Knowability and intuitionistic logic

Note de lecture, UE 904 EC 1

Léo Vacher*

2023

1 Introduction

La logique modale ajoute à la logique classique des opérateurs exprimant la nécessité $\Box p$ et la possibilité $\Diamond p$ d'une proposition p. Similairement, l'ajout des opérateurs de connaissance $\mathbf{K}p$ et de croyance $\mathbf{B}p$, permet de fonder la logique épistémique qui représente un outil formel puissant pour l'épistémologie, car elle permet de représenter l'information accessible à un sujet [1].

Certaines positions philosophiques et leurs conséquences peuvent être sondées en utilisant la logique épistémique. C'est par exemple le cas de l'anti-réalisme dit modéré, caractérisé par la vérité des propositions du type *toutes les vérités sont connaissables* à savoir que si une proposition est vraie, il est possible de la connaître et en conséquence *rien ne peut être vrai et inaccessible à la connaissance*. Une telle thèse est au fondement de nombreux systèmes épistémologiques et philosophiques sous des formes plus ou moins fortes. On pourra par exemple citer certaines approches constructivistes en philosophie des mathématiques (une proposition mathématique p est vraie uniquement si une preuve formelle ou mentale peut en être fournie, voir e.g. Bridges, Palmgren et Ishihara (2022)), le vérificationisme et l'empirisme logique (voir e.g. Creath (2022)) ou certaines formes d'idéalisme comme celui de Berkeley (1710) en métaphysique ("être c'est être perçu").

Étonnamment, une fois insérée dans un système de logique épistémique classique, la thèse anti-réaliste modérée semble amener à la conclusion beaucoup plus forte que *toute vérité doit être connue*, basculant vers l'anti-réalisme dit radical : c'est le paradoxe de Fitch (1963). Quelle que soit la position que l'on souhaite défendre, il apparaît alors urgent de comprendre et de questionner les mécanismes amenant à un tel basculement et éventuellement chercher à le contrer. Ainsi, le paradoxe de Fitch a occupé une place centrale dans la littérature épistémologique contemporaine [2].

Dans ce court essai, nous présenterons la discussion de Vidi et Solomon (2001), dans laquelle les auteurs proposent de contrer le paradoxe en remplaçant la logique classique par la logique intuitionniste et sa sémantique au sein des modèles de Kripke. En section 2, nous commencerons par résumer les arguments des auteurs en suivant l'ordre de l'article, puis nous discuterons son contenu et ses extensions possibles en section 3. En guise de compléments, les modèles de Kripke sont brièvement présentés en annexe A et une dérivation du paradoxe de Fitch est donnée en annexe B.

*vacher.leo.etu@gmail.com

1. Pour une introduction à la logique modale et épistémique, nous renvoyons à la lecture de Garson (2023).
2. Pour une vue d'ensemble sur le paradoxe de Fitch, nous renvoyons à la lecture de Brogaard et Salerno (2019)).

2 Résumé de l'article

2.1 Le paradoxe de Fitch

Au sein d'un système de pensée, la position de l'anti-réalisme modéré peut s'énoncer comme suit : *soit p une proposition. Si p est vraie, alors il est possible de connaître p.* Dans le cadre de la logique épistémique, on est alors tenté de formaliser cet énoncé comme

$$p \to \Diamond \mathbf{K}p, \tag{ARM}$$

où $\to$ est l'implication logique et la combinaison de l'opérateur de possibilité $\Diamond$ et de connaissance $\mathbf{K}$ permettent de former la proposition *il est possible de connaître p* ($\Diamond\mathbf{K}p$). Le paradoxe de Fitch (1963) apparaît comme une conséquence problématique de l'inclusion de cette proposition au sein d'un système de logique épistémique classique cohérent, en supposant uniquement quelques propriétés élémentaires pour l'opérateur $\mathbf{K}$. En effet, comme détaillé en appendice B, supposant (ARM), il est possible de dériver

$$p \to \mathbf{K}p, \tag{ARR}$$

s'interprétant comme : *si p est vraie, alors p est connue.* Cette dernière position est plus proche d'un anti-réalisme radical comme le solipsisme : tout ce qui n'est pas connu par un agent n'est pas vrai, voire n'existe pas [3]. Si (ARM) conduit logiquement à (ARR), doit-on alors conclure que le réalisme modéré se réduit au réalisme radical ? Il serait alors tentant d'utiliser cet argument contre toute forme d'anti-réalisme ou en faveur de l'anti-réalisme radical. Pour autant, il est légitime de douter que l'on puisse établir une conclusion si forte à partir d'une dérivation si simple. De nombreuses autres pistes ont été proposées afin de mieux cerner ce paradoxe et ses conséquences pour l'anti-réalisme. Naturellement, il est possible de remettre en question la formalisation proposée par (ARM), préférant une autre formulation ou l'utilisation de nouveaux opérateurs (voir e.g. Edgington (1985)). On peut également soutenir que la question du réalisme est indépendante de celle de la vérité (Horwich 1999) ou qu'elle échappe à la formalisation syntaxique (Morgan 1973), argumentant ainsi que les outils de la logique épistémique ne sont tout simplement pas les bons pour aborder correctement une telle question. Il est également envisageable de proposer un cadre interprétatif différent des opérateurs logiques, et en particulier de $\mathbf{K}$ comme étant un opérateur de connaissance implicite. Finalement, une suggestion largement discutée a été de remettre en question le cadre logique utilisé pour dériver (ARR) depuis (ARM). C'est cette dernière option que souhaite discuter plus en avant Vidi et Solomon (2001), proposant de remplacer le cadre de la logique classique par celui de la logique intuitionniste. Ce faisant, nous verrons que les auteurs proposent également un nouveau cadre interprétatif pour les opérateurs ainsi qu'une formalisation alternative pour le principe de l'anti-réalisme modéré.

2.2 La réponse de la logique intuitionniste

La logique dite intuitionniste propose une alternative à la logique classique dans laquelle la vérité d'une proposition dépend de la possibilité ou non d'en construire/établir une preuve (Brouwer 1952) [4]. Dans ce nouveau cadre, la validité de certaines propositions de la logique classique, comme l'identité de la double négation $\neg\neg p \Leftrightarrow p$, ne peut plus être établie. La position anti-réaliste est souvent associée à une forme de constructivisme, tout comme le choix d'adopter une logique intuitionniste, on peut ainsi facilement argumenter qu'un tel cadre logique

3. Pour une présentation des différentes formes plus ou moins radicales d'idéalisme, voir e.g. Guyer et Horstmann (2023).

