Une approche ordinale de la décision dans l'incertain : axiomatisation, représentation logique et application à la décision séquentielle

Régis Sabbadin

No d'ordre : 3277

THÈSE

présentée devant
L'UNIVERSITÉ PAUL SABATIER DE TOULOUSE
en vue de l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ PAUL SABATIER
Spécialité : Informatique

par

Régis SABBADIN

UNE APPROCHE ORDINALE DE LA DECISION DANS L'INCERTAIN : AXIOMATISATION, REPRESENTATION LOGIQUE ET APPLICATION A LA DECISION SEQUENTIELLE

Soutenue le Mercredi 9 Décembre 1998

Devant le jury composé de :

Mr Pierre BERNHARD	Professeur à l'ESSI (Nice)	Examinateur
Mr Michel CAYROL	Professeur à l'Université Paul Sabatier	Examinateur
Mr Didier DUBOIS	Directeur de recherche IRIT/CNRS	Directeur
Mr Hector GEFFNER	Professeur à l'Université de Caracas	Examinateur
Mr Malik GHALLAB	Directeur de recherche LAAS/CNRS	Examinateur
Mr Jean-Yves JAFFRAY	Professeur à l'Université Paris VI	Rapporteur
Mr Henri PRADE	Directeur de recherche au CNRS	Examinateur
Mr Marc ROUBENS	Professeur à l'Université de Liège	Rapporteur

sabbadin@irit.fr
Tél : (33)-(0)5-61-55-82-97
Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier - 118, Route de Narbonne - 31062 Toulouse Cedex 4

Résumé

La première partie de cette thèse consiste en un état de l'art de la décision dans l'incertain, considérée du point de vue de l'Intelligence Artificielle. En premier lieu sont décrites certaines approches classiques de la décision dans l'incertain, dont la théorie de l'utilité espérée puis certaines approches non-classiques, numériques ou ordinales. Ensuite, nous évoquons deux domaines où la décision dans l'incertain interagit avec l'IA : Les processus décisionnels Markoviens et leur application à la planification dans l'incertain et la representation logique des preferences.

La seconde partie constitue l'apport spécifique de cette thèse. Elle est divisée en trois sous-parties reprenant les thèmes de la première partie :

1) Nous étudions des critères qualitatifs de décision dans l'incertain prenant leurs valeurs dans une échelle finie, totalement ordonnée et nous en proposons une justification axiomatique. Ces critères sont basés sur une intégrale de Sugeno qui peut être considérée comme une contrepartie qualitative de l'intégrale de Choquet. Parmi les critères axiomatisés on retrouve, entre autres, les deux fonctions d'utilité qualitative possibiliste proposées par Dubois et Prade.

2) Nous étudions une contrepartie possibiliste des processus décisionnels Markoviens, totalement et partiellement observables et leur application à la planification sous incertitude et nous proposons un certain nombre d'algorithmes de résolution de type "programmation dynamique".

3) Nous proposons enfin un langage structuré de représentation des problèmes de décision sous incertitude qualitative. Ce langage est basé sur la logique propositionnelle (valuée) et les "systèmes de maintien de la cohérence basés sur les hypothèses" (ATMS). Le langage proposé s'adapte à la représentation de problèmes de décision sous incertitude, non seulement lorsque celle-ci est du type "possibiliste", mais aussi lorsqu'elle est représentée par des "fonctions de croyance". Des méthodes et algorithmes de résolution sont proposés, utilisant des procédures de recherche de modèles du type Davis et Putnam.

Remerciements

Je tiens à remercier Jean-Yves Jaffray et Marc Roubens pour l'intérêt qu'ils ont bien voulu porter à ma thèse en acceptant d'en être les rapporteurs. La version finale de cette thèse doit beaucoup aux nombreuses remarques du premier sur le manuscrit et aux discussions enrichissantes que j'ai pu avoir avec le second.

Pierre Bernhard, Michel Cayrol, Hector Geffner et Malik Ghallab m'ont fait l'honneur de participer à mon jury ; je leur en suis reconnaissant.

Didier Dubois et Henri Prade sont deux chefs d'une grande rigueur scientifique et d'une grande humanité. Je les remercie pour m'avoir initié à la recherche, pour leurs conseils permanents et pour l'intérêt qu'ils portent à leurs thésards.

Je remercie également Pierre Grenier, qui après m'avoir co-encadré pendant mon stage de DEA, a continué à suivre le déroulement de ces travaux, de très près, pendant ces trois années.

Hélène Fargier et Jérôme Lang m'ont fait une petite place dans leur bureau lorsque je suis arrivé en DEA : cela fait partie des petits coups de chance qui permettent de transformer le dur travail de thésard en partie de plaisir.

Merci à Christine pour sa relecture attentive du manuscrit, pour son soutien efficace et nécessaire pour la thèse et pour le reste... Et merci à tous les autres, membres du “troisième étage”, ancien ou nouveau militaires, ou pas militaires du tout !

Last, but not least, ma famille, toujours si proche malgré l'éloignement géographique.

Table des matières

Introduction

Traditionnellement, l'Intelligence Artificielle traite de la représentation des connaissances et du raisonnement à partir de celles-ci. Certaines branches de l'Intelligence Artificielle visent cependant à aider directement un agent à prendre des décisions, c'est-à-dire à choisir une action parmi un ensemble de décisions envisageables. Parfois, les informations dont dispose l'agent pour choisir une action sont entachées d'incertitude. Il peut n'avoir qu'une vision incomplète ou déformée de l'environnement du problème de décision que l'on qualifie alors d'incertain.
En IA, la prise de décision en environnement incertain intervient entre autres cas dans :

– La *planification sous incertitude*. En planification classique, un agent cherche une séquence d'actions devant mener un système donné dans un état satisfaisant. Dans la réalité, le système peut évoluer de manière imprévue sous l'influence du monde extérieur ou les actions peuvent ne pas avoir l'effet escompté (le monde extérieur évolue sous l'influence des actions de l'agent, de façon plus ou moins prévue).

– Le *diagnostic de pannes*, dans lequel les actions correspondent soit aux tests visant à repérer les pannes, soit aux réparations éventuelles.

– L'*ordonnancement de tâches dans un atelier*, par exemple lorsque celles-ci ont des durées incertaines et que des événements imprévus (pannes, commandes de "clients", ...) peuvent bouleverser les moyens de production ou les objectifs.

La problématique de la décision dans l'incertain fait intervenir des notions d'*incertitude* sur l'état réel du système considéré et de *préférences* sur les conséquences des actions dans les différents états possibles du système.
Les théories de la décision dans l'incertain visent à élaborer des critères de choix entre les actions. Ces critères de choix sont déterminés, entre autres, par la nature de l'incertitude et des préférences d'un agent.
Ces théories peuvent être envisagées selon deux aspects :

– L'aspect *normatif*. Les règles ou axiomes proposés par une théorie représentent des "postulats" qu'un agent rationnel doit respecter. Le respect de ces postulats entraîne l'utilisation d'un critère de choix donné entre les actions.

– L'aspect *descriptif*. Les postulats sont élaborés et testés à partir de l'observation du comportement de sujets différents face à des problèmes de décision réels. Ils doivent alors refléter le plus fidèlement possible le comportement d'une classe d'agents.

La théorie de la décision la plus communément utilisée est basée sur l'utilisation du critère de l'*utilité espérée*. Ce critère suppose une représentation de l'incertitude de l'agent par une distribution de probabilité sur les états possibles du monde et de ses préférences par une fonction d'utilité à valeurs réelles. Les actions sont modélisées soit par des distributions de probabilité sur les conséquences (von Neumann et Morgenstern, [vNM44]), soit par des applications de l'ensemble des états du monde dans celui des conséquences. Les actions sont comparées selon l'espérance mathématique de la fonction d'utilité.

La théorie de l'utilité espérée peut se révéler inapte à représenter le comportement d'un décideur face à certains problèmes. Des paradoxes ont été mis en évidence et des généralisations ou des alternatives à la théorie de l'utilité espérée ont été proposées (voir chapitre 2, partie I).
En Intelligence Artificielle, justement, le besoin de modèles non probabilistes pour traiter des connaissances incertaines a été exprimé depuis longtemps et des alternatives aux fonctions d'utilité classiques ont été proposées pour la représentation des préférences.
Plusieurs modèles alternatifs ont été proposés pour la décision dans l'incertain, faisant appel à des mesures d'incertitude/fonctions d'utilité non classiques : quantitatifs (chapitre 2, partie I) ou qualitatifs (chapitre 4, partie I). Par qualitatif on sous-entend que les échelles utilisées pour mesurer l'incertitude et les préférences sont ordinales. Dans les cas extrêmes, les relations d'incertitude et de préférence sont simplement binaires : on distingue les états du monde possibles ou impossibles et les conséquences satisfaisantes ou non satisfaisantes.

Notre étude s'insère dans cette dernière mouvance d'approches ordinales de la décision en environnement incertain. Nous ferons appel à des critères de décision ordinaux, pour lesquels nous proposerons une axiomatisation en termes de préférence entre actions. Ensuite, nous exploiterons ces critères originaux dans deux domaines d'étude privilégiés de l'Intelligence Artificielle :

- Dans le domaine de la planification sous incertitude, en premier lieu.
- Puis dans le domaine de la représentation logique pour la résolution (par une machinerie logique) de problèmes de décision dans l'incertain.

Ce manuscrit se compose d'une première partie décrivant un certain nombre de critères de décision dans l'incertain et mettant en lumière deux champs d'interaction entre Décision et Intelligence Artificielle et d'autre part d'une seconde partie (représentant notre apport) et traitant d'une approche ordinale de la décision dans l'incertain basée sur la théorie des possibilités. Elle intègre à la fois des aspects axiomatisation, décision séquentielle, méthodes logiques de représentation et de calcul de décisions optimales.

Dans la première partie, la problématique de la décision dans l'incertain sera décrite sous deux angles :

- Dans les chapitres 1 et 2, nous allons la décrire sous l'angle des *Sciences de la décision*. Après un rapide exposé de la problématique, montrant les deux aspects complémentaires incertitude-préférences dans les problèmes de décision dans l'incertain, nous décrirons un certain nombre de critères de décision et nous nous intéresserons aux *axiomatisations* de ces différents critères.
- Dans les chapitres 3 et 4, nous évoquerons la problématique de la décision dans l'incertain sous l'angle de *l'Intelligence Artificielle* (IA).

  Nous présenterons en particulier deux domaines où les Sciences de la Décision et l'IA peuvent interagir fortement :

  - Le domaine de la planification sous incertitude est en effet un domaine de l'IA pouvant faire appel aux Sciences de la Décision, en particulier par l'intermédiaire de la théorie des *processus décisionnels markoviens* (chapitre 3).

– Le domaine des langages de représentation des problèmes de décision dans l'incertain. Nous exposerons dans le chapitre 4 quelques formalismes logiques récents pour la représentation de problèmes de décision dans l'incertain.

L'IA a proposé des méthodes de représentation des préférences et des connaissances naturelles (implicites, granulaires et compactes) alors que ces problèmes de représentations sont quelque peu négligés par les Sciences de la Décision qui se concentrent plutôt sur les modèles de décision.

Dans la seconde partie nous présenterons tout d'abord les principaux résultats de cette thèse en ce qui concerne l'axiomatisation de critères "qualitatifs" pour la décision dans l'incertain. Le premier chapitre de cette partie sera découpé comme suit :

– Nous rappellerons l'approche de Dubois et Prade [DP95b], basée sur la théorie des possibilités, pour la décision dans l'incertain qualitatif. Nous décrirons l'axiomatisation, du type de celle de *von Neumann et Morgenstern*, des critères qu'ils ont proposée, considérant les actions comme des "loteries possibilistes".

– Ensuite, nous décrirons en détail l'approche que nous avons proposée pour l'axiomatisation de ces critères, qui, tout comme celle de Savage, considère les actions comme des applications de l'ensemble des états du système dans un ensemble de conséquences.

– Enfin, nous proposerons une axiomatisation pour des critères de décision basés sur des *degrés de confiance monotones*, généralisant les mesures de possibilité et de nécessité.

Ensuite, dans le chapitre 2 de cette seconde partie, nous allons étendre la théorie possibiliste de la décision sous incertitude proposée par Dubois et Prade [DP95b] à la décision séquentielle, c'est-à-dire lorsque plusieurs actions doivent être appliquées successivement, avant d'atteindre les objectifs d'un agent. La théorie proposée est une contrepartie possibiliste de la théorie des processus décisionnels markoviens en environnement observable. Le calcul d'une *police* (ou séquence d'actions) optimale est réalisé à l'aide de méthodes de *programmation dynamique*.

Enfin, dans le dernier chapitre nous présenterons deux approches logiques de la décision dans l'incertain s'appuyant sur la *logique possibiliste* : l'une pessimiste et l'autre optimiste. Les connaissances et les préférences seront exprimées dans des bases stratifiées de formules logiques propositionnelles. Nous montrerons alors que les deux approches sont respectivement en accord avec les deux fonctions d'utilité qualitative possibilistes, pessimiste et optimiste exposées dans le chapitre 1 de cette même partie.
Reprenant alors nos travaux présentés dans [DLPS98a], nous donnerons de brefs rappels sur le cadre des *Assumption-based Truth Maintenance Systems* (ATMS) de de Kleer [de 86] et nous montrerons comment ce cadre permet de coder un problème de décision qualitative possibiliste.
L'approche originale que nous proposons pour la décision qualitative possibiliste, peut être généralisée au calcul de décisions optimales vis-à-vis d'autres critères de décision. Nous terminerons ce chapitre en exposant le lien entre les ATMS et les approches de la décision dans l'incertain basées sur les fonctions de croyance de Dempster et Shafer [Dem67, Sha76]. Nous décrirons également un cadre général basé sur les ATMS qui permet d'intégrer la représentation de problèmes qualitatifs de décision sous incertitude, indépendamment des fonctions utilisées pour mesurer l'incertitude et les préférences.

Première partie

Décision dans l’incertain et Intelligence Artificielle

Chapitre 1

Décision dans l'incertain et décision sous risque

1.1 Introduction

L'objet d'une théorie de la décision dans l'incertain est de proposer un ou des critères permettant de modéliser le comportement d'un agent face à un problème de choix parmi un ensemble d'actions disponibles. Selon (Savage, 1954 [Sav54]), une action a est une application d'un ensemble d'états du monde possibles $S = \{s_1, s_2, \ldots\}$ vers un ensemble de conséquences $X = \{x_1, x_2, \ldots\}$. L'ensemble S représente l'ensemble des états du monde dans lesquels l'action peut être appliquée, et pour tout état $s \in S$, $a(s) \in X$ représente le résultat de l'action a appliquée à l'état du monde s. On parle de *décision dans l'incertain* lorsque les conséquences d'une action ne sont pas complètement connues avant qu'on exécute cette action et/ou lorsque l'état du monde n'est pas exactement connu au moment du choix de l'action. L'agent exprime des préférences sur les actions, reflétant ses préférences sur les conséquences des actions.

Illustrons tout d'abord les notions d'état du monde, d'action et de conséquence à l'aide de l'exemple suivant, tiré de (Savage, 1954 [Sav54]) :

Exemple 1.1.1 *Vous êtes en train de préparer une omelette. Vous avez déjà cassé cinq oeufs dans un bol, vous tenez le sixième en main et celui-ci vous semble suspect : est-il pourri ou sain?*
Pour simplifier le problème, supposons que vous ayez trois actions à votre disposition :

– *casser l'oeuf dans l'omelette,*

– *le jeter directement à la poubelle,*

– *sortir une tasse pour y casser l'oeuf et examiner son état de fraîcheur, avant d'éventuellement l'incorporer à l'omelette.*

Ce problème de décision peut être résumé par le tableau suivant :

	État du monde	
Action	oeuf sain (s_1)	oeuf pourri (s_2)
Casser l'oeuf dans le bol	omelette à 6 oeufs	pas d'omelette, 5 oeufs gâchés
Le casser dans une tasse	omelette à 6 oeufs une tasse à laver	omelette à 5 oeufs une tasse à laver
Le jeter	omelette à 5 oeufs un oeuf gâché	omelette à 5 oeufs

Les états du monde (pertinents au problème) sont :

– s_1 : "l'oeuf est sain",

– s_2 : "l'oeuf est pourri".

Le problème de décision revient à choisir quelle action appliquer parmi les trois disponibles. Ces trois actions "disponibles" forment un sous-ensemble de l'ensemble X^S des actions théoriquement envisageables (l'ensemble de ces dernières étant l'ensemble des applications de S vers X). Classiquement, on suppose l'existence d'un *préordre* (relation réflexive et transitive), *complet*, exprimant les préférences de l'agent entre les actions disponibles. L'existence de ce préordre impose que l'ordre de préférence entre deux actions ne dépend pas de l'ordre de préférence entre l'une de ces actions et une troisième. Dans l'exemple précédent, l'ordre de préférence entre les actions "casser l'oeuf dans l'omelette" et "le jeter" ne doit pas dépendre du fait que l'action "le casser dans une tasse" soit disponible ou non.

Comment choisir une action ? Le choix de l'agent est supposé guidé par une *relation de préférence* sur les conséquences éventuelles des actions applicables, et une *relation d'incertitude* entre les états possibles du monde. En effet, une action peut donner de meilleurs résultats qu'une autre dans certains états du monde et de moins bons dans d'autres. S'il en est ainsi, les "plausibilités" relatives des états du monde possibles doivent être prises en compte pour évaluer les actions.

Parlons des préférences, dans un premier temps. Dans ce problème, il est évident que la conséquence "omelette à 6 oeufs" est préférée à "pas d'omelette, 5 oeufs gâchés"...
Dans le cas général, cette préférence peut être complexe à déterminer entièrement. Dans le meilleur des cas, elle est exprimée par l'intermédiaire d'une fonction d'utilité, numérique ou ordinale, sur les conséquences. Dans d'autres cas, un agent exprime ses préférences suivant plusieurs critères et l'un des problèmes à résoudre est d'agréger ces différents critères : ce sont des *problèmes de décision multicritères* (Roy, 1985 [Roy85], Vincke, 1989 [Vin89]). Dans d'autres cas encore, un agent a un critère de décision unique, mais plusieurs agents sont impliqués dans la prise de décision : l'agrégation des préférences de plusieurs agents est le thème d'étude de *la théorie du choix social* (Arrow, 1951 [Arr51]). Le cadre de cette thèse étant restreint au domaine (déjà vaste) de la décision dans l'incertain, nous nous limiterons au cas où les préférences sont exprimées directement par une fonction d'utilité unique, numérique ou purement ordinale. Nous verrons tout de même que dans certains cas cette relation de préférence unique peut être issue de l'agrégation de relations de préférence locales, représentées par exemple par des contraintes ou des buts.

Le problème de décision peut éventuellement être entaché d'incertitude. Si l'état du monde est parfaitement connu, le problème de décision revient à choisir l'action dont la conséquence (dans cet état) est préférée à toutes les conséquences des autres actions. Cela revient à dire que les actions sont directement assimilées à leur conséquence et ordonnées d'après la relation de préférence entre ces dernières. Il est à noter que dans ce cas, peu importe que la fonction d'utilité soit numérique ou non. Seul compte l'ordre entre les conséquences. Dans l'exemple précédent, si on sait que l'oeuf est sain, on choisira de le casser directement dans l'omelette, ce qui aura pour effet de procurer une omelette à six oeufs (c'est la conséquence préférée). Si au contraire on sait qu'il est pourri, la meilleure conséquence que l'on peut obtenir (une omelette à cinq oeufs) le sera en choisissant de jeter l'oeuf.
Que se passe-t-il maintenant si l'état réel de fraîcheur de l'oeuf est inconnu ? Prendre la décision de casser l'oeuf dans l'omelette est excellent si l'oeuf est sain, mais déplorable si il est pourri. Le jeter est la meilleure chose à faire s'il est pourri, la pire s'il est sain. La meilleure action est-elle de le casser dans une tasse à part, alors que cette action n'est la meilleure ni dans le cas où l'on sait que l'oeuf est sain, ni dans celui où on sait qu'il est pourri ?

Dans ce chapitre et le suivant, nous allons passer en revue un certain nombre de théories pour la décision dans l'incertain. Ces théories diffèrent principalement par la *nature* de l'incertitude et par sa représentation. Dans ce chapitre, nous évoquerons les deux cas classiques d'incertitude : *l'incertitude non probabilisée* [LR57], [Mil54]et le *risque* [vNM44].

– En premier lieu, nous examinerons le cas où les connaissances de l'agent décideur sont les moins informatives. Il ne connaît pas l'état réel du monde $s_0 \in S$, mais il sait seulement que celui-ci appartient à un ensemble $A \subseteq S$ d'états possibles. Nous explorerons les critères de décision recensés par [LR57], pouvant être utilisés pour résoudre ces problèmes, ainsi que leurs justifications axiomatiques, proposées entre autres par Milnor [Mil54] ou Brafman et Tennenholtz [BT96, BT97].

– Ensuite, nous verrons un cas dans lequel, quoique toujours confronté à un problème de décision sous incertitude, l'agent dispose d'une information plus riche. Il dispose de données statistiques sur les différents états possibles du monde, résumées par une distribution de probabilité *objective*, p sur S. Lorsqu'il en est ainsi, on parle de *décision dans le risque* [vNM44]. Nous présenterons la *théorie de l'utilité espérée,* proposée pour traiter ces problèmes de décision sous risque. Dans cette théorie, l'agent classe les actions suivant l'espérance mathématique d'une fonction d'utilité sur les conséquences. De manière opérationnelle, la théorie permet de construire la fonction d'utilité de l'agent (sur l'ensemble des conséquences) à partir des préférences qu'il exprime entre les actions et de la distribution de probabilités objective. Elle suppose, entre autres, que le problème de décision en question est rencontré plusieurs fois et que les conséquences, bonnes ou mauvaises, se compensent, puisqu'elle évalue les actions en faisant une moyenne arithmétique pondérée des conséquences possibles.

1.2 Décision dans l'incertain non probabilisé

Les critères décrits dans cette section s'appliquent à des problèmes de décision pour lesquels l'information est très grossièrement représentée : les connaissances du décideur ne sont modélisées que par un sous-ensemble d'états possibles du monde. Parfois, la connaissance du décideur est représentée de façon plus élaborée. Celui-ci peut, par exemple, disposer de données *statistiques* sur les états possibles du monde, exprimées par une *distribution de probabilité* sur les états possibles. Lorsqu'il en est ainsi, on parle de *décision sous risque*. C'est ce cas que nous décrirons plus en détails dans la section 1.3 page 15.

1.2.1 Quelques critères classiques pour la décision dans l'incertain non probabilisé

Les critères de décision dans l'incertain que nous présentons dans ce paragraphe (et que nous retrouverons plusieurs fois dans cette thèse) sont décrits dans (Luce et Raiffa, 1957 [LR57]). L'ensemble des états possibles du monde est supposé fini, ainsi que l'ensemble des actions disponibles (nous préciserons pour toutes les théories de la décision exposées dans cette thèse si elles supposent des ensembles finis ou infinis d'états/de conséquences).
Dans ce paragraphe, les problèmes de décision sont représentés par des matrices $[u_{ij}]$, où $u_{ij} = u(a_i(s_j))$ représente l'utilité de la conséquence de l'action a_i lorsque l'état du monde est s_j (l'ensemble des états du monde est $S = \{s_1, \ldots, s_n\}$ et l'ensemble des actions est $\mathcal{A} = \{a_1, \ldots, a_p\}$).
Les quatre critères de décision dans l'incertain recensés par Luce et Raiffa [LR57] sont les critères *maximin* (ou critère de Wald), *minimax regret*, *Hurwicz* et *Laplace*.

– Le critère *maximin* évalue les actions selon l'utilité de leur pire conséquence possible. Ce critère est absolument pessimiste, et ne permet pas de compensation entre les conséquences d'actions répétées.

– Le critère *minimax regret* proposé par Savage [Sav51] évalue les actions en fonction des différences, pour chaque état possible du monde, entre les utilités de leur conséquence et de celle que la meilleure action permet d'obtenir. Il les évalue ainsi relativement aux autres actions disponibles, contrairement au critère *maximin* qui évalue les actions de manière "absolue".

– Le critère d'Hurwicz généralise le critère de Wald, puisqu'il évalue les actions en fonction d'une moyenne pondérée (par un indice de pessimisme) des utilités de leur pire et de leur meilleure conséquence possible.

– Enfin, le critère de Laplace évalue les actions en fonction de la moyenne arithmétique des utilités de toutes leurs conséquences possibles. Ici, la notion de compensation entre conséquences d'actions répétées prend tout son sens.

Ces quatre critères requièrent des échelles d'utilité de plus en plus raffinées : une échelle ordonnée suffit pour le critère de Wald, la notion de différence entre utilités est nécessaire pour le critère de Savage, et les critères d'Hurwicz et Laplace nécessitent de calculer des moyennes.

Critère *maximin* ou critère de Wald

Le critère *maximin*, ou critère de Wald [Wal50], ordonne les actions selon l'utilité de leur conséquence obtenue dans le pire cas, c'est à dire lorsque l'état du monde est le plus défavorable. Si, par exemple, le problème de décision est représenté par la matrice :

$$\begin{array}{c c} & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \left[\begin{array}{cc} 1 & 3 \\ 4 & 0 \end{array}\right] \end{array}$$

a_1 est préférée à a_2, car $u_W(a_1) = 1 > 0 = u_W(a_2)$.

Le critère *maximin* peut se révéler peu discriminant, comme le montre l'exemple suivant :

$$\begin{array}{cc} & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{bmatrix} 2 & 1 \\ 1 & 100 \end{bmatrix} \end{array}$$

où les deux actions sont trouvées équivalentes, alors que la seconde “semble” bien meilleure. Le critère *maximin* peut être raffiné, par exemple en utilisant le critère *leximin* de Moulin [Mou88] qui ordonne les conséquences de chaque action par ordre croissant d'utilité, puis compare les utilités des conséquences de chaque action, une par une : d'abord les utilités minimales, puis les suivantes et ainsi de suite. Dans le dernier exemple, $(1, 2) <_{lex} (1, 100)$, donc a_2 est préférée à a_1. Le critère *leximin* permet de définir un préordre complet, deux actions étant équivalentes si et seulement si les vecteurs ordonnés des utilités de leurs conséquences sont égaux. Un autre ordre strict sur les vecteurs, pouvant être utilisé pour comparer des actions a été étudié par Fargier, Lang et Schiex [FLS93] : l'ordre *discrimin*[1]. Pour deux actions a_1 et a_2, l'ordre *discrimin* compare les actions par leurs pires conséquences dans les états où ces actions diffèrent, uniquement : $a_1 >_{discr} a_2$ ssi $min_{s \in \delta(a_1,a_2)} a_1(s) > min_{s \in \delta(a_1,a_2)} a_2(s)$, où $\delta(a_1, a_2) = \{s \in S, a_1(s) \neq a_2(s)\}$. Notons que le préordre associé à *discrimin* n'est que partiel.

Critère *minimax regret*

Ce critère ordonne les actions suivant leur *regret maximal*, que l'on cherche à minimiser (Savage [Sav51]). Il transforme les utilités des actions en *regrets*, où le regret associé à une action dans un état du monde donné est la différence entre l'utilité de la conséquence que l'on obtiendrait en appliquant la *meilleure* action (pour cet état), et celle que l'on obtient effectivement.

$$\begin{array}{ccccc} & \text{Utilités} & & & \text{Regrets} \\ & \begin{array}{cc} s_1 & s_2 \end{array} & & & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{bmatrix} 0 & 100 \\ 1 & 1 \end{bmatrix} & \rightarrow & \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{bmatrix} 1 & 0 \\ 0 & 99 \end{bmatrix} \end{array}$$

La meilleure action est a_1, qui minimise le regret maximal. Notons que *minimax regret* n'est pas un raffinement de *maximin* (ici, les deux critères donnent des résultats opposés).

Critère de Hurwicz

Les critères *maximin* et *minimax regret* sont tous deux très prudents, puisqu'ils se focalisent sur le pire cas (le pire “regret” pour le critère de Savage). Le critère de Hurwicz [Hur51] autorise une compensation entre l'utilité obtenue dans le pire cas et l'utilité obtenue dans le meilleur cas, par l'intermédiaire d'un indice de pessimisme $\alpha \in [0, 1]$. Soient m_i et M_i les utilités respectivement de la pire et de la meilleure conséquence possible de a_i,

$$H_\alpha(a_i) = \alpha \cdot m_i + (1 - \alpha) \cdot M_i.$$

Lorsque $\alpha = 1$, on retrouve le critère *maximin*.

1. Cet ordre avait déjà été évoqué par Cohen et Jaffray [CJ80] dans un article proposant une axiomatisation des critères de décision dans l'incertain, rationnels, et montrant comment ils peuvent être *approchés* par des critères ne prenant en compte que les conséquences extrêmes des actions.

Critère de Laplace ou de Bernoulli

Le principe de ce critère consiste à supposer que, dans un cas d'ignorance totale sur l'état réel du monde, on peut associer à chaque action la *moyenne* des utilités de ses conséquences possibles (puisque rien ne nous permet de supposer qu'un état donné a plus de "chances" qu'un autre d'être l'état réel du monde) :

$$\text{si } S = \{s_1, \ldots, s_n\}, \text{ alors } u(a_i) = \frac{u_{i1} + \ldots + u_{in}}{n}$$

1.2.2 La caractérisation axiomatique de Milnor pour ces différents critères de décision dans l'incertain non probabilisé

Plusieurs critères de choix peuvent donc être utilisés pour la décision dans l'incertain. Lequel doit-on utiliser lorsque l'on se trouve face à un problème donné ? Ce problème du choix du critère est d'autant plus important à traiter que les différents critères peuvent classer de manière totalement contradictoire les différentes alternatives :

Exemple 1.2.1 (Luce et Raiffa [LR57])
Comparaison des différents critères :

$$\begin{array}{cc} & \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} & \begin{bmatrix} 2 & 12 & -3 \\ 5 & 5 & -1 \\ 0 & 10 & -2 \end{bmatrix} \end{array}$$

Sur cet exemple, les différents critères donnent les classements suivants :
- maximin : $a_2 \succ a_3 \succ a_1$,
- Savage : $a_1 \succ a_3 \succ a_2$,
- Hurwicz ($\alpha = 3/4$*) :* $a_3 \succ a_1 \succ a_2$,
- Laplace : $a_1 \succ a_2 \succ a_3$.

Cette sensibilité très grande de la solution du problème de décision vis-à-vis du critère choisi rend nécessaire une axiomatisation des différents critères : Milnor [Mil54] a proposé une caractérisation axiomatique des critères de décision dans l'incertain non probabilisé. Cette caractérisation est fondée sur dix axiomes. A partir du choix des axiomes que l'on souhaite voir respectés par la relation de préférence entre actions, certains des critères précédents vont être éliminés ou au contraire confortés pour représenter les préférences de l'agent. Les problèmes de décision sous incertitude seront considérés sous forme matricielle, les colonnes représentant les états possibles du monde et les lignes les actions à la disposition de l'agent. Cette caractérisation nécessite la définition suivante :

Définition 1.2.1 *Domination*
L'action a_i *domine fortement (resp. faiblement) l'action* a_j *si et seulement si,*

$$\forall s \in S, u(a_i(s)) > u(a_j(s)) (resp. \forall s \in S, u(a_i(s)) \geq u(a_j(s))).$$

Deux actions sont dites équivalentes si et seulement si elles se dominent faiblement l'une l'autre ($\forall s \in S, u(a_i(s)) = u(a_j(s))$*).*

Les axiomes de Milnor

1. **Ordre.** Toutes les actions doivent être ordonnées, i.e. le préordre entre les actions est complet.
2. **Symétrie.** L'ordre doit être indépendant de l'étiquetage des lignes et des colonnes.
3. **Domination forte.** a' est préférée à a'' si a' domine fortement a''.
4. **Continuité.** Si tout élément a'_i d'une suite convergente (a') d'actions est préféré à tout élément a''_i d'une autre suite convergente, alors la limite de la suite (a'') n'est pas préférée à celle de (a').
5. **Linéarité.** L'ordre est invariant par transformation linéaire des utilités ($u' = \alpha \cdot u + \beta, \alpha > 0$).
6. **Ajout de lignes.** L'ajout d'une nouvelle action ne modifie pas l'ordre entre les autres actions.
7. **Linéarité par rapport aux colonnes.** L'ordre ne change pas si on ajoute une même constante à toutes les colonnes.
8. **Duplication de colonnes.** L'ordre ne change pas si on duplique une colonne dans la matrice.
9. **Convexité.** Si a' et a'' sont classées indifférentes par l'ordre, alors aucune des deux n'est préférée à l'action $1/2 \cdot a' + 1/2 \cdot a''$.
10. **Ajout de lignes affaibli.** Ajouter une nouvelle action faiblement dominée par toutes les autres actions ne modifie pas l'ordre entre les autres actions.

L'axiome 1 impose que toutes les actions (hypothétiques ou réellement disponibles) puissent être comparées. L'axiome 2 accorde à tous les états possibles du monde la même importance dans l'évaluation des actions. D'après l'axiome 3, si une action donne un meilleur résultat qu'une autre quel que soit l'état réel du monde, alors elle lui est préférée. Les axiomes 4, 5 et 9 contraignent les propriétés mathématiques des fonctions d'utilité. L'axiome 6 stipule que les actions peuvent être évaluées indépendamment les unes des autres. Cet axiome est violé par le critère de Savage comme nous l'avons déjà signalé.
L'axiome 7 indique que si on offre un "bonus" pour un état donné du monde (le même pour toutes les actions), on ne modifie pas l'ordre entre ces actions. Bien entendu, si le bonus est offert pour un état dans lequel l'action a_i donne sa pire conséquence, on change l'utilité (au sens de Wald ou de Hurwicz) de a_i, et pas forcément celle des autres actions : ainsi, le classement de a_i peut changer. Les critères de Savage et de Laplace, eux, respectent cet axiome : Savage parce que l'ajout de bonus ne change pas la matrice de regrets, et Laplace parce que son caractère additif fait que le bonus apporte la même contribution à toutes les actions.
L'axiome 8 stipule encore qu'en cas d'ignorance, les critères de choix doivent être insensibles à la "redondance" des informations. Si deux états du monde se comportent de manière identique pour toutes les actions, ils peuvent être rassemblés pour l'évaluation de ces actions. Le critère de Laplace ne satisfait pas ce principe, puisqu'un état décrit deux fois dans la matrice d'utilité a un "poids" deux fois plus fort que s'il n'est décrit qu'une fois. L'axiome 10 est une version affaiblie de l'axiome 6.

Le tableau 1.1 page suivante résume la caractérisation des différents critères exposés précédemment. Un $\times$ ou un $\otimes$ signifie que l'axiome est vérifié par le critère correspondant.
Aucun des quatre critères ne vérifie les dix axiomes à la fois. En fait, ces dix axiomes sont même incompatibles entre eux. Le choix d'un sous-ensemble d'axiomes (compatibles entre eux) va permettre

Axiome	*maximin*	Savage	Hurwicz	Laplace
1. Ordre	$\otimes$	$\otimes$	$\otimes$	$\otimes$
2. Symétrie	$\otimes$	$\otimes$	$\otimes$	$\otimes$
3. Dom. forte	$\otimes$	$\otimes$	$\otimes$	$\otimes$
4. Continuité	$\times$	$\otimes$	$\otimes$	$\otimes$
5. Linéarité	$\times$	$\times$	$\otimes$	$\times$
6. Aj. ligne	$\otimes$		$\otimes$	$\otimes$
7. Lin. colonnes		$\otimes$		$\otimes$
8. Dup. colonnes	$\otimes$	$\otimes$	$\otimes$	
9. Convexité	$\otimes$	$\otimes$		$\times$
10. Aj ligne aff.	$\times$	$\otimes$	$\times$	$\times$

TAB. 1.1 – *Caractérisation axiomatique des différents critères de décision dans l'incertain non probabilisé.*

de déterminer l'utilisation de l'un ou l'autre des critères. L'ensemble des axiomes pointés par un $\otimes$ dans une colonne suffit à caractériser le critère correspondant : si un ordre entre les actions satisfait tous les axiomes d'une colonne pointés par un $\otimes$ (plus d'autres éventuellements, pointés par un $\times$) alors il est représentable par le critère correspondant à la colonne.
Ainsi, si on est prêt à abandonner l'axiome 6 (ajout de ligne), par exemple, c'est le critère de Savage que l'on choisira. Si l'on veut garder les axiomes 1, 2, 3, 4, 6 et 7 on est conduit à choisir le critère de Laplace (qui respecte aussi 5, 9 et 10).

1.2.3 Le modèle de Brafman et Tennenholtz pour la décision qualitative dans l'incertain

Dans ce paragraphe, nous allons revenir sur les critères de décision dans l'incertain non probabilisé, *maximin* et *minimax regret*. Nous avons présenté dans le paragraphe précédent les axiomes proposés par Milnor [Mil54], permettant de caractériser ces critères. Cependant, dans son axiomatisation, Milnor suppose que la matrice d'utilité $u(a_i, s_j)$ du problème de décision est une donnée de ce problème. Nous allons exposer les travaux de Brafman et Tennenholtz [BT96], [BT97] qui se placent dans une optique différente en proposant une axiomatisation permettant, lorsque l'on choisit le critère que l'on souhaite adopter, de déterminer la matrice d'utilité. Ce type d'approche est dite *constructive*.

L'ensemble des états possibles du monde est noté S. Un sous-ensemble l de S est un *état de connaissance*. Soit L, l'ensemble de ces états de connaissance ($L = 2^S$). Parallèlement à la notion d'action hypothétique ils utilisent la notion d'état de connaissance hypothétique. Ensuite, ils définissent la notion de *police* $\mathcal{P} : L \to TO(\mathcal{A})$, où TO représente l'ensemble des ordres totaux sur $\mathcal{A}$, l'ensemble des actions (les résultats exposés tiennent également pour des *préordres totaux*).
Une police n'est rien d'autre qu'un rangement des actions par ordre de préférence!

Dans (Brafman et Tennenholtz [BT96]), les auteurs proposent une justification du critère *maximin*. Cette justification requiert la définition de *maximin représentabilité* :

Définition 1.2.2 *Maximin représentabilité*
La police $\mathcal{P} : L \to TO(\mathcal{A})$ *est* maximin-représentable *si et seulement si il existe une fonction d'utilité* u *sur* $\mathcal{A} \times S$ *(à valeurs dans une échelle totalement ordonnée), telle que* a *est préférée à* a' *dans l'état*

de connaissance l, suivant l'ordre $\mathcal{P}(l)$ si et seulement si :

$$min_{s \in PW(l)} u(a, s) > min_{s \in PW(l)} u(a', s).$$

En termes simples, une police $\mathcal{P}$ est maximin-représentable si et seulement si on peut trouver une fonction d'utilité sur $\mathcal{A} \times S$ telle que $\mathcal{P}$ puisse être représentée par le critère *maximin*.
Remarquons que ce modèle fait intervenir une fonction d'utilité sur $\mathcal{A} \times S$, et non sur un ensemble de conséquences (ce dernier cas peut être vu comme un cas particulier du précédent).

Brafman et Tennenholtz [BT96] proposent de construire, à partir d'un ensemble d'ordres totaux $\{\succ_W, W \subseteq S\}$ sur $\mathcal{A}$, un ordre partiel $\succ$ sur $\mathcal{A} \times S$. Cet ordre partiel étend en particulier les ordres $\succ_{\{s\}}$ pour $s \in S$: $a \succ_{\{s\}} a' \Leftrightarrow (a, s) \succ (a', s)$.
Dans [BT96] les auteurs prouvent que si la famille d'ordres totaux vérifie une propriété simple de *fermeture sous union* ($a \succ_V a'$ et $a \succ_W a' \Leftrightarrow a \succ_{V \cup W} a'$) et une certaine forme de transitivité dite *transitivité faible*, alors $\succ$ est *maximin-représentable*.

En prouvant ce résultat, les auteurs donnent une méthode de construction de la fonction d'utilité u sur $\mathcal{A} \times S$.
En outre, dans (Brafman et Tennenholtz [BT97]), ils montrent qu'une police est *maximin-représentable* si et seulement si elle est *minmax-regret représentable*. Les matrices d'utilités obtenues dans les deux représentations sont bien entendu différentes, mais ce résultat signifie que dans le cas où les utilités d'un problème de décision sous incertitude doivent être déterminées à partir des préférences d'un agent, les deux critères pourront être indifféremment utilisés pour résoudre le problème de décision.
Enfin, dans le même article, les auteurs étendent leurs résultats aux cas où les connaissances d'un agent sont représentées par des préordres sur S (et non de simples sous-ensembles). Dans ce cas, ils prouvent que si les deux propriétés évoquées plus haut sont toujours respectées, le critère *maximin* peut encore être invoqué, en limitant les ensembles d'états possibles aux états “les plus plausibles” suivant le préordre sur S.

Le même type d'approche, visant à ne prendre en compte que les états les plus plausibles pour résoudre un problème de décision sous incertitude, a été adopté par Boutilier [Bou94], ou Tan et Pearl [TP94], dans leurs approches logiques que nous allons décrire dans le chapitre 4 page 53.
Au contraire, Dubois et Prade [DP95b] proposent deux critères de décision qualitatifs qui ne se limitent pas aux états les plus plausibles (voir chapitre 2 page 21). En contrepartie, comme nous le verrons dans la seconde partie de cette thèse, leur méthode suppose que les ordres utilisés pour représenter les connaissances et préférences d'un agent sont *commensurables*, i.e. représentables sur une même échelle. L'un des deux critères est *maximin-représentable* alors que l'autre est *maximax-représentable* (propriété duale de la précédente, également présentée par Brafman et Tennenholtz).

1.3 Décision dans le risque

1.3.1 Présentation du critère de l'utilité espérée

Les problèmes de décision sous risque forment un sous-ensemble des problèmes de décision sous incertitude. Dans les problèmes de décision sous risque, on assimile les actions à des *loteries* (probabilistes), dont les lots sont les conséquences éventuelles des actions. Les probabilités des différentes conséquences (les lots) peuvent être déterminées à partir de données statistiques (si on dispose de telles données, c'est-à-dire si le problème de décision est répété fréquemment), ou à partir de considérations de symétrie. Les jeux de hasard de type roulette, jeux de dés, black jack,... sont des exemples typiques de problèmes de décision sous risque.

Exemple 1.3.1 *Le jeu de roulette*
Le jeu bien connu de roulette fait intervenir une roue comportant 37 cases (numérotées de 0 à 36). Le problème de décision consiste à miser une certaine somme (pouvant être nulle) selon l'un des modes suivants :

- *Pari sur un nombre de 1 à 36. Si le numéro choisi sort, on récupère 36 fois la mise, sinon on la perd.*
- *Pari sur* pair *ou* impair. *Si le numéro sortant possède la même parité (sauf 0), on récupère deux fois la mise, sinon on la perd.*
- *Pari sur* rouge *ou* noir. *Les numéros de 1 à 36 sont répartis équitablement dans les deux couleurs. Si le numéro sortant est de la couleur choisie, on gagne deux fois la mise, sinon on la perd.*
- *Pari sur l'une des trois douzaines : 1-12, 13-24, 25-36. Si le numéro sortant appartient à la douzaine choisie, on gagne 3 fois la mise, sinon on la perd.*

Si le zéro sort, la mise est perdue.

Dans les problèmes de décision sous risque les actions sont évaluées par l'espérance mathématique d'une fonction d'utilité sur les conséquences. Nous supposerons que la fonction d'utilité est égale au gain, mais ce choix est arbitraire : dans le modèle la fonction d'utilité d'un agent sur les conséquences exprime son attitude par rapport au risque et n'est pas forcément directement proportionnelle au gain. Les *utilités espérées* des différentes actions disponibles servent à classer directement les actions disponibles. Si on parie sur un numéro, on a une chance sur 37 (n'oublions pas le zéro!) qu'il sorte, et 36 chances sur 37 qu'il ne sorte pas. Si le numéro choisi sort, on gagne 35 fois la mise initiale m (36 fois en fait, mais il faut tenir compte de la somme misée). S'il ne sort pas, le "gain" est négatif et vaut $-m$ (somme misée et perdue). La formule de l'espérance mathématique du gain G est $E(G) = \sum p(G) \cdot G$. En l'appliquant ici pour obtenir l'utilité espérée de l'action $Pari(i)$ (pari sur le nombre i), on obtient : $UE(Pari(i)) = p(i) \cdot 35m + p(\bar{i}) \cdot (-m)$, soit $UE(Pari(i)) = (1/37) \cdot 35m + (36/37) \cdot (-m) = -(1/37)m$.
On peut faire le même calcul pour tous les autres paris, pour lesquels on trouve le même gain espéré, négatif si la somme misée est positive. La meilleure décision à prendre pour ce problème (si on suit la théorie "rationnelle" de l'utilité espérée), est de ne pas parier, ou mieux encore, de parier une somme négative (c'est-à-dire, de se mettre à la place du casino!).

Dans le prochain paragraphe nous allons décrire plus en détail la notion de loterie, puis l'axiomatique proposée par (von Neumann et Morgenstern, 1944 [vNM44]) pour la décision sous risque, permettant de justifier l'utilisation du critère de l'utilité espérée pour le choix de décisions optimales.

1.3.2 Loteries et utilité espérée : axiomatisation

Loteries

Une *loterie*, dans l'approche de von Neumann et Morgenstern [vNM44] est simplement une distribution de probabilité p sur un ensemble de conséquences X, pouvant être fini ou infini.

Rappelons que dans le cadre de la décision sous incertitude, une action a est une application d'un *ensemble d'états possibles du monde* S vers l'ensemble de conséquences X. Dans le cas particulier de la décision sous risque, l'incertitude sur l'état du monde est modélisée par une distribution de probabilité p (connue a priori) sur S. Il apparaît que l'on peut aisément associer à chaque action a une distribution de probabilité p_a sur les conséquences (une loterie).
Celle-ci est simplement définie par :

Définition 1.3.1 *Loterie associée à une action*

$$\forall x \in X, p_a(x) = P(a^{-1}(x)).$$

Dans le cas où S est fini, $P(a^{-1}(x)) = \Sigma_{s \in S, a(s)=x} p(s)$.

Ainsi, dans le cadre de la décision sous risque, comparer des actions revient à comparer des distributions de probabilité sur les conséquences, donc des loteries.

L'ensemble des loteries sur l'ensemble de conséquences X peut être défini à partir des éléments de X et d'une opération de combinaison de loteries construite à partir de deux opérateurs, $\oplus$ et $\odot$:

Définition 1.3.2 (Von Neumann et Morgenstern [vNM44])
$(\mathcal{L}(X), \oplus, \odot)$*, l'ensemble des loteries sur* X*, est défini par :*

1. $X \subseteq \mathcal{L}(X)$
2. *Pour tout couple* $(l, l') \in \mathcal{L}(X)^2$ *et pour tout* $\alpha, (0 < \alpha < 1)$*, on définit une opération de combinaison* $\mathcal{L}(X)^2 \to \mathcal{L}(X)$, $(l, l') \to l'' = (\alpha \odot l) \oplus ((1 - \alpha) \odot l')$.
3. *Toute combinaison de loteries est une loterie.*

Toute conséquence $x \in X$ est appelée *loterie élémentaire* (une loterie élémentaire x donne la conséquence x avec la probabilité 1). Attention, les éléments de $\mathcal{L}(X)$ ne sont pas des réels, ainsi $\odot$ et $\oplus$ ne sont pas les opérations classiques sur les réels, toutefois elles vérifient les mêmes propriétés d'associativité, distributivité...
Un exemple de loterie élémentaire est le jeu de pile ou face, dans lequel l'action de parier sur *pile* peut être assimilée à une loterie sur les conséquences *gagné* et *perdu* : $l = (0,5\odot$ *gagné*$) \oplus (0,5 \odot perdu)$.
La loterie en question est définie indépendamment des utilités associées à *gagné* et *perdu*.

Axiomatisation

Dans les problèmes de décision sous risque, les loteries représentant les actions peuvent être extrêmement complexes, ainsi les préférences entre de telles loteries peuvent être difficiles à déterminer. L'idée qui sous-tend l'axiomatique de von Neumann et Morgenstern est de permettre à un agent d'exprimer ses préférences simplement par une fonction d'utilité sur les loteries élémentaires (l'ensemble des conséquences, X). Si les préférences de l'agent sur l'ensemble des loteries respectent une série d'axiomes, alors elles peuvent être résumées par l'espérance mathématique de la fonction d'utilité sur les conséquences.

Soit $\preceq$ une relation d'ordre sur $\mathcal{L}(X)$. l, l' et l'' sont des loteries, $\alpha \in \ ]0,1[$. Von Neumann et Morgenstern ont proposé l'ensemble d'axiomes suivant, divisé en trois groupes, et portant sur la relation $\preceq$:

A : Préordre total

A.1 $\preceq$ est un préordre total sur $\mathcal{L}(X)$: il est réflexif, transitif et tous les éléments de $\mathcal{L}(X)$ sont comparables deux à deux. $\prec$ est la partie stricte de $\preceq$.

