

Analyse de quelques problèmes traditionnels

Le cadre et les postulats de Savage, qui permettent de parvenir aux résultats que nous venons d'exposer, ont alimenté la recherche et les controverses depuis plus d'un demi-siècle, si bien qu'il semble même vain de vouloir proposer une synthèse de la littérature qu'ils ont suscitée. Nous nous contenterons, dans cette partie, de rappeler quelques-uns des problèmes traditionnels que pose la représentation savagienne d'un problème de décision en situation d'incertitude, et de montrer que notre interprétation causale permet de leur apporter un éclairage nouveau.

Les problèmes que nous évoquons sont de deux ordres. D'une part, c'est la définition d'un problème de décision à partir de deux notions fondamentales, les conséquences et les événements, et du troisième, dérivé des deux premiers, d'actes qui ne semble pas correspondre, dans de nombreux cas, à l'expérience humaine de décision en situation d'incertitude. Si l'on s'en tient au cadre de Savage, cette construction peut donc paraître artificielle ou invraisemblable. D'autre part, les postulats de Savage, qui caractérisent les préférences du décideur à l'intérieur de ce cadre, imposent aussi des conditions particulières à la façon dont il définit les trois composantes de ce cadre. Or ces conditions rencontrent à leur tour certaines difficultés quand elles sont confrontées à la pratique.

Il faut noter que cet ensemble d'hypothèses, à première vue assez restrictif, est difficile à invalider

10. Des décisions contingentes sont des décisions conditionnées par des valeurs elles-mêmes incertaines *a priori*.

pour un observateur extérieur qui le mettrait à l'épreuve sur la base de cas concrets. En effet, la violation d'une hypothèse peut toujours, comme nous le verrons, être mise sur le compte d'un mauvais choix des conséquences, des événements ou des actes, puisque ces notions ne sont pas observables. Un observateur extérieur doit tenter de deviner les événements et conséquences par lesquels un décideur caractérise un problème donné¹¹. De ce fait, le cadre de Savage ne peut être considéré comme une théorie scientifique, dont l'expérimentation doit permettre d'établir la véracité selon les canons du positivisme logique, ou au moins de la rejeter selon le critère de falsifiabilité poppérien¹². Son examen se ramène souvent à des questions d'interprétation, et de ce fait, nous pouvons utilement mener cette discussion à partir de notre exemple précédent, sans avoir formellement développé notre modèle.

Nous commençons, dans la première section de cette partie, par analyser les difficultés liées à la définition même d'un problème de décision chez Savage, avant de consacrer la seconde section à trois implications particulières des postulats : l'indépendance des utilités par rapport aux états, l'absence d'aléa moral et le conséquentialisme.

7.3.1 La définition du problème de décision

Sans viser à l'exhaustivité, il est possible de distinguer quatre critiques fréquemment adressées à la façon dont le cadre de Savage rend compte d'un problème de décision :

- des situations relativement communes se traduisent par des problèmes de décision d'une grande complexité ;
- la distinction entre événements et conséquences est floue ;
- l'hypothèse de partition de l'ensemble des états par les événements est trop contraignante ;
- les actes virtuels sont absurdes.

Nous ne nous attarderons pas sur ce dernier argument, en notant simplement qu'en dépit d'une affirmation assez commune, le cadre de Savage n'exige pas que le décideur prenne effectivement les actes virtuels en considération. Le théorème de représentation nous dit simplement que le décideur se conduit *comme s'il* maximisait une utilité espérée, et la construction de Savage nous indique un moyen, impliquant notamment les actes virtuels, pour dériver la distribution de probabilités et une fonction d'utilité correspondant à cette conduite. Si l'existence d'actes virtuels entache la pertinence de la représentation de Savage, ce ne peut donc être sur le plan conceptuel mais tout au plus sur celui de la pratique, en l'occurrence des possibilités de dérivation des probabilités et utilités. À cette éventualité, Savage nous semble avoir répondu de façon aussi convaincante que possible

11. Dans les applications descriptives, qui nous intéressent ici moins que l'usage normatif, le modélisateur doit au préalable faire l'hypothèse quelque peu héroïque que le décideur use bien des *notions* d'événements et de conséquences pour se représenter un problème donné.

