

UNIVERSITÉ PARIS DESCARTES - ENS - EHESS

MASTER DE SCIENCES COGNITIVES

L'effet de l'unawareness dans la prise de décision

Sonia MARIN

Sous la direction de Mikael COZIC (IHPST)

et de Brian HILL (GREGHEC)

Septembre 2014

Déclaration d'originalité et de contribution

Ce mémoire rassemble une série d'idées sur la manière d'intégrer la notion d'unawareness aux modèles de théorie de la décision. Celles-ci sont issues d'une mise en perspective de deux pans de la littérature, logique épistémique et théorie de la décision, ainsi que de réflexions et de tentatives d'analyses formelles personnelles. J'ai sans doute manqué de l'expérience nécessaire pour regrouper ces idées en un modèle complet et satisfaisant de l'effet de l'unawareness sur la prise de décision, mais ce mémoire correspond au travail préparatoire nécessaire à l'élaboration d'un tel modèle. En dehors des références directes à la littérature, son contenu est entièrement personnel et original.

Je remercie Mikael Cozic et Brian Hill de m'avoir fait découvrir les questionnements liés à l'unawareness, d'avoir participé à la formulation de la problématique à l'origine de ce travail et d'avoir aiguillé mes recherches bibliographiques et mes essais de formalismes. Je remercie Aurélien pour son oreille attentive et ses commentaires constructifs, Gabrielle pour son soutien sans faille, Antoine pour ses conseils toujours pertinents et Jean pour l'organisation de la journée "Master" de l'équipe DRI.

Sommaire

| | |
|--|-----------|
| Introduction | 4 |
| 1 Pourquoi introduire l'unawareness en théorie de la décision? | 5 |
| 1.1 Intuitions de psychologie naïve | 5 |
| 1.2 Intuitions issues de l'approche expérimentale | 6 |
| 1.3 Intuitions issues de l'approche théorique | 8 |
| 2 De la logique de l'unawareness à la théorie de la décision | 13 |
| 2.1 Modélisation de l'unawareness en logique épistémique | 13 |
| 2.2 Utilisation de la logique de l'unawareness en théorie de la décision | 14 |
| 2.3 Un essai de formalisation | 15 |
| 3 L'analyse de l'imprévu en théorie de la décision | 19 |
| 3.1 La prise en compte des événements imprévus | 19 |
| 3.2 Dynamique de l'awareness | 21 |
| 3.3 Unawareness consciente | 24 |
| Conclusion | 27 |
| References | 29 |

“What you need is to be aware of being aware. Be aware deliberately and consciously, broaden and deepen the field of awareness. You are always conscious of the mind, but you are not aware of yourself as being conscious.”

Sri Maharaj Nisargadatta

Introduction

La théorie de la décision cherche à modéliser le comportement d'un agent dans une situation de choix. Dans la théorie classique de Savage [14] on considère qu'il ne connaît pas l'état exact du monde mais qu'il est capable de lister les caractéristiques le déterminent et d'attribuer une probabilité à chaque état possible défini par la valuation de ces caractéristiques. Pourtant, il est de nombreux cas dans la vie courante où les personnes ignorent involontairement certaines des données utiles au problème qu'elles sont en train de traiter. D'ailleurs, de nombreux théoriciens de la décision estiment qu'il n'est pas possible de comprendre correctement les théories économiques sans considérer le rôle des *contingences imprévues*, certaines possibilités auxquelles un agent "oublierait" de penser au moment où il prend sa décision. On constate d'ailleurs que les contrats, les documents qui définissent la structure d'une entreprise ou les lois constitutionnelles laissent certains points un peu flous de manière à ce qu'ils puissent prendre en compte un grand nombre de situations et notamment celles que l'on n'aurait pas anticipées.

Parallèlement s'est développée en logique l'étude de la notion d'*unawareness*, une forme de conscience partielle qui permet de s'abstraire de certaines contraintes de la logique épistémique, notamment l'*introspection négative*; en effet lorsqu'un agent ne sait pas quelque chose, il ne sais pas forcément qu'il ne le sait pas, il peut simplement ne pas être en mesure de le concevoir ou ne pas y penser car son point de vue n'est pas assez global. Il semble que cette notion issue de la logique pourrait être la formulation adéquate pour représenter les cas où, en théorie de la décision, tous les éléments pertinents ne sont pas perçus par le sujet ce qui le rend incapable de donner une description objective de la situation de choix. L'idée serait alors de formaliser celle-ci de la manière dont le sujet se la représente subjectivement sur la base de ce qu'il perçoit effectivement et c'est exactement la manière dont l'*unawareness* est traitée en logique, comme une restriction du langage aux informations dont dispose l'agent.

Certains théoriciens de la décision s'intéressent d'ailleurs à l'effet de l'*unawareness* d'un agent sur sa manière de prendre des décisions, mais cette notion de perception partielle n'a pas encore de modèle canonique dans le cadre de la théorie de la décision. On peut distinguer deux approches dans la manière de traiter cette question. La première serait de prendre comme point de départ les modèles issus de la logique épistémique pour reconstruire un modèle de théorie de la décision. L'autre consisterait à considérer directement les modèles de choix et à se demander où l'*unawareness* pourrait intervenir.

Dans une première partie nous essayerons de justifier la nécessité d'introduire la notion d'*unawareness* en théorie de la décision par des intuitions de psychologie naïve d'abord puis par des considérations plus techniques à la fois empiriques et théoriques. Dans une deuxième partie nous exposerons une sélection d'études logiques de la notion d'*awareness* et la manière dont celles-ci peuvent être utiles pour construire une théorie de la décision qui tienne compte de l'*awareness* des agents. Enfin dans une troisième partie, nous reviendrons sur l'étude plus directe de la notion d'*unawareness* dans les modèles de théorie de la décision.

1 Pourquoi introduire l'unawareness en théorie de la décision?

1.1 Intuitions de psychologie naïve

Dans un premier temps, nous allons essayer de donner une justification à notre sujet en se basant sur de simples intuitions de psychologie naïve. On peut évoquer trois types de situations qui mettent en évidence l'effet de l'unawareness dans la prise de décision: des situations d'unawareness *spécifique* où un sujet est susceptible de prendre une décision sous-optimale parce que son niveau d'awareness est insuffisant, des situations d'unawareness *consciente* où un sujet a conscience de cette insuffisance et en tient compte dans sa décision et des situations d'unawareness *dynamique* où le niveau d'awareness du sujet évolue au cours du temps ce qui peut faire évoluer sa décision.

Lorsque l'on doit prendre une décision dans une situation complexe, il est fréquent que l'on ne possède pas toutes les informations pertinentes pour analyser le problème auquel on doit faire face. On parlera alors d'unawareness *spécifique*. Par exemple, lorsque l'on doit choisir entre deux offres d'emploi, on n'est pas forcément à même d'envisager tous les événements susceptibles de se produire et toutes les conséquences auxquelles le choix de l'une ou l'autre des offres conduirait dans le cas où un tel événement se réaliserait. Il est possible qu'un observateur extérieur bien documenté sur la question soit capable d'analyser la situation dans son ensemble et notamment, qu'au regard des informations globales dont il dispose, l'option choisie lui semble sous-optimale. Cette déviation par rapport au choix optimal ne pourra pas être expliqué par les différents modèles de théorie de la décision dont dispose l'observateur. En effet, l'étude traditionnelle de la prise de décision en situation d'incertitude part implicitement du principe que les agents ont pleinement conscience de toutes les données pertinentes d'un problème et que la question se résume ensuite à prendre la meilleure décision en fonction de ces données et de certaines caractéristiques propres à l'agent (son évaluation du risque, de l'ambiguïté, de l'utilité des conséquences). On aimerait être capable de modéliser des agents qui n'ont pas nécessairement pleinement conscience de toutes les données pertinentes d'un problème. Plus précisément, on aimerait être capable de distinguer au sein du même modèle la description objective du problème, celle qui pourrait être faite par un observateur extérieur, et la représentation subjective que se fait l'agent à partir des données dont il dispose.

Dans certaines situations, nous avons conscience que notre perception du problème de décision est potentiellement erronée, que nous ne sommes pas capables d'envisager toutes les caractéristiques de la situation pouvant influencer le résultat de notre choix, ni toutes les conséquences que pourrait avoir ce choix. D'ailleurs l'utilisation courante de mémo ou de check-list montre bien que les agents sont habitués à gérer leur manque de vision globale d'une situation de décision. Dans ce cas, il est possible que cette conscience de notre propre unawareness nous amène à être plus prudent (ou plus imprudent) lors de la prise de décision. Le comportement qui en résulte ne sera alors pas expliqué non plus par les modèles classiques de théorie de la décision. Ainsi, on peut vouloir améliorer encore les modèles de théorie de la décision pour que non seulement il rendent compte des déviations par rapport au choix optimal dues à l'unawareness spécifique d'un agent mais plus généralement à son unawareness consciente. Dès lors, l'unawareness spécifique pourrait possiblement être considérée comme une forme faible d'unawareness consciente où l'agent est indifférent à sa propre unawareness, c'est-à-dire où cette conscience ne le rend ni plus ni moins prudent par rapport aux choix qu'il ne comprend pas parfaitement.

Enfin, il est également courant que notre perception d'un problème de décision évolue au cours du temps au fur à mesure de l'acquisition de nouvelles informations. On est relativement habitués, à l'échelle de l'histoire ou de notre vie individuelle, à ce que ce produisent des événements que l'on n'avait pas anticipé, comme la fermeture annuelle d'un lieu où l'on avait décidé de se rendre, ou que certains choix induisent des conséquences que l'on aurait été incapable de concevoir au préalable. Ces imprévus nous amènent bien souvent à modifier notre représentation du problème de décision auquel on était confronté et à réviser ce que l'on considérait comme le choix de l'option optimale. C'est sans doute cette forme dynamique d'unawareness qui s'ancre le plus fortement dans nos intuitions de psychologie naïve. S'il n'est pas forcément clair que la prise en compte de l'unawareness soit absolument nécessaire dans un modèle de choix statique, il est relativement évident que si l'on souhaite modéliser l'évolution de l'heuristique de décision d'un agent, il faille analyser son comportement vis à vis d'un gain d'information. De plus, l'évolution dynamique de l'unawareness peut également avoir une influence sur d'autres paramètres du choix de l'agent. En effet, lorsque de nouvelles informations parviennent à l'agent, celui-ci est bien forcé de se rendre compte qu'il n'avait pas anticipé ces nouvelles informations et donc de devenir conscient de sa propre unawareness. Cela peut jouer sur son niveau de confiance dans les informations qu'il possède déjà, sur son niveau d'aversion au risque, à l'ambiguïté; bref, tous les niveaux de la perception du problème de décision semblent pouvoir être influencés par cette évolution de l'awareness de l'agent. C'est pourquoi il semble nécessaire d'être capable d'intégrer cette notion d'unawareness aux modèles existants de théorie de la décision afin d'analyser ses interactions avec les différents paramètres qui jouent un rôle dans le choix.

