Le paradoxe de Fitch dans l'œil du positiviste

Sur le problème de l'inconnaissabilité

Paul Égré*

(Institut Jean-Nicod, CNRS)

Résumé : Toute vérité est-elle connaissable en principe ? Une réponse négative à cette question suit d'un argument logique dû à F. Fitch, voisin du paradoxe de Moore, et connu sous le nom de *paradoxe de la connaissabilité*. Le paradoxe de Fitch constitue un obstacle à la conception anti-réaliste de la vérité, et plus généralement, semble-t-il, à l'idéal positiviste d'après lequel toute vérité devrait nous être accessible en principe. Dans cet article, j'examine différentes tentatives pour préserver le principe selon lequel toute vérité est connaissable, chacune inspirée d'une forme propre d'anti-réalisme (Dummett, Tennant, Edgington). Ces différentes approches sont comparées à une conception réaliste du positivisme, évoquée récemment par Burgess, qui postulerait seulement que toute vérité *nécessaire* est connaissable. Selon cette conception, certaines vérités contingentes sont effectivement inconnaissables, mais l'accessibilité de principe des vérités nécessaires demeure suffisante pour garantir la confiance positiviste dans la science.

Abstract: Are all truths knowable? A negative answer to this question follows from a logical argument related to the Moore Paradox and due to F. Fitch, also known as Fitch's Knowability Paradox. Fitch's paradox is widely considered to be an obstacle to the anti-realist conception of truth, but even for a realist, it appears to threaten the positivist faith in the accessibility of all truths to our minds. In this paper, I first review different strategies to circumvent the paradox, each of them inspired by a different form of anti-realism (Dummett, Tennant, Edgington). I then compare these approaches to a realist version of positivism, discussed recently by Burgess, which would only postulate that all *necessary* truths are knowable. On that view, some contingent truths are indeed unknowable, but the idea that all necessary truths are within our reach remains sufficient to maintain the positivist faith in scientific knowledge.

Toute vérité est-elle connaissable en principe ? Une réponse négative à cette question peut paraître difficilement admissible à première vue. Car si l'on admet qu'il existe des vérités inconnaissables, il peut sembler qu'on limite indûment l'étendue de notre connaissance et la portée de nos facultés mentales. Nous savons pourtant, à y réfléchir d'un peu plus près, qu'il existe des sons que nous ne pouvons pas entendre, de même que des couleurs que nous ne pouvons pas voir à l'œil nu[1]. Toutefois, si nous ne pouvons pas percevoir de telles couleurs et de tels sons directement, du moins pouvons-nous les mettre en évidence par d'autres moyens que les ressources immédiates de notre appareil perceptif. Que faut-il entendre alors par « vérité inconnaissable » ? On entend manifestement par là une vérité telle qu'aucun esprit humain, doué de raison, ne puisse savoir qu'elle est vraie. Imaginons, par exemple, que personne n'ait compté le nombre de volumes consultés dans la salle W de la Bibliothèque nationale de France le 15 juin 2004, entendant par là les volumes rangés sur les rayons de la

* Je remercie Joe Salerno, Denis Bonnay, Mikaël Cozic, Michael Fara, John MacFarlane, Julien Murzi, Philippe Schlenker, Carlo Proietti, Marie Guillot, Benjamin Spector et Sylvain Bromberger pour les discussions stimulantes que j'ai pu avoir avec chacun d'eux au cours de la rédaction de cet essai. Je remercie particulièrement D. Bonnay, P. Schlenker, B. Spector et S. Bromberger pour leurs commentaires détaillés. Je remercie également très vivement Stéphane Chauvier pour son invitation et pour ses encouragements en vue de l'élaboration de ce texte. Enfin je tiens à exprimer ma gratitude à John P. Burgess et Timothy Williamson, dont les travaux sur le paradoxe de Fitch ont particulièrement éclairé ce travail, et attiré mon attention sur le statut des vérités nécessaires.

[1] C'est le cas des infra-rouges et des ultraviolets, ou encore, pour ce qui est des sons, des infrasons (<20Hz) et des ultrasons (>20kHz).

salle W. Du temps a passé, nous voilà en 2007, et il est désormais trop tard pour rassembler des indices concluants[2]. L'une des deux propositions « au plus 100 volumes ont été consultés en salle W le 15 juin 2004 » ou « plus de 100 volumes ont été consultés en salle W le 15 juin 2004 » est vraie. Pourtant, nul n'est désormais en mesure de déterminer laquelle de ces deux propositions est vraie. Il existe par conséquent au moins une de ces deux propositions qui est vraie, et dont la vérité est inconnaissable.

Par « inconnaissable », on entend ici qu'il n'existe *plus* aucun moyen de connaître cette vérité *désormais*. Malgré cela, on reste confiant sur le fait que le 14 juin 2004, il était *a priori* possible de déterminer laquelle des deux propositions « au plus 100 volumes seront consultés en salle W le 15 juin 2004 » et sa négation allait être vraie. On pourrait alors renforcer la question de la façon suivante : existe-t-il une vérité telle qu'*à aucun moment du temps*, il ne soit en principe possible de la connaître ? Un philosophe tel que Pascal concevrait sans doute la réponse comme positive : la proposition « Dieu existe » est selon Pascal vraie, mais inconnaissable d'un esprit fini par la raison à aucun moment du temps. Imaginons en effet que nécessairement, si Dieu existe, alors Dieu est tel qu'on ne pourra jamais avec aucune certitude savoir qu'il existe[3]. On aurait là l'exemple d'une vérité inconnaissable.

L'hypothèse d'un Dieu caché est concevable et cohérente mais elle n'établit pas qu'il existe une vérité inconnaissable : elle accrédite seulement la possibilité qu'il existe une vérité de cette nature[4]. La littérature logique suggère toutefois qu'il puisse exister de telles vérités inconnaissables. Il existe en effet un argument indirect en faveur de l'existence de vérités inconnaissables : l'argument, dû au logicien Frederic Fitch et désormais connu sous le nom de « paradoxe de Fitch », établit que si toute vérité est connaissable en principe, alors toute vérité doit d'ores et déjà être connue[5]. Pour quiconque admet qu'il existe des vérités que nous ne connaissons pas présentement, l'argument de Fitch suggère par conséquent qu'il existe des vérités inconnaissables.

L'interprétation de l'argument de Fitch fait l'objet d'un désaccord persistent entre réalistes et anti-réalistes sur le statut de la vérité. Pour les partisans de l'anti-réalisme et du vérificationnisme (la thèse selon laquelle la vérifiabilité de principe des énoncés est un critère de signification), le paradoxe de Fitch suggère non pas d'abandonner le principe de connaissabilité, mais plutôt d'en donner une formulation plus raffinée. La plupart des travaux consacrés au paradoxe de Fitch sont du même coup des tentatives pour préserver le

[2] L'exemple est inspiré de Williamson, *Knowledge and its Limits*, chap. 12, Oxford, 2000.

[3] Voir Pascal, *Pensées*, 242. Pour Pascal, le caractère caché de Dieu, insaisissable de la seule raison, est précisément l'une des raisons de croire: "On ne dit point que ceux qui cherchent le jour en plein midi, ou de l'eau dans la mer, en trouveront; ainsi il faut bien que l'évidence de Dieu ne soit pas telle dans la nature. Aussi elle nous dit ailleurs: *Vere tu es deus absconditus*". La position pascalienne est proche en cela d'une conception mystique de l'existence de Dieu – le mysticisme pouvant être compris, plus généralement, comme l'idée selon laquelle il existe des vérités essentiellement inaccessibles.

[4] Une telle vérité serait du moins concevable, bien que non connaissable. Toute vérité est-elle alors concevable? Si la réponse à cette question est négative, alors il devrait s'ensuivre *a fortiori* qu'il n'est pas vrai que toute vérité est connaissable.

[5] Voir F. Fitch (1963), "On the Logical Analysis of some Value Concepts", *The Journal of Symbolic Logic*, 28, pp. 135-142, où le résultat de Fitch apparaît comme Théorème 5. Fitch crédite un rapporteur anonyme du *JSL* pour l'inspiration de ce résultat. Joe Salerno et Julien Murzi ont récemment établi qu'il s'agissait d'Alonzo Church. Les rapports anonymes de Church et sa correspondance avec l'éditeur du *JSL* à propos de l'article de Fitch sont à paraître dans le volume édité par J. Salerno, *New Essays on the Knowability Paradox*. Voir aussi, dans le même volume, l'excellent article de J. Salerno, "Knowability Noir: 1945-1963", sur la genèse du paradoxe et les liens méconnus avec la théorie de la valeur proposée par Fitch. Pour une vision synthétique de la littérature désormais abondante sur le paradoxe de Fitch, les textes de références sont l'article de B. Brogaard et J. Salerno "Fitch's Paradox of Knowability", *The Stanford Encyclopedia of Philosophy (Summer 2004 Edition)*, Edward N. Zalta (ed.), http://plato.stanford.edu/archives/sum2004/entries/fitch-paradox/, et le livre de J. Kvanvig, *The Knowability Paradox*, Oxford, 2006. Pour une introduction en français au paradoxe de Fitch, cf. J. Vidal-Rosset, *Qu'est-ce qu'un Paradoxe ?*, Vrin, 2004.

vérificationnisme sous une forme ou sous une autre. Mais ce faisant, on néglige en général de considérer que pour un réaliste, le paradoxe de Fitch impose également une limitation, cette fois de l'idéal positiviste selon lequel toute vérité est en principe ouverte à la connaissance. Quel fondement reste-t-il, dans ce cas, à l'attitude positiviste elle-même ? L'objet de cet essai est de répondre à cette question. Pour cela, nous examinons en détail différents amendements touchant la formulation du principe de connaissabilité, proposés respectivement par Michael Dummett, Neil Tennant et Dorothy Edgington. Comme nous le verrons, aucun de ces amendements ne permet de préserver sans reste le principe de connaissabilité initial. Pour le réaliste, cependant, nous verrons qu'il existe différentes façons de préciser la position positiviste, tout en prenant acte de l'idée que toute vérité n'est pas connaissable.