12. Gilboa et Schmeidler (2001) discutent de ce point dans le prologue de leur ouvrage, et proposent d'appeler une telle construction théorique un *cadre conceptuel*.

en affirmant que l'expérience mentale de l'association entre un événement et une conséquence inconciliables en pratique n'avait rien d'irréalisable, en tout cas dès lors que l'on admettait la possibilité du raisonnement contrefactuel¹³. Mais, bien sûr, il reste que les probabilités et utilités du décideur savagien doivent être élicités, et ne peuvent être révélés par ses choix spontanés en raison de l'existence d'actes virtuels.

7.3.1.1 La complexité des problèmes de décision

Revenons cette fois au point de départ de notre exemple : un décideur en charge de la politique de santé publique sait qu'une usine rejette des substances cancérigènes, peut-être à des doses toxiques, mais que son activité procure aussi des bénéfices ; que le population consomme les produits contaminés en quantité inconnue, et peut de ce fait s'exposer au développement du cancer, mais que cette consommation a également des avantages économiques ; que les produits non consommés pourraient être exportés, et que le cancer entraîne des conséquences sanitaires incertaines ; enfin que les actes qu'il peut engager sont de limiter les rejets de l'usine, d'inciter la population à réduire sa consommation, de favoriser les exportations et de mener une campagne de dépistage du cancer. Imaginons que le décideur essaie de représenter cette situation par un problème de décision unique, où tous les enjeux feraient partie du même ensemble de conséquences, où la description de chaque événement lèverait entièrement l'incertitude, et où les actes seraient reformulés comme des fonctions entre les ensembles d'états et de conséquences correspondants. La seule élaboration d'un tel problème, sans même parler de sa résolution, représenterait une tâche d'une complexité extrême.

S'appuyant sur un exemple similaire, Karni estime :

"in many instances, the state-space is too large and complex to be compatible with the limited cognitive ability of decision makers to grasp, let alone be invoked in the decision-making process. In such instances, as the following examples illustrate, the notion of states of the world as an uncertainty-resolving device seems unintuitive, non-compelling, and outright useless, for the purpose of modeling decision-making under uncertainty" (Karni, 2006, p.326).

Gilboa et Schmeidler proposent aussi deux cas de cette nature dans lesquels :

"[Expected Utility Theory] does not describe a plausible cognitive process. Should the decision maker attempt to think in the language of EUT, she would have to imagine all possible outcomes and all relevant states. Often the definition of states of the world would involve conditional statements, attaching outcomes to acts. Not only would the

13. 'To some - perhaps to you - it will seem grotesque if I say that I should not mind being hung so long as it be done without damage to my health or reputation, but I think it desirable to adopt such language so that the danger of being hung can be contemplated in this framework. [...] Any statement of the form, "I would do this if that," is somewhat mysterious philosophically; are the ones you emphasize so much worse than the others?' (Aumann et Savage, 1971).

number of states be huge, the states themselves would not be defined in an intuitive way" (Gilboa et Schmeidler, 2001, p.32).

Ces critiques paraissent excessives en ce qu'elles laissent à penser qu'un cadre conceptuel tel que celui de Savage devrait s'appliquer de façon uniforme à tous les problèmes, indépendamment de leur degré de complexité. Il semble naturel de penser que confronté à une telle situation, un décideur rationnel chercherait à distinguer les différentes sources d'incertitude, identifier les aspects de la situation qui lui importent, évaluer comment ses actions pourraient influencer ces derniers et dans quelle mesure leur issue est affectée par l'une ou l'autre des sources d'incertitude. Il chercherait ainsi à utiliser une partie de ses capacités cognitives à décomposer son problème en une succession de problèmes de décision interdépendants. Dans chacun de ceux-ci, le cadre de Savage pourrait apparaître comme un outil de décision pertinent. La méthode que nous avons exposée à partir d'un exemple relativement complexe suit précisément ce cheminement. Elle permet, comme nous l'avons vu, de réduire considérablement le nombre d'événements, de conséquences et d'actes à considérer par rapport à une approche séquentielle traditionnelle.