1.2 Intuitions issues de l'approche expérimentale

Nous avons essayé de donner quelques intuitions de psychologie naïve afin de justifier notre intérêt pour la question de l'unawareness dans le cadre de la théorie de la décision. Ce genre d'intuitions simples ont guidé Savage dans l'élaboration de sa théorie et sont tout à fait pertinentes pour avoir une idée globale de la manière optimale de prendre une décision dans une situation donnée. Cependant, les modèles de la théorie de la décision ne se sont pas uniquement développés en tant que guides normatifs, mais également comme modèles descriptifs afin d'expliquer les choix réels des individus. Dans ce second cadre, la question qui se pose est de prévoir le résultat du choix d'un agent dans le cas où il agit conformément à l'un ou l'autre des modèles. Ceci présuppose que les modèles sont suffisamment efficaces pour décrire le comportement des agents. Il est donc indispensables de les tester expérimentalement et/ou empiriquement. L'étude de la question de l'aversion des agents rationnels à l'ambiguïté a été par exemple directement stimulée par le paradoxe expérimental mis en évidence par Ellsberg [8]. D'autres déviations par rapport aux théories classiques peuvent être mises en évidence expérimentalement: effets de cadre, intransitivité des préférences, instabilité du comportement au cours du temps... Ces données expérimentales sont autant de challenges pour les théoriciens de la décision pour adapter leurs modèles ou en développer de plus généraux de manière à mieux décrire ce qui est observé expérimentalement.

Dans le cas de l'unawareness, s'il est clair que cette notion est fortement susceptible de jouer un rôle lors de la prise de décision, on ne bénéficie pas d'un paradoxe similaire au paradoxe d'Ellsberg pour l'aversion à l'incertitude, c'est-à-dire un paradoxe qui fasse ressortir un comportement dû exclusivement au fait que l'agent est unaware d'un événement. Cependant, certaines expériences d'économie

comportementales mettent déjà en évidence que l'unawareness peut avoir une influence sur certains paramètres de la décision. C'est le cas de l'étude menée par Mengel, Tsakas et Vostroknutov [7]. Ces auteurs montrent que le niveau d'information dont l'agent dispose lors de prise de décisions ultérieures peu influencer ces décisions futures et la manière dont il percevra le risque par la suite, potentiellement dans des domaines apparemment sans correspondance.

Ils étudient comment évolue l'aversion au risque d'agents ayant été soumis à des niveaux différents de connaissance au préalable. Pour cela, ils réalisent une expérience en deux parties qui consistent chacune en une série de choix entre une loterie fixe et différentes sommes d'argent (sur le modèle de la méthode BDM).

| probabilité | résultat |
|-------------|----------|
| 0,1% | - 20 |
| 5% | - 1 |
| 5% | Twix |
| 20% | 6 |
| 25% | 8 |
| 37,9% | 10 |
| 7% | 14 |

Figure 1: Loterie présentée lors de la première phase de l'expérience

Dans la première phase de l'expérience, trois groupes de sujet sont soumis à trois niveaux d'information différents alors que dans la seconde tous les groupes bénéficient du niveau d'information maximal. Les trois niveaux d'information correspondent aux conditions expérimentales suivantes:

- la condition *Risk* où les participants ont connaissance de toutes les issues de la loterie ainsi que des probabilités qui y sont attribuées
- la condition *Ambiguity* où seuls les issues de la loterie ne sont connues de l'agent
- la condition *Unawareness* où l'on ne communique au sujet que certaines des issues de la loterie et pas de probabilités. S'il choisit de jouer la loterie, il bénéficie néanmoins d'informations supplémentaires au fur et à mesure

Il est systématiquement expliqué au sujet de l'expérience à quel type de tâche il va être confronté donc en particulier dans la condition *Unawareness*, le sujet est conscient de cette unawareness et dans les deux autres conditions il sait qu'il n'y a pas d'autres issues que celles qui lui sont directement présentées.

Les auteurs s'intéressent au comportement des sujets dans les étapes 17 à 32 qui sont identiques pour tous. En effet, d'après eux, leur design ne permettrait pas d'élucider les attitudes vis-à-vis de l'ambiguïté puisque l'on ne connaît pas la probabilité que le sujet attribue a priori aux différentes issues de la loteries.

Les auteurs parviennent à montrer que les sujets ayant bénéficié d'une information imparfaite concernant l'espace d'états font preuve pas la suite d'une aversion au risque plus importante dans les situations de décision que les sujets ayant bénéficié d'une connaissance complète de l'espace d'états. Plus précisément, ils observent que moins les sujets ont d'information dans la première phase de

l'expérience et moins ils vont vouloir choisir des loteries à forte variance dans la seconde phase. Ainsi, les attitudes par rapport au risque des différents groupes sont d'autant plus marquées par l'aversion au risque (β) que le sujet a été exposé à un faible niveau d'information.

$$\beta_{Unawareness} > \beta_{Ambiguity} > \beta_{Risk}$$

D'autre part, Mengel et al. tentent également de développer l'analyse de leur expérience de manière à étudier directement les différences de comportement entre les différentes conditions dans la première phase de l'expérience. Ils observent que dans la condition *Unawareness*, les sujets choisissent plutôt la somme obtenue de manière certaine que la loterie et leurs temps de réponse sont aussi généralement plus rapides, mais ils ne parviennent pas à distinguer les deux autres conditions, *Ambiguity* et *Risk* sur ces points. Il ne nous semble pas évident que leur design expérimental soit suffisamment sophistiqué pour permettre une analyse fine de la différence de comportement dans la première phase de l'expérience. Cependant, comme les auteurs le suggèrent eux-mêmes, leur méthode pourrait très certainement être développée dans cette direction.

Ainsi, même s'il n'existe pas encore d'équivalent du paradoxe d'Ellsberg pour la question de l'unawareness, on possède déjà des données expérimentales qui montrent que les sujets sont sensibles à cette forme d'incertitude portant sur les données du problème de décision. Cela permet de justifier notre intérêt pour intégrer la notion d'unawareness aux modèles existants de théorie de la décision.

1.3 Intuitions issues de l'approche théorique

Même si les données expérimentales ne nous permettent pas encore de déterminer exactement l'effet de l'unawareness de l'agent sur sa prise de décision, elles montrent qu'il est nécessaire de s'intéresser à cette question. Nous pensons qu'une étude approfondie des modèles théoriques permet de guider l'expérimentation et que ces deux méthodes doivent être exploitées en parallèle afin de déterminer avec précision le rôle de l'awareness dans la prise de décision. Notre approche sera donc d'explorer les modèles de théorie de la décision, les plus basiques dans un premier temps, et de se demander de quelle manière on peut y intégrer une notion d'unawareness.

En théorie de la décision, on considère deux types de situation de décision, la décision dans le risque et la décision dans l'incertitude. Dans le risque, le décideur doit choisir entre des loteries, c'est-à-dire des actions dont les conséquences ne sont pas certaines mais dont la probabilité de chacune d'entre elles est connue. La théorie de Von Neumann et Morgenstern [11] permet de représenter le comportement de choix d'un agent dans une telle configuration par une maximisation de l'utilité espérée. Dans l'incertitude en revanche, le décideur doit choisir entre des actions dont l'issue dépend de la réalisation d'événements dont les probabilités ne sont pas données. La théorie de Savage [14] permet de modéliser le comportement des agents comme s'ils attribuaient en fait des probabilités subjectives aux différents événements.

Nous allons nous poser la question de la modélisation de l'unawareness de l'agent dans ces deux cadres de base mais d'abord quelle notion d'unawareness voulons nous représenter exactement? Nous voulons considérer un agent qui est partiellement aware des données du problème de décision, c'est-à-dire qu'il existe un décalage entre la représentation objective de la situation et la représentation que

s'en forme l'agent. Malgré cela, celui-ci met en place une stratégie de décision qui lui semble optimale et que nous voulons identifier. Du côté du décideur, deux attitudes sont envisageables:

- *agent 1* : soit il agit comme s'il possédait toutes les informations nécessaires à la mise en place de cette stratégie; c'est un cas d'unawareness spécifique
- *agent 2* : soit il est conscient de sa propre unawareness, il estime qu'il a une compréhension insuffisante de la situation et il adopte une stratégie potentiellement plus prudente, en laquelle il a peut-être moins confiance mais cherche malgré tout à optimiser son utilité dans la situation telle qu'il peut l'appréhender.

Que se passe-t-il ensuite si l'agent obtient de nouvelles informations sur la situation de choix? Dans les deux cas, il doit certainement modifier sa stratégie. Il est susceptible de vouloir également modifier sa position de principe c'est-à-dire,

- l'*agent 1* peut se rendre compte qu'il est partiellement unaware et vouloir prendre en compte cela dans ses décisions futures
- l'*agent 2* peut considérer que cette information suffit à lui faire comprendre le problème de choix dans son intégralité ou au contraire cette nouvelle découverte peut accroître sa sensation d'unawareness

Dans un premier temps, nous n'allons considérer que des agents de type *agent 1* et nous allons simplement supposer que les agents conservent leur position de principe. Nous allons nous intéresser aux deux formalismes précités et à la manière dont ils pourraient être modifiés afin que les agents soient capables d'intégrer des informations concernant les données du problème de choix au cours du temps.

Dans le formalisme de Von Neumann et Morgenstern [11]:

Soit Z un ensemble fini de conséquences. Soit $\Delta(Z)$ l'ensemble des mesures de probabilités sur Z , c'est-à-dire ici les fonctions $p: Z \rightarrow [0, 1]$ telles que $\sum_{z \in Z} p(z) = 1$. On muni cet ensemble d'une relation binaire \succ .

Soient $p, q \in \Delta(Z)$ et $a \in [0, 1]$, la *combinaison convexe* $p_a q$ est définie par:

$$p_a q(z) = ap(z) + (1 - a)q(z) \text{ pour tout } z \in Z$$

Il suffit des trois axiomes suivants pour énoncer le théorème de Von Neumann et Morgenstern.

