



## **Philosophia Scientiæ**

Travaux d'histoire et de philosophie des sciences

19-1 | 2015

Logic and Philosophy of Science in Nancy (II)

---

# Insaisissable Belle au bois dormant

Laurent Delabre et Léo Gerville-Réache

---



### **Édition électronique**

URL : <http://journals.openedition.org/philosophiascientiae/1064>

DOI : [10.4000/philosophiascientiae.1064](https://doi.org/10.4000/philosophiascientiae.1064)

ISSN : 1775-4283

### **Éditeur**

Éditions Kimé

### **Édition imprimée**

Date de publication : 1 mars 2015

Pagination : 251-269

ISBN : 978-2-84174-707-8

ISSN : 1281-2463

### **Référence électronique**

Laurent Delabre et Léo Gerville-Réache, « Insaisissable Belle au bois dormant », *Philosophia Scientiæ* [En ligne], 19-1 | 2015, mis en ligne le 23 avril 2015, consulté le 02 novembre 2020. URL : <http://journals.openedition.org/philosophiascientiae/1064> ; DOI : <https://doi.org/10.4000/philosophiascientiae.1064>

---

Tous droits réservés

# Insaisissable Belle au bois dormant

*Laurent Delabre*

Université Paris 1 Panthéon-Sorbonne  
UMR 8590 – IHPST (France)

*Léo Gerville-Réache*

Université de Bordeaux  
UMR 5251 – Institut de Mathématiques de Bordeaux  
(France)

**Résumé :** Dans cet article sont commentées trois tentatives de résolution de la Belle au bois dormant, la célèbre énigme d'auto-localisation qui engendre un paradoxe probabiliste troublant : l'arbitrage fréquentiste propose une lecture ontologique des probabilités lorsque l'approche bayésienne atteint une limite ; l'analogie avec le Monty Hall s'efforce de penser de nouvelles règles de révision doxastique communes aux deux problèmes ; enfin, la leçon des compagnons encourage le partage et l'harmonisation d'estimations probabilistes entre agents rationnels. Chacun des trois arguments est rappelé brièvement, puis certaines parties sont examinées plus attentivement, critiquées ou au contraire appuyées par des idées neuves, telles que la variante des Quatre Belles.

**Abstract:** In this paper, three pathways are discussed so as to resolve the famous Sleeping Beauty problem, a self-location puzzle that causes a troubling probabilistic paradox: the frequentist arbitration proposes an ontological reading of probabilities when the Bayesian approach reaches a limit; the analogy with the Monty Hall game tries to conceive new doxastic revision rules that apply to both problems; finally, the lesson of the companions encourages the sharing and harmonization of probabilistic estimates between rational agents. All three arguments are briefly recalled, then some portions are scrutinized, criticized or on the contrary supported by new ideas like the Four Beauties variation.

# 1 Introduction

Le paradoxe probabiliste de la Belle au bois dormant n'est pas une récréation mathématique. C'est une véritable boîte de Pandore qui nous ramène toujours aux zones d'ombre de chaque théorie invoquée en vue de sa résolution. En ajoutant une drogue à effet amnésique au protocole anodin d'une expérience de pensée, nous ranimons soudain des controverses très profondes liées aux concepts de probabilité, de rationalité, d'identité, de connaissance, de temps ; nous revisitons des principes, moins solides qu'ils en avaient l'air, et participons à une aventure épistémologique captivante.

Rappelons le problème. Ce dimanche soir, des chercheurs vont endormir pendant quelques jours la Belle au bois dormant ; puis ils lanceront une pièce de monnaie équilibrée. Selon le résultat, ils interrompront brièvement le sommeil de la Belle soit une, soit deux fois : si face, un réveil le lundi ; si pile, un réveil le lundi, un autre le mardi. Chaque fois, ils auront un entretien avec elle, puis la rendormiront à l'aide d'une drogue qui lui fera complètement oublier ce réveil. Voici que la Belle, qui connaît tout ce protocole, se réveille au cours de l'expérience, incapable de savoir si c'est lundi ou mardi. On lui demande alors : « À quel degré devez-vous croire que la pièce est tombée sur face ? »

Complétons l'énoncé : si on est lundi, dès que la Belle a répondu à cette question principale, on lui annonce qu'on est lundi et on lui repose la question.

Selon les « tiéristes<sup>1</sup> », qui semblent majoritaires, une Belle parfaitement rationnelle doit répondre  $1/3$  à la question principale, puis  $1/2$  à la question subsidiaire du lundi. Peu nombreux, les demistes répondent  $1/2$  à la première question,  $2/3$  à la seconde. Ce demisme traditionnel est concurrencé aujourd'hui par le « double-demisme », ou demisme de la conservation de la croyance, qui répond  $1/2$  aux deux questions. Dans tous les camps se trouvent des philosophes attachés au bayésianisme, mais aussi des insatisfaits. Aucun consensus ne s'est dégagé, toutes ces contrariétés révèlent un paradoxe redoutable. Quelques travaux, qui ne seront pas présentés ici, prétendent le résoudre par une désambiguïsation de son énoncé ou l'acceptation de plusieurs réponses<sup>2</sup>.

Nous allons éclairer, questionner, prolonger et critiquer trois des nombreux arguments de la littérature : l'arbitrage fréquentiste, l'analogie avec le problème de Monty Hall, et la leçon des « compagnons » de la Belle. Ces pistes de réflexion donneront un petit aperçu du vaste et riche débat en cours et introduiront aussi des idées inédites.

---

1. « Tierciste » serait plus juste, mais le mot « tiériste » l'emporte dans les rares articles en français et dans les correspondances des chercheurs.

2. [Delabre 2008] propose une vue d'ensemble du débat plus détaillée.

## 2 L'arbitrage fréquentiste

Ce premier argument, déjà évoqué par Adam Elga, le philosophe qui a popularisé la Belle au bois dormant [Elga 2000], est notamment complété et défendu par le mathématicien Jean-Paul Delahaye [Delahaye 2003]. Il s'agit d'un raisonnement tiériste à l'origine, même si plus tard il inspire des avis moins tranchés.

Reformulons d'abord l'argument demiste bayésien de David Lewis, au centre de la critique tiériste : la Belle, qui estimait dimanche à  $1/2$  la probabilité de face (ontologique puis épistémique), n'apprend rien de pertinent pour l'issue pile/face à son réveil. Elle est réveillée, l'expérience est en cours, rien de surprenant. Elle sait que « ce jour est lundi ou ce jour est mardi », mais ce n'est qu'une évidence indexicale impuissante. L'inertie doxastique apparaît donc nécessaire et rationnelle : la Belle doit encore croire au degré  $1/2$  en l'obtention de face. En revanche, l'information « on est lundi » est pertinente, la Belle qui la reçoit doit augmenter la probabilité de face jusqu'à  $2/3$ , suivant la loi de conditionalisation bien connue [Lewis 2001].

