg\varphi)$?

(Q2) $QV\varphi \leftrightarrow \varphi$?

Selon Bueno, ni (Q1) ni (Q2) ne sont valables pour le concept de *quasi-vérité*. Notre sémantique quadrivalente donne une réponse positive à la première question, en raison de la validité de (P1). Rappelons que ce resultat n'a rien de contre-intuitif, contrairement à l'objection posée par Dubois et discutée plus tôt (section 3.1, page 5). En revanche, notre réponse négative à la seconde question posée par Bueno repose sur une explication entièrement différente de la sienne. En effet, (Q2) n'est pas valide dans $\boldsymbol{AR}_{4_{\blacksquare}}$ parce que la vérité de φ n'implique pas sa quasi-vérité (malgré la validité de (P4)). Pour Bueno, (Q2) n'est pas valable parce que la quasi-vérité de φ n'implique pas de vérité. Cette divergence de résultat s'explique par la nature justifiée des croyances dans $\boldsymbol{AR}_{4_{\blacksquare}}$, par opposition à l'analyse de Bueno.

Un autre moyen de caractériser la quasi-vérité consisterait à "temporaliser" l'ensemble des croyances de l'agent correspondant, afin d'expliquer que tout argument admis à un certain moment doit pouvoir être rejeté en principe par la suite. Une telle distinction temporelle a déjà été effectuée en logique modale, notamment. Mais nous ne voyons ici aucune raison de préférer cette méthode à notre cadre multivalent, la notion de quasi-vérité reposant davantage sur la diversité des arguments que sur le moment ou un changement de point de vue se produit.

5 Conclusion

Cet article a voulu mettre en évidence la distinction fondamentale entre les concepts d'information et de croyance : sans cette distinction de base, il est impossible de clarifier les débats relatifs à la frontière entre ontologie et épistémologie. De plus, l'utilisation d'une sémantique quadrivalente produit une épistémologie formelle capable de décrire une variété d'agents dont les critères de justification sont variables. A contre-courant des logiques modales et de la sémantique relationnelle qui les caractérise, notre sémantique repose sur un cadre dialogique qui interprète les réponses de locuteurs sous forme de valeurs logiques structurées. Notre objectif a été de montrer les résultats suivants :

1. Les concepts d'information, de justification et de croyance peuvent être corrélés les uns aux autres afin de produire une logique des croyances multivalente $\boldsymbol{AR}_{4_{\blacksquare}}$;
2. La différence entre les types d'agents repose sur une différence dans les critères qualitatifs de justification des croyances (par opposition aux critères quantitatifs de la sémantique relationnelle, basée sur des opérateurs analogues aux quantificateurs) ;
3. Il existe une différence fondamentale entre les notions de rationalité et d'intelligibilité chez un agent : à la différence du second, le premier a un comportement prévisible qui peut être décrit par un opérateur vérifonctionnel ;
4. L'opérateur de quasi-vérité trouve son expression dans $\boldsymbol{AR}_{4_{\blacksquare}}$.

D'autres thèmes seront abordés à la suite de cet article, afin de développer notre analyse logique de l'épistémologie sociale au sein de $\boldsymbol{AR}_{4_{\blacksquare}}$:

– Combinaison d'opérateurs de croyance distincts : que signifie et quelles sont les propriétés logiques d'une formule itérée telle que $\boxtimes\boxminus$, par exemple ?

– Traduction des systèmes logiques : comment reformuler les systèmes multivalents standards au sein du cadre dialogique $\boldsymbol{AR_{m^n}}$?
– Typologie des agents : peut-on caractériser des agents propres aux systèmes de logique classique, intuitionniste, et paraconsistante ?
– Oppositions non-standards : peut-il y avoir d'autres types d'opposition entre agents épistémiques, au-delà des quatre oppositions aristotéliciennes ?

Notons enfin qu'aucune preuve de complétude n'a été produite dans cet article. Notre recherche a concerné uniquement les aspects sémantiques de la croyance justifiée, et la recherche d'un système d'axiomes susceptible de caractériser les énoncés d'information puis de justification sera l'objet d'un travail ultérieur.

Références

[1] ARTEMOV, S. (2008). The logic of justification. In : *The Review of Symbolic Logic*, vol.1, 477-513.

[2] BELNAP, N. (1977). A useful four-valued logic. In : *Modern Uses of Multiple-Valued Logic*, J.M. Dunn and G. Epstein (eds.), Dordrecht : D. Reidel Publishing , 8-37.

[3] BUENO, O. (1999). Truth, quasi-truth and paraconsistency. In : *Contemporary Mathematics*, vol.235, 275-293.

[4] BUENO, O ; DE SOUZA, E. (1996). The concept of quasi-truth. In : *Logique et Analyse*, vol.153-154, 183-199.

[5] CONIGLIO, M ; SILVESTRINI, L. (2014). An alternative approach for quasi-truth. *Logic Journal of IGPL*, vol. 22(2), 387-410.

[6] COSTA-LEITE, A. (2014). Lógicas da justificação e quase-verdade. In : *Principia : an international journal of epistemology*, vol. 18(2), 175-186.

[7] DA COSTA, N. (1999). *O conhecimento científico*, 2a. edição, São Paulo : Discurso Editorial.

[8] DA COSTA, N. (1986). Pragmatic Probability. In *Erkenntnis*, vol.25, 141-162.

[9] DUBOIS, D. (2008). On ignorance and contradiction considered as truth-values. In : *Logic Journal of the IGPL*, vol. 16, 195-21.

[10] DUGUNDJI, J. (1940). Note on a property of matrices for Lewis and Langford's calculi of propositions. In : *The Journal of Symbolic Logic*, vol. 5 : 150-151.

[11] GINSBERG, M. (1988). Multivalued logics : A Uniform Approach to Reasoning in Artificial Intelligence. In : *Computational intelligence*, vol. 4 : 265-316.

[12] SCHANG, F. (201X). A four-valued logic of strong implication. Submitted.

[13] SCHANG, F. (201X). Epistemic pluralism. À paraître in *Logique et Analyse*.

[14] SCHANG, F. (2009). Relative charity. *Revista Brasileira de Filosofia*. Vol. 233. 159-172.

[15] WANSING, H ; BELNAP, N. (2010). Generalized truth-values : A reply to Dubois. *Logic Journal of the IGPL*, vol. 18(6), 921-935.