7.3.1.2 La distinction entre événements et conséquences

Nombre de situations d'incertitude se décrivent le plus naturellement à partir de l'ensemble d'actes effectifs à la disposition du décideur, ainsi que les conséquences de ces actes qu'il est possible d'entrevoir. En revanche, il est souvent malaisé de résumer l'incertitude portant sur ces conséquences par un ensemble d'événements exogènes. La démarche de Savage conduit alors à associer aux conséquences identifiées un ensemble *ad hoc* d'événements exhaustifs et mutuellement exclusifs, ce qu'Arrow (1971) a formalisé en définissant les événements comme des fonctions de l'ensemble des actes vers celui des conséquences¹⁴. L'espace des états du monde n'est plus alors qu'un artefact « endogène et construit » par le décideur, selon la formule de Machina (2003), ce qui met sérieusement en cause sa pertinence, au moins sur le plan descriptif.

Pour examiner cette question, considérons une variante de notre problème de décision en S et supposons même, dans un premier temps, que le décideur ignore tout des causes des cas de décès ; il connaît simplement un ensemble de valeurs possibles de X_S en présence et en absence d'action de sa part (respectivement sous γ_S et γ'_S), tel que $\{\textit{négligeables, modérées, importantes, catastrophiques}\}$. La seule source d'incertitude qu'il puisse associer à son problème est alors un aléa général \bar{Y}_S dont les valeurs sont définies par $x_S = \gamma_S(\bar{y}_S)$ et $x_S = \gamma'_S(\bar{y}'_S)$. Inversement, la seule grandeur à l'aide de laquelle les événements, c'est-à-dire les valeurs de \bar{Y}_S , puissent être décrites est X_S : l'événement sera *grave* lorsque les conséquences sanitaires associées, en l'absence de dépistage, seront *importantes, etc.*

14. Selon Gilboa (2009, p. 114), c'est Savage lui-même qui aurait le premier proposé cette définition.

Notons que notre décideur, disposant des actes γ_S et γ'_S sous forme fonctionnelle, peut calculer les valeurs \bar{y}_S comme antécédents des valeurs x_S . Il lui est donc inutile de recourir à l'artifice dispendieux d'événements définis comme fonctions des actes vers les conséquences, qui l'encombrerait d'une grande quantité d'états virtuels¹⁵.

De telles situations sont fréquentes dans notre expérience quotidienne. Par exemple, quand, confrontés à un phénomène dont les sources nous échappent ou sont trop intriquées, nous disons : « si les choses continuent ainsi, alors... », la continuation des « choses » n'est qu'un événement abstrait associé aux conséquences observées, ce qui ne nous interdit pas de lui attribuer une probabilité. Des situations similaires se présentent fréquemment dans des problèmes de décision où les probabilités sont données, et où l'on peut donc les utiliser pour décrire les événements correspondants. Pour la gestion des risques liés à certaines catastrophes naturelles, à titre d'exemple, l'événement de référence est un aléa centennal, défini par ses conséquences et sa fréquence estimée, ce qui n'en fait aucunement un phénomène « exogène ».

Supposons à présent que notre décideur ait établi un lien causal entre le coût de la maladie et sa prévalence, considérée comme une variable exogène ($X_M = Y_M$). Ce lien change la nature de la représentation *dans la seule mesure* où la prévalence de la maladie est un phénomène distinct des conséquences sanitaires. Il devient alors possible de mettre en relation les conséquences avec des événements ayant une réalité propre, comme dans l'équation $x_S = \gamma_S(y_M)$ ou la proposition « Une maladie endémique a , en l'absence de mesure de dépistage, des conséquences modérées ». Nombre d'exemples classiques de la théorie de la décision correspondent à ce cas : l'œuf avarié, lorsque l'on choisit de le casser, est une cause autonome de l'omelette gâchée ; la foudre est une cause autonome de l'incendie de la maison en l'absence de paratonnerre, etc.

Cependant, *du fait même de l'existence autonome de la cause et de l'effet*, leur relation ne peut pas, en toute rigueur, être supposé déterministe, de sorte qu'elle n'est complète qu'en la présence d'une variable aléatoire supplémentaire : $x_S = \gamma_S(x_M, y_S)$. On peut, naturellement, choisir de négliger cette variable, ce dont nous discuterons dans le point suivant.

En résumé, les événements constitutifs d'un problème de décision n'apparaissent comme des artefacts que lorsqu'ils ne sont pas fondés sur les réalisations d'une cause identifiée. Dans ce cas, la description des événements ne peut reposer que sur les conséquences, sans toutefois que cela constitue un obstacle à la conduite du raisonnement savagien.