4. Pour une présentation plus complète de la logique intuitionniste voir e.g. Moschovakis (2023).

est préférable et plus naturel pour analyser le paradoxe de Fitch. Williamson (1982) propose ainsi que le paradoxe n'a plus lieu d'être dans un cadre intuitionniste. En effet, retirant la possibilité d'éliminer les doubles négations, il n'est alors plus possible d'utiliser R4 comme détaillé en annexe B, rendant impossible la démonstration de (ARR) à partir de (ARM). La seule conclusion à laquelle on peut ainsi aboutir en suivant la démonstration devient

$$p \to \neg\neg \mathbf{K}p. \tag{ARI}$$

On peut alors débattre de la "gravité" épistémique d'une telle conclusion (classiquement : *si p est vraie, alors on ne peut pas ne pas connaître p*), qui semblerait à priori aussi terrible que (ARR) pour l'anti-réalisme modéré. Cependant, comme nous allons le discuter plus avant, les propositions comme (ARI) ne sont problématiques que si l'on fait une lecture classique des opérateurs logiques et adopter la logique intuitionniste demande également une interprétation intuitionniste de ceux-ci. En suivant Williamson (1992), Vidi et Solomon (2001) proposent alors de revoir l'interprétation des opérateurs logiques au sein de la sémantique donnée par les modèles de Kripke, afin de montrer comment un tel changement de logique peut dissoudre le paradoxe.

2.3 Une relecture intuitionniste de la logique épistémique

La logique modale propositionnelle est généralement présentée dans le cadre des modèles de Kripke, détaillés en appendice A. Il est alors crucial de chercher à interpréter et à raffiner ces modèles pour les rendre compatible avec 1) une interprétation des modèles favorable à une compréhension épistémique en terme d'information accessible à un sujet 2) l'application possible de la logique intuitionniste[5]. Pour cela, Vidi et Solomon (2001) suggèrent de penser les mondes comme des "états d'information" plutôt que des mondes possibles. Si il existe une relation $R(w_1 w_2) \in \mathcal{R}$ entre deux mondes, on comprendra que w_1 a accès à w_2 signifiant que w_2 est une extension possible de l'état d'information du sujet donnée par w_1.

Dans ce nouveau cadre, une définition alternative des opérateurs logiques devient possible. En particulier, on considèrera $\neg p$ comme vraie dans un monde w_1 ($\mathcal{M}, w_1 \vDash \neg p$) si et seulement si il n'existe aucun monde w_2 accessible depuis w_1 tel que p soit vraie ($\mathcal{M} w_i \vDash \neg p$ ssi pour tout $w_j \in W$ tel que $R(w_i, w_j) \in \mathcal{R}$ existe, $w_j \nvDash p$[6]). Ainsi $\neg$ exprime une impossibilité plutôt qu'une fausseté. Similairement on considèrera $p \to q$ comme vraie dans un état d'information w_1 donné lorsque pour tous les états w_j liés à w_1, si p est vraie alors q est aussi vraie ($\mathcal{M}, w_i \vDash p \to q$ ssi pour tout w_j tel que $R(w_i, w_j)$ existe, si $w_j \vDash p$ alors $w_j \vDash q$).

Finalement, on changera la notation pour la relation entre deux mondes $R(w_1 w_2)$ en $w_1 \leq w_2$ pour désigner que w_2 est une extension de l'état d'information de w_1, possiblement à un instant ultérieur. Pour mettre en place une telle gradation entre les états, on imposera à $w_1 \leq w_2$ d'être réflexive – $\forall w_i, w_i \leq w_i$ i.e. il existe toujours des relations des mondes sur eux-mêmes – et transitive – si $w_1 \leq w_2$ et $w_2 \leq w_3$ alors $w_1 \leq w_3$ (identiquement à la relation d'ordre $\leq$ en théorie des ensembles). On encode alors la possibilité que l'information d'un sujet puisse grandir avec le temps, en s'appuyant sur des connaissances passées. La transitivité combinée à la définition de la négation assure que, lorsque $\neg p$ est vérifiée dans un monde donné, elle le sera également dans tous les mondes qui peuvent lui succéder. On dira que la proposition est *forcée*.

5. Voir également à ce sujet la définition plus formelle mais limpide des Intuitionistic Kripke models (IKM) donnée par Proietti (2012).

6. A comparer à la définition usuelle $\mathcal{M} w_i \vDash \neg p$ ssi $\mathcal{M} w_i \nvDash p$

2.4 La double négation comme opérateur et la solution au paradoxe

Cherchons alors à comprendre comment interpréter (ARI) à l'aide du nouveau cadre proposé dans la section précédente afin d'établir si elle est épistémologiquement moins problématique que (ARR). Dans la succession de Došen (1984), les auteurs cherchent alors à interpréter la double négation $\neg\neg p$ comme un opérateur modal, idéalement traduisant une forme de possibilité. Cependant, $\neg\neg$ semble plutôt obéir aux propriétés associées à $\Box$ [7]. Malgré cela, comme $\neg\neg$ exprime l'impossibilité d'une impossibilité, on cherchera plutôt à l'interpréter comme une possibilité forte, ou une sorte d'opérateur hybride comme une "nécessité particulière". Revenons alors au cadre sémantique proposé en section 2.3 et aux définitions données pour $\to$ et $\neg$. Dans ce cadre, la vérité de (ARR) dans un état d'information donné ($\mathcal{M}, w \vDash p \to \mathbf{K}p$) signifie que dans tous les autres états d'informations accessibles depuis celui-ci, si p est vraie, alors $\mathbf{K}p$ l'est aussi. Autrement dit : la vérité implique la connaissance et il n'y a pas d'état d'information possible dans lequel il y a des informations vraies qui ne soient pas connues. Cette position est donc bien celle de l'anti-réalisme radical et doit être rejetée par l'anti-réalisme modéré. (ARI) au contraire encode que : lorsque $p \to \neg\neg\mathbf{K}p$ est vraie dans un monde w_1, alors dans tous les mondes w_i accessibles depuis w_1, si p est vraie dans w_i alors n'y a pas d'état ω tel que $w_i \leq \omega$ dans lesquels $\neg\mathbf{K}p$ est vrai, c'est-à-dire qu'il est impossible que nous puissions nier que p sera connu. Autrement dit, (ARI) affirme que : sachant que p est vraie, alors il est possible de connaître p, c'est à dire qu'il existe des extensions possible de l'état d'information présent dans lequel on pourra établir $\mathbf{K}p$ (même si le nombre d'étapes intermédiaires ou le fait que cela soit un jour réalisé n'est pas déterminé). Ainsi si p est vraie, alors la possibilité que p soit un jour connue reste toujours ouverte [8]. On comprend alors que en utilisant cette sémantique, (ARI) permet de capturer parfaitement l'anti-réalisme modéré pour lequel (ARM) a été originellement invoqué. Pour arriver à une telle conclusion, il faut implicitement assumer le principe de indiscernabilité faible : à partir d'un état d'information donné, il sera possible de connaître p ou $\neg p$ à un instant ultérieur et indéterminé, sans pour autant qu'il ne soit jamais réalisé, représentant une contrepartie à la contrainte sur les valeurs de vérité des propositions ($p \vee \neg p$). Autrement dit, il n'existe pas de proposition pour laquelle il sera impossible d'imaginer ni une forme de preuve de vérité ni une forme de preuve d'impossibilité de dériver une preuve. Cette hypothèse semble bien également en accord avec le principe d'anti-réalisme modéré tel que l'on souhaite l'implémenter [9].