Axiome 1.1. \succ est une relation transitive et complète.

Axiome 1.2. Pour tous $p, q, r \in \Delta(Z)$, si $p \succ q \succ r$ alors il existe $a, b \in (0, 1)$ tels que $p_a r \succ q \succ p_b r$.

Axiome 1.3. Pour tous $p, q, r \in \Delta(Z)$ et $a \in (0, 1]$, si $p \succ q$ alors $p_a r \succ q_a r$.

Le théorème établit qu'un agent dont la relation de préférence satisfait ces axiomes est un maximisateur d'utilité espérée.

Théorème 1. *La relation binaire \succ sur $\Delta(Z)$ satisfait les axiomes 1.1, 1.2 et 1.3 si et seulement s'il existe une fonction $u: Z \rightarrow \mathbb{R}$ telle que:*

$$p \succ q \text{ ssi } \sum_{z \in Z} u(z) \cdot p(z) > \sum_{z \in Z} u(z) \cdot q(z)$$

et u est unique à transformation affine positive près.

Supposons maintenant que l'agent gagne en awareness. La seule possibilité ici est que de nouveaux éléments doivent être ajoutés à l'ensemble de conséquences Z . Considérons donc un ensemble $Z^* \supseteq Z$ et $\Delta(Z^*)$ l'ensemble des mesures de probabilités sur Z^* muni de la relation binaire \succ^* . On a directement que $\Delta(Z) \subseteq \Delta(Z^*)$. On peut donc considérer toute distribution de probabilité sur Z comme une distribution de probabilité sur Z^* .

Soit $p \in \Delta(Z)$, on définit son *extension* à Z^* par:

$$\tilde{p}(z^*) = \begin{cases} p(z^*) & \text{si } z^* \in Z \\ 0 & \text{sinon} \end{cases}$$

Intuitivement, il semble acceptable, au moins normativement, de supposer que les préférences de l'agent quant à l'ensemble de loteries $\Delta(Z)$ ne sont pas modifiées lors du gain d'awareness. Cela peut se traduire par l'axiome suivant:

Axiome 1.4. Soient $p, q \in \Delta(Z)$, $p \succ q$ si et seulement si $\tilde{p} \succ^* \tilde{q}$.

Supposons que les relations de préférence \succ sur $\Delta(Z)$ et \succ^* sur $\Delta(Z^*)$ satisfont les axiomes 1.1, 1.2 et 1.3 et qu'elles satisfont conjointement l'axiome 1.4..

Soient $p, q \in \Delta(Z)$.

$$\begin{aligned} p \succ q & \leftrightarrow \tilde{p} \succ^* \tilde{q} \text{ d'après l'axiome 1.4} \\ & \leftrightarrow \sum_{z \in Z^*} u^*(z) \cdot \tilde{p}(z) > \sum_{z \in Z^*} u^*(z) \tilde{q}(z) \text{ d'après le théorème 1} \\ & \leftrightarrow \sum_{z \in Z} u^*(z) p(z) > \sum_{z \in Z} u^*(z) q(z) \text{ d'après la définition d'extension} \end{aligned}$$

Il semble que toute fonction d'utilité $u^*: Z^* \rightarrow \mathbb{R}$ pour Z^* puisse ainsi donner une fonction d'utilité $u|_Z: Z \rightarrow \mathbb{R}$ pour Z trivialement par réduction à l'ensemble de départ.

Dans le formalisme de Savage [14]

Le design du modèle de Savage est un peu plus sophistiqué. Les briques de base sont:

- S l'espace des états
- Z l'espace des conséquences, qui décrivent le résultat d'un choix de l'agent
- des fonctions de S dans Z , les *actes*, qui sont les choix possibles de l'agent et qui forment l'ensemble F
- une relation de préférence \succ sur F

C'est dans ce cadre que Savage énonce le théorème de représentation selon lequel les choix des agents peuvent être réduits à une maximisation de l'utilité espérée. Plus précisément,

Théorème 2. *La relation de préférence vérifie les axiomes de Savage si et seulement s'il existe une fonction d'utilité $u: Z \rightarrow \mathbb{R}$ et une mesure de probabilité p sur Z telles que:*

$$f \succ g \text{ ssi } \mathbb{E}_p u(f) > \mathbb{E}_p u(g) \text{ pour tous actes } f, g \in F$$

Dans ce cas, il semble qu'un gain d'awareness puisse correspondre à une extension de l'ensemble des conséquences $Z \subseteq Z^*$, cas trivial comme précédemment puisque le nouvel ensemble d'actes contient l'ancien $Z^S \subseteq Z^{*S}$, ou bien à une extension de l'espace d'états $S \subseteq S^*$. Dans ce deuxième cas, ce n'est plus aussi simple puisque $Z^S \not\subseteq Z^{*S}$.

Considérons l'exemple où l'espace d'états initial contient deux éléments $S = \{s_1, s_2\}$ tout comme l'ensemble de conséquences $Z = \{x, y\}$. Dès lors l'ensemble des actes contient quatre éléments que l'on notera $F = \{\alpha, \beta, \gamma, \delta\}$. La situation se représente donc par la matrice suivante:

| | s_1 | s_2 |
|----------|-------|-------|
| α | x | x |
| β | x | y |
| γ | y | x |
| δ | y | y |

Figure 2: Matrice de choix avec ensemble endogène d'actes-fonctions et avant gain d'awareness

Maintenant, supposons que l'agent devienne aware d'un nouvel état s_3 . Cela conduit naturellement à l'expansion de l'ensemble d'actes et à la matrice suivante:

| | s_1 | s_2 | s_3 |
|------------|-------|-------|-------|
| α_1 | x | x | x |
| α_2 | x | x | y |
| β_1 | x | y | x |
| β_2 | x | y | y |
| γ_1 | y | x | x |
| γ_2 | y | x | y |
| δ_1 | y | y | x |
| δ_2 | y | y | y |

Figure 3: Matrice de choix avec ensemble endogène d'actes-fonctions et après gain d'awareness

Cependant, cela semble un peu contredire nos intuitions. Alors même que l'on cherche à représenter un agent dont les capacités sont limitées dans le sens où il n'est pas forcément aware de toutes les informations pertinentes lors de sa prise de décision, l'ajout de nouveaux états fait exploser la complexité de l'ensemble d'actes que l'agent doit considérer. On ne peut donc pas rester dans le cadre strict de Savage, il nous faut considérer un ensemble exogène de fonctions qui correspondent à la représentation que l'agent se fait des actions parmi lesquelles il va choisir.

Supposons que dans notre exemple cet ensemble soit $F = \{\alpha, \beta\}$. La situation de choix est donc perçue par l'agent conformément à la matrice suivante:

Que se passe-t-il lorsque l'agent devient aware du nouvel état s_3 ?

La première question qui se pose est celle de l'issue des actions considérées dans le cas où l'état s_3 se réalise. En effet, en sortant du cadre de Savage, on perd l'aspect systématique de la construction des

| | s_1 | s_2 |
|----------|-------|-------|
| α | x | x |
| β | x | y |

Figure 4: Matrice de choix avec ensemble exogène d’actes-fonctions et avant gain d’awareness

| | s_1 | s_2 | s_3 |
|----------|-------|-------|-------|
| α | x | x | ?? |
| β | x | y | ?? |

Figure 5: Matrice de choix avec ensemble exogène d’actes-fonctions et après gain d’awareness

actes. comme fonctions de l’espace d’états dans l’ensemble des conséquences. Ghirardato [21] suggère de redéfinir les actes comme des correspondances plutôt que des fonctions, c’est-à-dire des applications qui à chaque état du monde associe un ensemble de conséquences (voir Fig.6). Cela demande ensuite de considérer que l’agent évalue la vraisemblance de chaque issue à l’aide d’une mesure de probabilité (distincte de celle qu’il attribue déjà aux différents états).

| | s_1 | s_2 | s_3 |
|----------|-------|-------|------------|
| α | x | x | $\{x, y\}$ |
| β | x | y | $\{x, y\}$ |

Figure 6: Matrice de choix avec ensemble exogène d’actes-correspondances et après gain d’awareness

La deuxième question concerne justement la mesure de probabilité qui représente, chez Savage, l’incertitude sur l’espace d’états. Comment le sujet doit-il la modifier lorsque le support de probabilité sui était $S = \{s_1, s_2\}$ devient $S^* = \{s_1, s_2, s_3\}$? La règle classique d’*update bayésien* n’a pas cours ici puisqu’elle s’applique dans les cas où une information permet la réduction de l’espace d’états à considérer et non pas lors de son expansion (par gain d’awareness). A priori, il ne semble pas y avoir de règles canoniques qui puissent s’appliquer dans le cadre que nous avons détaillé ici.

Ces quelques investigations techniques des modèles standard de la décision permettent donc de mettre en évidence les premiers problèmes qui se posent lorsque l’on cherche à introduire la notion d’unawareness à ces modèles. Nous continuerons de nous appuyer sur ces différents niveaux d’intuition pour explorer les modèles existants et possibles de l’effet de cette forme d’ignorance forte sur la décision. Par ailleurs, le terme *unawareness* ayant été introduit d’abord par des logiciens, il nous a paru intéressant de chercher à savoir si leurs travaux pouvaient être transposer en théorie de la décision.

2 De la logique de l'unawareness à la théorie de la décision

2.1 Modélisation de l'unawareness en logique épistémique

En logique épistémique, la notion d'awareness a été étudiée pour affaiblir l'hypothèse d'*omniscience logique* implicitement présente dans les modèles standards de la connaissance. En effet, dans ces modèles il est supposé que les agents connaissent d'une part toutes les formules valides et d'autre part toutes les conséquences logiques des propositions qu'ils savent, c'est-à-dire si un agent sait que p et sait que $p \rightarrow q$ alors il sait que q . Les premières contributions majeures sur ce sujet ont montré qu'une notion d'unawareness satisfaisant un certain nombre d'axiomes "intuitifs" ne pouvait être obtenue dans un modèle standard. (Dekel, Lipman et Rustichini [15]). La solution proposée afin de dépasser ces difficultés techniques est de faire du raisonnement un processus "*awareness-dépendant*". Cela revient à réduire le langage de l'agent aux faits dont il est aware et de n'autoriser le raisonnement que dans les limites de ce langage. Différents modèles permettant cette limite ont été proposés dans lesquels l'awareness est définie soit comme un opérateur syntaxique introduit dans le langage, soit directement à partir de l'opérateur de connaissance déjà présent dans le langage de la logique épistémique.