B : Ordre et combinaison

B.1 $l \prec l' \Rightarrow l \prec [(\alpha \odot l) \oplus ((1-\alpha) \odot l')]$ et $[(\alpha \odot l) \oplus ((1-\alpha) \odot l')] \prec l'$

B.2 $l \prec l' \prec l'' \Rightarrow$
$\exists \alpha \in \ ]0,1[$ tel que $[(\alpha \odot l) \oplus ((1-\alpha) \odot l'')] \prec l'$
et $\exists \beta \in \ ]0,1[$ tel que $l' \prec [(\beta \odot l) \oplus ((1-\beta) \odot l'')]$

C : Algèbre des combinaisons

C.1 $(\alpha \odot l) \oplus ((1-\alpha) \odot l') = ((1-\alpha) \odot l') \oplus (\alpha \odot l)$

C.2 $(\alpha \odot ((\beta \odot l) \oplus ((1-\beta) \odot l')) \oplus ((1-\alpha) \odot l') = (\gamma \odot l) \oplus ((1-\gamma) \odot l')$, où $\gamma = \alpha\beta$

Le premier axiome indique que toutes les actions peuvent être comparées vis-à-vis des préférences de l'agent. L'axiome B.1 précise que si deux conséquences ne sont pas équivalentes, alors toute combinaison des deux est comprise strictement entre les deux (au sens de la relation de préférence). C'est un axiome d'indépendance. L'axiome B.2 est un axiome de continuité qui stipule que si deux loteries ne sont pas équivalentes, on peut trouver une combinaison des deux qui soit arbitrairement proche de l'une ou l'autre. Les axiomes du groupe C montrent comment les loteries peuvent être combinées.

Le théorème suivant affirme que pour toute relation d'ordre $\preceq$ entre les loteries qui satisfait ce groupe d'axiomes il existe une fonction d'utilité sur les loteries permettant de la représenter. L'utilité d'une loterie quelconque est égale à l'espérance mathématique d'une fonction d'utilité sur l'ensemble des conséquences.

Théorème 1.3.1 *Axiomatisation de l'utilité espérée*
Soit X un ensemble de conséquences et $\preceq$ une relation d'ordre sur $\mathcal{L}(X)$ vérifiant les axiomes A.1, B.1, B.2, C.1 et C.2. Il existe une fonction d'utilité réelle u sur $\mathcal{L}(X)$ telle que

$$\forall l, l', l \preceq l' \Leftrightarrow u(l) \leq u(l').$$

$$u(l) = \sum_{x \in X} p_l(x) \cdot u(x),$$

où p_l est la distribution de probabilités sur X associée à l.

Remarque : u n'est pas unique. Toute transformation affine $u' = a.u + b$ avec $a > 0$ représente la même relation de préférence.

1.4 Conclusion

La principale limitation de la théorie de von Neumann et Morgenstern provient du fait qu'elle ne se préoccupe que de la décision sous risque. Dans ce cadre, l'incertitude est représentée par une distribution *objective* de probabilités sur l'ensemble des états possibles du monde. Cela veut dire que ces probabilités, déterminées par les conditions extérieures du problème de décision sont connues par le "décideur". Cette théorie n'est donc applicable que lorsque les probabilités sont réellement objectives.
Le problème se complique lorsque ces probabilités ne sont pas disponibles ou lorsqu'elles ne peuvent être déterminées de manière objective. Dans ce cas, l'incertitude ne peut être mesurée par la simple observation de phénomènes physiques extérieurs au décideur (statistiques, symétries du problème...). C'est au contraire l'agent (éventuellement un "expert") qui devra délivrer ses "impressions" à propos des états du monde qu'il considère plausibles ou non. Dans ce cas, on n'est plus en face d'un problème de décision *sous risque*, mais sous *incertitude généralisée* (les problèmes de décision sous risque sont des cas particuliers de problèmes de décision sous incertitude). Dans les deux cas, les connaissances sont résumées en une distribution de probabilité, *objective* lorsqu'on est "sous risque" et *subjective* lorsqu'on est sous "incertitude généralisée".

Le problème revient donc à faire révéler au décideur les probabilités *subjectives* qu'il attache aux états possibles du monde, avant d'appliquer éventuellement le critère de l'utilité espérée au nouveau problème de décision sous risque obtenu. Les méthodes proposées classiquement pour révéler ces probabilités procèdent d'une analyse de la relation de préférence de l'agent entre les actions mises à sa disposition.

Une première méthode, proposée par Anscombe et Aumann [AA63] consiste en une extension du cadre de von Neumann et Morgenstern à la décision sous incertitude. Ces auteurs ont proposé de considérer, en plus des loteries classiques (pour lesquelles les probabilités sur les conséquences sont objectives), d'autres loteries qu'ils nomment "horse lotteries". Ces "horses lotteries" ont pour conséquences des loteries classiques (éventuellement des conséquences au sens classique, cas particuliers de loteries). Les "horse lotteries" et les loteries classiques peuvent être composées : une loterie classique peut avoir à son tour des "horses lotteries" comme conséquences.
Anscombe et Aumann supposent ensuite que la relation de préférence du décideur entre toutes les loteries (objectives et subjectives) satisfait, en plus des axiomes de von Neumann et Morgenstern deux axiomes supplémentaires :

– Le premier concerne une forme de dominance entre loteries subjectives (horse lotteries). Si deux loteries subjectives ne diffèrent que par l'un de leur lots (les lots des lotteries subjectives sont des loteries classiques) alors l'ordre de préférence entre ces loteries subjectives est le même qu'entre les deux loteries classiques qui diffèrent.

– Le second précise que si une loterie composite fait intervenir pour une phase une loterie classique et pour une autre une "horse lottery", peu importe l'ordre dans lequel est définie cette loterie composite. Ce peut être une "horse lottery" dont les lots sont des loteries classiques, ou au contraire une loterie classique dont les lots sont des "horses lotteries", les deux loteries obtenues sont équivalentes.

Grâce à ces deux axiomes supplémentaires, on peut déduire une distribution de probabilité (subjective), représentant chaque "horse lottery" et les préférences du décideur sont à nouveau représentables par un critère d'utilité espérée.

La théorie de Anscombe et Aumann représente une avancée par rapport à celle de von Neumann et Morgenstern, puisqu'elle étend l'usage du critère de l'utilité espérée à la décision sous incertitude. Seulement, elle fait aussi intervenir la notion de loterie, et pose encore l'usage des probabilités pour représenter l'incertitude comme un postulat. L'approche que nous allons décrire dans le chapitre suivant est plus riche, puisqu'elle considère directement une relation de préférence entre actions et non entre loteries. C'est seulement à partir de ces préférences que sont déterminées les probabilités (subjectives ou objectives) des états du monde. Ces probabilités ne sont pas des données du problèmes comme dans l'approche de von Neumann et Morgenstern (et de Anscombe et Aumann, dans une moindre mesure).

Chapitre 2

Le critère de l'utilité espérée et ses variantes

2.1 Introduction

Dans ce chapitre nous allons décrire plusieurs théories (et critères de décision associés) pour la décision dans l'incertain. Certaines de ces théories sont des extensions de celle de von Neumann et Morgenstern (cf. Chap. 1), d'autres ont un cadre d'application différent (description qualitative ou ordinale des problèmes de décision, problèmes de décision sans compensation entre les conséquences possibles).

Ce chapitre est composé comme suit :

– Nous décrirons en premier lieu l'approche de Savage [Sav54] : des données statistiques sont rarement disponibles. Pourtant, la théorie de l'utilité espérée est un outil puissant et attrayant pour la représentation et la résolution de problèmes de décision sous incertitude. C'est pourquoi elle a été étendue au cas où l'incertitude n'est pas de nature statistique. Dans ce cas, on utilise toujours une distribution de probabilité pour représenter les connaissances de l'agent, mais celle-ci est maintenant *subjective*. Elle est construite, au même titre que la fonction d'utilité, directement à partir des préférences entre actions, en supposant toujours que ces préférences sont représentables par comparaison de l'utilité espérée des différentes actions.

– La théorie de l'utilité espérée est basée sur certains postulats, or ces postulats ne sont pas toujours validés par tous les agents faisant face à un problème de décision sous incertitude. La théorie de l'utilité espérée peut aussi être mise en défaut sur des exemples précis de problèmes de décision. De tels exemples peuvent être trouvés, dans lesquels l'incertitude comporte à la fois une composante statistique et une composante non-statistique. Des théories basées sur les *fonctions de croyances* (par exemple) ont été proposées pour traiter ce type de cas.

– Enfin, l'incertitude et les préférences d'un agent ne sont pas toujours modélisables par des probabilités numériques sur les états possibles du monde et des utilités additives sur les conséquences. L'agent peut être seulement capable d'ordonner (en termes de plausibilité, ou de préférence) les états du monde et les conséquences. A la fin de ce chapitre nous exposerons une théorie *qualitative* de la décision utilisant des *distributions de possibilité* à la fois pour représenter les connaissances et les préférences d'un agent. Nous verrons que cette théorie utilise des données *ordinales*, et non additives, et qu'elle rejette la notion de moyenne, centrale à la théorie de l'utilité espérée. Une grande partie de nos travaux concerne cette théorie, c'est pourquoi un chapitre entier (Partie II, chap. 1 page 73) sera consacré à la description de l'axiomatique que nous avons proposée pour elle, dans la seconde partie de cette thèse.

2.2 Décision dans l'incertain et utilité espérée : l'approche de Savage

Comme nous l'avons déjà dit, le point faible de la théorie de [vNM44] est qu'elle suppose que l'incertitude pesant sur un problème de décision est représentée par une distribution de probabilité connue a priori. Ainsi, on peut déterminer la fonction d'utilité du décideur sur l'ensemble des conséquences X, à partir d'une distribution de probabilité, et de ses préférences sur l'ensemble des loteries.
L'approche de Savage [Sav54] est plus générale :

- Elle permet, à partir d'une relation de préférence entre les actions (et non les loteries), de déterminer à la fois la fonction d'utilité du décideur sur les conséquences, mais aussi "sa" distribution de probabilité *subjective* sur les états possibles du monde.

- Elle utilise la notion d'action, plus générale que celle de loterie (de toute action on peut déduire une loterie, mais plusieurs actions différentes peuvent générer les mêmes loteries).

Cette approche ressemble à celle de Anscombe et Aumann, à ceci près que la nature probabiliste de l'incertitude n'est pas supposée a priori mais découle d'axiomes concernant les préférences entre les actions. Du fait que les axiomes proposés par Savage concernent directement les actions, ils peuvent être vérifiés par l'observation du comportement du décideur sur des problèmes réels.

S et X sont encore une fois deux ensembles, respectivement d'états et de conséquences. Savage suppose que S est infini. Cette supposition est nécessaire pour pouvoir construire une distribution de probabilité représentant l'incertitude de l'agent. X peut être fini ou infini. $\mathcal{A} = X^S$ constitue l'ensemble des actions "potentielles", applications associant une conséquence à chaque état possible du monde.
Dans un problème de décision, les actions réellement disponibles ne constituent en général qu'un sous-ensemble de $\mathcal{A}$, néanmoins nous supposerons que le décideur est capable d'ordonner toutes les actions potentielles. Sa relation de préférence est représentée par le préordre $\preceq$ sur $\mathcal{A}$. La première hypothèse posée par Savage est que $\preceq$ est un préordre *complet* sur $\mathcal{A}$: toutes les actions peuvent être comparées (rangées).

Sav1 Rangement
$(\mathcal{A}, \preceq)$ est un préordre complet, i.e. $\preceq$ est réflexif, transitif et complet.

Deux sous-familles d'actions sont particulièrement utiles pour déterminer d'une part la fonction d'utilité du décideur sur les conséquences, et d'autre part sa distribution de probabilité subjective sur les états possibles du monde : les actions *constantes* et *binaires*, respectivement.

Les conséquences des actions sont généralement notées $x, y, x', y' \dots$ (éléments de X). Les actions constantes donnent la même conséquence, quel que soit l'état réel du monde. Pour distinguer une action constante de son unique conséquence x, on la note en gras (**x**) : pour tout $x \in X$, **x** représente l'action définie par $\forall s \in S$, $\mathbf{x}(s) = x$. L'ensemble des actions constantes est noté $\mathcal{C}$.

Puisque $\preceq$ est un préordre complet sur $\mathcal{A}$, il en est aussi un sur $\mathcal{C} \subseteq \mathcal{A}$. Comme $\mathcal{C}$ et X sont bijectivement reliés, on peut définir, à partir de $\preceq$, un préordre complet sur X : $\leq_P$.

Définition 2.2.1 *Préférences sur X induites par les préférences sur $\mathcal{A}$*
$\forall x, y \in X$, si pour tout $s \in S$, $\mathbf{x}(s) = x$ et $\mathbf{y}(s) = y$, alors $x \leq_P y \Leftrightarrow \mathbf{x} \preceq \mathbf{y}$.

Afin d'éviter le cas dégénéré où toutes les conséquences sont équivalentes, auquel cas toutes les actions le sont également, l'axiome suivant est introduit (les axiomes ne sont pas présentés dans le même ordre que dans (Savage [Sav54]) mais par souci de clarté nous conservons la même numérotation) :

Sav5 Non trivialité
$\exists x, y \in X, y <_P x$, où $<_P$ représente la partie stricte de $\leq_P$.

La notion d'*action binaire* est utile afin de déterminer les probabilités attachées par le décideur aux différents états possibles du monde. Une action binaire est une action qui a au plus deux conséquences possibles, x et y telles que $y <_P x$.
Pour toute action binaire, il existe un ensemble d'états $A \subseteq S$ (dans le cadre de la décision sous incertitude, un ensemble d'états est appelé *événement*), tel que la conséquence de l'action binaire est x pour tout élément s de A, et y pour tout élément de $\overline{A}$. Cette action binaire est notée xAy.

Définition 2.2.2 *Action binaire*
Une action binaire, notée xAy, $x, y \in X, A \subseteq S$ peut être vue comme un pari sur l'événement A, donnant la conséquence x si A "arrive", et y sinon :

$$\forall s \in A, xAy(s) = x \text{ et } \forall s \in \overline{A}, xAy(s) = y.$$

Définissons maintenant $\leq_V$, relation de *vraisemblance* entre événements, par :

$$A \leq_V B \Leftrightarrow \forall x, y \in X, y <_P x, xAy \preceq xBy.$$

$A \leq_V B$ revient à dire que tout *pari* sur B est préféré à un pari aux conséquences identiques sur A.

Il va sans dire que, dans le cas général, rien ne garantit que $\leq_V$ soit un préordre complet : changer les conséquences peut modifier l'ordre entre les paris. Afin de rendre $\leq_V$ complet, Savage propose l'axiome suivant :

Sav4 Projection sur l'ensemble des événements
$\forall x, y, x', y' \in X, y <_P x, y' <_P x', \forall A, B \subseteq S, xAy \preceq xBy \Leftrightarrow x'Ay' \preceq x'By'$.

Cet axiome assure que la restriction de la relation de préférence sur les actions aux événements (paris sur les événements) est bien définie et ne dépend que de l'ordre entre les conséquences des paris.

La notion d'action binaire est un cas particulier de la notion d'*action mixte* :

Définition 2.2.3 *Action mixte*
Si $\mathbf{f}$ et $\mathbf{g}$ sont deux actions, et $A \subseteq S$ est un événement, l'action mixte $\mathbf{f}A\mathbf{g}$ est définie par : $\forall s \in A$, $\mathbf{f}A\mathbf{g}(s) = \mathbf{f}(s)$ et $\forall s \in \overline{A}$, $\mathbf{f}A\mathbf{g}(s) = \mathbf{g}(s)$.

Une action binaire est donc une action mixte "constante". L'axiome suivant est un axiome clé de la théorie de Savage. C'est lui qui permet de définir le conditionnement des préférences par rapport à un événement.

Sav2 Principe de la chose certaine [1]
$\forall$ **f**, **g**, **h**, **h'** $\in \mathcal{A}, \forall A \subseteq S$, **f**$A$**h** $\preceq$ **g**A**h** $\Leftrightarrow$ **f**A**h'** $\preceq$ **g**A**h'**.

Cet axiome stipule que si deux actions **f**A**h** et **g**A**h** donnent les mêmes conséquences sur un événement donné alors leur classement relatif ne dépend pas de ces conséquences.

Le "principe de la chose certaine" permet de définir une relation de *préférence conditionnelle* par rapport à un événement A par :

Définition 2.2.4 *Préférence conditionnelle*
f *est préférée à* **g** *conditionnellement à* A*, noté (***f** $\succeq$ **g**$)_A$*, si et seulement si*
$\forall$ **h** $\in \mathcal{A}$, **f**A**h** $\succeq$ **g**A**h**.

Cette définition est bien fondée, grâce au "principe de la chose certaine". Qui plus est, la propriété (**f** $\succeq$ **g**$)_A$ ne dépend pas du choix de **h**, et la relation de préférence conditionnelle est un préordre complet sur $\mathcal{A}$.
Un ensemble d'événements peut être distingué, celui conditionnellement auxquel toutes les actions sont équivalentes : les *événements nuls*. Un événement A est dit nul si et seulement si, pour tout triplet d'actions **f**, **g** et **h**, **f**A**h** $\preceq$ **g**A**h**. On peut montrer que A est nul si et seulement si $A \sim_L \emptyset$.

Si on restreint la relation de préférence conditionnelle par rapport à un événement aux actions constantes, le préordre obtenu doit coïncider avec le préordre sur les conséquences (sauf si on conditionne par rapport à un événement nul, bien entendu). Cette propriété découle de l'axiome suivant :

Sav3 Conditionnement restreint aux actes constants :
Soient $x, y \in X$ et A un événement *non nul*. Soient les actions constantes **x** $\equiv x$ et **y** $\equiv y$. Alors, (**x** $\preceq$ **y**$)_A \Leftrightarrow x \leq_P y$.

Faisons une pause à cet endroit pour observer les propriétés de la relation $\leq_L$ sur les événements, induite par une relation $\preceq$ sur les actions, satisfaisant les axiomes **Sav1** à **Sav5**. Savage a montré qu'une telle relation sur les événements est une *relation de probabilité qualitative*, c'est à dire, vérifie les propriétés suivantes (de Finetti [dF37]) :

A1 $\leq_L$ est un préordre complet,

A2 $\emptyset <_L S$ (non trivialité),

A3 $\forall A \subseteq S, \emptyset \leq_L A$ (consistance),

P Si $A \cap (B \cup C) = \emptyset$, alors $B \leq_L C$ si et seulement si $A \cup B \leq_L A \cup C$ (additivité).

Une relation d'ordre $\leq_L$ sur les événements basée sur une mesure de probabilité P ($A \leq_L B \Leftrightarrow P(A) \leq P(B)$) est une relation de probabilité qualitative. Cependant, une relation de probabilité

1. Le nom original donné par Savage à cet axiome est *the Sure Thing principle* que nous traduisons par principe de la chose certaine, faute de mieux.

qualitative n'est pas forcément représentable par une distribution de probabilité comme l'on montré Kraft, Pratt et Seidenberg [KPS59].

Pour pouvoir déterminer la mesure de probabilité, Savage introduit l'axiome suivant, qui sous-entend que S est *infini* :

Sav6 Probabilité quantitative
Soient **f** et **g** $\in \mathcal{A}$ tels que **f** $\prec$ **g**, soit $x \in X$. Il existe une partition $\bigcup A_i$ de S telle que pour tout i, $xA_i\mathbf{f} \prec \mathbf{g}$ et $\mathbf{f} \prec xA_i\mathbf{g}$.

Sav6 implique une forme de "continuité" de l'espace d'états : il est possible de partitionner S en de nombreux sous-ensembles de probabilité très faible...

Alors, si $\preceq$ est une relation de préférence sur $\mathcal{A}$ satisfaisant **Sav1** à **Sav6**, elle est représentable par une fonction d'utilité EU sur $\mathcal{A}$ à valeurs réelles. Pour toute action **f**, $EU(\mathbf{f})$ est l'espérance mathématique d'une fonction d'utilité sur les conséquences, au sens d'une distribution de probabilité p sur S :

Théorème 2.2.1 *Axiomatisation de l'utilité espérée*
Soit S un ensemble infini d'états du monde, et X un ensemble de conséquences. X peut être infini, mais $\mathcal{A} \subset X^S$ est un ensemble fini d'actions potentielles et $\preceq$ est une relation de préférence sur $\mathcal{A}$ vérifiant les axiomes **Sav1** *à* **Sav6**.
Alors il existe une distribution de probabilité p, unique, sur S et une fonction d'utilité u (non unique[2]) sur X, telles que :

$$\forall \mathbf{f}, \mathbf{g} \in \mathcal{A}, \mathbf{f} \preceq \mathbf{g} \Leftrightarrow \sum_{s \in S} p(s) \cdot u(\mathbf{f}(s)) \leq \sum_{s \in S} p(s) \cdot u(\mathbf{g}(s)).$$

Le théorème précédent peut enfin être étendu à des actes quelconques en ajoutant un dernier axiome :

Sav7 Dominance
Soient **f** et **g** $\in \mathcal{A}, A \subseteq S$. $(\mathbf{f} \preceq \mathbf{g}(s))_A, \forall s \in A \Rightarrow (\mathbf{f} \preceq \mathbf{g})_A$.

Cet axiome précise que si toutes les conséquences possibles de l'action **g** sont préférées (ou équivalentes) à l'action **f**, alors **g** est préférée (ou équivalente) à **f**.

Corollaire 2.2.1 *Si $\preceq$ vérifie* **Sav1** *à* **Sav7** *alors le théorème précédent est valide pour des actes quelconques.*

Depuis l'époque de von Neumann et Morgenstern, le critère de l'utilité espérée a été équipé d'une base axiomatique très solide, que ce soit pour la décision sous risque ou sous incertitude. Cependant, plusieurs auteurs ont exhibé des exemples de problèmes de décision sous risque ou sous incertitude pour lesquels l'utilisation du critère de l'utilité espérée donne des résultats contre-intuitifs. Des généralisations de ce critère ont alors été proposées. Dans la section suivante nous allons exposer certains contre-exemples avant de montrer les généralisations proposées.

2. Comme dans le cas de la décision sous risque, u est définie à une transformation affine près.

2.3 Limitations et généralisations de la théorie de l'utilité espérée

2.3.1 Deux contre-exemples de la théorie de l'utilité espérée

Paradoxe d'Allais pour la décision sous risque

Cet exemple a été proposé par Allais [All53] et présente un problème de décision sous risque pour lequel de nombreux décideurs auxquels il l'a soumis ont fait un choix incompatible avec le critère de l'utilité espérée. Nous présentons ici une version simplifiée de cet exemple (tirée de (Kast [Kas93])) :

Exemple 2.3.1 *Paradoxe d'Allais*
Soient les quatre loteries suivantes :

– l_1 *donne accès à un gain de 15000 Frs avec la probabilité* $0,09$ *et* 0 *Frs avec la probabilité* $0,91$*,*

– l_2 *donne accès à un gain de 10000 Frs avec la probabilité* $0,1$ *et* 0 *Frs avec la probabilité* $0,9$*,*

– l_3 *donne accès à un gain de 15000 Frs avec la probabilité* $0,9$ *et* 0 *Frs avec la probabilité* $0,1$*,*

– l_4 *offre un gain certain de 10000 Frs.*

Il ne semble pas illogique de préférer strictement l_1 à l_2 (car l_1 permet d'obtenir un meilleur gain que l_2, avec une probabilité assez voisine) et de préférer strictement l_4 à l_3 car l_4 représente un gain certain, alors que la probabilité de ne rien gagner avec l_3, quoique faible, est non nulle.
Or un tel comportement ($l_1 \succ l_2$ et $l_4 \succ l_3$) est en contradiction avec le critère de l'utilité espérée : cherchons des utilités $u(15000)$, $u(10000)$ et $u(0)$ compatibles avec ces deux préférences. Comme dans la théorie de von Neumann et Morgenstern u n'est déterminée qu'à une transformation affine près par les préférences du décideur, on peut librement poser $u(0) = 0$. Exprimées en termes d'utilités, les deux préférences donnent les inégalités suivantes : $0,09 \times u(15000) > 0,1 \times u(10000)$ et $u(10000) > 0,9 \times u(15000)$.

Ce système d'inéquations n'admet pas de solution, ce qui signifie que pour ce problème il n'existe aucune fonction d'utilité sur les conséquences telle que les préférences de l'utilisateur soient modélisables par l'espérance mathématique de cette fonction.

Nous citons, simplement pour mémoire, le critère proposé par Allais [All53], généralisant le critère de l'utilité espérée et permettant de résoudre ce paradoxe : $X = \{x_1, \ldots, x_n\}$ et $u(x_1) < \ldots < u(x_n)$,

$$u(l) = u(x_1) + q(p_2 + \ldots + p_n) \cdot (u(x_2) - u(x_1)) + \ldots + q(p_n) \cdot (u(x_n) - u(x_{n-1}))$$

où q est un fonction croissante et continue de $[0, 1]$ dans lui même.
Nous n'irons pas plus loin dans cette voie qui généralise la théorie de l'utilité espérée pour la *décision sous risque*. En revanche, nous allons nous arrêter plus longuement sur des théories généralisant le critère de l'utilité espérée pour la décision dans l'incertain. Là, c'est la notion de probabilités subjectives qui est remise en cause, et généralisée.

Paradoxe d'Ellsberg : risque *vs* incertitude

Le contre-exemple proposé par Ellsberg [Ell61] montre comment la théorie de l'utilité espérée, et l'usage des probabilités (additives) pour représenter l'incertitude peut être mise en défaut par des problèmes mélangeant risque et incertitude. Ceci a mené certains chercheurs (Gilboa [Gil87], Schmeidler, [Sch89]) à proposer une généralisation de l'utilité espérée, utilisant des probabilités non-additives, ou *capacités* (Choquet [Cho54]).

Exemple 2.3.2 *Paradoxe d'Ellsberg*
Nous avons une urne remplie de 90 boules de couleurs rouge (R), blanche (B) ou jaune (J). Nous savons qu'exactement 30 boules sont rouges, et que 60 sont de couleur blanche ou jaune. Les proportions de boules blanches et jaunes sont indéterminées. Une boule est tirée au hasard et le problème de décision consiste à ordonner les deux paires de paris suivantes :

Pari	*Rouge*	*Blanc*	*Jaune*
f	*1000 Frs*	*0 Frs*	*0 Frs*
g	*0 Frs*	*1000 Frs*	*0 Frs*
f'	*1000 Frs*	*0 Frs*	*1000 Frs*
g'	*0 Frs*	*1000 Frs*	*1000 Frs*

Le pari **f** *rapporte 1000 Frs si la boule tirée est rouge, et rien sinon, etc...*

Il apparaît raisonnable de préférer **f** à **g**, car la proportion de boules rouges vaut exactement un tiers alors que la proportion de boules blanches est imprécisément connue : elle peut être de deux tiers, mais elle peut également être nulle. Préférer **f** à **g** signifie que le décideur a une *aversion pour l'ambiguïté* (ou l'incertitude).
Maintenant, si il préfère **f** à **g**, il doit pour la même raison (aversion pour l'ambiguïté) préférer **g'** à **f'** car **g'** offre exactement deux chances sur trois de gagner alors que **f'** offre entre une chance sur trois et cent pour cent de chances de gain. Ces préférences (**f** $\succ$ **g** et **g'** $\succ$ **f'**) sont une nouvelle fois incompatibles avec la théorie de l'utilité espérée. On observe facilement que le "principe de la chose certaine" (**Sav2**) est violé : **f** et **g** donnent la même conséquence pour l'état "J" (0 Frs), et la modification de cette conséquence (1000 Frs) conduit à un renversement de préférences.
Notons que l'attitude négative du décideur envers l'ambiguïté n'est pas le point crucial : cette attitude doit simplement être cohérente pour tous les problèmes de décision qui s'offrent à lui. Un raisonnement similaire pour un décideur ayant un "attrait" pour l'ambiguïté conduirait également à une violation de **Sav2**. En effet, les préférences seraient cette fois (**f** $\prec$ **g** et **g'** $\prec$ **f'**).

Détaillons les contraintes imposées aux probabilités $P(B)$, $P(J)$ et $P(R)$ par les préférences **f** $\succ$ **g** et **g'** $\succ$ **f'**. Les deux contraintes obtenues sont : $P(R) > P(B)$ et $P(R \cup J) < P(B \cup J)$, donc $P(B) + P(R \cup J) < P(B \cup J) + P(R)$. Or si on suppose que les probabilités sont additives, puisque B, R et J sont disjoints, on obtient $P(R \cup B \cup J) < P(R \cup B \cup J)$, ce qui est impossible. Ainsi, aucun jeu de probabilités subjectives (même différentes des probabilités objectives) ne peut exister.
Pour résoudre ce paradoxe il nous faut donc abandonner l'hypothèse d'additivité des probabilités subjectives. Nous décrirons dans le paragraphe 2.3.2 page 32 une théorie de la décision utilisant des mesures de *capacité* qui sont une généralisation non additive des mesures de probabilité.

Mais avant d'étudier cette théorie dans toute sa généralité, nous allons nous intéresser à deux cas particuliers, utilisant deux capacités particulières, généralisant tout de même les probabilités : les fonctions de croyance et de plausibilité, introduites par Dempster [Dem67] et Shafer [Sha76].

2.3.2 Décision avec fonctions de croyance, décision basée sur l'intégrale de Choquet

Introduction

Comme le montre l'exemple d'Ellsberg, la théorie classique de la décision peut être mise en défaut sur des problèmes de décision comportant à la fois risque et ambiguïté.
Pour de tels problèmes l'approche classique (Savage [Sav54], von Neumann et Morgenstern [vNM44], mais surtout Anscombe et Aumann [AA63]) propose de traiter l'information disponible en deux temps : d'abord l'information "probabilisable" (objective), puis "ambiguë" (en utilisant des probabilités subjectives), puis de combiner les deux distributions de probabilité. Dans l'exemple d'Ellsberg, l'information probabilisable est donnée par l'intermédiaire de la répartition (connue) des boules : 30 rouges et 60 blanches ou jaunes. L'information ambiguë concerne la répartition des boules blanches et jaunes. Or c'est là que le bât blesse : les choix "intuitifs" du décideur d'Ellsberg sont incompatibles avec une représentation probabiliste (subjective) de l'ambiguïté. Au contraire, les solutions proposées sont dictées par *l'attitude* du décideur envers l'ambiguïté : aversion ou attrait. Si il a une aversion pour l'ambiguïté, il considérera que le pire cas possible se présentera pour chaque décision. Lorsqu'il évalue le pari (l'action) **g** (1000 Frs pour une boule blanche), c'est la situation "aucune boule blanche" qui gouverne son évaluation. Lorsqu'il évalue **f'** (1000 Frs pour une boule J ou R), c'est au contraire sur la situation "aucune boule jaune" qu'il se concentre.

Ce type d'attitude est incompatible avec une représentation de la connaissance du décideur par une distribution de probabilité unique, indépendante des actions considérées. En revanche, il est représentable par des critères comme *maximin* ou *maximax* (voir Section 1.2.1 page 10).
A partir de ce constat, Jaffray [Jaf89]), puis Jaffray et Wakker [JW94] ont proposé un mode de décision utilisant une représentation de l'incertitude en deux étapes, l'étape "subjective" ne faisant pas appel à une représentation probabiliste. Ce mode d'évaluation est illustré par l'exemple suivant :

Exemple 2.3.3 *Le vendeur de téléviseurs* (Jaffray et Wakker [JW94])
A l'heure de fermeture de son magasin, un vendeur de TV se demande s'il va servir un dernier client qui se présente devant sa boutique. S'il le sert, il va sans doute rater le concert de Jazz auquel il avait prévu de se rendre. D'un autre côté, il est certain de vendre un téléviseur supplémentaire. En fait, le profit du vendeur va dépendre de la catégorie de prix, basse (B), moyenne (M) ou élevée (E) du poste qu'il va vendre.
Il possède les connaissances suivantes :

(1) 60 % des clients ont une TV de catégorie basse (b), 30 % de catégorie moyenne (m), et 10 % de catégorie élevée (e) ;

(2) lorsque les gens rachètent une TV, soit ils restent dans la même catégorie de prix, soit ils passent à la catégorie immédiatement supérieure.

Le problème de décision en question (servir ou ne pas servir) est illustré par la figure 2.1 page suivante. $S = \{B, M, E\}$ représente l'ensemble des états du monde immédiatement pertinents pour la décision. $\Omega = \{b, m, e\}$ est un ensemble *d'états de connaissance*, équipé d'une distribution de probabilités objectives, relié à l'ensemble S : à chaque état de connaissance correspond un sous-ensemble de l'espace d'états ($b \to \{B, M\}$; $m \to \{M, E\}$; $e \to \{E\}$).
Des deux informations (1) et (2), la première est probabilisable, mais pas la seconde. Le décideur peut tout de même obtenir des encadrements de probabilités à partir de (1) et (2) : $0 \leq P(\{B\}) \leq 0,6$ car le client n'achètera une TV à bas prix que s'il en possède déjà une ($p(b) = 0,6$ et encore dans ce cas

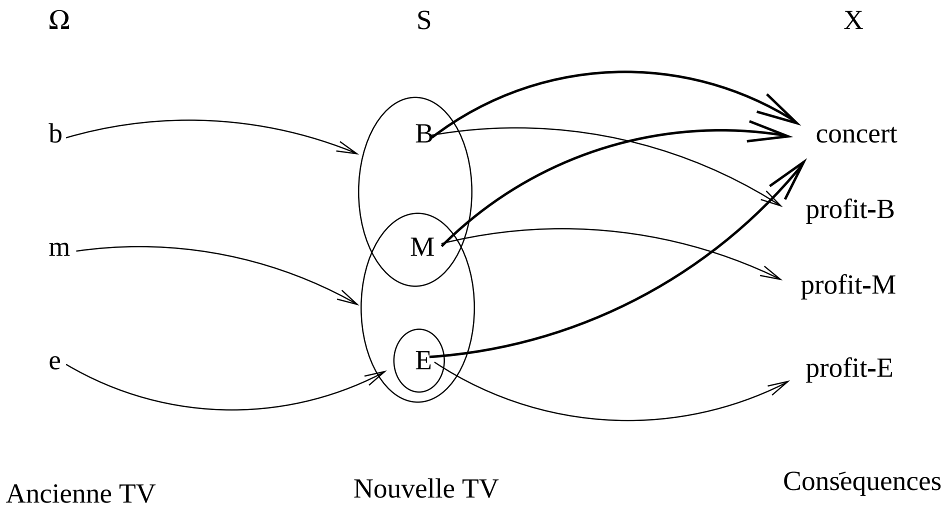

FIG. 2.1 – *Problème du vendeur de TV : risque et ambiguïté*

il peut tout aussi bien acheter une TV à prix moyen). De même, il sait que $0,1 \leq P(\{E\}) \leq 0,4$, etc...

Dans le paragraphe suivant, nous décrivons les fonctions de croyance, adaptées à ce type de représentation de l'incertitude.

Fonctions de croyance

L'information sur S est représentée par une application $m\ :\ 2^S\ \rightarrow\ [0,1],\ A\ \rightarrow\ m(A)$, telle que $\Sigma_{A \subseteq S} m(A) = 1$.
Dans l'exemple précédent, m est obtenue à partir de l'information (1) (p sur Ω) et de (2) (application de Ω vers 2^S). Ainsi, m est définie par :

- $m(\{B, M\}) = 0,6$,
- $m(\{M, E\}) = 0,3$ et
- $m(\{E\}) = 0,1$.

Une telle application m est appelée *fonction de masse* par Shafer [Sha76].
A partir de m on définit deux mesures sur S, appelées respectivement fonctions de croyance et de plausibilité.

Définition 2.3.1 *Fonctions de croyance et de plausibilité*
Soit m une fonction de masse sur S, les applications suivantes :

$$Bel : 2^S \rightarrow [0,1], A \rightarrow \Sigma_{B \subseteq A} m(B)$$

$$Pl : 2^S \rightarrow [0,1], A \rightarrow \Sigma_{B \cap A \neq \emptyset} m(B)$$

sont respectivement les fonctions de croyance et de plausibilité associées à m.

Remarquons que Pl et Bel sont reliées : $\forall A \subseteq S, Pl(A) = 1 - Bel(\overline{A})$. On peut remarquer également que Pl et Bel ne sont pas des représentations numériques de probabilités qualitatives (elles violent l'axiome **P**), c'est d'ailleurs la raison pour laquelle elles permettent de résoudre le paradoxe d'Ellsberg...
Néanmoins, les fonctions de croyance et de plausibilité sont des généralisations des probabilités : si m vérifie la propriété $\forall A \subseteq S, m(A) \neq 0 \Rightarrow A$ est un singleton (les ensembles A tels que $m(A) \neq 0$ sont appelés *éléments focaux* de m), alors les fonctions de croyance et de plausibilité associées à m sont égales et possèdent toutes les caractéristiques d'une mesure de probabilités (de distribution associée m).

Wong et al. [WYBB91], [WYL93] ont proposé deux généralisations (duales) de l'axiome **P**, caractérisant [3] les relations (sur 2^S) de croyance et de plausibilité qualitatives :

Bel $B \subseteq A, A \cap C = \emptyset \Rightarrow (A \succ B \Rightarrow A \cup C \succ B \cup C)$

Pl $B \subseteq A, B \cup C = S \Rightarrow (A \succ B \Rightarrow A \cap C \succ B \cap C)$

Les fonctions Bel et Pl vérifient également (parmi d'autres) les propriétés suivantes :

$$Bel(A \cup B) + Bel(A \cap B) \geq Bel(A) + Bel(B), \text{ et } Pl(A \cup B) + Pl(A \cap B) \leq Pl(A) + Pl(B).$$

Rappelons que les mesures de probabilités vérifient l'égalité : $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.

Dans le paragraphe suivant, nous allons décrire le modèle de décision basé sur les fonctions de croyance (et de plausibilité), proposé par Jaffray et Wakker. Nous verrons qu'il s'agit d'une généralisation de l'utilité espérée, et nous montrerons brièvement un relâchement de l'axiome du *principe de la chose certaine*, satisfait par ce critère.

Décision basée sur les fonctions de croyance

Revenons à la modélisation en deux étapes d'une décision sous *risque-ambiguïté* (Fig. 2.2 page ci-contre) :
B_ω représente l'ensemble des états du monde associés à l'état de connaissance ω. Du fait que Ω est probabilisable, une extension "raisonnable" du critère de l'utilité espérée peut être de considérer a non comme une application de S vers X, mais de Ω vers X. Seulement, dans ce cas a est une application de Ω vers 2^X, car à chaque état de connaissance correspond un ensemble de conséquences possibles par a. Du coup, pour appliquer le critère de l'utilité espérée, on doit étendre la fonction d'utilité u (sur X dans le cas classique) à 2^X. C'est dans cette extension que se reflète l'attitude du décideur envers l'ambiguïté car la "non-détermination" de la conséquence de l'action appliquée dans un état donné de connaissance provient de l'ambiguïté de cette dernière.
Afin d'étendre le critère de l'utilité epérée, Jaffray [Jaf89] a proposé le critère suivant,

$$U : a \to \Sigma_{B \subseteq S} m(B) \cdot V_a(B) \tag{1}$$

où $V_a(B) = V(a(B))$, V étant une extension de u sur 2^X. Si $a(B) = \{x \in X, \exists s \in B, a(s) = x\}$ est le singleton $\{x_0\}$, alors $V_a(B) = u(x_0)$. Jaffray a montré que si l'on souhaite que la relation de

3. En fait, ces deux propriétés ne sont pas exactement des caractérisations des mesures de croyance et de plausibilité : (1) toute fonction de croyance (resp. de plausibilité) satisfait **Bel** (resp. **Pl**) (2) toute relation satisfaisant **Bel** (resp. **Pl**) est représentable par une fonction de croyance (resp. de plausibilité), mais une relation satisfaisant **Bel** (resp. **Pl**) peut parfois être représentée par d'autres fonctions.

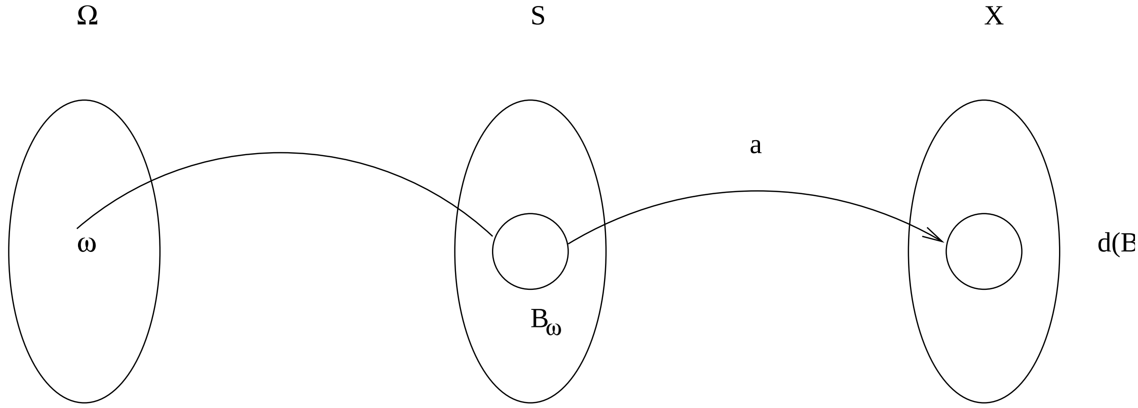

FIG. 2.2 – *Risque et ambiguïté : décision en deux étapes*

préférence sur $\mathcal{A}$ représentée par U respecte la propriété naturelle de "respect de la dominance" (si une action domine faiblement une autre (cf. Déf. 1.2.1 page 12), alors elle n'est pas "moins préférée"), alors U doit prendre la forme :

$$U : a \to \Sigma_{B \subseteq S} m(B) \cdot \psi(u_*(a(B)), u^*(a(B))) \qquad (2)$$

où ψ est une fonction non décroissante de ses deux arguments et $u_*(a(B))$ et $u^*(a(B))$ sont respectivement les utilités de la pire et de la meilleure conséquence de $a(B)$.

Jaffray et Wakker [JW94] ont proposé une caractérisation axiomatique du critère U qui consiste en un affaiblissement des axiomes de Savage. En effet, U représente une généralisation du critère de l'utilité espérée, ce dernier pouvant être retrouvé lorsque m est une distribution de probabilités (on retrouve un problème de décision sous risque). Nous n'allons pas décrire l'axiomatique de Jaffray et Wakker, mais simplement donner la propriété suivante qui montre un lien avec la théorie classique :

Propriété 2.3.1 *U satisfait le principe de la chose certaine, restreint aux événements non ambigus.*

Rappelons que le principe de la chose certaine stipule que :
$\forall$ **f**, **g**, **h**, **h'**, $A \subseteq S$, **f**A**h** $\preceq$ **g**A**h** $\Rightarrow$ **f**A**h'** $\preceq$ **g**A**h'**.
L'utilité basée sur les fonctions de croyance satisfait ce principe pour tout événement A non ambigu, i.e. tel que : $\nexists B \subseteq S$, $m(B) > 0$, $A \cap B \neq \emptyset$ et $\overline{A} \cap B \neq \emptyset$. Les événements non ambigus sont également caractérisés par la propriété : A *non ambigu ssi* $Bel(A) = Pl(A)$.

Dans la seconde partie de cette thèse, nous évoquerons deux cas particuliers de fonction U se prêtant particulièrement bien à un traitement logique des problèmes de décision. Ces deux cas sont obtenus en choisissant comme fonction V dans (1), les fonctions min et max respectivement (correspondant aux critères $maximin$ et $maximax$ de la décision sous incertitude *non probabilisée*) :

- $U_* : a \to \Sigma_{B \subseteq S} m(B) \cdot u_*(a(B))$ et
- $U^* : a \to \Sigma_{B \subseteq S} m(B) \cdot u^*(a(B))$.

Les fonctions d'utilité basées sur les fonctions de croyance généralisent à la fois le critère de l'utilité espérée et les critères de décision dans l'incertain (prendre pour unique élément focal $m = S$). Néanmoins, ces critères (à part le critère (2)) ne sont que des cas particuliers d'un critère encore plus général, basé sur *l'intégrale de Choquet* et la notion de *capacité*. C'est ce critère que nous allons brièvement présenter dans le paragraphe suivant.

Décision basée sur l'intégrale de Choquet

L'intégrale de Choquet [Cho54] est basée sur une mesure de l'incertitude ν très générale, possédant les propriétés suivantes :

- $\nu(\emptyset) = 0$, $\nu(S) = 1$,
- $B \subseteq A \Rightarrow \nu(A) \geq \nu(B)$.

La seconde propriété est appelée propriété de *monotonie*. Elle est très générale et vérifiée par bon nombre de mesures d'incertitude (probabilités, fonctions de croyance, possibilités et nécessités, ...). Nous appellerons *mesure monotone* toute fonction satisfaisant ces deux propriétés. Plusieurs noms ont été donnés à de telles mesures, suivant le domaine dans lequel elles ont été étudiées : *capacités* (par Choquet qui les a utilisées pour modéliser les caractéristiques de certains composants électriques), ou *mesures floues* (par Sugeno, [Sug77]).

L'intégrale de Choquet d'une fonction $\phi : S \rightarrow \mathbb{R}$ relativement à une mesure monotone ν, notée $\int_S \phi d\nu$ est définie par :

$$\int_S \phi d\nu = \int_{\mathbb{R}^+} \nu(\Phi_\alpha) d\alpha + \int_{\mathbb{R}^-} [\nu(\Phi_\alpha) - 1] d\alpha$$

où $\Phi_\alpha = \{s \in S, \phi(s) \geq \alpha\}$.

L'intégrale de Choquet a été appliquée à la décision par plusieurs auteurs (Gilboa [Gil87], Schmeidler [Sch89], Sarin et Wakker [SW92]). Si u est une fonction d'utilité sur l'ensemble de conséquences X, et **f** est une action (une application de S vers X), l'utilité de Choquet de **f** vaut :

$$U_C(\mathbf{f}) = \int_{\mathbb{R}^+} \nu(F_\alpha) d\alpha + \int_{\mathbb{R}^-} [\nu(F_\alpha) - 1] d\alpha$$

où $F_\alpha = \{s \in S, u(\mathbf{f}(s)) \geq \alpha\}$.

Dans le cas où $X = \{x_0, \ldots, x_n\}$ est fini, et $x_n \geq \ldots \geq x_0$, on obtient l'expression suivante de U_C :

$$U_C(\mathbf{f}) = \alpha_0 + \sum_{i=1}^{n} (\alpha_i - \alpha_{i-1}) \cdot \nu(F_{\alpha_i}) \tag{3}$$

où $\alpha_i = u(x_i)$.

On peut remarquer la généralité du critère basé sur l'intégrale de Choquet : "en spécialisant" ν en une fonction de croyance ou de plausibilité, on obtient les critères particuliers U^* et U_* proposés par Jaffray et Wakker. En effet, l'expression (3) est alors égale à l'expression (1) (avec $\psi(u_*(a(B)), u^*(a(B))) = u_*(a(B))$), si on remplace ν par une fonction de croyance Bel, et à l'expression duale si ν est une mesure de plausibilité.

Sarin et Wakker [SW92] ont proposé une axiomatisation de ce critère, basée sur un relâchement des axiomes de Savage. Comme dans l'axiomatisation de Jaffray et Wakker [JW94], il est fait appel à la notion d'événements *non-ambigus* ("probabilisables"). Les axiomes **Sav 2**, **Sav 3** et **Sav 6** sont restreints aux événements non-ambigus, et **Sav 4** est remplacé par un *axiome de dominance* :

DC *Dominance cumulative*
$\forall$ **f**, **g**, $(\forall \alpha \in$IR, $F_\alpha \geq_L G_\alpha) \Rightarrow$ **f** $\succeq$ **g**
où $F_\alpha - \{s \in S, u(\mathbf{f}(s)) \geq \alpha\}$ et $G_\alpha - \{s \in S, u(\mathbf{g}(s)) \geq \alpha\}$.

2.4 Une approche ordinale de la décision dans l'incertain basée sur la théorie des possibilités

2.4.1 Quelques rappels de théorie des possibilités

Une distribution de possibilité π sur un ensemble d'états du monde possibles S est une fonction de l'ensemble S vers une échelle simplement ordonnée $(L, <)$, bornée (Zadeh [Zad65]). Cette échelle peut être quelconque : par exemple l'intervalle unité $[0, 1]$ équipé des opérations *maximum* et *minimum* et d'une opération de renversement n_L telle que $n_L(0) = 1$ et $n_L(1) = 0$. Dans le cas où $L = [0, 1]$, en général $n_L = 1 - \cdot$. Lorsque L est finie, on continue à noter 0 et 1 ses plus petit et plus grand éléments.

Une distribution de possibilité décrit les connaissances de l'agent sur l'état du monde (mal connu). Elle distingue les états du monde *plausibles* de ceux qui le sont moins. Elle permet de distinguer les états "normaux" des états "anormaux". La fonction $\pi : S \to L$ représente une restriction *flexible* sur les états du monde. Par convention, $\pi(s) = 0$ signifie que s est un état *impossible*, alors que $\pi(s) = 1$ signifie que s est entièrement possible (plausible).