Avant d'entrer dans le vif du sujet, deux clarifications préliminaires s'imposent, d'une part sur ce qu'il convient d'appeler « vérité » lorsque nous nous demandons si toute vérité est connaissable en principe, et d'autre part sur ce que doit signifier « connaissable en principe ». S'agissant du terme de « vérité », dans ce qui suit il s'agira toujours de propositions exprimables dans un langage dénombrable (tel que le Français), de sorte que nous nous demandons si toute vérité exprimable dans un langage tel que le Français est connaissable en principe. Sans cette restriction, on s'expose directement à la remarque selon laquelle certaines vérités doivent être inconnaissables parce qu'il y a alors nécessairement plus de vérités possibles que de vérités exprimables[6]. Par « connaissable en principe », par ailleurs, nous entendons ici : abstraction faite des limitations physiques et computationnelles qui s'appliquent de toute façon à la connaissance, et abstraction faite de l'imprécision de nos concepts[7]. Par exemple, il se peut qu'il soit pratiquement impossible de déterminer à quoi est égal le produit des cent mille premiers nombres entiers, du fait de l'explosion combinatoire à quoi donnerait lieu un tel calcul. Il se peut aussi qu'il soit pratiquement impossible de déterminer quel est le nombre de galaxies dans l'univers, du fait du vague des termes de « galaxie » et « univers », et compte tenu de l'immensité du domaine de recherche. Bien qu'il s'agisse là de limites drastiques à la connaissance, nous les considérons ici comme des limitations de fait. En parlant de connaissabilité en principe, nous entendons ici « ce qui est connaissable sans qu'il en résulte de contradiction », soit le connaissable au sens logique.

---

[6] Voir J. P. Burgess, « Can Truth Out? », à paraître in J. Salerno (ed.), *New Essays on the Knowability Paradox*. Un langage comme le Français ne peut produire qu'un ensemble dénombrable d'énoncés. En revanche, l'ensemble des nombres réels est indénombrable. En ce sens, il y a plus d'énoncés vrais de la forme « $x > 2$ », lorsque la variable x parcourt les nombres réels, qu'il n'est possible d'en exprimer dans un langage tel que le Français (en y incluant les expressions et symboles mathématiques, qui restent en nombre dénombrable).

[7] Sur la question des limitations physiques et computationnelles à la connaissance, cf. H. Zwirn, *Les limites de la connaissance*, Odile Jacob, 2000, et D. Harel, *Compters Ltd. : What They Really Can't Do*, Oxford University Press, 2000. Je n'aborderai pas ici le problème de l'insolubilité en mathématiques, notamment parce qu'il s'agit d'une question en partie orthogonale à celle qui nous occupe. Par exemple, comme le souligne M. Minsky dans son ouvrage *Computation : Finite and Infinite Machines*, Prentice-Hall, 1967, pp. 153-154 et p. 164, la preuve de l'indécidabilité du problème de l'arrêt des machines de Turing prouve uniquement qu'il n'existe pas de machine de Turing qui puisse décider de toute autre machine si elle s'arrêtera ou pas. Cela n'implique en aucun cas qu'il existerait une machine de Turing dont il soit *absolument* impossible de décider si elle s'arrêtera ou pas. De la même façon, les résultats d'incomplétude de Gödel n'impliquent nullement qu'il existerait des vérités mathématiques particulières qu'il soit impossible de prouver *par quelque moyen que ce soit*. En ce sens, ces résultats ne vont pas à l'encontre de l'idée selon laquelle toute vérité est connaissable en principe (bien qu'ils aillent manifestement à l'encontre de l'idée selon laquelle toute vérité serait connaissable par une méthode uniforme). Comme le remarque en outre Minsky, on pourrait à la rigueur concevoir qu'il existe une machine de Turing qui ne s'arrête pas et dont il soit impossible de prouver qu'elle ne s'arrêtera pas. Mais on ne peut prouver qu'il soit impossible de prouver si une machine donnée s'arrêtera ou pas. Car si elle doit s'arrêter, on peut le constater en principe, et cela contredirait la preuve que le problème est insoluble. Si donc il existait une telle preuve d'insolubilité, elle devrait impliquer que la machine ne s'arrêtera pas, et par là même prouver que le problème a une solution, contrairement à l'hypothèse.

1. Le paradoxe de Fitch et le paradoxe de Moore

Le paradoxe de Fitch est dérivable dans un système de logique modale faible comportant un opérateur de connaissance, représenté par le symbole K, et un opérateur de possibilité, représenté par le symbole $\Diamond$. Si p représente l'énoncé « il pleut », alors $\Diamond p$ représente l'énoncé « il est possible qu'il pleuve ». De la même façon, Kp représente l'énoncé « p est connu » (ou encore « je sais que p »), et $\Diamond Kp$ l'énoncé « p est connaissable » (« je peux savoir que p »). Muni de ces ressources, le principe selon lequel toute vérité est connaissable peut s'énoncer de la façon schématique suivante[8] :

(V) $p \rightarrow \Diamond Kp$ (« si p, alors je peux savoir que p »)

Pour comprendre comment procède le paradoxe de Fitch, il est utile de dire d'abord un mot du paradoxe de Moore. Le paradoxe de Moore n'est pas un paradoxe logique, mais correspond plutôt à un paradoxe pragmatique ou performatif. Moore a fait remarquer qu'il est contradictoire d'affirmer simultanément : « il pleut, mais je ne crois pas qu'il pleut ». La contradiction provient du fait qu'en affirmant « il pleut », un locuteur sincère implique qu'il croit ce qu'il dit ; mais en affirmant « je ne crois pas qu'il pleut », le locuteur contredit ce à quoi il s'engage implicitement en affirmant « il pleut ». De la même façon, si l'on considère un énoncé comme « il pleut et je ne sais pas qu'il pleut », on a typiquement affaire à un énoncé qu'il semble impossible de connaître sans contradiction. Supposons en effet qu'il pleuve, et que j'en sois ignorant. Puis-je savoir qu'il pleut et que je ne sais pas qu'il pleut ? Cela implique manifestement que je sache qu'il pleuve, et aussi que je sache que je ne sais pas qu'il pleut. Mais pour savoir que je ne sais pas qu'il pleut, il faut qu'il soit vrai que je ne sache pas qu'il pleut. Si donc il pleut et que je ne sais pas qu'il pleut, il m'est impossible de savoir ce fait complexe sans contradiction. Pour aboutir à cette conclusion, les seules hypothèses en jeu sont que savoir exprime une attitude véridique (on ne peut savoir ce qui est faux), et que pour savoir une conjonction, il faut savoir chacun des conjoints. Plus formellement, l'argument peut être présenté comme suit :

(1)	$K(p \wedge \neg Kp)$	(hypothèse)	(« je sais qu'il pleut et que je ne sais pas qu'il pleut »)
(2)	$Kp \wedge K\neg Kp$	(par distributivité de K sur $\wedge$)	(« je sais qu'il pleut et je sais que je ne le sais pas »)
(3)	$Kp \wedge \neg Kp$	(par la véridicité de K)	(« je sais qu'il pleut et je ne sais pas qu'il pleut »)

Le paradoxe de Fitch, qui s'appuie sur le paradoxe de Moore, suppose ainsi que l'opérateur de connaissance est véridique et satisfait la contrainte de distributivité sur la conjonction, et en outre que l'opérateur de possibilité est *normal* : ce qui signifie que de $A \rightarrow B$ on peut inférer $\Diamond A \rightarrow \Diamond B$, et qu'un énoncé contradictoire (comme $Kp \wedge \neg Kp$), qu'on notera $\perp$, est impossible, c'est à dire que $\Diamond\perp \rightarrow \perp$ (« l'absurde est impossible »). Supposons dans ce cas que (V) soit vrai sans restriction, alors de (V) il suit :

(4) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$ (« s'il pleut et que je ne le sais pas, je peux savoir qu'il pleut et que je ne le sais pas »)

---

[8] Dans ce qui suit, « $\rightarrow$ » symbolise le conditionnel matériel « si…alors… », « $\neg$ » symbolise la négation, et « $\wedge$ » la conjonction.

or en vertu de (1)-(3) :

(5) $K(p \wedge \neg Kp) \rightarrow \bot$ (« il est contradictoire que je sache qu'il pleut et que je ne le sais pas »)

et par conséquent :

(6) $\Diamond K(p \wedge \neg Kp) \rightarrow \Diamond\bot$ (« s'il est possible que je sache qu'il pleut et que je ne le sais pas, alors l'absurde est possible)

mais puisque l'absurde est impossible, il en résulte :

(7) $\Diamond K(p \wedge \neg Kp) \rightarrow \bot$ (« je ne peux pas savoir qu'il pleut et que je ne le sais pas »)

De (4)-(7), il suit donc que :

(8) $p \wedge \neg Kp \rightarrow \bot$ (« il est contradictoire qu'il pleuve et que je ne le sache pas)

ce qui équivaut à :

(9) $\neg(p \wedge \neg Kp)$ (« il n'est pas vrai qu'il pleut et que je ne le sais pas »)

ou de façon équivalente en logique classique :

(10) $p \rightarrow Kp$ (« s'il pleut, alors je sais qu'il pleut »)

Si donc on admet (V), qui énonce que toute vérité est connaissable, alors il suit de ce principe et des hypothèses que nous avons faites sur les notions de connaissance et de possibilité que toute vérité est connue. La conclusion (10) semble à première vue beaucoup trop forte, cependant. Se peut-il en effet qu'il n'existe pas plus de vérités connaissables, à un moment donné, que de vérités connues ? Il semble que seule une forme radicale d'anti-réalisme sur le statut de la vérité nous autorise à une telle conclusion[9]. L'anti-réaliste soutient que la vérité n'existe pas indépendamment de notre appréhension de la vérité. Pour un anti-réaliste radical, il est donc concevable qu'il n'existe pas plus de vérités connaissables à un moment donné que de vérités connues.

L'antiréalisme radical semble cependant difficilement tenable, comme le remarque par exemple Williamson : si, à l'heure qu'il est, personne ne sait si le nombre des livres consultés en salle W le 15 juin 2004 était inférieur à 100, ou supérieur à 100, il semble pourtant qu'exactement une des deux propositions correspondantes doive être vraie. Celle des deux propositions qui est vraie est présentement inconnue, mais on peut supposer qu'elle est connaissable : par exemple, qu'une recherche nous permettra de décider laquelle des deux propositions est vraie. Pour donner un autre exemple : à l'heure qu'il est j'ignore si j'ai laissé

[9] Ce pourrait être, à la rigueur, une forme de platonisme épistémologique, fidèle à l'enseignement du *Ménon*, d'après lequel toute vérité que nous venons à connaître doit nous être déjà connue, de façon implicite. Dans l'optique socratique du *Ménon*, rien n'est connaissable qui ne soit déjà connu préalablement. Dans cette perspective (peu vraisemblable), il suit directement de la thèse que toute vérité est connaissable que toute vérité est déjà connue.

une pièce d'un euro dans la poche de mon blouson. Soit il y a une pièce d'un euro dans la poche de mon blouson, soit il n'y en a pas. L'une de ces deux propositions est manifestement vraie, et inconnue de moi. Mais je puis cependant la connaître : il suffit que je fouille les poches de mon blouson. Supposons donc qu'à 11h je fouille les poches de mon blouson et réalise qu'il s'y trouve une pièce d'un euro. A 10h, cette vérité m'était encore inconnue. A 10h, l'énoncé « il y a une pièce d'un euro dans mon blouson » était donc vrai, inconnu de moi, mais bel et bien connaissable.