2.5 Des conséquences suspicieuses

L'utilisation de la logique intuitionniste permet donc d'éviter (ARR) tout en acceptant la conclusion donnée par (ARI). Cependant, il a été rétorqué que d'autres conclusions aussi problématiques que (ARR) peuvent être dérivées à partir de (ARM) ou de (ARI) dans un

7. En logique intuitionniste, on peut montrer que $\neg\neg\neg p \leftrightarrow \neg p$. De cette propriété, on peut déduire que $\neg\neg$ se distribue sur la conjonction mais pas la disjonction, propriété associée à $\Box$ mais pas à $\Diamond$. De plus, $(\neg\neg)p \wedge \neg(\neg\neg)p$ est toujours contradictoire, ce qui est le cas de $\Box p \wedge \neg\Box p$ mais pas – en général – de $\Diamond p \wedge \neg\Diamond p$. Comme $\neg\neg p$ n'implique pas p (il ne vérifie pas l'axiome (**T**)), il ne peut pas toujours être identifié $\Box$, mais pourrait être interprété comme un opérateur de nécessité déontique. Pour comprendre $\neg\neg$ comme une possibilité forte plutôt qu'une nécessité, il est cependant possible de l'interpréter comme étant équivalent à $\Box\Diamond$. A ce sujet voir Goldblatt (2003).

8. Une excellente explication de $\neg\neg p$ et de ses conséquences dans un cadre intuitionniste peut également être trouvée dans Dummett (2009).

9. Si p est vraie alors il peut être connu, il doit en aller de même pour l'impossibilité de prouver p : $\neg p$ (si $\neg p$ est vraie alors il peut être connu).

contexte intuitionniste. Percival (1990) dérive trois conséquences de ce type données par [10]

$$\neg \mathbf{K}p \rightarrow \neg p, \tag{I}$$
$$\neg \mathbf{K}p \leftrightarrow \neg p, \tag{II}$$
$$\neg(\neg \mathbf{K}p \wedge \neg \mathbf{K}\neg p). \tag{III}$$

Voyons comment le cadre donné par les modèles intuitionniste de Kripke permet d'interpréter ces propositions :

(I) Une lecture classique interprétait cette proposition comme problématique car ce n'est pas parce qu'on ne connait pas p que $\neg p$ est vraie. Comme le fait remarquer Percival (1991), cette proposition est particulièrement problématique pour les propositions empiriques : supposons que nous ne connaissons pas p et que p est un fait empirique, devrait on conclure pour autant que $\neg p$ est vraie ? Pour répondre à cela, rappelons nous désormais de la définition de $\neg$ donnée en section 2.3 : celle-ci dépend de la possibilité ou non qu'une proposition soit vraie dans des états d'informations futurs. $\neg \mathbf{K}p$ ne doit pas être compris comme *p n'est pas connue* mais plutôt comme *p est inconnaissable*[11]. Dans ce cadre, on retrouve donc bien la contraposée du principe de l'anti-réalisme modéré désiré : *si p est inconnaissable alors p n'est pas vrai.*

(II) Ici encore, la proposition semble en adéquation avec l'anti-réalisme modéré pour les propositions mathématiques, mais problématique pour les fait empiriques. Percival (1990) propose l'exemple suivant : imaginons un chercheur menant une étude sur un groupe de lapins. Lorsque celui-ci a le dos tourné, une mère lapine appelée Mabel met un de ses petits appelé Peter dans la portée d'une autre lapine appelée Doris. Si le chercheur ne se retourne pas, il ne saura pas que $p =$*Peter est la progéniture de la lapine Mabel.* Ainsi $\neg \mathbf{K}p$ est vraie. Pour autant, il est clair que $\neg p$ est faux. Ainsi $\neg p$ ne peut pas être logiquement équivalent à $\neg \mathbf{K}p$ comme affirmé par (II). Ici encore, il faut alors comprendre $\neg \mathbf{K}p$ comme *p est inconnaissable.* On pourra alors assumer que même si p n'est pas connu, d'autres moyens (e.g. une analyse ADN) permettraient de conclure p, de sorte que $\neg \mathbf{K}p$ soit faux. Pour cela, il faut bien comprendre que $\leq$ ne peut pas être interprété comme une relation d'ordre temporel i.e. $w_1 \leq w_2$ ne peut pas être interprété comme *w_2 est un état d'information futur pour w_1* mais plutôt comme *w_2 est un état d'information représentant une extension cohérente de w_2.*

(III) Percival (1990) interprète (III) classiquement comme *aucune proposition ne reste pour toujours non décidée*, considéré comme inacceptable face à des contre-exemples tels que le suivant : imaginons que je tire dix fois à pile ou face sans regarder le résultat. Ni moi ni personne ne saura jamais si face est tombé 5 fois, laissant p pour toujours non décidée. A nouveau, comprenant $\neg \mathbf{K}p$ comme *p est inconnaissable*, on comprend $\neg \mathbf{K}p \wedge \neg \mathbf{K}\neg p$ comme *p est inconnaissable et $\neg p$ est inconnaissable*, en contradiction avec le principe de décidabilité faible discuté en section 2.4 qui a été clairement identifié comme une hypothèse pour la validité de notre système de pensée. La négation de cette proposition apparaît donc comme nécessaire dans le cadre de que nous avons imposé. Ainsi (III) ne doit pas être compris comme *aucune proposition ne reste pour toujours non décidée* mais plutôt *Il n'est pas possible que p et $\neg p$ soient tous deux inconnaissables.* Une fois les pièces lancées sans regarder, on peut admettre que l'on soit entrainé dans des états

10. Partant de (ARI), $p \Leftrightarrow \neg\neg p$, on trouve $\neg\neg\neg p \rightarrow \neg p$ par contraposition (R3). Comme en logique intuitionniste la triple négation est équivalente à une négation unique, on obtient bien (I). Prenant maintenant la contraposée (R3) de (I) on retrouve l'axiome (**T**) : $\mathbf{K}p \rightarrow p$ que l'on sait valide, ainsi on a donc bien montré (II). Enfin et utilisant deux fois (R3) sur (I), il est possible de dériver $p \wedge \neg p \leftrightarrow \bot$, montrant donc que (III).