Fagin et Halpern [23] ont construit le premier modèle de logique épistémique autorisant l'unawareness de l'agent. L'opérateur d'unawareness y est introduit indépendamment de l'opérateur de connaissance. L'intuition derrière ce modèle repose sur la distinction entre deux formes de connaissances: la connaissance *implicite*, notion de connaissance habituellement considérée en logique épistémique et permettant à l'agent de raisonner sur n'importe quelles propositions du langage, et la connaissance *explicite* qui capture exactement ce dont l'agent est explicitement aware et donc les raisonnements qu'il peut faire explicitement.

Ils considèrent le langage suivant, basé sur l'ensemble de variables propositionnelles $\mathcal{P} = \{p, q, \dots\}$:

$$\phi, \psi \dots ::= p, q, \dots \mid \phi \wedge \psi \mid \neg \phi \mid K\phi \mid X\phi \mid A\phi$$

où K est l'opérateur de connaissance explicite, X est l'opérateur de connaissance implicite et A est l'opérateur d'awareness.

Une *structure d'awareness* est alors définie comme un vecteur $\langle \mathfrak{W}, \rho, \mathcal{K}, \mathcal{A} \rangle$ où:

- \mathfrak{W} est un ensemble de mondes possibles
- ρ est une application qui à chaque $p \in \mathcal{P}$ associe l'ensemble des mondes $\rho(p) \subseteq \mathfrak{W}$ où p est vrai
- \mathcal{K} est une relation binaire sur \mathfrak{W} (sérielle, transitive et euclidienne)
- \mathcal{A} est une application qui associe à tout monde $w \in \mathfrak{W}$ un ensemble de formules $\mathcal{A}(w)$ qui correspond intuitivement à l'ensemble des formules dont l'agent est aware dans ce monde

Modica et Rustichini [25, 26] ont proposé une logique où l'opérateur d'awareness est défini de manière syntactique à partir de l'opérateur de connaissance c'est-à-dire, un agent est aware de p s'il sait que p ou bien s'il sait qu'il ne sait pas p .

$$A(p) = K(p) \vee K\neg K(p)$$

Ils généralisent le modèle standard de la logique des mondes possibles afin d'avoir la possibilité de restreindre l'ensemble de variables propositionnelles disponible pour décrire le raisonnement de l'agent en un monde aux seules variables propositionnelles dont l'agent est aware. Par la suite, Halpern [12] a montré que les Modèles Standard Généralisés ainsi définis par Modica et Rustichini peuvent en fait être considérées comme des cas particuliers de *structures d'awareness* où l'ensemble de propositions dont l'agent est aware est généré par un ensemble de propositions primitives et où l'awareness est close par sous-formules c'est-à-dire que si l'agent est aware de ϕ , il est aware de toutes les propositions primitives qui composent ϕ .

2.2 Utilisation de la logique de l'unawareness en théorie de la décision

Ainsi, on a pu constater d'une part que la théorie standard de la décision ne permet pas de représenter l'unawareness de l'agent et d'autre part que la littérature sur l'unawareness [23, 25] vise avant tout à en étudier les propriétés épistémiques sans en donner une caractérisation en terme de comportement. Une idée naturelle serait de faire dialoguer les deux littératures. Certains auteurs tentent donc de construire un pont entre ces deux littératures et d'utiliser les idées issues de l'étude épistémique de la notion d'unawareness afin de développer des théories économiques qui prennent en compte l'influence de cette forme forte d'ignorance sur le comportement. Par exemple, Heifetz, Meier et Schipper ont dans un premier temps défini et étudié un certain type de structures d'awareness, basées non plus sur un unique espace d'état mais sur un ensemble partiellement ordonné d'espaces d'état qui représentent les différents niveaux d'awareness possibles d'un agent [16]; puis ils ont utilisés ces structures dans l'étude des échanges spéculatifs entre des agents ne bénéficiant pas des mêmes informations [18] ou des jeux dynamiques sous forme extensive [17].

Cependant, ces modèles considèrent systématiquement les croyances des agents sur la représentation du monde et leur évaluation des différentes options comme des données primitives du problème alors que le but de la théorie de la décision est justement de déduire ces éléments de la relation de préférence de l'agent. Schipper [2] cherche donc à construire un modèle qui corresponde à cette attente fondamentale de la théorie de la décision tout en autorisant l'agent à être unaware de certaines caractéristiques de l'espace d'états, Il remplace donc l'espace d'états standards de la théorie d'Anscombe-Aumann par une structure d'awareness telle que définie par [16] afin d'obtenir un théorème de représentation des préférences par une maximisation de l'utilité espérée qui dépende du niveau d'awareness de l'agent.

Ce modèle a deux défauts majeurs:

- Il ne prend en compte que l'unawareness que l'agent peut avoir à propos de l'espace d'états et n'autorise pas l'agent être unaware des autres données du problème de décision, soient les conséquences et les actes
- D'autre part, c'est un modèle qui serait impossible à tester expérimentalement car il faudrait être capable de déterminer une relation de préférence pour chaque niveau d'awareness du même sujet, ce qui ne semble pas réaliste puisqu'une fois qu'on pose une question au sujet, on le rend aware de force du sujet de la question.

Dans la suite, nous proposons une utilisation différente de la logique de l'awareness appliquée à la théorie de la décision qui essaye de répondre à cette première critique au moins. Malheureusement,

ce n'est pas sûr que cette approche soit encore la meilleure pour l'étude empirique de la question de l'unawareness.

2.3 Un essai de formalisation

Notre proposition est différente de l'approche de Schipper qui cherche à importer les structures logiques d'unawareness dans le formalisme de la théorie de la décision. Ici, nous voulons plutôt importer les éléments de base de la théorie de la décision dans une modélisation logique afin de pouvoir exprimer l'unawareness des données d'un problème de décision à l'aide d'opérateurs logiques.

Dans cette démarche, on cherche à s'inspirer de la théorie épistémique des jeux pour construire le modèle logique qui sous-tendrait l'axiomatique de la théorie de la décision. On peut construire un modèle épistémique d'un jeu sous forme stratégique si le jeu est donné sous la forme $G = \langle Ag, \{S_i, \pi_i\}_{i \in Ag} \rangle$. On construit la structure $\langle \mathfrak{W}, \{R_i\}_{i \in Ag}, \{\sigma_i\}_{i \in Ag} \rangle$ où $\langle \mathfrak{W}, \{R_i\} \rangle$ est un cadre de Kripke qui représente les croyances des agents et pour tout $i \in Ag$, $\sigma_i : \mathfrak{W} \rightarrow S_i$ est une fonction qui indique la stratégie choisie par le joueur i dans chaque monde de \mathfrak{W} . Cette fonction vérifie une condition qui assure que le joueur est toujours certain de la stratégie qu'il choisit lui-même.

$$\text{si } w' \in R_i(w) \text{ alors } \sigma_i(w') = \sigma_i(w)$$

L'idée serait de construire un modèle similaire pour un problème de décision. Cela paraît à la fois plus simple car on n'a seulement besoin de se placer dans un cadre avec un seul agent et plus compliqué car cet agent doit avoir connaissance d'un plus grand nombre de données sur le problème.

Un modèle épistémique du problème de décision pourrait alors se présenter comme une structure $\langle \mathfrak{W}, R, \Sigma \rangle$ où $\langle \mathfrak{W}, R \rangle$ est une structure de Kripke et σ attribue à chaque monde une description de l'état de la nature, l'action choisie par l'agent et la conséquence qu'il considère que ça va lui apporter. Cette fonction devra satisfaire une condition parallèle à la précédente assurant que l'agent est certain de l'action qu'il choisit. En revanche, il n'est pas nécessaire qu'il soit certain de la description de la nature et a plus forte raison de la conséquence de l'action qu'il choisit.

Afin de pouvoir parler de choix entre des actions dans une telle structure, il est nécessaire de préciser le langage sur lequel va se baser notre logique. Ce langage doit permettre d'exprimer les différentes données d'un problème de décision. Pour cela on va reprendre les éléments primitifs de la théorie de Savage, les états, les conséquences et les actes. On choisira de garder ces trois espaces et non d'en exprimer l'un à l'aide des deux autres comme on l'a expliqué précédemment.

Cependant, il y a encore deux possibilités de formalismes pour représenter l'espace d'états. On peut choisir de représenter l'espace d'états comme un simple ensemble d'états possibles donnés et lorsqu'un nouvel état est découvert cela revient simplement à ajouter un élément à l'ensemble précédent. C'est cette approche que nous avons adoptée dans la première partie du mémoire. Cela peut correspondre aux cas où l'on fait une liste d'états différents de la nature qui s'excluent mutuellement. Par exemple la liste des états

$$\{s_1 : \text{il fait beau}, s_2 : \text{il pleut}\}$$

. Dès lors, devenir aware d'un nouvel état reviendra à rallonger sa liste, par exemple de la manière suivante:

$$\{s_1 : \text{il fait beau}, s_2 : \text{il pleut}, s_3 : \text{il grêle}\}$$

. Mais on peut aussi décrire chaque état comme la liste de ses différentes propriétés. On peut par exemple considérer la propriété “ p : il pleut”. Celle-ci permet de définir deux états différents de la nature soit

$$\{[p] : \text{état où il pleut}, [\bar{p}] : \text{état où il ne pleut pas}\}$$

. Dans ce cas, on ne devient pas directement aware d’un nouvel élément de l’ensemble mais plutôt d’une nouvelle propriété caractéristiques de l’environnement, par exemple “ q : il y a du vent”. Cela engendre un plus grand nombre d’états de la nature définis de la même manière que précédemment:

$$\{[p, q], [p, \bar{q}], [\bar{p}, q], [\bar{p}, \bar{q}]\}$$

Awareness spécifique

On va donc considérer trois ensembles primitifs de propositions:

- un ensemble \mathcal{P} de caractéristiques qui décrivent les états de la nature (et leur négation)
- un ensemble \mathcal{F} d’actions
- un ensemble \mathcal{C} de conséquences

Ensuite, il est aussi nécessaire de pouvoir parler de préférence entre deux actions donc il est possible qu’il faille ajouter une relation de préférence \geq à notre langage.