15. Contrairement aux actes virtuels, ces événements virtuels sont totalement inutiles dans la dérivation des probabilités et utilités du décideur, mais, dans un cadre savagien où la partition des états du monde viendrait à manquer et devrait être construite, ils seraient impossibles à distinguer des événements réels.

7.3.1.3 L'exhaustivité des événements

L'hypothèse selon laquelle le décideur pourrait, en toute circonstance, se donner une partition de l'ensemble des états du monde constitue une troisième source de difficultés. Comme une partition est exhaustive, elle signifie en effet que le décideur attribue une probabilité de 1 à l'union des événements répertoriés. Karni en déduit que le décideur doit estimer connaître le monde dans sa totalité et ignorer la simple possibilité d'une surprise (Karni, 2006, p.326). Cette conclusion mérite toutefois d'être pondérée, puisque l'hypothèse d'exhaustivité laisse le décideur libre d'inclure dans sa partition des événements destinés en quelque sorte à le prémunir contre les surprises. Machina reconnaît cette possibilité, mais considère qu'il s'agirait alors d'un événement résiduel « attrape-tout », dont les conséquences seraient nécessairement très vagues (Machina, 2003, p.31).

À ce point, il peut être utile de revenir à notre exemple. Nous considérons à présent à un problème statique où le décideur ne s'intéresserait qu'à la seule prévalence X_M de la maladie, qu'il saurait liée aux rejets X_R , à la consommation de produits contaminés X_C et à un ensemble d'autres facteurs Y_M . Nous supposons qu'il n'a d'autre acte à sa disposition que la seule observation de X_M (disons qu'il s'agit d'un problème de prévision). Supposons de surcroît que le décideur ne soit pas certain de la valeur du paramètre μ , et qu'en conséquence il le considère comme une variable $X_\mu = f_\mu(Y_\mu) = Y_\mu$, qui devient un parent supplémentaire de X_M dans le graphe causal. Les événements en X_M correspondent alors à des quadruplets (y_R, y_C, y_μ, y_M) , dont chacun est associé à une conséquence :

$$\begin{aligned} x_M &= \gamma_M(x_R, x_C, x_\mu, y_M) \\ &= \gamma_M(\gamma_R(y_R), \gamma_C(y_C), \gamma_\mu(y_\mu), y_M) \\ &= \delta \cdot y_\mu \cdot y_C \cdot 1_{y_R > \rho} + y_M \end{aligned}$$

Cela suffit à notre décideur pour parer à toute éventualité, depuis l'absence de la maladie jusqu'à son extension à l'ensemble de la population, et quelle que soit la combinaison de causes identifiées (X_R , X_C et X_μ) et non identifiées (Y_M) à son origine. Constatons que la conséquence d'aucun événement n'est définie de façon vague.

Parer à toute éventualité, dans ce contexte, signifie assigner une probabilité *a priori* non nulle à tout événement envisageable, c'est-à-dire admettre la possibilité de surprises, pour peu que celles-ci soient imaginables. Les surprises, nous le savons, sont un sujet épineux pour la théorie bayésienne, puisqu'aucune observation ne permet de corriger la probabilité d'un événement selon la règle de révision de Bayes lorsque sa probabilité *a priori* a été estimée nulle. De ce fait, un décideur bayésien qui pense connaître la totalité du monde ne peut jamais revenir sur cette opinion. *A contrario*, un décideur dont la partition initiale de l'ensemble des états du monde est suffisamment riche pourra

toujours réviser ses croyances si celles-ci sont infirmées par l'expérience.

Pour construire cette partition, le décideur sera amené, comme au point précédent, à engendrer des événements à partir des conséquences, c'est-à-dire dans un premier temps à identifier (x_R, x_C, x_μ, y_M) tels que $x_M = f_M(x_R, x_C, x_\mu, y_M)$, puis y_R tel que $x_R = f_R(y_R)$, etc. Cette démarche d'inférence des effets aux causes sera, à nouveau, plus économique que celle consistant à définir les événements comme toutes les fonctions des actes vers les conséquences. Gilboa constate que la résolution du problème de Newcombe passe par un tel « élargissement contrôlé » de l'ensemble d'événements. Il en tire l'enseignement suivant :

"in order to avoid surprises, one should take into account all possible "causal" relationships in the formulation of the states of the world" (Gilboa, 2003, p.12).

mais il n'indique pas comment y parvenir dans le cas général en évitant une « explosion de la cardinalité » des ensembles d'actes et d'états du monde.