11. Rappelons nous ici que $\leq$ est transitive, ainsi si $\mathcal{M}, w_1 \vDash \neg p$, alors $\neg p$ est vraie dans tous les états d'informations accessibles ultérieurs (forcée).

d'information où la vérité de p est devenu inconnaissable : $\neg\mathbf{K}p$, impliquant d'après (I) qu'une preuve de p ne pourra jamais être construite ($\neg p$) mais on pourra alors admettre que l'on sait que aucune preuve de p ne pourra être construite ($\mathbf{K}\neg p$), de sorte que l'on n'ait pas à reconnaître que $\neg\mathbf{K}p \wedge \neg\mathbf{K}\neg p$ à savoir que p est à la fois inconnaissable et qu'il est impossible de savoir si l'on ne peut pas construire de preuve de p.

Ainsi, les trois propositions problématiques proposées par Percival (1990) peuvent être contrées dans le nouveau cadre épistémique intuitionniste proposé ici.

2.6 Logique modale et intuitionnisme : un nouveau principe pour l'anti-réalisme modéré

Nous avons vu que, à défaut d'être problématique, (ARI) encode précisément la position de l'anti-réalisme modéré dans le cadre de la logique sémantique intuitionniste. Vidi et Solomon (2001) nous invitent alors à penser (ARI) comme une formalisation appropriée de la thèse de l'anti-réalisme modéré, préférable à (ARM). (ARI) n'utilise que des opérateurs d'implications et de négations bien connus et définis dans le cadre de la logique intuitionniste. De l'autre côté, (ARM) combine les opérateurs épistémiques et ceux de la logique modale classique en faisant appel à l'opérateur modal $\Diamond$, combinaison encore mal comprise dans le cadre de la logique intuitionniste. En effet, à partir de l'inter-définition des opérateurs $\Diamond$ et $\Box$ ((POS) et (NES)) en logique modale classique, on peut en effet dériver [12]

$$\neg\neg\Diamond p \Leftrightarrow \Diamond p. \tag{1}$$

$$\neg\neg\Box p \Leftrightarrow \Box p. \tag{2}$$

Il semble alors que le principe de l'annulation des doubles négations que nous souhaitions rejeter originalement en logique intuitionniste soit valide pour les propositions modales. Il s'agit là d'un problème majeur en quête d'une logique intuitionniste modale alternative à la logique classique. Face à cela, plusieurs pistes ont pu être proposées : comme suggéré par les auteurs dans un travail antérieur (DeVidi et Solomon 1997), on peut choisir de rejeter la sémantique de Kripke et les interprétations intuitionnistes des connecteurs logiques associés à la recherche d'un système alternatif ne permettant plus l'annulation de la double négation des propositions modales. Cependant, une telle approche ne pourrait pas aider ici car la signification des opérateurs $\Box$ et $\Diamond$ se voit significativement transformée. On pourrait également rester dans le cadre des modèles de Kripke mais rejeter la définition des opérateurs modaux à travers (POS) et (NES), préférant une définition sous forme des énoncés donnés en appendice A. Dans ce cas, on se trouve avec deux relations d'accessibilités : une pour $\mathbf{K}$ et une pour $\Box/\Diamond$ et il n'est pas techniquement évident de savoir comment gérer ces deux relations. Il est également possible de changer l'interprétation de $\Box$ et $\Diamond$ au sein des modèles de Kripke, au risque de perdre l'intérêt principal de ces modèles pour étudier la nécessité et la possibilité en philosophie. Pour toutes ces raisons, on préfèrera donc (ARI) plutôt que (ARM), car cette formalisation n'utilise que des connecteurs logiques non modaux, bien définis et connus dans le cas des modèles intuitionniste de Kripke. Dans tous les cas, l'étude du paradoxe de Fitch dans le cadre de la logique intuitionniste ne fait que commencer et la littérature est maigre sur ce sujet. Il restera encore beaucoup à faire afin de cerner comment bien l'aborder dans ce cadre.

12. Partant de (POS) et appliquant deux fois la contraposition (R3), on obtient $\neg\neg\Box\neg p \Rightarrow \neg\Diamond p$ puis $\neg\neg\Diamond p \Rightarrow \neg\neg\neg\Box\neg p$. La triple négation étant équivalente à la négation simple en logique intuitionniste, on retrouve $\neg\neg\Diamond p \Leftrightarrow \Diamond p$ en réinsérant la définition donnée par (POS). Identiquement, on peut trouver $\neg\neg\Box p \Leftrightarrow \Box p$ pour la nécessité en partant de (NES).

3 Discussion

L'article présenté ci-dessus soulève de nombreux points qui méritent d'être discutés. En un premier temps, on peut se demander si le choix d'une approche intuitionniste ne rend pas trivial le paradoxe de Fitch. Après tout, choisir la logique intuitionniste c'est également choisir le cadre interprétatif qui lui est associé, à savoir que p est vraie si il est possible d'en construire une preuve (ou de le vérifier autrement pour les propositions non mathématiques e.g. par les sens). On peut alors se demander si des propositions comme (ARM), (ARI) ou même (ARR) ne se retrouvent pas trivialisées, la logique épistémique incluant déjà par essence l'anti-réalisme modéré. En effet, pour que p soit vraie, je dois pouvoir en construire une preuve et ainsi il me sera possible de connaître p. Ce point a déjà été soulevé par exemple par Martino et Usberti (1994) ou Khlentzos (2004). Si tel est le cas, alors il n'est pas nécessaire de développer une sémantique particulièrement riche pour la logique épistémique intuitionniste face au paradoxe de Fitch, puisqu'elle inclut déjà naturellement en elle la position de l'anti-réalisme modéré et la solution au paradoxe. Cependant, comme discuté par Murzi (2010), un tel raccourci est peu convaincant et sa validité dépend fortement du cadre dans lequel on essaie d'appliquer cette logique (mathématique, connaissance empirique d'un sujet ...) ainsi que ce que l'on entend par *établir/construire un preuve de p* soit finalement quelle ontologie on associe aux preuves. De nombreux arguments sont avancés par Murzi (2010), que nous ne retranscrirons pas tous ici, mais on évoquera par exemple le problème de savoir si on doit penser les preuves exigées pour affirmer p comme des types abstraits ou des réalisations particulières ("token"). Penser les preuves comme des types, i.e. comme des classes d'objets, potentiellement "Platonicien" i.e. indépendant d'un esprit humain, ou comme des classes d'objets possibles *en principe pour un être omniscient* rend impossible la trivialisation du paradoxe de Fitch, car l'existence d'une preuve ne s'identifie plus avec la connaissance de celle-ci ni même avec la possibilité d'un certain sujet particulier d'en acquérir la connaissance. Si on pense aux preuves comme des 'tokens', des objets actuellement réalisés, on doit alors se confronter à d'autres problèmes. Par exemple, si la vérité d'une proposition est identifiée avec l'existence d'une preuve pensée comme un objet (i.e. dans un livre), on peut imaginer qu'une proposition devienne fausse ou du moins non vraie si toutes les preuves finissent par être perdues pour une certaine raison. Ainsi la vérité pourrait se "gagner" ou se "perdre" dans le temps. On peut alors se demander si considérer que la vérité puisse être une fonction du temps est acceptable.