Si donc on refuse l'antiréalisme radical, il peut être tentant, au vu du paradoxe de Fitch, d'adopter l'attitude exactement opposée, c'est-à-dire de rejeter le principe de connaissabilité en même temps que la conclusion qui en découle. Dans ce cas, on doit admettre qu'il est faux que toute vérité est connaissable, et par là même concéder qu'il existe des vérités inconnaissables. N'est-ce pas toutefois, à bien y réfléchir, l'enseignement même du paradoxe de Moore ? Le paradoxe de Moore enseigne qu'au moment où il pleut et où j'ignore qu'il pleut, il m'est impossible de savoir ce fait même, car il faudrait déjà que je sache qu'il pleut. Une autre façon de le présenter est la suivante : si je ne sais pas que p est vrai, alors je ne peux pas savoir, *au même moment*, que ce que j'ignore est vrai. Il faut noter ici une subtilité : le paradoxe de Moore ne dit pas que, si je suis incertain quant à la vérité de p, il m'est impossible de savoir que je suis incertain, et par là même que je suis ignorant de la vérité. Ce dernier principe, qu'on appelle couramment le principe d'*introspection négative*, s'énonce sous la forme schématique suivante :

$\neg Kp \rightarrow K\neg Kp$ (« si je ne sais pas qu'il pleut, alors je sais que je le sais pas »)

Ce dernier principe ne requiert que la capacité pour un agent de savoir qu'il est incertain quand il est incertain, c'est-à-dire une capacité d'introspection sur soi-même, indépendamment de la vérité ou de la fausseté de p. Un énoncé mooréen comme $p \wedge \neg Kp$ requiert plus pour être connu, en revanche, puisqu'il exige que l'agent sache ce qu'il en est du monde. Considérons, pour le voir, le modèle très simple suivant[10] :

w	——	w'
p		$\neg$p
$\neg$Kp		$\neg$Kp
K$\neg$Kp		

Supposons que *w* désigne le monde réel, et que l'agent en *w* se représente les deux éventualités *w* et *w'* comme également possibles. Supposons que p représente « il pleut ». En *w* l'agent considère possible qu'il ne pleuve pas, et envisage la possibilité *w'* suivant laquelle il ne pleut pas. Il ignore donc s'il pleut ou pas. En *w*, on a donc : $p \wedge \neg Kp$, il pleut, et l'agent ne sait pas qu'il pleut. Mais par ailleurs, l'agent sait qu'il ne sait pas que p est vrai, $K\neg Kp$. Pour décrire adéquatement la situation en langage naturel, il faudrait donc plutôt dire qu'il sait qu'il ignore *si* p est vrai ou pas, $K(\neg Kp \wedge \neg K\neg p)$. Ce que proscrit le paradoxe de Moore, en

[10] On adopte ici la sémantique de Hintikka usuelle pour l'opérateur de connaissance. Etant donné un modèle M=<W,R,V>, où W est un ensemble d'états, R une relation de possibilité épistémique entre les états de W, et V une valuation (fonction qui associe à chaque énoncé atomique une valeur de vérité en chaque monde), Kp est vrai en un monde *w* de M si et seulement si p est vrai dans tous les mondes *w'* accessibles depuis *w* par la relation R. (La relation est supposée réfléxive, symmétrique et transitive dans le modèle ci-dessus).

revanche, c'est la condition plus forte exprimée par $K(p \wedge \neg Kp)$, c'est-à-dire que l'agent sache à la fois que p est vrai et qu'il ne le sait pas.

Pour un philosophe comme Williamson, critique du vérificationnisme, le paradoxe de Moore décrit un mode d'inconnaissabilité proprement « structurel », au sens où un sujet ignorant de la vérité d'une proposition ne peut pas savoir qu'il ignore *que cette proposition est vraie* ; au mieux, il saura qu'il est dans l'ignorance. Le principe de connaissabilité, par conséquent, ne peut valoir sans restriction. La question qui se pose alors est celle de l'étendue qui demeure pour le principe de connaissabilité si l'on admet qu'il ne vaut pas sans restriction. Certaines des résolutions avancées au paradoxe de Fitch proposent de modifier la logique classique sous-jacente à la dérivation exposée précédemment, afin de préserver le principe de connaissabilité sans restriction; d'autres proposent de préserver à la fois la logique classique et le principe de connaissabilité sans restriction, mais de remettre en question certains des postulats relatifs au comportement des opérateurs – par exemple la distributivité de la connaissance sur la conjonction [11]. Chacune de ces approches part donc de l'hypothèse qu'il est souhaitable de préserver le principe de connaissabilité sans restriction. Une telle hypothèse ne va pas de soi, cependant : pour qui admet la logique classique et les hypothèses faites sur la connaissance et la possibilité, ce qu'établit l'argument de Fitch est que certaines propositions, comme les propositions mooréennes, ne sont pas connaissables. Mais s'agit-il nécessairement d'une limitation insensée de l'étendue de notre connaissance ? Avant d'entrer plus avant dans la question, il sera utile de dire un mot du débat qui oppose réalistes et antiréalistes sur la question de la connaissabilité.

2. Réalisme et anti-réalisme

Le postulat selon lequel toute vérité est connaissable en principe est couramment attribué aux anti-réalistes, à ceux qui considèrent que le concept même de vérité n'est pas dissociable de la capacité de la pensé humaine à produire des concepts, à formuler des énoncés, et à en établir la vérité ou la fausseté[12]. On peut noter, cependant, qu'il ne serait pas contradictoire d'être à la fois réaliste touchant la vérité, et de souscrire néanmoins au principe selon lequel toute vérité est connaissable en principe. Pour tout réaliste qui admet que nous ne sommes pas omniscients, certaines vérités, à un moment du temps, ne nous sont pas connues. La conjecture de Goldbach, par exemple, énonce que tout nombre pair supérieur à 2 peut s'écrire comme la somme de deux nombres premiers. On ignore encore aujourd'hui si cette conjecture est vraie ou fausse. Certains anti-réalistes diraient que la conjecture de Goldbach n'est ni vraie ni fausse présentement, tant que son statut de vérité ou de fausseté n'a pas été décidé au moyen d'une preuve. Pour le réaliste, à l'inverse, la conjecture est présentement vraie, ou présentement fausse, mais nous sommes ignorants quant à sa vérité ou sa fausseté. Cela n'empêche nullement un tel réaliste d'être optimiste du point de vue épistémologique et de considérer que la vérité ou la fausseté de la conjecture nous est connaissable en principe, à défaut d'être connue présentement.

Ce réalisme optimiste est probablement celui du sens commun, mais aussi bien, à n'en pas douter, celui d'une majorité de savants, pour lesquels résoudre un problème consiste à découvrir une vérité cachée. On sait que pour un mathématicien comme Hilbert, par exemple,

---

[11] Cf. T. Williamson 2000, *op. cit*, pp. 275 sqq., et J. Kvanvig, *op. cit*, chap. 4 sur ce type de stratégie, que nous laissons ici de côté.

[12] Cf. D. Edgington (1985), "The Paradox of Knowability", *Mind*, vol. 94 (376), pp. 557-568: "the items to which truth can be ascribed are not *in rerum natura*. They are the products of human thought – linguistic items, or what is conveyed by linguistic items, or beliefs."

les mathématiques ne devaient comporter aucun *ignorabimus*[13]. Tout en refusant l'intuitionnisme, Hilbert refusait également le réalisme naïf, pensant que toute vérité mathématique était décidable par des moyens élémentaires ; mais le vérificationnisme de Hilbert, bien que lié dans son programme de fondation des mathématiques à une forme spécifique d'antiréalisme, était aussi ancré dans la confiance positiviste dans les capacités de l'esprit humain à résoudre les énigmes scientifiques et dans la réalité de la notion de progrès en science. Il y a, de ce point de vue, une différence de principe importante entre l'antiréaliste vérificationniste et le réaliste optimiste eu égard au principe de connaissabilité : pour l'antiréaliste, le principe de connaissabilité est la conséquence directe d'une conception du vrai qui le fait dépendre de notre capacité à reconnaître le vrai : en ce sens le principe de connaissabilité ne relève en principe d'aucun acte de foi, d'aucune forme d'optimisme, mais suit rationnellement de l'analyse du concept de vérité. Pour souscrire au principe de connaissabilité, le réaliste est quant à lui contraint de faire état d'indices indirects en sa faveur. Il s'agira, la plupart du temps, d'indices inductifs, comme le constat du progrès dans les sciences, la capacité de l'esprit humain à progressivement mettre en lumière les réponses à un problème, etc.

Que l'adoption du principe de connaissabilité relève d'une attitude antiréaliste sur la vérité, ou au contraire d'une confiance positiviste dans les capacités de l'esprit humain à reconnaître le vrai, l'argument de Fitch impose chaque fois une limite. Comme l'a noté Tennant, pour un antiréaliste *modérément radical*, disposé à admettre que toute vérité empirique ne nous est pas *ipso facto* connue, et qui donc refuserait de souscrire à la conclusion de l'argument de Fitch, une issue possible est de contester la validité de l'argument, au nom par exemple de l'adoption d'une logique comme la logique intuitionniste. L'inférence de (9) à (10), en particulier, c'est-à-dire de $\neg(p \wedge \neg Kp)$ à $p \rightarrow Kp$, n'est pas correcte en logique intuitionniste, au contraire de chacune des étapes (1)-(9) de l'argument de Fitch tel que nous l'avons exposé plus haut. L'antiréaliste modérément radical est donc prêt à accepter (9), ce qui revient, en langage intuitionniste, à accepter qu'on ne peut pas prouver qu'il existe une vérité qui nous soit inconnue, tout en refusant (10), c'est-à-dire sans pour autant affirmer que toute vérité nous est connue[14]. Cette réponse au problème est loin d'être satisfaisante, cependant : bien que moins catégorique que l'affirmation que toute vérité est connue, elle conduit à une forme d'agnosticisme touchant l'existence de vérités inconnues qui répugne encore au sens commun[15]. Par ailleurs, et de façon plus directe, cette solution manque de reconnaître le cœur même de l'argument de Fitch, c'est-à-dire sa dépendance à l'égard du paradoxe de Moore.