Pour Delahaye, ces résultats, obtenus au mépris du décompte le plus élémentaire des réveils en tant qu'événements réalisés lors de très nombreuses répétitions de l'expérience, sont faux et montrent les limites des révisions bayésiennes pourtant salutaires en d'autres circonstances. La Belle qui se sait engagée dans une série d'expériences sait aussi qu'elle existe en double dans les expériences-pile, elle sait par exemple qu'il est dans son intérêt de parier sur pile à chaque réveil. Son espace privé de probabilité se déforme en partie, s'adapte aux probabilités fréquentielles. Les réponses tiéristes s'imposent alors à elle. Lorsqu'on lui annonce qu'on est lundi, un glissement bayésien a bien lieu mais ne fait qu'annuler cette anamorphose dite « effet de loupe » qui avait rendu pile deux fois plus probable que face.

Ce retour au fréquentisme est louable mais pour le moins téméraire, puisqu'il prétend démêler un problème fortement noué de localisation d'un sujet (potentiellement) amnésique dans des mondes possibles *et dans le temps* en répétant une expérience qui se voulait unique. Mais ce qui nous intéresse est un point précis. Il est un fait que si l'expérience de la Belle est répétée un grand nombre de fois, la fréquence de réveils-face sera deux fois moindre que la fréquence de réveils-pile. Mais dans quelle mesure ces fréquences sont-elles des probabilités et constituent-elles un argument en faveur de la position tiériste ?

La répétition du scénario sur une infinité de semaines produit deux séries de résultats :

- L'une est une suite d'expériences hebdomadaires et donc de tirages équiprobables et indépendants d'une pièce. La série obtenue est par exemple :

$$P - F - P - P - F - F - F - P - F - P - P - P - F - F...$$

Elle contiendra, sur une infinité de semaines, une fréquence égale de piles et de faces ( $1/2 - 1/2$ ).

- L'autre est une succession de réveils, conséquence directe, déterminée par le protocole, de la série d'expériences précédente. En partant de notre exemple, on obtient la série de réveils suivante :

$$PP-F-PP-PP-F-F-F-PP-F-PP-PP-PP-F-F..$$

Elle contiendra sur une infinité de semaines une fréquence deux fois moindre de réveils-face que de réveils-pile ( $1/3 - 2/3$ ).

Une lecture attentive des deux séries dévoile deux différences importantes. Si les fréquences calculées sur la première série le sont sur une suite de réalisations *indépendantes* dont le nombre est égal au nombre de semaines, les fréquences calculées sur la série de réveils le sont sur une suite de réalisations *dépendantes* dont le nombre total est *aléatoire*. Si, pour la première série, on est dans le cadre traditionnel de la loi des grands nombres et de la convergence de la fréquence vers la probabilité, la seconde série pose une question : dans quelle mesure la dépendance des réveils et le nombre aléatoire de réveils permettent-ils d'affirmer que la fréquence converge vers une probabilité ?

L'interprétation du calcul de la fréquence de réveils-face et/ou de réveils-pile n'est pas claire. Rappelons que la définition de la probabilité comme fréquence limite d'un événement trouve son origine dans la loi des grands nombres, qui concerne une somme de variables aléatoires (de Bernoulli par exemple) indépendantes et identiquement distribuées divisée par le nombre déterministe de ces variables. La probabilité est donc définie comme un quotient vérifiant certaines conditions. Notre problème est de comprendre, dans le cas de la Belle, si la fréquence proposée entre dans le cadre d'une définition (éventuellement élargie) de la probabilité au sens fréquentiste.

Initialement, on définit la probabilité comme la limite suivante :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = p.$$

où les  $X_i$  sont des variables aléatoires de Bernoulli indépendantes et identiquement distribuées. Ici,  $X_i = 1$  si la pièce tombe sur face la semaine  $i$ , 0 sinon. Pour la Belle, aux réveils, on s'intéresse à :

$$\lim_{n \rightarrow \infty} \frac{1}{\sum_{i=1}^n X_i + 2 \sum_{i=1}^n (1 - X_i)} \sum_{i=1}^n X_i = \lim_{n \rightarrow \infty} \frac{1}{2n - \sum_{i=1}^n X_i} \sum_{i=1}^n X_i = \frac{p}{2-p}.$$

Cette limite vaut assurément  $1/3$  si  $p = 1/2$ . Mais la question est de savoir si :

$$\lim_{n \rightarrow \infty} \frac{1}{2n - \sum_{i=1}^n X_i} \sum_{i=1}^n X_i$$

est une probabilité, et en quel sens. Pour résumer, quand on divise une somme de variables aléatoires de Bernoulli indépendantes et identiquement distribuées par  $n$ , cette statistique converge vers une probabilité qui est la probabilité que

chaque variable aléatoire prenne la valeur 1. Mais quand on divise cette même somme par le nombre aléatoire de réveils  $2n - \sum_{i=1}^n X_i$ , qu'obtient-on ?

*In fine*, d'un côté, à chaque expérience  $P(\text{Face}) = 1/2$ ; de l'autre côté, la fréquence de réveils-face sur une infinité de réveils est  $P(\text{Face})/(2 - P(\text{Face})) = 1/3$ . Reste à savoir si  $P(\text{Face})/(2 - P(\text{Face}))$  est une probabilité. Si oui, est-ce une probabilité au sens fréquentiste ? Si oui, est-ce la « bonne » réponse à la question posée à la Belle ?

La rivalité des interprétations ontologique et épistémique ne doit pas faire oublier les multiples conceptions philosophiques de la probabilité. Même au sein du fréquentisme, plusieurs axiomatiques existent. Selon von Mises [Bienvenu 2007, 145–146], une probabilité fréquentielle doit vérifier que :

- la fréquence relative d'un ensemble d'éléments caractérisés dans une suite de référence admet une limite (axiome d'existence de limites) ;
- cette limite est invariable dans toute suite partielle issue de la suite de référence au moyen d'une méthode de sélection des éléments de la suite partielle par « choix de position » (axiome d'irrégularité).

Manifestement, si l'axiome d'existence d'une limite est vérifié sur la série des réveils répétés, l'axiome d'irrégularité ne l'est pas. Aussi,  $1/3$  n'est pas une probabilité-fréquence au sens de von Mises. Un « sens » pas toujours retenu. Dans sa proposition de définition, Reichenbach ne fait aucunement appel à l'axiome d'irrégularité [Bienvenu 2007, 145–146]. La probabilité est pour lui fondamentalement associée à des paris : la mise doit nécessairement découler du constat d'une fréquence sur une section finie d'une suite d'événements supposée prolongeable à l'infini. Ainsi, au sens de Reichenbach,  $1/3$  est bien une probabilité fréquentielle... mais de quel événement ?

La réponse se trouve peut-être dans la position de Weatherford à propos du principe d'indifférence :

Les difficultés empiriques qui ont discrédité le principe d'indifférence sont entièrement dues au fait que beaucoup ont confondu la probabilité de choisir un succès avec la probabilité d'être un succès. [De Scheemaekere 2011, 64]

Weatherford pointe, ici pour la question du principe d'indifférence, les différences de conception entre von Mises et Reichenbach. Pour ce dernier, la probabilité de choisir un succès et la probabilité d'être un succès ne font qu'un. Pour von Mises, une différence existe et c'est peut-être cela que traduit l'axiome d'irrégularité.