Finalement il apparaît donc assez clair que la logique épistémique intuitionniste (et certainement tout autre cadre formel épistémique) requiert absolument une implémentation de la temporalité pour pouvoir modéliser la connaissance, justifiant ainsi l'argumentaire proposé par DeVidi et Solomon (1997) et la nécessité d'utiliser des diagrammes de Kripke pour la retranscrire. Cependant, on peut également se demander si cette implémentation répond complètement aux problèmes posés par cette temporalité. De plus, il me semble également qu'en réinterprétant la négation comme une impossibilité plutôt qu'une fausseté ($\neg p$ devient : *il n'est pas possible de construire une preuve de p* plutôt que *p est faux*), la logique intuitionniste permet de donner des interprétations beaucoup plus acceptables de l'ensemble des conséquences du réalisme modéré mais l'affaiblit en partie, car il encode bien que *si p est vraie alors il est connaissable* mais avec la nouvelle sémantique associée à $\neg$, il n'y a plus de distinction entre *p est faux* et *p ne peut pas être prouvé*, brouillant les pistes sur la distinction entre impossibilité et fausseté. Pourtant, on peut bien imaginer des propositions vraies sans pouvoir en construire de preuves, en particulier si elles ont eu lieu dans le passé. Ces deux problèmes sont particulièrement apparents lors de l'utilisation nécessaire du principe de "décidabilité faible" et la réponse correspondante apportée à (III). Concrètement il s'agit de savoir comment gérer un problème comme celui du lancé de pièces aveugle, dans lequel aucune information n'est accessible après le lancé i.e. le cas des vérités à jamais inconnues. Dans ce cas, il faudra s'assurer que

$\neg\mathbf{K}p \wedge \neg\mathbf{K}\neg p$ soit bien une contradiction i.e. qu'il soit impossible que p et $\neg p$ soient tout deux inconnaissables. Si tout élément de preuve est perdu et que p est inconnaissable ($\neg\mathbf{K}p$), on est forcé de conclure que $\neg p$ par (I), que l'on comprendra comme *il est impossible de prouver* p [13] i.e. tous les états d'informations possible qui nous sont accessibles ne prouvent pas p et non comme p *est faux* [14]. Si on conclut de cela que $\neg p$, on peut alors sans mal accepter que on sait $\mathbf{K}\neg p$, satisfaisant à la contrainte imposée par (III). Mais peut on vraiment prouver en pratique que des propositions sont inconnaissables ? Si tel n'est pas le cas, il faudra accepter que p n'est pas inconnaissable et qu'il reste toujours des embranchements dans les mondes – jamais réalisés – où l'on pourrait remonter au résultat du pile ou face en calculant la trajectoire exacte de chaque atome pour savoir si face est tombé 5 fois [15] (comme le suggèrent DeVidi et Solomon (1997), ces états n'ont pas à être des états futurs mais simplement des états cohérents, même si les critères de cohérence restent à définir clairement). Ainsi on peut questionner sur l'applicabilité de la sémantique intuitionniste de Kripke pour des vérités contingentes ou éphémères (e.g. *il fait beau aujourd'hui*) et chercher à clarifier explicitement ses conséquences.

Dans ce sens, il serait également intéressant de se demander comment penser la discernibilité faible (et donc de (III)) face à l'existence d'énoncés indécidables au sens de Gödel, à savoir des énoncés pour lesquels il n'est ni possible de montrer p, ni $\neg p$ (et ainsi pour lesquels $\neg\mathbf{K}p \wedge \neg\mathbf{K}\neg p$ ne semble pas être une contradiction).

Dans la continuité de DeVidi et Solomon (1997), nous avons ici discuté le rôle crucial que peuvent jouer les interprétations intuitionnistes de $\neg$. Cependant, il faudrait également discuter plus avant la relecture proposée de $\rightarrow$, également centrale dans le paradoxe de Fitch et sa résolution. Par soucis de concision, nous ne nous aventurerons pas plus loin dans cette direction ici.

Quoi qu'il en soit, il est souhaitable de chercher à formaliser complètement les modèles intuitionnistes de Kripke introduit par Vidi et Solomon (2001) afin d'en explorer plus profondément tous les aspects et toutes les conséquences. En particulier, il faudrait mieux comprendre comment interpréter les états d'informations w et surtout la relation $\leq$ qu'il peut y avoir entre eux : qu'est-ce qu'un état d'information cohérent pouvant succéder à un autre ? Ne devrait-on pas établir une distinction entre les états d'informations possiblement accessibles par une évolution temporelle et les états d'informations accessibles en principe par le sujet ? Pour cela, il faudrait également chercher à définir formellement $\mathbf{K}$ et à établir toutes les propriétés auxquelles il peut satisfaire dans un contexte intuitionniste. En particulier, il faudra éviter que cette définition conduise à l'effondrement de la connaissance en vérité et se demander quels axiomes parmi les usuels ($\forall$ propositions ϕ, ψ)

$$\mathbf{K}(\phi \rightarrow \psi) \rightarrow \mathbf{K}\phi \rightarrow \mathbf{K}\psi \tag{\textbf{K}}$$

$$\mathbf{K}\phi \rightarrow \phi \tag{\textbf{T}}$$

$$\mathbf{K}\phi \rightarrow \neg\mathbf{K}\neg\phi \tag{\textbf{D}}$$

$$\mathbf{K}\phi \rightarrow \mathbf{K}\mathbf{K}\phi \tag{\textbf{4}}$$

$$\neg\mathbf{K}\phi \rightarrow \mathbf{K}\neg\mathbf{K}\phi \tag{\textbf{5}}$$

13. Ou toutes tentatives de prouver p conduit à une antilogie $p \rightarrow \perp$.

14. Il faudra alors se demander dans quel contexte pourrait on vraiment conclure que p est faux ? J'imagine qu'une forme d'anti-réalisme relativement fort pourrait argumenter que si une proposition n'est absolument pas prouvable (e.g. par ce qu'elle a eu lieu dans le passé sans laisser de trace), alors elle doit être considérée comme fausse voire même qu'il n'y a pas lieu de se questionner sur sa vérité. Il ne me semble pas qu'il soit possible d'opérer explicitement une telle distinction ici et il serait intéressant d'identifier comment implémenter une telle position.