Notre représentation permet de mettre en œuvre une version légèrement amendée de cette recommandation : il n'est pas possible de tenir compte de toutes les relations causales, au moins dans la vision de la causalité que nous défendons ici ; mais il est possible d'estimer leurs effets agrégés, et de construire le nombre strictement nécessaire d'événements sur cette base.

Notre réponse à la critique initiale de Karni et Machina est certes incomplète. En premier lieu, elle repose à ce stade sur l'hypothèse que le décideur connaît une structure causale adaptée à son problème. Nous remettons l'analyse de situations où il ne peut pas tenir une telle croyance au chapitre suivant¹⁶. Par ailleurs, le fait même que les événements considérés doivent avoir une probabilité *a priori* non nulle signifie que leur nombre doit être fini. Dans toute application réaliste, il ne pourra même être que très limité.

Pour autant, il semble excessif d'en déduire que le décideur doit croire connaître le monde ; il doit plutôt penser connaître les limites de ce qu'un événement inattendu pourrait signifier pour son état. Ce faisant, naturellement, il s'exposera à quelques surprises.

7.3.2 Les conditions portant sur les préférences du décideur

Les postulats de Savage requièrent du décideur :¹⁷

- qu'il accorde à chaque conséquence une valeur intrinsèque, indépendante de l'état du monde et de l'acte sous lesquels elle se réalise (postulats *P3* et *P4*) ;
- qu'il attribue à chaque événement une probabilité subjective, quel que soit l'acte choisi et la conséquence associée à sa réalisation (postulat *P4*) ;

16. Voir la discussion sur l'incertitude causale (section 8.4.2.2)

17. Nous laissons ici de côté la question de la représentation de l'incertitude elle-même par une distribution de probabilités, qui fera l'objet d'un traitement plus détaillé dans la dernière partie du chapitre suivant (voir en particulier la section 8.4.1).

- qu'il n'évalue un acte que sur la base de ses conséquences et des événements auxquels il les associe (postulat $P2$).

Nous discutons à présent de situations qui mettent en question la pertinence de ces conditions.

7.3.2.1 La dépendance aux états

Supposons que nous formulions le problème de décision dans les termes du point précédent, où le décideur ne semble s'intéresser qu'à la prévalence de la maladie X_M , mais qu'il soit en outre sensible à la qualité de l'offre de soins. Toutefois, ne voyant pas comment lier causalement cette variable à X_M , le décideur aurait choisi de la considérer comme une composante des facteurs exogènes Y_M (il va de soi que cette construction est erronée). Il se retrouverait alors dans une situation où il apprécierait une conséquence x_M différemment selon l'événement qui s'est réalisé, en d'autres termes, où sa fonction d'utilité serait dépendante des états. De telles préférences contreviennent aux postulats $P3$ et $P4$ de Savage, et interdisent une représentation en utilité espérée avec une distribution de probabilités unique sur l'ensemble des états du monde (Drèze, 1961). Supposons cependant qu'après mûre réflexion, le décideur trouve l'origine de cette anomalie : son état personnel n'est en fait pas déterminé par la prévalence de la maladie mais par son coût social, dont la prévalence et la qualité de l'offre de soins sont des causes. Il peut alors modifier son RBC, en y intégrant deux variables supplémentaires X_O et X_S , et reprendre la définition de son problème de décision.

Imaginons à présent que dans la même situation, le décideur se montre en outre sensible au fait que lui-même pourrait être affecté par la maladie. N'étant pas un pur altruiste, son appréciation des conséquences sanitaires de l'épidémie serait très différente selon que cet événement se produirait ou pas, et il serait donc confronté à un nouveau problème de dépendance aux états. Comme dans le cas précédent, celui-ci résulterait d'une sous-spécification de sa représentation causale, mais cette fois d'un type différent. La variable manquante dans le cas précédent représentait une cause identifiée d'un aspect de la situation important aux yeux du décideur. Ici, c'est la variable représentant un tel aspect qui fait défaut et dont la prise en compte résout le problème de dépendance.