Cela nous donne le langage suivant:

$$E ::= p, \bar{p} \in \mathcal{P} \mid E \wedge E$$

$$\mathcal{L} ::= p, \bar{p} \in \mathcal{P} \mid E \in At(\mathcal{P}) \mid f, g \in \mathcal{F} \mid c \in \mathcal{C} \mid E \wedge f \wedge c \mid f \geq g$$

Ensuite, on veut que la relation syntaxique \geq corresponde, dans la structure sémantique que nous nommes en train de construire, aux préférences de l’agent qui sont déterminées par sa fonction d’utilité et sa manière d’attribuer des probabilités aux états du monde.

On va donc considérer un modèle $\mathcal{M} = \langle W, \pi, \{\mu_f\}_{f \in \mathcal{F}}, u \rangle$.

- W est un ensemble de mondes possibles qui correspond à l’ensemble $At(\mathcal{P}) \times \mathcal{F} \times \mathcal{C}$
- $\pi: W \rightarrow \mathcal{P} \cup \mathcal{F} \cup \mathcal{C} \rightarrow \{0, 1\}$ est l’attribution de valeur de vérité aux trois ensembles primitifs
- pour tout $f \in \mathcal{F}$, $\mu_f: At(\mathcal{P}) \times \mathcal{C} \rightarrow [0, 1]$ est une mesure de probabilité
- $u: \mathbb{R} \rightarrow \mathbb{R}$ est une fonction continue strictement croissante

Hypothèse 2.1. La probabilité marginale sur les états est indépendante de l’acte choisi.

On va donc noter pour tout $E \in At(\mathcal{P})$: $\mu(E) := \sum_{c \in \mathcal{C}} \mu_f(E, c)$ quelque soit $f \in \mathcal{F}$.

On définit ensuite une relation de satisfaction qui s’inspire du modèle issu de la théorie des jeux ainsi que du théorème de représentation de Savage pour l’interprétation de la relation de préférence.

$$\begin{array}{llll}
\text{pour } p \in \mathcal{P} : & \mathcal{M}, w \models p & \text{ssi} & \pi(w)(p) = 1 \\
\text{pour } E, E' \in At(\mathcal{P}) : & \mathcal{M}, w \models E \wedge E' & \text{ssi} & w \models E \text{ et } w \models E' \\
\text{pour } f \in \mathcal{F} : & \mathcal{M}, w \models f : & \text{ssi} & \pi(w)(f) = 1 \\
\text{pour } c \in \mathcal{C} : & \mathcal{M}, w \models c : & \text{ssi} & \pi(w)(c) = 1 \\
\text{pour } E \in At(\mathcal{P}), f \in \mathcal{F} \text{ et } c \in \mathcal{C} : & w \models E \wedge f \wedge c & \text{ssi} & w \models E \text{ et } w \models f \text{ et } w \models c \\
\text{pour } f, g \in \mathcal{F} : \mathcal{M} \models f \geq g & \text{ssi} & \sum_{\{w \models E \wedge f \wedge c\}} \mu_f(E, c)u(c) \geq \sum_{\{w \models E \wedge g \wedge c\}} \mu_g(E, c)u(c)
\end{array}$$

On peut essayer de représenter une forme d'awareness dans notre structure de la manière suivante:

$$\text{pour } E \in At(\mathcal{P}), f \in \mathcal{F}, c \in \mathcal{C} : \mathcal{M}, w \models A(E \wedge f \wedge c) \text{ ssi } \mu_f(E, c) > 0$$

Malheureusement, il semble que l'on fait fausse route ici. En logique épistémique, on représente l'unawareness en restreignant le langage de l'agent aux seules propositions dont il est aware. En considérant comme ensemble primitif à la fois les actes, les états et les conséquences, on a réduit tout le langage aux seules données dont l'agent serait aware. On ne peut plus distinguer entre l'espace de l'agent et l'espace de l'observateur, ce qui ne permet pas de considérer un opérateur d'awareness non trivial.

Awareness dynamique

En revanche, il est peut-être possible d'utiliser cette construction théorique pour s'intéresser à la dynamique de l'awareness et à la manière dont l'agent peut faire évoluer son attribution de probabilité et comment il déduit les conséquences de ses actions lorsqu'il devient aware d'un nouvel état de la nature. On va donc considérer deux telles structures, $\mathcal{M} = \langle W, \pi, \{\mu_f\}_{f \in \mathcal{F}}, u \rangle$ représentera la situation avant gain d'awareness et $\mathcal{M}^* = \langle W^*, \pi^*, \{\mu_f^*\}_{f \in \mathcal{F}^*}, u^* \rangle$ après.

Considérons le cas suivant; l'agent considère un acte f , qui peut avoir pour conséquence c ou d selon que l'état du monde est $[p]$ ou $[\bar{p}]$.

$$\begin{array}{l}
w_1 : p \wedge f \wedge c \\
w_2 : \bar{p} \wedge f \wedge d
\end{array}$$

Figure 7: Représentation subjective avant gain d'awareness

A présent, l'agent devient aware d'une nouvelle caractéristique environnementale q . Comment va-t-il adapter sa représentation du monde? On se retrouve plus ou moins dans la situation décrite dans la première partie de ce mémoire. Cependant on peut essayer de se restreindre à quelques cas simples.

| Cas 1 | Cas 2 | Cas 3 |
|--|---|--|
| $w_1^* : (p \wedge q) \wedge f \wedge c$ | $w_1^* : (p \wedge q) \wedge f \wedge c$ | $w_1^* : (p \wedge q) \wedge f \wedge ??$ |
| $w_2^* : (p \wedge \bar{q}) \wedge f \wedge c$ | $w_2^* : (p \wedge \bar{q}) \wedge f \wedge ??$ | $w_2^* : (p \wedge \bar{q}) \wedge f \wedge c$ |
| $w_3^* : (\bar{p} \wedge q) \wedge f \wedge d$ | $w_3^* : (\bar{p} \wedge q) \wedge f \wedge d$ | $w_3^* : (\bar{p} \wedge q) \wedge f \wedge ??$ |
| $w_4^* : (\bar{p} \wedge \bar{q}) \wedge f \wedge d$ | $w_4^* : (\bar{p} \wedge \bar{q}) \wedge f \wedge ??$ | $w_4^* : (\bar{p} \wedge \bar{q}) \wedge f \wedge d$ |

Figure 8: 3 Représentations subjectives possibles après gain d'awareness

- Dans le cas 1, l'agent considère l'événement comme non informatif et ne change rien à sa représentation du monde malgré la nouvelle information
- Dans le cas 2, l'agent se comporte comme s'il avait considéré la réalisation de q comme un fait acquis, à l'étape précédente. Il doit maintenant déterminer quelles seront les conséquences du choix de l'action f dans les états où q n'est pas réalisé
- Dans le cas 3, c'est l'exact opposé, l'agent considèrerait la non réalisation de q comme un fait acquis et doit donc déterminer les conséquences de son choix dans les états où q est effectivement réalisé.

Ces trois cas sont assez simples à analyser individuellement. Cependant, rien ne force un agent à agir systématiquement selon l'un ou l'autre de ces cas simples. Il peut simplement employer l'une ou l'autre des stratégies correspondantes en fonction du contexte, voire des stratégies plus compliquées encore. On ne peut pas imposer de règles de conduite aux agents que l'on prétend modéliser mais on peut structurer notre modèle autour de certaines contraintes qui nous semblent rationnelles. On peut notamment imaginer des règles que doivent respecter les attributions de probabilité lors d'un gain d'awareness.

Règles possibles:

- Les rapports de probabilités marginales sur l'espace initial sont conservés au cours du temps:
 - soit sur l'ensemble des événements $At(\mathcal{E})$

$$\frac{\mu_f^*(\phi)}{\mu_f(\phi)} = \frac{\mu_f^*(\psi)}{\mu_f(\psi)} \quad \forall \phi, \psi \in At(\mathcal{E}) \quad (1)$$

- soit sur l'ensemble des conséquences \mathcal{C}

$$\frac{\mu_f^*(c)}{\mu_f(c)} = \frac{\mu_f^*(d)}{\mu_f(d)} \quad \forall c, d \in \mathcal{C} \quad (2)$$

- soit sur l'espace produit

$$\frac{\mu_f^*(\phi, c)}{\mu_f(\phi, c)} = \alpha \quad \forall \phi \in At(\mathcal{E}) \quad \forall c \in \mathcal{C} \quad (3)$$

- Les probabilités elles-mêmes sont conservées lors du passage à la plus grande structure:
 - soit sur l'ensemble des événements $At(\mathcal{E})$

$$\mu_f(\phi) = \mu_f^*(\phi) \quad \forall \phi \in At(\mathcal{E}) \quad (4)$$

- soit sur l'ensemble des conséquences \mathcal{C}

$$\mu_f(c) = \mu_f^*(c) \quad \forall c \in \mathcal{C} \quad (5)$$

Là encore le choix entre ces différentes règles n'est pas imposé par notre modèle. Cependant, en choisir certaines impose des contraintes sur les différents cas, les différentes stratégies d'adaptation au

gain d'awareness que l'on peut traiter. Choisir les hypothèses (2) et (3) simultanément par exemple empêche le cas suivant de se produire.¹

$$\begin{aligned} w_1^* &: p \wedge q \wedge f \wedge e \\ w_2^* &: p \wedge \bar{q} \wedge f \wedge c \\ w_3^* &: \bar{p} \wedge q \wedge f \wedge d \\ w_4^* &: \bar{p} \wedge \bar{q} \wedge f \wedge d \end{aligned}$$

Figure 9: Une autre représentation subjective après gain d'awareness

En effet, l'hyp (2) donne un certain réel α tel que: $\mu_f(w_2^*) = \alpha\mu_f(w_1)$ et $\mu_f(w_3^*) + \mu_f(w_4^*) = \alpha\mu_f(w_2)$. L'hyp (3) donne $\mu_f(w_1) = \mu_f(w_1^*) + \mu_f(w_2^*)$ et $\mu_f(w_2) = \mu_f(w_3^*) + \mu_f(w_4^*)$. Ce qui implique que $\mu_f(w_1^*) = 0$ et trivialise le raisonnement.

Nous avons donc présenté ici différentes petites applications à l'étude de la notion d'unawareness que pourrait avoir un modèle logique de la théorie de la décision. Ce n'est certes pas entièrement satisfaisant, nous n'avons pas vraiment réussi à analyser les questions liées l'awareness dans toute leur généralité mais ce genre de croisement entre différentes disciplines est toujours intéressant à explorer, ne serait-ce que pour mettre en évidence les différences et les impossibilités à faire coïncider parfaitement deux domaines distinctes.