Contrairement aux probabilités, les possibilités permettent d'exprimer que plusieurs états sont entièrement plausibles (ont un degré de possibilité valant 1). Bien sûr, si S représente l'ensemble des états possibles du monde, l'état réel du monde appartient forcément à S, de telle sorte qu'il existe au moins un état du monde de degré de possibilité 1. Cette propriété est appelée "normalisation" de la distribution de possibilités. Dans cette thèse nous considérerons des distributions de possibilité normalisées, sinon nous le mentionnerons explicitement. D'après Zadeh [Zad78], une distribution de possibilité peut être vue comme la fonction caractéristique "généralisée" d'un ensemble flou.

Une distribution de possibilité π est dite "au moins aussi spécifique" qu'une distribution π' ssi $\forall s \in S, \pi(s) \leq \pi'(s)$ (Yager [Yag79]). Dans ce cas, π est au moins aussi restrictive, et donc informative que π'. La théorie des possibilités permet d'exprimer des formes extrêmes de connaissance partielle :

- connaissance complète : $\exists s_0, \pi(s_0) = 1$ et $\forall s \neq s_0, \pi(s) = 0$,
- ignorance totale : $\forall s \in S, \pi(s) = 1$.

Dans le premier cas, un seul état du monde est possible, tous les autres sont impossibles : l'état réel du monde est donc parfaitement connu. Dans le second cas, tous les états du monde sont également plausibles, nous n'avons aucune information sur l'état réel du monde.

Si $A \subseteq S$ est un sous-ensemble d'états du monde, et π est la distribution de possibilité exprimant les connaissances de l'agent, on peut se poser des questions du genre : "est-ce que l'état réel du monde appartient à A" ?
Pour répondre à cette question on dispose des deux indices suivants :

- mesure de possibilité,
$$\Pi(A) = sup_{s \in A} \pi(s),$$

– mesure de nécessité,

$$N(A) = n_L(\Pi(\overline{A})) = inf_{s \in \overline{A}} n(\pi(s)).$$

$\Pi(A)$ représente à quel point A est *compatible* avec les connaissances de l'agent (exprimées par π). $\Pi(A)$ est défini par le degré de possibilité de l'état le plus plausible parmi ceux de A, supposant ainsi que si l'événement A se produit, l'état réel du monde se trouve parmi les plus plausibles, compatibles avec A.
L'axiome de base des mesures de possibilité (dans le cas fini) est le suivant :

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B)),$$

cet axiome est justifié par le fait que l'on se concentre sur les situations les plus plausibles pour évaluer le degré de possibilité d'un événement.

$N(A)$ est appelé *degré de nécessité* de A, et représente à quel point $\overline{A}$ est impossible. $N(A) \geq \alpha > 0$ implique que dans toutes les situations les plus plausibles (de possibilité supérieure ou égale à $n_L(\alpha)$), A est vraie. Ainsi, $N(A) > 0$ signifie que A est une connaissance "acceptable", i.e. A peut être supposée vraie. Néanmoins, ce type de raisonnement est révisable et A peut être remise en cause par de nouvelles connaissances. Dans le cas fini, les mesures de nécessité sont définies par un axiome dual de celui des mesures de possibilités [4],

$$N(A \cap B) = \min(N(A), N(B)).$$

Jusqu'ici nous avons vu comment ces mesures peuvent être interprétées pour décrire les connaissances incomplètes d'un agent. Elles peuvent être interprétées différemment, en termes de préférences (entre conséquences de décisions, par exemple).
Dans ce cadre, la possibilité d'une conséquence peut être vue comme le degré auquel elle satisfait les préférences de l'agent. Ainsi, les possibilités modélisent des *préférences graduelles*. Une distribution de possibilité peut alors être vue comme une fonction d'utilité (voir (Dubois, Fargier et Prade [DFP96]) pour une discussion détaillée sur le sujet).

Nous allons voir dans le paragraphe suivant comment l'utilisation conjointe des deux types de possibilités permet d'obtenir une théorie qualitative (ordinale) de la décision dans l'incertain.

2.4.2 Théories possibilistes de la décision dans l'incertain

Dans le cadre proposé par Savage pour la décision dans l'incertain, l'incertitude de l'agent est modélisée par une distribution de probabilité p sur l'ensemble S des états possibles du monde et ses préférences par une fonction d'utilité u, à valeurs réelles, sur l'ensemble X des conséquences éventuelles de ses actions.
Dans le cadre possibiliste, les connaissances de l'agent sont modélisées par une distribution de possibilité π, normalisée, fonction de S vers une échelle simplement ordonnée L. L'échelle L est une échelle de "plausibilité", et comme il a été dit dans le paragraphe précédent, $\pi(s) \in L$ représente le degré de "vraisemblance" que s soit l'état réel du monde.
Si l'on considère que les informations que possède l'agent sur le problème de décision sont purement ordinales, il est raisonnable de penser que non seulement ses connaissances peuvent être exprimées

4. La nature ordinale des mesures de possibilité et nécessité doit être soulignée.

par une distribution de possibilité, mais également ses préférences. Soit donc μ, la distribution de possibilité représentant ses préférences, μ prend ses valeurs dans l'échelle simplement ordonnée, L'. Une action est, comme dans la théorie de Savage, représentée par une fonction a, de S vers X. L'utilité d'une action a, de conséquence $a(s) \in X$ dans l'état s peut être évaluée en combinant les degrés de possibilité $\pi(s)$ et les utilités $\mu(a(s))$ de manière adéquate, pour tous les états possibles du monde.

Dans la littérature des ensembles flous, deux critères d'évaluation ont été proposés, permettant de réaliser de telles combinaisons en supposant une certaine forme de commensurabilité entre les échelles de plausibilité (L) et d'utilité (L') :

– un critère pessimiste v_*,
$$v_*(a) = \inf_{s \in S} \max(n(\pi(s)), \mu(a(s)))$$
qui généralise le critère de Wald pour la décision dans l'incertain (n est une fonction décroissante de L vers L', telle que $n(1_L) = 0_{L'}$ et $n(0_L) = 1_{L'}$),

– un critère optimiste v^*,
$$v^*(a) = \sup_{s \in S} \min(m(\pi(s)), \mu(a(s)))$$
qui généralise le critère optimiste $maximax$ (m est une fonction croissante de L vers L', telle que $m(0_L) = 0_{L'}$ et $m(1_L) = 1_{L'}$).

Le critère optimiste à d'abord été proposé par Yager [Yag79] et le pessimiste d'abord par Whalen [Wha84] puis réutilisé par Inuiguichi, Ichihashi et Tanaka [IIT89]. Ces critères représentent respectivement les degrés de nécessité et de possibilité de l'ensemble flou sur S dont la fonction d'appartenance est $\mu(a(\cdot))$. Ce sont aussi les opérations de base du "fuzzy pattern matching" (en bases de données ou en systèmes experts).

Nous exposerons dans le chapitre 1 page 73 de la partie II l'axiomatique "à la von Neumann et Morgenstern" proposée par Dubois et Prade [DP95b] pour l'utilisation de ces deux critères en décision dans l'incertain. Dans ce même chapitre, nous exposerons également la principale contribution axiomatique de cette thèse : une justification de ces critères en termes de préférence entre actions (à la Savage).

Ces deux critères sont des formes particulières d'une intégrale plus générale : *l'intégrale de Sugeno* [Sug77]. Nous proposerons également, dans le chapitre 1 page 73 de la partie II une axiomatisation de l'intégrale de Sugeno en tant que critère de décision ordinal (Dubois, Prade et Sabbadin [DPS98]).

2.5 Conclusion

L'objet de ce chapitre était de donner un aperçu de différentes théories pour la décision dans l'incertain. Ces théories ont toutes la caractéristique commune de faire intervenir des relations de préférence sur un ensemble de conséquences et de "plausibilité" sur un ensemble d'états du monde, voire d'événements.
Ces théories diffèrent par la nature des relations envisagées. Celles-ci peuvent être numériques : préférences et plausibilités, dans les théories basées sur l'intégrale de Choquet ou ses dérivées. Les critères de décision dans l'incertain "non probabilisé" que nous avons décrit différencient simplement états du

monde possibles et impossibles, par contre la structure de la relation de préférence peut être simplement ordinale (critère de Wald) ou plus raffinée. Nous avons également évoqué des critères "ordinaux" de décision dans l'incertain, basés sur la théorie des possibilités.

Dans le chapitre 1 page 73 de la partie II nous poursuivrons dans cette voie, utilisant des *structures ordinales* pour représenter à la fois les relations de préférence et de plausibilité. Nous exposerons et justifierons axiomatiquement ces derniers critères généralisant le critère de Wald par l'utilisation, non pas de l'intégrale de Choquet, mais de l'intégrale de Sugeno qui en est une contrepartie ordinale (où les seules opérations utilisées sont min, max et l'opération de renversement d'échelle). Nous décrirons en détails les critères d'*utilité qualitative possibiliste* qui sont à l'intégrale de Sugeno ce que les critères basés sur les fonctions de croyance et de plausibilité sont à l'intégrale de Choquet.

Dans ce chapitre, nous avons montré que l'intérêt de la justification axiomatique de critères de décision n'est pas seulement théorique. En effet, grâce à l'observation des justifications des différents critères on peut déterminer celui qui est le plus approprié pour résoudre un problème de décision donné.
Ces justifications axiomatiques peuvent également servir (celle de Savage pour l'utilité espérée, par exemple) à construire, ou au moins à vérifier, les relations de plausibilité sur les états du monde et de préférence sur les conséquences, en observant seulement un sous-ensemble des actions potentielles, disponibles pour un agent.

Il est un domaine de l'*Intelligence Artificielle* qui fait appel à la théorie de l'utilité espérée : celui de la planification sous incertitude. Dans le prochain chapitre, nous allons parler de décision séquentielle dans l'incertain. Le problème de décision ne se limite plus au choix d'une action unique à appliquer, mais d'une suite d'actions visant à atteindre un but donné. Plus particulièrement, nous allons parler de *processus décisionnels markoviens* (l'incertitude est de nature probabiliste et les différentes étapes de décision sont indépendantes).
Nous exposerons une méthode de résolution de tels problèmes : la *programmation dynamique*. Nous évoquerons ensuite la notion d'*observabilité*, qui peut intervenir dans ce type de problèmes : une fois qu'une décision est prise et que l'action choisie est appliquée, l'agent a-t-il une connaissance exacte de l'état (conséquence) résultant ? Si oui, on parlera de problèmes *totalement observables*, sinon de problèmes *partiellement observables*, voire à *observabilité nulle*.

Chapitre 3

Problèmes de décision séquentielle

3.1 Introduction

Dans les deux chapitres précédents, nous avons parlé de décision dans l'incertain. Nous avons exposé un certain nombre de théories dont la plus connue est celle de l'utilité espérée.
Nous y avons fait l'hypothèse qu'un problème de décision ne se limitait qu'au choix et à l'application d'une seule action. Néanmoins cette hypothèse est restrictive. Dans de nombreux domaines, un agent est confronté à des choix répétitifs et interdépendants. De ses actions résultent des conséquences ayant une certaine utilité pour lui, mais elles modifient également l'état du monde d'une manière dont vont dépendre ses décisions futures.
Dans le domaine de la *planification*, en particulier, un agent est confronté à ces problèmes de choix répétitifs (ou séquentiels), mais on retrouve également de tels problèmes en *ordonnancement*, *allocation de ressources*...

Dans ce chapitre, nous allons nous intéresser à ces problèmes de décision séquentiels. Après avoir présenté une formulation des problèmes de décisions *markoviens*, nous allons exposer la méthode de résolution appelée *programmation dynamique*, basée sur le *Principe d'Optimalité* de Bellman [Bel57]. Ensuite, nous effectuerons une petite digression sur la notion *d'observation* dans un problème de décision sous incertitude (probabiliste), avant de décrire les processus décisionnels markoviens *partiellement observables*.

3.2 Décision séquentielle, processus décisionnels markoviens et programmation dynamique

3.2.1 Formulation des problèmes décisionnels markoviens

Lorsqu'un agent commande (par ses décisions) un système dynamique, à chaque instant il observe l'état de ce système afin de choisir une action à appliquer, parmi un ensemble d'actions disponibles. La conséquence de l'application de l'action choisie est double.

– L'agent reçoit une récompense (ou une pénalité) immédiate résultant de l'application de l'action dans l'état courant du monde, comme dans les théories que nous avons décrites dans les chapitres précédents. Ces récompenses peuvent s'additionner ou non, au cours du temps.

– Le système évolue (i.e. l'état du monde change).

L'objectif de l'agent est de maximiser les performances du système, i.e. de maximiser la récompense totale qu'il recevra au cours du temps. Une méthode très simple pour guider les choix de l'agent consiste à choisir, à chaque étape, l'action qui maximise la récompense immédiate de l'agent. Malheureusement, cette méthode ne garantit pas que la récompense totale obtenue soit maximale. En effet, l'action appliquée immédiatement influe sur l'évolution future du système, qui à son tour va influer sur le total des récompenses que l'agent va recevoir, pouvant remettre en cause le bénéfice immédiat résultant de l'application de l'action choisie.

Ainsi, il est important de prendre en compte l'évolution future du système pour choisir l'action suivante. Qu'en est-il de l'évolution passée du système ?
Sous *l'hypothèse markovienne*, l'état suivant du système ne dépend que de l'état courant et de l'action mise en oeuvre dans cet état : l'histoire passée du système n'a aucune influence sur son évolution future.

Ces deux propriétés plaident pour une *recherche arrière* de la séquence d'actions optimale. L'hypothèse markovienne permet de plus une recherche *par induction* (l'état suivant se déduit de l'état courant et de l'action appliquée dans cet état uniquement).
Nous exposerons une méthode de recherche par induction arrière ("backward induction" en anglais), mais avant cela, nous devrons donner un certain nombre de notations qui seront utilisées dans ce chapitre.

T représente l'ensemble des instants (ou étapes) auxquels les décisions peuvent être prises. T peut être fini ou infini, et dans le dernier cas, discret ou continu. Suivant les cas, on parle de décision séquentielle à *horizon fini* ou *infini*. Nous ne traiterons pas dans cette thèse le cas où T est infini et continu. Dans les deux autres cas, T sera représenté respectivement par les ensembles $\{1, 2, \ldots, N\}$ (cas fini) ou $\{1, 2, \ldots\}$ (cas infini). L'étape courante sera notée t et la suivante, $t + 1$.

L'ensemble des états possibles du monde à l'étape t est noté S_t. Si le système se trouve à l'étape t dans l'état $s \in S_t$, l'ensemble des actions disponibles est noté $A_{s,t}$. Tout comme T, S_t et $A_{s,t}$ peuvent être finis ou infinis, discrets ou continus.

Lorsque l'agent effectue l'action $a \in A_{s,t}$ dans l'état s et à l'instant t, il reçoit une récompense $r_t(s, a)$, pouvant être positive ou négative, additive ou non.
Dans le même temps, la connaissance de l'évolution du système peut changer de manière *déterministe*, *non déterministe* ou *probabiliste* (dans la seconde partie de cette thèse il pourra évoluer de manière "possibiliste"). On note $w_t(s, a)$ la *fonction de transfert* du système.

- Dans le cas déterministe, l'état successeur est parfaitement connu : $w_t(s, a) \in S_{t+1}$ est l'état résultant de l'application de l'action a dans l'état s à l'instant t. Dans le cas non déterministe, $w_t(s, a)$ est un sous-ensemble de S_{t+1}.
- Dans le cas probabiliste, comme dans la théorie de von Neumann et Morgenstern [vNM44], l'état successeur est mal connu et l'incertitude est représentée par une distribution de probabilité $w_t(s, a)$ sur S_{t+1}, encore notée $p_t(\cdot|s, a)$. Notons que le cas déterministe peut être vu comme un cas particulier du cas probabiliste (mais pas le cas non-déterministe).

On appelle *règle de décision* une fonction $d_t : S_t \to A_{s,t}$ qui associe une action à chaque état possible du monde à l'étape t. L'ensemble de ces règles de décision est noté D_t. Une *stratégie* est

une séquence de règles de décision, applicables tout au long de l'horizon du problème. Une stratégie δ est une séquence, finie ou infinie, de règles de décision : $\delta = \{d_1, d_2, \ldots, d_N\}$ où $d_t \in D_t$ pour $t = 1, \ldots, N$. $\Delta = D_1 \times D_2 \times \ldots \times D_N$ est l'ensemble des stratégies applicables (dans le cas fini).
Dans le cas particulier où les ensembles et fonctions précédemment décrits ne dépendent pas du paramètre t on parle de problème décisionnel markovien *stationnaire*. Un exemple typique est celui de la navigation d'un robot dans un environnement statique.

Adopter la notion de règle de décision suppose qu'à chaque étape l'agent *observe* l'état du monde s_t dans lequel il se trouve. On parle dans ce cas *d'observabilité totale*.
Il se peut également que l'agent doive construire la séquence d'actions qu'il va appliquer, avant même d'appliquer la première d'entre elles. Dans ce cas, il ne possède aucune autre information pour choisir l'action a_t à appliquer à l'étape t que la donnée de l'état du monde initial s_1 et de la séquence d'actions déjà choisies $\{a_1, \ldots, a_{t-1}\}$. On parle dans ce cas *d'observabilité nulle*.

Enfin, des cas intermédiaires peuvent se présenter dans lesquels l'agent n'a pas accès à la connaissance de l'état réel du monde s_t, mais peut observer une manifestation du système $o_t \in O_t$, O_t représentant toutes les manifestations possibles du système à cette étape. L'état réel du système, s_t, et la manifestation o_t sont reliés, par exemple par une distribution de probabilité sur $S_t \times O_t$: $(s_t, o_t) \to p_t(s_t, o_t)$ dans le cas probabiliste. On parle dans ce cas *d'observabilité partielle*. Cette observation peut également dépendre de l'action a_t exécutée.

En résumé, les différents éléments d'un problème de décision séquentiel markovien peuvent avoir les caractéristiques suivantes :

- *Horizon* : une seule étape, nombre fini ou infini d'étapes (discret ou continu).
- *Actions* : détermistes, non-déterministes, probabilistes...
- *Utilités* : réelles ou non, additives ou non, positives ou négatives...
- *Observabilité* : totale, partielle ou nulle.

3.2.2 Valeur d'une stratégie dans le cas observable, déterministe ou probabiliste

Lorsque le problème est déterministe, l'agent spécifie entièrement l'évolution du système en choisissant une stratégie δ. Une fois δ choisie, si l'état initial du monde est s_1^δ, la *trajectoire* du système $\{s_1^\delta, \ldots, s_N^\delta\}$ (l'ensemble des états qu'il occupe tout au long de l'horizon du problème) est entièrement déterminée par l'équation fonctionnelle :

$$s_{t+1}^\delta = \delta(s_t^\delta).$$

Comment évaluer une stratégie δ dans le cas déterministe à horizon fini $\{1, 2, \ldots, N\}$?
Dans l'état s_1^δ l'agent reçoit une récompense $r_1(s_1^\delta, d_1(s_1^\delta)) = r_1(s_1^\delta, s_2^\delta)$, à la seconde étape il reçoit une récompense $r_2(s_2^\delta, d_2(s_2^\delta))$, etc...
Plusieurs évaluations sont possibles suivant que l'on considère que les récompenses s'additionnent au cours du temps (éventuellement pondérées par un facteur d'amortissement avec le temps lorsque l'on considère un horizon infini) ou que l'on prend la moyenne sur un nombre fini d'étapes (nombre éventuellement plus petit que l'horizon du problème) ou encore, dans une optique pessimiste, que l'on se concentre sur la "pire" récompense obtenue le long de la trajectoire.

Dans le cas où les récompenses s'additionnent, la valeur totale d'une stratégie δ est définie par :

Définition 3.2.1 *Valeur d'une stratégie δ dans le cas déterministe :*

$$v_N^\delta(s_1) = \sum_{t=1}^{N} r_t(s_t^\delta, d_t(s_t^\delta)).$$

Dans le cas probabiliste la stratégie δ ne détermine pas une simple trajectoire, mais permet de définir un ensemble de probabilités de transition entre états : les états successifs occupés par le système (l'état réel du monde) sont décrits par une *chaîne de Markov*. Ainsi, δ génère un ensemble de trajectoires de la forme $\{s_1^\delta, s_2^\delta, \ldots, s_N^\delta\}$ qui seront éventuellement parcourues par l'état réel du système, avec probabilité $p(s_2^\delta|s_1, d_1(s_1^\delta)) \times \ldots \times p(s_{N+1}^\delta|s_N, d_N(s_N^\delta))$, où s_{N+1}^δ est un état fictif, ajouté au problème.

Dans ce cas la valeur d'une stratégie n'est pas la simple somme des récompenses obtenues le long d'une trajectoire, mais l'espérance mathématique de la somme de ces récompenses, le long de la chaine de Markov :

Définition 3.2.2 *Valeur d'une stratégie δ dans le cas probabiliste :*

$$v_N^\delta(s_1) = E_{\delta,s_1}(\sum_{t=1}^{N} r_t(s_t^\delta, d_t(s_t^\delta))).$$

E_{δ,s_1} représente l'espérance mathématique, relativement à la distribution de probabilité sur l'ensemble des trajectoires, générée par δ, conditionnellement à l'état initial s_1.

Dans tous les cas, l'objectif de l'agent est de trouver une stratégie δ^* maximisant la valeur de $v_N^\delta(s_1)$, i.e. telle que $v_N^{\delta^*}(s_1) = sup_{\delta\in\Delta} v_N^\delta(s_1)$. Dans les cas où S_t et $A_{s,t}$ sont finis pour tout s et t on sait qu'une telle stratégie optimale existe.
Dans le paragraphe suivant nous exposerons une méthode de "programmation dynamique" permettant de calculer une stratégie optimale dans le cas probabiliste par recherche arrière par induction, mais avant voyons un exemple de problème de décision séquentielle tiré de (Puterman [Put87]).

Exemple 3.2.1 *Chaque mois le directeur d'un entrepôt doit déterminer quelle quantité d'un certain produit il commande, afin de garnir son stock pour satisfaire la demande (aléatoire) de ses clients. Son objectif est de maximiser l'espérance mathématique de son profit (revenu des ventes moins coûts d'inventaire et de commande). La demande mensuelle est aléatoire, de distribution de probabilité connue.*

Les hypothèses simplificatrices suivantes sont posées :

- *Les commandes et livraisons au stock sont exécutées simultanément, en début de mois.*
- *Les demandes des clients sont livrées en fin de mois.*
- *Si la demande des clients excède le stock, seule la valeur du stock est livrée et payée.*
- *Le produit est acheté et vendu par lots (unités entières).*
- *La capacité de l'entrepôt est de M unités (lots).*
- *L'horizon du problème est $\{1, 2, \ldots, N\}$.*

s_t *représente la quantité de stock à l'étape* t, a_t *représente la quantité additionnelle commandée au fournisseur et* D_t *la demande totale des clients à l'étape* t. $p_d = P\{D_t = d\}$ *représente la probabilité (connue) que la demande totale soit de* d *unités à l'étape* t. $c(u)$ *représente le coût d'une commande de* u *unités au fournisseur et* $h(u)$ *le coût de stockage de* u *unités. Enfin,* $f(u)$ *représente le revenu de la vente de* u *unités.*
L'espérance mathématique du revenu de la vente sachant que le stock initial est de u *unités, notée* $F(u)$, *est déterminée par la formule suivante :*

$$F(u) = \sum_{d=0}^{u-1} f(d) \cdot p_d + f(u) \cdot P\{D_t \geq u\}.$$

Le gain espéré $r_t(s_t, a_t)$ *est donné par la formule :*

$$r_t(s_t, a_t) = F(s_t + a_t) - c(a_t) - h(s_t + a_t).$$

Si l'horizon est $T = \{1, 2, \ldots, N\}$, $r_{N+1}(s_{N+1}, a_{N+1}) = g(s_{N+1})$ *est la valeur du stock à la fin de la période.*

Prenons les valeurs numériques proposées par Puterman [Put87] :

- *Le coût d'une commande de* u *unités au fournisseur est* $c(u) = 4 + 2 \cdot u$ *si* $u > 0$ *et* $c(u) = 0$ *sinon,*
- *le coût de stockage de* u *unités est* $h(u) = u$,
- *la capacité de stockage est* $M = 3$,
- *l'horizon est* $T = \{1, 2, 3\}$,
- *le revenu de la vente de* u *unités est* $f(u) = 8 \cdot u$ *et*
- *les probabilités que 0, 2 ou 1 unités soient commandés par les clients sont* $p_0 = p_2 = 1/4$, $p_1 = 1/2$.

Après calcul, on obtient $F(0) = 0$, $F(1) = 6$, $F(2) = F(3) = 8$.
$r_t(s, a)$ *est donnée par le tableau suivant :*

s	*a = 0*	*a = 1*	*a = 2*	*a = 3*
0	*0*	*-1*	*-2*	*-5*
1	*5*	*0*	*-3*	$\times$
2	*6*	*-1*	$\times$	$\times$
3	*5*	$\times$	$\times$	$\times$

et $p_t(s'|s, a)$ *par :*

s+a	*s' = 0*	*s' = 1*	*s' = 2*	*s' = 3*
0	*1*	*0*	*0*	*0*
1	*3/4*	*1/4*	*0*	*0*
2	*1/4*	*1/2*	*1/4*	*0*
3	*0*	*1/4*	*1/2*	*1/4*

Nous verrons comment cet exemple peut être résolu par une méthode de programmation dynamique dans le paragraphe suivant.

3.2.3 Programmation dynamique

Les méthodes de programmation dynamique se basent sur un principe simple : décomposer un problème de décision séquentiel en une série de problèmes simples (d'une seule étape), interdépendants. Pour ce faire on utilise le *Principe d'Optimalité* de Bellman [Bel57] selon lequel une solution optimale peut être trouvée en résolvant l'équation fonctionnelle reliant la valeur d'un problème comportant $t+1$ étapes à un problème en reliant t.
Le principe de Bellman est valide dès qu'il existe une certaine forme de séparabilité de la fonction de valeur.

Parmi les méthodes de programmation dynamique existantes, nous ne décrirons que la méthode de recherche par induction arrière qui est la seule à fonctionner pour les processus décisionnels markoviens probabilistes à horizon fini. D'autres méthodes, notamment basées sur une recherche avant fonctionnent pour le cas déterministe. Le cas où l'horizon est infini requiert des méthodes de résolution spécifiques, et il ne sera pas évoqué ici. On pourra se référer à (Puterman [Put87]) pour des exemples de méthodes de programmation dynamique autres que par recherche par induction arrière.

Equation fonctionnelle de la programmation dynamique

On note $v^t(s)$ la somme *maximale* des récompenses que peut recevoir l'agent entre les étapes t et $N+1$ (en choisissant une stratégie optimale). Lorsque l'évolution du problème est probabiliste et non déterministe, $v^t(s)$ représente l'espérance mathématique maximale de cette somme.

L'équation fonctionnelle de la programmation dynamique, encore appelée équation de Bellman relie les valeurs des stratégies optimales $v^t(s)$ et $v^{t+1}(s)$ aux étapes t et $t+1$. Dans le cas déterministe, elle est définie par :

Définition 3.2.3 *Equation de Bellman dans le cas déterministe.*
$\forall t \in 1 \dots N, \forall s \in S_t$,

$$v^t(s) = \max_{a \in A_{s,t}} \{r_t(s,a) + v^{t+1}(w_t(s,a))\}$$

et $v^{N+1}(s) = 0, \forall s \in S_{N+1}$

Cette équation est la même que celle que l'on rencontrerait dans un problème comportant une seule étape, dans lequel l'agent (initialement dans l'état s) doit choisir une action $a^* \in A_{s,t}$ maximisant la somme de la récompense immédiate $r_t(s,a)$ et de la “valeur” de l'état résultant $w_t(s,a)$.

Cette méthode de choix diffère de la méthode “naïve” évoquée au début de ce chapitre dans la mesure où, pour choisir une action à l'étape t, elle prend en compte non seulement la récompense $r_t(s,a)$ obtenue immédiatement, mais aussi les récompenses qui pourront être obtenues aux étapes suivantes, par l'intermédiaire de $v^{t+1}(w_t(s,a))$.

Dans le cas probabiliste, l'équation est modifiée afin de tenir compte des probabilités de transition entre états :

Définition 3.2.4 *Equation de Bellman dans le cas probabiliste.*
$\forall t \in 1, \dots, N, \forall s \in S_t$,

$$v^t(s) = \max_{a \in A_{s,t}} \{r_t(s,a) + \sum_{s' \in S_{t+1}} p_t(s'|s,a) \cdot v^{t+1}(s')\}$$

Dans ce cas, on cherche à maximiser, à chaque étape, la somme de la récompense obtenue immédiatement, et de l'espérance mathématique des valeurs des états successeurs possibles de s par a.

Dans le paragraphe suivant nous allons voir comment on peut évaluer une stratégie δ et comment on peut calculer une stratégie optimale δ^* dans les cas déterministe et probabiliste, par une méthode de recherche arrière par induction.

Calcul de stratégies optimales par recherche arrière par induction

Puterman [Put87] a proposé l'algorithme suivant, utilisant de manière itérative les définitions 3.2.3 page précédente et 3.2.4 page ci-contre afin de calculer la valeur d'une stratégie $\delta = \{d_1, \ldots, d_N\}$ ou une stratégie optimale δ^* dans les cas déterministe et probabiliste :

Algorithme 3.1 : Evaluation d'une stratégie $\delta = \{d_1, \ldots, d_N\}$

début
- $t \leftarrow N + 1$;
- $v^t \leftarrow 0$;
- $\forall s, v^{N+1}(s) \leftarrow 0$;
- **tant que** $N \geq 2$ **faire**
  - $t \leftarrow t - 1$;
  - **pour** $s \in S_t$ **faire**
    - % Dans le cas déterministe ;
    - $v^t(s) \leftarrow r(s, d_t(s)) + v^{t+1}(w_t(s, d_t(s)))$;
    - % Dans le cas probabiliste ;
    - $v^t(s) \leftarrow r(s, d_t(s)) + \sum_{s' \in S_{t+1}} p(s'|s, d_t(s)) \cdot v^{t+1}(s')$;
- **retourner** $v^1(s_1)$

fin

Cet algorithme évalue la valeur d'une stratégie par recherche arrière par induction.
Il peut être modifié afin de pouvoir déterminer une stratégie optimale.

Algorithme 3.2 : Calcul de l'ensemble des stratégies optimales

début
- $t \leftarrow N + 1$;
- $v^t \leftarrow 0$;
- $\forall s, v^{N+1}(s) \leftarrow 0$;
- **tant que** $N \geq 2$ **faire**
  - $t \leftarrow t - 1$;
  - **pour** $s \in S_t$ **faire**
    - `Déterminer l'ensemble` $A^*_{s,t}$ `des actions maximisant` $v_t(s)$
- **retourner** $v^1(s_1)$, $\{\delta^* = \{a^*_1, \ldots, a^*_N\}, a^*_t \in A^*_{s,t}\}$

fin

Afin de déterminer l'ensemble $A^*_{s,t}$ des actions optimales dans l'état s, à l'étape t, on calcule $v^t(s)$ par l'une des définitions 3.2.3 page précédente ou 3.2.4 page ci-contre, et on conserve les actions permettant d'atteindre la valeur $v^t(s)$.

Exemple 3.2.2 *(suite de l'exemple 3.2.1) :*

Dans l'exemple (probabiliste) de l'entrepôt on définit tout d'abord $v^t(s,a)$, la valeur de l'action a à l'etape t, dans l'état s, par :

$$v^t(s,a) = r_t(s,a) + \sum_{s' \in S} p_t(s'|s,a) \cdot v^{t+1}(s')$$

Appliquons l'algorithme 3.1. Au départ, $t = 4, v^4(s) = 0, \forall s \in \{0,1,2,3\}$. Ensuite, en appliquant l'algorithme, on obtient les résultats suivants :

– $t = 3$

s	$v^3(s)$	$A^*_{s,3}$
0	0	$\{0\}$
1	5	$\{0\}$
2	6	$\{0\}$
3	5	$\{0\}$

– $t = 2$

s	$v^2(s)$	$A^*_{s,2}$
0	2	$\{2\}$
1	$6,25$	$\{0\}$
2	10	$\{0\}$
3	$10,5$	$\{0\}$

– $t = 1$

s	$v^1(s)$	$A^*_{s,1}$
0	$67/16$	$\{3\}$
1	$129/16$	$\{0\}$
2	$194/16$	$\{0\}$
3	$227/16$	$\{0\}$

La stratégie optimale peut être exprimée très simplement dans cet exemple : "Tant qu'il reste du produit en stock on n'en commande pas. Lorsqu'il n'en reste plus, à l'étape 1 on commande 3 unités et à l'étape 2, 2 unités.

Conclusion

Nous avons décrit dans cette section une méthode de programmation dynamique par recherche arrière inductive, permettant de résoudre des problèmes de décision séquentielle :

– à horizon fini,

– avec actions déterministes ou probabilistes,

– avec fonctions d'utilité (récompenses) additives,

– et observabilité totale.

Dans la seconde partie de cette thèse nous conserverons l'hypothèse d'horizon fini, mais nous examinerons les cas où les actions sont non-déterministes ou possibilistes. Dans ces cas, les actions déterminent des *possibilités* de transition et non des probabilités. Les fonctions d'utilité que nous considérerons ne seront pas additives, mais seulement ordinales, donc exprimables elles aussi par des distributions de possibilités.

Jusqu'ici nous avons fait l'hypothèse que les problèmes de décision concernés sont totalement observables. Cela signifie qu'après l'exécution d'une action, l'agent a une connaissance complète de l'état du monde résultant.
Dans la suite de ce chapitre, nous allons relâcher cette hypothèse en supposant que l'agent n'acquiert pas une connaissance complète à chaque étape mais qu'il observe seulement certaines manifestations du monde réel en résultat de ses actions.

3.3 Observation, processus décisionnels markoviens partiellement observables

3.3.1 Qu'est-ce qu'une observation ?

Dans la plupart des problèmes de décision séquentielle dans l'incertain, un agent n'a pas une connaissance complète du monde réel mais il acquiert, au fur et à mesure de ses actions, des informations sur le monde. Un robot évoluant dans un environnement mal connu et devant atteindre un objectif fixé, par exemple, devra alterner des phases de déplacement et d'acquisition d'information. Les informations acquises par l'intermédiaire de senseurs lui permettront de choisir ensuite des actions plus efficaces pour atteindre son but.

Que sont ces phases d'acquisition d'information ? Sont-elles des actions comme les autres ? Peuvent-elles également être évaluées en mesurant une forme d'utilité espérée ?
Nous allons donner certaines réponses à ces questions dans le cadre de la décision sous incertitude probabiliste et des processus décisionnels markoviens *partiellement observables*.

En premier lieu, demandons nous ce qu'est une observation. Tout comme l'on distingue les actions déterministes des actions probabilistes (entre autres), on distingue aussi les observations déterministes des observations probabilistes.
Considérons une observation obs. On peut supposer, par exemple, que celle-ci, en fonction de l'état réel $s \in S_t$, donne un résultat (une manifestation du monde réel) $o \in O_t$, O_t étant l'ensemble des manifestations possibles de toutes les observations disponibles à l'étape t.

Dans un exemple donné par Savage [Sav54], un consommateur souhaitant acheter du raisin demande à l'épicier de goûter un grain. Sur la base du résultat de cette observation, il va décider de la quantité de raisin qu'il va acheter, en fonction de la satisfaction qu'il espère en retirer. Dans cet exemple, le lien entre le résultat de l'observation (qualité du grain goûté) et l'état du monde réel (qualité moyenne du raisin) n'est pas forcément "déterministe" : la qualité du grain goûté ne reflète pas forcément la qualité moyenne du raisin...

Reprenons nos considérations générales sur la notion d'observation en commençant par les observations appelées “problèmes de partition” chez Savage. Une telle observation consiste à réduire l'ensemble des états possibles du monde en observant la valeur d'une variable observable qui dépend de manière déterministe de l'état réel du monde (chaque état possible du monde renvoie une seule valeur de la variable observable). Nous qualifierons ces observations de *déterministes*.
Formellement, une observation déterministe obs est une application de l'ensemble des états possibles du monde, S_t, dans l'ensemble des résultats d'observation $O_t = \{o_1, \dots, o_p\}$. Cette application réalise une *partition* de S_t :

$$S_t = \bigcup_{i \in 1..n} S_i, \forall s \in S_i, obs(s) = o_i.$$

obs est déterministe dans la mesure où, répétée plusieurs fois, elle donnera toujours le même résultat.

Une observation probabiliste, par contre, ne donnera pas toujours un résultat identique dans le même état du monde. Le résultat de l'observation est lié à l'état du monde par une distribution de probabilité jointe $p_{obs}(s_i, o_j)$ sur $S_t \times O_t$.
Le cas déterministe peut être vu comme un cas particulier, dans lequel $p_{obs}(s_i, o_j) = 1$ si $o_j = obs(s_i)$ et 0 sinon.

Après cet aperçu de la notion d'observation nous allons voir comment s'en servir afin de choisir des actions plus “utiles” en environnement partiellement observable.
Deux approches différentes sont envisageables :

- l'approche de Savage qui pour évaluer l'utilité d'une observation obs, fait la différence entre les utilités espérées d'un couple observation-action optimal et d'une action optimale sans observation,

- l'approche utilisée dans le cadre de *processus décisionnels markoviens partiellement observables* (PDMPO) (Cassandra et al. [CKL94]), (Monahan [Mon82]), (Lovejoy [Lov91]) qui ne distinguent pas les observations des autres actions mais suggèrent que toute action procure en retour une observation. Ces auteurs proposent de transformer un PDMPO discret et fini en PDM continu dont les *états de connaissance* (distributions de probabilité sur les états possibles du monde) remplacent les états du monde classiques.

3.3.2 Observer, puis agir

Dans cette approche exposée par Savage [Sav54] l'agent observe d'abord, puis en fonction du résultat de son observation il choisit une action à appliquer.

Rappelons en premier lieu la définition de $v^t(s, a)$, la valeur de l'action a dans l'état s, dans un processus décisionnel markovien totalement observable :

$$v^t(s, a) = r_t(s, a) + \sum_{s' \in S_{t+1}} p(s'|s, a) \cdot v^{t+1}(s').$$

Lorsque l'état du monde réel est inconnu, mais que l'observation obs a été effectuée et a donné le résultat o_i, on peut déterminer une “valeur moyenne” des valeurs des états compatibles avec le résultat d'observation o_i :

$$v^t(o_i, a) = \sum_{s \in S_t} p(s|o_i, obs) \cdot v^t(s, a),$$

où $p(s|o_i, obs)$ se déduit de p_{obs} par $p(s|o_i, obs) = \frac{p_{obs}(s, o_i)}{p_{obs}(o_i)}$ (où $p_{obs}(o_i) = \sum_{s \in S_t} p_{obs}(s, o_i)$).

$v^t(o_i, a)$ représente l'utilité de l'action a une fois que le résultat o_i de l'observation obs a été obtenu. On note ensuite $v^t(o_i) = \max_{a \in A_{s,t}} v^t(o_i, a)$, la "valeur" du résultat o_i, soit l'utilité de la meilleure action applicable après que o_i ait été observé.

Enfin Savage définit l'utilité d'une observation comme l'espérance mathématique des valeurs de ses résultats possibles :

$$v^t(obs) = \sum_{o_i \in O_t} p_{obs}(o_i) \cdot v^t(o_i).$$

L'utilité réelle d'une observation obs est déterminée en comparant la valeur de $v^t(obs)$ (observation puis action optimale fonction du résultat) et $v^t(s)$, l'utilité d'une action optimale *sans observation*.

Exemple 3.3.1 *Valeur d'une observation* (Savage [Sav54]).

Dans le problème de l'achat de raisin, celui-ci peut être de qualité inférieure (1), moyenne (2) ou supérieure (3). Le consommateur évalue a priori les probabilités de ces trois qualités :

Q(*qualité*)	1	2	3
P(*probabilité*)	$1/4$	$1/2$	$1/4$

Il a quatre actions à sa disposition : acheter 0, 1, 2 ou 3 grappes de raisin. Les valeurs des conséquences de ses actions suivant la qualité du raisin sont résumées dans le tableau suivant, ainsi que l'espérance mathématique des différentes actions :

	Q			
a	1	2	3	$v(a)$
0	0	0	0	0
1	-1	1	3	1
2	-3	0	5	$1/2$
3	-6	-2	6	-1

L'épicier lui propose de goûter un grain de raisin. Etant un connaisseur, le client est capable de classer la qualité du raisin en cinq catégories que nous noterons, par qualité croissante de 1 à 5.
Le tableau suivant contient la distribution de probabilité jointe sur $O \times Q$, où O est l'ensemble des 5 qualités observables par le consommateur. Pour éviter de manipuler des fractions, les probabilités sont multipliées par un même facteur 128.

	Q			
o	1	2	3	$128 \cdot p(o)$
1	15	5	1	21
2	10	15	2	27
3	4	24	4	32
4	2	15	10	27
5	1	5	15	21
	32	64	32	128
	$128 \cdot p(Q)$			

De ces données on déduit aisément les valeurs $v^t(o_i, a_j)$ des différentes actions suivant l'observation faite :

a \ o	1	2	3	4	5
0	$0/21$	$0/27$	$0/32$	$0/27$	$0/21$
1	$-7/21$	$11/27$	$32/32$	$43/27$	$49/21$
2	$-40/21$	$-20/27$	$8/32$	$44/27$	$72/21$
3	$-94/21$	$-78/27$	$-48/32$	$18/27$	$74/21$

De cette table on déduit les actions optimales après observation : ne pas acheter de raisin si la qualité observée est $o = 1$, une grappe si c'est $o = 2$ ou 3, deux si c'est 4 et trois si c'est 5.

L'utilité de l'observation est évaluée à $(0 + 11 + 32 + 44 + 74)/128 \sim 1,26$ à laquelle on doit retrancher 1, soit la valeur de l'action optimale sans observation (acheter une grappe).

Ainsi, en connaissant la valeur des états successeurs $s' \in S_{t+1}$ on peut calculer, par recherche arrière, une observation optimale obs, et un ensemble d'actions optimales, respectivement associées aux différents résultats possibles de l'observation obs.

Le problème de la méthode de Savage est qu'elle est difficilement itérable : de v^{t+1} sur S_{t+1} on déduit v^t sur O_t, mais pas sur S_t !
En d'autres termes, la méthode proposée par Savage permet de passer d'une utilité sur l'espace des états du monde à l'étape $t + 1$ à une utilité sur l'espace des observations (ou *états de connaisance*) à l'étape t. Dans le paragraphe suivant, nous allons exposer la méthode de Cassandra et col. [CKL94] qui est itérative et dont le principe consiste à modifier l'espace d'états S_t en un espace *d'états de connaissance*, qui devient malheureusement infini. Ainsi, chez Cassandra et col. on résout le problème en restant au niveau des états de connaissance.

3.3.3 Etat du monde contre état de connaissance

Le concept de *Processus Décisionnel Markovien Partiellement Observable* (PDMPO) a été introduit par Monahan [Mon82] et Lovejoy [Lov91] dans le cadre de la *Recherche Opérationnelle* puis Cassandra et col. [CKL94] ont proposé d'utiliser ce modèle pour la décision séquentielle sous incertitude avec observabilité partielle. Un PDMPO comprend, outre les ensembles d'états S_t et d'actions $A_{s,t}$ pour chaque étape, un ensemble *d'observations* $O_t = \{o_1, \ldots, o_k\}$.

Contrairement à l'approche de Savage, Cassandra et col. ne mettent pas en jeu d'actions d'acquisition d'informations mais considèrent que chaque action a dans l'état s_t ne change pas seulement l'état du monde (avec une distribution de probabilité $p(\cdot|s, a)$), mais produit également un résultat d'observation qui est une manifestation du système en retour de l'action. Cette manifestation, fonction de l'action effectuée et de l'état résultant, peut elle-même être non-déterministe, auquel cas elle est représentée par une distribution de probabilité $p'(\cdot|s', a)$ sur O_{t+1}.

A chaque étape, l'agent ne connait pas l'état exact du système. La seule information à laquelle il a accès est ce résultat d'observation. En revanche, il peut estimer à chaque étape une probabilité de se trouver dans tel ou tel état du monde. Sa connaissance est résumée par une distribution de probabilité b_t (pour "belief state" ou *état de croyance* en français) sur S_t. L'ensemble des états de croyance à

1	2	BUT	4

FIG. 3.1 – *Environnement partiellement observable*

l'étape t est noté B_t.

Cassandra et col. proposent de transformer le PDMPO en un PDM totalement observable dont les états sont les états de croyance. Ils commencent par déterminer une nouvelle équation fonctionnelle. La première chose à faire pour cela est de déterminer le nouvel état de connaissance en fonction du précédent, de l'action exécutée et de l'observation reçue en retour. La nouvelle distribution de probabilité est déterminée par :

$$p(s'|b_t, a, o) = \frac{p(o|s', a) \cdot \sum_{s \in S_t} p(s'|s, a) \cdot b_t(s)}{p(o|b_t, a)}$$

où $p(o|b_t, a) = \sum_{s' \in S_{t+1}} p(o|s', a) \cdot (\sum_{s \in S_t} p(s'|s, a) \cdot b_t(s))$ est un facteur de normalisation. $p(\cdot|b_t, a, o)$ n'est autre que la fonction b_{t+1} sur S_{t+1} ou état de connaissance à l'étape $t + 1$, résumant la nouvelle connaissance de l'agent qui était initialement b_t, après l'action a et l'observation o.

Exemple 3.3.2 *(Cassandra et col. [CKL94])*
Un agent se trouve dans le milieu dessiné Figure 3.1. Celui-ci comporte quatre états dont l'un est un état but. L'agent se trouve, à chaque étape, dans un de ces états et peut aller dans une des cases adjacentes, soit à gauche (g) soit à droite (d). Si au cours d'un de ses déplacements il rencontre un mur, il reste à sa place initiale. Si un de ses déplacements l'amène sur la case but, il reçoit une récompense additionnelle de 1 puis se retrouve aléatoirement sur une des cases non-but (avec une même probabilité $1/3$). Notons que ce problème est stationnaire.
L'agent ne sait pas où il se trouve à chaque étape, mais observe s'il se trouve sur la case but ou non. Initialement il sait qu'il n'est pas dans l'état but : $b_0 =< 1/3, 1/3, 0, 1/3 >$. Après un déplacement à droite, si l'agent n'observe pas l'état but, son état de connaissance est $b_1 =< 0, 1/2, 0, 1/2 >$ (il ne peut plus être dans la case de gauche). Après un nouveau déplacement à droite, s'il n'observe toujours pas le but, il est dans l'état de connaissance $b_2 =< 0, 0, 0, 1 >$, i.e. il est certain d'être dans la case de droite. Si à un moment il observe le but, il reçoit une récompense de 1 et se retrouve dans l'état de connaissance b_0.

Construire une stratégie optimale pour un PDMPO est un problème difficile. Il est même difficile de spécifier une stratégie quelconque puisque l'espace d'états fini S_t est remplacé par un espace d'états de croyance infini, B_t!
Watkins et Dayan [WD92] ou Chrisman [Chr92] suggèrent une méthode simple pour trouver une solution à un PDMPO : il s'agit de choisir, dans l'état de connaissance b_t, l'action maximisant la moyenne des valeurs $v^t(s, a)$ (calculées comme dans le cas totalement observable), pondérées par les coefficients $b_t(s)$,

$$a^* = argmax_{a \in A_{s,t}} \sum_{s \in S_t} b_t(s) \cdot v(s, a).$$

Malheureusement cette méthode ne prend pas en compte les observations et ne donne pas, en général, une stratégie optimale.

Le point-clé pour résoudre un PDMPO est de le transformer en un PDM totalement observable dont les espaces d'états sont les B_t (Cassandra et col. [CKL94]). La probabilité de transition entre états de croyance peut être calculée :

$$p(b'|b,a) = \sum_{o \in O, p(\cdot|a,o,b)=b'(\cdot)} p(o|b,a)$$

où $p(o|b,a)$ se calcule comme précédemment.

La fonction récompense ρ sur les états de croyance se déduit de la fonction récompense r sur les états par :

$$\rho(b,a) = \sum_{s \in S_t} b(s) \cdot r(s,a).$$

Enfin, la valeur d'un état de connaissance s'exprime grâce à une équation fonctionnelle similaire au cas observable :

$$v^t(b_t) = \max_{a \in A}(\rho(b,a) + \sum_{b' \in B_{t+1}} p(b'|b,a) \cdot v^{t+1}(b')).$$

Voilà donc comment on peut transformer un PDMPO en un PDM totalement observable. Malheureusement, le PDM obtenu est continu, ce qui rend sa résolution pratique très difficile (problème PSPACE-complet dans la hiérarchie polynomiale (Papadimitriou et Tsitsiklis [PT87])).
De nombreux auteurs ont proposé des algorithmes de résolution approchée de PDMPO : Smallwood et Sondik [SS73], Monahan [Mon82], Cheng [Che88]. Ces algorithmes construisent une fonction valeur approchée $\hat{v}^t(b)$, puis améliorent peu à peu cette valeur. Cassandra et col. proposent un algorithme de cette lignée. D'autres auteurs comme Parr et Russel [PR95], Brafman [Bra97] ou Hauskrecht [Hau97] discrétisent le PDM continu et se contentent de prendre en compte un sous-ensemble fini de B_t.