15. On suppose alors l'existence d'un "démon de Laplace", en conflit certain avec l'augmentation nécessaire de l'entropie, les évolutions hautement non-linéaires (chaotique) et/ou la mesure en mécanique quantique, mais là n'est pas le sujet de notre discussion (voir e.g. Hoefer (2023)).

pourront être attribués à $\mathbf{K}$. On pourra ainsi se demander comment traiter les problèmes de l'introspection et de l'omniscience dans ces modèles et également savoir si on souhaite intégrer des opérateurs modaux comme $\Box$ et $\Diamond$ malgré les difficultés évoqués en section 2.6. Ce programme ambitieux a été entrepris par Proietti (2012), dans la suite de Božić et Došen (1984). L'auteur souhaite apporter un cadre formel solide pour (re)discuter les problèmes épistémiques comme le paradoxe de Fitch dans un cadre intuitioniste. Pour cela, il défini une nouvelle relation d'accessibilité $R_{\mathbf{K}}$ sur les modèles, différente de $\leq$, cette dernière étant réellement interprétée comme une relation temporelle entre les états d'informations. On peut alors définir $\Box$ grâce à $\leq$, en utilisant une définition indépendante de $\Diamond$ [16] comme celle donnée en appendice A, évitant le problème discuté à la fin de la section 2.6. On définira ensuite $\mathbf{K}$, similairement à $\Box$, comme :

$$\mathcal{M}, w_i \vDash \mathbf{K}\phi \text{ ssi pour tout } w_j \in W \text{ tel que } R_{\mathbf{K}}(w_i, w_j) \in \mathcal{R} \text{ existe, alors } \mathcal{M}, w_j \vDash \phi.$$

Notant $w_1 \equiv w_2$ si les deux mondes w_1 et w_2 forcent les mêmes propositions atomiques et définissant $Pr(w)$ comme l'ensemble des prédécesseur à w par $\leq$, incluant w lui même, on peut alors définir $R_{\mathbf{K}}$ comme

> $R_{\mathbf{K}}(w_1, w_2)$ existe ssi il existe une fonction bijective f allant de $Pr(w_1)$ au sous ensemble descendant fermé [17] par rapport à $\leq$ de $Pr(w_2)$, tel que, pour tout $z \in Pr(w_1)$, $f(z_1) \equiv z$.

On a alors $\leq \subseteq R_{\mathbf{K}}$, la relation temporelle étant ainsi incluse dans la relation de connaissance : tous les futurs temporels accessibles sont aussi des états d'informations accessibles mais l'agent peut accéder à des états d'informations qui ne sont pas des états futurs. Cette définition largement abstraite permet d'interpréter $\mathbf{K}\phi$ comme ϕ *est vérifié dans toutes les expansions de l'état d'information d'un agent cohérentes avec son évolution*. De plus, cette définition permet au sujet de garder la mémoire de l'ordre dans lequel les informations ont été acquises. On peut également vérifier que $R_{\mathbf{K}}$ satisfait aux propriétés usuellement demandées aux relations modales. Sous cette définition, $\mathbf{K}$ satisfait à (**K**) et le sujet est donc omniscient c'est à dire que si il connait une proposition, il connait aussi toutes ses conséquences logiques. Nous accepterons cela comme une idéalisation acceptable, car il s'agit ici de modéliser en un premier temps l'information accessible à un sujet idéal. $\mathbf{K}$ satisfait également à (**T**), condition dite de factivité ou de véridicité, comme souhaité dans une définition de $\mathbf{K}$ (si je sais p, alors p est vraie). (**D**) est également satisfait : *si* ϕ *est connaissable alors* $\neg\phi$ *est inconnaissable*. Une telle proposition semble bien nécessaire pour éviter des contradictions. Finalement, $\mathbf{K}$ satisfait aussi à (**4**) et ainsi à l'introspection positive : *si* ϕ *est connaissable, la connaissance de* ϕ *est connaissable*, axiome qui peut être largement questionnable (e.g. dans le cadre des animaux et des êtres non linguistiques ou tout simplement toute formes de connaissances qui pourraient être "inconscientes" ou "intuitives"), mais on est facilement prêt à reconnaître (**4**) dans le cas d'un sujet idéal, hypothèse clairement formulée ici. Par contre, $\mathbf{K}$ ne satisfait pas à (**5**) d'introspection négative : ϕ *est inconnaissable implique qu'on peut savoir que* ϕ *n'est pas connaissable*. Cela rejoint directement notre discussion plus haut sur l'exemple du pile ou face et il faudrait examiner avec soin les conséquences du non respect de (**5**) dans un tel cas.

Dans ce nouveau cadre, (ARI) fournit bien une proposition interprétable permettant d'implémenter la position de l'anti-réalisme modéré comme suggéré par Vidi et Solomon (2001) et ne se réduisant pas à (ARR). Une telle conclusion n'est réalisée que pour certains modèles satisfaisant à des conditions particulières [18] (dans lesquelles (I), (II) et (III) sont aussi satisfaites). L'anti-réalisme modéré présente donc des restrictions sur les modèles possibles représentant

16. Ici, $\Diamond$ n'est pas défini.

17. X est un sous ensemble descendant fermé si pour tout $x \in X$ tel que $y \leq x$, alors $y \in X$

18. (ARI) n'est satisfait que pour tous les modèles satisfaisant $\forall w_1 \forall w_2 (w_1 \leq w_2 \rightarrow \exists z (w_2 \leq z \wedge \forall z' (R_{\mathbf{K}}(z, z') \rightarrow z' \equiv z))$ si $R_{\mathbf{K}}$ est définit comme suggéré plus haut.

l'évolution temporelle de l'information accessible à un sujet et il est intéressant de pouvoir identifier lesquelles.

Comme nous l'avons déjà mentionné, la présente construction des modèles intuitionnistes de Kripke cherche à modéliser un penseur idéal, omniscient et avec une parfaite mémoire. Il faudrait alors se demander comment assouplir la définition de $R_{\mathbf{K}}$ pour permettre par exemple l'oubli d'informations, ou la non inclusion des alternatives temporelles dans les alternatives épistémiques, sans rendre la théorie inextricablement complexe. Un développement conjoint du formalisme et de son interprétation permettrait notamment de clarifier le traitement des phénomènes contingents évoqués plus haut. Dans tous les cas, le cadre ainsi proposé est robuste et mériterai d'être discuté plus en avant afin d'établir précisément ces limites et dans quelle mesure cette approche est préférable à la logique classique.

Par soucis de concision et de clarté, nous n'aborderons pas ici les approches alternatives proposées face au paradoxe de Fitch et comment elles se comparent à la proposition de DeVidi et Solomon (1997) et de Proietti (2012). La littérature à ce sujet est foisonnante et pour un tour d'horizon de différentes approches, je renvoie le lecteur à e.g. Salerno (2008). Notons que au niveau des révisions logiques, la logique intuitionniste n'est pas le seul choix qui ait été considéré (voir e.g. les révisions paraconsistantes), mais il apparaît comme le plus justifié, justement par ce que l'approche intuitionniste de la logique est si proche de la position anti-réaliste. Il est certain que le paradoxe de Fitch n'a pas fini de faire parler de lui et a déjà été extrêmement fructueux pour toutes les approches formelles de l'épistémologie car il fournit un cadre riche pour questionner de près l'ensemble de nos outils logiques, leur significations et leurs conséquences épistémiques.