Ainsi, dans le problème de la Belle, si l'on reprend la série de réveils et si l'on tire un réveil avec équiprobabilité, la probabilité de tirer un réveil-face est bien  $1/3$ . Cela pose pourtant deux derniers problèmes :

- Si l'on considère une série infinie de réveils, une telle procédure de sélection équiprobable est impossible. En effet, il n'existe pas de loi de probabilité uniforme sur un ensemble infini d'objets. On peut néanmoins

lever cette difficulté en considérant un grand nombre de réveils plutôt qu'une infinité.

- Dans quelle mesure, à son réveil, la Belle doit se référer à la probabilité de « choisir un succès » et pas à la probabilité « d'être un succès » ?

Cette question nécessiterait encore bien des développements. L'important est de voir que, quand il prétend répondre à la question principale de l'entretien, l'argument fréquentiste tiériste n'a pas de fondement sous l'axiomatique de von Mises, que la fréquence de réveils-face est une probabilité au sens de Reichenbach, mais qu'une Belle inspirée par Reichenbach n'est pas encore tirée d'affaire. Le demisme traditionnel a donc un espoir. Pas pour longtemps.

Transformons la série de réveils que nous avons précédemment donnée en éliminant ceux qui ne satisfont pas la condition « a lieu lundi » :

$$P_P - F - P_P - P_P - F - F - F - P_P - F - P_P - P_P - P_P - F - F$$

Naturellement, cette série de réveils-lundi est celle des face/pile du dimanche soir :

$$P - F - P - P - F - F - F - P - F - P - P - P - F - F$$

Cela signifie que la Belle qui sait qu'on est lundi et croit en face au degré  $1/2$  peut sereinement affirmer que ce degré correspond à une probabilité fréquentielle au sens de von Mises et de Reichenbach. Ce n'est pas un hasard si très peu de philosophes après Lewis argumentent en faveur de  $P(\text{Face}|\text{Lundi}) = 2/3$  et s'ils sont peu suivis par la communauté : certes, les approches objectivistes ne semblent pas pouvoir départager demistes et tiéristes en ce qui concerne la question principale, mais cette réponse  $2/3$  à la question subsidiaire paraît si étrangère à l'analyse fréquentiste qu'elle devient également suspecte aux yeux des bayésiens.

D'autres arguments discréditent la réponse  $2/3$ . Reprenons et prolongeons le plus fameux d'entre eux, exposé pour la première fois par Elga. Lorsque la Belle apprend qu'on est lundi, elle ne fait que récupérer un savoir que l'expérience et son protocole lui ont ôté et doit simplement croire en l'obtention de face à un degré égal à la propension à tomber sur face d'une pièce équilibrée. En effet, la voici réveillée lundi, événement qui devait assurément arriver ; elle sait qu'elle n'est pas encore amnésique, qu'aucun accident cognitif ne la trompe ; la pensée que demain mardi peut avoir lieu un réveil ne peut pas la gêner. Plus rien d'important ne la différencie d'un sujet extérieur au protocole qui doit obéir à un principe d'alignement des probabilités ontologique et épistémique. Remarquons que le déroulement de l'expérience ne dépend pas du résultat du lancer de la pièce tant que mardi matin n'est pas venu. Imaginons alors un scénario similaire où la pièce n'est lancée que dans la nuit de lundi à mardi afin de déterminer si un autre réveil doit avoir lieu avant que l'expérience ne prenne fin : cette légère variation des règles, connue par la Belle, ne peut pas changer ses réponses et en particulier, quand elle est réveillée lundi et qu'elle en est informée, elle doit clairement croire que la pièce qui sera lancée dans

quelques heures a une chance sur deux de tomber sur face. On peut même l'inviter à lancer elle-même la pièce, puisque de toute façon la drogue va lui faire oublier la journée du lundi : une fois que la pièce est dans sa main, la Belle peut-elle raisonnablement estimer à  $2/3$  la probabilité d'obtenir face ? Nous pensons que non.

Le prolifique demiste bayésien Darren Bradley [Bradley 2011] a développé une argumentation qui, selon lui, clarifie la critique trop hâtive que Lewis opposait à cet argument solide tel qu'il était présenté par Elga. Quand on annonce à la Belle qu'on est lundi, celle-ci acquerrait en fait une *inadmissible evidence*, c'est-à-dire une information qui peut justifier que ses degrés de croyance s'écartent des chances objectives. Un exemple plus clair est celui d'une boule de cristal infaillible qui prédit qu'une pièce donnée va tomber sur pile. Tout agent ayant cette prédiction en tête est certain que la pièce va tomber sur pile : une telle information sur le futur disqualifie les probabilités habituellement associées à l'objet « pièce de monnaie équilibrée ». Autre exemple extrême : la Belle qui apprend qu'on est mardi devient tout simplement certaine que la pièce est tombée sur pile, donc ne fait plus coïncider probabilités épistémique et ontologique. Puisque « on est mardi » est une information inadmissible, alors l'information concurrente « on est lundi » l'est aussi et justifie que la Belle puisse se désolidariser de la probabilité  $1/2$  en adoptant à la place  $2/3$ . Bradley explique encore que l'annonce « on est lundi » ressemble à une prédiction : jusqu'alors incapable de se repérer dans le temps, un agent qui soudain apprend la date ou l'heure apprend en même temps, en quelque sorte, des informations sur le futur, puisqu'il découvre quelles sont, parmi toutes les positions temporelles possibles, ou plutôt parmi toutes les parties temporelles de l'agent (tous les *individuals-at-times* au sens lewisien), celle qui est actuelle, celles qui appartiennent au passé et celles qui appartiennent à l'avenir.

Bradley fait de notre pouvoir de localisation dans le temps une prescience insoupçonnée. Mais sa critique échoue. Le protocole de l'expérience de la Belle interdit les réveils-face-mardi, donc la série des réveils-mardi produite par répétition de l'expérience est tout entière une succession de réveils-pile. Cela signifie que les chances objectives d'avoir un réveil-face-mardi sont nulles. Apprendre qu'on est mardi rend la Belle certaine qu'elle est dans un réveil-pile, justement parce qu'elle calque ses degrés de croyance sur des chances objectives. Bradley se focalise sur la pièce de monnaie comme si elle représentait tout le dispositif aléatoire : c'est une première erreur. Il oublie que les événements produits sont des réveils et, à tort, ne considère comme « chances » que la propension de la pièce à tomber sur un côté plutôt que sur l'autre. Sa deuxième erreur consiste à croire que les *inadmissible evidences* se rencontrent facilement dans la vie d'un agent rationnel. En réalité, le demiste dénonce des écarts entre probabilités épistémiques et ontologiques dont il a lui-même forcé l'existence, soit en occultant des éléments du dispositif aléatoire, soit en brochant une histoire où le surnaturel est roi (la boule de cristal).