Exemple 3.3.3 *(suite et fin de l'exemple 3.2.2)*
L'exemple de Cassandra et al. est particulièrement simple : il ne comporte que trois états et les actions sont déterministes. C'est pourquoi il peut être résolu de façon exacte, d'autant plus que l'on peut aisément vérifier qu'en partant de l'état de connaissance initial b_0, seuls cinq états de connaissance sont accessibles : $b_0, b_1, b_2, b_3 =< 1, 0, 0, 0 >$ et $b_4 =< 0, 1, 0, 0 >$.
Les actions "aller à droite" (d) et "aller à gauche" (g) sont représentées par les fonctions de transition stochastiques entre états de connaissance suivantes :

g	b_0	b_1	b_2	b_3	b_4
b_0	$1/3$	0	0	$2/3$	0
b_1	$1/2$	0	0	$1/2$	0
b_2	1	0	0	0	0
b_3	0	0	0	1	0
b_4	0	0	0	1	0

d	b_0	b_1	b_2	b_3	b_4
b_0	$1/3$	$2/3$	0	0	0
b_1	$1/2$	$1/2$	0	0	0
b_2	0	0	1	0	0
b_3	0	0	0	0	1
b_4	1	0	0	0	0

Plaçons nous en horizon fini, comportant quatre étapes. A l'étape quatre, l'état b_0 rapporte une unité ($v^4(b_0) = 1$) et les autres ne rapportent rien ($v^4(b_i) = 0, \forall i \neq 0$).
En appliquant la méthode de recherche arrière par induction, on obtient les stratégies optimales et

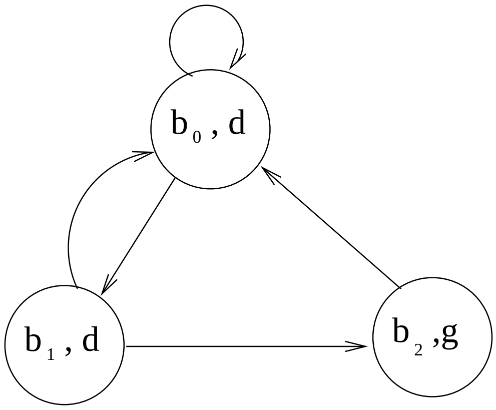

FIG. 3.2 – *Stratégie stationnaire optimale*

les valeurs associées :

$v^t(b_i)$	$t = 4$	$t = 3$	$t = 2$	$t = 1$
b_0	$1, \{g, d\}$	$4/3, \{g, d\}$	$16/3, \{d\}$	$23/9, \{d\}$
b_1	$0, \{g, d\}$	$1/2, \{g, d\}$	$7/6, \{d\}$	$10/3, \{d\}$
b_2	$0, \{g, d\}$	$1, \{g\}$	$4/3, \{g\}$	$16/3, \{g\}$
b_3	$0, \{g, d\}$	$0, \{g, d\}$	$1, \{d\}$	$4/3, \{d\}$
b_4	$0, \{g, d\}$	$1, \{d\}$	$4/3, \{d\}$	$16/3, \{d\}$

Lorsque l'horizon croît, on trouve rapidement une stratégie stationnaire optimale (celle des colonnes $t = 2$ *et* $t = 1$*) qui peut se résumer dans la figure 3.2 :*

3.4 Conclusion

Dans ce chapitre, nous avons vu comment on pouvait étendre la théorie de l'utilité espérée à la prise de décisions *séquentielles*. Cette extension, basée sur la théorie des *processus décisionnels markoviens* (PDM), permet une application de la théorie de la décision dans l'incertain à la planification sous incertitude.

Passer d'une seule étape de décision à une séquence de décisions interdépendantes pose de nouveaux problèmes demandant, entre autres, l'introduction d'une propriété nouvelle *d'observabilité*. Nous avons vu comment Savage décrit cette notion et évalue la "valeur" d'une observation en termes de valeur ajoutée à une action optimale. Ensuite, nous nous sommes replacés du point de vue des PDM *partiellement observables* dans lesquels la notion d'observation est intégrée directement à la notion d'action, permettant une itération du processus de décision, au prix d'un accroissement significatif du coût d'évaluation (en temps de calcul et en espace mémoire) des *stratégies* optimales.

Dans la seconde partie de cette thèse et en particulier dans le chapitre 2 page 109, nous étendrons la théorie de *l'utilité qualitative possibiliste* à la prise de décisions séquentielles (voir aussi [FLS96a], [DFL$^+$96], [FLS98]). Nous proposerons des méthodes de programmation dynamique permettant une évaluation inductive par recherche arrière de stratégies optimales au sens du critère possibiliste, dans les cas observable et partiellement observables.

Nous montrerons en particulier une propriété intéressante : lorsque l'espace d'états d'un PDM partiellement observable est fini, l'espace *d'états de connaissance* du PDM totalement observable correspondant reste fini. Cette propriété permet dans le cadre possibiliste, d'évaluer des stratégies optimales en environnement partiellement observable à un coût "raisonnable".

Mais avant de passer à la seconde partie de cette thèse qui décrit notre propre travail, il reste un point à évoquer dans cet état de l'art, qui se trouve au coeur de l'intelligence artificielle (IA) : la représentation structurée des connaissances et des préférences d'un agent et son application à la décision dans l'incertain.
C'est ce que nous allons faire dans le chapitre suivant, en nous limitant toutefois à une catégorie de représentation largement utilisée pour cela ces dernières années dans la communauté IA et basée sur les *préférences conditionnelles*.

Chapitre 4

Représentation logique de problèmes de décision dans l'incertain

4.1 Introduction

Dans les chapitres précédents nous avons considéré les problèmes de décision dans l'incertain sous l'angle *Théories de la décision*. Dans ce chapitre nous allons les considérer sous l'angle *Intelligence Artificielle* (IA) et en particulier sous l'angle de la logique.

Cela fait longtemps que l'IA, à travers le domaine de la planification, s'intéresse à la représentation logique des actions (voir les approches de type STRIPS (Fikes et Nilsson, 1971) [FN71] ou le calcul des situations (Mac Carthy, 1969, [MH69])).
Notre objectif est en quelque sorte complémentaire : nous souhaitons exprimer de manière "compacte" les connaissances plus ou moins certaines d'un agent et ses préférences graduelles. Ensuite, en utilisant un langage de représentation des actions"classique" (comme l'un des deux ci-dessus ou, plus simple encore, basé sur la logique classique), on se propose de trouver une "machinerie" logique permettant de déterminer les actions "optimales" vis-à-vis d'un critère de décision dans l'incertain donné.

Les *logiques pondérées* peuvent être utilisées pour exprimer non seulement des connaissances incertaines, mais aussi des préférences. Nous avons évoqué dans le chapitre 2 le cas de la logique possibiliste, et nous en reparlerons plus en détails dans la seconde partie de cette thèse. La *logique des pénalités* de Pinkas [Pin91] qui affecte aux formules, non pas des degrés de priorité, mais des coûts de violation peut également être interprétée en termes de préférences (Dupin de Saint-Cyr [Dup96]).
Dans le chapitre 3 page 131 de la partie II, nous proposerons un langage logique (valuée) ainsi que la machinerie associée permettant de calculer des décisions optimales vis-à-vis des critères de l'utilité qualitative possibiliste, ou des fonctions d'utilité basées sur les fonctions de croyance.
En attendant, dans ce chapitre nous allons décrire un certain nombre de langages logiques proposés récemment pour modéliser et résoudre des problèmes de décision dans l'incertain *qualitatifs*.
Ce chapitre sera centré autour d'approches logiques basées sur la notion de *préférences conditionnelles*. Les préférences conditionnelles sont des moyens compacts d'exprimer les préférences d'un agent. Dans le cas général, elles symbolisent des préférences du type : "dans le contexte ϕ, l'agent préfère que φ soit vraie" (où ϕ et φ sont des propositions). Dans ce chapitre nous présenterons deux interprétations différentes des préférences conditionnelles :

– Les *préférences ceteris paribus*.

Doyle et Wellman [DW94] ou Doyle, Shoham et Wellman [DSW91] interprètent la formule "dans le contexte ϕ, l'agent préfère que φ soit vraie" en imposant que l'on préfère les interprétations satisfaisant φ et ϕ à celles satisfaisant $\neg\varphi$ et ϕ, "toutes choses étant égales par ailleurs", c'est-à-dire lorsque les variables n'étant pas impliquées dans ϕ et φ ont la même valeur de vérité. Nous décrirons leur modèle, ainsi qu'un langage logique de représentation des préférences et des connaissances, proposé par Tan et Pearl [TP94], basé sur les préférences ceteris paribus.

– Les opérateurs de type *idéalité*.
  Boutilier [Bou94] interprète la même préférence conditionnelle en termes d'idéalité : Parmi les interprétations satisfaisant ϕ, celles qui satisfont également φ sont idéales, c'est-à-dire préférées à toutes les autres. A partir de cette interprétation, Boutilier établit un lien avec une logique bimodale, CO, ce qui lui permet de proposer un mode de représentation de préférences et de raisonnement à partir de celles-ci. Etendant la logique CO à la représentation et au raisonnement sur des connaissances incertaines, Boutilier définit la logique QDT, apte à la représentation et à la résolution de problèmes de décision sous incertitude.

  Lang [Lan96] utilise le même type d'opérateur d'idéalité que Boutilier, mais l'interprète différemment. Les opérateurs d'idéalité de Lang sont utilisés comme des contraintes sur une fonction d'utilité numérique sur les interprétations. Les fonctions d'utilité respectant ces contraintes sont "distinguées" et un préordre partiel sur les interprétations est défini (un monde étant préféré à un autre si et seulement si il est meilleur au sens de toutes les fonctions d'utilité distinguées). La théorie est étendue à la décision sous incertitude dans le cas simple de connaissance incomplète : une base de connaissance ne définit qu'un ensemble de mondes possibles.

Dans le prochain paragraphe, nous allons décrire l'approche proposée par Doyle et Wellman [DW94].

4.2 Préférences *ceteris paribus*

$L = \{p, q, r, \ldots\}$ représentera généralement un langage propositionnel, équipé des connecteurs habituels : $\vee, \wedge, \ldots$ L sera utilisé pour décrire des *faits*, et suivant les théories que nous décrirons, L sera étendu, afin de pouvoir décrire préférences et connaissances par défaut.

Avant de passer à la description des logiques pour la décision dans l'incertain, un premier problème doit être résolu, concernant les préférences entre formules : si φ et ψ sont des formules de L, que signifie "φ est préférée à ψ" (noté $\varphi \trianglerighteq \psi$) ?
Si Ω représente l'ensemble des interprétations de L ou *états du monde*, on peut très bien imaginer que les préférences de l'agent sont exprimées par un préordre $\succeq$ sur Ω, appelé *préférence faible*.
Mais comment définir une relation de préférence $\trianglerighteq$ sur les formules à partir d'une relation $\succeq$ sur les interprétations, ou au contraire, comment définir $\succeq$ à partir d'une relation de préférence $\trianglerighteq$ exprimée sur les formules ? C'est à ce type de question que nous allons apporter une réponse dans cette section.

4.2.1 Les préférences ceteris paribus : principe général

Les préférences ceteris paribus sont un mode de représentation d'un ordre de préférence entre les interprétations, à partir de préférences entre des formules.

Soit une proposition φ de L, $[\varphi]$ représente le sous-ensemble de l'ensemble d'interprétations $\Omega = \{\omega_1, \omega_2, \ldots\}$, correspondant à l'ensemble des interprétations "satisfaisant" φ.

Nous supposons donné un préordre complet $\succeq$ sur Ω (i.e. réflexif, transitif et complet). Une première définition de la *préférence étendue* aux propositions pourrait être la suivante :

Définition 4.2.1 *Préférence étendue simple.*
Soient $\varphi, \psi \in L$, φ est (faiblement) préférée à ψ ssi

$$\varphi \trianglerighteq \psi \Leftrightarrow \forall(\omega, \omega') \in [\varphi] \times [\psi], \omega \succeq \omega'$$

D'après cette définition, φ est faiblement préférée à ψ ssi toute interprétation satisfaisant φ est faiblement préférée à n'importe quelle interprétation satisfaisant ψ. On définit également *l'indépendance logique* de deux propositions :

Définition 4.2.2 *Indépendance logique. [DSW91]*

– *φ et ψ sont dites logiquement indépendantes ssi $[\varphi \wedge \psi]$, $[\varphi \wedge \bar{\psi}]$, $[\bar{\varphi} \wedge \psi]$, et $[\bar{\varphi} \wedge \bar{\psi}]$ sont non-vides (où $[\varphi \wedge \psi] = [\varphi] \cap [\psi]$, et $[\bar{\varphi}] = \Omega - [\varphi]$).*

– *φ et ψ sont dites semi-indépendantes ssi $[\varphi \wedge \bar{\psi}]$ et $[\bar{\varphi} \wedge \psi]$ sont non vides.*

Cette définition peut être étendue à plus de deux propositions, en considérant toutes les conjonctions booléennes.

Lorsque plusieurs propositions sont indépendantes, savoir que l'une d'entre elle est *vraie* (c'est-à-dire l'ensemble de ses modèles contient le monde réel) ou *fausse* ne garantit rien sur la valeur de vérité des autres. Si deux propositions sont semi-indépendantes, constater que l'une d'entre elles est vraie (ou l'imposer) ne permet pas de garantir que l'autre le soit également.

Le problème de la *préférence étendue simple* est qu'elle est très peu décisive. Supposons que l'on parte d'un ensemble de propositions semi-indépendantes $\varphi_1, \ldots, \varphi_n$ et de l'ensemble de préférences $(\varphi_i \succeq \overline{\varphi_i})_{i \in 1 \ldots n}$. Les φ_i peuvent être vues comme des contraintes, ou des propriétés que l'on souhaite satisfaire. La relation induite sur l'ensemble des mondes, Ω, ne distingue que trois classes d'équivalence : Les mondes ω satisfaisant toutes les contraintes φ_i, ceux qui en satisfont au moins une (mais pas toutes), et ceux qui n'en satisfont aucune.

A partir de ce constat d'échec, Doyle et Wellman [DW94] ont proposé de définir une relation plus faible de préférence entre les propositions, à partir de la comparaison de propositions *ceteris paribus*, i. e., toutes choses égales par ailleurs. Que signifie "toutes choses égales par ailleurs"? Pour le comprendre, il nous faut d'abord définir une relation *d'équivalence contextuelle* sur Ω.

Définition 4.2.3 *Equivalence contextuelle.*
Une fonction d'équivalence contextuelle η est une fonction qui associe à tout ensemble de propositions $\{\varphi_1, \ldots, \varphi_n\}$ une relation d'équivalence $\eta(\varphi_1, \ldots, \varphi_n)$ sur Ω.

Lorsque $\omega\eta(\varphi_1, \ldots, \varphi_n)\omega'$, on écrira $\omega \equiv \omega' mod_\eta \varphi_1, \ldots, \varphi_n$.

Exemple 4.2.1 *η telle que $\forall \varphi_1, \ldots, \varphi_n$, $\omega\eta(\varphi_1, \ldots, \varphi_n)\omega'$ ssi $\omega \in [\varphi_1] \cap \ldots \cap [\varphi_n] \Leftrightarrow \omega' \in [\varphi_1] \cap \ldots \cap [\varphi_n]$ est une fonction d'équivalence contextuelle.*
Les deux classes d'équivalence de $\eta(\varphi_1, \ldots, \varphi_n)$ sont respectivement l'ensemble des interprétations qui satisfont $\varphi_1 \wedge \ldots \wedge \varphi_n$ et l'ensemble des interprétations qui satisfont $\neg\varphi_1 \vee \ldots \vee \neg\varphi_n$.

On peut maintenant définir la relation de préférence *ceteris paribus* de deux propositions. Cette relation de préférence dépend bien sûr de la fonction d'équivalence contextuelle η choisie.

Définition 4.2.4 *Préférence ceteris paribus définie relativement à une fonction d'équivalence contextuelle η.*
φ est préférée (faiblement) ceteris paribus à ψ, suivant η, noté $\varphi \trianglerighteq \psi$ ssi

$$\forall(\omega, \omega') \in [\varphi] \cap [\overline{\psi}] \times [\overline{\varphi}] \cap [\psi], \omega \equiv \omega' mod_{\eta} \varphi \wedge \overline{\psi}, \overline{\varphi} \wedge \psi \Rightarrow \omega \succeq \omega'$$

La relation stricte $\triangleright$ et la relation d'équivalence $\bowtie$ associées sont définies classiquement à partir de $\trianglerighteq$: $\varphi \triangleright \psi$ ssi $\varphi \trianglerighteq \psi$ et non $\psi \trianglerighteq \varphi$; $\varphi \bowtie \psi$ ssi $\varphi \trianglerighteq \psi$ et $\psi \trianglerighteq \varphi$.

Enfin, cette définition peut être restreinte à un contexte donné, représenté par une formule ϕ. Dans ce cas, l'ensemble des mondes considérés pour effectuer la comparaison entre des formules est restreint au "contexte" défini par ϕ, c'est-à-dire à l'ensemble des interprétations satisfaisant ϕ.

Définition 4.2.5 *Préférence ceteris paribus contextuelle.*

$$\varphi \trianglerighteq \psi | \phi \textit{ ssi } \forall(\omega, \omega') \in [\varphi] \cap [\overline{\psi}] \cap [\phi] \times [\overline{\varphi}] \cap [\psi] \cap [\phi], \omega \equiv \omega' mod_{\eta} \varphi \wedge \overline{\psi}, \overline{\varphi} \wedge \psi \Rightarrow \omega \succeq \omega'$$

Dans la suite de ce chapitre, nous nous limiterons au cas de la logique propositionnelle. Dans ce cas, Doyle, Shoam et Wellman [DSW91] ont donné une définition de l'équivalence contextuelle propositionnelle. Soit $Q = \{p, q, r, \ldots\}$ l'ensemble des propositions atomiques du langage. Pour toute formule φ, $S(\varphi)$ représente l'ensemble des propositions atomiques dont *dépend* φ. φ dépend de p ssi toute proposition logiquement équivalente à φ contient p ou sa négation. Cette définition de S peut être étendue à un ensemble quelconque de formules, en notant que $S(\varphi, \psi) = S(\varphi) \cup S(\psi)$.
Une interprétation ω est simplement définie par l'ensemble des valeurs de vérité des propositions atomiques du langage. Deux mondes ω et ω' sont dits "en accord sur S" (où $S \subseteq Q$) ssi ils affectent les mêmes valeurs de vérité aux atomes de S, noté $\omega \equiv_S \omega'$.
Alors, l'équivalence contextuelle propositionnelle peut être définie :

Définition 4.2.6 *Equivalence contextuelle propositionnelle.*

$$\omega \eta(\varphi_1, \ldots, \varphi_n) \omega' \textit{ ssi } \omega \equiv_{Q - S(\varphi_1, \ldots, \varphi_n)} \omega'$$

Exemple 4.2.2 *Soit $Q = \{pluie, parapluie, \textit{café}\}$, $\varphi_1 = pluie \wedge parapluie$, $\varphi_2 = \neg pluie \vee \neg parapluie$, $\omega = \{pluie, parapluie, \textit{café}\}$, $\omega' = \{pluie, parapluie, \neg\textit{café}\}$.*
φ_1 et φ_2 sont les expressions des préférences d'un agent :

- φ_1 *l'agent souhaite à la fois qu'il pleuve et porter un parapluie,*
- φ_2 *l'agent souhaite porter un parapluie lorsqu'il pleut.*

Le fait de boire ou non un café lui est indifférent, puisque les préférences qu'il exprime ne contraignent pas la variable "café".
En d'autres termes, $\omega \eta(\varphi_1, \varphi_2) \omega'$.

Dans le paragraphe suivant, nous allons décrire l'approche logique de Tan et Pearl [TP94] pour la décision qualitative dans l'incertain, basée sur la notion de *désir conditionnel*, utilisant la relation de préférence ceteris paribus.

4.2.2 Principe de l'approche de Tan et Pearl basée sur les préférences ceteris paribus

L'approche de Tan et Pearl [TP94] pour la décision dans l'incertain est basée sur les notions de :

– *Défauts quantifiés* pour exprimer les connaissances incertaines. $\varphi \overset{\delta}{\Rightarrow} \psi$, où δ est un entier positif et est interprété comme : "si φ est vraie, alors je crois que ψ l'est également, avec force δ". Plus δ est grand, plus il est vraisemblable que ψ soit vraie lorsque φ l'est : le défaut $\varphi \overset{0}{\Rightarrow} \psi$ ne donne aucune information, alors que le défaut $\varphi \overset{+\infty}{\Rightarrow} \psi$ correspond à l'implication matérielle $\varphi \rightarrow \psi$: lorsque φ est vraie, ψ l'est forcément.

– *Désirs conditionnels* pour exprimer les préférences (graduelles). $D_\varepsilon(\alpha|\beta)$, où ε est un entier, est interprété comme : "α est désirée au degré ε si β est vraie". Les forces des désirs conditionnels sont interprétées de la même manière que celles des défauts : $\varepsilon = 0$ n'exprime aucune contrainte alors que $\varepsilon = +\infty$ correspond à une contrainte inviolable.

L'objectif de la méthode de Tan et Pearl est de déterminer à quel point et avec quel degré de confiance, une action σ_1 donne de "meilleurs" résultats qu'une action σ_2, dans un contexte donné, décrit par une formule ϕ.

Exemple 4.2.3 *Le parapluie.*
Un agent se demande s'il doit prendre son parapluie... La connaissance incertaine est exprimée par le défaut : $nuage \overset{1}{\Rightarrow} pluie$ ("généralement, si le ciel est nuageux, il va pleuvoir"). On sait de plus que l'agent préfère qu'il ne pleuve pas, ne pas porter de parapluie sauf si il pleut, auquel cas il préfère en porter un : $\mathcal{D} = \{D_1(\neg pluie), D_1(\neg para), D_2(para|pluie)\}$. Dans cet exemple le désir de porter un parapluie lorsqu'il pleut est plus fort que les autres désirs ($\varepsilon = 2$).

Défauts et ordre de "normalité" entre les mondes

De l'ensemble $\{\varphi_i \overset{\delta_i}{\Rightarrow} \psi_i\}$ des défauts représentant la connaissance de l'agent, Tan et Pearl, en utilisant la méthode proposée par Goldszmidt et Pearl [GP91] déduisent un préordre complet sur Ω, exprimé par une fonction κ sur Ω, à valeurs entières.
$\kappa(\omega) = 0$ signifie que le monde ω est normal, i. e. entièrement plausible, et plus $\kappa(\omega)$ est grand, plus ω est surprenant ($\kappa(\omega) = +\infty$ ssi ω est impossible). κ peut être interprété comme un *degré de surprise*. Un tel ordre est défini par :

$$\kappa(\omega) = max_{i,\omega \models \varphi_i \wedge \neg\psi_i} \delta_i$$

Si ω viole un défaut de degré δ_i, ω est surprenant au moins au degré δ_i, et si ω ne viole aucun défaut, il est entièrement plausible ($\kappa(\omega) = 0$). Tan et Pearl [TP94] notent $\kappa(\phi, \sigma_i)$ la fonction κ induite sur les mondes après que l'action σ_i ait été accomplie, dans le contexte ϕ. $\kappa^\delta(\phi, \sigma_i)$ représente l'ensemble des mondes de degré de surprise (fonction κ) inférieur ou égal à δ.

Exemple 4.2.4 *Le parapluie, suite.*
L'agent observe que le ciel est nuageux ($\phi = nuage$). L'ordre induit sur les mondes satisfaisant ϕ par le défaut $nuage \overset{1}{\Rightarrow} pluie$ peut être décrit par la fonction κ suivante :

ω	$\kappa(\omega)$
$\{pluie, nuage, para\}$	*0*
$\{pluie, nuage, \neg para\}$	*0*
$\{\neg pluie, nuage, para\}$	*1*
$\{\neg pluie, nuage, \neg para\}$	*1*

Désirs conditionnels, préférences ceteris paribus et relation d'ordre entre les mondes

Tan et Pearl [TP94] proposent d'interpréter un désir conditionnel $D(\alpha|\beta)$ (exprimant une préférence pour que la formule α soit vraie lorsque la formule β l'est) comme une préférence ceteris paribus $(\alpha \trianglerighteq \neg\alpha)|\beta$. Ils notent $D(\alpha)$ le désir conditionnel $D(\alpha|\top)$. Dans un certain sens, $D(\alpha)$ peut être vu comme un désir "inconditionnel" : on préfère que α soit vraie quel que soit le contexte et pas seulement lorsqu'une formule β est vérifiée. Notons que le terme "inconditionnel" donne une impression exagérée de force du désir. Au contraire, dans l'approche de Tan et Pearl, les désirs inconditionnels seront les premiers que l'agent accepte de voir violés en cas de conflit entre plusieurs désirs ($D(\neg\alpha|\beta)$ prendra toujours le pas sur $D(\alpha)$).

Les désirs conditionnels de Tan et Pearl sur les formules induisent des relations de *préférence* sur Ω. Ces désirs conditionnels peuvent être vus comme un ensemble de contraintes sur la relation de préférence $\succeq$ de l'agent sur Ω.
Dans l'exemple du parapluie, si on considère les désirs de l'agent $D(\neg pluie)$ et $D(\neg para)$, on obtient les contraintes :
- $D(\neg pluie) \Rightarrow \{\neg pluie, \neg para\} \succ \{pluie, \neg para\}$ et $\{\neg pluie, para\} \succ\succ \{pluie, para\}$
- $D(\neg para) \Rightarrow \{\neg pluie, \neg para\} \succ \{\neg pluie, para\}$ et $\{pluie, \neg para\} \succ \{pluie, para\}$.

Ces deux contraintes induisent un ordre partiel sur les interprétations défini par (les interprétations $\{\neg pluie, para\}$ et $\{pluie, \neg para\}$ sont incomparables) :
- $\{\neg pluie, \neg para\} \succ \{\neg pluie, para\} \succ \{pluie, para\}$
- $\{\neg pluie, \neg para\} \succ \{pluie, \neg para\} \succ \{pluie, para\}$.

Si on ajoute la contrainte $D(para|pluie)$, on obtient $\{pluie, para\} \succ \{pluie, \neg para\}$, qui est en conflit avec $D(\neg para)$. Pour résoudre ce type de conflit, Tan et Pearl [TP94] proposent qu'en cas de conflit le désir conditionnel *le plus spécifique* prenne le pas sur l'autre, et que les contraintes contradictoires du désir le moins spécifique soient éliminées.
La relation de spécificité entre désirs est encore basée sur la notion de contexte :

Définition 4.2.7 *Contexte d'un désir conditionnel.*
On définit tout d'abord $C(\alpha,\omega)$ par :

$$C(\alpha,\omega) = \{\omega', \omega' \sim_{Q-S(\alpha)} \omega\}.$$

Le contexte d'un désir conditionnel $D(\alpha|\beta)$, noté $C(\alpha,\beta)$ est défini par :

$$C(\alpha,\beta) = \bigcup_{\omega \models \beta} C(\alpha,\omega)$$

Le contexte d'un désir conditionnel représente l'ensemble des mondes "contraints" par le désir : $\omega' \in C(\alpha,\beta)$ ssi $\exists\omega \models \beta, \omega' \sim_{Q-S(\alpha)} \omega$. En particulier, il est aisé de montrer que le contexte d'un désir absolu $D(\alpha)$ est l'ensemble Ω lui même ($C(\alpha) = \bigcup_{\omega\in\Omega} C(\alpha,\omega) = \Omega$). On peut également montrer que lorsque $S(\alpha) \cap S(\beta) = \emptyset$, $C(\alpha,\beta) = \{\omega, \omega \models \beta\}$. Le contexte du désir conditionnel n'est dans ce cas que l'ensemble des interprétations satisfaisant "l'antécédent" du désir.

Définition 4.2.8 *Spécificité d'un désir conditionnel.*
$D(\alpha|\beta)$ est dit plus spécifique que $D(\alpha'|\beta')$ ssi $C(\alpha,\beta) \subseteq C(\alpha',\beta')$.

Exemple 4.2.5 *Le parapluie (suite).*
Soient les deux désirs conditionnels conflictuels $D(\neg para)$ et $D(para|pluie)$.
$C(para) = \Omega$*, et* $C(para, pluie) = \{\{pluie, para\}, \{pluie, \neg para\}\}$.
$C(para, pluie) \subset C(para)$*, donc* $D(para|pluie)$ *est plus spécifique.*

Une fois que l'on a identifié tous les conflits entre désirs, et éliminé toutes les préférences conflictuelles issues du désir le moins spécifique, il nous reste un ensemble cohérent de contraintes entre mondes. Cela signifie que l'on peut ordonner les mondes, i.e. trouver une fonction π à valeurs entières, telle que $\forall(\omega, \omega'), \omega \succeq \omega' \Leftrightarrow \pi(\omega) \geq \pi(\omega')$. Cette fonction n'est bien entendu pas unique : si π est une telle fonction, les fonctions $k \cdot \pi$ et $k + \pi$, avec k entier naturel positif, reflètent le même ordre.
Les désirs conditionnels induisent des contraintes sur les interprétations. Ces contraintes sont appelées *contrainte Ceteris Paribus (CP)* et sont définies par :

Définition 4.2.9 *Contrainte CP.*
$\omega \succ_\varepsilon \omega'$ *est une contrainte CP d'un désir conditionnel* $D(\alpha|\beta)$ *ssi* $\exists \nu \models \beta, (\omega, \omega') \in C(\alpha, \nu)^2$,
$\omega \models \alpha, \omega' \models \neg\alpha$.

Encore une fois, lorsque deux contraintes CP issues de deux désirs conflictuels sont incompatibles, on supprime celle qui provient du désir le moins spécifique.

Nous n'avons pas encore utilisé la notion de *force* d'un désir conditionnel. Celle-ci intervient dans la définition des rangements admissibles des interprétations, en accord avec les désirs conditionnels : Lorsqu'il ne reste plus que des contraintes CP compatibles, on peut définir la notion de *rangement admissible* :

Définition 4.2.10 *Rangement admissible.*
Une fonction π *sur* Ω *à valeurs entières est appelée rangement admissible pour un ensemble de désirs ssi, pour toute contrainte CP on a* $\omega \succ_\varepsilon \omega', \pi(\omega) \geq \pi(\omega') + \varepsilon$.

Parmi tous les rangements possibles, Tan et Pearl proposent de ne garder que les plus compacts, i. e. minimisant la somme $\sum_{\omega,\omega' \in \Omega} |\pi(\omega) - \pi(\omega')|$. De tels rangements sont notés π^+.
Cette approche consiste à privilégier une *indifférence maximale* entre les mondes : si deux mondes ne peuvent être différenciés par une contrainte CP, alors ils doivent avoir la même évaluation.

Exemple 4.2.6 *Le parapluie (suite).*
Dans l'exemple du parapluie, $\mathcal{D} = \{D_1(\neg pluie), D_1(para|pluie), D_1(\neg para)\}$.
On note $\{pluie, para\} = \omega_1$, $\{\neg pluie, para\} = \omega_2$, $\{pluie, \neg para\} = \omega_3$ *and* $\{\neg pluie, \neg para\} = \omega_4$.
$D_1(\neg pluie)$ *nous donne les contraintes* $\pi(\omega_2) \geq \pi(\omega_1) + 1$ *et* $\pi(\omega_4) \geq \pi(\omega_3) + 1$. $D_1(para|pluie)$ *nous donne la contrainte* $\pi(\omega_1) \geq \pi(\omega_3) + 1$*, et* $D_1(\neg para)$ *nous donne les contraintes* $\pi(\omega_4) \geq \pi(\omega_2) + 1$*, et* $\pi(\omega_3) \geq \pi(\omega_1) + 1$. *Cette dernière contrainte est incompatible avec celle provenant de* $D_1(para|pluie)$ *qui est plus spécifique. Elle est donc éliminée. Les rangements les plus compacts sont décrits dans le tableau suivant :*

ω	$\pi^+(\omega)$
$\{\neg pluie, \neg para\}$	*m+3*
$\{\neg pluie, para\}$	*m+2*
$\{pluie, para\}$	*m+1*
$\{pluie, \neg para\}$	*m*

Application à l'évaluation des actions

Nous avons construit les deux préordres sur les mondes (incertitude et préférences) résultant d'une action σ dans un contexte ϕ.
Comment comparer deux actions à partir de ces préordres? Tan et Pearl suggèrent que les deux échelles ne sont pas commensurables. Pour comparer deux actions σ_1 et σ_2, ils proposent de comparer les ensembles $\kappa^i(\phi, \sigma_1)$ et $\kappa^i(\phi, \sigma_2)$ pour tout i. Ces ensembles représentent respectivement les conséquences possibles de σ_1 et σ_2 au degré de confiance i.

Dans un premier temps, comparons $\kappa^0(\phi, \sigma_1)$ et $\kappa^0(\phi, \sigma_2)$ (les conséquences “normales”). Tout d'abord, on définit une relation de préférence $\succ_\pi^\varepsilon$, où π est un rangement sur 2^Ω. Pour tout $A \subseteq \Omega$, $\pi^*(A)$ et $\pi_*(A)$ représentent respectivement les degrés de préférence de la meilleure et de la pire conséquence (éléments de A).

Définition 4.2.11 *$W \succ_\pi^\varepsilon V$ ssi*

1. *$\varepsilon = 0$ si $V = W$,*
2. *$\pi_*(W) \geq \pi^*(V - W) + \varepsilon$ si $W \subset V$,*
3. *$\pi_*(W - V) \geq \pi^*(V) + \varepsilon$ si $V \subset W$,*
4. *$\pi_*(W - V) \geq \pi^*(V - W) + \varepsilon$ sinon.*

Si l'un des ensembles est inclus dans l'autre, on compare les conséquences qu'il contient à celles de son complémentaire dans l'autre ensemble. Sinon, on compare les deux ensembles d'après les éléments qu'ils n'ont pas en commun.

Alors, Tan et Pearl définissent une relation de préférence (quantifiée) entre actions par :

Définition 4.2.12 *Préférence quantifiée entre actions.*
Soit $\mathcal{D}$ un ensemble de désirs conditionnels, et σ_1 et σ_2, deux actions. σ_1 est préférée à σ_2 sachant ϕ avec force ε et confiance δ, noté $\phi \vdash_\delta (\sigma_1 \succ_\varepsilon \sigma_2)$ ssi $\kappa^i(\phi, \sigma_1) \succ_{\pi^+}^\varepsilon \kappa^i(\phi, \sigma_2)$ pour tout rangement π^+ maximalement compact déduit de $\mathcal{D}$ et pour tout $i \in 0, \ldots \delta$.

En résumé, cela signifie que les conséquences les plus plausibles de σ_1 sont préférées aux conséquences les plus plausibles de σ_2, et ce avec au moins avec force ε.

Exemple 4.2.7 *Le parapluie, suite et fin.*
$\kappa^0(nuage, para) = \{\{para, nuage, pluie\}\}$ de rang minimal $m + 1$,
$\kappa^0(nuage, \neg para) = \{\{\neg para, nuage, pluie\}\}$ de rang maximal m.
Ainsi, $\kappa^0(nuage, para) \succ_{\pi^+}^1 \kappa^0(nuage, \neg para)$, i. e. $nuage \vdash_0 (para \succ^1 \neg para)$.
Par contre, $\kappa^1(nuage, para) = \{\{para, nuage, \neg pluie\}\}$ de rang $m + 2$, et $\kappa^1(nuage, \neg para) = \{\{\neg para, nuage, \neg pluie\}\}$ de rang $m+3$. L'ordre de préférence entre les deux actions est renversé : $nuage \not\vdash_1 (para \succ^1 \neg para)$.

Conclusion

Le modèle logique de Tan et Pearl pour la décision qualitative dans l'incertain fait intervenir la notion de préférence *ceteris paribus*. $D(\alpha|\beta)$ signifie que l'on préfère que α soit vraie, plutôt que $\neg\alpha$, dans *tous* les états du monde satisfaisant β. Mais des désirs *plus spécifiques* que $D(\alpha|\beta)$ peuvent prendre le pas sur celui-ci. Par la suite, les mondes sont regroupés en classes d'équivalence suivant leurs degrés de préférence, de la manière la plus compacte possible (deux mondes sont équivalents dès lors qu'aucun désir conditionnel n'oblige à préférer l'un à l'autre) : la méthode de Tan et Pearl privilégie donc une *indifférence maximale* entre les mondes.

Dans le paragraphe suivant, nous allons décrire l'approche de Boutilier [Bou94] qui fait également appel à un connecteur du type $D(\cdot|\cdot)$, mais qui est interprété différement. Ce connecteur, noté $I(\cdot|\cdot)$ est un opérateur exprimant une idée de but conditionnel *idéal*. $I(\beta|\alpha)$ signifie : "Idéalement, β doit être vérifiée si α l'est".
La différence principale avec l'approche de Tan et Pearl est qu'alors que $D(\beta|\alpha)$ contraint *tous* les mondes satisfaisant α, $I(\beta|\alpha)$ ne contraint que les mondes *préférés* (au sens d'une relation que nous allons définir) satisfaisant α.

4.3 Préférences exprimées par des opérateurs d'idéalité

4.3.1 La logique QDT de Boutilier

Dans ce paragraphe, nous allons d'abord décrire la logique modale CO proposée par Boutilier, puis son utilisation pour la définition du connecteur $I(\cdot|\cdot)$. Ensuite nous décrirons la logique QDT (pour "Qualitative Decision Theory) qui étend CO pour permettre de décrire à la fois préférences et "défauts" (ou connaissances incertaines). Enfin, nous décrirons la méthode de représentation des décisions dans l'approche de Boutilier [Bou94].

La logique modale CO

La logique CO est une logique bimodale, comportant les deux opérateurs modaux $\Box$ et $\overleftarrow{\Box}$. La sémantique de la logique CO est basée sur un triplet $M =< \Omega, \geq, \nu >$, où Ω est un ensemble de mondes possibles, ν est une fonction de valuation, et $\geq$ est une relation *d'accessibilité*, qui est un préordre sur Ω. $\geq$ sera utilisé dans ce paragraphe pour ordonner les mondes suivant leurs degrés de préférence. $\omega \geq \omega'$ signifie que le monde ω est préféré ou équivalent au monde ω' [1].

Les opérateurs modaux de la logique CO sont définis comme suit à partir de la relation d'accessibilité :

- $M \models_\omega \Box\alpha$ ssi $\forall \omega' \geq \omega, M \models_{\omega'} \alpha$,
- $M \models_\omega \overleftarrow{\Box}\alpha$ ssi $\forall \omega' < \omega, M \models_{\omega'} \alpha$.

$\Box\alpha$ est vraie dans le monde ω ssi α est vraie dans tous les mondes préférés ou équivalents à ω. $\overleftarrow{\Box}\alpha$ est vraie dans le monde ω ssi α est vraie dans tous les mondes strictement moins préférés au monde ω.

1. Boutilier adopte la convention inverse : ω est préféré (ou équivalent) à ω' ssi $\omega \leq \omega'$. Cette convention est habituellement utilisée dans le domaine des logiques modales, dans lequel on s'intéresse le plus souvent à des mondes *minimaux* pour une certaine propriété. A l'inverse, dans les théories de la décision on cherche plutôt à maximiser (une fonction d'utilité, par exemple), c'est pourquoi nous n'utiliserons pas ici la convention de Boutilier.

On définit classiquement les opérateurs duaux $\Diamond$ et $\overleftarrow{\Diamond}$ par :

$$\Diamond \equiv_{def} \neg\Box\neg \text{ et } \overleftarrow{\Diamond} \equiv_{def} \neg\overleftarrow{\Box}\neg$$

$\Diamond\alpha$ signifie qu'il existe un monde préféré ou équivalent au monde courant dans lequel α est vraie, et $\overleftarrow{\Diamond}\alpha$ qu'il existe un monde strictement moins préféré dans lequel α est vraie.

Enfin, Boutilier définit deux derniers opérateurs modaux : $\overleftrightarrow{\Box}$ et $\overleftrightarrow{\Diamond}$.

$$\overleftrightarrow{\Box}\,\alpha \equiv_{def} \Box\alpha \wedge \overleftarrow{\Box}\alpha \text{ et } \overleftrightarrow{\Diamond}\,\alpha \equiv_{def} \Diamond\alpha \vee \overleftarrow{\Diamond}\alpha$$

$\overleftrightarrow{\Box}\,\alpha$ (resp. $\overleftrightarrow{\Diamond}\,\alpha$) signifie que α est vraie dans *tous* les mondes (resp. dans *au moins un* monde).

Le connecteur $I(\cdot|\cdot)$ et la logique CO

Comme nous l'avons déjà précisé, $I(\beta|\alpha)$ signifie " dans les mondes préférés satisfaisant α, β est également satisfaite". Cette proposition se traduit en logique CO par la définition suivante :

Définition 4.3.1 *Opérateur modal d'idéalité.*

$$I(\beta|\alpha) \equiv_{def} \overleftrightarrow{\Box}\,\neg\alpha \vee \overleftrightarrow{\Diamond}\,(\alpha \wedge \Box(\alpha \to \beta)).$$

Ce que l'on pourrait traduire par : "Soit α n'est jamais vraie, soit tous les mondes préférés satisfaisant α satisfont également β. En particulier, le désir *inconditionnel* $I(\beta|\top)$ (abrégé $I(\beta)$) s'ecrit aussi $\overleftrightarrow{\Diamond}\,\Box\beta$ et signifie que β est vraie dans tous les mondes préférés.
La préférence étendue simple de Doyle et Wellman (voir définition 4.2.1) peut être représentée, dans un cas particulier, dans la logique CO : $\alpha \succeq \neg\alpha$ ssi $\overleftrightarrow{\Box}\,(\alpha \to \Box\alpha)$, qui signifie que α est préférée à $\neg\alpha$ ssi α est vraie dans tous les mondes préférés (la formule modale précédente se traduit sémantiquement par $M \models_\omega \alpha \Rightarrow \forall\omega' \geq \omega, M \models_{\omega'} \alpha$, i.e. si α admet un modèle, tous les modèles qui la satisfont sont préférés aux modèles qui ne la satisfont pas).

La logique CO ne permet pas seulement d'exprimer de manière compacte des préférences, elle permet également de raisonner avec ces préférences, i.e. de dériver de nouvelles préférences avec les préférences exprimées par un agent. La propriété suivante, notamment, permet la dérivation de préférences nouvelles :

Proposition 4.3.1 *Détachement préférentiel.*

$$I(\beta|\alpha) \wedge I(\alpha) \to I(\beta)$$

Par contre, la propriété $I(\beta|\alpha) \wedge \alpha \to I(\beta)$ n'est pas valide dans le cas général!

Nous avons défini le connecteur $I(\cdot|\cdot)$ à partir d'une relation de préférence $\geq$ entre mondes. Il est clair cependant que dans les applications de la logique CO la relation $\geq$ qui est la sémantique associée à l'expression des préférences reste implicite et n'est pas donnée a priori. Au contraire, les préférences d'un agent sont exprimées par l'intermédiaire du connecteur $I(\cdot|\cdot)$, contraignant la relation $\geq$. En fait, celle-ci n'est en général pas entièrement définie, mais en tout cas l'ensemble des mondes préférés par la relation est, lui, défini. Nous verrons un peu plus loin que la méthode de Boutilier pour choisir des décisions optimales, solutions de problèmes de décision dans l'incertain, ne se sert que de ces mondes préférés.

Exemple 4.3.1 *Expression des préférences dans le problème du parapluie (simplifié).*
Pour l'instant nous considérons le problème du parapluie avec les seuls symboles $pluie$ et $para$. L'agent exprime les préférences $I(\neg para)$, $I(\neg pluie)$ et $I(para|pluie)$. Ces désirs sont incohérents au sens de Tan et Pearl (voir la section précédente) qui vont alors éliminer le désir le moins spécifique. Dans l'approche de Boutilier nous allons voir que ces désirs ne sont pas incompatibles, car la notion de préférence qu'il utilise est plus faible que la préférence ceteris paribus.
$I(\neg para) \equiv \overleftrightarrow{\Diamond} \Box \neg para$, i.e. $\exists \omega, \forall \omega' \geq \omega M \models_{\omega'} \neg para$. Cela implique que l'un des deux mondes $\{pluie, \neg para\}, \{\neg pluie, \neg para\}$ est un monde idéal (non dominé par un autre monde, au sens de $\geq$). De même, $I(\neg pluie)$ implique que l'un des deux mondes $\{\neg pluie, para\}, \{\neg pluie, \neg para\}$ est idéal. Enfin, $I(para|pluie) \equiv \overleftrightarrow{\Box} \neg pluie \vee \overleftrightarrow{\Diamond} (pluie \wedge \Box(pluie \rightarrow para))$. Comme il existe au moins un monde violant $\neg para$, cela signifie qu'il existe un monde ω, satisfaisant $pluie$, tel que tous les mondes préférés ou équivalents à ω satisfont $pluie \rightarrow para$.
Finalement, $I(para|pluie)$ est vérifié ssi $\{pluie, para\} \geq \{pluie, \neg para\}$. Ceci entraine que $\{pluie, \neg para\}$ ne peut être un monde idéal, soit que le seul monde idéal est $\{\neg pluie, \neg para\}$.

Logique CO et connaissances incertaines : la logique QDT

L'extension de la logique des préférences CO à la logique QDT pour la décision dans l'incertain est directe. Boutilier utilise le connecteur $I(\cdot|\cdot)$ pour exprimer les préférences d'un agent, il propose d'utiliser de la même manière un connecteur $\Rightarrow$ pour exprimer les connaissances incertaines, ou défauts.

Le modèle sémantique de la logique QDT est un quadruplet $M =< \Omega, \geq_P, \geq_N, \nu >$. Le langage de QDT est étendu par rapport à celui de CO, et compte quatre opérateurs modaux : $\Box_N$, $\overleftarrow{\Box}_N$, $\Box_P$, $\overleftarrow{\Box}_P$ (N concerne la "normalité", i.e. les connaissances par défaut, et P les prf́érences). Le connecteur $I(\cdot|\cdot)$ est défini comme précédement, à partir de $\Box_P$, et le connecteur de normalité $\Rightarrow$ est défini comme suit :

$$\alpha \Rightarrow \beta \equiv_{def} \overleftrightarrow{\Box}_N \neg\alpha \vee \overleftrightarrow{\Diamond}_N (\alpha \wedge \Box_N(\alpha \rightarrow \beta)).$$

Une règle par défaut a donc une interprétation similaire à un désir conditionnel d'idéalité.

En plus des défauts et des préférences, on ajoute au modèle QDT une base de faits $\mathcal{F}$ (cohérente), exprimant des connaissances certaines. Les connaissances (défauts) servent à étendre la base de faits, et ainsi à restreindre l'ensemble des mondes possibles à un ensemble de mondes *les plus plausibles*. L'extension de la base $\mathcal{F}$, ou clôture de $\mathcal{F}$ est notée $Cl(\mathcal{F})$, et définie par :

Définition 4.3.2 *Clôture d'une base de connaissances.*

$$Cl(\mathcal{F}) = \{\alpha, \mathcal{F} \Rightarrow \alpha\}$$

La clôture de $\mathcal{F}$ représente donc l'ensemble des faits normaux (les plus plausibles), déductibles de $\mathcal{F}$.

Boutilier définit également la notion de *but idéal* :

Définition 4.3.3 *But idéal.*
α est un but idéal ssi $M \models I(\alpha|Cl(\mathcal{F}))$.

Un but idéal est, en quelque sorte, une proposition qui est vraie si le monde réel est un monde préféré dans le cas "normal".

Exemple 4.3.2 *Dans l'exemple du parapluie, on rajoute le symbole $nuage$, et la connaissance par défaut $nuage \Rightarrow pluie$. Si $\mathcal{F} = \{nuage\}$, $Cl(\mathcal{F}) = nuage \wedge pluie$. Si on considère les préférences exposées dans le paragraphe précédent, on trouve que $para$ est un but idéal.*
On peut noter l'aspect non-monotone du calcul des buts idéaux : si on ajoute $\neg pluie$ à la base de faits, on trouve maintenant que $\neg para$ est un but idéal.

Ensuite, Boutilier définit la notion de condition suffisante d'idéalité pour un but (CSI) :

Définition 4.3.4 *Condition suffisante d'idéalité pour un but.*
γ est une CSI pour β ssi $\gamma \wedge \beta$ est cohérente, et $M \models \overset{\leftrightarrow}{\Box}_P \ (\beta \rightarrow \overset{\leftarrow}{\Box}_P(\beta \rightarrow \neg\gamma))$.

Cette expression, en apparence complexe, signifie tout simplement que les mondes préférés (au sens de $\geq_P$) parmi ceux satisfaisant β satisfont également $\beta \wedge \gamma$.

Alors, la proposition suivante peut être démontrée :

Proposition 4.3.2 *Si γ est une CSI pour $Cl(\mathcal{F})$, alors pour tout but idéal α, $M \models \gamma \wedge Cl(\mathcal{F}) \rightarrow \alpha$.*

On entrevoit déjà le rôle des CSI dans l'approche de Boutilier de la décision qualitative dans l'incertain : si l'agent à le pouvoir de décider de la valeur de vérité d'une CSI, alors, si le monde réel est normal, il est certain de se retrouver dans la meilleure situation possible.
Nous allons revenir en détail sur les CSI, mais d'abord, donnons la définition d'une action dans l'approche de Boutilier.