La réponse intuitionniste au paradoxe de Fitch n'est pas la seule issue pour l'anti-réaliste, comme nous le verrons, mais considérons alors l'attitude inverse, celle du réaliste positiviste, attaché à la logique classique comme aux principes utilisés dans la dérivation du paradoxe de Fitch, et disposé à admettre qu'il existe des vérités inconnues : l'argument de Fitch, logiquement valide et fondé sur des prémisses correctes, selon cette perspective,

---

[13] Voir D. Hilbert 1925, "Sur l'Infini », in J. Largeault (ed.), *Logique Mathématique, Textes*, Armand Colin, et l'adage radiophonique de Hilbert resté célèbre : « *wir wollen wissen, wir werden wissen* ». Tennant, dans *The Taming of the True*, p. 166, parle quant à lui d'« optimisme gödelien », mais en référence à l'attitude réaliste de Gödel.

[14] Rappelons que $\neg(p \wedge \neg Kp)$ doit s'interpréter comme un énoncé quantifié universellement, soit $\forall p \neg(p \wedge \neg Kp)$, ce qui équivaut en logique intuitionniste à $\neg\exists p(p \wedge \neg Kp)$. Si on donne à la négation le sens constructiviste de « il n'existe pas de preuve que », alors l'énoncé dit qu'il n'existe pas de preuve qu'il existe une vérité qui nous soit inconnue, ce qui est moins fort que de dire que toute vérité est connue.

[15] Pour plus de détails sur la critique de la réponse intuitionniste, cf. notamment H. Rückert, "A Solution to Fitch's Paradox of Knowability", in S. Rahman, J. Symons, D. Gabbay, J.P. van Bendegem, (eds), *Logic, Epistemology and the Unity of Science*, vol.1, Kluwer 2004, pp. 351-380, et T.Williamson, "Troubles with Tennant", à paraître in J. Salerno (ed.), *op. cit.*

l'oblige à réfréner son optimisme naïf, et à admettre qu'il existe des vérités inconnaissables. Mais quel genre de limitation cela impose-t-il à l'attitude positiviste qui le caractérise ? A examiner la question de plus près, le positiviste (qu'il soit ou non réaliste[16]) peut fort bien se dire : « ce que montre l'argument de Fitch, c'est d'abord ce qu'établit le paradoxe de Moore : au moment où il pleut et où j'ignore qu'il pleut, je ne peux pas savoir *à la fois* qu'il pleut *et* que j'ignore qu'il pleut. Cela je l'admets bien volontiers. Mais cela ne saurait faire offense à mes convictions positivistes. Car ce qui importe, s'il pleut et que je l'ignore, c'est :

(i) *à ce moment même ou à un moment ultérieur*, que je puisse être conscient du fait que j'ignore *s*'il pleut ou pas
(ii) *à un moment ultérieur*, que je puisse savoir qu'il pleuvait au moment où je l'ignorais ».

Si cette réaction au paradoxe de Fitch est fondée, elle suggère que le renoncement au principe de connaissabilité (V) n'implique pas pour autant de renoncer à l'attitude positiviste qui lui est sous-jacente. Plus précisément, elle suggère que l'on puisse substituer au principe $p \to \Diamond Kp$ d'autres principes, conformes à l'inspiration positiviste du principe de connaissabilité, et néanmoins indemnes vis-à-vis de la conclusion paradoxale de Fitch.

Ce que préconise le postulat (i), c'est en effet qu'un agent soit conscient de son ignorance chaque fois qu'il ignore si une proposition est vraie ou fausse : c'est le contenu même du principe d'introspection négative $\neg Kp \to K\neg Kp$, discuté plus haut. Il s'agit là d'une maxime de rationalité interne du point de vue épistémique, conforme aussi bien à la maxime socratique du « connais-toi toi-même » qu'à la recommandation cartésienne d'éviter prévention et précipitation. Le postulat (ii) , inversement, correspond à un cas particulier du principe de connaissabilité initial, mais relatif cette fois à la partie *objective* ou non-épistémique de l'énoncé de Moore : le fait qu'il pleuve, selon cette perspective, n'est pas par lui-même inconnaissable. Ce qui est inconnaissable à un moment donné, c'est en effet un *fait complexe*, le fait qu'il pleuve et que j'ignore qu'il pleut, décomposable en un fait relatif à ma connaissance et en un fait indépendant de ma connaissance ; pour un rationaliste, la partie épistémique de ce fait, à savoir mon ignorance, doit cependant être accessible à la conscience ; et de la même manière, la partie objective du fait demeure elle-même connaissable en principe.

La réaction qui précède, aussi convaincante soit-elle, ne suffit pas encore à résoudre le problème : ce que suggèrent les postulats (i) et (ii), c'est en effet une manière de concevoir la connaissabilité de certaines propositions complexes *particulières*, à savoir les propositions mooréennes. Il reste à préciser, cependant, comment il convient d'articuler *en toute généralité* la notion de connaissabilité pertinente pour le positiviste. Dans ce qui suit, nous examinons plusieurs restrictions possibles au principe de connaissabilité initial. Nous reviendrons dans la section suivante sur les limitations qui subsistent à la connaissabilité, telle qu'on peut vouloir articuler la notion dans une perspective positiviste.

3. Restrictions au principe de connaissabilité

[16] Soulignons que le positivisme, si l'on entend par là avant tout la confiance dans la capacité de l'esprit humain à faire progresser rationnellement la science et les connaissances, est une attitude compatible aussi bien avec le réalisme qu'avec l'anti-réalisme. Ici et dans ce qui suit, nous opposons le réaliste positiviste et l'anti-réaliste vérificationniste, mais il faut garder en tête que le vérificationnisme, directement issu des thèses du positivisme logique, est l'une des manières de fonder et d'interpréter la notion même de positivisme.

L'objet de cette section est d'examiner plusieurs illustrations de ce que N. Tennant a appelé la « stratégie de restriction » en réponse au paradoxe de Fitch[17]. Plus précisément, nous examinons ici trois stratégies de restriction en réponse au paradoxe. La première stratégie, proposée notamment par Dummett, restreint le principe de connaissabilité aux énoncés *atomiques*. La seconde, défendue par Tennant, restreint le principe de connaissabilité aux énoncés non-mooréens, et que Tennant appelle *cartésiens*, c'est-à-dire aux énoncés tels que la connaissance de ces énoncés n'aboutisse pas à une contradiction. La troisième enfin, défendue à l'origine par Edgington, suggère de restreindre le principe de connaissabilité aux énoncés préfixés d'un opérateur d'actualité, c'est-à-dire aux vérités dont le contexte est rendu explicite. Chacune de ces stratégies a été présentée par son auteur comme une façon de défendre le vérificationnisme contre le réalisme. Nous examinerons dans la section qui suit une forme de restriction encore différente, inspirée en partie de l'approche d'Edgington, mais affranchie de tout souci anti-réaliste.

3.1. « Toute vérité de base est connaissable ».

La réaction au paradoxe de Moore esquissée dans le paragraphe précédent consistait à remarquer que le fait qu'il pleuve, à lui seul, ne semble pas en soi inconnaissable, au contraire de l'énoncé de Moore. La première réponse qu'on puisse envisager au paradoxe de Fitch consisterait par conséquent à restreindre le principe de connaissabilité aux énoncés atomiques, en particulier aux énoncés qui ne comportent pas de vocabulaire épistémique, censés décrire des faits élémentaires, de nature non-épistémique. Un énoncé comme $(p \wedge \neg Kp)$ n'est pas un énoncé atomique de ce point de vue, puisque d'une part il s'exprime à l'aide d'une conjonction, et que d'autre part il comporte comme sous-énoncé l'énoncé épistémique $\neg Kp$. Selon cette perspective, la formulation adéquate du principe de connaissabilité devrait donc être :

(V_{at}) $p \rightarrow \Diamond Kp$, pour p un énoncé atomique

En formulant ainsi le principe de connaissabilité, on exclut du domaine de la connaissabilité les énoncés mooréens, et on évite ainsi le paradoxe. Cette stratégie, comme l'ont souligné la plupart des critiques, risque cependant d'être trop restrictive, car il semble qu'on exclut du champ de la connaissabilité certaines vérités intuitivement connaissables et qui s'exprimeraient à l'aide d'énoncés complexes, comme « il pleut et il fait froid », ou encore des vérités relatives à la connaissance elle-même : si je sais qu'il pleut, c'est manifestement là un fait connaissable sans contradiction, mais qui ne s'exprime pas sans faire intervenir de vocabulaire épistémique.

L'objection n'est pas insurmontable, puisqu'on peut imaginer de libéraliser le principe (V_{at}), de façon à prendre en compte certains énoncés complexes. C'est ainsi que Dummett, soucieux de défendre l'anti-réalisme contre la conclusion de Fitch, propose une théorie épistémique de la vérité qui étend le principe (V_{at}) de façon inductive[18]. Dans la théorie de Dummett, un énoncé *de base* est vrai si et seulement si, par définition, il est connaissable. Par extension, une conjonction est vraie si chacun des conjoints est vrai, une disjonction est vraie si l'un des disjoints est vrai, et la négation d'un énoncé est vraie si cet énoncé n'est pas vrai. La théorie de Dummett a pour conséquence que la conjonction $(p \wedge q)$ de deux énoncés de base est vraie si chacun des énoncés p et q est connaissable ; de même elle a pour conséquence qu'un énoncé de base n'est *pas* vrai si et seulement si cet énoncé n'est pas connaissable. Deux aspects de la théorie de Dummett méritent par ailleurs mention : le

[17] Cf. également le chap. 3 de J. Kvanvig, *op. cit.*, intitulé "Syntactic Restriction Strategies".
[18] M. Dummett, "Victor's Error", *Analysis*, 61, pp. 1-2.

premier est que Dummett traite les connecteurs de son métalangage comme des connecteurs intuitionnistes ; le second, c'est qu'il laisse ouverte la définition précise des énoncés qui comptent comme « énoncés de base ». Dans un premier temps, cependant, Dummett semble admettre qu'un énoncé épistémique de la forme Kp compte comme un énoncé de base [19].