$P(\text{Face}|\text{Lundi}) = 1/2$  : voilà la seule estimation acceptable. Certaines analyses, certaines remarques du demisme traditionnel sont sans doute éclair-

rantes, mais l'estimation  $2/3$  (sachant lundi) est à rejeter. Le tiérisme, qu'il soit fréquentiste ou bayésien, n'est pourtant pas forcément vainqueur...

### 3 L'analogie avec le problème de Monty Hall

Quelques auteurs ont abordé la question des similitudes entre le Monty Hall et la Belle au bois dormant. Halpern [Halpern 2004] et Jenkins [Jenkins 2005] considèrent que les tiéristes font la même erreur que les défenseurs du  $1/2 - 1/2$  pour le Monty Hall. Weatherson ne partage pas ces analyses et estime que la trop grande différence entre les deux problèmes ne permet pas de trancher [Weatherson 2011]. Il est vrai que le Monty Hall n'est pas un problème d'auto-localisation et qu'il ne fait plus peur à nombre d'adeptes des mathématiques récréatives, mais il fait partie, avec la Belle, de ces paradoxes qui entraînent les philosophes insatisfaits par l'inférence bayésienne sur des voies polémiques, à la recherche de lois de dynamique doxastique originales.

Dans le jeu de Monty Hall interviennent un présentateur et un candidat placé devant trois portes fermées (A, B et C) ; derrière l'une d'elles se trouve une voiture et derrière chacune des deux autres une chèvre. Le candidat doit d'abord désigner une porte (par exemple, A). Puis le présentateur ouvre une porte qu'il sait n'être ni celle qui cache la voiture, ni celle choisie par le candidat (ce sera, mettons, C). Celui-ci a alors le droit ou bien d'ouvrir la porte qu'il a choisie initialement, ou bien d'ouvrir la troisième porte (B) : il gagnera le lot correspondant. Que doit-il faire et quelles sont ses chances de gagner la voiture ? Deux réponses rapides et intuitives s'opposent :

- Après ouverture de la porte par le présentateur, il reste deux portes qui ont autant de chances l'une que l'autre de cacher la voiture. On aurait donc  $1/2$  de chances de gagner sans changer,  $1/2$  de chances en changeant.
- Mais si l'on ne change pas de porte, on gagne si on avait fait le bon choix au départ ; or ce choix avait une chance sur trois d'être bon. Il y aurait donc  $1/3$  de chances de gagner sans changer contre  $2/3$  en changeant.

La réponse  $2/3$  en changeant fait aujourd'hui consensus. La réponse  $1/2$  correspond aux chances de gagner en sélectionnant au hasard une des deux portes restantes et pas en changeant de porte rationnellement. Les défenseurs du  $1/2$  reprochaient aux avocats du  $2/3$  de considérer que l'ouverture d'une mauvaise porte laisse inchangée la probabilité que la porte initialement choisie soit la bonne. Il est effectivement légitime de se demander pourquoi l'ouverture d'une porte ne modifie la probabilité que d'une des deux autres. En particulier, il est évident que si deux portes étaient ouvertes, la probabilité de gagner deviendrait 0 ou 1.

Comprendre que le candidat est certain, juste avant l'intervention du présentateur, qu'une chèvre va être « découverte » derrière une des portes B et C (mais pas A) est essentiel pour faire disparaître le paradoxe. Le dimanche soir, la Belle est elle aussi dans un tel état de certitude, sûre d'être réveillée durant l'expérience mais aussi d'être réveillée lundi. Un réveil mardi constituerait la seule surprise. Ce qui est commun aux problèmes de Monty Hall et de la Belle, c'est d'abord cette situation d'un agent qui sait en partie ce qui va se passer et qui refuse de croire que la réalisation des événements prévus puisse d'une quelconque façon le renseigner sur l'état d'un objet donné (l'orientation d'une pièce de monnaie ou l'éventuelle présence d'un gain derrière une porte choisie).

Dans son analyse du problème de la Belle, Cozic accepte la partie du raisonnement demiste lewisien qui conclut la réponse  $1/2$  à la question principale, mais aussi le raisonnement du tiériste Elga qui conclut la réponse  $1/2$  à la question subsidiaire du lundi [Cozic 2007]. Chacune de ces réponses est intuitive, et pourtant leur égalité rebute les bayésiens persuadés que l'annonce « on est lundi » confirme face. Cozic étudie alors les concepts d'*imaging* et de mise à jour, qu'il distingue de la conditionalisation et de la révision. Il montre en particulier que l'*imaging*, règle de dynamique des croyances esquissée sans optimisme par Lewis en 1976, pourrait, en invitant la Belle à conserver lundi la probabilité de face, réconcilier demistes et tiéristes autour d'un compromis double-demiste. Pour autant, l'*imaging*, que Cozic lie étroitement à la mise à jour (variation de l'intensité des croyances d'un sujet qui apprend que son environnement a changé), ne correspondrait pas précisément à la règle que la Belle devrait suivre.

L'habitude de la conditionalisation rend l'*imaging* suspect. Pourtant, si l'on regarde le problème de Monty Hall en groupant les portes B et C, les seules que le présentateur est susceptible d'ouvrir, la solution  $2/3 - 1/3$  est celle que l'on obtiendrait par application de l'*imaging* ; la conditionalisation classique y parvient aussi mais en suscitant l'étonnement et en contribuant à l'effet paradoxal. Certes, nous sommes dans un contexte de révision des croyances, pas de mise à jour, mais l'*imaging* ne pourrait-il pas être aussi un outil de révision ?

Les schémas de la figure 1 montrent que la réponse  $1/2 - 1/2$  résulte d'une conditionalisation modifiant jusqu'aux branches principales de l'arbre (nous l'appellerons désormais conditionalisation profonde ou C.P.), alors que la réponse  $2/3 - 1/3$  correspond à un *imaging* qui, par définition, n'affecte que les branches secondaires sœurs.