Les actions dans la logique QDT

La notion d'action est relativement simple dans l'approche de Boutilier. L'ensemble des symboles propositionnels $\mathcal{Q}$ est séparé en deux : les variables contrôlables (i.e. dont le décideur peut décider de la valeur de vérité) $\mathcal{C}$, et les autres $\overline{\mathcal{C}}$. Si u et v sont respectivements des affectations complètes des valeurs de vérité des variables de $\mathcal{C}$ et $\overline{\mathcal{C}}$, $u; v$ représente l'affectation complète des valeurs de vérité des variables de $\mathcal{Q}$, compatible avec les deux précédentes. $V(\mathcal{P})$ représente l'ensemble des affectations de valeurs de vérité des variables de $\mathcal{P}$.
On peut définir les notions de *contrôlabilité* et *d'influençabilité* d'une proposition α quelconque :

Définition 4.3.5 *Contrôlabilité, influençabilité.*
α est contrôlable ssi $\forall u \in \overline{\mathcal{C}}, \exists (v, w) \in \mathcal{C}, v; u \models \alpha, w; u \models \neg\alpha$.
α est influençable ssi $\exists u \in \overline{\mathcal{C}}, \exists (v, w) \in \mathcal{C}, v; u \models \alpha, w; u \models \neg\alpha$.

Nous allons pouvoir, maintenant, grâce à cette notion de variable contrôlable, définir les notions d'action, et d'action optimale.

Dans la suite, on notera $NI(\mathcal{F})$ l'ensemble des croyances non influencables, ($NI(\mathcal{F}) = \{\alpha \in Cl(\mathcal{F}), \alpha$ n'est pas influencable). $\mathcal{A}$ représentera l'ensemble des actions disponibles pour l'agent. Nous considérerons que ces actions sont représentées par l'ensemble des affectations de valeurs de vérités aux variables de $\mathcal{C}$ ($\mathcal{A} = V(\mathcal{C})$).

Dans un premier temps, on va se limiter au cas où la connaissance de l'agent est complète : il connaît les valeurs de vérité de toutes les propositions élémentaires incontrôlables. Alors, on peut définir un but en environnement entièrement connu (EEC-But), comme étant une proposition *contrôlable* qui permet de se retrouver dans une situation idéale, connaissant la valeur des variables non contrôlables :

Définition 4.3.6 *But en environnement entièrement connu.*
α est un EEC-But ssi $M \models I(\alpha | NI(\mathcal{F}))$ et α est contrôlable.

Maintenant, Boutilier définit une action *satisfaisante* en environnement connu, comme une action permettant d'obtenir tous les EEC-Buts, sachant $NI(\mathcal{F})$:

Définition 4.3.7 *$a \in \mathcal{A}$ est une action satisfaisante ssi :*

$$\forall EEC - But\ \alpha, M \models NI(\mathcal{F}) \wedge a \rightarrow \alpha$$

Boutilier prouve également que les actions satisfaisantes sont exactement les CSI pour $NI(\mathcal{F})$.

Dans le cas où la connaissance est incomplète, Boutilier définit une relation de préférence entre actions (instanciations complètes des variables de $\mathcal{C}$), $\succeq$, par :

Définition 4.3.8 *Préférence entre actions.*

$$a_1 \succeq a_2 \textit{ ssi } M \models \overleftrightarrow{\Diamond}_P\ (a_2 \wedge NI(\mathcal{F}) \wedge \neg \overleftarrow{\Diamond}_P (a_1 \wedge NI(\mathcal{F})))$$

En d'autres termes : "Il existe un monde normal satisfaisant a_2 tel qu'aucun monde moins préféré ne satisfait a_1". Ou encore, que la pire conséquence possible de a_1 est au moins aussi bonne que la pire conséquence possible de a_2.

En termes de sémantique, ceci peut se traduire par : $\exists \omega \in PW \models a_2, \forall \omega' \models a_1, \omega' \geq_P \omega$ où PW représente l'ensemble des mondes normaux ("possible worlds"). Soit encore,

$$\min_{\omega \in PW, \omega \models a_1} \omega \geq_P \min_{\omega' \in PW, \omega' \models a_2} \omega'$$

où min est associé à la relation $\geq_P$.
Ce type de sémantique de la représentation logique proposée par Boutilier correspond tout à fait à l'approche *maximin* de Brafman et Tennenholtz exposée dans le chapitre 1.

Résumé de l'approche de Boutilier

L'approche de Boutilier pour la décision dans l'incertain donne la priorité à l'expression des défauts sur l'expression des préférences. Dans un premier temps, on détermine l'ensemble des faits "normaux" déductibles d'une base de faits et de connaissances par défaut. Ensuite, on détermine les actions (instanciations des variables contrôlables) optimales, c'est-à-dire telles que leurs pires conséquences éventuelles sont les meilleures possibles.
Boutilier [Bou94] note que sa méthode possède tout de même une faiblesse du fait de son approche en deux temps : aucun compromis n'est possible entre préférences et connaissances incertaines. Supposons qu'un agent doive traverser une autoroute à pied, pour aller chercher un café dans une station-service en face. S'il pense avoir de bonnes chances de s'en sortir vivant, la méthode de Boutilier va lui conseiller d'aller chercher son café, en négligeant totalement la possibilité d'accident...

Dans la seconde partie de cette thèse, nous rediscuterons de ce problème de commensurabilité des échelles de préférence et de certitude.

Avant cela, nous allons décrire une dernière approche logique de la décision qualitative dans l'incertain basée sur les désirs conditionnels, proposée par Lang [Lan96].

4.3.2 Idéalité et utilité numérique

Lang [Lan96] a proposé une méthode pour construire un préordre partiel sur les mondes, basée sur la représentation des désirs conditionnels de Boutilier.
Alors que la méthode de Boutilier [Bou94] ne calcule que l'ensemble des mondes "idéaux", i.e. préférés au sens d'une relation induite par un ensemble de désirs conditionnels, la méthode proposée par Lang [Lan96] permet de construire une relation de préférence (un préordre partiel) entre les mondes.

Il existe une autre différence entre les deux approches. Cette différence se situe au niveau de la représentation des connaissances de l'agent. Les deux approches utilisent une base de faits $\mathcal{F}$, mais alors que Boutilier détermine l'ensemble des mondes "plausibles" en utilisant des connaissances par défaut et en "clôturant" la base de faits, Lang utilise une base de connaissances "génériques", BC cohérente (également cohérente avec $\mathcal{F}$: un fait ne peut contredire la base de connaissances génériques). L'approche de Lang pour la représentation des connaissances est *monotone* (apprendre un nouveau fait ne peut que restreindre l'ensemble des mondes possibles), alors que celle de Boutilier est *non monotone*.

Désirs conditionnels et fonction d'utilité

Dans (Lang [Lan96]), les désirs conditionnels sont notés $D(\beta|\alpha)$, comme dans (Tan et Pearl [TP94]), mais sont interprétés comme dans (Boutilier [Bou94]) : une relation de préférence $\geq$ sur l'ensemble Ω des mondes est compatible avec $D(\beta|\alpha)$ ssi les mondes préférés au sens de $\geq$, satisfaisant α, satisfont également β.
Chez Lang, la relation de préférence entre les mondes est représentée par une *fonction d'utilité* u, à valeurs réelles. La relation de préférence déduite de u est notée $\geq_u$: $\omega \geq_u \omega'$ ssi $u(\omega) \geq u(\omega')$. Si $S \subseteq \Omega$ est un ensemble de mondes, $Max_{\geq_u} S$ représente l'ensemble des mondes préférés (au sens de $\geq_u$) de S. Si $D = \{D(\beta_1|\alpha_1), \ldots D(\beta_n|\alpha_n)\}$, on dit que la fonction u *satisfait* D (noté $u \models D$) ssi $\forall i \in 1 \ldots n, Max_{\geq_u}[\alpha_i] \subseteq [\beta_i]$, où $[\alpha_i]$ (resp. $[\beta_i]$) représente l'ensemble des mondes satisfaisant α_i (resp. β_i). Un ensemble de désirs conditionnels est dit *cohérent* ssi il existe une fonction d'utilité qui le satisfasse.

Lang [Lan96] utilise des fonctions d'utilité d'un type particulier. Avant tout, il associe à chaque défaut $D(\beta|\alpha)$ une fonction d'utilité *locale* $u_{\beta|\alpha}$, définie comme suit :

Définition 4.3.9 *Fonction d'utilité locale.*
Soit $D(\beta|\alpha)$ un désir conditionnel. $u_{\beta|\alpha}$ est une fonction d'utilité locale associée à $D(\beta|\alpha)$ ssi il existe $a > 0$ tel que : $u_{\beta|\alpha}(\omega) = -a$ si $\omega \models \alpha \wedge \neg\beta$ et $u_{\beta|\alpha}(\omega) = 0$ sinon.

Exemple 4.3.3 *Le parapluie.*
$D = \{D(\neg para|\top), D(\neg pluie|\top), D(para|pluie)\}$. *Les utilités locales sont de la forme :*

- $u_{\neg para|\top}(\omega) = -a$ *si* $\omega \models para$ *et 0 sinon,*
- $u_{\neg pluie|\top}(\omega) = -b$ *si* $\omega \models pluie$ *et 0 sinon,*

– $u_{para|pluie}(\omega) = -c$ *si* $\omega \models \neg para \wedge pluie$ *et 0 sinon.*

Lang définit ensuite la notion de fonction d'utilité *distinguée* pour un ensemble de désirs conditionnels :

Définition 4.3.10 *Fonction d'utilité distinguée pour* D.
u *est une fonction d'utilité distinguée pour* D *ssi*

– $u \models D$,

– $u = u_1 + \ldots + u_n$,
où $u_1, \ldots, u_n$ *sont des fonctions d'utilité locales associées à* $D(\beta_1|\alpha_1) \ldots D(\beta_n|\alpha_n)$.

On peut remarquer que les fonctions d'utilité distinguées ne sont jamais positives, puisque ce sont des sommes de fonctions d'utilités locales qui ne le sont pas non plus.
Une fonction d'utilité locale $u_{\beta|\alpha}$ associe une pénalité (ou coût de violation) à tout monde violant l'implication matérielle $\alpha \rightarrow \beta$. Une fonction d'utilité distinguée associe à chaque monde la somme des pénalités associées aux implications $\alpha_i \rightarrow \beta_i$ qu'il viole.

Exemple 4.3.4 *Le parapluie : calcul des fonctions d'utilité distinguées.*
On note $\omega_0 = \{\neg para, \neg pluie\}, \omega_1 = \{\neg para, pluie\}, \omega_2 = \{para, \neg pluie\}, \omega_3 = \{para, pluie\}$.
Soit $u = u_{\neg para|\top} + u_{\neg pluie|\top} + u_{para|pluie}$. $u(\omega_0) = 0$, $u(\omega_1) = -(b+c)$, $u(\omega_2) = -a$, $u(\omega_3) = -(a+b)$.

$u \models D(\neg para|\top)$ *et* $u \models D(\neg pluie|\top)$ *ne donnent pas d'information supplémentaire, en revanche,* $u \models D(para|pluie)$ *ssi* $Max_{\geq_u}\{\omega_1, \omega_3\} \subseteq \{\omega_2, \omega_3\}$, *i.e.* $u(\omega_3) > u(\omega_1)$, *soit* $c > a$.

En fin de compte, u *est distinguée pour* D *ssi* $u(\omega_0) = 0$, $u(\omega_1) = -(b+c)$, $u(\omega_2) = -a$ *et* $u(\omega_3) = -(a+b)$, *avec* $c > a$.

Préordre partiel entre les mondes

Nous venons de voir comment sont calculées les fonctions d'utilité distinguées pour un ensemble de désirs conditionnels dans (Lang [Lan96]). A partir de l'ensemble de ces fonctions, l'auteur détermine un préordre (partiel) entre les mondes, noté $\geq_D$. Celui-ci est défini par :

Définition 4.3.11 *Préordre entre les mondes, obtenu à partir d'un ensemble de fonctions d'utilité distinguées.*

$$\omega \geq_D \omega' \text{ ssi } \forall u, \text{ distinguée pour } D, \omega \geq_u \omega'$$

L'ordre strict et la relation d'équivalence associés à $\geq_D$ sont notés $>_D$ et $\sim_D$.

Exemple 4.3.5 *Suite.*
On ajoute à D *(qui devient* D'*) le désir* $D(pluie|para)$, *signifiant que si l'agent prend son parapluie, il préfère qu'il pleuve (pour ne pas le prendre pour rien).*
$u_{pluie|para} = -d$ *si* $\omega \models para \wedge \neg pluie$ *et 0 sinon.* $u' = u + u_{pluie|para}$ *:* u' *n'est modifiée par rapport à* u *qu'en* ω_2, *où* $u'(\omega_2) = -(a+d)$. $u' \models D(pluie|para)$ *ssi* $Max_{\geq_{u'}}\{\omega_2, \omega_3\} \subseteq \{\omega_1, \omega_3\}$, *i.e.* $d > b$.

Les contraintes d'ordre entre les mondes deviennent : $\omega_0 >_{D'} \omega_3$, $\omega_3 >_{D'} \omega_1$ *et* $\omega_3 >_{D'} \omega_2$. ω_1 *et* ω_2 *ne sont plus comparables : l'ordre est devenu partiel.*

Ensuite, Lang montre que les désirs conditionnels les plus spécifiques (au sens de Doyle et Welman et de Tan et Pearl) prennent le pas sur les autres pour la détermination de la relation de préférence sur les mondes.

Application à la décision dans l'incertain

Lang adopte la même séparation que Boutilier [Bou94] entre variables contrôlables ($\mathcal{C}$) et non contrôlables ($\overline{\mathcal{C}}$) pour représenter les actions de l'agent.
En revanche, pour représenter les connaissances de l'agent, Lang utilise une base de connaissances BC (un ensemble de formules propositionelles). Les définitions de la satisfaction d'un ensemble de désirs et des fonctions d'utilité locales sont légèrement modifiées pour tenir compte de la présence de la base de connaissances :

Définition 4.3.12 *Satisfaction d'un ensemble de désirs conditionnels D en présence d'une base de connaissances BC.*
$u \models < D, BC >$ *ssi*

- $\forall D(\beta|\alpha) \in D$, $Max_{\geq_u}[BC \wedge \alpha] \subseteq [\beta]$.
- $\forall \omega \models \neg BC, u(\omega)$ *est indéfinie.*

Définition 4.3.13 *Fonction d'utilité locale en présence d'une base de connaissances.*
$u_{\beta|\alpha}$ *est une fonction d'utilité locale associée à* $D(\beta|\alpha)$ *ssi il existe* $a > 0$ *tel que* $u_{\beta|\alpha}(\omega) = -a$ *si* $\omega \models BC \wedge \alpha \wedge \neg\beta$, $u_{\beta|\alpha}(\omega) = 0$ *si* $\omega \models BC \wedge (\alpha \to \beta)$, *et* $u_{\beta|\alpha}(\omega)$ *est indéfinie si* $\omega \models \neg BC$.

Dans un problème de décision sous incertitude, outre les connaissances génériques (BC), l'utilisateur dispose de connaissances *factuelles* (F), qui viennent compléter les connaissances génériques. Répétons que dans l'approche de Lang [Lan96] les connaissances factuelles ne doivent jamais remettre en cause les connaissances génériques.
La distinction entre connaissances génériques et factuelle que l'on retrouve chez Lang [Lan96] est similaire à celle que l'on retrouve chez Geffner [Gef92] ou van der Torre [van94].
Dans un premier temps, supposons que la connaissance factuelle est *complète*, i. e. $BC \wedge F$ admet un modèle unique ω_F sur $\overline{\mathcal{C}}$. Une action d correspond (comme chez Boutilier) à une interprétation sur $\mathcal{C}$. Lorsque la connaissance est complète, l'ordre entre les actions est déterminé par l'ordre $\geq_{<D,BC>}$.

Définition 4.3.14 *Ordre entre les actions lorsque la connaissance est complète.*
Soient d *et* d', *deux actions.* $d \geq_{<D,BC,F>} d'$ *ssi* $(\omega_F; d) \geq_{<D,BC>} (\omega_F; d')$.

Exemple 4.3.6 *Suite.*
$\mathcal{C} = \{para\}, \overline{\mathcal{C}} = \{pluie\}$. *Supposons que l'agent constate qu'il pleuve* ($F = pluie$) *et qu'il n'ait aucune autre information générique* ($BC = \top$)*. Alors la connaissance est complète* ($\omega_F = \{pluie\}$)*.*
L'ensemble des désirs conditionnels est D, *et* $\omega_0 \geq_D \omega_2 \geq_D \omega_3 \geq_D \omega_1$.
Or $(\omega_F; para) = \omega_3$ *et* $(\omega_F; \neg para) = \omega_1$, *on en déduit que* $para >_{<D,BC,F>} \neg para$.

Dans le cas où la connaissance est incomplète, $BC \wedge F$ ne détermine pas un monde unique, mais un sous-ensemble de mondes possibles. Alors, on se retrouve dans une situation d'incertitude non quantifiée, et Lang propose des méthodes de décisions, respectivement en accord avec les critères de Wald (ou *maximin*) et de Pareto (ou de *domination*) :

Définition 4.3.15 *Critère de Wald.*

$$d \geq^{Wald}_{<D,BC,F>} d' \textit{ ssi } \forall\omega \models BC \wedge F, \exists\omega' \models BC \wedge F, \textit{ tel que } (\omega; d) \geq_{<D,BC>} (\omega'; d').$$

Ce critère compare les pires interprétation satisfaisant chacune des décisions, au sens de $\geq_{<D,BC>}$.
Dans l'exemple du parapluie, $para \geq^{Wald}_{<D,BC,F>} \neg para$.
Cette méthode de décision génère un préordre complet (en accord avec le critère de Wald de la décision dans l'incertain) lorsque $\geq_{<D,BC>}$ est un préordre complet.

Lang définit enfin le critère de Pareto, similaire à la relation de dominance dans la théorie de la décision dans l'incertain :

Définition 4.3.16 *Critère de Pareto.*

$$d \geq^{Pareto}_{<D,BC,F>} d' \textit{ ssi } \forall\omega \models BC \wedge F, (\omega; d) \geq_{<D,BC>} (\omega; d').$$

Alors que le critère précédent représentait le critère de Wald, celui-ci est en accord avec le critère de Pareto. Dans l'exemple du parapluie, $\neg para$ et $para$ sont incomparables suivant l'ordre de Pareto.
Dans le cas général, $d \geq^{Pareto}_{<D,BC,F>} d' \Rightarrow d \geq^{Wald}_{<D,BC,F>} d'$: si une décision est meilleure qu'une autre au sens de Pareto, elle est également meilleure au sens de Wald.

4.4 D'autres approches logiques de la décision dans l'incertain

D'autres approches "logiques" de la décision dans l'incertain ont été proposées. Bien que toutes basées sur la logique, elles diffèrent par leur sémantique, i.e. par le critère de décision sous-jacent. (Bonet et Geffner, 1996 [BG96]), par exemple, proposent une approche logique basée sur les *systèmes d'argumentation*. Les problèmes de décision dans l'incertain sont modélisés dans un langage propositionnel *typé* (les granules de connaissances et les préférences sont ainsi identifiables). Ils proposent alors de calculer les connaissances qui *étayent* les préférences (i.e. permettant de les satisfaire lorsqu'une action donnée est effectuée). En associant des *degrés* (valeurs entières, *commensurables*) aux préférences (priorités) et aux connaissances (plausibilités), ils calculent *l'utilité espérée qualitative* de chaque action. La sémantique qui sous-tend leur approche est basée sur les travaux de (Pearl, 1993 [Pea93]) et (Wilson, 1995 [Wil95]) utilisant des fonctions kappa pour mesurer l'incertitude. Cette approche nécessite d'exprimer les degrés de préférence et de connaissance sur une même échelle, au moins additive.
Dans la seconde partie de cette thèse nous décrirons une approche ressemblant à celle de Bonet et Geffner, basée sur les *Assumption-based Truth Maintenance Systems* (ATMS) (de Kleer, 1986 [de 86]). La sémantique que nous présenterons pour cette approche est différente de celle de Bonet et Geffner. Nous verrons qu'elle est en accord avec le critère de *l'utilité qualitative possibiliste* que nous avons évoqué dans le chapitre 2 et que nous décrirons en détails bientôt.

4.5 Conclusion

Nous avons décrit un certain nombre d'approches récentes de la représentation structurée des problèmes de décision sous incertitude, basées sur la notion de "préférence conditionnelle".
Ces approches ont de nombreux points communs avec un autre domaine de l'Intelligence Artificielle : celui des *logiques déontiques*. Ces dernières traitent de ce qu'un agent *doit* faire, et non de ce qu'il *veut*. Le pendant des opérateurs d'idéalité, ou de désir conditionnel, est l'opérateur *d'obligation contextuelle*, $O(\cdot|\cdot)$. $O(\varphi|\psi)$ signifie "si la formule ψ est vraie, alors l'agent *doit* se retrouver dans une situation vérifiant φ" (voir (Makinson, 1993 [Mak93]) pour une étude détaillée de la notion d'obligation contextuelle). Evidemment, les notion d'obligation et de désir sont étroitement liées : l'agent cherche dans les deux cas à satisfaire un ensemble de buts. Dans le premier cas les buts lui sont imposés, alors que dans le second il se les fixe lui même (par l'intermédiaire de ses préférences). Dans (Lang, [Lan96]) ou (van der Torre,[van94], [van97] on peut trouver une discussion plus profonde sur les liens entre logiques pour la décision et logiques déontiques.

Le problème des approches logiques pour la décision dans l'incertain que nous avons décrit dans ce chapitre est que les sémantiques qui leur sont associées sont relativement pauvres. Les décisions calculées "syntaxiquement" ne sont optimales, dans le meilleur des cas, que suivant un critère du type *maxmin*. Cette "pauvreté" sémantique résulte, dans les modèles que nous avons exposés, de l'absence *d'hypothèse de commensurabilité* entre échelles de préférence et d'incertitude. Dans les approches de Boutilier et de Lang, on isole les situations les plus plausibles, puis on raisonne à partir d'elles seules pour déterminer des actions optimales, en fonction du seul ordre de préférence sur les conséquences. Tan et Pearl proposent bien de construire deux échelles distinctes pour mesurer les préférences et l'incertitude, mais ne comparent (au sens de la relation de préférence) que les conséquences d'égale "plausibilité". Qui plus est, leur méthode de décision ne permet pas de construire un préordre complet sur l'ensemble des actions, mais simplement de les comparer deux à deux. Au final, la relation de préférence entre actions qu'ils obtiennent n'est pas transitive.

Deuxième partie

Une approche ordinale de la décision dans l'incertain basée sur la théorie des possibilités

Chapitre 1

Des approches ordinales de la décision basées sur des mesures monotones

1.1 Introduction

Dans ce chapitre seront présentés les principaux résultats de cette thèse en ce qui concerne l'axiomatisation de critères "qualitatifs" pour la décision dans l'incertain.

Ici, le terme *qualitatif* est utilisé par opposition à l'aspect *quantitatif* de la théorie de l'utilité espérée et de ses extensions. Les approches quantitatives utilisent des représentations additives des préférences et de l'incertitude (quoique dans les approches basées sur l'intégrale de Choquet les mesures de l'incertitude peuvent être non-additives). Dans les approches qualitatives que nous allons décrire ici, une échelle ordinale unique sera utilisée pour mesurer à la fois l'incertitude et les préférences d'un agent.

Contrairement à l'approche de Savage [Sav54], nous ferons l'hypothèse que l'espace d'états est *fini*. Nous verrons que cette hypothèse est nécessaire pour les axiomatisations que nous proposerons, comme est nécessaire l'hypothèse d'un espace d'états infini, chez Savage.

Ce chapitre sera découpé comme suit :

- Dans un premier temps nous rappellerons l'approche de Dubois et Prade [DP95b], basée sur la théorie des possibilités, pour la décision dans l'incertain qualitatif. Nous décrirons leur axiomatisation, du type de celle de *von Neumann et Morgenstern* [vNM44], des critères qu'ils ont proposés.

- Ensuite, nous passerons plus de temps sur l'approche du type *Savage* que nous avons proposée pour l'axiomatisation de ces critères (voir [DPS97a]).

- Enfin, toujours dans un cadre du type *Savage*, nous proposerons une axiomatisation pour des critères de décision basés sur des *mesures monotones*, généralisant les mesures de possibilité et de nécessité (voir [DPS98]).

En premier lieu, nous allons commencer par décrire le concept de *loterie possibiliste*, et montrer l'axiomatisation des relations de préférence entre loteries possibilistes, proposée par Dubois et Prade [DP95b].

1.2 Préférences entre loteries possibilistes

Nous avons présenté les critères possibilistes de décision dans l'incertain, pessimiste et optimiste, dans le chapitre 2 page 21 de la partie I.
Rappelons l'expression de ces deux critères :

– Critère pessimiste
$$v_*(a) = inf_{s \in S} \max(n(\pi(s)), \mu(a(s)))$$

– Critère optimiste
$$v^*(a) = sup_{s \in S} \min(m(\pi(s)), \mu(a(s)))$$

On rappelle également que π est une fonction de S vers une échelle ordonnée, finie, $(L, \leq)$, et μ de X vers une échelle U. n est une fonction de renversement d'ordre de L vers U, et m est une fonction non-décroissante de L vers U. π est une distribution de possibilité normalisée, i.e. $\exists s_0 \in S$, $\pi(s_0) = 1_L$ (plus grand élément de $(L, \leq)$).

Ces deux critères ont été justifiés axiomatiquement, dans le style de von Neumann et Morgenstern [vNM44], par Dubois et Prade [DP95b].
Dans l'approche de von Neumann et Morgenstern le comportement d'un agent en face du risque est entièrement déterminé par ses préférences sur l'ensemble des distributions de probabilité sur les conséquences ou "loteries". Ces préférences sur les loteries probabilistes doivent respecter un certain nombre d'axiomes décrivant l'attitude d'un agent "rationnel" en face de l'incertitude. L'Utilité Espérée représente ainsi un critère simple pour ordonner les loteries et donc les actions puisqu'à chaque action en environnement incertain correspond une loterie.

L'idée principale des théories possibilistes pour la décision consiste à représenter les connaissances incertaines d'un agent par une distribution de possibilité π sur un ensemble d'états du monde S. Chaque action a détermine une distribution de possibilité π_a sur X à partir de π, par : $\pi_a(x) = \Pi(a^{-1}(x)) = \sup\{\pi(s) : a(s) = x\}$. Par conséquent, ordonner les actions revient à ordonner les distributions de possibilité. Cet ordre entre les loteries, déterminé par l'agent, exprime son attitude envers l'incertitude et ainsi le critère de décision qu'il adopte.

Soit $\pi_a(x)$, la "plausibilité" d'obtenir la conséquence x, une fois l'action a exécutée. Notre problème revient à déterminer quels axiomes l'ordre de préférence entre les distributions de possibilité sur X doit satisfaire afin de pouvoir être représenté par l'un des deux critères (pessimiste ou optimiste) exposés plus haut.
Soient x et y, deux éléments de X et λ et ν, deux éléments de l'échelle L tels que $\max(\lambda, \nu) = 1_L$, où 1_L représente l'élément maximal de l'échelle L (0_L représente son élément minimal).
On appelle *loterie possibiliste binaire* une distribution de possibilité π_a telle que

– $\pi_a(x) = \lambda$, $\pi_a(y) = \nu$

– $\pi_a(z) = 0_L, \forall z \neq x, z \neq y$.

Une telle loterie sera notée $(\lambda/x, \nu/y)$ et peut être interprétée par : "l'action a n'a que deux conséquences possibles, x et y, de degrés de possibilité respectifs λ et ν". a est bien entendu un *acte binaire*.
Plus généralement, une distribution de possibilité π_a sur X peut être vue comme une loterie à conséquences multiples $(\lambda_1/x_1, \ldots, \lambda_m/x_m)$, où $X = \{x_1, \ldots, x_m\}$ et $\lambda_i = \pi_a(x_i)$. Pour simplifier les notations nous omettrons l'indice a dans la suite de ce paragraphe.

La notation $(\lambda/\pi, \nu/\pi')$ symbolise une loterie "d'ordre supérieur" dont les conséquences sont deux autres loteries π et π'. On suppose toujours que $\max(\lambda, \nu) = 1_L$ et on peut montrer que si π et π' sont deux loteries normalisées, $(\lambda/\pi, \nu/\pi')$ en est également une.
A tout singleton $\{x_0\}$ correspond une distribution de possibilité π valant 0 partout sauf en x_0 (où π vaut 1). Une telle distribution de possibilité correspond à un acte *constant* dont l'unique conséquence possible est x_0. Par simple extension, à tout sous-ensemble A de X correspond une distribution de possibilité π_A telle que $\forall x \in A, \pi_A(x) = 1$ et $\forall x \in \bar{A}, \pi_A(x) = 0$.

Soit $\succeq$ la relation de préférence entre loteries possibilistes déterminée à partir des préférences d'un agent. $\succeq$ étend les préférences de l'agent sur X à l'ensemble des distributions de possibilité normalisées $Pi(X) \subseteq L^X$.
Les axiomes suivants déterminent la relation $\succeq$ dans le cas pessimiste [DGPZ98] :

– **Axiome 1** $\succeq$ est un préordre complet.

– **Axiome 2** (indépendance) : $\pi_1 \sim \pi_2 \Rightarrow (\lambda/\pi_1, \nu/\pi') \sim (\lambda/\pi_2, \nu/\pi')$.

– **Axiome 3** (continuité) : $\pi \succeq \pi' \Rightarrow \exists \lambda \in V, \pi' \sim (1/\pi, \lambda/X^1)$.

– **Axiome 4** (réduction des loteries) :

$$(\lambda/x, \nu/(\alpha/x, \beta/y)) \sim (\max(\lambda, \min(\nu, \alpha))/x, \min(\nu, \beta)/y).$$

– **Axiome 5** (aversion pour l'incertitude) : $\pi \leq \pi' \Rightarrow \pi \succeq \pi'$.

L'axiome 1 permet de représenter l'utilité des loteries sur une échelle complètement ordonnée. Les axiomes 2, 3 et 4 sont des contreparties des axiomes proposés par von Neumann et Morgenstern. L'axiome 4 permet de réduire des loteries d'ordre supérieur en loteries "standards". L'axiome 3 est motivé par la forme particulière des "mixtures" en théorie des possibilités (Dubois et Prade, [DP95a], Dubois et col. [DFPR96]). Il signifie que l'utilité d'une distribution de possibilité π baisse "sans saut" lorsque l'incertitude représentée par π augmente. L'axiome 5 stipule que plus l'incertitude est grande, plus la situation est risquée : le pire état de connaissance est celui d'incertitude totale.

Des axiomes 3 et 5, on peut déduire que si $A \subseteq X$ est un ensemble de conséquences, alors $\exists x \in A, x \sim A^2$ (Dubois et col. [DGPZ98]). Cette propriété (violée par l'utilité espérée) suggère que l'utilité qualitative pessimiste v_* n'est pas basée sur la notion de moyenne et d'actions répétées, mais s'applique à des problèmes de décision non répétitifs. L'idée sous-jacente est que lorsqu'une action est appliquée elle mènera automatiquement à une conséquence unique, $x \in A$, et l'utilité de l'action sera celle de la conséquence x. Cela revient à rejeter la notion de *valeur moyenne*. De l'axiome 5 on déduit que la loterie représentant A est équivalente à celle représentant l'action constante dont l'unique conséquence est le "pire" élément de A.

Le critère possibiliste pessimiste v_* est donc une extension du critère de Wald qui évalue les actions selon leur pire conséquence, quelle que soit sa plausibilité. Cependant il est moins pessimiste que le critère de Wald. Il se repose sur l'utilité des pires conséquences *un tant soit peu* plausibles de

1. X correspond à la distribution de possibilité π_X : $\forall s \in X, \pi_X(s) = 1$.

2. x correspond à la distribution de possibilité valant 1 sur x et 0 ailleurs, A correspond à la distribution de possibilité π_A : $\pi_A(s) = 1 \forall s \in A$ et 0 ailleurs.

l'action. Les états dont la plausibilité est inférieure à un certain seuil variable (déterminé à partir de comparaisons entre les distributions π et μ) sont négligés dans l'évaluation de l'action par v_*.
En fait, l'action a aura une faible utilité $v_*(a)$ dès lors qu'il existe une conséquence x *suffisamment plausible* et d'utilité faible.

Un ensemble d'axiomes dual du précédent permet de définir l'utilité possibiliste *optimiste* v^* (voir (Dubois et col. [DGPZ98])).

- **Axiome 1** $\succeq$ est un préordre complet.
- **Axiome 2** (indépendance) : $\pi_1 \sim \pi_2 \Rightarrow (\lambda/\pi_1, \nu/\pi') \sim (\lambda/\pi_2, \nu/\pi')$.
- **Axiome 3'** (continuité) : $\pi \succeq \pi' \Rightarrow \exists \lambda \in V, \pi \sim (1/\pi', \lambda/X)$.
- **Axiome 4** (réduction des loteries) :

$$(\lambda/x, \nu/(\alpha/x, \beta/y)) \sim (\max(\lambda, \min(\nu, \alpha))/x, \min(\nu, \beta)/y).$$

- **Axiome 5'** (propension pour l'incertitude) : $\pi \geq \pi' \Rightarrow \pi \succeq \pi'$.

Le critère v^* est très optimiste dans la mesure où il est élevé dès lors qu'il existe une conséquence plausible de l'action qui a une utilité élevée. C'est pourquoi ce dernier critère n'est utilisé, en général, que pour départager des actions indifférentes au sens du critère pessimiste.

Exemple 1.2.1 *L'omelette de Savage, version possibiliste.*

Reprenons l'exemple de l'omelette exposé dans le chapitre 1 de la partie I, résumé par le tableau suivant :

	États du monde	
Actions	*oeuf sain (s)*	*oeuf pourri (p)*
Casser l'oeuf dans le bol (CB)	*omelette à 6 oeufs (6O)*	*pas d'omelette, 5 oeufs gâchés (G)*
Le casser dans une tasse (CT)	*omelette à 6 oeufs une tasse à laver (6T)*	*omelette à 5 oeufs une tasse à laver (5T)*
Le jeter (J)	*omelette à 5 oeufs un oeuf gâché (5G)*	*omelette à 5 oeufs (5O)*

L'échelle d'utilité symbolique $L = \{0, 1, 2, 3, 4, 5\}$ est utilisée pour représenter les niveaux de certitude et de préférence. Les nombres sont utilisés ici uniquement pour refléter la structure ordinale de L : des qualificatifs linguistiques du type faible, moyen, élevé... auraient tout aussi bien pu être utilisés.
L'ordre de préférence entre les conséquences peut être exprimé par la fonction d'utilité symbolique μ, affectant les utilités suivantes aux conséquences : $\mu(6O) = 5$, $\mu(6T) = 4$, $\mu(5O) = 3$, $\mu(5T) = 2$, $\mu(5G) = 1$ et $\mu(G) = 0$.

Dans cet exemple la distribution de possibilité π_a contraignant les valeurs des conséquences possibles de a ne dépend que des degrés de possibilité respectifs des deux événements s et p, soit $\Pi(s)$

et $\Pi(p)$. $N(s) = n(\Pi(p))$ *et* $N(p) = n(\Pi(s))$ *sont les degrés de nécessité respectifs des deux événements* s *et* p. *Puisque la distribution de possibilité sur* $\{s, p\}$ *est normalisée,* $\min(N(s), N(p)) = 0$.

Les utilités (pessimistes) des différentes actions peuvent être évaluées en fonction des degrés de nécessité de s *et* p *:*

- $v_*(CB) = \min(\max(n(\Pi(p)), \mu(G)), \max(n(\Pi(s)), \mu(6O)))$,
  qui se simplifie en $v_*(CB) = N(s)$.
- $v_*(CT) = \min(\max(n(\Pi(p)), \mu(5T)), \max(n(\Pi(s)), \mu(6T)))$,
  soit $v_*(CT) = \min(\max(N(s), 2), 4)$.
- $v_*(J) = \min(\max(n(\Pi(p)), \mu(5O)), \max(n(\Pi(s)), \mu(5G)))$.

Les meilleures décisions selon le critère pessimiste sont donc :

- *de casser l'oeuf dans l'omelette si* $N(s) = 5$*, i.e. si l'agent est absolument certain qu'il est sain,*
- *de le casser dans l'omelette ou alors dans une tasse si* $N(s) \in \{2, 3, 4\}$*, i.e. si l'agent a de bonnes raisons de penser qu'il est sain,*
- *de le casser dans une tasse si l'agent est ignorant de son état de fraîcheur (*$N(s) < 2$ *et* $N(p) < 2$*),*
- *de le jeter, enfin, si l'agent pense qu'il a des chances d'être pourri (*$N(p) > 2$*).*

1.3 Axiomatisations des utilités qualitatives possibilistes

1.3.1 Rappel de l'axiomatisation de Savage pour le critère de l'utilité espérée

Nous rappelons pour mémoire les axiomes **Sav 1** à **Sav 7** de Savage qui permettent de justifier le critère de l'utilité espérée pour la décision dans l'incertain :

Sav1 Pr'eordre complet
$(\mathcal{A}, \preceq)$ est un préordre complet, i.e. $\preceq$ est réflexif, transitif et complet.

Sav2 Principe de la chose certaine
$\forall$ **f**, **g**, **h**, **h'** $\in \mathcal{A}$, $\forall A \subseteq S$, **f**A**h** $\preceq$ **g**A**h** $\Leftrightarrow$ **f**A**h'** $\preceq$ **g**A**h'**.

Sav3 Conditionnement restreint aux actes constants
Soient $x, y \in X$ et A un événement *non impossible*. Soient les actions constantes **x** $\equiv x$ et **y** $\equiv y$. Alors, (**x** $\preceq$ **y**)$_A \Leftrightarrow x \leq_P y$.

Sav4 Projection sur l'ensemble des événements
$\forall x, y, x', y' \in X, y <_P x, y' <_P x', \forall A, B \subseteq S, xAy \preceq xBy \Leftrightarrow x'Ay' \preceq x'By'$.

Sav 5 Non trivialité
$\exists x, y \in X, y <_P x$, où $<_P$ représente la partie stricte de $\leq_P$.

Sav6 Probabilité quantitative
Soient **f** et **g** $\in \mathcal{A}$ tels que **f** $\prec$ **g**, soit $x \in X$. Il existe une partition $\bigcup A_i$ de S telle que pour tout i, xA_i**f** $\prec$ **g** et **f** $\prec xA_i$**g**.

Sav7 Extension à un ensemble infini de conséquences
Soient **f** et **g** $\in \mathcal{A}$, $A \subseteq S$. $(\mathbf{f} \preceq \mathbf{g}(s))_A, \forall s \in A \Rightarrow (\mathbf{f} \preceq \mathbf{g})_A$.

1.3.2 Propriétés des utilités qualitatives possibilistes

Un des postulats les plus importants de Savage est le *Principe de la chose certaine* qui exprime que si une action **f** est préférée ou équivalente à une autre action **g** et que ces deux actions donnent des résultats identiques sur un sous ensemble d'états B, alors si les résultats des deux actions sont modifiés de manière identique sur B l'ordre entre les deux nouvelles actions n'est pas modifié.

Cependant, deux actions peuvent être équivalentes aux yeux d'un agent parce qu'elles donnent des résultats identiques (extrêmement bons ou mauvais) sur $B \subseteq S$, alors que si ces résultats étaient modifiés (devenant moins "dramatiques") l'agent ne serait plus indifférent entre les nouvelles actions. Bien sûr, modifier uniformément des conséquences identiques ne doit pas inverser l'ordre de préférence entre les actions.
Ce constat suggère une nouvelle forme de *principe de la chose certaine* que nous appellerons *principe d'indépendance faible* (IF) :

IF : Indépendance faible.

Soient **f**, **g**, **h** et **h'** $\in \mathcal{A}$. $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h}$ implique $\mathbf{f}A\mathbf{h'} \succeq \mathbf{g}A\mathbf{h'}$.

IF permet une sorte "d'effet de noyade" : la préférence stricte entre deux actions peut être transformée en indifférence si on change de manière identique leurs conséquences communes.

Proposition 1.3.1 *Les utilités possibilistes v_* et v^* satisfont au principe d'indépendance faible mais pas au principe de la chose certaine.*

Preuve :

$v_*(\mathbf{f}A\mathbf{h}) = \min(inf_{s\in A} \max(n(\pi(s)), \mu(\mathbf{f}(s))), inf_{s\in \bar{A}} \max(n(\pi(s)), \mu(\mathbf{h}(s))))$. Notons $v_*^B(\mathbf{f}) = inf_{s\in B} \max(n(\pi(s)), \mu(\mathbf{f}(s))), \forall B \subseteq S$.

- Si le terme $v_*^{\bar{A}}(\mathbf{h})$ est plus petit que les deux termes $v_*^A(\mathbf{f})$ et $v_*^A(\mathbf{g})$ alors on trouve $\mathbf{f}A\mathbf{h} \sim \mathbf{g}A\mathbf{h}$. Si on change maintenant **h** en l'action constante donnant toujours la conséquence x^*, $v_*(\mathbf{f}Ax^*) = v_*^A(\mathbf{f})$ et $v_*(\mathbf{g}Ax^*) = v_*^A(\mathbf{g})$, puisque $v_*^{\bar{A}}(x^*) = 1$. Ainsi le principe de la chose certaine est violé si $v_*^A(\mathbf{g}) \succ v_*^A(\mathbf{f})$.
- Il ne peut y avoir de renversement des préférences strictes : $v_*(\mathbf{f}A\mathbf{h}) = \min(v_*^A(\mathbf{f}), v_*^{\bar{A}}(\mathbf{h}))$. donc $v_*(\mathbf{f}A\mathbf{h}) > v_*(\mathbf{g}A\mathbf{h}) \Rightarrow v_*^A(\mathbf{g}) < \min(v_*^A(\mathbf{f}), v_*^{\bar{A}}(\mathbf{h}))$, soit, $v_*^A(\mathbf{g}) < v_*^A(\mathbf{f})$.
  Si on suppose l'existence de **h'** tel que $v_*(\mathbf{f}A\mathbf{h'}) < v_*(\mathbf{g}A\mathbf{h'})$, alors $v_*^A(\mathbf{f}) < \min(v_*^A(\mathbf{g}), v_*^{\bar{A}})\mathbf{h'}))$, donc $v_*^A(\mathbf{f}) < v_*^A(\mathbf{g})$, ce qui entraîne une contradiction.

On peut montrer de manière similaire que v^* viole également ce principe.

La notion de *Préférence conditionnelle*, définie dans le chapitre 2 de la partie I doit être revue dans le cadre possibiliste. Trois situations doivent être envisagées :

– Si deux actions **f** et **g** sont telles que pour toute troisième action **h**, $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h}$ on peut dire que $(\mathbf{f} \succ \mathbf{g})_A$ tient "fortement".

– Si **f** et **g** sont telles que pour toute troisième action **h**, $\mathbf{f}A\mathbf{h} \sim \mathbf{g}A\mathbf{h}$ on peut dire que $(\mathbf{f} \sim \mathbf{g})_A$ tient "fortement".

– Enfin, si **f** et **g** sont telles que pour certaines actions **h**, $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h}$ et pour d'autres $\mathbf{f}A\mathbf{h} \sim \mathbf{g}A\mathbf{h}$ on peut dire que $(\mathbf{f} \succ \mathbf{g})_A$ tient "faiblement".

L'exposé de ces trois situations souligne que la notion de préférence conditionnelle doit être prise avec précaution en l'absence du principe de la chose certaine. En particulier la vision de Savage du conditionnement est hypothétique puisqu'il compare deux actions **f** et **g** sur un sous-ensemble d'états A, sans se préoccuper si l'événement $\bar{A}$ est possible ou non, et pour n'importe quelle action **h**. Ce type de conditionnement peut poser des problèmes si les utilités considérées sont non-additives puisque l'ordre entre les actions $\mathbf{f}A\mathbf{h}$ et $\mathbf{g}A\mathbf{h}$ peut dépendre de l'action **h**.

Au contraire, si le conditionnement est effectué après révision des connaissances par une information prouvant que $\bar{A}$ est impossible, toutes les actions deviennent équivalentes sur $\bar{A}$ (l'événement $\bar{A}$ devient "nul") et les préférences conditionnelles doivent être définies en restreignant la comparaison aux éléments de A. Dans ce cas il est possible de définir la relation de préférence conditionnelle $(\mathbf{f} \succ \mathbf{g})_A$ par $\mathbf{f}Ax_* \succ \mathbf{g}Ax_*$ dans le cas pessimiste (x_* est la pire conséquence de X) et par $\mathbf{f}Ax^* \succ \mathbf{g}Ax^*$ dans le cas optimiste. Ce type d'approche correspond à une *révision des préférences*.
Le traitement axiomatique du conditionnement hypothétique requiert l'étude d'une famille de relations de préférence $\succeq_A$ sur X^A pour tout sous-ensemble $A \subseteq S$, représentant directement les préférences conditionnelles de l'agent. Lehmann [Leh96] propose quelques voies de recherche dans cette direction. Nous ne poursuivons pas dans cette thèse dans la voie du conditionnement hypothétique mais nous nous restreindrons aux préférences entre *actions mixtes* du type $\mathbf{f}A\mathbf{h}$.

Revenons maintenant aux liens entre l'axiomatique de Savage et le décision possibiliste. Le non-respect de l'axiome de la chose certaine par les utilités v_* et v^* entraîne également le non-respect des axiomes **Sav 3** et **Sav 4**. En fait, celles-ci satisfont des versions affaiblies de ces deux axiomes.

L'axiome de *Cohérence faible avec les actes constants* (CFAC) est une version affaiblie de l'axiome **Sav 3** :

CFAC *Cohérence faible avec les actes constants*.
Soient les deux conséquences x et y. $x \geq_P y \Rightarrow xA\mathbf{h} \succeq yA\mathbf{h}$, $\forall A \subseteq S$ et $\forall$ $\mathbf{h} \in \mathcal{F}$.

L'axiome **CFAC** est satisfait par les deux formes d'utilité qualitative possibiliste. Cependant, celles-ci violent **Sav 3** pour la même raison pour laquelle elles violent **Sav 2** : l'effet de "noyade" d'une action **h** dont les conséquences sont extrêmes sur des actions mixtes du type $xA\mathbf{h}$.

Si les utilités possibilistes ne sont pas strictement cohérentes avec la relation de préférence correspondante entre actions constantes, elles vérifient tout de même une cohérence avec la relation de *domination* ou relation de Pareto, dont nous avons donné la définition dans le chapitre 1 de la partie I et que nous redonnons ici :

Définition 1.3.1 *Pareto-domination.*
On dit que **f** *domine (au sens de Pareto) une action* **g**, *noté* $\mathbf{f} \geq_P \mathbf{g}$ *si et seulement si* $\forall s \in S$, $\mathbf{f}(s) \geq_P \mathbf{g}(s)$ *(au sens de la relation de préférence sur* X*).*

La relation de domination qui est une extension simple de la relation de préférence $\geq_P$ sur X est notée de la même manière. Dans la terminologie des ensembles flous la relation de domination correspond à l'inclusion (des ensembles flous). On vérifie facilement que les deux fonctions d'utilité v_* et v^* sont en accord avec la relation de domination : $\mathbf{f} \geq_P \mathbf{g} \Rightarrow v_*(\mathbf{f}) \geq v_*(\mathbf{g})$ et $v^*(\mathbf{f}) \geq v^*(\mathbf{g})$.

De la même manière, les utilités qualitatives possibilistes sont en accord avec une certaine forme de *domination stochastique*[3].

Définition 1.3.2 *Domination possibiliste.*
On dit que **f** *domine une action* **g** *au sens d'une mesure de possibilité* Π *(noté* $\mathbf{f} \geq_\Pi \mathbf{f}$*) si et seulement si* $\Pi(F_x) \geq \Pi(G_x), \forall x$, *où* $F_x = \{s \in S, \mathbf{f}(s) \geq_P x\}$ *et* $G_x = \{s \in S, \mathbf{g}(s) \geq_P x\}$ *sont respectivement les alpha-coupes au niveau* x *de* **f** *et* **g**. *De même, on dit que* **f** *domine une action* **g** *au sens d'une mesure de nécessité* N *(noté* $\mathbf{f} \geq_\Pi \mathbf{f}$*) si et seulement si* $\mathsf{N}(F_x) \geq \mathsf{N}(G_x), \forall x$.

On montre encore que v_* est en accord avec la relation $\geq_N$ et que v^* est en accord avec la relation $\geq_\Pi$.

En fait, les utilités possibilistes satisfont l'axiome **Grant** qui, pour Grant, Kajii et Polak [GKP97] est une forme affaiblie du *principe de la chose certaine* :

Grant *Principe de la chose certaine affaibli.*
Si $\mathbf{f}A\mathbf{g} \succ \mathbf{f}$ et $\mathbf{g}A\mathbf{f} \succ \mathbf{f}$ alors $\mathbf{g} \succ \mathbf{f}$.