Considérons alors la prédiction que fait l'analyse de Dummett dans le cas de l'énoncé mooréen $p \wedge \neg Kp$. L'énoncé est vrai si et seulement si, par définition, p est connaissable, et si Kp n'est pas vrai, c'est-à-dire si Kp n'est pas connaissable. Cela revient à dire que $p \wedge \neg Kp$ est vrai si et seulement si l'on a : $\Diamond K(p) \wedge \neg\Diamond K(Kp)$, ce qui doit signifier : p est connaissable, mais il n'est pas connaissable que p est connu. L'interprétation que donne Dummett de ce résultat est la suivante : si p est une vérité inconnue de tous à tous les moments du temps, alors on ne sait pas que p est faux. Mais selon Dummett, pour tout énoncé p dont on ne sache pas qu'il soit faux, « on ne peut jamais complètement exclure la possibilité » qu'on en vienne à savoir qu'il est vrai. Dans ce cas il est rationnel de concevoir que p pourrait être connu à un moment ultérieur. D'un autre côté, l'énoncé est tel que personne ne pourra jamais savoir que quiconque sait qu'il est vrai, et selon Dummett, si la possibilité de savoir que p apparaît trop lointaine, alors il est aussi rationnel d'affirmer que l'on ne pourra jamais savoir qu'on le sait.

La conclusion de Dummett n'en demeure pas moins énigmatique, car il semble que tout repose sur le bon comportement des notions de possibilité et de connaissance. Ainsi, comme le remarquent Brogaard et Salerno, l'énoncé $\Diamond K(p) \wedge \neg\Diamond K(Kp)$ est très proche d'être une contradiction ; si l'on suppose que la notion de possibilité exprimée par l'opérateur $\Diamond$ est transitive (obéissant au schéma $\Diamond\Diamond p \rightarrow \Diamond p$: « s'il est possible que p soit possible, alors p est possible »), il est en effet facile de montrer que du principe de connaissabilité pour les énoncés de base et de $\Diamond Kp$, il suit $\Diamond\Diamond KKp$, et par transitivité, $\Diamond KKp$, en contradiction avec $\neg\Diamond KKp$[20]. Or rien dans l'argument de Fitch ne justifierait de rejeter le principe de transitivité pour la possibilité, si ce n'est de chercher à échapper, de façon alors circulaire, à la menace de contradiction relevée par Brogaard et Salerno. La même contradiction s'ensuit, en outre, si l'on suppose que l'opérateur de connaissance K obéit au principe d'introspection positive ($Kp \rightarrow KKp$).

On pourrait imaginer, néanmoins, de modifier la définition inductive de la vérité proposée par Dummett, de façon à valider autrement l'idée que chacun des conjoints de l'énoncé de Moore est connaissable indépendamment. Pour cela, il faudrait un moyen de valider l'inférence $p \wedge \neg Kp$ à la conclusion voulue $\Diamond Kp \wedge \Diamond K\neg Kp$, elle-même plus faible que l'énoncé contradictoire $\Diamond(Kp \wedge K\neg Kp)$[21]. L'approche la plus directe consisterait à supposer qu'un énoncé comme $\neg Kp$ fait partie des énoncés de base, et donc à traiter la négation comme directement liée à l'opérateur K (comme dans le verbe « ignorer » par opposition à la locution périphrastique « ne pas savoir »). Il n'est pas évident, cependant, qu'on puisse obtenir de la sorte une théorie adéquate de la connaissabilité[22]. La principale faiblesse d'une approche de

---

[19] Dummett écrit en particulier (à propos de Victor, le théoricien anti-réaliste de la vérité qu'il imagine): "There is a good deal of work for Victor to do, particularly in specifying what is to count as a basic statement; if his inductive characterization of truth is to be comprehensive, the basic statements must include all those that cannot be represented as in any of the forms governed by clauses (ii) to (vii), or by any supplementary clauses", les clauses (ii)-(vii) étant celles qui concernent les énoncés complexes par opposition aux énoncés de base. C'est le cas des énoncés de la forme "Kp", qui ne rentrent pas sous les clauses (ii)-(vii) en question.

[20] Voir B. Broogard & J. Salerno, "Clues to the Paradoxes of Knowability: Reply to Dummett and Tennant", *Analysis* 66 (2), 2002, pp. 143-150.

[21] Rappelons que de $\Diamond A \wedge \Diamond B$, il ne suit pas $\Diamond(A \wedge B)$. Par exemple, il est possible que demain matin il pleuve et il est possible que demain matin il ne pleuve pas, mais il n'est pas possible qu'il pleuve et ne pleuve pas.

[22] Faudra-t-il, en particulier, traiter tous les énoncés de la forme $\neg K\phi$ comme des énoncés de base, de même que tous les énoncés de la forme $K\phi$?

ce genre, on le voit, est qu'elle oblige à multiplier les stipulations, sans livrer une caractérisation sémantique indépendante des énoncés qui comptent comme énoncés de base.

3.2. « Toute vérité cartésienne est connaissable »

Dans la conception de Dummett, l'énoncé de Moore $p \wedge \neg Kp$ fait obstacle au principe de connaissabilité essentiellement du fait de sa forme conjonctive. Avant Dummett, Tennant a défendu une vision plus holistique du caractère pathologique des énoncés de type Moore. Ainsi, pour Tennant, toute vérité qui n'aurait pas le caractère d'inconnaissabilité structural typique des énoncés de Moore doit en principe être connaissable. Tennant appelle *cartésien* tout énoncé ϕ tel que ϕ puisse être connu sans contradiction, c'est-à-dire tout énoncé ϕ tel que $K\phi$ soit un énoncé cohérent ou satisfaisable. Selon Tennant, la formulation adéquate du principe vérificationniste devrait donc être :

($\Diamond$KC) $\phi \rightarrow \Diamond K\phi$, pour tout énoncé cartésien ϕ

Comme l'a fait remarquer Tennant contre Dummett[23], chaque fois qu'un énoncé de la forme $K\phi$ est contradictoire, ϕ doit nécessairement être un énoncé complexe du point de vue syntaxique. Cela signifie, de façon équivalente, que tout énoncé atomique est nécessairement cartésien au sens de Tennant. Mais la réciproque est fausse : il existe des énoncés cartésiens non-atomiques, comme par exemple l'énoncé $p \wedge q$. De ce fait, le principe de Tennant est d'emblée plus général que le principe (V_{at}) énoncé plus haut, et le principe est également censé être plus général que le principe de Dummett qui affirme que toutes les vérités *de base* sont connaissables. On pourrait néanmoins se demander si la théorie de Dummett n'est pas extensionnellement équivalente à celle de Tennant quant à la caractérisation des énoncés connaissables, du fait des clauses inductives proposées par Dummett. Ce n'est pas le cas : si une conjonction du type $p \wedge q$ est vraie, où p et q sont des énoncés atomiques, alors la théorie de Tennant prédit qu'il est possible de savoir que la conjonction est vraie, soit $\Diamond K(p\wedge q)$, $(p\wedge q)$ étant un énoncé cartésien. La théorie de Dummett, en revanche, fait d'emblée une prédiction plus faible du point de vue logique, à savoir qu'il est possible de connaître p *et* qu'il est possible de connaître q, soit $\Diamond Kp \wedge \Diamond Kq$. Dans la perspective constructiviste de Dummett, ce serait manifestement le sens à donner à l'énoncé $\Diamond K(p\wedge q)$ lui-même, mais *stricto sensu*, on voit que « connaissable au sens de Dummett » et « connaissable au sens de Tennant » ne sont pas des notions coextensionnelles.

La restriction du principe de connaissabilité aux énoncés cartésiens présente l'avantage d'exclure d'emblée du champ de la connaissabilité tous les énoncés analogues à l'énoncé de Moore : ces énoncés que, pour des raisons qui tiennent à la factivité de la connaissance, il est impossible de connaître sans contradiction. La conjonction $(p \wedge \neg Kp)$ n'est pas cartésienne, mais en revanche chacun des conjoints p et $\neg Kp$ est cartésien, et par conséquent chacun des conjoints est connaissable d'après le principe de Tennant : s'il pleut et que j'ignore qu'il pleut, ce fait là n'est pas connaissable, mais chacun des faits plus élémentaires qui le composent est bien connaissable.

Comme le souligne van Benthem, « être un énoncé cartésien » est une propriété décidable en logique modale épistémique, ce qui veut dire que pour tout énoncé ϕ, on peut calculer de façon effective si $K\phi$ est cohérent ou pas[24]. En ce sens, la règle ($\Diamond$KC) est bien une règle effective. On peut également observer que si un énoncé ϕ *n'est pas cartésien*, alors par définition on a $K\phi \rightarrow \bot$, et donc $\Diamond K\phi \rightarrow \bot$, ce qui veut dire que ϕ *est inconnaissable*. Cela

[23] N.Tennant, "Victor Vanquished", *Analysis* 62 (2), 2002, pp. 135-142.
[24] J. van Benthem, "What one may come to know", *Analysis* 64(2), 2002, pp. 95-105.

signifie, par contraposition, que tout énoncé connaissable est nécessairement un énoncé cartésien, ce qui est le contenu même de la règle ($\Diamond$KC). Mais qu'en est-il de la réciproque ? Se pourrait-il qu'il y ait des énoncés cartésiens qui ne soient pas connaissables ? Si tel était le cas, cela signifierait qu'il y a un énoncé ϕ tel que $K\phi$ est satisfaisable, mais tel que $\Diamond K\phi$ n'est pas satisfaisable : mais si l'on suppose que la notion métaphysique de possibilité est au moins réflexive (satisfaisant le schéma $p \rightarrow \Diamond p$), alors $\Diamond K\phi$ est vrai en tout contexte dans lequel $K\phi$ est vrai, et par conséquent tout énoncé cartésien est connaissable. Moyennant l'hypothèse communément admise que la possibilité est réflexive, on voit que « cartésien » et « connaissable » sont des concepts coextensionnels.

Parmi les critiques formulées à l'encontre de Tennant, on trouve du même coup l'idée selon laquelle la restriction de Tennant est *ad hoc*, voire circulaire : le principe de Tennant reviendrait à affirmer que toute vérité est connaissable, sauf les vérités qui ne sont pas connaissables. Le reproche est en partie injustifié, cependant, puisque « cartésien » et « connaissable » ne sont pas des notions *intensionnellement* équivalentes. En particulier, l'idée que les seules vérités inconnaissables soient des vérités non-cartésiennes suggère que les vérités en questions sont inconnaissables pour des raisons structurales, et qu'on ne perd aucune vérité substantielle en restreignant de cette façon le vérificationnisme.