Les positions tiériste, demiste et double-demiste peuvent être présentées sous forme d'arbres probablisés similaires à ceux du Monty Hall. Dans un premier temps, le protocole peut être décrit par la figure 2. La Belle sait qu'une pièce est lancée le dimanche soir, que si face sort, elle sera réveillée lundi (RL) avec probabilité 1, que si pile sort, elle sera réveillée lundi avec probabilité 1 et mardi (RM) avec probabilité 1. Aussi, à chaque expérience, la probabilité d'être réveillée est 1, la probabilité d'être réveillée lundi est 1 et

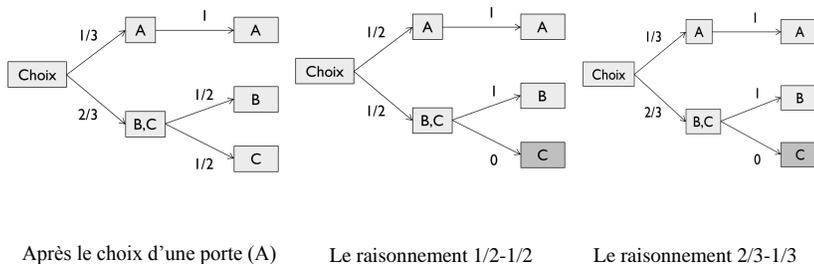


FIGURE 1 – Les schémas du Monty Hall

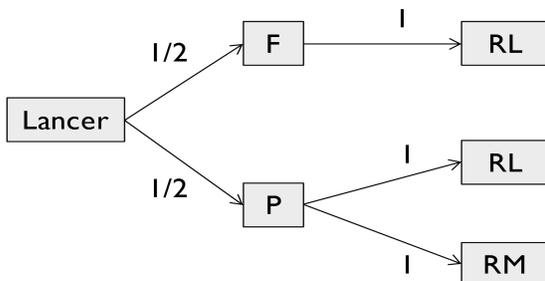


FIGURE 2 – Le protocole de la Belle

la probabilité d'être réveillée mardi est  $1/2$ . Au réveil, la Belle doit cependant adapter ces valeurs car elle ne peut croire qu'au degré 1 qu'elle est réveillée lundi ou mardi. Le problème est de déterminer la manière correcte de ventiler au réveil les probabilités du dimanche soir, puis de répartir ces probabilités après l'annonce « on est lundi ».

- Un tiériste, conditionnellement au réveil, réajuste les probabilités face-pile et les probabilités RL et RM (si pile). En apprenant que c'est lundi, il modifie encore les probabilités face-pile et les probabilités RL et RM (si pile) [cf. Fig. 3]. Au réveil, une Belle tiériste modifie ses croyances par conditionalisation profonde. Dès qu'elle sait que ce réveil a lieu lundi, elle utilise également la C.P.
- Un demiste, conditionnellement au réveil, ne modifie pas les probabilités face-pile mais uniquement les probabilités RL et RM (si pile). En apprenant que c'est lundi, il modifie les probabilités face-pile et à nouveau les probabilités RL et RM (si pile) [cf. Fig. 4]. Au réveil, une Belle demiste

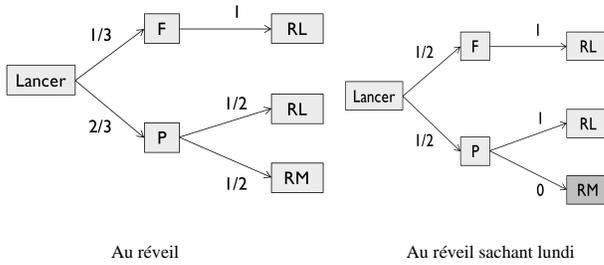


FIGURE 3 – La Belle tiériste

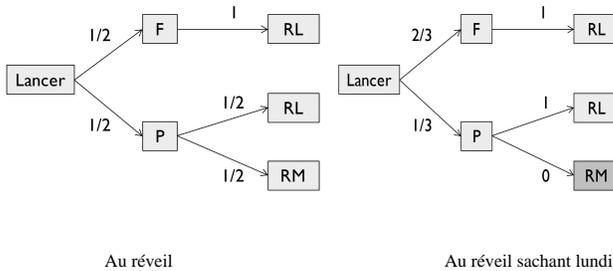


FIGURE 4 – La Belle demiste

modifie ses croyances par imaging . Puis, en apprenant que c'est lundi, elle utilise la C.P.

- Un double-demiste, conditionnellement au réveil, ne modifie pas les probabilités face-pile mais uniquement les probabilités RL et RM (si pile). En apprenant que c'est lundi, il ne modifie toujours pas les probabilités face-pile mais modifie à nouveau les probabilités RL et RM (si pile) [cf. Fig. 5]. Au réveil, une Belle double-demiste modifie ses croyances par imaging , règle qu'elle suit à nouveau dès qu'elle sait qu'on est lundi.

On peut donc considérer que le tiérisme applique une double C.P., le double-demisme un double imaging ; le demisme traditionnel, lunatique et décidément suspect, applique un imaging puis une C.P. L'imaging réunit les schémas résultant des raisonnements double-demistes et le schéma des raisonnements qui résolvent le Monty Hall, quand bien même certains de ces raisonnements, à l'origine, ne reconnaîtraient que la conditionalisation clas-

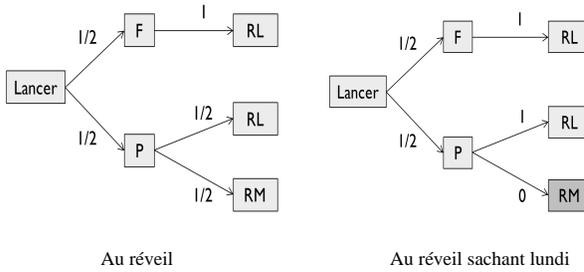


FIGURE 5 – La Belle double-demiste

sique. Si l'on accepte les similitudes entre les deux paradoxes, si l'on tient la C.P. pour une règle inadaptée à leur situation commune, le double-demisme prend le dessus sur le tiérisme... qui n'a cependant pas dit son dernier mot.

## 4 La leçon des compagnons de la Belle

La variante suivante est au cœur d'un argument tiériste apprécié [Neal 2006, 15–17], [Stalnaker 2008, 63]. La Belle s'endort dimanche soir avec son Prince. Elle est réveillée, en compagnie du Prince avec lequel elle peut alors s'entretenir, selon le protocole de l'expérience originale (lundi si face, lundi et mardi si pile). Comme la Belle, le Prince est endormi à l'aide de la drogue à effet amnésique ; il ignore de quel côté est tombée la pièce et ne peut dater le jour de son réveil que si on lui donne l'information ; il connaît le protocole entier. La différence, c'est qu'il est réveillé lundi et mardi *quel que soit le résultat du lancer de la pièce* ; il peut donc arriver qu'il soit réveillé alors que la Belle reste endormie. Voici que le Prince se réveille et comprend que l'expérience est en cours. Respectueux du principe d'indifférence, il attribue la probabilité épistémique *a priori*  $1/4$  à chacune des quatre hypothèses centrées, exclusives et conjointement exhaustives :

**T1** : « Aujourd'hui est lundi et la pièce est tombée sur pile »

**T2** : « Aujourd'hui est mardi et la pièce est tombée sur pile »

**H1** : « Aujourd'hui est lundi et la pièce est tombée sur face »

**H2** : « Aujourd'hui est mardi et la pièce est tombée sur face ».

Il découvre alors que la Belle est aussi réveillée. Bien qu'il sût que cette situation arriverait durant l'expérience, il reçoit une nouvelle information indexicale :  $\neg H_2$ , autrement dit  $T_1 \vee T_2 \vee H_1$ . La probabilité de  $H_2$  est maintenant

annulée et, par conditionnalisation équivalente à un partage de l'ancienne probabilité  $1/4$  de  $H_2$  entre les trois hypothèses restantes, le Prince doit maintenant attribuer à chacune des trois la probabilité  $1/3$ , ne voyant pas pourquoi en favoriser une au détriment des autres. Par conséquent, il doit à présent croire au degré  $1/3$  que la pièce est tombée sur face.