3 L'analyse de l'imprévu en théorie de la décision

3.1 La prise en compte des événements imprévus

En théorie de la décision, on modélise la manière dont le sujet se représente l'incertitude par une fonction de probabilité sur les ensembles de mondes possibles. S'intéresser à la problématique de l'unawareness revient à se demander si la représentation de l'ignorance se réduit à cette incertitude probabiliste et s'il n'est pas possible de représenter également une forme plus forte d'ignorance qui est justement l'ignorance sur la composition-même de l'espace d'états du monde.

Kreps [19] est le premier à observer que les modèles standards de théorie de la décision échouent à prendre en compte les événements imprévus (*unforeseen contingencies*). Il propose donc un théorème de représentation qui permette de modéliser un agent qui prend en compte la potentialité d'événements imprévus lors de son choix. Dans la terminologie employée jusqu'à présent, on dirait que cet auteur vise à modéliser l'unawareness consciente portant sur l'espace d'états. De plus, pour Kreps, la meilleure manière de tester si un individu anticipe le fait qu'il peut avoir une vision erronée du problème auquel il fait face, est d'évaluer sa volonté de préserver une certaine flexibilité, de ne pas s'engager complètement dans sa décision au cas où il pourrait bénéficier d'informations supplémentaires. La notion de contingences imprévues a donc aussi dans son travail un très fort lien avec la dynamique du choix. Dekel, Lipman et Rustichini [15] définissent une notion de contingences imprévues plus précise qui se rapprocherait plutôt de ce que nous avons appelé l'unawareness spécifique, toutefois portant toujours uniquement sur l'espace d'états et non sur les autres éléments pertinents d'un problème de décision.

¹La présence d'une conséquence e peut surprendre mais à aucun moment je n'ai précisé de quels éléments étaient constitués les espaces primitifs ici donc on peut par exemple supposer que depuis le départ $\mathcal{C} = \{c, d, e\}$ puisque l'on n'a considéré ici que des exemples

Une question reste pourtant en suspens: est-ce qu'être unaware d'un événement ne conduit pas au même comportement observable qu'être aware de cet événement mais en lui attribuant une probabilité nulle? En effet, si la notion peut être montrée équivalente à une construction qui existerait déjà dans les modèles classiques de la prise de décision, cela n'aurait que peu d'intérêt de tenter de formaliser de nouvelles théories qui intègrent cette notion. Or, un individu sera certes incapable de prendre en considération les conséquences d'un événement dont il est *unaware* mais il ne tiendra pas non plus compte des conséquences d'un événement auquel il attribue une probabilité nulle, dans la mesure où il est rationnel. Il y a pourtant une distinction conceptuelle très claire entre l'*unawareness* qui est l'ignorance de la possibilité même qu'un événement se produise et le fait d'attribuer une probabilité nulle qui revient à considérer cet événement comme impossible, ce qui nécessite bien sûr d'en être *aware*.

La notion d'*unawareness* ne peut pas être facilement transcrite en probabilités classiques car elle est par nature symétrique. En effet, lorsque l'on est unaware d'un événement E , on est aussi unaware de son complémentaire $\neg E$ et inversement.

$$unaware(E) \leftrightarrow unaware(\neg E)$$

On va montrer qu'il est impossible d'attribuer une probabilité nulle à la fois à un événement et à son complémentaire. L'attribution d'une probabilité nulle à un événement se caractérise en termes de préférences de la même manière qu'un événement 'nul au sens de Savage'.

Un événement est *Savage-nul* si pour tous actes f et g , $f_E h \sim g_E h$, c'est-à-dire que l'agent est indifférent entre deux actes dont les issues sont identiques sur la partie de l'espace d'états où E ne se réalise pas.

Dès lors, si l'*unawareness* d'un événement se définit comme l'attribution d'une probabilité nulle à cet événement, on a:

$$unaware(E) \leftrightarrow f_E h \sim g_E h \rightarrow_{h=f} f \sim g_E f$$

Et donc également:

$$unaware(\neg E) \leftrightarrow f_{\neg E} h \sim g_{\neg E} h \leftrightarrow h_E f \sim h_E g \rightarrow_{f=g} g_E f \sim g$$

Donc il aurait indifférence entre tous actes f et g : $f \sim g_E f \sim g$, ce qui est absurde.

Modica, Rustichini et Tallon [1] proposent également une expérience de pensée inspirée de modèles concrets de spéculation afin de démontrer qu'il n'est pas possible d'interpréter un modèle où un agent est unaware de certains états comme un modèle où il attribue une probabilité nulle à ces états. D'après ces auteurs, on peut différencier ces deux niveaux d'ignorance par observation du comportement d'un agent vis à vis d'un état possible dans lequel il risque la faillite. Si cet agent attribue une probabilité nulle à cet état, il lui attribue en fait une probabilité infinitésimale ce qui peut conduire à deux conséquences. Dans le cas où cette probabilité, multipliée à la punition supposée infinie de la faillite, rend l'utilité de cet état infiniment négative, il pourra éviter la faillite mais dans le cas où le produit des deux reste un infiniment petit, il risque de perdre une quantité arbitrairement grande d'argent. En revanche, un agent qui est unaware d'un tel état ne pourra jamais éviter la faillite, il n'y aura tout

simplement "pas pensé", mais il perdra systématiquement une quantité finie d'argent déterminée par ses investissements sur les états dont il est effectivement aware.

Nous soutenons l'idée selon laquelle pour qu'un modèle de l'unawareness spécifique en situation de décision soit pertinent, il faudrait que la représentation de celle-ci soit distincte de l'attribution de probabilités nulles, qu'elle porte sur les états de la nature ou sur les liens entre un acte et ses conséquences. Cependant, les modèles actuels de l'unawareness ne permettent en fait pas de rendre compte de cette différence conceptuelle pour le moment, comme le souligne Li [10]. En revanche, la modélisation de la dynamique de l'awareness semble plus accessible pour les théoriciens de la décision car les caractéristiques que doit vérifier le modèle sont sans doute plus intuitives.

Approche intéressante parce qu'on veut rendre compte de contingences imprévues : au moment où le sujet prend une décision, les contingences que tu prévoies pas elles interviennent pas. Unawareness devient intéressant à utiliser au moment où la décision de l'agent dévie par rapport aux prédictions de la théorie. Ça veut dire que la formulation subjective du problème de décision est différente de celle de l'observateur.

3.2 Dynamique de l'awareness

Le modèle le plus abouti en ce qui concerne la dynamique de l'awareness est celui présenté par Karni et Vierø [5]. Il ne se base ni sur la théorie de VonNeumann et Morgenstern, ni sur la théorie de Savage, mais sur une théorie développée par Schmeidler et Wakker [3] et Karni et Schmeidler [4]. Celle-ci considère les ensembles de conséquences et d'actes comme primitifs et permet de construire un espace d'états à partir de ces ensembles. Un état sera alors un lien entre un acte et une conséquence ce qui permet d'après les auteurs d'obtenir un espace d'états plus abstraits, donc une résolution de l'incertitude qui ne dépend que des éléments concrets du choix, déconnectée de toutes contingences dues à l'environnement.

Ils étudient quelle dynamique est impliquée par un gain d'awareness au sein du modèle suivant.

Eléments de construction du modèle [5]

C : ensemble fini, non vide de conséquences

F : ensemble fini, non vide d'actions

C^F : espace des états concevables, l'ensemble des applications de l'espace des actes dans l'espace des conséquences.

$\hat{F} = \{f: C^F \rightarrow \Delta(C)\}$: ensemble des actes concevables

$\succeq_{\hat{F}}$ est la relation de préférences qui caractérise le comportement de choix du décideur

$S(C, F) \subseteq C^F$: les états que le décideur considère comme non nuls au sens de Savage, définissent l'espace des états faisables

On pourrait interpréter l'espace des états concevables comme l'espace de l'observateur alors que l'espace des états faisables serait l'espace du décideur. D'ailleurs, pour Karni et Vierø, les états dont on est aware sont les états que l'on considère faisables. Lorsque l'on devient aware d'un nouvel état, il s'agit d'un état contenu dans l'espace des états concevables.

Ils distinguent trois types de gains d'awareness (en pointillés sur Fig. 10):

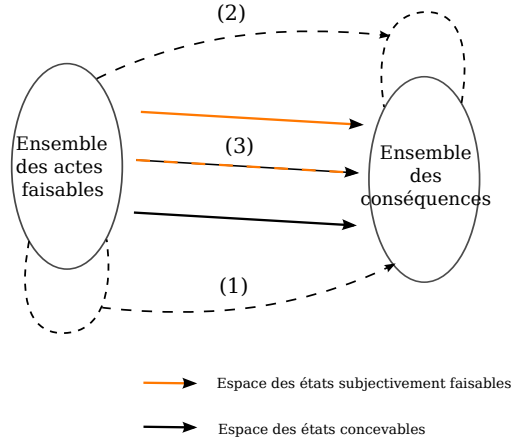


Figure 10: Représentation figurative de la construction du modèle

- découverte d'une nouvelle conséquence
- découverte d'un nouvel acte
- découverte d'un nouveau lien acte-conséquence (état)

Pour Karni et Vierø, la découverte d'une nouvelle conséquence conduit à l'extension de l'espace d'état, ce qui peut affecter le choix du décideur. En effet, deux actes qui coïncident sur l'ancien espace d'état peuvent devenir distincts lorsque l'on les associe à de nouvelles conséquences et le décideur peut ne plus être indifférent entre ces deux actes. Cependant, il faut que la nouvelle conséquence découverte soit effectivement perçue par le décideur comme une conséquence d'au moins un de ces actes pour que leur utilité soit modifiée par cette découverte et donc potentiellement les préférences de l'agent.

Dynamique du modèle

Hypothèse 3.1 (AA-EU). $\succeq_{\hat{F}}$ satisfait les axiomes de la théorie de l'espérance d'utilité formulés par Anscombe et Aumann [9].

Hypothèse 3.2 (Act-independence). L'ordre de préférence des lotteries conditionnellement à n'importe quel état dans l'espace originel est indépendant de l'ensemble des actes sur lequel la relation de préférence est définie, c'est-à-dire la relation de préférence conditionnelle aux états de l'espace d'états originel n'est pas affecté par un gain d'awareness.

Hypothèse 3.3 (Awareness consistency). Un gain d'awareness n'altère pas la relation de préférence sur l'espace d'état originel.