Ces deux cas de reformulation du problème de décision correspondent, en substance, à ceux proposés par Savage en réponse au problème de dépendance aux états dans le problème du choix entre maillot de bain et raquette de tennis (Savage, 1954, p.25) et, de façon peut-être moins claire, dans celui de l'opération chirurgicale de l'épouse de M. Jones (Aumann et Savage, 1971, p. 309). Naturellement, l'argument de Savage ne reposait pas sur une interprétation causale. Mais celle que nous proposons permet à la fois de clarifier le débat et d'étayer son raisonnement.

Dans notre modèle, la dépendance aux états ne peut survenir que si la transposition d'une situation incertaine en RBC a été incomplète. L'utilité d'une conséquence donnée ne peut dépendre d'un élément extérieur que dans deux cas de figure : si celui-ci en est une cause, directe ou indirecte

(c'est-à-dire un ascendant markovien), *qui est connue du décideur* (puisque l'utilité de celui-ci ne peut pas dépendre *a priori* d'aspects du monde qu'il n'a pas identifiés comme causes du problème considéré) ; ou si celui-ci est un aspect du problème qui détermine l'état du décideur et, de surcroît, un aspect dont l'utilité est inséparable de celle de la conséquence considérée. Le problème de la maladie personnelle du décideur (ou de l'opération chirurgicale de sa femme dans le cas de M. Jones) est parfois considérée comme un cas d'inapplicabilité de la théorie de la décision en raison de la modification de la capacité même du décideur à apprécier des conséquences (ce que Gilboa appelle son « moteur hédonique » (Gilboa, 2009, p. 130)). Notre représentation ne permet pas de répondre à la question de validité du raisonnement contrefactuel dans un cas où le monde contrefactuel est radicalement différent du monde actuel. Mais elle permet de distinguer ce problème de celui de séparation de conséquences dont les utilités sont inséparables, ce qui ne nous semble pas avoir été toujours le cas dans les contributions passées à cette question.

L'analyse causale affine aussi notre compréhension d'un autre aspect épineux du cadre de Savage : celui des limites du « petit monde » représentant un problème de décision, et plus précisément des conséquences ultimes à considérer. Dans la mesure où le décideur verrait dans chaque variable qu'il prendrait en compte une influence causale vers une autre variable qui lui importerait, sa représentation pourrait s'étendre à l'infini. Comment lui imposer des limites adéquates pour une situation de décision donnée ? Savage avouait être lui-même incapable de formuler un critère pour cela (Savage, 1954, p. 16). Le décideur peut effectivement arrêter son exploration prospective lorsque l'utilité d'une conséquence lui fournit une mesure satisfaisante de l'utilité des problèmes de décision qui en découlent, comme lorsque l'utilité d'une somme monétaire est considérée approximer l'utilité espérée des usages futurs qui en seront faits. Dans notre modèle, cet « élagage » s'avère toutefois problématique si les problèmes de décision éliminés font parties des conséquences indirectes d'autres problèmes de décision qui sont, eux, conservés. Il apparaît ainsi que le décideur peut éviter de développer une ramification du réseau si ses choix peuvent être représentés par une espérance d'utilité récursive et si aucun problème de décision à venir ne mène à la branche élaguée, c'est-à-dire si celle-ci intègre tous les problèmes de décision contingents au dernier problème considéré.

7.3.2.2 L'aléa moral

Imaginons qu'un observateur ait pour but de dériver les probabilités et utilités du décideur à partir de ses choix dans le problème de décision en X_S . Supposons que cet observateur connaisse le modèle causal que le décideur a en tête et les interventions qu'il peut effectuer, sauf au nœud X_C où il ignore la possibilité pour le décideur d'influer sur le niveau de consommation par le biais du mécanisme f'_c . De ce fait, l'observateur peut incorrectement attribuer l'action du décideur sur X_C à la variable exogène Y_C . Supposons maintenant que dans une situation où les rejets sont im-