Les demistes et *certain*s double-demistes acceptent cette partie du raisonnement tiériste, mais ils soulignent, à la manière de leur critique d'une variante similaire décrite par le tiériste Dorr en 2002, que la Belle, contrairement au Prince, ne peut jamais avoir l'occasion d'apprendre que celui des deux sujets rationnels qui dans cette expérience peut ne pas être réveillé mardi, à savoir elle-même, est ou n'est pas réveillé, puisqu'un tel apprentissage nécessite la conscience, l'état de veille, donc l'événement dont la réalisation ou la non-réalisation, justement, est à apprendre. En d'autres termes, selon le demisme, comprendre que la construction mentale « j'apprends que je suis réveillé » est aussi absurde que « j'apprends que je dors », c'est comprendre que la dynamique des croyances de la Belle engagée dans l'expérience est différente de celle des croyances du Prince.

Poursuivons le raisonnement tiériste. La Belle et son Prince peuvent à présent discuter et surtout échanger leurs estimations de la probabilité de face. Si la Belle est tiériste, les estimations se rejoignent, ce qui ne produit chez elle aucun questionnement particulier. Si elle est demiste, les estimations divergent : ce désaccord de deux agents rationnels n'est pas rare, mais il serait bien étrange qu'ils ne parviennent pas à faire concorder leurs estimations après avoir mis en commun leurs connaissances. Et si la demoiselle hésite entre  $1/2$  et  $1/3$ ? Ne doit-elle pas repenser au raisonnement tiériste du Prince, ne peut-elle pas adopter sa conclusion? Si nous admettons que la Belle doit calquer son degré de croyance sur celui de son compagnon, nous pouvons encore objecter que la présence de celui-ci modifie de façon significative l'expérience originale. Alors, par variations successives, plaçons-nous progressivement dans les conditions de l'expérience originale, d'abord en supposant que le couple est séparé par un rideau, puis que le Prince se situe dans une chambre isolée sans aucun moyen de dialoguer avec sa bien-aimée; finalement, examinons le cas où le compagnon n'est plus du tout là et où la Belle ne peut que l'imaginer et comprendre le raisonnement tiériste qu'il pourrait tenir. À quel moment doit-elle cesser de croire au degré  $1/3$  en l'obtention de face? Jamais, répondront les convaincus.

La variante du Prince pourrait mettre le double-demisme en difficulté : l'idée d'une divergence persistante de deux estimations probabilistes rationnelles portant sur la même croyance au même moment ne semble pas supportable. Pourtant, nous devrions avoir de sérieux doutes. Le Prince serait plus informé, plus savant que la Belle, ses probabilités à lui feraient autorité? Contestable, étant donné l'asymétrie subtile de leurs cheminements mentaux. Toutefois, mettre en relation la Belle avec un compagnon est une bonne idée; l'argument est perfectible s'il repose sur une variante plus ingénieuse que nous allons à présent introduire et analyser.

## La variante des Quatre Belles

Quatre Belles participent à une expérience dont elles connaissent toutes les règles. Le dimanche soir, elles s'endorment. On tire alors au sort une des quatre, équitablement : elle est ainsi désignée pour être réveillée lundi et seulement lundi. Puis on tire au sort une Belle parmi les trois restantes : elle est désignée pour être réveillée mardi et seulement mardi. Les deux sujets restants seront réveillés lundi et mardi. Le lundi, on laisse dormir profondément la Belle désignée pour ne pas être réveillée ce jour-là, on réveille les trois autres, on les réunit pour un entretien, puis on les rendort après leur avoir administré une drogue qui leur fait oublier tout ce qui s'est passé dans la journée. Le mardi, on réveille les trois Belles qui doivent être réveillées ce jour-là, on les réunit pour un entretien, on les rendort avec la même drogue. Les Belles n'ont aucun moyen de savoir si on est lundi ou mardi, ni de connaître le résultat précis des tirages au sort, sauf si on les renseigne dans la conversation (néanmoins, elles savent évidemment que la Belle qui dort a été désignée par le sort). Le mercredi, on les réveille toutes les quatre et l'expérience est terminée.

Dans ce scénario, en quelque sorte, quatre expériences de Belle au bois dormant sont menées simultanément et imbriquées de telle façon que, chaque jour, exactement trois Belles sont réveillées et peuvent facilement communiquer entre elles. Les mondes centrés dont la Belle au bois dormant faisait originellement correspondre ses hypothèses semblent rendus, par impossible, physiquement accessibles et non plus seulement mentalement accessibles, ils sont comme réalisés en mêmes monde, lieu et temps pour permettre à leurs trois ambassadrices, conscientes de l'absence d'une quatrième restée inconsciente, de se rencontrer.

Suivons Aurore, une des Belles engagées dans l'expérience. Elle se réveille avec deux autres Belles, toutes trois sont incapables de se repérer dans le temps au jour près ; la quatrième Belle continue à dormir. À quel degré Aurore doit-elle croire qu'elle est une « désignée », c'est-à-dire que le hasard, au premier ou au second tirage, l'a désignée pour qu'elle ne soit réveillée qu'un jour sur les deux jours de l'expérience, autrement dit juste aujourd'hui ?

Raisonnons d'abord ainsi : deux Belles sur quatre sont désignées, et comme les tirages sont équitables, Aurore avait une chance sur deux, dimanche soir, d'être désignée. Il semble que ce soit différent maintenant, car Aurore comprend comme tout le monde que la Belle qui est en train de dormir est une des deux désignées. Il reste trois Belles, une seule d'entre elles est une désignée, il n'y a aucune raison apparente de ne pas partager les probabilités équitablement : Aurore devrait donc croire au degré  $1/3$  qu'elle est une désignée. Un argument fréquentiste naïf supporte cette probabilité : si l'expérience des Quatre Belles était répétée de très nombreuses fois, un réveil d'Aurore sur trois serait un réveil dans une expérience où Aurore est désignée.