Références

Berkeley, George (1710). *Principes de la connaissance humaine.*

Božić, Milan et Kosta Došen (1984). « Models for Normal Intuitionistic Modal Logics ». In : *Studia Logica* 43.3, p. 217-245. DOI : 10.1007/bf02429840.

Bridges, Douglas, Erik Palmgren et Hajime Ishihara (2022). « Constructive Mathematics ». In : *The Stanford Encyclopedia of Philosophy*. Sous la dir. d'Edward N. Zalta et Uri Nodelman. Fall 2022. Metaphysics Research Lab, Stanford University.

Brogaard, Berit et Joe Salerno (2019). « Fitch's Paradox of Knowability ». In : *The Stanford Encyclopedia of Philosophy*. Sous la dir. d'Edward N. Zalta. Fall 2019. Metaphysics Research Lab, Stanford University.

Brouwer, L. E. J. (1952). *Historical background, principles, and methods of intuitionism.*

Creath, Richard (2022). « Logical Empiricism ». In : *The Stanford Encyclopedia of Philosophy*. Sous la dir. d'Edward N. Zalta et Uri Nodelman. Winter 2022. Metaphysics Research Lab, Stanford University.

DeVidi, David et Graham Solomon (1997). « $\Box$ in Intuitionistic Modal Logic ». In : *Australasian Journal of Philosophy* 75.2, p. 201-213. DOI : 10.1080/00048409712347801.

Došen, Kosta (1984). « The Paradox of Knowability ». In : *Intuitionistic Double Negation as a Necessity Operator* 35(49), p. 15-20.

Dummett, Michael (2009). « Fitch's Paradox of Knowability ». In : *New Essays on the Knowability Paradox*. Sous la dir. de Joe Salerno. Oxford University Press.

Edgington, Dorothy (1985). « The Paradox of Knowability ». In : *Mind* 94.376, p. 557-568. ISSN : 00264423, 14602113. URL : http://www.jstor.org/stable/2254726 (visité le 14/08/2023).

Fitch, Frederic B. (1963). « A logical analysis of some value concepts ». In : *The Journal of Symbolic Logic* 28.2, p. 135-142. DOI : 10.2307/2271594.

Garson, James (2023). « Modal Logic ». In : *The Stanford Encyclopedia of Philosophy*. Sous la dir. d'Edward N. Zalta et Uri Nodelman. Spring 2023. Metaphysics Research Lab, Stanford University.

Goldblatt, Robert (2003). « Mathematical Modal Logic : A View of its Evolution ». In : *Journal of Applied Logic* 1.5-6, p. 309-392. DOI : 10.1016/s1570-8683(03)00008-9.

Guyer, Paul et Rolf-Peter Horstmann (2023). « Idealism ». In : *The Stanford Encyclopedia of Philosophy*. Sous la dir. d'Edward N. Zalta et Uri Nodelman. Spring 2023. Metaphysics Research Lab, Stanford University.

Hoefer, Carl (2023). « Causal Determinism ». In : *The Stanford Encyclopedia of Philosophy*. Sous la dir. d'Edward N. Zalta et Uri Nodelman. Spring 2023. Metaphysics Research Lab, Stanford University.

Horwich, Paul (1999). « Truth ». In : *Meaning*. Sous la dir. de Paul Horwich. Oxford University Press, p. 261-272.

Khlentzos, Drew (2004). *Naturalistic Realism and the Antirealist Challenge*. National Geographic Books.

Kripke, Saul (1963). « Semantical Considerations on Modal Logic ». In : *Acta Philosophica Fennica* 16, p. 83-94.

Martino, Enrico et Gabriele Usberti (1994). « Temporal and Atemporal Truth in Intuitionistic Mathematics ». In : *Topoi* 13.2, p. 83-92. DOI : 10.1007/bf00763507.

Morgan, Charles G. (1973). « DRAWING DICHOTOMIES VIA FORMAL LANGUAGES ». In : *The Southern Journal of Philosophy* 11.3, p. 216-227. DOI : https://doi.org/10.1111/j.2041-6962.1973.tb01136.x. eprint : https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2041-6962.1973.tb01136.x. URL : https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2041-6962.1973.tb01136.x.

Moschovakis, Joan (2023). « Intuitionistic Logic ». In : *The Stanford Encyclopedia of Philosophy*. Sous la dir. d'Edward N. Zalta et Uri Nodelman. Summer 2023. Metaphysics Research Lab, Stanford University.

Murzi, Julien (2010). « Knowability and Bivalence : Intuitionistic Solutions to the Paradox of Knowability ». In : *Philosophical Studies* 149.2, p. 269-281. DOI : 10.1007/s11098-009-9349-y.

Percival, Philip (1990). « Fitch and Intuitionistic Knowability ». In : *Analysis* 50.3, p. 182-187. ISSN : 00032638, 14678284. URL : http://www.jstor.org/stable/3328541 (visité le 14/08/2023).

— (1991). « Knowability, Actuality, and the Metaphysics of Context-Dependence ». In : *Australasian Journal of Philosophy* 69.1, p. 82-97. DOI : 10.1080/00048409112344531.

Proietti, Carlo (2012). « Intuitionistic Epistemic Logic, Kripke Models and Fitch's Paradox ». In : *Journal of Philosophical Logic* 41.5, p. 877-900. ISSN : 00223611, 15730433. URL : http://www.jstor.org/stable/41653761 (visité le 12/08/2023).

Rebuschi, M. (2023). *Logique épistémique*.

Salerno, Joe (2008). *New Essays on the Knowability Paradox*. Sous la dir. de Joe Salerno. Oxford, England et New York, NY, USA : Oxford University Press.

Vidi, David De et Graham Solomon (2001). « Knowability and Intuitionistic Logic ». In : *Philosophia* 28.1-4, p. 319-334. DOI : 10.1007/bf02379783.

Williamson, Timothy (1982). « Intuitionism Disproved ? » In : *Analysis* 42.4, p. 203-7. DOI : 10.1093/analys/42.4.203.

— (1992). « On Intuitionistic Modal Epistemic Logic ». In : *Journal of Philosophical Logic* 21.1, p. 63-89. ISSN : 00223611, 15730433. URL : http://www.jstor.org/stable/30226465 (visité le 13/08/2023).

Wójcik, Arkadiusz (2020). « The Knowability Paradox and Unsuccessful Updates ». In : *Studies in Logic, Grammar and Rhetoric* 62.1, p. 53-71. DOI : doi:10.2478/slgr-2020-0013. URL : https://doi.org/10.2478/slgr-2020-0013.