D'après ce principe, si le fait de changer **f** en **g** sur A ou sur $\bar{A}$ améliore l'utilité de **f**, alors **g** est forcément préférée à **f**, quel que soit l'événement A.
Les utilités possibilistes pessimiste et optimiste respectent l'axiome **Grant**, mais de deux manières particulières et différentes :

PES Pessimisme :
$\forall$ **f**, **g**, $\forall A \subseteq S$, [$(\mathbf{f}A\mathbf{g} \succ \mathbf{f}) \Rightarrow (\mathbf{f} \succeq \mathbf{g}A\mathbf{f})$].

OPT Optimisme :
$\forall$ **f**, **g**, $\forall A \subseteq S$, [$(\mathbf{f} \succ \mathbf{f}A\mathbf{g}) \Rightarrow (\mathbf{g}A\mathbf{f} \succeq \mathbf{f})$].

Les axiomes **PES** et **OPT** impliquent tous deux l'axiome **Grant**.
Preuve :

- Si $\mathbf{f}A\mathbf{g} \succ \mathbf{f}$, d'après **PES**, "$\mathbf{g}A\mathbf{f} \succ \mathbf{f}$" n'est pas vérifié, donc la précondition de l'axiome **Grant** est toujours fausse.

3. La définition originelle de la domination stochastique est reliée à la notion de *mesure de probabilité* mais peut être généralisée à toute mesure d'incertitude monotone.

– Pour prouver que **OPT** $\Rightarrow$ **Grant**, posons **f'** = **f**A**g** et **g'** = **g**A**f**. **OPT** s'écrit **f'**A**g'** $\succ$ **f'**$\Rightarrow$ **g'** $\succ$ **f'**A**g'**. Par transitivité de $\succ$, **f'**A**g'** $\succ$ **f'** $\Rightarrow$ **g'** $\succ$ **f'**. Donc **Grant** est vérifié.

L'axiome de pessimisme peut être interprété de la manière suivante : Étant donnée l'action **f**, si le fait de changer **f** en l'action **g** lorsque l'événement $\bar{A}$ intervient "améliore" l'action **f** aux yeux de l'agent alors il n'y a aucun moyen d'améliorer l'action **f** en la modifiant lorsque A intervient. Ceci, tout simplement parce que l'agent considère que A est au moins aussi plausible que $\bar{A}$ et qu'à cause de son "pessimisme", il néglige totalement les éventuelles bonnes conséquences, obtenues lorsque $\bar{A}$ intervient.
Illustrons cette notion de pessimisme par l'exemple suivant :

Exemple 1.3.1 *Un conducteur doit se rendre par la route à un rendez-vous important à l'autre bout de la ville. Il souhaite bien entendu arriver à l'heure et toute l'incertitude du problème est résumée dans l'événement incertain A : "il y a des embouteillages". Il peut choisir trois itinéraires différents :*

– *Action* **f** *: passer à travers le centre-ville. S'il n'y a pas d'embouteillages il atteint son rendez-vous à l'heure, s'il y en a, il ratera son rendez-vous.*

– *Action* **g** *: contourner la ville par l'Est. Même si il n'y a pas d'embouteillages l'agent aura du retard à son rendez-vous. Si il y en a, il le manquera.*

– *Action* **h** *: contourner la ville par l'Ouest. Quel que soit l'état du trafic l'agent aura du retard à son rendez-vous, mais pas suffisamment pour le manquer.*

L'attitude typique de l'agent pessimiste est, par crainte des embouteillages, d'être indifférent entre **f** *et* **g***, et de leur préférer strictement* **h***.*
Au contraire, un agent optimiste préférera l'action **f** *qui est la seule à lui permettre (dans le meilleur des cas) d'être à l'heure à son rendez-vous et sera indifférent entre les actions* **g** *et* **h** *qui induisent le même retard dans le meilleur des cas.*

Nous allons voir que la relation de préférence entre actions induite par l'utilité pessimiste satisfait **PES**, alors que l'utilité optimiste satisfait **OPT** (les deux satisfont donc l'axiome **Grant**.

Avant de le vérifier, introduisons les notions de *conjonction* et *disjonction* d'actions. Soient **f** et **g**, deux actions. La conjonction de **f** et **g** (resp. la disjonction), notée **f** $\wedge$ **g** (resp.**f** $\vee$ **g**) est l'action qui donne, dans tout état possible du monde s, la moins bonne des deux conséquences **f**(s) et **g**(s) (resp. la meilleure), au sens de la relation de préférence $\geq_P$ sur X. En termes d'ensembles flous il s'agit tout simplement de l'intersection (resp. l'union).
Il est facile de montrer (grâce aux propriétés élémentaires des opérateurs min et max) que les deux propriétés suivantes sont vérifiées par les utilités possibilistes, alors qu'elles ne le sont pas par l'utilité espérée :

Lemme 1.3.1 *min et max-décomposabilité.*

$$v_*(\mathbf{f} \wedge \mathbf{g}) = \min(v_*(\mathbf{f}), v_*(\mathbf{g})), et$$

$$v^*(\mathbf{f} \vee \mathbf{g}) = \max(v^*(\mathbf{f}), v^*(\mathbf{g})).$$

Les propriétés duales ne sont pas vérifiées, sauf dans les cas où l'une des deux actions **f** et **g** est une action constante (on parle dans ce cas de *semi-décomposabilité*). Ceci peut à nouveau être facilement vérifié, grâce aux propriétés élémentaires de min et max.

Soient **f** et **g** deux actions quelconques et **x** l'action dont l'unique conséquence est x. Soit $\succ$, l'ordre strict induit par l'une des fonctions d'utilité (pessimiste ou optimiste). Les deux axiomes suivants sont liés aux propriétés de décomposabilité des fonctions d'utilité possibilistes :

DCR : *Domination conjonctive restreinte.*
g $\succ$ **f** et **x** $\succ$ **f** $\Rightarrow$ **g** $\wedge$ **x** $\succ$ **f**.

DDR : *Domination disjonctive restreinte.*
f $\succ$ **g** et **f** $\succ$ **x** $\Rightarrow$ **f** $\succ$ **g** $\vee$ **x**.

Un ordre entre les actions basé sur l'utilité espérée ne vérifie pas forcément ces deux axiomes. Pour le montrer dans le cas de **DCR** il suffit de trouver des valeurs réelles a, b, a', b' et un nombre α compris entre 0 et 1, tels que $a \cdot \alpha + b \cdot (1-\alpha) > a' \cdot \alpha + b' \cdot (1-\alpha)$, $c > a' \cdot \alpha + b' \cdot (1-\alpha)$ et $\min(a,c) \cdot \alpha + \min(a,b) \cdot (1-\alpha) \leq a' \cdot \alpha + b' \cdot (1-\alpha)$. De telles valeurs existent et le lecteur peut vérifier que $a = 1000$, $b = 2$, $a' = 3$, $b' = 100$ et $\alpha = 0,93$ forment un contre-exemple.

La propriété de décomposabilité de la fonction d'utilité optimiste v^* pour l'opérateur max, que nous appellerons *maxitivité* est une contrepartie de la propriété *d'additivité* de l'utilité espérée : l'utilité espérée d'une action dont les conséquences ont pour utilité la somme des utilités des conséquences de deux actions **f** et **g** est la somme des utilités espérées de **f** et **g**.
De même, la propriété de décomposabilité restreinte de v^* pour l'opérateur min est une contrepartie de la décomposabilité de l'utilité espérée pour la multiplication par une constante.

Ces propriétés ont été utilisées par Campos et Bolaños [dB92] pour caractériser la possibilité d'un événement flou (ils n'étudient pas le cas dual de la nécessité d'un événement flou, i.e. de la minitivité et de la semi décomposabilité de v_*).
Ces propriétés de décomposabilité sont reliées aux axiomes de pessimisme et d'optimisme :

Proposition 1.3.2 *L'utilité pessimiste v_* satisfait* **PES** *et l'utilité optimiste v^* satisfait* **OPT**.

Preuve (pour v_*) :

> Supposons $v_*(\mathbf{f}A\mathbf{g}) > v_*(\mathbf{f})$ et $v_*(\mathbf{g}A\mathbf{f}) > v_*(\mathbf{f})$. Alors $\min(v_*(\mathbf{f}A\mathbf{g}), v_*(\mathbf{g}A\mathbf{f})) > v_*(\mathbf{f})$.
>
> Or si on utilise la propriété de décomposabilité par le min de l'utilité pessimiste, la dernière inégalité peut s'écrire $v_*(\mathbf{f}A\mathbf{g} \wedge \mathbf{g}A\mathbf{f}) > v_*(\mathbf{f})$. Comme $\mathbf{f}A\mathbf{g} \wedge \mathbf{g}A\mathbf{f} = \mathbf{f} \wedge \mathbf{g}$ on obtient finalement $v_*(\mathbf{f} \wedge \mathbf{g}) > v_*(\mathbf{f})$ ce qui est impossible puisque **f** domine (au moins faiblement) $\mathbf{f} \wedge \mathbf{g}$ et que l'utilité pessimiste respecte la relation de domination.
>
> Pour finir il suffit d'observer que la négation de $v_*(\mathbf{f}A\mathbf{g}) > v_*(\mathbf{f})$ et $v_*(\mathbf{g}A\mathbf{f}) > v_*(\mathbf{f})$ est précisément l'axiome de pessimisme.
>
> Une preuve similaire peut être trouvée pour l'utilité optimiste (qui vérifie l'axiome d'optimisme **OPT**).

Attardons-nous maintenant quelque peu sur les actions binaires du type $xAy (x >_P y)$. En premier lieu, notons que

$$v_*(xAy) = \max(\mu(y), \min(N(A), \mu(x))) = \min(\mu(x), \max(N(A), \mu(y)))$$

Cette forme de pessimisme est facile à comprendre : si l'agent est suffisamment sûr que A intervient ($N(A) > \mu(x)$) alors l'utilité de l'action xAy est $\mu(x)$.
S'il a trop peu de connaissances ($\max(N(A), N(\bar{A})) \leq \mu(y)$) alors l'utilité est $\mu(y)$ i.e. celle de la pire conséquence. Bien sûr c'est également le cas lorsque l'agent pense que $\bar{A}$ est vrai. Si la certitude que A intervient est positive, mais pas extrême ($\mu(y) < N(A) < \mu(x)$), l'utilité de l'action reflète cette certitude.
On notera une nouvelle analogie avec l'utilité espérée qui est une moyenne, dans le fait que l'utilité qualitative pessimiste est la *médiane* de $\{\mu(x), \mu(y), N(A)\}$.

De la même manière, l'utilité optimiste d'une action binaire prend la forme

$$v^*(xAy) = \max(\min(\Pi(A), \mu(x)), \mu(y))$$

qui se trouve être aussi la médiane de $\{\mu(x), \mu(y), \Pi(A)\}$. L'utilité d'une action binaire xAy est $\mu(x)$ dès lors que x devient suffisamment plausible ($\Pi(A) \geq \mu(x)$).

Les utilités pessimiste et optimiste violent l'axiome **Sav 4** à cause de l'effet "noyant" d'événements relativement certains aux conséquences extrêmes. En effet, considérons les actions binaires xAx', xBx', yAy' et yBy'. Rien n'empêche d'avoir $v_*(xAx') = N(A) > v_*(xBx') = N(B)$ et $v_*(yAy') = v_*(yBy') = \mu(y)$, par exemple lorsque $\mu(y) \leq \min(N(A), N(B))$.

Comme dans l'approche de Savage on peut considérer $\mathcal{F}_{xy}$ (avec $x > y$), l'ensemble des actes de la forme xAy, isomorphe à $\mathcal{P}(S)$, l'ensemble des parties de S. On peut encore définir $\trianglerighteq^{xy}$ le préordre complet sur les événements induit par la relation de préférence entre actions binaires $\succeq$ par $A \trianglerighteq^{xy} B \Leftrightarrow xAy \succeq xBy$.
Savage [Sav54] requiert, par son axiome **Sav 4** que l'ordre induit entre les événements ne dépende pas de x et y. Dans l'approche possibiliste cette propriété n'est pas vérifiée. Toutefois les utilités possibilistes satisfont une version affaiblie de **Sav 4** dans laquelle les préférences entre événements restent faiblement cohérentes lorsque x et y varient.

Sav 4' *Cohérence faible entre actions.*
Soient $x > x', y > y'; A, B \subseteq S : xAx' \succ xBx' \Rightarrow yAy' \succeq yBy'$. Si on a en plus $x' \leq y' < y \leq x$, alors : $xAx' \succeq xBx' \Rightarrow yAy' \succeq yBy'$.

Ceci signifie que si deux actions binaires aux conséquences identiques sont équivalentes, changer les conséquences en conséquences *moins extrêmes* conserve les deux actions équivalentes. De plus, dans tous les cas, changer les conséquences de manière identique (en conservant l'ordre entre les conséquences des actions binaires) ne peut créer un renversement de préférences : on ne peut avoir à la fois $v_*(xAx') > v_*(xBx')$ et $v_*(yAy') < v_*(yBy')$.

Proposition 1.3.3 *L'axiome* **Sav 4'** *est vérifié par les deux fonctions d'utilité possibilistes.*

Preuve (pour v_*) :

Si $v_*(xAy) > v_*(xBy)$ plusieurs cas peuvent intervenir :

– $v_*(xAy) = N(A) > v_*(xBy) = \mu(y)$. Ceci implique que $\mu(x) \geq N(A) > \mu(y) \geq N(B)$, donc $N(A) > N(B)$.

– Le même résultat est valable si $v_*(xAy) = N(A)$ et $v_*(xBy) = N(B)$.
– $v_*(xAy) = \mu(x)$ et $v_*(xBy) = \mu(y)$ ou $N(B)$. On a également $N(A) > \mu(x) \geq N(B)$.

On a ainsi montré que si $v_*(xAy) > v_*(xBy)$ alors $N(A) > N(B)$. On en déduit que $v_*(x'Ay') \geq v_*(x'By')$ puisque la fonction $\min(a, \max(b, c))$ est non-décroissante en ses variables.

Si on choisit en plus $\mu(x') > \mu(x) > \mu(y) > \mu(y')$ on peut montrer que dans les trois cas $v_*(xAy)$ ne peut que croître et $v_*(xBy)$ ne peut que décroître, ainsi $v_*(xAy) > v_*(xBy) \Rightarrow v_*(x'Ay') > v_*(x'By')$.

Le même type de raisonnement peut être tenu pour l'utilité optimiste.

1.3.3 Axiomatisations des utilités qualitatives possibilistes, théorème de représentation

Ici nous allons montrer que les utilités possibilistes pessimistes et optimistes peuvent être axiomatisées dans le style de Savage, tout comme l'utilité espérée. La principale différence réside dans le fait qu'un cadre fini est suffisant pour prouver les résultats.
De plus, nous décrirons un cadre général décrivant plusieurs familles de mesures monotones sur les événements en termes d'actions, et nous fournirons par la même occasion un cadre axiomatique pour plusieurs théories non-probabilistes de l'incertain.
Ces cadres sont "testables" : en demandant à un agent d'ordonner un certain nombre d'actions en environnement incertain on peut déterminer avec quelle mesure de l'incertain il travaille implicitement.

Mesures d'incertitudes induites par les préférences entre les actions

Le type de représentation de l'incertitude le plus général que nous utiliserons est basé sur des fonctions d'ensemble $\sigma : S \to L$ qui sont connues sous les noms de *mesures de Sugeno* ou *mesures monotones* [Sug77]. Les propriétés de ces mesures sont les suivantes :

$$\sigma(\emptyset) = 0_L, \sigma(S) = 1_L, \text{ et } A \subseteq B \Rightarrow \sigma(A) \leq \sigma(B)$$

Ce type de fonction d'ensemble est très général et ses propriétés sont en quelque sorte les propriétés minimales que l'on peut exiger d'une fonction de représentation de connaissances partielles. La dernière condition, appelée condition de *monotonie* est vérifiée par les mesures de probabilité et de nombreuses autres mesures.
Les mesures de Sugeno peuvent être définies en termes de préférence entre actions, en considérant à nouveau la restriction de la relation de préférence sur les actes binaires.
Très peu d'axiomes sont nécessaires pour obtenir le résultat, néanmoins la notion de préférence conditionnelle évoquée précédemment est utilisée.

Rappel : **f** est dite "faiblement préférée à **g**, conditionnellement à A, noté **(f** $\succeq$ **g**$)_A$ si et seulement si $\forall$ **h**, **f**A**h** $\succeq$ **g**A**h**. Remarquons que la préférence conditionnelle n'est pas forcément bien définie pour une paire d'actions quelconques (ce n'est le cas que si **Sav 2** est vérifié).

Lemme 1.3.2 *Si* **f** *est faiblement préférée à* **g** *conditionnellement à l'ensemble A et à son complémentaire alors l'axiome* **Sav 1** *est suffisant pour prouver que* **f** *est préférée à* **g**.

Preuve :

Supposons que ($\mathbf{f} \succeq \mathbf{g})_A$, soit $\forall$ **h**, $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ et que ($\mathbf{f} \succeq \mathbf{g})_{\bar{A}}$. On a entre autres $\mathbf{f} \succeq \mathbf{g}A\mathbf{f}$ (en posant **h** = **f**). Or $\mathbf{g}A\mathbf{f}$ = $\mathbf{f}\bar{A}\mathbf{g} \succeq \mathbf{g}$ (en utilisant la préférence conditionnelle sur A, avec **h** = **g**). Alors, par transitivité (axiome **Sav 1**), $\mathbf{f} \succeq \mathbf{g}$.

Lemme 1.3.3 *(Monotonie)*
Si l'ensemble $\mathcal{F} = X^S$ est équipé d'une relation de préférence $\succeq$ respectant **Sav 1** *et* **CFAC** *alors la domination (au sens du chapitre 1 de la partie I)) implique la préférence au sens de $\succeq$:* $\mathbf{f} \geq_P \mathbf{g} \Rightarrow \mathbf{f} \succeq \mathbf{g}$.

Preuve :

$\geq_P$ est le préordre complet sur X obtenu en restreignant $\succeq$ aux actions constantes. Par extension, $\mathbf{f} \geq_P \mathbf{g}$ signifie que **f** domine **g**, i.e. $\forall s \in S, \mathbf{f}(s) \geq_P \mathbf{g}(s)$.

Si $\mathbf{f} \geq_P \mathbf{g}$ on peut trouver une suite d'actions $\{\mathbf{f}_0, \ldots, \mathbf{f}_n\}$ telle que $\mathbf{f}_0 = \mathbf{f}$, $\mathbf{f}_n = \mathbf{g}$ et pour tout $i \in \{1, \ldots, n\} \exists s_i$, $\mathbf{f}_{i-1}(s) = \mathbf{f}_i(s), \forall s \in S, s \neq s_i$ et $\mathbf{f}(s_i) geq_P\ \mathbf{g}(s_i)$.

Par **CFAC**, on a $\mathbf{f}_{i-1} \succeq \mathbf{f}_i$, et par transitivité de $\succeq$ (**Sav 1**), $\mathbf{f} \succeq \mathbf{g}$.

Grâce à ce lemme, le théorème suivant devient facile à démontrer :

Théorème 1.3.1 *Représentation des mesures de Sugeno.*
Si l'ensemble d'actions $\mathcal{F} = X^S$ est équipé d'une relation de préférence $\succeq$ vérifiant **Sav 1**, **CFAC** *et* **Sav 5** *alors la relation d'incertitude déduite de $\succeq$ restreinte aux actions binaires (pour un même couple de conséquences) est une mesure de Sugeno.*

Preuve :

De **Sav 5** on déduit l'existence de $x >_P y$. La relation $\trianglerighteq^{xy}$ est définie par $A \trianglerighteq^{xy} B$ si et seulement si $xAy \succeq xBy$. Cette relation est un préordre complet et peut être projetée sur une échelle finie, simplement ordonnée, L_{xy} dont les éléments sont les classes d'équivalence de $\mathcal{F}_{xy}$: $[xAy]$ représente la classe d'équivalence de xAy. Soit σ la fonction d'ensemble telle que $\sigma(A) = [xAy]$. $\forall A \subseteq B, xBy \geq_P xAy$ et par **Sav 1** et **CFAC**, via le lemme 1.3.3, on montre que $xBy \succeq xAy$, soit $\sigma(B) \geq \sigma(A)$.

Ainsi, σ est une mesure monotone. La dernière propriété, $\sigma(\Omega) > \sigma(\emptyset)$ se déduit de **Sav 5**.

A cet instant se pose la question de savoir si les ensembles $\mathcal{F}_{xy}$ sont cohérents entre eux, i.e. si les préordres induits sur $\mathcal{F}_{xy}$ ou $\mathcal{F}_{x'y'}$ pour deux paires de conséquences $x >_P y$ et $x' >_P y'$ ne se contredisent pas.
Un minimum de cohérence est assuré par l'axiome **Sav 4'**. Grâce à celui-ci on sait que la relation d'ordre $\trianglerighteq^{xy}$ raffine la relation $\trianglerighteq^{x'y'}$ lorsque $x \geq_P x' >_P y' \geq_P y$. De plus le cas de contradiction $xAy \succ xBy$ et $x'Ay' \succ x'By'$ ne pourra jamais être observé : le précédent théorême le montre dans le cas où $A \subseteq B$ et **Sav 4'** suffit à le prouver dans le cas général.

Notons x_* et x^* respectivement le plus petit et plus grand des éléments de $(X, \geq_P)$. Par les axiomes **Sav 1**, **CFAC** et **Sav 5** via le lemme 1.3.3 on prouve que les actions constantes $\mathbf{x}_*$ et $\mathbf{x}^*$ sont respectivement les moins et les plus préférées. Si L représente l'échelle ordonnée isomorphe à l'ensemble des classes d'équivalence de $(\mathcal{F}, \succeq)$ alors $\mathbf{x}_*$ et $\mathbf{x}^*$ correspondent aux bornes 0 et 1 de l'échelle L. Ainsi,

la relation d'incertitude sur $(\mathcal{F}, \succeq)$ la plus raffinée que l'on puisse obtenir l'est à partir de $\mathcal{F}_{x^* x_*}$, que l'on notera dans la suite $\mathcal{F}_{10}$ et dont les éléments sont les actes binaires notés $1A0$, donnant la meilleure conséquence possible si A "arrive" et la pire sinon.
On notera $\trianglerighteq$ la relation de préférence sur les événements, définie par $A \trianglerighteq B$ si et seulement si $1A0 \succeq 1B0$. $\trianglerighteq$ est *la* relation d'incertitude représentant les connaissances d'un agent dont les préférences sur $\mathcal{F}$ satisfont les axiomes **Sav 1**, **CFAC**, **Sav 4'** et **Sav 5**.

Nous allons continuer dans cette voie d'axiomatisation de relations d'incertitudes en termes de préférence entre actions binaires, afin de "capturer" un certain nombre de mesures de Sugeno, autres que les probabilités (voir le chapitre 2 de la partie I).
Pour ce faire nous allons considérer en premier lieu des relaxations du *principe de la chose certaine* de Savage et d'abord le principe *d'indépendance faible* (**IF**) qui évite simplement les renversements de préférence du type **f**A**h** $\succ$ **g**A**h** et **g**A**h'** $\succ$ **f**A**h'**, sans éviter un éventuel effet de noyade en passant de **h** à **h'** sur $\bar{A}$.

Dubois [Dub86] A proposé une relaxation de l'axiome **P** des probabilités qualitatives qui est également satisfait par les possibilités qualitatives :

MD : $\forall A, B, C, A \cap (B \cup C) = \emptyset$ et $B \trianglerighteq C \Rightarrow B \cup A \trianglerighteq C \cup A$.

Ainsi qu'un axiome dual de **MD** qui est satisfait à la fois par les probabilités qualitatives et les nécessités qualitatives :

DMD : $\forall A, B, C, A \cup (B \cap C) = S$ et $B \trianglerighteq C \Rightarrow B \cap A \trianglerighteq C \cap A$.

Chateauneuf [Cha96] a amélioré les résultats de Dubois [Dub86] et prouvé qu'une relation d'incertitude satisfaisant **A1**, **A2**, **A3** et **MD** pouvait être représentée par une *mesure décomposable*, i.e. une fonction d'ensemble à valeurs dans une échelle $L = \sigma(2^S)$ munie d'une opération $\oplus$ satisfaisant les propriétés suivantes :

- $1 \oplus \lambda = 1$,
- $0 \oplus \lambda = \lambda$,
- $\oplus$ est commutative et associative,
- $\sigma(A \cup B) = \sigma(A) \oplus \sigma(B)$ pour toute paire d'événements disjoints A et B.

Ces mesures décomposables ont été introduites par Dubois et Prade [DP82] et Weber [Web84] dans le cas où $\oplus$ est une conorme triangulaire au sens de Schweitzer et Sklar [SS83]. Dans le cas où $\oplus \equiv \max$ on retrouve les mesures de possibilité et dans celui où L est numérique et $a \oplus b = \min(a+b, 1)$ (somme bornée), on retrouve les mesures de probabilité.

Les mesures satisfaisant **A1**, **A2**, **A3** et **DMD** sont appelées *mesures décomposables duales*, et peuvent être représentées par une fonction d'ensemble ρ sur $L = \rho(2^S)$ équippé d'une opération $\odot$ vérifiant :

- $1 \odot \lambda = \lambda$,
- $0 \odot \lambda = 0$,

– $\odot$ est commutative et associative,

– enfin, $\rho(A \cap B) = \rho(A) \odot \rho(B)$ pour toute paire d'événements A et B tels que $A \cup B = S$.

Les mesures décomposables duales sont de la forme $\rho(A) = n_L(\sigma(\bar{A}))$ où n_L est une fonction de renversement d'ordre de L, involutive. On peut choisir pour $\odot$ une norme triangulaire au sens de Schweitzer et Sklar [SS83]. Lorsque $\odot \equiv \min$ on retrouve les mesures de nécessité et lorsque $a \odot b \equiv \max(0, a+b-1)$ (conjonction de Lukasiewicz) on retrouve à nouveau les probabilités. Ainsi les mesures décomposables duales comprennent à la fois les mesure de probabilité et de necessité.

Les mesures décomposables et leurs duales englobent déjà un certain nombre de mesures d'incertitude, mais l'axiome d'indépendance faible délimite une classe de mesures d'incertitude encore plus vaste. Pour définir cette classe il nous faut introduire l'axiome suivant relachant l'axiome des mesures décomposables :

MDF : $\forall A, B, C, A \cap (B \cup C) = \emptyset, B \rhd C \Rightarrow B \cup A \unrhd C \cup A$.

On peut donner une forme différente mais équivalente à **MDF** :
$\forall A, B, C, A \cup (B \cap C) = \emptyset, B \rhd C \Rightarrow B \cap A \unrhd C \cap A$.

Pour prouver l'équivalence des deux formes il suffit de poser $E = \overline{B \cap A}, F = \overline{C \cap A}, G = A$ et considérer la forme contraposée de **MDF**.

Le théorême suivant peut être montré :

Théorème 1.3.2 *Soit $\succeq$, une relation de préférence sur les actions satisfaisant* **Sav 1**, **IF**, **CFAC** *et* **Sav 5**. *La relation d'ordre $\unrhd$ sur les événements, induite par la relation de préférence sur les actions binaires $1A0$ satisfait les axiomes* **A1**, **A2**, **A3** *et* **MDF**.

Preuve :

> Il suffit de prouver **MDF**. Ecrivons **IF** (**f**D**h** $\succ$ **g**D**h** $\Rightarrow$ **f**D**h'** $\succeq$ **g**D**h'**) pour **f** = $1B0$, **g** = $1C0$, $D = A \cup B$ et **h** = $\mathbf{x}_*$. **f**D**h** $\succ$ **g**D**h** s'écrit ainsi $1B0 \succ 1C0$, i.e. $B \rhd C$. Si on pose **h'** = $1A0$ avec $A \cap D = \emptyset$ alors **f**D**h'** $\succeq$ **g**D**h'** s'écrit $B \cup A \unrhd C \cup A$.

Notons que **Sav 4'** est inutile pour prouver ce résultat.

MDF peut sembler très faible à première vue, néanmoins il n'est satisfait ni par les mesures de croyance ni par les mesures de plausibilité de Shafer [Sha76].
Rappelons (voir chapitre 2, partie I) les deux axiomes satisfaits par les fonctions de croyance et de plausibilité proposés par Wong et col. [WYBB91] :

Bel $B \subseteq A, A \cap C = \emptyset \Rightarrow (A \succ B \Rightarrow A \cup C \succ B \cup C)$

Pl $B \subseteq A, B \cup C = S \Rightarrow (A \succ B \Rightarrow A \cap C \succ B \cap C)$

Proposition 1.3.4 *Les relations d'incertitude basées sur les mesures de croyance et de plausibilité ne satisfont pas forcément* **MDF**.

Preuve :

Il suffit d'exhiber un contre exemple. m est une fonction de masse sur S.
Soient $A, B, C, E_1, E_2 \subseteq S$ tels que

- $A \cap (B \cup C) = \emptyset$,
- $E_1 \subset (A \cup B) \cap \bar{C}, E_1 \cap \bar{B} \neq \emptyset$,
- $E_2 \subset (A \cup C) \cap \bar{B}, E_2 \cap \bar{C} \neq \emptyset$,
- $m(B) > m(C) > 0, m(E_1) > 0, m(E_2) > 0$ et $m(E) = 0, \forall E \notin \{B, C, E_1, E_2\}$,
- $m(C) + m(E_2) > m(B) + m(E_1)$.

Alors $Bel(B) = m(B) > Bel(C) = m(C)$ et $Bel(A \cup B) = m(B) + m(E_1) < m(C) + m(E_2) = Bel(A \cup C)$.

MDF n'est donc pas satisfait par la relation d'incertitude basée sur la mesure de croyance Bel.

Le même type de contre exemple peut être trouvée pour une relation d'incertitude basée sur une mesure de plausibilité.

Si nous souhaitons définir les mesures décomposables en termes de préférence entre les actions, **IF** doit être renforcé de la manière suivante :

D : Décomposabilité :
Soient **f**, **g**, **h** et **h'**, éléments de $\mathcal{F}$, soit $A \subseteq S$. Si **f**A**h** $\succeq$ **g**A**h** et **h'** $\geq_P$ **h** alors **f**A**h'** $\succeq$ **g**A**h'**.

Théorème 1.3.3 *Soit $\succeq$ une relation de préférence sur les actions satisfaisant* **Sav 1**, **D**, **CFAC** *et* **Sav 5**. *La relation d'incertitude $\trianglerighteq$ déduite de $\succeq$ satisfait* **MD**.

Preuve :

Il faut prouver **MD**. Soient $B, C, D \subseteq S$ tels que $D \cap (B \cup C) = \emptyset$. **f** = $1B0$, **g** = $1C0$, **f'** = **f** et **g'** = **g** sur $B \cup C$ et **f'** = **g'** = $1D0$ sur $\overline{B \cup C}$. Alors **f'** $\succeq$ **f** sur $\overline{B \cup C}$ et par **D** on obtient **f** $\succeq$ **g** $\Leftrightarrow$ **f**$(B \cup C)0 \succeq$ **g**$(B \cup C)0 \Rightarrow$ **f**$(B \cup C)$**f'** $\succeq$ **g**$(B \cup C)$**g'** $\Leftrightarrow$ $1(D \cup B)0 \succeq 1(D \cup C)0$ (i.e. $B \trianglerighteq C \Rightarrow B \cup D \trianglerighteq C \cup D$).

En changeant l'axiome **D** en son dual **DD**, on s'assure d'obtenir une mesure décomposable duale :

DD Décomposabilité duale.
Soient **f**, **g**, **h** et **h'**, éléments de $\mathcal{F}$, soit $A \subseteq S$. Si **f**A**h** $\succeq$ **g**A**h** et **h** $\geq_P$ **h'** alors **f**A**h'** $\succeq$ **g**A**h'**.

Nous pourrions bâtir une théorie ordinale pour la décision dans l'incertain à partir de mesures décomposables, en nous servant des axiomes **D** et **DD**. Ce n'est pourtant pas l'option que nous avons choisie : nous allons utiliser les axiomes **PES** et **OPT** introduits dans la section précédente, qui sont plus forts.
Dans la section suivante, au contraire, nous exposerons l'axiomatisation que nous avons proposée dans (Dubois et col. [DPS98]) pour une théorie basée sur des mesures de Sugeno, pas forcément décomposables.

Mais d'abord, revenons à notre axiomatisation des utilités qualitatives possibilistes, et en premier lieu, tâchons de récupérer des mesures de possibilité et de nécessité pour représenter l'incertitude. Pour cela les axiomes caractéristiques seront **PES** et **OPT**.

Lemme 1.3.4 *Soit $\succeq$ une relation de préférence entre les actions, vérifiant* **Sav 1**, **CFAC**, **Sav 5** *et* **PES**. *$\succeq$ vérifie la propriété :*

$$\forall \mathbf{f}, \mathbf{h}, \forall A \subseteq S, [(\exists \mathbf{g} \text{ t.q. } \mathbf{g}A\mathbf{f} \succ \mathbf{f}) \Rightarrow \mathbf{f} \succeq \mathbf{f}A\mathbf{h}].$$

Preuve

> **PES** s'écrit : $\forall$ **f**, **g**, $\forall A \subseteq S$ $\mathbf{g}A\mathbf{f} \succ \mathbf{f} \Rightarrow \mathbf{f} \succeq \mathbf{f}A\mathbf{g}$. Maintenant, puisque $1A\mathbf{f} \geq_P \mathbf{g}A\mathbf{f}$, $1A\mathbf{f} \succ \mathbf{f}$. Supposons qu'il existe **h** tel que $\mathbf{f}A\mathbf{h} \succ \mathbf{f}$. Pour la même raison, $\mathbf{f}A1 \succ \mathbf{f}$. Or **PES** interdit d'avoir à la fois $\mathbf{f}A1 \succ \mathbf{f}$ et $1A\mathbf{f} \succ \mathbf{f}$, donc l'hypothèse de départ est fausse et $\mathbf{f} \succeq \mathbf{f}A\mathbf{h}$ pour toute action **h**.

Dubois [Dub86] a proposé les deux axiomes suivants, caractérisant les relations d'incertitude basées respectivement sur des mesures de possibilité ou de nécessité :

- **Pos** $B \trianglerighteq C \Rightarrow B \cup A \trianglerighteq C \cup A$.
- **N** $B \trianglerighteq C \Rightarrow B \cap A \trianglerighteq C \cap A$.

Dans [Dub86], Dubois donne également une forme différente de ces axiomes :

Lemme 1.3.5 *Sous les axiomes* **A1**, **A2**, **A3**, *les axiomes* **Pos** *et* **N** *peuvent être mis sous la forme :*

- **Pos** $\Leftrightarrow [(A \sim A \cup B$ *ou* $B \sim A \cup B)$ *et* $A \subseteq B] \Rightarrow B \trianglerighteq A$.
- **N** $\Leftrightarrow [(A \sim A \cap B$ *ou* $B \sim A \cap B)$ *et* $A \subseteq B] \Rightarrow B \trianglerighteq A$.

Preuve :

> Pour montrer que **Pos** implique la monotonie, on écrit $C \trianglerighteq \emptyset$ et en supposant que $B = A \cup C$ **Pos** donne $C \trianglerighteq \emptyset \Rightarrow B \trianglerighteq A$. Posons $A = B$ dans **Pos**. Celui-ci se lit $B \trianglerighteq C \Rightarrow B \trianglerighteq C \cup B$ mais puisque $\trianglerighteq$ est monotone, $B \sim B \cup C$. Alors on a $C \sim C \cup B$ ou $B \sim C \cup B$ puisque soit $C \trianglerighteq B$ soit $B \trianglerighteq C$. Réciproquement, si $C \sim C \cup B$ ou $B \sim C \cup B$, si on suppose la monotonie et que $B \trianglerighteq C$ et $C \cup A \triangleright B \cup A$, alors, $C \cup A \sim A$ et $B \cup A \sim B$, sinon on obtiendrait une contradiction. Finalement, $A \triangleright B$ et $A \sim A \cup B$. Donc, $A \triangleright B \trianglerighteq C$ et $C \cup A \sim A$. En fin de compte $C \cup A \triangleright B \cup A$ implique $A \triangleright A$, une contradiction !
>
> Un même raisonnement peut être tenu en ce qui concerne **N**.

De ce lemme on déduit qu'une fonction d'ensemble satisfaisant **Pos** est une mesure de possibilité ($\sigma(A \cup B) = \max(\sigma(A), \sigma(B))$), et qu'une fonction d'ensemble satisfaisant **N** est une mesure de nécessité ($\sigma(A \cap B) = \min(\sigma(A), \sigma(B))$).

Finalement on prouve le théorème suivant :

Théorème 1.3.4 *(Représentation des mesures de nécessité) Si $\mathcal{F} = X^S$ est équipé d'une relation de préférence $\succeq$ satisfaisant* **Sav 1**, **CFAC**, **Sav 5** *et l'axiome de pessimisme* **PES** *alors la relation d'incertitude induite par restriction de $\succeq$ aux actions binaires (à conséquences fixées) est représentable par une mesure de nécessité uniquement.*

Preuve :

Posons **f** = $1B0$, **g** = $1C0$ et **h** = $1D0$. Alors **g**A**f** $\succ$ **f** correspond à $(A\cap C)\cup(\overline{A}\cap B)\rhd B$, et **f** $\succeq$ **f**A**h** s'écrit $B \unrhd (A\cap B)\cup(\overline{A}\cap D)$, pour tout D. En particulier, en posant $C = A$ et $D = \overline{A}$, le premier terme s'écrit $A\cup B \rhd B$ et le second $B \unrhd \overline{A}\cup B$. Par les lemmes précédents, l'axiome de pessimisme implique : $A\cup B \rhd B \Rightarrow B \unrhd \overline{A}\cup B$.

En posant $E = A\cup B$ et $F = \overline{A}\cup B$, $B = E\cap F$ et la propriété s'écrit: $E \rhd E\cap F \Rightarrow E\cap F \unrhd F$. Puisque $\unrhd$ est monotone, $F \unrhd E\cap F$ et donc, soit $F \sim E\cap F$ ou $E\cap F \sim F$. Cette dernière propriété, accompagnée de la monotonie est équivalente à **N**, l'axiome caractéristique des mesures de nécessité.

Évidemment, un raisonnement similaire peut être tenu pour montrer que la mesure d'incertitude correspondant à une relation de préférence sur les actions satisfaisant **Sav 1**, **CFAC**, **Sav 5** et **OPT** est une mesure de possibilité.

Théorème de représentation pour les critères d'utilité qualitative possibiliste

Finalement, nous proposons deux théorèmes de représentation, respectivement pour les utilités qualitatives pessimistes et optimistes. Pour cette représentation nous allons voir que l'on doit ajouter un des axiomes **DDR** ou **DCR**, qui permettent d'imposer une semi-décomposabilité des fonctions d'utilité, et une forme $minmax$ de leur expression.

Théorème 1.3.5 représentation de l'utilité qualitative pessimiste *:*
Soit $\succeq$ une relation de préférence sur $\mathcal{F}$, l'ensemble des actions **f** *de S vers X, satisfaisant* **Sav 1**, **CFAC**, **Sav 5**, **PES** *et* **DDR**. *Alors il existe une échelle qualitative L, une fonction d'utilité μ de X vers L et une distribution de possibilité π sur S, à valeurs dans L et une fonction v_* à valeurs dans L telle que :* **f** $\succeq$ **f'** $\Leftrightarrow v_*(\mathbf{f}) \geq v_*(\mathbf{f'})$.
De plus, on peut choisir v_* de la forme $v_*(\mathbf{f}) = \min_{s\in S}\max(n(\pi(s)),\mu(\mathbf{f}(s)))$ sur X, où n est la fonction de renversement de L.

Pour prouver ce théorème plus facilement, on commence par prouver les lemmes suivants :

Lemme 1.3.6 *Sous* **Sav 1**, **CFAC**, **Sav 5** *et* **PES**, *si* **h** =**f** $\wedge$ **g**, *alors* **h** $\sim$ **f** *ou* **h** $\sim$ **g**.

Preuve :

Soit $A = \{s, \mathbf{f}(s) >_P \mathbf{g}(s)\}$. Alors **h** = **g**$A$**f** = **g**$A$**h** = **h**$A$**f**. **g** $\geq_P$ **h** et **f** $\geq_P$ **h**, donc **f** $\succeq$ **h** et **g** $\succeq$ **h**. Supposons qu'à la fois **f** $\succ$ **h** et **g** $\succ$ **h**. Alors **f**A**h** $\succ$ **h** et **h**A**g** $\succ$ **h**, ce qui est impossible, d'après le lemme 1.3.4.

Lemme 1.3.7 *Sous* **Sav 1**, **CFAC**, **Sav 5**, *et* **DDR** *si* **h** =**f** $\vee$ **x**, *où* **x** *est l'action constante de conséquence x, alors* **h** $\sim$ **f** *ou* **h** $\sim$ **x**.

Preuve :

L'axiome **DDR** indique que **f** $\succ$ **g** et **f** $\succ$ **x** impliquent **f** $\succ$ **g**$\vee$ **x** . Mais, par les autres axiomes, **h** = **f**$\vee$ **x** $\succeq$ **f** et **h** $\succeq$ **x** (domination). Supposons à la fois **h** $\succ$ **f** et **h** $\succ$ **x**. Alors, par **DDR**, **h** $\succ$ **h**, ce qui est impossible. Donc **h** $\sim$ **f** ou **h** $\sim$ **x**.

Nous pouvons maintenant prouver notre théorème. Nous avons déjà montré que l'utilité possibiliste pessimiste satisfait les axiomes **Sav 1**, **CFAC**, **Sav 5**, **DDR** et **PES**. L'autre direction de la preuve du théorème , i.e. qu'une relation de préférence satisfaisant ces quatre axiomes est représentable par une fonction d'utilité qualitative pessimiste, se décompose de quatre étapes :

1. Construction de l'échelle d'utilité.
   De **Sav 1** on déduit que $(\mathcal{F} = X^S, \succeq)$ est un préordre complet. Puisque X et S sont finis, on peut le décomposer en un ensemble de classes d'équivalence $\mathcal{F}/\sim$, que l'on peut projeter (bijectivement) sur une échelle finie, simplement ordonnée, L, de plus petit élément 0 et de plus grand, 1, appelée *échelle d'utilité*. Pour chaque action **f**, l'image de la classe d'équivalence [**f**] dans L est appelé utilité de **f** et notée $v_*(\mathbf{f})$.

   Considérons maintenant les actions constantes **x** (valant x partout). On définit la fonction d'utilité μ sur X par $\mu(x) = v_*(\mathbf{x})$. Grâce à la propriété de domination, $\mu(x_*) = 0$ et $\mu(x^*) = 1$.

2. Construction de la distribution de possibilité qualitative.
   Considérons maintenant la relation d'incertitude $\trianglerighteq$ induite par la restriction de $\succeq$ sur $\mathcal{F}_{10}$ l'ensemble des actions binaires de la forme $1A0$. Depuis le paragraphe précédent on sait qu'il s'agit d'une nécessité qualitative. Alors l'utilité $v_*(1A0)$ de tels actes définit une mesure de nécessité N lorsque A varie, telle que $N(A) = v_*(1A0)$. Soit n la fonction de renversement d'ordre de L. La fonction π de S vers L définie par $\pi(s) = n(v_*(1(S \setminus \{s\})0))$ est la distribution de possibilité associée à N (telle que $N(A) = inf_{s \in \overline{A}} n(\pi(s)))$.

3. Utilité des actions binaires de la forme xAy
   Considérons une action $1Ax$. On peut l'écrire comme une disjonction $1A0 \vee$ **x**. Du lemme 1.3.7, $v_*(1Ax) = N(A)$ ou $\mu(x)$. Mais les relations de domination $1Ax \geq_P$ **x** et $1Ax \geq_P 1A0$ impliquent $v_*(1Ax) \geq \max(N(A), \mu(x))$. Donc $v_*(1Ax) = \max(N(A), \mu(x))$. Maintenant, toute action xAy avec $x \geq_P y$ est de la forme $1Ay \wedge$ **x**. Utilisant le lemme 1.3.6, et un raisonnement similaire au précédent, on prouve facilement que l'utilité est min-décomposable et que $v_*(xAy) = \min(v_*(1Ay), \mu(x)) = \min(\max(N(A), \mu(y)), \mu(x))$, et plus généralement, $v_*(\mathbf{f} \wedge \mathbf{g}) = \min(v_*(\mathbf{f}), v_*(\mathbf{g}))$.

4. Utilité d'une action quelconque.
   Finalement, on étend la fonction v_* à tout l'ensemble des actions, X^S, et on prouve que $v_*(\mathbf{f}) = \min_{s \in S} \max(n(\pi(s)), \mu(\mathbf{f}(s)))$.

   Toute action peut être vue comme une conjonction **f**$= \wedge_{s \in S} 1(S \setminus \{s\})\mathbf{f}(s)$. D'après le point précédent, $v_*(1(S \setminus \{s\})\mathbf{f}(s)) = \max(N(S \setminus \{s\}), \mu(\mathbf{f}(s))) = \max(n(\pi(s)), \mu(\mathbf{f}(s)))$. Il ne reste plus qu'à appliquer la décomposabilité pour le min pour obtenir le résultat.

Nous pouvons prouver un théorème correspondant pour la représentation de l'utilité qualitative possibiliste optimiste :

Théorème 1.3.6 représentation de l'utilité qualitative optimiste.
Soit $\succeq$ une relation de préférence sur $\mathcal{F}$, satisfaisant **Sav 1**, **CFAC**, **Sav 5**, **OPT** *et* **DCR**. *Alors il existe une échelle qualitative finie L, une fonction d'utilité μ de X vers L et une distribution de possibilité π sur S, également à valeurs dans L, et une fonction d'utilité v^* à valeurs dans L telle que :* **f** $\succeq$ **f'**$\Leftrightarrow v^*(\mathbf{f}) \geq v^*(\mathbf{f'})$. *De plus, v^* peut être choisie sous la forme* $v^*(\mathbf{f}) = \max_{s \in S} \min(\pi(s), \mu(\mathbf{f}(s)))$.

Preuve :

Les seules différences avec la preuve précédente sont les suivantes :

– Construction d'une distribution de possibilité qualitative sur S.
La relation d'incertitude $\trianglerighteq$ induite par la restriction de $\succeq$ à $\mathcal{F}_{10}$, l'ensemble des actions binaires $1A0$ est une relation de possibilité (grâce à **OPT**). Ensuite, l'utilité $v^*(1A0)$ de telles actions définit une mesure de possibilité Π lorsque A varie, telle que $\Pi(A) = v^*(1A0)$. Alors, la fonction π de S vers L définie par $\pi(s) = v^*(1s0)$ est la distribution de possibilité associée à Π (telle que $\Pi(A) = max_{s \in A}\pi(s))$.

– Utilité des actions binaires de la forme $xA0$.
Considérons une action $xA0$. Elle s'écrit comme une conjonction $1A0\wedge$ **x**. D'après **DCR**, $v^*(xA0) = \Pi(A)$ ou $\mu(x)$. Les relations de domination $xA0 \leq_P$ **x** et $xA0 \leq_P 1A0$ impliquent $v^*(xA0) \leq \min(\Pi(A), \mu(x))$. Donc $v^*(xA0) = \min(\Pi(A), \mu(x))$. En utilisant **OPT** et un raisonnement similaire, on prouve facilement que l'utilité optimiste est décomposable pour le max, et que $v^*(\mathbf{f} \vee \mathbf{g}) = \max(v^*(\mathbf{f}), v^*(\mathbf{g}))$.

– Toute action peut être vue comme une disjonction $\mathbf{f} = \vee_{s \in S}\mathbf{f}(s)\{s\}0$. Du point précédent on déduit que $v^*(\mathbf{f}(s)\{s\}0) = \min(\pi(s), \mu(\mathbf{f}(s)))$. Il nous suffit alors d'appliquer la décomposabilité pour le max pour obtenir le résultat.

Ce dernier théorème a déjà été démontré par Campos et Bolaños [dB92], mais dans une forme différente, non basée sur la notion de préférence entre actions.

Ce type de justification axiomatique peut être généralisé à un critère étendant à la fois les utilités possibilistes optimistes et pessimistes. Le critère en question est une *intégrale de Sugeno* [Sug77]. Dans la section suivante, nous allons reprendre nos résultats proposés dans [DPS98], qui constituent une justification axiomatique de l'utilisation de l'intégrale de Sugeno pour la décision dans l'incertain. Les résultats exposés reposent largement sur la semi-décomposabilité de l'intégrale de Sugeno pour deux actions dont l'une est une action constante, ou plus généralement pour deux actions *comonotones*.
L'axiomatisation proposée est proche de celles proposées par de Campos et col. [dLM91], [dB92] ou Ralescu et Sugeno [RS96], mais contrairement à ces dernières, elle se place dans un cadre *à la Savage* de décision dans l'incertain.

1.4 Décision dans l'incertain et intégrale de Sugeno : Utilité Qualitative Monotone

Dans cette section nous allons élargir le cadre possibiliste dans lequel nous nous sommes places jusqu'ici et nous allons exposer l'axiomatisation proposée dans [DPS98]. Nous allons utiliser une mesure de l'incertitude très générale puisque simplement monotone pour l'inclusion et un critère correspondant pour la décision dans l'incertain, qui est une intégrale de Sugeno.