En distinguant les trois types d'awareness énoncés ci-dessus, les auteurs peuvent énoncer trois théorèmes de représentation correspondants à chacune de ces situations. Ici nous allons nous contenter d'une formulation globale simplifiée.

Théorème 3. Soient C_0 et F_0 (resp. C_1 et F_1) les ensembles de conséquences et d'actes avant (resp. après) augmentation du niveau d'awareness de l'agent. Les propositions suivantes sont équivalentes:

- (1) $\succeq_{\hat{F}_0}$ et $\succeq_{\hat{F}_1}$ satisfont chacune l'Hypothèse 3.1 et les Hypothèses 3.2 et 3.3 conjointement.
- (2) Il existe une fonction U affine non constante sur $\Delta(C)$ et deux mesures de probabilités $\pi_{\hat{F}_0}$ sur $C_0^{F_0}$ et $\pi_{\hat{F}_1}$ sur $C_1^{F_1}$ telles que:

pour tout $f, g \in \hat{F}_0$:

$$f \succeq_{\hat{F}_0} g \iff \sum_{s \in C_0^{F_0}} U(f(s)) \pi_{\hat{F}_0}(s) \geq \sum_{s \in C_0^{F_0}} U(g(s)) \pi_{\hat{F}_0}(s)$$

pour tout $f, g \in \hat{F}_1$:

$$f \succeq_{\hat{F}_1} g \iff \sum_{s \in C_1^{F_1}} U(f(s)) \pi_{\hat{F}_1}(s) \geq \sum_{s \in C_1^{F_1}} U(g(s)) \pi_{\hat{F}_1}(s)$$

De plus, U est unique à une transformation affine positive près et $\pi_{\hat{F}_0}$ et $\pi_{\hat{F}_1}$ sont uniques et vérifient:

$$\text{pour tous } s, s' \in C_0^{F_0} \text{ non nuls : } \frac{\pi_{\hat{F}_0}(s)}{\pi_{\hat{F}_0}(s')} = \frac{\pi_{\hat{F}_1}(E(s))}{\pi_{\hat{F}_1}(E(s'))}$$

Les auteurs qualifient ce résultats de *bayésianisme inverse*. En effet, ce résultat très élégant fait écho à la règle de Bayes qui est classiquement utilisée en théorie de la décision pour décrire la révision de croyances probabilistes qui s'opère lorsqu'une information signale la réduction de l'espace d'état à un événement plus restreint. En effet, comme la règle de Bayes, leur énoncé permet de connaître que les rapports entre probabilités avant et après le gain d'information. Lorsque l'on applique la règle de Bayes, pour calculer le niveau absolu des probabilités *a posteriori*, il faut en plus du ratio de proportionnalité, la valeur des probabilités *a priori*. De manière parallèle dans le modèle de Karni et Vierø, il est nécessaire de connaître les probabilités *a posteriori* pour pouvoir calculer les valeurs précises des probabilités *a priori*. Ce modèle permet de répondre à la question que l'on posait à la fin de la première section puisque, par ce résultat de bayésianisme inverse, Karni et Vierø fournissent un moyen de calculer les probabilités attribuer par l'agent aux différents états du monde après gain d'awareness.

Cependant, d'autres interrogations surgissent une fois que l'on est face à ce modèle:

- Ne serait-il pas pertinent d'introduire un espace d'état exogène, de la même manière qu'il est nécessaire de faire appel à un ensemble d'actions différents des actes tels que définis par Savage? Cela permettrait de considérer des cas où le décideur ignore ce qui détermine l'issue d'une action et envisage plusieurs conséquences possibles
- L'agent a-t-il conscience qu'il ne maîtrise pas tous les éléments de la situation de choix? Dans quelle mesure cette conscience correspond-elle à la réalité? Quelle est son attitude vis à vis de sa propre unawareness: est-ce que l'agent a tendance à choisir les actions qu'il pense connaître le mieux?

3.3 Unawareness consciente

La difficulté de modéliser de l'unawareness consciente vient du paradoxe suivant, mis en évidence par Halpern et Rêgo [13]. Ces trois propositions qui individuellement paraissent des hypothèses simples et intuitives s'avèrent être incompatibles:

- Il paraît clair que dans un modèle économique représentant un agent qui a conscience de son unawareness, on veut que lorsqu'un agent est aware de quelque chose, il croit effectivement qu'il en est aware.
- Ce n'est pas parce que l'agent est effectivement complètement aware des informations pertinentes pour la situation de choix, qu'il le sait. Il peut penser qu'il y a malgré tout des choses dont il est unaware.
- Parallèlement, ce n'est pas parce qu'un agent n'est pas entièrement aware de la situation de choix, qu'il le sait. Il peut penser qu'il est complètement aware alors même qu'il y a des éléments pertinent qui lui échappent.

Il y a deux manières d'adresser ce problème se distinguent dans la littérature. La première repose sur une intuition linguistique assez intuitive et la seconde est plutôt issue de la réflexion sur la structure des modèles de choix eux-mêmes.

Comme la logique standard utilisée pour traiter de la connaissance et de l'awareness ne permet pas d'exprimer le fait qu'un agent sait qu'il y a des choses dont il est unaware, Halpern et Rego [13] proposent d'étendre la logique d'awareness de Fagin et Halpern [23] pour permettre la quantification (sur les variables propositionnelles) de manière à ce qu'une formule du langage puisse exprimer "un agent sait explicitement qu'il existe une chose dont il est aware".

$$X(\exists p \neg Ap) \text{ où } X \text{ désigne la connaissance explicite}$$

Cette formule peut être vraie sans qu'il n'y ai aucune formule A telle que l'agent sache explicitement qu'il est unaware de A .

Cette alternative permet de résoudre cette incohérence en faisant en sorte que l'ensemble des choses dont l'agent pourrait être aware varie d'un monde à l'autre, c'est-à-dire qu'un agent peut être aware des mêmes choses dans tous les mondes qu'il considère possibles sans pour autant que cela corresponde systématiquement à l'ensemble des choses dont il aurait pu être aware dans chaque monde et donc ce qui empêche qu'il se croit toujours pleinement aware. Cependant, cela oblige à supposer que l'agent accorde des croyances à des mondes 'impossibles' c'est-à-dire des mondes rendant vraies des propositions qui sont fausses dans le monde réel.

Dans leur système qui ne se base pas sur la logique, Karni et Vierø [6] utilise la même idée mais n'ont pas besoin de l'axiomatique sophistiquée développée par Halpern et Rêgo. Pour modéliser le fait que l'agent a conscience qu'il ne perçoit pas toutes les conséquences possibles de ses actions, il leur suffit d'ajouter à l'espace subjectif de l'agent une conséquence correspondant à "aucune des conséquences dont je suis aware". Ils obtiennent alors un théorème de représentation similaire à celui que l'on a détaillé dans la section précédente, dans ce contexte plus général.

Walker et Dietz [20] adopte une posture différente qui ressemble plutôt au travail de Ghirardato [21] que l'on a évoqué dans la première partie. L'idée est que le décideur estime que l'espace subjectif d'états

qu'il a en tête est insuffisamment riche pour permettre d'identifier toutes les conséquences des actions qu'il peut choisir. Il y a donc certaines actions présente dans la situation de choix qu'il ne "comprend" pas complètement. Il est possible que cela l'amène à être plus prudent (ou plus imprudent) vis à vis des actions qu'il ne comprendrait pas totalement.

Description du modèle [20]

X : espace des conséquences

\mathcal{A} : ensemble des actions

\succeq^* : relation de préférence sur \mathcal{A}

$\Sigma = X^{\mathcal{A}}$: ensemble des vecteurs de récompenses possibles, correspond à l'espace d'état chez Karni et Vierø

S : espace d'état subjectif

$W = S \times \Sigma$: ensemble de toutes les permutations états de la nature/vecteurs de conséquences

$W_s = \{s\} \times \Sigma$: sous-ensemble des permutations pour lesquelles s est réalisé

\mathcal{D} : espace des fonctions de W dans X

\succeq : relation de préférences sur \mathcal{D}

Pour $c, c' \in \mathcal{D}$ et $s \in S$, $c_s c'(w) = c(w)$ si $w = \{s, g\}$ et $c'(w)$ sinon

Axiome 3.1 (Monotonicity). Pour tout $c, c', c'' \in \mathcal{D}$, $c \succeq c'_s c$ ssi $c_s c'' \succeq c'_s c''$.

Cela permet de définir $c \succeq_s c'$ ssi $c_s c'' \succeq c'_s c''$ pour tout $c'' \in \mathcal{C}$.

Hypothèse 3.4 (Derivative SEU). $s \in S$ est soit nul, soit tel qu'il existe une fonction continue, bornée, strictement croissante $\phi_s : X \rightarrow \mathbb{R}$ et une mesure de probabilité μ_s telles que:

$$\text{pour tout } c, c' \in \mathcal{D} : c \succeq c' \iff \int_{W_s} \phi_s(c(w)) d\mu_s \geq \int_{W_s} \phi_s(c'(w)) d\mu_s$$

Hypothèse 3.5 (State Independence). Soient $s, s' \in S$ non nuls et $c_1, c_2, c_3, c_4 \in \mathcal{D}$ tels que, pour tout $Z \subseteq X$, $\mu_s(c_1^{-1}(Z)) = \mu_{s'}(c_2^{-1}(Z))$ et $\mu_s(c_3^{-1}(Z)) = \mu_{s'}(c_4^{-1}(Z))$, alors:

$$c_1 \succeq_s c_3 \iff c_2 \succeq_{s'} c_4$$

\mathcal{F} : ensemble des fonctions de S dans $\Delta(X)$, l'ensembles des distributions de probabilité sur X

\succeq^{AA} : relation de préférence sur \mathcal{F}

Hypothèse 3.6 (AA-EU). Il existe une fonction continue, bornée, strictement croissante $v : X \rightarrow \mathbb{R}$ et une mesure de probabilité π sur S telles que:

$$\text{pour tous } f, f' \in \mathcal{F} : f \succeq^{AA} f' \iff \sum_{s \in S} \pi(s) \mathbb{E}_{f(s)}[v(x)] \geq \sum_{s \in S} \pi(s) \mathbb{E}_{f'(s)}[v(x)]$$