Parmi les autres objections faites à Tennant figurent cependant différentes formes de « revanche » du paradoxe de Fitch. Williamson, par exemple, soutient contre Tennant que même en adoptant la restriction voulue aux énoncés cartésiens, on peut encore dériver la conclusion de Fitch. L'argument de Williamson part de l'observation qu'un énoncé *non-cartésien* comme l'énoncé de Moore ($p \wedge \neg Kp$) implique logiquement l'énoncé *cartésien* ($p \wedge (Kp \rightarrow q)$), de même que l'énoncé cartésien ($p \wedge (Kp \rightarrow \neg q)$). En vertu du principe de Tennant, chacun de ces énoncés est connaissable et on a donc $\Diamond K(p \wedge (Kp \rightarrow q))$ et $\Diamond K(p \wedge (Kp \rightarrow q))$, et chacun de ces énoncés implique respectivement $\Diamond q$ et $\Diamond \neg q$. Cette conclusion n'est pas contradictoire, mais le nerf de l'objection de Williamson est dans la supposition que q et $\neg q$ puissent être des propositions telles que $\Diamond q$ et $\Diamond \neg q$ soient vraies si et seulement si q et $\neg q$ sont vraies. Par exemple, si 9 était impair, cela impliquerait qu'il n'est pas possible que 9 soit pair, car ce serait là une propriété nécessaire du nombre 9. Mais par contraposition, cela signifie que s'il était possible que 9 soit pair, alors 9 *ne serait pas* impair, et donc 9 *serait* pair. Si donc q exprime une vérité nécessaire de ce type, alors de $\Diamond q \wedge \Diamond \neg q$ on devrait pouvoir inférer $q \wedge \neg q$, une contradiction. Comme cette contradiction suit de l'énoncé mooréen ($p \wedge \neg Kp$) et du principe de Tennant, Tennant doit en principe nier ($p \wedge \neg Kp$), ce qui revient à accepter ($p \rightarrow Kp$), le principe qui affirme que toute vérité est connue.

L'objection de Williamson n'est pas absolument satisfaisante, notamment parce que tout repose sur l'idée que si de tout énoncé q, on peut démontrer $\Diamond q \wedge \Diamond \neg q$, comme c'est le cas sous l'hypothèse de Tennant, on aurait là une conclusion trop forte. On voit que l'argument repose sur des contraintes logiques beaucoup plus fortes que celles utilisées par Fitch dans son argument initial. Une objection d'un autre type a été proposée par Brogaard et Salerno, mais là encore en invoquant un axiome supplémentaire, le principe $K\neg p \rightarrow \neg\Diamond p$, d'après lequel s'il est connu que p n'est pas vrai, alors p n'est pas possible[25]. Brogaard et Salerno proposent d'interpréter la notion de possibilité comme une notion de possibilité épistémique, et non plus métaphysique, ce qui rend le principe plausible. Sous cette hypothèse ils montrent que le principe ($\Diamond$KC) de Tennant oblige à renoncer à l'existence d'énoncés *indécidés*, c'est-à-dire tels qu'on ne sache ni si p est vrai, ni si p n'est pas vrai. La conjecture de Goldbach est un exemple d'énoncé indécidé, puisqu'on ne sait ni si elle est vraie, ni si elle est fausse.

[25] Cf. B. Brogaard & J. Salerno 2002, *loc. cit.*

L'objection de Brogaard et Salerno peut sembler à nouveau dirimante pour l'antiréalisme défendu par Tennant, mais sa principale faiblesse provient du principe $K\neg p \rightarrow \neg\Diamond p$: car si l'opérateur $\Diamond$ est interprété au sens de la possibilité épistémique, comme le suggèrent Brogaard et Salerno, le principe est certes plausible, mais dans ce cas on devrait s'attendre à ce que $\Diamond p$ soit synonyme de l'expression $\neg K\neg p$, qui exprime la possibilité épistémique de façon usuelle (« je n'exclus pas que p »). Or pour des raisons d'uniformité le principe de connaissabilité ($\Diamond$KC) devrait à son tour s'exprimer comme suit :

($\Diamond$KC*) $\phi \rightarrow \neg K\neg K\phi$, pour tout énoncé cartésien ϕ

Le principe signifie que si ϕ est vrai et cartésien, alors je n'exclus pas que je sache ϕ, ou encore je ne sais pas que je ne sais pas ϕ. A l'heure qu'il est, cependant, je ne sais pas s'il fait beau en Argentine, et même je sais que je le sais pas. Mais à supposer qu'il fasse beau en Argentine à l'heure qu'il est, c'est là une vérité cartésienne qui est telle que je sais que je ne sais pas si elle est vraie, ce qui contredit de façon évidente le principe ($\Diamond$KC*). Selon toute vraisemblance, par conséquent, ce n'est pas l'existence d'énoncés indécidés que devrait rejeter Tennant, mais plutôt l'interprétation de l'opérateur de possibilité qui résulte du principe $K\neg p \rightarrow \neg\Diamond p$ mis en avant par Brogaard et Salerno. *A contrario*, on peut remarquer que si l'opérateur de possibilité est lu de façon métaphysique (au sens de « il pourrait se faire que » ou « il pourrait arriver que »), alors le principe invoqué perd cette fois de sa plausibilité : je sais qu'il n'existe pas de licornes, par exemple, mais il pourrait se faire qu'il en existe (un jour, ou sur une autre planète) : il y a plus de possibilités métaphysiques, en principe, que de possibilités épistémiques à un moment donné.

Ni l'objection de Williamson, par conséquent, ni celle de Brogaard et Salerno ne sont décisives. Dans les deux cas, on voit que les critiques de Tennant sont obligés de faire des hypothèses spécifiques sur le comportement ou sur l'interprétation de l'opérateur de possibilité, dont la généralité est chaque fois discutable. La principale énigme qui demeure pour la stratégie de Tennant, malgré tout, et qui rejoint partiellement le reproche d'adhocité déjà évoqué, concerne l'ampleur de la restriction aux énoncés cartésiens qu'il propose. Ainsi, se pourrait-il que certaines des vérités non-cartésiennes qui ne nous sont pas connaissables soient des vérités substantielles ? Un présupposé de la stratégie de Tennant semble en effet être que les vérités non-cartésiennes sont inconnaissables en vertu d'un trait structural, à la manière de l'énoncé de Moore, et qu'elles peuvent alors se résoudre en vérités cartésiennes plus élémentaires. Mais est-ce nécessairement le cas ?

Dans *The Taming of the True*, Tennant se demande : « est-ce que la restriction aux vérités cartésiennes garantit la vérité du principe de connaissabilité par l'effet d'une pure stipulation ? Est-ce que cela revient à affirmer que le principe est valable sauf là où il ne l'est pas ? ». La réponse de Tennant est la suivante : « Pas du tout. Pensez à toutes les propositions cartésiennes des mathématiques et de la science empirique, les propositions qui n'impliquent aucune mention, en leur sein, de notions épistémiques. Affirmer que chaque vérité de ce type est connaissable en principe c'est encore conjurer le réalisme métaphysique ».

On pourrait se demander, toutefois, s'il pourrait exister des vérités non-épistémiques qui soient structurellement inconnaissables. Imaginons, par exemple, un mécanisme infaillible, tel que si j'appuie sur le bouton, je ne saurai pas que j'ai appuyé sur le bouton (parce que l'effet du mécanisme est d'effacer ma mémoire). Puis-je savoir si j'ai appuyé sur le bouton ou pas ? Demandons-nous d'abord si l'énoncé « j'ai appuyé sur le bouton » est cartésien ou pas. A première vue, il semble que l'énoncé soit un énoncé cartésien, au sens où la phrase « je sais que j'ai appuyé sur le bouton » n'est pas par elle-même contradictoire[26]. A la différence de l'énoncé de Moore, rien dans la structure logique interne de « j'ai appuyé sur le bouton » ne rend l'énoncé directement inconnaissable. Malgré cela, le postulat en vertu duquel « si j'ai appuyé sur le bouton, alors je ne saurai pas que j'ai appuyé sur le bouton » rend l'énoncé inconnaissable. Car pour savoir que j'ai appuyé sur le bouton, il faut que j'aie appuyé sur le bouton. Mais si tel est le cas, je ne sais pas que j'ai appuyé sur le bouton. Supposons inversement que je n'aie pas appuyé sur le bouton : dans ce cas je peux savoir en principe que je n'ai pas appuyé sur le bouton. Mais la nature du mécanisme a de quoi me faire douter : si je n'ai pas appuyé sur le bouton, alors je ne peux pas savoir que j'ai appuyé sur le bouton ; dans ce cas, cependant, il se pourrait fort bien que j'aie appuyé sur le bouton et que je ne m'en souvienne pas. Pour symboliser les choses, on peut représenter par p l'énoncé : « j'ai appuyé sur le bouton ». Alors le mécanisme est tel que $(p \rightarrow \neg Kp) \wedge (\neg p \rightarrow (\neg Kp \wedge \neg K \neg p))$. Notons que la structure même de cet énoncé complexe est typique des hypothèses sceptiques : si par exemple je suis un cerveau dans une cuve, alors par la nature même de l'hypothèse je ne sais pas que je suis un cerveau dans une cuve. Mais si je ne suis pas un cerveau dans une cuve, ce qui me semble être le cas la plupart du temps, je n'exclus pas pour autant catégoriquement que ce soit le cas.

De tels exemples montrent que certains énoncés, comme « je suis un cerveau dans un cuve », qui ne comportent pas de vocabulaire épistémique, pourraient s'avérer inconnaissables et même indécidables en vertu de liens logiques extrinsèques avec d'autres énoncés relatifs à la connaissance, et non pas en raison de leur structure logique interne. De tels énoncés mettent en péril la position anti-réaliste de Tennant : intuitivement, l'une des deux propositions « je suis un cerveau dans une cuve » ou « je ne suis pas un cerveau dans une cuve » est vraie ; or soit Tennant concèderait que chaque énoncé considéré isolément est cartésien ; soit il doit considérer que les énoncés ne sont pas cartésiens, mais d'une manière qui ne plaide pas nécessairement en faveur de l'anti-réalisme.