L'intégrale de Sugeno est une contrepartie qualitative de l'intégrale de Choquet qui ne requiert qu'une échelle simplement ordonnée pour sa définition. Alors que la théorie classique de la décision utilise la notion de moyenne pour évaluer les actions, l'intégrale de Sugeno est une *médiane*, qui peut être vue comme contrepartie qualitative de cette notion.

L'analogie entre l'intégrale de Choquet et l'intégrale de Sugeno est encore plus poussée puisque que les deux présentent un comportement similaire en face d'une certaine catégorie de fonctions (donc d'actions) : les *actions comonotones* (deux actions **f** et **g** sont dites comonotones si pour toute paire d'états (s, s'), $\mathbf{f}(s) >_P \mathbf{f}(s') \Rightarrow \mathbf{g}(s) \geq_P \mathbf{g}(s')$ et $\mathbf{f}(s) <_P \mathbf{f}(s') \Rightarrow \mathbf{g}(s) \leq_P \mathbf{g}(s')$).
Alors que l'intégrale de Choquet est additive pour deux actions comonotones, l'intégrale de Sugeno est min et max-décomposable.

Avant toute chose, nous allons exposer le critère de l'utilité qualitative monotone et certaines de ses propriétés dans le paragraphe suivant.

1.4.1 Utilité qualitative monotone d'une action

Considérons à nouveau l'ensemble X^S des actions potentielles, fonctions de S vers X. Comme nous l'avons déjà fait remarquer, si nous nous plaçons dans un cadre *à la* Savage, il est normal de supposer l'existence d'un échelle simplement ordonnée $(L, \geq)$, unique et finie puisque nous supposons S et X finis, pour représenter à la fois les préférences et l'incertitude d'un agent. Une telle échelle peut être construite à partir des classes d'équivalence (pour $\succeq$) de l'ensemble constitué des actions constantes et des actions binaires (à conséquences extrêmes).
$\mu : X \to L$ est la fonction d'utilité sur X, associant à chaque conséquence x la classe d'équivalence de la fonction constante associée. On supposera dans ce paragraphe que les bornes 0_L et 1_L de l'échelle L appartiennent à l'ensemble $\mu(X) = \{\mu(x), x \in X\}$. Si ce n'est pas le cas pour un problème donné, il suffit d'ajouter deux conséquences hypothétiques (pas forcément atteintes dans le problème), x_* et x^* telles que $\mu(x_*) = 0_L$ et $\mu(x^*) = 1_L$.
L'incertitude sera capturée par une fonction d'ensemble $\sigma : S \to L$, qui est une *mesure monotone* (encore appelée *mesure floue* par Sugeno). Ce type de fonction d'ensemble déjà rencontré plus haut est très général et constitue les conditions minimales pour représenter les connaissances incertaines. La condition de monotonie, en particulier, est vérifiée par la plupart des mesures d'incertitude les plus répandues (probabilités, fonctions de croyance, possibilités, etc.).

L'utilité d'une action **f** peut être définie par une intégrale de Sugeno [Sug77], contrepartie de l'utilité espérée dans laquelle la somme est remplacée par un sup (un max dans le cas fini) et le produit par un min.

Définition 1.4.1 *Utilité qualitative monotone d'une action.*

$$u_S(\mathbf{f}) = \max_{\lambda \in L} \min(\lambda, \sigma(F_\lambda))$$

où $F_\lambda = \{s \in S, \mu(\mathbf{f}(s)) \geq \lambda\}$ *et* σ *est une mesure monotone.*

Cette intégrale de Sugeno est appelée *utilité qualitative monotone*. u_S peut être réécrite en faisant varier la conséquence $x \in X$ plutôt que le niveau λ dans l'échelle L :

Proposition 1.4.1 $u_S(\mathbf{f}) = \max_{x \in X} \min(\mu(x), \sigma(F_x))$ *où* $F_x = \{s \in S, \mu(\mathbf{f}(s)) \geq_P \mu(x)\}$

Preuve :

$\max_{x \in X} \min(\mu(x), \sigma(F_x)) = \max_{\lambda \in \mu(X)} \min(\lambda, \sigma(F_\lambda))$.
Donc $\max_{x \in X} \min(\mu(x), \sigma(F_x)) \leq \max_{\lambda \in L} \min(\lambda, \sigma(F_\lambda))$. Soit x^λ tel que $\mu(x^\lambda) =$

$\min\{\mu(x), x \in \mathbf{f}(F_\lambda)\}$. x^λ est la pire conséquence d'une action **f** parmi celles dont l'utilité est au moins λ. Cette conséquence existe forcément puisque le plus grand élément de L est l'utilité de la meilleure conséquence x^*. Il est clair que $F_\lambda = F_{x^\lambda}$; de plus $\mu(x^\lambda) \geq \lambda$. Donc $\forall \lambda \in L, \exists x \in X$, tel que $F_\lambda = F_x$ et $\mu(x) \geq \lambda$. Finalement $\max_{\lambda \in L} \min(\lambda, \sigma(F_\lambda)) \leq \max_{x \in X} \min(\mu(x), \sigma(F_x))$.

Regardons les propriétés de l'utilité qualitative monotone. En premier lieu, l'utilité d'une action constante **x** est $\mu(x)$. Plus généralement, pour les actions binaires xAy, la propriété suivante est vérifiée :

Proposition 1.4.2 *Si $\mu(x) \geq \mu(y)$, alors $u_S(xAy) = \max\{\mu(y), \min(\mu(x), \sigma(A))\}$.*

Preuve :

Notons que si $\mu(x) \geq \mu(y)$, alors $\forall z \in X$ tel que $\mu(z) \leq \mu(y), F_z = S$, $\forall z$ tel que $\mu(y) < \mu(z) \leq \mu(x), F_z = A$, $\forall z, \mu(z) > \mu(x)$ tel que $F_z = \emptyset$.
Alors, de manière évidente : $u_S(xAy) = \max\{min(1_L, \mu(y)), \min(\mu(x), \sigma(A))\}$.

Par application de cette proposition, $u_S(x^*Ax_*) = \sigma(A)$. De plus, $u_S(xAy)$ est la médiane de l'ensemble $\{\mu(y), \mu(x), \sigma(A)\}$ si $\mu(x) \geq \mu(y)$, et la médiane de $\{\mu(y), \mu(x), \sigma(\overline{A})\}$ sinon.
Pour comprendre l'intuition qui est derrière cette expression, observons à nouveau le cas d'un conducteur devant se rendre à une réunion et devant choisir une route. Encore une fois x signifie "arriver avec un peu de retard", y signifie "arriver avec un retard important" et A signifie "il n'y a pas d'embouteillages". Choisir l'utilité de Sugeno $u_S(xAy)$ pour évaluer ce type d'action revient à supposer que si l'agent pense qu'il n'y aura pas d'embouteillages ($\sigma(A)$ est assez grand) alors il pense qu'il aura peu de retard ($u_S(xAy) = \mu(x)$). S'il n'a pas vraiment d'opinion ($\sigma(A)$ moyen), l'utilité de l'action reflète son incertitude ($u_S(xAy) = \sigma(A)$). Enfin, s'il ne pense pas que la route sera libre , $u_S(xAy) = \mu(y)$, il agira comme si il y avait des embouteillages.

Notons que dans le cas des utilités qualitatives possibiliste, le caractère pessimiste ou optimiste du décideur est représenté par le choix de la mesure σ. L'utilisation de l'intégrale de Sugeno permet, par le choix de la fonction σ, de modéliser toute une gamme d'attitudes, de la plus pessimiste à la plus optimiste. En choisissant une mesure de nécessité ($\sigma = \mathsf{N}$) on modélise un comportement pessimiste et en choisissant une mesure de possibilité ($\sigma = \Pi$) on modélise un comportement optimiste.

Remarquons également que l'utilité qualitative monotone se comporte comme une médiane, ce qui est satisfaisant car cela met en lumière son analogie avec l'utilité espérée qui se comporte comme une moyenne.

Proposition 1.4.3
Si X a n éléments $\{x_1 = x_*, \ldots, x_n = x^*\}$ avec $\mu(x_1) \leq \mu(x_2) \leq \mu(x_{n-1}) \leq \mu(x_n)$, alors $u_S(\mathbf{f})$ est la médiane des $2n-1$ nombres $\{\sigma(F_x), x \in X, x \neq x_*\} \cup \mu(X)$.

Preuve: Voir (Dubois et Prade [DP80]).

D'après la Proposition 1.4.1 page précédente, $u_S(\mathbf{f}) = \max_{x \in X} \min(\mu(x), \sigma(F_x))$ où $F_x = \{s \in S, \mu(\mathbf{f}(s)) \geq_P \mu(x)\}$. Si $X = \{x_1 = x_*, \ldots, x_n = x^*\}$ avec $\mu(x_1) \leq \mu(x_2) \leq \mu(x_{n-1}) \leq \mu(x_n)$, $u_S(\mathbf{f}) = \max_{i \in 1 \ldots n} \min(\mu(x_i), \sigma(F_i))$ où $F_i = \mathbf{f}^{-1}(\{x_i, \ldots, x_n\})$.
Soit i_0 tel que $u_S(\mathbf{f}) = \min(\mu(x_{i_0}), \sigma(F_{i_0}))$. On peut noter que les $n-1$ termes $\{\sigma(F_i), i >$

$i_0\} \cup \{\mu(x_i), i < i_0\}$ sont tous inférieurs ou égaux à $u_S(\mathbf{f})$ (croissance / décroissance de $\mu(x_i)$ / $\sigma(F_i)$). De même, les $n-1$ termes $\{\sigma(F_i), i < i_0\} \cup \{\mu(x_i), i > i_0\}$ sont tous supérieurs ou égaux à $u_S(\mathbf{f})$.

On en déduit que $u_S(\mathbf{f})$ est bien la médiane des $2 \cdot n - 1$ termes $\{\sigma(F_i), i \in 2, \ldots, n\} \cup \{\mu(x_i), i \in 1, \ldots, n\}$.

Voyons maintenant, parmi les axiomes de Savage et ceux présentés pour justifier les fonctions d'utilité qualitative possibilistes, lesquels sont satisfaits par l'utilité qualitative monotone :

– **Sav 1** est respecté puisque u_S est une application de X^S vers une échelle complètement ordonnée, $(L, \leq)$.

– L'utilité qualitative monotone ne respecte pas le principe de la chose certaine, **Sav 2**. Elle ne respecte même pas les versions affaiblies comme **PES**, **OPT** ou **Grant** : il est possible, comme le montre l'exemple suivant, de trouver **f**, **g**, **h** et **h'** tels que $\mathbf{f}A\mathbf{h} \prec \mathbf{g}A\mathbf{h}$ et $\mathbf{f}A\mathbf{h'} \succ \mathbf{g}A\mathbf{h'}$.

  Exemple 1.4.1 *Il est possible d'avoir une mesure de Sugeno σ telle que, pour trois ensembles A, B et C, on ait à la fois $\sigma(A) > \sigma(B)$ et $\sigma(A \cup C) < \sigma(B \cup C)$, car ceci ne contredit aucun des axiomes caractéristiques des mesures de Sugeno (sauf si $B \subseteq A$).*

  Prenons par exemple B, C, D et E quatre sous-ensembles disjoints de S. Choisissons une mesure de Sugeno σ telle que $\sigma(B \cup D) < \sigma(C \cup D)$ et $\sigma(B \cup E) > \sigma(C \cup E)$. Rien n'empêche de trouver une telle mesure puisque ni $C \cup D \subseteq B \cup D$, ni $B \cup E \subseteq C \cup E$.

  Définissons ensuite $\mathbf{f} = 1B0$, $\mathbf{g} = 1C0$, $\mathbf{h} = 1D0$ *et* $\mathbf{h'} = 1E0$. *Posons enfin $A = B \cup C$. Il est facile de vérifier que $u_S(\mathbf{f}A\mathbf{h}) = \sigma(B \cup D)$, $u_S(\mathbf{g}A\mathbf{h}) = \sigma(C \cup D)$, $u_S(\mathbf{f}A\mathbf{h'}) = \sigma(B \cup E)$, et $u_S(\mathbf{g}A\mathbf{h'}) = \sigma(C \cup E)$.*

  Nous obtenons donc $\mathbf{f}A\mathbf{h} \prec \mathbf{g}A\mathbf{h}$ *et* $\mathbf{f}A\mathbf{h'} \succ \mathbf{g}A\mathbf{h'}$.

– L'utilité qualitative monotone ne respecte pas non plus **Sav 3** et **Sav 4**. Le même exemple convient pour le prouver. Ces violations correspondent au fait que la préférence conditionnelle ne peut être définie de la même manière que chez Savage, sans le principe de la chose certaine.

  Néanmoins, l'utilité qualitative monotone satisfait le *principe de cohérence faible avec les actes constants* (**CFAC**) que nous rappelons :

  CFAC *principe de cohérence faible avec les actes constants* :
  Soient les deux conséquences x et y. $x \geq_P y \Rightarrow x A\mathbf{h} \succeq y A\mathbf{h}$, $\forall A \subseteq S$ et toute action **h**.

  Pour le prouver, nous avons besoin du lemme suivant, stipulant que les ordres de Pareto-domination et d'utilité qualitative monotone sont cohérents.

  D'abord, posons $X = \{x_0 <_P \ldots <_P x_n\}$ l'ensemble des conséquences possibles, **f** est une action, et $F_n \subseteq \ldots \subseteq F_1 \subseteq F_0 = S$ sont les sous-ensembles de S tels que $F_i = \{s \in S, \mathbf{f}(s) \geq x_i\}$.

  Lemme 1.4.1 $\mathbf{f} \geq_P \mathbf{g} \Rightarrow u_S(\mathbf{f}) \geq u_S(\mathbf{g})$

  Preuve :

  $\mathbf{f} \geq_P \mathbf{g} \Rightarrow G_i \subseteq F_i, \forall i \Rightarrow \sigma(F_i) \geq \sigma(G_i), \forall i$. De plus $u_S(\mathbf{f})$ est non-décroissante en $\sigma(A)$ et $\mu(x)$.

Ce résultat est déjà connu : il montre que l'intégrale de Sugeno est monotone vis-à-vis de l'inclusion d'ensembles flous. Maintenant, d'après **CFAC**, si $x \leq y$, alors $\forall A \subseteq S, \forall$ **h**, **x**A**h** $\leq_P$ **y**A**h**. Par le lemme précédent, nous obtenons $u_S($**x**A**h**$) \leq u_S($**y**A**h**), donc **CFAC** est satisfait.

Répétons que le conditionnement par rapport à un événement ne permet plus, comme chez Savage, de définir un préordre complet sur les actions, puisque **Sav 2** et **Sav 3** ne sont pas respectés. Il n'est même pas possible d'en définir une forme restreinte aux actions à conséquences extrêmes comme dans le cas possibiliste, puisque comme le montre l'exemple précédent, modifier de manière identique les conséquences communes de deux actions peut inverser l'ordre de préférence entre ces deux actions.

Tout comme les utilités qualitatives possibilistes pessimiste et optimiste, l'utilité qualitative monotone u_S ne respecte pas **Sav 4**, mais respecte sa version affaiblie que nous rappelons :

Sav 4' *Cohérence faible entre actions*.
Soient $x > x', y > y'; A, B \subseteq S : xAx' \succ xBx' \Rightarrow yAy' \succeq yBy'$. Si on a en plus $x' \leq y' < y \leq x$, alors : $xAx' \succeq xBx' \Rightarrow yAy' \succeq yBy'$.

Proposition 1.4.4 u_S *respecte* **Sav 4'**.

Preuve :

- $u_S(xAx') = \max(\mu(x'), \min(\sigma(A), \mu(x)))$. Donc, $xAx' \prec xBx' \Rightarrow \sigma(A) < \sigma(B) \Rightarrow yAy' \preceq yBy'$.
- Pour prouver la deuxième partie on va montrer que si $x' \leq y' < y \leq x$, alors $xAx' \sim xBx' \Rightarrow yAy' \sim yBy'$.
  Tout d'abord, remarquons que si $\sigma(A) \leq \mu(x')$, $u_S(xAx') = \mu(x')$, si $\mu(x') < \sigma(A) \leq \mu(x)$, $u_S(xAx') = \sigma(A)$ et si $\sigma(A) \geq \mu(x)$, $u_S(xAx') = \mu(x)$.
  Alors, $xAx' \sim xBx' \Leftrightarrow u_S(xAx') = u_S(xBx')$, ce qui impose les conditions suivantes : soit $\sigma(A) \leq \mu(x')$ et $\sigma(B) \leq \mu(x')$, soit $\sigma(A) \geq \mu(x)$ et $\sigma(B) \geq \mu(x)$, soit enfin $\sigma(A) = \sigma(B)$.
  Il est alors facile de montrer que $u_S(yAy') = u_S(yBy')$, puisque $\mu(y') \leq \sigma(A) \leq \mu(y) \Rightarrow \mu(x') \leq \sigma(A) \leq \mu(x)$ et $\mu(y') \leq \sigma(B) \leq \mu(y) \Rightarrow \mu(x') \leq \sigma(B) \leq \mu(x)$.

– Nous allons voir que l'utilité qualitative monotone satisfait les deux axiomes de domination conjonctive ou disjonctive restreinte.

On déduit tout d'abord aisément du lemme 1.4.1, que $u_S(\mathbf{f} \wedge \mathbf{g}) \leq min(u_S(\mathbf{f}), u_S(\mathbf{g}))$ et $u_S(\mathbf{f} \vee \mathbf{g}) \geq max(u_S(\mathbf{f}), u_S(\mathbf{g}))$, puisque **f** $\wedge$ **g** $\leq_P$ **f** et **g**, et **f** $\vee$ **g** $\geq_P$ **f** et **g**.

Ensuite on prouve une forme restreinte de décomposabilité de l'utilité qualitative monotone :

Lemme 1.4.2 *Soit* **f**, *une action et* **y** *une action constante de valeur* y. *Alors* $u_S(\mathbf{f} \wedge \mathbf{y}) = u_S(\mathbf{f})$ *ou* $\mu(y)$.

Preuve :

Soit $FY_x = \{s \in S,\ u_S(\mathbf{f} \wedge \mathbf{y})(s) \geq \mu(x)\}$. Il est facile de voir que puisque $\forall s \in S$, $\mathbf{f} \wedge \mathbf{y}(s) \leq y$, on a $\forall x$ tel que $\mu(x) > \mu(y), FY_x = \emptyset$ alors que $\forall x$ tel que $\mu(x) \leq \mu(y), FY_x = F_x = \{s \in S,\ u_S(\mathbf{f}\ (s)) \geq \mu(x)\}$.
Donc $u_S(\mathbf{f} \wedge \mathbf{y}) = \max_{\mu(x) \leq \mu(y)} \min(\mu(x), \sigma(F_x))$
$= \max(\min(\mu(y), \sigma(F_y)), max_{\mu(x) \leq \mu(y), x \neq y} \min(\mu(x), \sigma(F_x))$.
Deux cas doivent être envisagés :

1. $\mu(y) > \sigma(F_y)$
   Notons alors que si $\mu(x) > \mu(y)$ on a $F_x \subseteq F_y$ et $\sigma(F_x) \leq \sigma(F_y)$. Alors, $\forall x$ tel que $\mu(x) > \mu(y), \min(\mu(x), \sigma(F_x)) \leq \sigma(F_y) = \min(\mu(y), \sigma(F_y))$. Donc $u_S(\mathbf{f} \wedge \mathbf{y}) = u_S(\mathbf{f})$ puisque le terme $max_{\mu(x) > \mu(y), x \neq y} \min(\mu(x), \sigma(F_x))$ ne détermine pas la valeur de $u_S(\mathbf{f})$.
2. $\mu(y) \leq \sigma(F_y)$.
   Alors, puisque $u_S(\mathbf{f} \wedge \mathbf{y}) \geq \min(\mu(y), \sigma(F_y))$
   $= \mu(y) \geq max_{\mu(x) \leq \mu(y), x \neq y} \min(\mu(x), \sigma(F_x))$, on obtient $u_S(\mathbf{f} \wedge \mathbf{y}) \geq u_S(\mathbf{y})$.
   Cependant, puisque $u_S(\mathbf{y}) = \mu(y) \geq \mu((\ \mathbf{f} \wedge \mathbf{y})(s)), \forall s \in S$ (relation de domination), $u_S(\mathbf{y}) \geq u_S(\mathbf{f} \wedge \mathbf{y})$. Donc $u_S(\mathbf{y}) = u_S(\mathbf{f} \wedge \mathbf{y})$.

Cette propriété indique qu'en faisant décroître progressivement l'utilité des meilleures conséquences (en majorant la fonction d'utilité par une borne constante), dans un premier temps on n'affecte pas l'utilité de l'action du point de vue du décideur, jusqu'au moment où celui-ci commence à négliger l'incertitude pesant sur l'action, se focalisant entièrement sur cette borne supérieure.

De la même manière on prouve une forme duale de décomposabilité restreinte :

Lemme 1.4.3 *Soit* **f** *une action et* **y** *une action constante de valeur* y*. Alors,* $u_S(\mathbf{f} \vee \mathbf{y}) = u_S(\mathbf{f})$ *ou* $\mu(y)$

Preuve :

Du Lemme 1.4.1 nous déduisons que $u_S(\mathbf{f} \vee \mathbf{y}) \geq u_S(\mathbf{f})$ et $u_S(\mathbf{f} \vee \mathbf{y}) \geq \mu(y)$. Puis, $u_S(\mathbf{f} \vee \mathbf{y}) = \max_{x \in X} \min(\mu(x), \sigma(FY_x))$, mais $\forall x \preceq y, FY_x = S$, donc $\sigma(FY_x) = 1$. Finalement, $u_S(\mathbf{f} \vee \mathbf{y}) = \max(\mu(y), \max_{x \succ y} \min(\mu(x), \sigma(F_x)))$.
Alors, si $u_S(\mathbf{f} \vee \mathbf{y}) > u_S(\mathbf{f})$, puisque $\max_{x \succ y} \min(\mu(x), \sigma(F_x)) \leq u_S(\mathbf{f})$ nous obtenons $u_S(\mathbf{f} \vee \mathbf{y}) = \mu(y)$.

Enfin, grâce à ces deux lemmes on peut prouver que l'utilité qualitative monotone satisfait les axiomes **DCR** et **DDR** dont on rappelle les expressions :

DCR Soient **f** et **g** deux actions quelconques et **y** une action constante de valeur y:
$\mathbf{g} \succ \mathbf{f}$ et $\mathbf{y} \succ \mathbf{f} \Rightarrow \mathbf{g} \wedge \mathbf{y} \succ \mathbf{f}$,

DDR Soient **f** et **g** deux actions quelconques et **y** une action constante de valeur y:
$\mathbf{f} \succ \mathbf{g}$ et $\mathbf{f} \succ \mathbf{y} \Rightarrow \mathbf{f} \succ \mathbf{g} \vee \mathbf{y}$

Proposition 1.4.5 u_S *satisfait* **DCR**.

Preuve : La propriété se prouve facilement, grâce au lemme 1.4.2, car $u_S(\mathbf{g} \wedge \mathbf{y}) = u_S(\mathbf{g})$ ou $\mu(y)$.

Et de même,

Proposition 1.4.6 *u_S satisfait* **DDR**.

Preuve : De la même manière, grâce au lemme 1.4.3, on a $u_S(\mathbf{f} \vee \mathbf{y}) = u_S(\mathbf{f})$ ou $u_S(\mathbf{f} \vee \mathbf{y}) = u_S(\mathbf{y})$.

Une propriété plus forte de l'intégrale de Sugeno peut être montrée : la propriété de décomposabilité pour le max et le min, pour des actions comonotones. Rappelons que deux actions **f** et **g** sont dites comonotones si et seulement si $\forall s, s' \in S$, $\mathbf{f}(s) >_P \mathbf{f}(s') \Rightarrow \mathbf{g}(s) \geq_P \mathbf{g}(s')$.
Alors, dans le cas où $\succeq$ est représentée par une fonction d'utilité qui est une intégrale de Sugeno, les propriétés suivantes sont valables :

Proposition 1.4.7 *Si* **f** *et* **g** *sont comonotones, elles vérifient :*
$u_S(\mathbf{f} \vee \mathbf{g}) = \max(u_S(\mathbf{f}), u_S(\mathbf{g}))$

Preuve :

Remarquons d'abord que : **f** = $\vee_{i=1\ldots n} x_i F_{x_i} x_*$.
Ensuite, on prouve le lemme suivant :

Lemme 1.4.4 **f** *et* **g** *comonotones* $\Leftrightarrow \forall i$, $F_{x_i} \subseteq G_{x_i}$ *ou* $G_{x_i} \subseteq F_{x_i}$.

Preuve :

- $\Rightarrow$ Par l'absurde : Supposons l'existence de i tel que ni $F_{x_i} \subseteq G_{x_i}$ ni $G_{x_i} \subseteq F_{x_i}$. Soient alors $s \in F_{x_i} - G_{x_i}$ et $s' \in G_{x_i} - F_{x_i}$. Puisque $s \in F_{x_i}$ et $s' \notin F_{x_i}$, $\mathbf{f}(s) >_P \mathbf{f}(s')$ et puisque $s' \in G_{x_i}$ et $s \notin G_{x_i}$, $\mathbf{g}(s') >_P \mathbf{f}(s)$, ce qui contredit l'hypothèse de comonotonie.
- $\Leftarrow$ Montrons que $\mathbf{f}(s) >_P \mathbf{f}(s') \Rightarrow \mathbf{g}(s) \geq_P \mathbf{g}(s')$.
  $\mathbf{f}(s) >_P \mathbf{f}(s') \Leftrightarrow \exists i, s \in F_{x_i}, s' \notin F_{x_i}$. Supposons $\mathbf{g}(s') >_P \mathbf{g}(s)$, i.e. $\exists j, s' \in G_{x_j}, s \notin G_{x_j}$. Or, $s \in F_{x_i}$ et $s \notin G_{x_j} \Rightarrow x_j >_P x_i$ et $s' \notin F_{x_i}$ et $s' \in G_{x_j} \Rightarrow x_i >_P x_j$, d'où une contradiction qui implique que $\mathbf{g}(s) \geq_P \mathbf{g}(s')$.

Remarquons que pour toute action **f**, $u_S(\mathbf{f}) = \max_{x_i \in X} u_S(x_i F_{x_i} x_*)$ puisque d'après la proposition 1.4.1 page 93, $u_S(\mathbf{f}) = \max_{x_i \in X} \min(\mu(x_i, \sigma(F_{x_i}))$ et d'après la proposition 1.4.2 page 94, $u_S(x_i F_{x_i} x_*) = \min(\mu(x_i, \sigma(F_{x_i}))$.

Ensuite, si **h** = $\mathbf{f} \vee \mathbf{g}$, $\forall i, H_{x_i} = F_{x_i} \cup G_{x_i}$. Si **f** et **g** sont comonotones, d'après le lemme précédent, $\forall i, F_{x_i} \subseteq G_{x_i}$ ou $G_{x_i} \subseteq F_{x_i}$. Alors, $\forall i, H_{x_i} = \max_{\subseteq}\{F_{x_i}, G_{x_i}\}$. Donc, $\forall i, u_S(x_i H_{x_i} x_*) = \max\{u_S(x_i F_{x_i} x_*), u_S(x_i G_{x_i} x_*)\}$. En permutant les deux max dans la nouvelle expression obtenue de $u_s(\mathbf{h})$, on prouve la proposition, soit :

f et **g** comonotones $\Rightarrow u_S(\mathbf{f} \vee \mathbf{g}) = \max(u_S(\mathbf{f}), u_S(\mathbf{g}))$.

Proposition 1.4.8 *Si* **f** *et* **g** *sont comonotones, elles vérifient :*
$u_S(\mathbf{f} \wedge \mathbf{g}) = \min(u_S(\mathbf{f}), u_S(\mathbf{g}))$

Preuve :

Seule la seconde partie de cette preuve change par rapport à la précédente. Pour prouver cette proposition, il nous faut utiliser l'expression duale de l'utilité qualitative monotone que nous ne définirons qu'un peu plus tard (définition 1.4.2 page suivante). Par une proposition semblable à la proposition 1.4.1 page 93, cette forme duale s'écrit :

$u_S(\mathbf{f}) = \min_{x_i \in X} \max(\mu(x_i), \sigma(F_{x_{i+1}}))$ (proposition 1.4.10 page 101).

Si **h** = **f**$\wedge$**g**, on note que si **f** et **g** sont comonotones, $\forall i, H_i = F_i \cap G_i = \min_{\subseteq}\{F_i, G_i\}$. Donc, $\sigma(H_{i+1}) = \min\{\sigma(F_{i+1}), \sigma(G_{i+1})\}$, et en reportant dans l'expression duale de u_S, on prouve le résultat, soit :

f et **g** comonotones $\Rightarrow u_S(\mathbf{f} \wedge \mathbf{g}) = \min(u_S(\mathbf{f}), u_S(\mathbf{g}))$.

Voila une revue d'axiomes *à la Savage* satisfaits par l'utilité qualitative monotone. Dans la suite nous allons montrer comment certains de ces axiomes permettent de justifier axiomatiquement la fonction d'utilité qualitative monotone.

1.4.2 Axiomatisation de la fonction d'utilité qualitative monotone

Dans ce paragraphe nous allons montrer le théorème suivant, proposant une axiomatisation de l'utilité qualitative monotone en tant que critère de décision dans l'incertain :

Théorème 1.4.1 *Axiomatisation de l'utilité qualitative monotone.*
Soit $\preceq$ une relation de préférence entre actions satisfaisant **Sav 1**, **CFAC**, **Sav 5**, **DCR**, **DDR**. *Il existe une échelle simplement ordonnée, finie $(L, \leq)$, une fonction d'utilité $\mu : X \to L$ et une fonction d'ensemble monotone $\sigma : 2^S \to L$, telles que*

f $\preceq$ **f'** $\Leftrightarrow u_S(\mathbf{f}) \leq u_S(\mathbf{f'})$, *où* $u_S(\mathbf{f}) = \max_{x \in X} \min(\mu(x), \sigma(F_x))$.

Preuve :

- Nous avons déjà montré, par le théorème 1.3.1 page 85, que si la relation $\preceq$ satisfait **Sav 1**, **CFAC** et **Sav 5** alors sa restriction aux actions binaires définit une relation d'incertitude qui est une mesure de Sugeno.
  Nous avons également montré par le lemme 1.3.3 page 85 qu'une telle relation de préférence est compatible avec la relation de *domination*, $\leq_P$.
- On montre ensuite le lemme suivant, concernant les actions de la forme $\mathbf{x}Ax_*$.

  Lemme 1.4.5 *Sous* **Sav 1**, **CFAC**, **Sav 5**, **DCR**, $u_S(\mathbf{x}Ax_*) = \min(\sigma(A), \mu(x))$.

  Preuve :

  Remarquons tout d'abord que $\mathbf{x}Ax_* = x^*Ax_* \wedge \mathbf{x}$. Par **DCR** on déduit que $\forall$ **f**, **f** $\prec x^*Ax_*$ et **f** $\prec$ **x** $\Rightarrow$ **f** $\prec \mathbf{x}Ax_*$. Maintenant, supposons que $\mathbf{x}Ax_*$

$\prec x^*Ax_*$ et $\mathbf{x}Ax_* \prec \mathbf{x}$. Par **DCR** nous obtiendrions $\mathbf{x}Ax_* \prec \mathbf{x}Ax_*$, ce qui est impossible. Ainsi, soit $\mathbf{x}Ax_* \succeq x^*Ax_*$ soit $\mathbf{x}Ax_* \succeq \mathbf{x}$.
Mais, par définition de $\mathbf{x}Ax_*$ nous savons aussi que $\mathbf{x}Ax_* \leq_P \mathbf{x}$ et $\mathbf{x}Ax_* \leq_P x^*Ax_*$, alors, par **CFAC** nous obtenons $\mathbf{x}Ax_* \preceq \mathbf{x}$ et $\mathbf{x}Ax_* \preceq x^*Ax_*$.
Finalement,
$u_S(\mathbf{x}Ax_*) = \min(u_S(x^*Ax_*), u_S(\mathbf{x}))$, soit $u_S(\mathbf{x}Ax_*) = \min(\sigma(A), \mu(x))$.

– Maintenant, étendons u_S à l'ensemble des actions, et prouvons le théorème, en ajoutant l'axiome **DCR** : En premier lieu, notons que toute action **f** de X^S (où $X = \{x_0, \ldots x_n\}$, avec $x_0 < x_1 < \ldots < x_n$) peut s'écrire sous la forme suivante : **f** = $\max_{\leq_P, i=0,\ldots,n}(x_i F_{x_i} x_*)$.
Alors, du fait que **f** domine tous les $x_i F_{x_i} x_*$ on déduit que $u_S(\mathbf{f}) \geq max_i(u_S(x_i F_i x_*))$.

L'inégalité réciproque peut être prouvée en suivant le cheminement suivant :
  – On prouve que $\mathbf{f} \preceq \mathbf{x}$ et $\mathbf{f} \preceq x^*Ax_* \Rightarrow \mathbf{f} \preceq \mathbf{x}Ax_*$.
  Preuve:
  Evidente, en appliquant le lemme 1.4.5 page précédente, puisque $u_S(xAx_*) = \min(\sigma(A), \mu(x))$.
  – Maintenant, montrons que $\forall x \in X$, $\mathbf{f} \preceq \mathbf{x}$ et $\mathbf{f} \preceq \mathbf{f}F_x x_* \Rightarrow \mathbf{f} \preceq \mathbf{x}F_x x_*$.
  Preuve :
  Par le lemme 1.3.3 page 85 (monotonie), on montre : $\mathbf{f}F_x x_* \preceq x^*F_x x_*$. En appliquant le point précédent, si $\mathbf{f} \preceq x^*F_x x_*$ et $\mathbf{f} \preceq \mathbf{x}$ alors on a $\mathbf{f} \preceq \mathbf{x}F_x x_*$.
  – Montrons une dernière propriété inductive qui va nous permettre de prouver le théorème : $\forall x_i \in X\,(x_0 < \ldots < x_n)$, $\mathbf{f} \preceq \mathbf{x}_i \Rightarrow \mathbf{f} \preceq \mathbf{x}_i F_{x_i} x_*$ ou $\mathbf{f} \preceq \mathbf{x}_{i-1}$.
  Preuve :
  Supposons que $\mathbf{f} \preceq \mathbf{x}_i$, $\mathbf{f} \succ \mathbf{f}F_{x_i}x_*$ et $\mathbf{f} \succ \mathbf{x}_{i-1}$. Alors, de **DDR** nous obtiendrions $\mathbf{f} \succ \mathbf{f}F_{x_i}\mathbf{x}_{i-1}$ ce qui est en contradiction avec le lemme 1.3.3 puisque par hypothèse, $\mathbf{f} \preceq \mathbf{f}F_{x_i}x_{i-1}$.
  Donc, soit $\mathbf{f} \preceq \mathbf{x}_i F_{x_i} x_*$ soit $\mathbf{f} \preceq \mathbf{x}_{i-1}$. La propriété d'induction est montrée.
  – **Conclusion:**
  En itérant ce dernier point, sachant que $\mathbf{f} \preceq \mathbf{x}_n$, (car $\mathbf{f} \leq_P \mathbf{x}_n$) on montre que $\exists i$, $\mathbf{f} \preceq \mathbf{x}_i F_{x_i} x_*$. Alors : $u_S(\mathbf{f}) = \max_{x\in X} u_S(\mathbf{x}_i F_{x_i} x_*)$.
  Finalement,
$$u_S(\mathbf{f}) = \max_{x\in X} \min(\mu(x), \sigma(F_x))$$

Nous avons utilisé comme critère de décision la forme *maxmin* de l'intégrale de Sugeno : $u_s(\mathbf{f}) = \max_{\lambda\in L} \min(\lambda, \sigma(F_\lambda))$, mais comme nous allons le voir, nous aurions tout aussi bien pu utiliser une forme *minmax*, équivalente.

Définition 1.4.2 *Utilité qualitative monotone duale :*

$$u'_S(\mathbf{f}) = \min_{\lambda\in L} \max(\lambda, \sigma(F'_\lambda))$$

où $F'_\lambda = \{s \in S, \mu(\mathbf{f}(s)) > \lambda\}$.

Les deux formes d'utilité qualitative monotone coïncident :

Proposition 1.4.9 $\forall$ **f**, $u_S(\mathbf{f}) = u'_S(\mathbf{f})$.

Preuve:

Soit **f** une action de S vers X.
Remarquons d'abord que $\forall \lambda, \sigma(F'_\lambda) \leq \sigma(F_\lambda)$ car $F'_\lambda \subseteq F_\lambda$.

- $\forall \lambda, \min(\lambda, \sigma(F_\lambda)) \leq \lambda \leq \max(\lambda, \sigma(F'_\lambda))$. En passant à la limite : $u_S(\mathbf{f}) \leq u'_S(\mathbf{f})$.
- Prouvons l'inégalité réciproque :
  - Si il existe λ^* tel que $\lambda^* = \sigma(F_{\lambda^*})$, alors $\min(\lambda^*, \sigma(F_{\lambda^*})) = \lambda^* = \max(\lambda^*, \sigma(F'_{\lambda^*}))$, car $\sigma(F'_{\lambda^*}) \leq \sigma(F_{\lambda^*})$. So $u_S(\mathbf{f}) = u'_S(\mathbf{f})$.
  - Si ce n'est pas le cas alors il existe λ^* tel que $\lambda^* < \sigma(F_{\lambda^*})$ and $\forall \lambda > \lambda^*$, $\lambda > \sigma(F_\lambda)$. $u_S(\mathbf{f})$ peut être réécrite :
  $u_S(\mathbf{f}) = \max\{\max_{\lambda \leq \lambda^*} \min(\lambda, \sigma(F_\lambda)), \max_{\lambda > \lambda^*} \min(\lambda, \sigma(F_\lambda))\}$.
  Mais, $\max_{\lambda \leq \lambda^*} \min(\lambda, \sigma(F_\lambda)) = \lambda^*$, et $\max_{\lambda > \lambda^*} \min(\lambda, \sigma(F_\lambda)) = \max_{\lambda > \lambda^*} \sigma(F_\lambda)$ $= \sigma(F'_{\lambda^*})$. Alors, $u_S(\mathbf{f}) = \max\{\lambda^*, \sigma(F'_{\lambda^*})\}$. Donc, de manière évidente, $u_S(\mathbf{f}) \geq u'_S(\mathbf{f})$.

Maintenant que nous avons une expression duale de l'utilité qualitative monotone, montrons que l'on peut aussi l'écrire en faisant varier x, à la place de λ:

Proposition 1.4.10 $u'_S(\mathbf{f}) = \min_{x \in X} \max(\mu(x), \sigma(F'_x))$,
où $F'_x = \{s \in S, \mu(\mathbf{f}(s)) > \mu(x)\}$.

Preuve :
La preuve est la même que pour u_S:

1) Puisque $\mu(X) \subseteq L$, alors $u'_S(\mathbf{f}) \leq \min_{x \in X} \max(\mu(x), \sigma(F'_x))$.

2) Soit x_λ tel que $\mu(x_\lambda) = \max\{\mu(x), \mu(x) \leq \lambda)\}$. Évidemment $F'_\lambda = F'_{x_\lambda}$ (car $\mu(\mathbf{f}(s)) > \lambda \Leftrightarrow \mu(\mathbf{f}(s)) > \mu(x_\lambda)$). Mais comme $\mu(x_\lambda) \leq \lambda$, alors $\forall \lambda \in L$, $\exists x \in X$, tel que $F'_x = F'_\lambda$ et $\mu(x) \leq \lambda$, donc $\min_{x \in X} \max(\mu(x), \sigma(F'_x)) \leq u'_S(\mathbf{f})$.

1.4.3 D'autres jeux d'axiomes

DCR et **DDR** peuvent être remplacés par l'axiome suivant qui est une contrepartie de la propriété d'additivité restreinte aux actions comonotones des fonctions d'utilité de Choquet :

Décomposabilité comonotone
Si **f** et **g** sont des actions comonotones $\mathbf{f} \wedge \mathbf{g} \sim \mathbf{f}$ ou $\mathbf{g}$, $\mathbf{f} \vee \mathbf{g} \sim \mathbf{f}$ ou $\mathbf{g}$.

D'autres jeux d'axiomes peuvent être utilisés pour justifier l'utilité qualitative monotone. Conservons les axiomes basiques **Sav 1**, **CFAC** et **Sav 5** qui sont communs à toutes le théories que nous avons passées en revue et ajoutons les deux axiomes suivants qui peuvent recevoir une justification plus *intuitive* que les axiomes **DCR** et **DDR** :

NC *Non-compensation*:
$\forall A \subseteq S, \forall x \in X, xAy \sim \mathbf{f} \in \{\mathbf{x}, \mathbf{y}, 1A0, 1\bar{A}0\}$.

COM *Commensurabilité*:
$\forall A \subseteq S, \exists x \in X, x^*Ax_* \sim \mathbf{x}$.

L'axiome de non-compensation stipule que l'utilité d'une action binaire reflète forcément exactement soit l'utilité d'une de ses conséquences, soit la confiance dans l'obtention de l'une de ces conséquences. Il n'est pas question d'évaluer une action en faisant la moyenne de ses conséquences.
L'axiome de commensurabilité montre que l'on peut projeter l'échelle d'incertitude sur l'échelle de préférence : pour chaque événement A il existe une conséquence x que l'on est prêt à *échanger* avec un pari sur A aux conséquences extrêmes.
Si l'on suppose **NC** et **COM**, on peut montrer qu'il existe une conséquence x telle que l'agent est prêt à échanger **f**, soit avec la conséquence x, soit avec un pari sur F_x de conséquences 1 et 0 : $u_S(\mathbf{f}) = \mu(x) = \sigma(F_x)$.

Ainsi, les axiomes **NC** et **COM** peuvent sembler plus forts que les axiomes **DCR** et **DDR** puisque ces derniers n'imposent pas explicitement, pour chaque action **f**, l'existence d'une conséquence x telle que $x \sim F_x$. Cependant, on peut montrer que les deux jeux d'axiomes sont équivalents en présence des axiomes **Sav 1**, **CFAC** et **Sav 5**, et on peut prouver le théorème suivant :

Théorème 1.4.2 *Axiomatisation de l'utilité qualitative monotone :*
Soit $\preceq$ une relation de préférence sur les actions satisfaisant **Sav 1**, **CFAC**, **Sav 5**, **NC** *et* **COM**. *Il existe une échelle finie qualitative $(L, \leq)$, une fonction d'utilité $\mu : X \to L$ et une fonction d'ensemble monotone $\sigma : 2^S \to L$, telle que* **f** $\preceq$ **f'** $\Leftrightarrow u_S(\mathbf{f}) \leq u_S(\mathbf{f'})$

Preuve :

- Les deux premiers points de la démonstration sont les mêmes que pour la précédente, car ils n'utilisent que **Sav 1**, **CFAC** et **Sav 5**.
- Ensuite, pour toute action **f** et toute conséquence $x \in X$, rappelons que $F_x = \{s \in S, \mu(\mathbf{f}(s)) \geq \mu(x)\}$. Les deux propriétés $x^* F_x \mathbf{x} \geq_P$ **f**, et **f** $\geq_P \mathbf{x} F_x x_*$ découlent de cette définition.
  Alors, grâce au lemme 1.3.3 on trouve : $\forall \mathbf{f}, \forall x,\ u(x^* F_x \mathbf{x}) \geq u(\mathbf{f}) \geq u(\mathbf{x} F_x x_*)$.

  Maintenant, puisque ces deux inégalités tiennent pour tout $x \in X$, on a :
  $\min_{x \in X} u(x^* F_x \mathbf{x}) \geq u(\mathbf{f}) \geq \max_{x \in X} u(\mathbf{x} F_x x_*)$.

- Appliquons **NC** : $u(x^* F_x \mathbf{x}) \in \{\sigma(F_x), \mu(x)\}$, et $u(\mathbf{x} F_x x_*) \in \{\sigma(F_x), \mu(x)\}$.
  En appliquant le lemme 1.3.3 quelques fois supplémentaires, on prouve que $u(\mathbf{1} F_x \mathbf{x}) \leq \max\{\sigma(F_x), \mu(x)\}$ et $u(\mathbf{x} F_x x_*) \geq \min\{\sigma(F_x), \mu(x)\}$.
  En fin de compte on a montré que
  $\min_{x \in X} \max\{\sigma(F_x), \mu(x)\} \geq u(\mathbf{f}) \geq \max_{x \in X} \min\{\sigma(F_x), \mu(x)\}$.
- Appliquons maintenant **COM** :
  $\exists x_0, \sigma(F_{x_0}) = \mu(x_0)$. $\forall x/\mu(x) > \mu(x_0), \max\{\sigma(F_x), \mu(x)\} > \mu(x_0) = \sigma(F_{x_0})$
  et $\forall x/\mu(x) \leq \mu(x_0)$, puisque $\mu(x) \leq \mu(x_0) \Rightarrow \sigma(F_x) \geq \sigma(F_{x_0})$,
  alors $\max\{\sigma(F_x), \mu(x)\} \geq \mu(x_0) = \sigma(F_{x_0})$.
  Finalement, puisque X est fini, on obtient $\mu(x_0) = \sigma(F_{x_0})$
  $= \min_{x \in X} \max\{\sigma(F_x), \mu(x)\}$.
- De manière similaire on peut montrer que $\mu(x_0) = \sigma(F_{x_0})$
  $= \max_{x \in X} \min\{\sigma(F_x), \mu(x)\}$.

Finalement, on a montré que si $\preceq$ satisfait **Sav 1**, **CFAC**, **Sav 5**, **NC**, **COM**, alors il existe une échelle qualitative finie L, une fonction d'utilité $\mu : X \to L$, et une mesure monotone $\sigma : 2^S \to L$, telles que $\preceq$ peut être représentée par une fonction d'utilité u telle que :

$\forall \mathbf{f}, u(\mathbf{f}) = \min_{x \in X} \max\{\sigma(F_x), \mu(x)\} = \max_{x \in X} \min\{\sigma(F_x), \mu(x)\}$.

1.5 Liens entre l'approche possibiliste de la décision dans l'incertain et les approches $(\max, +)$

1.5.1 Approche $(\max, +)$ de la décision dans l'incertain

Akian et col. [AQV94, Aki95, Ber96] ont développé une approche de la décision dans l'incertain (étendue à la décision séquentielle et au contrôle de systèmes) basée sur une algèbre de type $(\max, +)$, contrepartie de la théorie classique de l'utilité espérée, de type $(+, \times)$.
La mesure de probabilité P est remplacée par une *mesure de coût*[4] K. Les mesures de coût, cas particuliers de mesures de Maslov [Mas87], sont définies par [AQV94] :

Définition 1.5.1 *Mesure de coût.*
$K : S \to]-\infty; 0] \cup \{-\infty\}$ *est une mesure de coût si elle vérifie les axiomes suivants :*

- $K(\emptyset) = -\infty$
- $K(S) = 0$
- *pour toute famille* (A_i) *de sous-ensembles de* S, $K(\cup A_i) = \sup_i K(A_i)$.

A toute mesure de coût K correspond une *distribution de coût*, Γ, telle que

$$\forall A \subseteq S, K(A) = sup_{s \in A} \Gamma(s).$$

On peut définir une contrepartie de l'utilité espérée d'une décision (appelée *frayeur mathématique* par les auteurs),

$$F(\mathbf{f}) = \max_{s \in S}\{u(\mathbf{f}(s)), \Gamma(s)\}.$$

La frayeur mathématique vérifie les propriétés suivantes :

- $F(\mathbf{f} \vee \mathbf{g}) = \max\{F(\mathbf{f}), F(\mathbf{g})\}$
- $F(\mathbf{f} + x) = u(x) + F(\mathbf{f})$.

On peut vérifier que ce critère de décision dans l'incertain est étroitement relié aux critères de type $(max, \top)$ et en particulier que l'ordre de préférence entre les actions exprimé par F est exprimable par un critère de type $(max, \times)$, l'incertitude étant représentée par une distribution de possibilité. Ceci n'est pas surprenant puisque les mesures de Maslov et les mesures de possibilité coïncident.

4. Le nom de *mesure de coût* est trompeur : il s'agit bien d'une mesure d'incertitude, et nom d'une fonction mesurant l'utilité, cependant, les degrés d'incertitude et de préférence sont additionnés.

1.5.2 Lien entre l'approche $(max, +)$ (avec distributions de coût) et l'approche possibiliste de type $(max, \times)$

Si les préférences d'un agent entre les actions sont modélisées par une *frayeur mathématique*, F, elles le sont également par la fonction $F' = e^F$, qui en est une transformée (continue) strictement croissante. $\forall$ **f**, $F'(\mathbf{f}) = e^{\max_{s \in S}\{u(d(s)+\Gamma(s)\}} = \max_{s \in S} e^{u(d(s))} \cdot e^{\Gamma(s)}$.
Plaçons nous dans le cas où les utilités des conséquences sont toutes négatives ($u < 0$). Si nous considérons que les utilités des conséquences et les relations de plausibilité entre états sont mesurées par les fonctions u' et Γ', transformées de u et Γ par : $u'(x) = e^{u(x)}$ et $\Gamma'(s) = e^{\Gamma'(s)}$, alors F' s'écrit :

$$F'(\mathbf{f}) = \max_{s \in S} u'(d(s)) \times \Gamma'(s)$$

où u' et Γ' sont à valeurs dans $[0, 1]$.
On peut vérifier aisément que Γ' et $K' = e^K$ sont respectivement une distribution de possibilité et sa mesure associée.