Changeons de cap et soyons bayésiens quelque temps. Intuitionsnons et raisonnons à la manière de Lewis dans la première partie de son analyse en

faveur de la réponse  $1/2$  à la question principale posée à la Belle au bois dormant, partie louée par tous les demistes et double-demistes. Demandons-nous ce qu'Aurore a appris entre dimanche soir et maintenant. Elle est réveillée et l'expérience est en cours? Mais elle savait déjà dimanche qu'elle devait être réveillée durant l'expérience. Aucun bayésien ne voit comment utiliser cette information pour modifier un degré de croyance. Elle a appris que c'est Lucie, et non ses deux autres amies Estelle et Norah, qui est désignée pour rester endormie aujourd'hui? Mais, là encore, elle savait d'avance qu'elle serait réveillée en compagnie de deux autres Belles et que le protocole, qui exige qu'une des deux désignées reste endormie, allait du coup lui révéler, non pas une désignée quelconque, mais une désignée qui ne peut en aucun cas être elle-même, Aurore! Que cette désignée restée endormie soit Lucie, Estelle ou bien Norah, cela ne fait aucune différence et n'a aucune incidence sur les croyances. Par conséquent, Aurore n'apprend rien de pertinent. Ses certitudes du dimanche sont celles d'aujourd'hui, à ceci près, par exemple, qu'elle sait aujourd'hui que Lucie dort, et non Estelle ou Norah, ce qui, répétons-le, est sans importance. Les demistes comme Bradley feraient remarquer qu'une Belle qui reste endormie toute la journée ne peut pas avoir conscience qu'elle est une désignée. Aurore n'a jamais l'occasion de se dire, au moment de son ou ses réveils, qu'elle est certainement une désignée, mais elle a toujours l'occasion de se croire (sans aucune certitude) désignée. Donc elle ne peut pas être surprise, elle ne peut pas apprendre. Elle croyait dimanche au degré  $1/2$  qu'elle serait désignée; maintenant que l'expérience est en cours, elle doit croire au même degré qu'elle est une désignée.

Notre réflexion sur la situation d'Aurore à son réveil nous a permis d'envisager deux raisonnements aux conclusions différentes : le premier ne peut être tenu que par les tieristes (de l'énigme originale de la Belle au bois dormant) et il mène à la probabilité  $1/3$ ; le second ne peut être tenu que par les demistes (et les double-demistes) et il mène à la probabilité  $1/2$ . En outre, gageons que si le second raisonnement, d'une façon ou d'une autre, s'avérait inconsistant, ou si sa conclusion avait des effets embarrassants, le demisme et le double-demisme en pâtiraient parce que, manifestement, ce raisonnement subjectiviste n'est ni plus ni moins que le raisonnement fondateur de tous les courants demistes, mais appliqué à une variation du problème original. Celle-ci, par certains aspects, s'éloigne davantage du problème original que ne le faisait la variante du Prince, mais ce n'est pas grave du tout tant que les demistes se reconnaissent.

Poursuivons l'analyse. Pour Aurore, continuer à croire au degré  $1/2$  qu'elle est une désignée a une conséquence fâcheuse : étant donné qu'une et une seule des trois Belles éveillées est une désignée et qu'Aurore n'a pas de raison de penser qu'Estelle a plus (ou moins) de chances que Norah d'en être une, Aurore doit estimer à seulement  $(1 - 1/2)/2 = 1/4$  la probabilité que chacune de ses deux amies éveillées soit une désignée. En d'autres termes, le demisme oblige un sujet (Aurore) à faire de lui-même une sorte d'élu qui, au milieu des gens réveillés, aurait plus de chances qu'un autre d'être celui qui ne sera réveillé qu'une fois pendant l'expérience! Si ses deux compagnons de veille raisonnent

comme lui, ils vont tirer des conclusions identiques mais centrées sur eux-mêmes. Au final, chacun des trois demistes croit au degré  $1/2$  qu'il est un désigné et réduit ce degré à  $1/4$  pour chacun de ses compagnons. « J'ai plus de chances que les autres d'être un désigné », pense l'un. « J'ai plus de chances que les autres d'être un désigné », pense un autre. Mais ces « Je » ne représentent pas la même personne.

Soyons encore plus démonstratifs : Aurore a maintenant l'occasion de discuter avec Estelle et lui affirme : « Tu as une chance sur quatre d'être une désignée. » Estelle répond : « Moi ? Une chance sur deux plutôt ! » Ainsi, les Belles ne peuvent qu'échanger des estimations contradictoires. Quand une probabilité subjective estimée ne peut pas être « interpersonnalisée », elle ne peut qu'être essentiellement personnelle, si intime et incommunicable qu'elle n'inspire plus confiance.

La grande force de la variante des Quatre Belles, c'est que l'étrangeté des conséquences de la position demiste confine à l'absurdité. L'aventure de la Belle et de son Prince ne provoquait pas un tel rejet des estimations conflictuelles. Au contraire, dans ce scénario, désaccords et contradictions avaient une excuse de taille : le protocole de l'expérience faisait en sorte qu'un agent établisse son propre espace de probabilité, à partir de données différentes des données du compagnon, et il faisait en sorte que le centrage de chaque univers probabilisé le rende plus résistant à la transformation raisonnée généralement observée lorsque deux agents partagent leurs informations pour affiner et harmoniser leurs croyances. La variante des Quatre Belles se distingue. En effet, les trois Belles réveillées ne participent pas seulement à une même expérience, elles sont aussi à égalité face au protocole, leurs états cognitifs et leurs croyances ne devraient pas présenter de différences significatives. Après avoir suivi strictement le même raisonnement demiste, comment peuvent-elles continuer à croire en privé à un certain degré alors que leurs voisines croient à un autre degré ? N'y a-t-il pas une incohérence au sein du (double-)demisme ? Le tiérisme, lui, ne semble pas manquer de cohésion dans la même situation, bien au contraire, et les trois Belles devraient se sentir obligées d'abandonner le demisme si elles veulent croire ensemble au même degré que chacune d'elles est une désignée.

Première objection : le protocole offre aux Belles désignées une journée entière de repos, donc leur réserve un sort différent de celui des Belles non désignées qui sont réveillées une fois de plus. Nous répondons : et alors ? Les Belles ignorent le sort qui leur est réservé, c'est justement là-dessus qu'elles s'interrogent, alors que, dans la variante du Prince, les deux agents connaissent leurs rôles. Le protocole de notre variante fait en sorte que les Belles considèrent au même moment les mêmes possibilités centrées exclusives et conjointement exhaustives : « Je suis une désignée et on est lundi », « Je suis une désignée et on est mardi », « Je ne suis pas une désignée et on est lundi » et « Je ne suis pas une désignée et on est mardi ». Que ces quatre hypothèses soient équiprobables ou pas, à chacune d'elles chaque Belle réveillée doit accorder une crédibilité, la même que celle estimée par ses deux amies éveillées. C'est en ce sens que nous affirmons que les Belles sont à égalité face au protocole.