Pour tout $a \in A$, il existe une fonction c_A dans \mathcal{D} telle que:

$$c_a(\{s, g\}) = g(a) \text{ pour tout } s \in S$$

Pour tout fonction $f \in \mathcal{F}$ telle que: $\forall s \in S. \exists x \in X. f(s) = x$, il existe une fonction c_f dans \mathcal{D} telle que:

$$c_f(\{s, g\}) = x \text{ ssi } f(s) = x \text{ pour tout } s \in S$$

Axiome 3.2 (Reduction). Les relations \succeq^* , \succeq^{AA} et \succeq sont cohérentes entre elles c'est-à-dire,

a) pour tout $a, a' \in A$: $a \succeq^* a'$ ssi $c_a \succeq c_{a'}$

b) pour tout $f, f' \in \mathcal{F}_\delta$: $f \succeq^{AA} f'$ ssi $c_f \succeq c_{f'}$

Théorème 4. *Les deux propositions suivantes sont équivalentes:*

- (1) \succeq^* , \succeq et \succeq^{AA} satisfont les Axiomes 3.1 et 3.2 ainsi que les Hypothèses 3.4, 3.5 et 3.6
- (2) Il existe des fonctions u et ϕ continues, bornées et strictement croissantes, une mesure de probabilité μ_s sur chaque \mathfrak{W}_s et une mesure de probabilité π sur S telles que:

$$\text{pour tous } a, a' \in A: a \succeq^* a' \iff \sum_{s \in S} \pi(s) \cdot u \left(\int_{W_s} \phi(c_a(w)) d\mu_s \right) \geq \sum_{s \in S} \pi(s) \cdot u \left(\int_{W_s} \phi(c_{a'}(w)) d\mu_s \right)$$

Ce modèle permet donc de répondre aux deux critiques que nous avons formulées à la fin de la section précédente. L'espace que considèrent les auteurs ici correspond effectivement à un produit de l'espace d'états abstrait considéré par Karni et Vierø [5] et d'un espace d'états exogène qui permet de prendre en compte la manière dont le sujet se représente effectivement les caractéristiques de l'état de la nature sur lesquelles porte l'incertitude.

D'autre part, le but de ce modèle est exactement de pouvoir représenter l'unawareness consciente d'un agent. Cependant, il s'agit uniquement de l'unawareness portant sur les liens entre actes et conséquences, c'est-à-dire la potentialité qu'un acte mène à une conséquence ou une autre. En revanche, ce modèle ne cherche pas à représenter explicitement l'unawareness consciente à propos de l'espace des états ou de l'ensemble des conséquences. Pour représenter cela, Walker et Dietz suggèrent d'ailleurs de l'intégrer à leur modèle exactement de la manière dont le font Karni et Vierø, c'est-à-dire en ajoutant un état ou une conséquence qui vaille pour "aucune des autres options".

Enfin, ce modèle est un modèle statique et il serait sans doute très intéressant de le développer en modèle dynamique, par exemple en étoffant le processus utilisé par Karni et Vierø, pour essayer de retrouver des résultats similaires sur l'espace d'états plus complexes considéré ici. Sachant que la plus ou moins forte concavité de la fonction ϕ du Théorème 4 indique la plus ou moins grande aversion de l'agent envers les actes qu'il ne comprend pas, on peut imaginer obtenir un résultat dynamique portant sur cet aspect; si devenir aware rend l'agent plus prudent, cela devrait se traduire sur la concavité de la fonction ϕ , possiblement en terme de dominance stochastique.

Conclusion

Nous avons globalement échoué à proposer un modèle satisfaisant de l'effet de l'unawareness dans la prise de décision. En effet, cette notion qui repose pourtant sur des intuitions psychologiques très basiques s'avère finalement compliquée à caractériser dans le cadre des théories de la décision. Cette étude préliminaire nous a néanmoins permis de mettre en évidence les problèmes qui surviennent quand on tente de prendre en compte cette notion de perception partielle dans les modèles théoriques.

D'ailleurs nous pensons que l'approche entièrement théorique de ce travail est peut-être une marque de trop grande ambition. Nous estimions que le fait que la logique parvienne à circonscrire le concept d'unawareness et ses propriétés épistémiques suffirait à rendre simple la tâche d'importer cette notion dans la modélisation de la prise de décision. Il est fort possible que l'approche expérimentale soit plus efficace pour déterminer quel effet a l'unawareness d'un agent sur son comportement de choix. Même si, bien sûr, l'exploration des modèles nous permet également de concevoir des designs expérimentaux efficaces et bien contrôlés.

En effet, nous l'avons vu, les données expérimentales dont nous disposons pour le moment concernant l'effet de l'unawareness sur la prise de décision sont plutôt insatisfaisantes. Dans l'expérience de Mengel et al., la condition d'*Unawareness* correspond à une absence totale d'information sur les probabilités et une information partielle sur les conséquences et la condition d'*emphAmbiguïté* est réalisée par une absence d'information au niveau des probabilités tout en maintenant une connaissance parfaite de l'espace de conséquences. Il ne semble pas que cela corresponde exactement à la différence entre risque, ambiguïté et unawareness telles que traitées dans la littérature théoriques. Pour traduire la notion d'ambiguïté (voir par exemple []) peut-être qu'il vaudrait mieux donner des intervalles potentiels de probabilités. D'autre part, l'unawareness semble plutôt être une forme d'ignorance orthogonale à la distinction risque/ambiguïté; elle pourrait donc intervenir ou non en sus dans chacune des conditions expérimentales. Beaucoup de pistes restent donc à explorer dans l'étude expérimentale de l'unawareness.

D'autre part, il pourrait être intéressant de mettre l'effet de l'unawareness en parallèle avec d'autres observations empiriques de violations des théories standards de la décision. Il a été notamment proposé par Ahn et Ergin que l'effet de cadre (*framing effect*) pourrait être dû à une mauvaise perception d'un gain d'awareness. Lorsque l'agent obtient une information, il l'a perçoit généralement comme pertinente pour sa représentation de la situation de choix et modifie le modèle qu'il en a. L'ajout d'une information non pertinente entraînerait alors une modification non justifiée de ce modèle et donc un effet de cadre. Cependant, d'autres causes psychologiques peuvent être invoquées pour expliquer les effets de cadre au-delà de la perception erronée: les limites de la mémoire, donc l'incapacité de rappeler l'ensemble des informations pertinentes au moment voulu ou bien le fait de rappeler des informations non pertinentes, un manque d'attention ou plus généralement le fait que certaines données apparaissent plus ou moins saillantes.

D'ailleurs, il pourrait être intéressant d'explorer encore plus loin les théories qui sous-tendent la notion de perception partielle en psychologie cognitive. Par exemple les théories des *modèles mentaux* (Johnson-Laird [22]) qui correspondent à une représentation du monde par l'individu dont la portée représentative varie en fonction du contexte et de la *catégorisation* (Goldstone et Kersten [24]) qui est supposée filtrer les observations qui arrivent au niveau de perception consciente des individus pourraient être de bons candidats.

References

- [1] Modica S. Rustichini A. and Tallon J.-M. Unawareness and bankruptcy: a general equilibrium model. *Economic Theory*, 12:259–292, 1998.
- [2] Schipper B.C. Preference-based unawareness. *Mathematical Social Science*, 2014.
- [3] Schmeidler D. and Wakker P. *Expected Utility and Mathematical Expectation*. The New Palgrave: A Dictionary of economics. Macmillan Press, 1987.
- [4] Karni E. and Schmeidler D. *Utility theory with uncertainty*. Handbook of Mathematical Economics vol.IV. Elsevier Science Publishers B.V., 1991.
- [5] Karni E. and Vierø M.-L. Reverse bayesianism: A choice-based theory of growing awareness. *American Economic Review*, 103:2790–2810, 2013.
- [6] Karni E. and Vierø M.-L. Awareness of unawareness: a theory of decision making in the face of ignorance. Queen’s Economics Department working paper, January 2014.
- [7] Mengel F. Tsakas E. and Vostroknutov A. An experiment on how past experience of uncertainty affects risk preferences. work in progress, October 2012.
- [8] Ellsberg. Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, 75(4):643–669, 1961.
- [9] Anscombe F.J. and Aumann R.J. A definition of subjective probability. *Annals of Mathematical Statistics*, 43:199–205, 1963.
- [10] Li J. A note on unawareness and zero probability. PIER working paper, January 2008.
- [11] Von Neumann J? and Morgenstern O. *Theory of game and economic behavior*. Princeton University Press, Princeton, NJ, 1953.
- [12] Halpern J.Y. Alternative semantics for unawareness. *Games and Economic Behaviour*, 37:321–339, 2001.
- [13] Halpern J.Y. and Rêgo L.C. Reasoning about knowledge of unawareness. *Games and Economic Behavior*, 67:503–525, 2009.
- [14] Savage L.A. *The foundations of statistics*. Dover Publications, Inc., New York, 1954.
- [15] Dekel E. Lipman B.L. and A. Rustichini. Recent developments in modeling unforeseen contingencies. *European Economic Review*, 42:523–542, 1998.
- [16] Heifetz A. Meier M. and Schipper B.C. Interactive unawareness. *Journal of Economic Theory*, 130:78–94, 2006.
- [17] Heifetz A. Meier M. and Schipper B.C. Dynamic unawareness and rationalizable behavior. *Games and Economic Behaviour*, 81:50–68, 2013.

- [18] Heifetz A. Meier M. and Schipper B.C. Unawareness, beliefs, and speculative trade. *Games and Economic Behaviour*, 77:100–121, 2013.
- [19] Kreps D. M. “Static choice in the presence of unforeseen contingencies” in *Economic Analysis of Markets and Games*. Dasgupta P. Gale D. Hart O and Maskin E., Cambridge, MA; MIT Press, 1992.
- [20] Walker O. and Dietz S. A representation result for choice under conscious unawareness. unpublished manuscript, November 2012.
- [21] Ghirardato P. Coping with ignorance: unforeseen contingencies and non-additive uncertainty. *Economic Theory*, 17:247–276, 2001.
- [22] Johnson-Laird P.N. *Mental models*. Harvard University Press, 1983.
- [23] Fagin R. and Halpern J.Y. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.
- [24] Goldstone R.L. and Kersten A. “Concepts and categorization” in *Comprehensive handbook of psychology, Volume 4: Experimental psychology*. Healy A.F. and Proctor R.W., 2003.
- [25] Modica S. and Rustichini A. Awareness and partitional information structures. *Theory and Decision*, 37:107–124, 1994.
- [26] Modica S. and Rustichini A. Unawareness and partitional information structures. *Games and Economic Behaviour*, 27:265–124, 1999.