3.3. « Toute vérité actuelle est connaissable »

L'énoncé de Moore « il pleut et je ne sais pas qu'il pleut » n'est pas connaissable au moment même où il pleut et où j'ignore qu'il pleut. En revanche, comme l'a souligné D. Edgington, il n'est pas contradictoire de considérer que s'il pleut au moment *t*, et si j'ignore qu'il pleut au moment *t*, je puisse néanmoins savoir à un moment *t'* ultérieur à *t* qu'au moment *t* il pleuvait, et qu'à ce moment *t* j'ignorais qu'il pleuvait[27]. Considérons à nouveau le principe de connaissabilité $p \rightarrow \Diamond Kp$. Si l'on interprète l'opérateur de possibilité $\Diamond$ comme un opérateur de possibilité temporel (« il existe un moment futur tel que... »), alors le principe de connaissabilité signifie que si p est vrai au moment *t*, il existe un moment ultérieur *t'* tel que, au moment *t'*, quelqu'un sait en *t'* que p est vrai en *t'*. Plus formellement, cela signifie :

$$(\Diamond F) \qquad p(t) \rightarrow (\exists t' > t)(K(t')p(t'))$$

Supposons par exemple qu'il ait plu lundi dernier à 10h du matin. Le principe implique qu'il y aura un moment ultérieur (par exemple le mardi) tel que je puisse savoir (le mardi) qu'il pleut à ce moment (le mardi). Mais comme le montre cette traduction, le principe n'est pas celui que nous recherchons : s'il pleuvait le lundi à 10h et qu'à ce moment là personne ne le sût, ce que l'on souhaite c'est qu'à un moment ultérieur quelqu'un sache qu'il

[26] Si l'on représente par l'atome "p" la phrase "j'ai appuyé sur le bouton", l'énoncé "Kp" est parfaitement satisfaisable, et p est donc un énoncé cartésien en ce sens.
[27] D. Edgington (1985), *loc.cit.*

pleuvait *le lundi à 10h*, et non pas qu'il pleut *à ce moment ultérieur*. En d'autres termes, le principe voulu est le suivant :

(◊M) $\quad p(t) \rightarrow (\exists t' > t)(K(t')p(t))$

Comme on peut le prouver, au contraire du principe (◊F), le principe (◊M) ne peut se traduire en logique modale uniquement en termes de l'opérateur de possibilité ◊. Pour obtenir une traduction satisfaisante, il faut avoir recours à un opérateur spécifique d'*actualité* : un opérateur d'actualité est un opérateur qui, comme l'opérateur « maintenant » en logique temporelle, permet de maintenir constante la référence au moment présent ou à la situation actuelle sous la portée des opérateurs derrière lesquels il est enchâssé. Si l'on représente par A cet opérateur (pour « actuellement »), alors la traduction de (◊M) en logique modale devient :

(◊KA) $\quad Ap \rightarrow \Diamond KAp$

Comme on peut le voir, le schéma est un cas particulier du schéma de connaissabilité initial : selon Edgington, ce schéma donne l'expression convenable du principe de connaissabilité, en énonçant que toute proposition *actuellement vraie* doit être telle qu'il est possible de savoir que cette proposition est *actuellement vraie*. Ainsi, si p est vrai *présentement*, mais que nul ne sait *présentement* que p est vrai, alors le principe, dans sa version temporelle, énonce qu'il existe un moment ultérieur tel qu'à ce moment quelqu'un saura que p était vrai *présentement*. Enoncé sous cette forme, le principe de connaissabilité ne donne plus lieu à la conclusion contre-intuitive de Fitch. Ainsi, de $A(p \wedge \neg Kp)$, on infère $\Diamond KA(p \wedge \neg Kp)$, ce qui équivaut à $\Diamond(KAp \wedge KA\neg Kp)$: or $KA\neg Kp$ implique $A\neg Kp$, mais KAp n'implique pas AKp (si quelqu'un sait au moment *t'* que p est vrai en *t*, cela n'implique pas qu'au moment *t* quelqu'un sait que p est vrai).

La restriction proposée par Edgington est certainement la plus naturelle de celles considérées jusqu'à présent, dans la mesure où elle restreint le principe de connaissabilité de façon uniforme, et sans courir le reproche de circularité. Contre cette solution, cependant, deux objections principales ont été avancées par Williamson[28]. La première est une objection technique relative au comportement sémantique de l'opérateur d'actualité. Elle consiste à remarquer que toute vérité actuelle est *nécessairement* actuellement vraie. Ainsi, de Ap, on peut inférer $\Box Ap$: cela signifie que si p est vrai dans la situation présente, alors relativement à toute situation possible, il reste vrai que p est vrai dans la situation présente. Ce fait n'est pas directement problématique, mais il le devient lorsque l'on considère les conditions de vérité usuelles associées à l'opérateur de connaissance K : dans ce cas, comme l'ont souligné Rabinowicz et Segerberg, la sémantique prédit que de Ap, on peut inférer de la même façon KAp : si p est actuellement vrai, alors p est actuellement vrai relativement à tout scénario *épistémique* possible[29]. Ainsi, il suit de la définition de l'opérateur d'actualité que $Ap \rightarrow KAp$, ce qui apparaît comme une « revanche » de la conclusion initiale de Fitch.

Le problème relevé par Segerberg et Rabinowicz n'est pas nécessairement insurmontable, si l'on peut imaginer de modifier la sémantique de l'opérateur de connaissance K d'une façon appropriée, comme Rabinowicz et Segerberg proposent de le faire par ailleurs. Du point de vue conceptuel, cependant, il reste une objection plus générale, relative à ce que Williamson appelle le problème de la connaissance non-actuelle des vérités actuelles. Williamson fait remarquer que s'il pleut à six heures, à six heures la pensée qu'il pleut présentement peut s'exprimer soit de façon indexicale par l'énoncé « il pleut maintenant »,

[28] T. Williamson (1987), "On the Paradox of Knowability", *Mind*, vol. 96, 382, pp. 256-261.
[29] W. Rabinowicz & K. Segerberg (1994), "Actual Truth, Possible Knowledge", *Topoi* 13, pp. 101-115.

soit de façon non-indexicale par l'énoncé « il pleut à six heures ». Comme John Perry et David Lewis l'ont montré à propos des pensées *de se*, ces deux pensées n'ont pas le même contenu : je peux penser qu'il pleut maintenant sans penser qu'il pleut à six heures, si j'ignore qu'il est maintenant six heures[30]. Or, la seule façon de savoir à un temps ultérieur, par exemple à sept heures, qu'il pleuvait à six heures, s'exprime par la pensée « il pleuvait à six heures », ou encore « il pleuvait *alors* », et non pas par la pensée « il pleut maintenant », qui voudrait dire qu'il pleut à sept heures. Plus précisément, « il pleuvait alors » est la façon adéquate de penser à sept heures la pensée qu'on aurait pu à six heures exprimer en disant « il pleut maintenant », mais implique de *se souvenir* à sept heures d'une situation vécue à six heures. Or selon Williamson, le lien causal qui existe entre ma situation à six heures et ma situation à sept heures n'existe pas forcément, de façon plus générale, « entre l'actuel et le non-actuel ».

Pour se représenter de façon plus éloquente encore le problème, faisons une expérience de pensée et imaginons que le 16 juin 2002 à 15h35, un arc-en-ciel se soit produit à tel endroit au dessus de la Seine. Imaginons cependant que personne n'ait été présent pour observer le phénomène. A 15h35 ce jour-là et au lieu dit, si un observateur avait été présent, il aurait eu la pensée indexicale : « il y a à présent un arc-en-ciel », car il aurait eu l'expérience de l'arc-en-ciel se produisant. Si l'on adopte un point de vue réaliste, cependant, il faut penser que « il y a maintenant un arc-en-ciel » était une proposition vraie le 16 juin 2002 à 15h35, indépendamment même de la présence de cet observateur. Mais puisque personne n'était alors présent pour en faire l'expérience, la seule façon de savoir désormais qu'il y avait un arc-en-ciel est au moyen de la proposition : « il s'est produit un arc-en-ciel à tel endroit *le 16 juin 2002 à 15h35* » : cette vérité est connaissable, mais le contenu de la proposition indexicale vraie « il y a à présent un arc-en-ciel » n'est pas connaissable à proprement parler, car cette proposition manque d'un analogue qui en préserverait le caractère indexical (cette fois personne ne peut *se souvenir* de l'arc-en-ciel se produisant *alors*).

Si donc la solution d'Edgington est fructueuse, c'est semble-t-il au prix de l'inconnaissabilité de certaines propositions indexicales. On pourrait objecter, néanmoins, qu'il n'y a d'indexicalité que relativement à un sujet connaissant, et de façon « token-réflexive » : dans ce cas, si personne n'a pu être témoin de l'arc-en-ciel au moment même, la proposition « il y a présentement un arc-en-ciel » ne peut avoir été vraie alors même que l'arc-en-ciel se produisait, puisqu'elle n'a pas pu être *proférée*. On pourrait répondre que si quelqu'un *avait proféré* « il y a présentement un arc-en-ciel », son assertion aurait été vraie. Mais dans ce cas, il est tentant de penser que si quelqu'un avait été présent pour faire cette assertion, il *aurait été à même d'observer* qu'il se produisait un arc-en-ciel, et donc il *pouvait savoir* qu'il se produisait un arc-en-ciel.

Aucune des deux objections présentées par Williamson n'est par conséquent absolument décisive contre la stratégie d'Edgington, mais chacune d'entre elle manifeste néanmoins un certain embarras de la restriction proposée. Comme la théorie de Dummett ou celle de Tennant, il semble que la théorie d'Edgington laisse un résidu de vérités inconnaissables. Que reste-t-il, par conséquent, comme domaine pour la connaissabilité ? Dans la section qui suit, je propose d'examiner la question non plus du point de vue du vérificationniste anti-réaliste, mais plutôt de celui du réaliste positiviste. Le problème peut être envisagé de la manière suivante : si l'on admet, ne serait-ce qu'au bénéfice du doute, que toute vérité n'est pas connaissable, reste-t-il néanmoins une classe de vérités substantielles dont on puisse postuler la connaissabilité sans courir le reproche de circularité (comme dans le cas de Tennant), d'adhocité (comme dans celui de Dummett), ou encore d'obscurité (comme dans le cas d'Edgington) ?

[30] Voir J. Perry (1979), "The Problem of the Essential Indexical", *Noûs* 13, pp. 3-21, and D. Lewis (1979), "Attitudes *De Dicto* and *De Se*", *The Philosophical Review*, 88 (4), pp. 513-43.