portants, le choix optimal du décideur soit de limiter la consommation et, se trouvant dans une situation simplement endémique, de choisir le laisser-faire en X_S ; mais qu'en revanche, lorsque les rejets sont moindres, la meilleure option consiste à ne pas agir sur le niveau de consommation, et de réduire les conséquences sanitaires d'une épidémie plus forte par le biais d'une campagne de dépistage en X_S . Du point de vue de l'observateur, les préférences du décideur en X_S présentent alors une incohérence : le dépistage paraît toujours plus avantageux que le laisser-faire sous l'événement ($Y_S = y'_S$) et moins avantageux sous l'événement ($Y_S = y_S$), mais le choix du décideur semble dépendre de l'ampleur des conséquences attendues. En d'autres termes, la probabilité subjective attribuée à ($Y_S = y_S$) semble changer avec les conséquences. L'observateur est amené à conclure que les préférences du décideur ne se conforment pas au postulat $P4$ de Savage, et ne peuvent de ce fait être représentées par une espérance d'utilité. Le problème, cependant, réside non pas dans le comportement du décideur, mais dans l'interprétation de l'observateur. Il a été qualifié dans la littérature d'aléa moral (Drèze, 1961 ; Arrow, 1963), par référence au problème que rencontre un assureur quand l'existence d'une couverture assurantielle conduit à un changement de comportement de l'assuré dans le sens d'une dégradation du risque. Dans une situation d'aléa moral, la séparation entre états du monde, conséquences et actes est donc entravée par la capacité du décideur à influencer les événements par une action extérieure.

Considérons à présent une situation où le décideur choisit de traiter en premier lieu le problème des débouchés à l'exportation. Si le résultat est positif, il peut laisser l'usine rejeter autant de substance que le nécessite son activité et mener une campagne nationale contre la consommation de produits contaminés. Dans le cas contraire, il choisit de limiter les rejets à la source. L'observateur ignore cette fois l'action en X_E , dont il attribue les effets à Y_E , et il conclut que les préférences du décideur en X_R dépendent non pas des conséquences directes, mais de conséquences assez lointaines. Cet exemple montre qu'il y a aléa moral non seulement lorsque le décideur peut influencer le problème de décision considéré par une action sur ses causes, mais plus généralement lorsqu'il peut en modifier le contexte au sens de l'observation 6.

Notre représentation caractérise le contexte causal d'une décision, et considère que le décideur et la Nature jouent des rôles symétriques dans son établissement. Comme la dépendance aux états, l'aléa moral ne peut y apparaître qu'en cas de sous-spécification du modèle. L'oubli d'une intervention sur une variable X_i affecte alors l'observation de tous les problèmes de décision contingents à X_i .

7.3.2.3 Le conséquentialisme

Dans la théorie de la décision dite « objective », c'est-à-dire celle qui considère des problèmes de décision où les probabilités des événements sont pré-établies, un axiome d'indépendance requiert

que les préférences du décideur entre deux actes soient préservées lorsque chacun de ces actes est « mélangé » par combinaison linéaire avec un troisième. L'un des corollaires de cet axiome, lorsque l'opération de mélange est elle-même interprétée comme une loterie, est une propriété que l'on a convenu de nommer *conséquentialisme* (Hammond, 1988a ; Karni and Schmeidler, 1991). Dans un problème de décision séquentielle, les choix du décideur conséquentialiste à un noeud donné ne dépendent pas de la trajectoire qui l'a amené jusque là, ni des alternatives qui ne se sont pas réalisées. Le principe est également à l'œuvre de façon sous-jacente dans le cadre de Savage, où il découle du postulat *P2* (principe de la chose sûre), et apparaît clairement dès lors que l'on considère un problème de décision dynamique sans le réduire au choix *ex ante* d'une stratégie (Hammond, 1988b). Un décideur qui ne se conforme pas à ce principe peut faire des choix dynamiques incohérents qui l'exposent aux paris hollandais.

En imposant au décideur de ne se préoccuper que des conséquences de ses choix à venir, le conséquentialisme véhicule effectivement un certain nombre de présupposés de la théorie qui s'opposent à une conduite fondée sur des principes déontologiques¹⁸. Le plus important de ces présupposés est peut-être l'absence de distinction entre des conséquences qui relèvent de l'action du décideur et d'autres, par ailleurs équivalentes, qui résultent des « choix de la Nature ». Une des formes particulièrement expressives est l'indifférence entre un bien (ou un mal) que l'on a causé et un autre auquel on a permis de se produire (Howard-Snyder, 2007). Machina (1989) décrit le dilemme d'une mère qui préfère voir le sort choisir lequel de ses enfants elle doit privilégier plutôt que d'avoir à en décider elle-même. Il qualifie ses préférences de « non séparables entre événements », mais ce faisant, il omet, nous semble-t-il d'indiquer la véritable origine du problème.