Deuxième objection : eh bien, justement ! même demistes, les Belles ne sont pas vraiment en conflit, elles s'entendent parfaitement sur la façon de probabiliser les quatre possibilités centrées mentionnées ci-dessus, et c'est seulement lorsque le pronom « je » est remplacé par « Aurore » ou un autre prénom que leur tranquillité prend fin. Nous répondons : si leur tranquillité prend fin, il y a bien conflit derrière une paix apparente, et c'est toute la subtilité de la variante ! Les quatre possibilités ne sont pas essentiellement centrées, dans le sens où, bien que chaque Belle s'interroge sur la partie temporelle d'elle-même qui est actuelle (« Moi-lundi ? Moi-mardi ? »), elle sait très bien qui est « moi », qui est « je ». D'une part, Aurore sait qu'elle est Aurore et il en va de même pour ses amies ; d'autre part, « je » ne représente aucune autre Belle que celle qui prononce le mot. Même si ces agents rationnels égaux devant le protocole évaluent exactement de la même manière la probabilité de « Je suis une désignée », ce « je » ne fait pas référence à la même personne si c'est Estelle et non plus Aurore qui le formule. Seule une Belle sur les trois en état de veille est une désignée : leur désaccord est inévitable si une probabilité autre que  $1/3$  est attribuée à l'hypothèse « Je suis une désignée ». Accord et désaccord semblent essentiellement liés aux estimations ; contrairement au protocole-piège du scénario du Prince, l'innocent protocole des Quatre Belles dirige les agents dans des réflexions centrées identiques qui devraient immédiatement les mettre d'accord, même sur des croyances décentrées. Les agents ne vivent pas dans la pure relativité : « je » est un individu concret identifié, « je » est à absolutiser et à nommer dès que possible, sinon « je » s'estime élu au milieu d'« eux », les autres, refuse un dialogue qu'il considère comme conflictuel, garde des croyances à jamais personnelles. Les (double-)demistes sont-ils des « je » extrémistes ?

## 5 Discussion

Les éléments de réflexion développés dans cet article sont de trois ordres. La question posée par l'arbitrage fréquentiste est essentielle : la possibilité de mesurer par une fréquence observée la probabilité recherchée serait un argument statistique déterminant, et la réconciliation du subjectivisme et de l'objectivisme permettrait de résoudre le paradoxe. Malheureusement, la répétition de l'expérience de la Belle ne produit pas l'effet escompté. La fréquence de réveils-face, clairement égale à  $1/3$ , ne mesure qu'un ersatz de probabilité. La spécificité de la suite de réveils pose en réalité la question également non résolue de la définition du concept de suite aléatoire. Néanmoins, la suite de réveils-lundi réunit les solutions tiériste et double-demiste et condamne pratiquement le demisme classique.

Nous avons ensuite discuté l'analogie avec le Monty Hall, problème dont la solution est tenue pour acquise par la plupart des bayésiens et des fréquentistes. La résolution du Monty Hall par imaging a pourtant une pertinence, et

des similitudes entre les deux paradoxes nous invitent à examiner la Belle également sous l'angle de l'imaging et de la conditionalisation profonde. Ces deux règles de répartition des probabilités caractérisent, l'une le double-demisme, l'autre le tiérisme, ce qui laisserait là encore le demisme classique sur le carreau. Prolonger l'analogie avec le Monty Hall avantage cependant le double-demisme. Un tel argument n'est pas irréprochable : imaging et C.P. étant peu théorisés et éprouvés, il est difficile de savoir s'ils fondent une analogie qui surmonte les différences objectives entre les deux paradoxes.

Enfin, l'étude d'un scénario retors est approfondie par la construction de variantes de plus en plus fines. La symétrie des « compagnons » et la confrontation des estimations personnelles et interpersonnelles font des Quatre Belles une variation aboutie, qui semble donner raison au tiérisme. Pourtant, la nécessité d'une convergence des estimations personnelles cohérentes en une probabilité interpersonnelle n'est pas totalement établie, des zones d'ombre persistent. Les Belles sont-elles en mesure de partager leurs connaissances au point qu'elles deviennent communes au sens lewisien ? Pour niveler les estimations des compagnons, ne faut-il pas élever en principe la thèse selon laquelle, à épistémès égales, deux agents rationnels ne peuvent être « d'accord pour être en désaccord » ? Nous reconnaissons la doctrine d'Harsanyi, séduisante mais pas consensuelle. Par exemple, Binmore, en rappelant l'hypothèse du « petit monde » de Savage, écrit :

Je préfère défendre la théorie d'Harsanyi comme étant une hypothèse de travail plutôt qu'un principe philosophique.  
[Binmore 1992, 470]

Les nombreuses pistes de résolution du paradoxe favorisent tantôt une école, tantôt une autre. Telle une ombre insaisissable, la Belle au bois dormant se rit de chaque nouvel éclairage et nous démontre son actuelle inaccessibilité. Ce miroir de vérité nous renvoie notre incapacité à donner des définitions univoques et à fonder les sciences sur d'indestructibles évidences. Comprendons que ce n'est pas pour nous humilier et nous décourager, c'est au contraire pour nous alerter, nous affermir et, en définitive, nous faire progresser.

## Bibliographie

- BIENVENU, Alexis [2007], *Un empirisme risqué : la philosophie des probabilités de Hans Reichenbach*, Paris : Université Panthéon-Sorbonne, thèse dirigée par Michel Bitbol.
- BINMORE, Kenneth [1992], *Fun and Games : A Text on Game Theory*, Lexington : D.C. Heath and Company, cité d'après la traduction française par Francis Bismans et Eulalia Damaso : *Jeux et théorie des jeux*, Bruxelles : De Boeck Université, 1999.

- BRADLEY, Darren J. [2011], Self-location is no problem for conditionalization, *Synthese*, 182(3), 393–411, doi :10.1007/s11229-010-9748-9.
- COZIC, Mikaël [2007], and Sleeping Beauty : A case for double halfers, dans *Proceedings of TARK 2007 Conference*, édité par D. Samet, Louvain, 112–117.
- DE SCHEEMAEKERE, Xavier [2011], *Les Fondements philosophiques du concept de probabilité*, Fernelmont : EME Editions.
- DELABRE, Laurent [2008], Sleeping Beauty : Debate on a paradox, URL <http://philsci-archive.pitt.edu/4342/>.
- DELAHAYE, Jean-Paul [2003], La Belle au bois dormant, la fin du monde et les extraterrestres, *Pour la Science*, 309, 98–103.
- DORR, Cian [2002], Sleeping Beauty : in defence of Elga, *Analysis*, 62(276), 292–296, doi :10.1111/1467-8284.00371.
- ELGA, Adam [2000], Self-locating belief and the Sleeping Beauty problem, *Analysis*, 60(266), 143–147, doi :10.1111/1467-8284.00215.
- HALPERN, Joseph [2004], Sleeping Beauty reconsidered : Conditioning and reflection in asynchronous systems, dans *Oxford Studies in Epistemology*, édité par T. Szabo Gendler & J. Hawthorne, Oxford University Press, 111–142.
- JENKINS, Carrie S. [2005], Sleeping Beauty : A wake-up call, *Philosophia Mathematica*, 13(2), 194–201, doi :10.1093/philmat/nki015.
- LEWIS, David [2001], Sleeping Beauty : reply to elga, *Analysis*, 61(271), 171–176, doi :10.1111/1467-8284.00291.
- NEAL, Radford [2006], Puzzles of anthropic reasoning resolved using full non-indexical conditioning, Technical report no. 0607, Dept. of Statistics, University of Toronto.
- STALNAKER, Robert [2008], *Our Knowledge of the Internal World*, Oxford : Oxford University Press.
- WEATHERSON, Brian [2011], Stalnaker on sleeping beauty, *Philosophical Studies*, 155(3), 445–456, doi :10.1007/s11098-010-9613-1.