A Sémantique des modèles de Kripke

La logique modale est généralement formulée en utilisant la sémantique des modèles de Kripke (Kripke 1963). Un modèle $\mathcal{M} = (W, \mathcal{R}, \mathcal{V})$ est un triplet composé

— D'un ensemble non vide de mondes possibles $w_i \in W$.

— D'une relation binaire d'accessibilité entre les mondes $R(w_i, w_j) \in \mathcal{R} \subseteq W \times W$.

— D'une valuation $\mathcal{V}(p)$ associant une valeur de vérité aux propositions p associées à chaque monde (Kripke 1963).

La validité d'une proposition ϕ pourra être estimée relativement à un monde $w_i \in W$ en notant $\mathcal{M}, w_i \vDash \phi$.

On définira alors les opérateurs $\Box$ et $\Diamond$ comme :

— $\mathcal{M}, w_i \vDash \Box\phi$ ssi pour tout $w_j \in W$ tel que $R(w_i, w_j) \in \mathcal{R}$ existe, alors $\mathcal{M}, w_j \vDash \phi$.

— $\mathcal{M}, w_i \vDash \Diamond\phi$ ssi il existe au moins un $w_j \in W$ tel que $R(w_i, w_j) \in \mathcal{R}$ existe et $\mathcal{M}, w_j \vDash \phi$.

On peut alternativement obtenir/définir $\Box$ à partir de $\Diamond$ et vice-versa à partir de

$$\Diamond p \Leftrightarrow \neg\Box\neg p \tag{POS}$$

$$\Box p \Leftrightarrow \neg\Diamond\neg p \tag{NES}$$

La définition que nous avons donnée d'un modèle peut être immédiatement généralisée pour inclure plusieurs relations d'accessibilités entre les mondes de sorte que $\mathcal{M} = (W, \mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3, ..., \mathcal{V})$. Les opérateurs de connaissance et de croyance – $\mathbf{K}\phi$ et $\mathbf{B}\phi$ – sont définis de manières analogues à $\Box\phi$, avec éventuellement des relations d'accessibilités $\mathcal{R}_\mathbf{K}$ et $\mathcal{R}_\mathbf{B}$ qui leur sont propres et des axiomes contraignant leur structure.

B Preuve du paradoxe de Fitch

Je me permet ici de reproduire la preuve donnée dans Wójcik (2020) et Rebuschi (2023), afin de pouvoir discuter son contenu.

Supposons un système de logique épistémique cohérent. Nous allons assumer deux propriétés que doit respecter l'opérateur **K**. Nous allons d'abord supposer que **K** est réflexif (axiome (**T**)) :

$$\mathbf{K}p \to p. \tag{\textbf{T}}$$

Autrement dit : *si je sais p alors p est vraie*. On peut considérer un tel axiome comme définitoire pour la connaissance.

Dans le système minimal de logique modale, on peut de plus montrer que **K** se distribue sur la conjonction. Soient p et q deux propositions, alors

$$\mathbf{K}(p \wedge q) \to (\mathbf{K}p \wedge \mathbf{K}q), \tag{3}$$

à savoir *si je sais que p et q alors je sais p et je sais q*.

Pour mener la démonstration à terme, on admettra de plus que (**T**) et (3) sont des propositions nécessaires, à savoir que $\Box(\mathbf{K}p \to p)$ (P1) et $\Box(\mathbf{K}(p \wedge q) \to (\mathbf{K}p \wedge \mathbf{K}q))$ (P2) c'est à dire que (**T**) et (3) sont toujours vérifiés. Comme (**T**) et (3) sont respectivement une propriété définitoire et une propriété élémentaire de **K**, on acceptera sans peine qu'elles puissent être nécessaires et que P1 et P2 soient bien valides.

A cela, on ajoutera la proposition (ARM) dont nous voulons connaître les conséquences, ainsi que quatre relations R_i qui peuvent facilement être dérivées en logique modale propositionnelle classique. En résumé, soient p, q, ψ et ϕ quatre propositions, on acceptera que

$$p \to \Diamond \mathbf{K}p \tag{ARM}$$
$$\Box(\mathbf{K}p \to p) \tag{P1}$$
$$\Box(\mathbf{K}(p \wedge q) \to (\mathbf{K}p \wedge \mathbf{K}q)) \tag{P2}$$
$$\Diamond\phi, \Box(\phi \to \psi) \Rightarrow \Diamond\psi \tag{R1}$$
$$\Diamond(\phi \wedge \psi), \Box(\psi \to \psi') \Rightarrow \Diamond(\phi \wedge \psi') \tag{R2}$$
$$\phi \to \psi, \neg\psi \Rightarrow \neg\phi \tag{R3}$$
$$\neg(\phi \wedge \neg\psi) \Rightarrow \phi \to \psi \tag{R4}$$

Utilisant uniquement ces relations, on peut alors dériver (ARR) en utilisant la sémantique de la logique propositionnelle classique.

Démonstration.

1.	$\Diamond \mathbf{K}(p \wedge \neg \mathbf{K}p)$	Hypothese
2.	$\Box(\mathbf{K}(p \wedge q) \to (\mathbf{K}p \wedge \mathbf{K}q))$	P2
3.	$\Box(\mathbf{K}(p \wedge \neg\mathbf{K}p) \to (\mathbf{K}p \wedge \mathbf{K}\neg\mathbf{K}p))$	$2, \mathrm{Sub}[\neg\mathbf{K}p/q]$
4.	$\Diamond(\mathbf{K}p \wedge \mathbf{K}\neg\mathbf{K}p)$	$3, 1, \mathrm{R1}$
5.	$\Box(\mathbf{K}p \to p)$	P1
6.	$\Box(K\neg\mathbf{K}p \to \neg\mathbf{K}p)$	$5, \mathrm{Sub}[\neg\mathbf{K}p/p]$
7.	$\Diamond(\mathbf{K}p \wedge \neg\mathbf{K}p)$	$4, 6, \mathrm{R2}$
8.	$\Diamond\bot$	$7, \mathrm{Def}[\bot]$
9.	$\bot$	$8, \mathrm{Def}[\bot]$

On a ainsi prouvé

$$\neg\Diamond \mathbf{K}(p \wedge \neg\mathbf{K}p) \tag{*}$$

Alors

10.	$\neg\Diamond \mathbf{K}(p \wedge \neg\mathbf{K}p)$	$*$
11.	$p \to \Diamond\mathbf{K}p$	ARM
12.	$(p \wedge \neg\mathbf{K}p) \to \Diamond\mathbf{K}(p \wedge \neg\mathbf{K}p)$	$11, \mathrm{Sub}[(p \wedge \neg\mathbf{K}p)/p]$
13.	$\neg(p \wedge \neg\mathbf{K}p)$	$10, 12, \mathrm{R3}$
14.	$p \to \mathbf{K}p$	$13, \mathrm{R4}$

$\Box$