4. Les vérités nécessaires et le positivisme

Dans un article sur les analogues temporels du paradoxe de Fitch, J.P. Burgess examine un principe qui ressemble dans une certaine mesure à la formulation du principe de connaissabilité proposée par Edgington, mais dont le sens est cette fois plus transparent, et que Burgess appelle le *principe de découverte*[31]. Il s'agit du principe selon lequel si, *à compter du moment présent*, p est vrai et continuera d'être vrai dans tous les moments qui suivent, alors il existe un moment futur tel que quelqu'un saura que p est vrai. En logique temporelle, le principe s'énonce de la façon suivante[32]:

(D) $Gp \rightarrow FKp$

A la différence du principe d'Edgington, dans l'antécédent duquel on avait Ap, Gp implique que p est vrai non seulement à l'instant présent, mais se perpétue à tous les instants qui suivent. Burgess montre que sous l'hypothèse que le temps est représenté de façon linéaire, et que l'opérateur K obéit aux axiomes habituels, le principe (D) ne donne pas lieu à la conclusion indésirable $p \rightarrow Kp$. En outre, le principe implique que si p est vrai à un moment donné, il y a un moment ultérieur où quelqu'un saura que p était vrai dans le passé ($p \rightarrow FKPp$). Si par exemple Smith a tué Jones, c'est là une vérité qui finira pas éclater au grand jour (« truth will out »), même si cela doit prendre du temps.

Notons que l'observation de Burgess ne vise pas à défendre le principe (D) comme un substitut du principe (V) initial, ni à défendre l'anti-réalisme : le contraire serait plutôt vrai, puisque le postulat implique que toute vérité contingente laisserait une trace derrière elle, susceptible de rendre manifeste cette vérité. Comme le souligne Burgess, la croyance en un principe tel que (D) dépend généralement d'une forme d'optimisme théologique et n'a plus qu'un rapport très lointain avec le vérificationnisme. Malgré cela, Burgess mentionne un analogue modal du principe de découverte, le principe selon lequel toute vérité nécessaire est connaissable :

(N) $\Box p \rightarrow \Diamond Kp$

Selon Burgess, « l'absence pour les possibilités d'un analogue évident des spécifications chronométriques pour les temps » rend le principe (N) « beaucoup moins satisfaisant » que (D). Mais est-ce nécessairement le cas ? Le principal attrait du principe (N) est qu'en déclarant connaissables toutes les vérités nécessaires, on déclare connaissables toutes les vérités mathématiques, et plus généralement toutes les vérités qui pourraient avoir le statut de lois fondamentales en vertu même de leur nécessité. Le principe est plus faible que le principe (D), puisqu'il n'implique pas que toute vérité contingente finira pas être connue, mais il a en revanche un fondement plus rationnel : c'est le caractère de permanence et de fondement des vérités nécessaires qui les rend accessibles à l'esprit en principe. Par ailleurs, dans une perspective rationaliste, les vérités nécessaires sont l'objet même de la science ; or on peut admettre de manquer de connaître certaines vérités contingentes, mais conformément au sentiment positiviste d'un Hilbert ou d'un Galilée, le domaine des vérités fondamentales doit du moins être accessible en principe.

---

[31] J.P. Burgess, *loc. cit*.

[32] L'opérateur F s'interprète comme: "il y a un instant futur tel que", et G, son dual ($\neg F \neg$), comme "tous les instants futurs sont tels que".

Qu'en est-il, alors, du paradoxe de Fitch ? Premièrement, on peut montrer que si l'on ajoute le schéma (N) à une logique de type S5 pour chacun des opérateurs $\Box$ et K, le schéma $p \to Kp$ n'en résulte pas logiquement[33]. Si toute vérité nécessaire est connaissable, il ne s'ensuit donc pas que toute vérité est connue. C'est le moins qu'on puisse exiger, cependant, dans la mesure où le principe (N) laisse ouverte la possibilité que certaines vérités contingentes ne soient pas connaissables. Mais par ailleurs, si toute vérité nécessaire est connaissable, il ne s'ensuit pas non plus que toute vérité *nécessaire* est connue *ipso facto*[34]. On évite de la sorte le problème que rencontrait Edgington dans le cas des vérités actuelles, qui était que toutes les vérités actuelles étaient par là même connues. Le principe (N) reste malgré tout relativement proche du principe ($\Diamond$KA) d'Edgington, qui impliquait que toutes les vérités actuelles étaient par là même nécessaires ; mais en faisant dépendre la connaissabilité de la nécessité, et non plus de l'actualité, on évite la seconde objection faite par Williamson au sujet de la connaissance non-actuelle des vérités actuelles. En effet, une proposition nécessaire est une proposition vraie dans tous les mondes possibles, y compris dans les mondes qui ne sont pas actuels. Il n'est plus mystérieux, dans ce cas, de postuler une connaissance non-actuelle des vérités nécessaires.

Moins ambitieux que le schéma d'Edgington, le schéma (N) apparaît donc comme mieux fondé. A la différence des schémas proposés par Dummett et Tennant, par ailleurs, le schéma n'est pas motivé purement par le souci d'éviter le paradoxe de Fitch ou de valoir comme un *substitut* au principe $p \to \Diamond Kp$, ce qui ne saurait être le cas. Le fait que le principe (N) ne donne pas lieu à une conclusion de type Fitch montre que le domaine des vérités connaissables est susceptible d'inclure *au moins* les vérités nécessaires. Mais (N) ne dit plus rien touchant la connaissabilité des vérités contingentes[35]. Considérons alors le positiviste optimiste décrit précédemment : pour celui-ci, l'argument de Fitch montre que certaines vérités contingentes sont inconnaissables pour des raisons structurelles, et oblige à renoncer à l'idée que toute vérité est connaissable. Pour autant, il reste cohérent de penser que toutes les vérités nécessaires sont connaissables par principe.

Naturellement, l'immunité du schéma (N) vis-à-vis d'un argument à la Fitch ne *prouve* pas que toute vérité nécessaire est *effectivement* connaissable. En tant que tel, le principe (N) demeure un postulat métaphysique au même titre que (V). Il se pourrait fort bien, en particulier, qu'il existe certaines vérités que nous appellerons *pascaliennes*, des vérités nécessaires et telles que nécessairement, personne ne sache qu'elles sont vraies. Si de telles vérités existent (comme par exemple « Dieu existe » pour Pascal), alors le schéma (N) est tout simplement contradictoire et faux. Rien, cependant, ne contraint à affirmer l'existence de telles vérités *nécessairement inconnaissables*, contrairement au cas des vérités *présentement inconnues* de nous, dont l'existence ne semble pas faire de doute.

---

[33] Pour le montrer, on peut vérifier que le schéma (N) est valide dans une classe de cadres $\langle W,R_K,R_\Box\rangle$ (où R_K est la relation d'accessibilité épistémique, et $R_\Box$ la relation pour la nécessité) si et seulement si les relations satisfont: $\forall x \exists y(xR_\Box y \wedge \forall z(yR_K z \to xR_\Box z))$. Or on peut construire un modèle basé sur ce cadre (voir la note qui suit), dans lequel les relations R_K et $R_\Box$ sont des relations d'équivalence (S5), qui ne valide pas $p \to Kp$.

[34] Considérer le modèle à quatre mondes, x,x',y,y', où p est vrai partout sauf en x', muni de la clôture réflexive, symétrique et transitive des relations R_K={xx'}, et $R_\Box$={xy,x'y'}. Le modèle valide le schéma $\Box p \to \Diamond Kp$, x satisfait $\Box p$, mais x ne satisfait pas Kp. Soulignons au passage que « nécessaire » signifie « vrai dans tous les mondes $\Box$-accessibles » (et non pas « vrai partout dans le modèle », ce qui aurait alors pour effet que tout ce qui est nécessaire est connu).

[35] Comparer à la remarque de Williamson (1987), *loc. cit.*, p. 257, qui écrit à propos du schéma ($\Diamond$KA) d'Edgington : « the only knowledge that [it] requires is of necessary truths. One might expect a robust form of verificationism to insist that at least some contingent truths are knowable".

5. Conclusion

Pour finir, ajoutons quelques mots des liens entre positivisme, réalisme et antiréalisme. Pour un anti-réaliste, le principe de connaissabilité est censé suivre de l'analyse du concept même de vérité. Pour un réaliste, en revanche, le principe selon lequel toute vérité est connaissable d'un esprit fini devrait apparaître comme une forme d'immodestie épistémologique, comme le soutient Williamson. Plus précisément, les énoncés de type Moore exhibent de telles vérités inconnaissables. En outre, l'inconnaissabilité des vérités de type Moore n'est pas éliminable à proprement parler, si l'on suit les critiques de Williamson à l'encontre de la solution d'Edgington : chaque fois que nous sommes dans l'ignorance d'une vérité contingente p, l'énoncé « p est actuellement vrai » ne peut pas nous être connu de façon non-actuelle, sauf à en perdre le caractère indexical. L'existence de telles vérités inconnaissables, et c'est là le propos de cet essai, n'exclut pas pour autant de retenir d'autres postulats métaphysiques, plus modestes et fondés indépendamment, comme le postulat selon lequel toute vérité nécessaire est connaissable. Notons, cependant, que ce postulat reste un postulat métaphysique, à l'égal du principe de découverte de Burgess, et que certains réalistes positivistes pourraient s'abstenir d'y adhérer. Un réaliste tel que Popper, comme le remarque John MacFarlane (c.p.), compte certainement comme positiviste. Malgré cela, Popper ne pensait pas toute vérité était connaissable, et pas même, sans doute, que toute vérité nécessaire était connaissable. Ainsi, pour Popper, nos meilleures théories scientifiques pourraient chaque fois manquer d'être vraies. Mais un philosophe comme Popper adhérait à un postulat méthodologique à certains égards plus utile que le principe métaphysique selon lequel toute vérité nécessaire est connaissable. Selon Popper, toute théorie empirique valable doit être formulée de telle manière que la théorie soit réfutable dans l'éventualité où elle est fausse. Dans la perspective méthodologique de Popper, il s'agit avant tout d'un réquisit d'informativité et de précision des théories scientifiques. Quant à ses implications métaphysiques, le principe popperien constitue cependant un cas particulier du principe de découverte formulé par Burgess, puisqu'il suppose que si une théorie scientifique est fausse, on doit pouvoir apprendre en principe qu'elle est fausse (« falsity will out »). Mais bien que plus modeste en apparence que le principe de découverte dans sa version positive, on voit que ce principe dépend lui aussi d'une forme minimale d'optimisme épistémologique, en excluant l'éventualité improbable qu'un mauvais génie nous laisse éternellement dans l'illusion de la vérité de théories fausses.