Pour étudier un problème équivalent dans notre modèle, considérons la situation de notre gestionnaire de santé publique lorsqu'il évalue ses actes en X_R . Il a alors la perspective d'obtenir le même niveau de rejets ρ s'il choisit de les limiter (f'_R) que s'il choisit le laisser-faire (f_R) et que la Nature choisisse de limiter l'activité de l'usine ($Y_R = \rho$). Le décideur peut alors souscrire à un principe déontologique selon lequel une conséquence (la faiblesse des rejets) n'aurait pas la même valeur, dans ses aspects positifs (l'absence de toxicité) comme négatifs (la faiblesse de l'activité), selon qu'il l'aurait provoquée ou qu'elle résulterait du cours des choses. Il paraît artificiel d'interpréter cette position comme une dépendance de l'utilité des conséquences à la valeur de la conséquence non réalisée, et d'attribuer l'exclusion de telles préférences au principe de conséquentialisme. Ce que le décideur signifierait est plutôt une appréciation différente de ses actes selon qu'ils respectent le mécanisme causal préexistant (f_R) ou qu'ils le changent (f'_R).

En considérant de la même façon le laisser-faire et les interventions actives du décideur, nous avons donc implicitement souscrit à un principe conséquentialiste important. Cette hypothèse a

18. Pour une discussion générale du conséquentialisme, on pourra se reporter à Sinnott-Armstrong (2006).

été faite lors de la définition de nos concepts initiaux, et non dans le cadre axiomatique, et il en est de même pour la théorie de la décision traditionnelle. Pour en changer, il nous suffit de permettre au décideur d'inclure dans la structuration de son problème une variable supplémentaire de gratification éthique, et d'ajouter des flèches entre tous les nœuds de décision pertinents et le nœud de cette variable.

L'hypothèse de conséquentialisme renvoie donc, de façon ultime, à la définition des conséquences. Si les conséquences d'un acte en X_i tiennent compte de tout ce qui détermine l'état du décideur, alors il est naturel de supposer que les préférences de ce dernier ne dépendent que des conséquences (Hammond, 1989). Comme nous en avons discuté plus haut¹⁹, ceci est assuré dans notre cadre par l'intégration au RBC de toutes les causes de X_i connues du décideur. Si cette condition est respectée, alors il semble y avoir peu d'arguments à opposer à l'hypothèse de conséquentialisme telle que formulée habituellement. Si ce n'est pas le cas, alors nous pouvons conclure, une nouvelle fois, que la structure causale proposée est incomplète.

7.4 Conclusion

À partir d'un exemple fictif, nous avons exposé dans ce chapitre les grandes lignes et les principaux résultats d'un modèle de décision fondé sur les RBC. Nous avons montré comment le dispositif d'évaluation du risque proposé au chapitre précédent pouvait être adapté à un contexte de décision. Par son intermédiaire, nous avons proposé une relecture du cadre de Savage, qui caractérise un problème de décision par des ensembles de conséquences, d'événements et d'actes. Nous avons montré que la prise en compte de ce dispositif causal accroissait très sensiblement la capacité de la théorie de la décision à traiter des problèmes complexes. Nous avons finalement survolé quelques-uns des problèmes classiques d'interprétation que pose la définition d'un problème de décision.

À la lumière de cette dernière discussion, il ne semble pas exagéré de dire que les rapports causaux entre états du monde, actes et conséquences constituent la part refoulée de la théorie bayésienne de la décision, notamment chez Savage. L'existence de tels rapports est évidente dans tous les exemples classiques où le cadre savagien semble bien adapté : l'état de l'œuf et l'action de le casser causent la qualité de l'omelette, la survenue de la foudre et l'absence de paratonnerre causent un incendie, etc. En contrepoint, la théorie doit faire des détours et accepter certaines ambiguïtés afin de ne pas avoir à rendre explicites ses soubassements causaux. Les problèmes que nous venons d'évoquer sont, à cet égard, à l'image de symptômes qui s'estompent largement dès que le non-dit est mis en lumière.

19. Voir la sous-section 7.2.2.3