Le critère de décision basé sur la notion de *frayeur mathématique* est donc très lié aux critères de décision de type $(max, \top)$ que nous avons évoqués dans ce chapitre. Il serait intéressant d'étendre les axiomatisations proposées dans ce chapitre à ce type de critère. Cela pourrait se faire en partant de l'axiomatisation du critère possibiliste optimiste, mais nécessiterait deux modifications :

– repasser d'une échelle purement ordinale à l'échelle $[0, 1]$ (où à un intervalle fermé quelconque),

– remplacer l'axiome d'optimisme par une sorte d'axiome de monotonie stricte : $\forall$ **f**, **g**, ($\forall s$, **f**$(s) >$ **g**$(s) \Rightarrow$ **f** $\succ$ **g**.

L'approche de Spohn [Spo88] est encore plus directement liée à l'utilisation des mesures de coût pour la modélisation de l'incertitude. Il utilise des fonctions κ, qui sont les opposées des mesure de coût.

1.6 Représentation de problèmes de décision dans l'incertain par des problèmes de satisfaction de contraintes flexibles

Dans [SDGP98] nous proposons de représenter un problème de décision dans l'incertain par un PSC flexible dont la solution (décision) optimale maximise l'utilité qualitative possibiliste pessimiste. Cette section résume certains des résultats proposés dans [SDGP98] et [SDGP95].

Le langage des *Problèmes de Satisfaction de Contraintes Flexibles* [Far94] peut être utilisé pour modéliser à la fois la flexibilité et les priorités entre contraintes, ainsi que l'incertitude des paramètres dans les problèmes de satisfaction de contraintes. En fait, on peut représenter à la fois :

– des préférences entre les valeurs possibles de n-uplets de variables, contraintes par des relations floues,

– l'effet sur le degré de satisfaction global d'une décision de la violation complète d'une contrainte,

– les valeurs plus ou moins plausibles de paramètres incertains.

On trouvera dans l'annexe B page 169 l'article [SDGP95] présentant une application à un problème d'ordonnancement de l'utilisation des PSC flexibles pour la décision dans l'incertain.

1.6.1 Contraintes flexibles et préférences

Lorsque les contraintes flexibles sont utilisées pour représenter l'incertitude, les variables sont *incotrôlables* et les contraintes décrivent les valeurs plus ou moins plausibles des variables. Le degré de satisfaction $\pi_C(v)$ représente dans ce cas à quel point la valeur v d'un n-uplet de variables est plausible, étant donnée la contrainte C.
Lorsque les contraintes flexibles sont utilisées pour modéliser les préférences, elles décrivent des valeurs plus ou moins *autorisées* pour des variables contrôlables. Dans ce cas le degré de satisfaction $\pi_C(d)$, ou d est un n-uplet de variables contrôlables, représente le degré auquel d satisfait la contrainte C.

Dans le cas où un PSC flexible ne comprend qu'un ensemble de contraintes de préférence (contraignant un n-uplet d de variables de décision), $\{C_1, \ldots, C_p\}$, l'ensemble R des solutions du problème est un sous-ensemble flou du domaine D des variables.
A chaque contrainte C_i correspond un ensemble flou de solutions satisfaisantes, R_i, défini par sa fonction d'appartenance $\mu_i : D \to L$:

- $\mu_i(d) = 1_L$ si d satisfait complètement C_i,
- $\mu_i(d) = 0_L$ si d viole complètement C_i et
- $0_L < \mu_i(d) < 1_L$ si d satisfait partiellement C_i.

Le degré de satisfaction global d'un ensemble de contraintes $C_1, \ldots, C_p$ par une décision d peut être calculé en réalisant l'agrégation conjonctive des relations floues $R_1, \ldots, R_p$. En supposant que tous les ensembles flous sont définis sur le même domaine (ce qui peut toujours être réalisé en utilisant une extension cylindrique [Zad75]), ce degré de satisfaction s'écrit

$$sat(d) = \min_{i \in 1 \ldots p} \mu_i(d).$$

Ce type d'agrégation, *égalitariste*, exprime que la violation d'une seule contrainte entraîne le rejet d'une solution.
Ce modèle permet de trouver des solutions partiellement satisfaisantes à un problème surcontraint, il permet également de choisir entre des solutions d'un problème sous-contraint, puisque les contraintes peuvent aussi représenter les préférences d'un agent.

Lorsque l'on utilise des contraintes classiques (non-flexibles), il est parfois nécessaire d'ignorer certaines contraintes de faible priorité afin de trouver une solution à un problème surcontraint. L'utilisation de la théorie des possibilités dans les PSC flexibles permet de représenter explicitement les niveaux de priorité de chaque contrainte. Si une contrainte C a un niveau de priorité $\alpha \in L$, toute solution d a un degré de satisfaction au moins égal à $n(\alpha)$ (n est la fonction de renversement de L) [Far94].
L'ensemble flou R' des solutions de la contrainte C de degré de priorité α peut être calculé à partir de l'ensemble flou R des solutions de C,

$$\mu_{R'}(d) = \max\{\mu_R(d), n(\alpha)\}.$$

D'après cette expression, si d viole totalement la contrainte C, son degré de satisfaction reste tout de même $n(\alpha)$. En particulier, remarquons que :

- si $\alpha = 1_L$, $\mu_{R'}(d) = \mu_R(d), \forall d$. La contrainte C ne peut jamais être ignorée,

– si $\alpha = 0_L$, $\mu(R'(d) = 1_L, \forall d$. La contrainte C n'est pas prise en compte pour l'évaluation globale des décisions.

1.6.2 Paramètres incertains, contraintes induites

Nous avons insisté sur le fait qu'il existe deux sortes de variables : les contrôlables dont l'agent peut préciser la valeur, afin de satisfaire les contraintes, et les incontrôlables correspondant à des paramètres dont la valeur est mal connue.
Nous allons voir que les paramètres mal connus induisent à leur tour des contraintes de préférence sur les variables de décision impliquées dans le mêmes contraintes.

Soit A, l'ensemble (pour l'instant non flou) des valeurs possibles du vecteur de variables mal connues, a, impliquées dans la contrainte (pour l'instant non flexible) C. Si D est l'ensemble des variables de décision impliquées dans C, on recherche l'ensemble des “bonnes décisions”, soit

$$\{d \in D, \forall a \in A, (d, a) \in R\}.$$

Dans le cas où A est un ensemble flou, représentant les valeurs “plus ou moins plausibles” des variables mal connues, on recherche une décision d telle que les valeurs plausibles de A soient bonnes, i.e. telles que $(d, a) \in R$. Notons $d^{-1}(R)$, l'ensemble des valeurs des paramètres incertains permettant de satisfaire les contraintes lorsque la décision d est prise. d est entièrement satisfaisante lorsque le support de A est inclus dans $d^{-1}(R)$. Lorsque le noyau de A n'est pas inclus dans $d^{-1}(R)$, il existe une valeur totalement plausible des paramètres incertains, incompatible avec la décision d. Dans ce cas, d n'est pas satisfaisante. En général, on recherche une décision d maximisant le degré d'inclusion (floue) de A dans $d^{-1}(R)$: $N(A \subseteq d^{-1}(R))$.

Dans le cas général où les paramètres sont mal connus et la (les) contrainte(s) flexibles, le degré d'inclusion $N(A \subseteq d^{-1}(R))$ est un degré d'inclusion d'ensembles flous. Si l'on définit l'utilité $v_*(d)$ d'une décision, par $v_*(d) = N(A \subseteq d^{-1}(R))$, on en obtient l'expression suivante :

$$v_*(d) = \min_a \max(\mu_R(d, a), n(\mu_A(a))).$$

En remarquant que $\mu_A(a)$ est le degré de possibilité de l'état du monde défini par la valeur a ($\pi(a)$) des paramètres et que $\mu_R(d, a)$ est le degré de satisfaction de l'action d dans l'état du monde a (que l'on peut assimiler à l'utilité de la conséquence de l'action d appliquée dans l'état du monde a, $\mu(d(a))$), on peut écrire

$$v_*(d) = \min_a \max(\mu(d(a)), n(\pi(a))).$$

Où l'on reconnait l'expression de l'utilité qualitative possibiliste pessimiste, ce qui n'est pas surprenant puisque celle-ci correspond également à un degré d'inclusion d'ensembles flous.

Dans l'article [SDGP95] présenté dans l'annexe B page 169 on montre un problème réel d'ordonnancement (rencontré dans l'industrie vinicole) faisant appel au mode de représentation des PSC flexibles. Cet article fait intervenir des contraintes (linéaires) à domaines continus et le problème est résolu par programmation linéaire.

1.7 Résumé, conclusion

Rappelons dans un premier temps les axiomes concernant les actions proposés dans ce chapitre :

IF *Indépendance faible* : $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h} \Rightarrow \mathbf{f}A\mathbf{h'} \succeq \mathbf{g}A\mathbf{h'}$.

CFAC *Cohérence faible avec les actes constant* : $x \geq_P y \Rightarrow xA\mathbf{h} \succeq yA\mathbf{h}$.

PES *Pessimisme* : $\mathbf{f}A\mathbf{g} \succ \mathbf{f} \Rightarrow \mathbf{f} \succ \mathbf{g}A\mathbf{f}$.

OPT *Optimisme* : $\mathbf{f} \succ \mathbf{f}A\mathbf{g} \Rightarrow \mathbf{g}A\mathbf{f} \succ \mathbf{f}$.

DCR *Domination conjonctive restreinte* : $\mathbf{g} \succ \mathbf{f}$ et $\mathbf{x} \succ \mathbf{f} \Rightarrow \mathbf{g} \wedge \mathbf{x} \succ \mathbf{f}$.

DDR *Domination disjonctive restreinte* : $\mathbf{f} \succ \mathbf{g}$ et $\mathbf{f} \succ \mathbf{x} \Rightarrow \mathbf{f} \succ \mathbf{g} \vee \mathbf{x}$.

D *Décomposabilité* : $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ et $\mathbf{h'} \geq_P \mathbf{h} \Rightarrow \mathbf{f}A\mathbf{h'} \succeq \mathbf{g}A\mathbf{h'}$.

DD *Décomposabilité duale* : $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ et $\mathbf{h} \geq_P \mathbf{h'} \Rightarrow \mathbf{f}A\mathbf{h'} \succeq \mathbf{g}A\mathbf{h'}$.

Nous avons tout d'abord proposé des caractérisations basées sur des préférences entre actions pour un certain nombre de mesures d'incertitude monotones. Ces caractérisations sont résumées dans le tableau suivant :

Axiomes caractéristiques	Mesures d'incertitude
Sav 1, **CFAC** et **Sav 5**	Mesures monotones
Sav 1, **CFAC**, **Sav 5** et **IF**	Mesures décomposables affaiblies
Sav 1, **CFAC**, **Sav 5** et **D**	Mesures décomposables
Sav 1, **CFAC**, **Sav 5** et **DD**	Mesures décomposables duales
Sav 1, **CFAC**, **Sav 5** et **POS**	Mesures de possibilité
Sav 1, **CFAC**, **Sav 5** et **N**	Mesures de nécessité

Puis nous avons proposé des caractérisations pour certains critères de décision dans l'incertain qualitatifs :

Axiomes caractéristiques	Critères de décision
Sav 1, **CFAC**, **Sav 5**, **DDR** et **DCR**	Utilité qualitative monotone
Sav 1, **CFAC**, **Sav 5**, **DDR** et **PES**	Utilité qualitative pessimiste
Sav 1, **CFAC**, **Sav 5**, **DCR** et **OPT**	Utilité qualitative optimiste

Les théories qualitatives (ou ordinales) de la décision sous incertitude que nous avons présentées en détail dans ce chapitre diffèrent significativement des théories basées sur la notion d'espérance mathématique. Elles ne requièrent qu'un cadre fini et une échelle ordinale alors que l'utilité espérée et ses dérivées basées sur l'intégrale de Choquet nécessitent des représentations numériques de l'incertitude et des préférences et un espace d'états infini.
De plus, les théories que nous avons exposées rejettent la notion de compensation entre éventuelles bonnes et mauvaises conséquences, puisque la notion de moyenne est abandonnée. Notre approche est également très générale puisqu'elle permet de se contenter de mesures monotones pour représenter l'incertitude d'un agent.
L'intégrale de Sugeno, utilisée pour représenter l'utilité qualitative monotone d'un agent est une contrepartie qualitative de l'intégrale de Choquet, axiomatisée par Schmeidler [Sch89] ou Sarin et

Wakker [SW92] pour la décision sous incertitude. La notion de médiane se substitue à la notion de moyenne (ou d'encadrement de moyennes pour les utilités basées sur les fonctions de croyance et de plausibilité). De plus, l'attitude de l'agent vis-à-vis de l'incertitude (pessimisme ou optimisme) peut être prise en compte pour mener aux critères de décisions proposés par Dubois et Prade [DP95b], généralisant le critère de Wald ou son dual optimiste, pour donner des critères plus "réalistes".

Nos résultats permettent de suggérer des critères de décision naturels lorsque l'information de l'agent et ses préférences sont exprimées de manière qualitative. Ces résultats permettent également d'observer l'attitude de l'agent en face de l'incertitude (sa propre mesure de l'incertitude) en observant simplement la manière dont il ordonne les actions et en relevant quels axiomes (**PES**, **OPT**, **RDD**, etc...) il satisfait ou viole.
Il est à noter que des résultats similaires à ceux que nous avons exposés ici ont été obtenus indépendamment, récemment, dans le cadre de la décision multicritère par Marichal [Mar97] entre autres.
Des critères de décision semi-quantitatifs de type *max-plus* ont été proposées, dans le cadre de la décision séquentielle (Akian et col. [AQV94]), (Bernhard, [Ber96]) ou non séquentielle (Akian, [Aki95]). Ces critères peuvent également être exprimés sous la forme $max - \top$, en prenant pour $\top$-norme le produit, et rentrer dans la famille des critères que nous avons étudiés.

Dans ce chapitre, et dans toute cette thèse, nous faisons une hypothèse de commensurabilité entre degrés de préférence et d'incertitude. Cette hypothèse, très forte, est également présente dans les théories que nous avons présentées dans la partie I. Des tentatives pour relâcher cette hypothèse ont été effectuées récemment, notamment par Dubois, Fargier et Prade [DFP97]. Dans cet article les auteurs font remarquer que l'abandon de l'hypothèse de commensurabilité les mène à des théories de l'incertitude proches des théories du *raisonnement non monotone*. Malheureusement, ils font également remarquer que les théories de la décision sous incertitude correspondantes sont peu intéressantes, puisque soit trop peu décisives, soit trop aventureuses.

Bien entendu, ce chapitre ne doit être vu que comme un premier pas vers une théorie complète de la décision qualitative sous incertitude dans un style *à la* Savage. La prochaine étape pour obtenir une telle théorie serait de permettre une prise en compte satisfaisante des actions conditionnelles, et de définir de manière "propre" le conditionnement par rapport aux événements, ce qui n'est pas un problème facile, en l'absence du *principe de la chose certaine* de Savage. Des jalons intéressants ont été posés par Lehmann [Leh96] pour cette voie et certaines de ses idées pourraient être reprises dans notre cadre.

Chapitre 2

Processus décisionnels markoviens possibilistes

2.1 Introduction

Dans ce chapitre nous allons étendre la théorie possibiliste de la décision sous incertitude proposée par Dubois et Prade [DP95b] à la décision séquentielle. La théorie proposée est une contrepartie possibiliste de la théorie des processus décisionnels markoviens. Le calcul d'une *stratégie* (ou séquence d'action) optimale est réalisé à l'aide de méthodes de *programmation dynamique*.
Afin d'avoir une vision synthétique des problèmes de décision séquentielle sous incertitude, nous utiliserons la taxonomie suivante qui permet de classer les problèmes (des plus élémentaires aux plus complexes) selon leurs caractéristiques. Nous donnerons le nom de “processus décisionnels markoviens généralisés” (PDMG en abrégé) à la classe la plus générale de problèmes, car nous nous placerons toujours sous l'hypothèse markovienne de l'indépendance des états futurs du monde (ou système) vis-à-vis des états passés : seuls l'état courant et la stratégie choisie pour le futur influent sur l'évolution à venir du monde.
Les propriétés suivantes permettent de caractériser un problème de décision sous incertitude séquentiel :

– *La structure temporelle des étapes de décision*. Il peut n'y avoir qu'une seule étape de décision ou un ensemble ordonné d'instants (ou étapes) auxquels des décisions doivent être prises. Cet ensemble peut être fini (on parle de décision séquentielle à horizon fini) ou infini et discret (horizon infini).

– *Connaissance disponible sur l'état initial du monde*. Cette connaissance peut être précise (il n'y a qu'un seul état du monde possible), probabiliste (la connaissance de l'état initial est décrite par une distribution de probabilités) ou possibiliste (cette fois, la connaissance est décrite par une distribution de possibilités π à valeurs dans une échelle ordonnée L). Lorsque π prend ses valeurs dans $\{0_L, 1_L\}$ (les bornes de L), on parle de connaissance incomplète : l'état initial appartient à un ensemble de fonction caractéristique π.

– *Connaissance sur les effets des actions*. Les actions peuvent être *déterministes*, ce qui signifie que pour tout état du monde s et pour toute action a, applicable en s, il n'existe qu'une seule conséquence possible, $a(s)$, appartenant à l'ensemble de conséquences, X.

Les actions sont dites *non-déterministes* lorsque pour chaque s et chaque a, il existe un *ensemble* de conséquences possibles (et non une seule).

Les action sont dites *stochastiques* lorsque leurs effets sont décrits par une distribution de probabilité $pr(\cdot|s,a)$ dépendant de l'état courant s et de l'action appliquée a. $pr(s'|s,a)$ représente la probabilité que l'action a appliquée à l'état du monde s résulte en s'.

Les actions sont dites *possibilistes* lorsque leurs effets sont décrits par une distribution de possibilités $\pi(\cdot|s,a)$. Encore une fois, une action non-déterministe peut être vue comme un cas particulier d'action possibiliste dans lequel $\pi(\cdot|s,a)$ représente la fonction caractéristique d'un ensemble.

– *Description des préférences* (buts). Dans les cas de décision séquentielle à horizon fini l'état final atteint par le système est souvent d'une importance capitale pour la satisfaction de l'agent. C'est même, dans la plupart des cas, le seul critère de décision. L'agent peut avoir des préférences *strictes* (non-flexibles) signifiant qu'il désire atteindre un état parmi les états buts. La notion de but peut être définie de façon plus flexible, par l'intermédiaire d'une fonction d'utilité, ou plus généralement, une fonction à valeurs dans une échelle ordonnée quelconque. Parfois, le degré de satisfaction attaché à une conséquence doit être évalué par l'intermédiaire de plusieurs quantités, hétérogènes (on parle dans ce cas de décision multicritère).

– *Rôle des états et actions intermédiaires dans le degré de satisfaction global d'un agent*. Les états intermédiaires par lesquels passe l'état du monde peuvent également être d'une certaine importance pour la satisfaction globale de l'agent. De la même manière, les actions effectuées par l'agent après chaque étape de décision peuvent avoir des coûts qui doivent également être pris en compte. Les degrés de satisfaction attachés aux états intermédiaires ainsi que les coûts des actions effectuées peuvent être agrégés additivement, comme dans la théorie des PDM classiques, mais également par le *minimum* ou par toute autre opération. Les utilités des états intermédiaires peuvent être affectés de facteurs d'amortissement, surtout dans le cas ou l'horizon est infini [Put94].

– *Critère de choix entre les stratégies*. Trouver une stratégie optimale revient à attacher à chaque état du monde pouvant être atteint la *meilleure* action disponible. Dans la théorie de l'utilité espérée ce critère correspond au choix de l'action maximisant l'espérance mathématique de la fonction d'utilité. Des critères plus "qualitatifs" peuvent être utilisés, dont les fonctions d'utilité qualitative possibiliste [DP95b], comme nous le verrons dans la section 2.3 de ce chapitre. D'autres méthodes sont également envisageables, pouvant conduire à des relations d'ordre partiel entre actions.

– *Observabilité*. Un PDMG est dit *complètement observable* si l'état résultant du monde est connu après l'exécution de chaque action. Il est dit *non-observable* si aucune information supplémentaire sur l'état du monde n'est reçue après l'exécution de chaque action. Un PDMG est enfin dit *partiellement observable* lorsque l'agent possède une connaissance incomplète du monde : dans ce dernier cas, l'agent peut exécuter des actions d'acquisition d'information, ou tests, tout au long de l'évolution du monde. Les résultats de ces actions peuvent lui être utiles pour maintenir ses connaissances sur l'état réel du monde (représentées par un ensemble d'états possibles, une distribution de probabilités, de possibilités, ...)

Les différents modèles de décision séquentielle que nous évoquerons seront identifiés par un quadruplet résumant certain des critères cités ci-dessus : en premier lieu, la structure temporelle du problème

($1, N$ et ∞ respectivement pour une, un nombre fini ou infini discret d'étapes), ensuite, le modèle utilisé pour la représentation de l'incertitude concernant les effets des actions (D pour déterministe, pr pour stochastique, π pour possibiliste), ensuite la description des buts (B pour des préférences binaires, $+$ pour des utilités additives, min pour des utilités ordinales agrégées par le min) et enfin le critère de choix utilisé ($\bar{u}$ pour l'espérance mathématique de la fonction d'utilité, v^* et v_* pour les fonctions d'utilité qualitatives possibilistes). En utilisant ces notations, les processus décisionnels markoviens à horizon fini correspondent au quadruplet $< N, pr, +, \bar{u} >$. Les modèles de décisions décrits dans le chapitre 1 de cette partie, pessimiste et optimiste, correspondent aux quadruplets $< 1, \pi, min, v_* >$ et $< 1, \pi, max, v^* >$.

Dans ce chapitre nous allons généraliser les critères de décision possibilistes à la décision séquentielle. En premier lieu nous nous limiterons à des préférences binaires, l'incertitude étant de nature possibiliste et l'observabilité totale. Ensuite, nous considérerons le cas où les préférences, graduelles, sont exprimées par une fonction d'utilité qualitative. En dernier lieu, nous nous intéresserons à des considérations d'observabilité, évoquant brièvement les travaux de Da Costa Pereira [Da 98], concernant les problèmes de décision séquentielle en environnement non observable.

2.2 Décision séquentielle sous incertitude possibiliste avec observabilité totale et préférences non-flexibles

2.2.1 Rappel : états, stratégies et trajectoires

Reprenons une nouvelle fois les notations de Puterman [Put94]. $T = \{1, 2, \ldots, N\}$ représente l'horizon (fini) du problème. L'ensemble des états possibles du monde à l'étape t est noté S_t[1]. L'état initial du monde est s_1 et puisque la dernière décision doit être prise à l'étape N, les derniers états que l'on peut atteindre (sur lesquels portent en général les préférences de l'agent) sont représentés par l'ensemble S_{N+1}.
Sous l'hypothèse d'observabilité totale à l'étape t, l'agent observe l'état $s \in S_t$ du système et choisit une action a parmi l'ensemble $A_{s,t}$ des actions disponibles à l'instant t dans l'état s (il peut arriver aussi que $A_{s,t}$ soit indépendant de s et t). Les effets des actions sont mal connus et la connaissance sur l'état résultant d'une action est décrite par une fonction de transition possibiliste π_t : $\pi_t(s'|s, a)$ est le degré de possibilité que l'état s' soit atteint à l'étape $t + 1$, l'action a ayant été effectuée dans l'état s, à l'étape t.
Une *stratégie* est une collection de règles de décision d_t (une pour chaque étape), associant à chaque état possible $s \in S_t$ une action disponible $d_t(s_t) \in A_{s,t}$. L'ensemble des règles de décision à l'étape t est noté D_t et appelé ensemble de décision : c'est l'ensemble des applications de S_t vers $A_{s,t}$. Une *stratégie partielle* $d_{t\to N} = \{d_t, d_{t+1}, \ldots, d_N\}$ spécifie la séquence de règles de décision à appliquer, depuis l'étape t jusqu'à l'horizon du problème. L'ensemble des stratégies partielles depuis l'étape t est noté $D_{t\to N}$. Une stratégie complète $d = \{d_1, d_2, \ldots, d_N\}$ spécifie entièrement la séquence de règles de décision à appliquer tout au long du problème. Notons que nous définissons les stratégies partielles jusqu'à l'horizon du problème. Ceci est dû au fait que les stratégies (complètes) optimales seront déterminées par induction arrière, et que la nature de l'état final (but, non but, utilité) est d'importance primordiale dans cette détermination.

1. Dans certains cas, l'ensemble des états possibles ne varie pas au cours du temps : $\forall t, t', S_t = S_{t'}$. Dans ces cas nous omettrons l'indice t.

Une trajectoire (complète) $\tau = \{s_1, s_2, \ldots, s_n\}$ est une séquence d'états atteints durant les étapes 1 à N. Nous utiliserons parfois la notation $\tau(i) = s_i$. L'ensemble $S_1 \times \ldots \times S_{N+1}$ des trajectoires concevables est noté $TRAJ$. Si $t < t'$ sont deux étapes de l'horizon du problème, une trajectoire partielle de t à t', notée $\tau_{t \to t'}$ est une séquence d'états $\{s_t, s_{t+1}, \ldots, s_{t'}\}$ atteints entre les étapes t et t'. L'ensemble des trajectoires partielles concevables entre les états t et t', noté $TRAJ_{t \to t'}$ est $S_t \times S_{t+1} \times \ldots \times S_{t'}$.

Dans le reste de cette section, l'ensemble des états buts sera représenté par un sous-ensemble (classique) $B \subseteq S_{N+1}$. Cette représentation ne permet pas de prendre en compte des préférences graduelles : les états appartenant à B sont satisfaisants alors que ceux appartenant à $S_{N+1} - B$ ne le sont pas.

Dans la suite de cette section nous allons commencer par revenir sur le cas extrêmement simple d'une seule étape de décision (de N à $N+1$) et nous définirons, pour chaque état $s \in S_N$ la possibilité et la nécessité qu'une action a mène à un état satisfaisant, ce qui nous permettra de définir des critères de choix. Ensuite, nous examinerons le cas d'un horizon comportant plusieurs étapes et nous verrons comment toute stratégie induit une distribution de possibilité sur l'ensemble des trajectoires. Nous en déduirons une définition de l'optimalité d'une stratégie. Ensuite nous verrons comment des stratégies optimales peuvent se calculer récursivement par des méthodes de type “programmation dynamique”. L'algorithme sera basé sur le calcul (pour t variant de N jusqu'à 1) des quantités $\Pi(But_t(s, a))$ et $\mathsf{N}(But_t(s, a))$ mesurant respectivement la possibilité et la nécessité qu'appliquer l'action a dans l'état s et à l'étape t mène à un état but, *à condition qu'une stratégie optimale soit appliquée à partir de l'étape* $t + 1$.

2.2.2 Le cas d'une seule étape

Considérons l'état du monde s, à l'étape N et une action $a \in A_{s,N}$ que l'on applique en s. Puisque l'évolution du système est décrite par la fonction de transition possibiliste $\pi(\cdot|s, a)$ il est possible de calculer la possibilité et la nécessité de l'événement “l'état successeur est un état but”, noté $But_N(s, a)$:

Définition 2.2.1 *Evaluations d'une action.*
$\Pi(But_N(s, a)) = \max_{s' \in B} \pi_N(s'|s, a)$;
$\mathsf{N}(But_N(s, a)) = \min_{s' \notin B}(1 - \pi_N(s'|s, a))$

$\Pi(But_N(s, a))$ et $\mathsf{N}(But_N(s, a))$ sont respectivement la possibilité et la nécessité du sous-ensemble (classique) B, induites par la distribution de possibilité $\pi(\cdot|s, a)$. $\Pi(But_N(s, a))$ et $\mathsf{N}(But_N(s, a))$ sont donc des degrés de possibilité et de nécessité standards, satisfaisant la propriété suivante :

$$\mathsf{N}(But_N(s, a)) > 0 \Rightarrow \Pi(But_N(s, a)) = 1. \tag{1}$$

Une approche prudente ou pessimiste consiste à préférer une action maximisant la nécessité d'atteindre le but. Bien que ce critère semble naturel, il n'est pas toujours suffisamment discriminant, puisque dans de nombreux cas on ne peut trouver aucune action menant au but avec un degré de nécessité strictement positif. En conséquence, il convient de discriminer les actions ex aequo en utilisant le critère optimiste $\Pi(But_N(s, a))$. Ceci conduit au préordre entre les actions noté $\geq_s$ ($>_s$ pour sa partie stricte), $a \geq_s a'$ se lisant “a est préférée ou équivalente à a' dans l'état s :

Définition 2.2.2 *Préordre entre les actions.*

1. $a >_s a'$ *si et seulement si une des conditions suivantes est satisfaite :*
   - $\mathsf{N}(But_N(s,a)) > \mathsf{N}(But_N(s,a'))$
   - $\mathsf{N}(But_N(s,a)) = \mathsf{N}(But_N(s,a'))$ *et* $\Pi(But_N(s,a)) > \Pi(But_N(s,a'))$ *;*
2. $a \geq_s a'$ *si et seulement si on n'a pas* $a' >_s a$ *;*
3. $a \sim_s a'$ *si et seulement si*
   $\mathsf{N}(But_N(s,a)) = \mathsf{N}(But_N(s,a'))$ *et* $\Pi(But_N(s,a)) = \Pi(But_N(s,a'))$.

Les propriétés suivantes sont vérifiées :

- $\geq_s$ est un préordre complet ;
- $a \geq_s a'$ si et seulement si
  $\mathsf{N}(But_N(s,a)) \geq \mathsf{N}(But_N(s,a'))$ et $\Pi(But_N(s,a)) \geq \Pi(But_N(s,a'))$;
- $a \sim_s a'$ si et seulement si
  $\mathsf{N}(But_N(s,a)) = \mathsf{N}(But_N(s,a'))$ et $\Pi(But_N(s,a)) = \Pi(But_N(s,a'))$.

Une action a est dite *optimale pour s* si et seulement si il n'existe aucune action a' telle que $a' >_s a$. L'ensemble des actions optimales pour l'état s (à l'étape N) est noté $A^*_{s,N}$. Grâce à (1), les actions optimales peuvent être caractérisées plus intuitivement par la propriété suivante :

Proposition 2.2.1 *Caractérisation de* $A^*_{s,N}$.
S'il existe une action a qui, appliquée en s mène au but avec un degré de certitude positif (soit $\mathsf{N}(But_N(s,a)) > 0$*) alors $A^*_{s,N}$ est l'ensemble des actions pour lesquelles* $\mathsf{N}(But_N(s,a))$ *est maximal. Sinon, puisque* $\mathsf{N}(But_N(s,a)) = 0$ *pour tout a, $A^*_{s,N}$ est l'ensemble des actions pour lesquelles* $\Pi(But_N(s,a))$ *est maximal.*

Une stratégie optimale assignera donc à chaque état $s \in S_N$ une action arbitraire de $A^*_{s,N}$. A partir de là on peut définir les degrés de possibilité et de nécessité d'atteindre le but depuis un état s. Ces degrés correspondent à ceux que l'on obtient en appliquant une action optimale en s.

Définition 2.2.3
Soit a^* une action quelconque de $A^*_{s,N}$:

- $\Pi(But_N(s)) = \Pi(But_N(s,a^*))$,
- $\mathsf{N}(But_N(s)) = \mathsf{N}(But_N(s,a^*))$

Cette définition est bien fondée dans la mesure où $\Pi(But_N(s,a^*))$ et $\mathsf{N}(But_N(s,a^*))$ sont constants pour toutes les actions de $A^*_{s,N}$. Grâce à (1) on peut vérifier facilement que
$\Pi(But_N(s)) = \max_{a \in A_{s,N}} \Pi(But_N(s,a))$ et $\mathsf{N}(But_N(s)) = \max_{a \in A_{s,N}} \mathsf{N}(But_N(s,a))$.
En d'autres termes, $\Pi(But_N(s))$ et $\mathsf{N}(But_N(s))$ représentent respectivement la possibilité et la nécessité de l'événement “il existe une stratégie menant de l'état s vers un état but”, ou en d'autres termes, “effectuer une action optimale en s permet d'atteindre un état but”. Sachant que, pour tout s, $\Pi(But_N(s))$ et $\mathsf{N}(But_N(s))$ sont égaux à $\Pi(But_N(s,a))$ et $\mathsf{N}(But_N(s,a))$ pour une certaine action a (optimale pour s), il suit que $\Pi(But_N(s))$ et $\mathsf{N}(But_N(s))$ sont des degrés de possibilité et de nécessité standards, vérifiant donc

$$\mathsf{N}(But_N(s)) > 0 \Rightarrow \Pi(But_N(s)) = 1.$$

2.2.3 Stratégies optimales dans le cas séquentiel

Nous allons maintenant généraliser la notion de stratégie optimale dans le cas où plusieurs étapes de décision sont rencontrées. Pour cela, nous allons devoir en premier lieu définir la possibilité et la nécessité qu'une stratégie mène d'un état initial donné jusqu'à un état but.

Définition 2.2.4 *Possibilité d'une trajectoire étant donnée une stratégie.*
Soit un état $s_t \in S_t$, une stratégie partielle $d_{t\to N} \in D_{t\to N}$ et une trajectoire $\tau_{t+1\to N+1} = \{s_{t+1}, \ldots, s_{N+1}\} \in TRAJ_{t+1\to N+1}$. Le degré de possibilité que la trajectoire $\tau_{t+1\to N+1}$ résulte de l'application de la stratégie $d_{t\to N}$ à s_t est définie par :

$$\pi(\tau_{t+1\to N+1}|s_t, d_{t\to N}) = \min_{i=t\ldots N} \pi(s_{i+1}|s_i, d_i(s_i))$$

Cette définition mérite quelques commentaires. D'abord, $\pi(\cdot|s_t, d_{t\to N})$ est une distribution de possibilité sur $TRAJ_{t+1\to N+1}$ normalisée puisque toutes les fonctions de transition sont normalisées et que toute trajectoire dont les possibilités de transition entre deux éléments successifs sont égales à 1_L a elle-même un degré de possibilité 1_L. Alors, définir une distribution de possibilité sur les trajectoires à partir des possibilités de transition élémentaires revient à définir une distribution de possibilité jointe sur un produit Cartésien d'ensembles, à chacun desquels est attachée une distribution de possibilité. En fait, une trajectoire peut être vue comme un vecteur de transitions élémentaires (s_i, s_{i+1}), non-interactives puisque l'hypothèse de Markov stipule que la distribution de possibilité à l'étape t ne dépend que de s_t et de l'action appliquée, et non du passé du système. Le choix le plus classique pour l'agrégation des possibilités de transition est celui de l'opérateur $minimum$ [DP88] : une trajectoire est exactement aussi "possible" que son maillon le plus faible. Pour des raisons de simplicité nous utiliserons cet opérateur, d'autant qu'il est utilisable pour une échelle de possibilité purement ordinale. Dans le cas où l'échelle utilisée pour mesurer les degrés de possibilité est l'intervalle $[0, 1]$, toute T-norme[2] conviendrait également et les algorithmes proposés dans ce chapitre se généralisent aisément.

Il nous reste maintenant à définir la possibilité et la nécessité qu'une stratégie donnée, appliquée en un état initial, mène à un état but. Ces deux facteurs se déduisant de la distribution de possibilité sur les trajectoires sont respectivement les mesures de possibilité et de nécessité de l'ensemble des *bonnes trajectoires* (i.e. menant à un état but).

Définition 2.2.5 *Bonne trajectoire.*
Une trajectoire $\tau_{t\to N+1}$ est bonne si et seulement si son dernier état, $\tau(N+1)$ est dans B. $BT_{t\to N+1} \subseteq TRAJ_{t\to N+1}$ représente l'ensemble des bonnes trajectoires de t à $N+1$.

Définition 2.2.6 *Possibilité et nécessité qu'une trajectoire soit bonne.*

$$\Pi(But_t(s_t, d_{t\to N})) = \Pi(BT_{t\to N+1}|s_t, d_{t\to N})$$

$$\mathsf{N}(But_t(s_t, d_{t\to N})) = \mathsf{N}(BT_{t\to N+1}|s_t, d_{t\to N})$$

2. On rappelle qu'une T-norme $*$ est une application de $[0, 1] \times [0, 1]$ vers $[0, 1]$, commutative, associative, monotone et dont 1 est un élément neutre. Les T-normes usuelles sont le $minimum$, le $produit$ et la T-norme de Lukasiewicz $(a, b) \to \max\{0, a + b - 1\}$

En utilisant l'expression de la possibilité d'une trajectoire étant donnée une stratégie, l'expression complète de ces degrés devient :

$$\Pi(But_t(s_t, d_{t\to N})) = \max_{\tau_{t\to N+1}\in BT_{t\to N+1}} \pi(\tau_{t\to N+1}|s_t, d_{t\to N})$$

$= \max_{\{s_t,\dots,s_{N+1}\}\in TRAJ_{t\to N+1} s_{N+1}\in B} \min_{i=t\dots N} \pi(s_{i+1}|s_i, d_i(s_i))$

$$\mathsf{N}(But_t(s_t, d_{t\to N})) = \min_{\tau_{t\to N+1}\in TRAJ_{t\to N+1}-BT_{t\to N+1}} n(\pi(\tau_{t\to N+1}|s_t, d_{t\to N}))$$

$= \min_{\{s_t,\dots,s_{N+1}\}\in TRAJ_{t\to N+1} s_{N+1}\in S_{N+1}-B} n(\min_{i=t\dots N} \pi(s_{i+1}|s_i, d_i(s_i)))$

On peut maintenant ordonner les stratégies et définir la notion de stratégie optimale.

Définition 2.2.7 *Ordre entre deux stratégies (partielles).*
Soient deux stratégies $d_{t\to N}$ et $d'_{t\to N}$ et soit $s \in S_t$, $d_{t\to N} >_s d'_{t\to N}$ si et seulement si l'une des deux conditions suivantes est vérifiée :

– $\mathsf{N}(But_t(s_t, d_{t\to N})) > \mathsf{N}(But_t(s_t, d'_{t\to N}))$

– $\Pi(But_t(s_t, d_{t\to N})) > \Pi(But_t(s_t, d'_{t\to N}))$

Les relations $\geq_s$ et $\sim_s$ sont définies comme dans le cas d'une seule étape. Et comme dans ce cas, $\Pi(But_t(s_t, d_{t\to N}))$ et $\mathsf{N}(But_t(s_t, d_{t\to N}))$ sont des degrés de possibilité et de nécessité standards.

Une stratégie optimale est définie de manière évidente :

Définition 2.2.8 *stratégie optimale.*
$d_{t\to N}$ est une stratégie optimale pour $s \in S_t$ si et seulement si il n'existe aucune stratégie $d'_{t\to N}$ telle que $d'_{t\to N} >_s d_{t\to N}$.

2.2.4 Calcul récursif d'une stratégie optimale

Nous allons maintenant calculer par induction arrière, pour chaque étape t, état $s \in S_t$ et chaque action $a \in A_{s,t}$:

(i) les degrés de possibilité et de nécessité qu'appliquer a en s, suivie d'une stratégie optimale pour les états $t+1$ à N conduise finalement à un état but ; une telle éventualité est notée $But_t(s, a)$;

(ii) La possibilité $\Pi(But_t(s))$ et la nécessité $\mathsf{N}(But_t(s))$ qu'un état but puisse être atteint depuis s en appliquant une stratégie optimale de t à N.

Définition 2.2.9 *Possibilité et nécessité d'atteindre un état but depuis s en appliquant a* [3].

$$\Pi(But_t(s, a)) = \max_{s'\in S_{t+1}} \min(\pi_t(s'|s, a), \Pi(But_{t+1}(s')))$$

$$\mathsf{N}(But_t(s, a)) = \min_{s'\in S_{t+1}} \max(n(\pi_t(s'|s, a)), \mathsf{N}(But_{t+1}(s')))$$

3. Nous prouverons dans les lignes suivantes que les expressions notées Π et N sont bien des degrés de possibilité et de nécessité. Nous utilisons d'ores et déjà ces notations pour ne pas en ajouter de nouvelles inutilement.

On peut comparer deux actions en un état $s \in S_t$ comme on l'a fait précédemment pour les états de S_N :

Définition 2.2.10 *Comparaison de deux actions.*
$a >_s a'$ *si et seulement si une des deux conditions suivantes est satisfaite :*

- $\mathsf{N}(But_t(s,a)) > \mathsf{N}(But_t(s,a'))$,
- $\mathsf{N}(But_t(s,a)) = \mathsf{N}(But_t(s,a'))$ *et* $\Pi(But_t(s,a)) > \Pi(But_t(s,a'))$.

Une action a est dite optimale pour un état s si et seulement si il n'existe aucune action a' telle que $a' >_s a$. L'ensemble des actions optimales pour s à l'étape t est noté $A^*_{s,t}$. Une stratégie optimale affecte à chaque état s pour chaque étape t une action optimale.
Enfin, on définit $\Pi(But_t(s))$ et $\mathsf{N}(But_t(s))$ mesurant à quel point il est possible (resp. certain) d'atteindre un état but à l'étape $N+1$ en appliquant une stratégie optimale à partir de s, à l'étape t. Ce sont en fait les mesures de possibilité et de nécessité de l'événement "il existe une stratégie menant à un état but depuis s".

Définition 2.2.11 *Possibilité et nécessité d'atteindre un état but depuis* s.

- $\Pi(But_t(s)) = \Pi(But_t(s,a^*))$,
- $\mathsf{N}(But_t(s)) = \mathsf{N}(But_t(s,a^*))$, *où* a^* *est une action optimale.*

Théorème 2.2.1 $\Pi(But_t(s,a))$, $\mathsf{N}(But_t(s,a))$, $\Pi(But_t(s))$ *et* $\mathsf{N}(But_t(s))$ *sont des degrés de possibilité ou de nécessité standards.*

Preuve :

On prouve par récurrence descendante que $\forall t, \Pi(But_t(s)) < 1 \Rightarrow \mathsf{N}(But_t(s)) = 0$. Ceci a déjà été montré pour $t = N$, on montre maintenant que si cette propriété est vérifiée pour $t+1$ elle l'est également pour t :

Soit $s \in S_t$, supposons que $\Pi(But_t(s)) < 1$. Par définition de $\Pi(But_t(s))$, ceci est équivalent à $\max_{a \in A_{s,t}} \Pi(But_t(s,a)) < 1$, soit $\forall a \in A_{s,t}$, $\Pi(But_t(s,a)) < 1$. Par définition de $\Pi(But_t(s,a))$, cela donne :

$$\forall a \in A_{s,t}, \forall s' \in S_{t+1}, \pi_t(s'|s,a) = 1 \Rightarrow \Pi(But_{t+1}(s')) < 1 \qquad (1)$$

En appliquant (1) et l'hypothèse de récurrence on montre que $\forall a \in A_{s,t}, \forall s' \in S_{t+1}, \pi_t(s'|s,a) = 1 \Rightarrow \mathsf{N}(But_{t+1}(s')) = 0$.

Soit alors $a \in A_{s,t}$. Soit $s' \in S_{t+1}$ tel que $\pi_t(s'|s,a) = 1$ (l'existence d'un tel état est garanti car π_t est normalisée). Puisque $\mathsf{N}(But_{t+1}(s')) = 0$ et $\pi_t(s'|s,a) = 1$ alors $\mathsf{N}(But_t(s,a)) = \min_{s' \in S_{t+1}} \max(n(\pi_t(s'|s,a)), \mathsf{N}(But_{t+1}(s'))) = 0$.

On a donc montré que $\forall a \in A_{s,t}, \mathsf{N}(But_{t+1}(s'))) = 0$, on en conclut que $\mathsf{N}(But_t(s)) = 0$.

On montre maintenant que cette méthode de calcul récursive donne bien des stratégies optimales.

Théorème 2.2.2 *Les stratégies calculées récursivement sont optimales.*

- $\Pi(But_t(s)) = \max_{d_{t\to N} \in D_{t\to N}} \Pi(But_t(s, d_{t\to N}))$,

– $\mathsf{N}(But_t(s)) = \max_{d_{t\to N}\in D_{t\to N}} \mathsf{N}(But_t(s, d_{t\to N}))$.

Preuve :

Remarquons tout d'abord que
$\Pi(But_t(s)) = \max_{a\in A_{s,t}} \Pi(But_t(s,a)) = \max_{d_t\in D_t} \Pi(But_t(s, d_t(s)))$ et $\mathsf{N}(But_t(s)) = \max_{a\in A_{s,t}} \mathsf{N}(But_t(s,a)) = \max_{d_t\in D_t} \mathsf{N}(But_t(s, d_t(s)))$.

- On prouve par induction arrière que
  $\Pi(But_t(s)) = \max_{d\in D_{t\to N}} \max_{\{s_{t+1},\dots,s_{N+1}\}\in BT_{t+1\to N+1}} \min_{j=t,\dots,N} \pi(s_{j+1}|s_j, d(s_j))$.
  Le résultat est évident pour $t = N$ (par définition de $\Pi(But_N(s))$). On montre que si il est vrai pour $t+1 \in 2,\dots N+1$ alors il est vrai pour t :
  $\Pi(But_t(s)) = \max_{d_t\in D_t} \max_{s_{t+1}\in S_{t+1}} \min(\pi_t(s_{t+1}|s, d_t(s)), \Pi(But_{t+1}(s_{t+1})))$.
  Si on suppose le résultat vrai à l'étape $t+1$, on obtient :
  $\Pi(But_t(s)) = \max_{d_t\in D_t} \max_{s_{t+1}\in S_{t+1}} \min(\pi_t(s_{t+1}|s, d_t(s)), \max_{d\in D_{t+1\to N}} \max_{\{s_{t+2},\dots,s_{N+1}\}\in BT_{t+2\to N+1}} \min_{j=t+1,\dots,N} \pi(s_{j+1}|s_j, d(s_j))$.
  Comme d'après l'hypothèse markovienne, $\pi_t(s_{t+1}|s, d_t(s))$ ne dépend pas des états suivants, on a :
  $\Pi(But_t(s)) = \max_{d_t\in D_t} \max_{s_{t+1}\in S_{t+1}} \max_{d\in D_{t+1\to N}} \max_{\{s_{t+2},\dots,s_{N+1}\}\in BT_{t+2\to N+1}} \min_{j=t,\dots,N} \pi(s_{j+1}|s_j, d(s_j))$.
  On intervertit $\max_{s_{t+1}\in S_{t+1}}$ et $\max_{d\in D_{t+1\to N}}$ (on peut le faire car ce sont deux max) :
  $\Pi(But_t(s)) = \max_{d_t\in D_t} \max_{d\in D_{t+1\to N}} \max_{s_{t+1}\in S_{t+1}} \max_{\{s_{t+2},\dots,s_{N+1}\}\in BT_{t+2\to N+1}} \min_{j=t,\dots,N} \pi(s_{j+1}|s_j, d(s_j))$.
  Le résultat est montré.
- La démonstration est du même type pour $\mathsf{N}(But_t(s))$.

Corollaire 2.2.1 *Toute stratégie calculée récursivement par induction arrière est optimale.*

Preuve :

Ce résultat se déduit du théorème précédent et de la définition d'une stratégie optimale.

Ainsi, cet algorithme récursif est sain. Notons cependant qu'il n'est pas complet et qu'il existe des stratégies optimales qui ne peuvent se trouver par induction arrière comme le montre le problème représenté par la Figure 2.1 page suivante (avec $N = 2$ et $G = \{s_6\}$).
La Figure 2.2 page suivante montre quatre stratégies possibles (les actions affectées à chaque état sont en gras et les degrés de possibilité de transition sont associés aux flèches correspondantes).
Dans les quatre cas la trajectoire la plus plausible menant à un état non but est (s_1, s_2, s_4). Les mesures de possibilité de cette trajectoire lorsque d, d', d'' ou d''' sont respectivement appliquées sont 0,8 ; 0,7 ; 0,7 et 1. C'est pourquoi,
$\mathsf{N}(But_1(s, d_{1\to3})) = 0.2;$
$\mathsf{N}(But_1(s, d'_{1\to3}))$ = $\mathsf{N}(But_1(s, d''_{1\to3})) = 0.3$
$\mathsf{N}(But_1(s, d'''_{1\to3})) = 0$ (et $\Pi(But_1(s, d'''_{1\to3})) = 1)$).
d' et d'' sont des stratégies optimales. Pourtant d'' ne peut être obtenue par calcul récursif arrière à cause de sa sous-stratégie entre les étapes 2 et 3 qui affecte une action sous-optimale à s_3. Cet

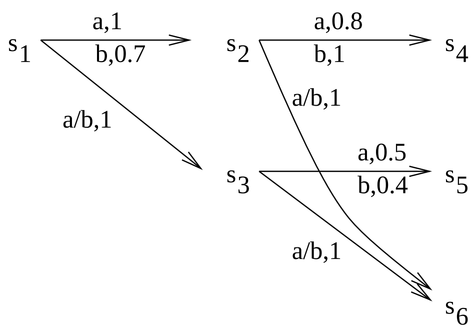

FIG. 2.1 – *Modèle d'action*

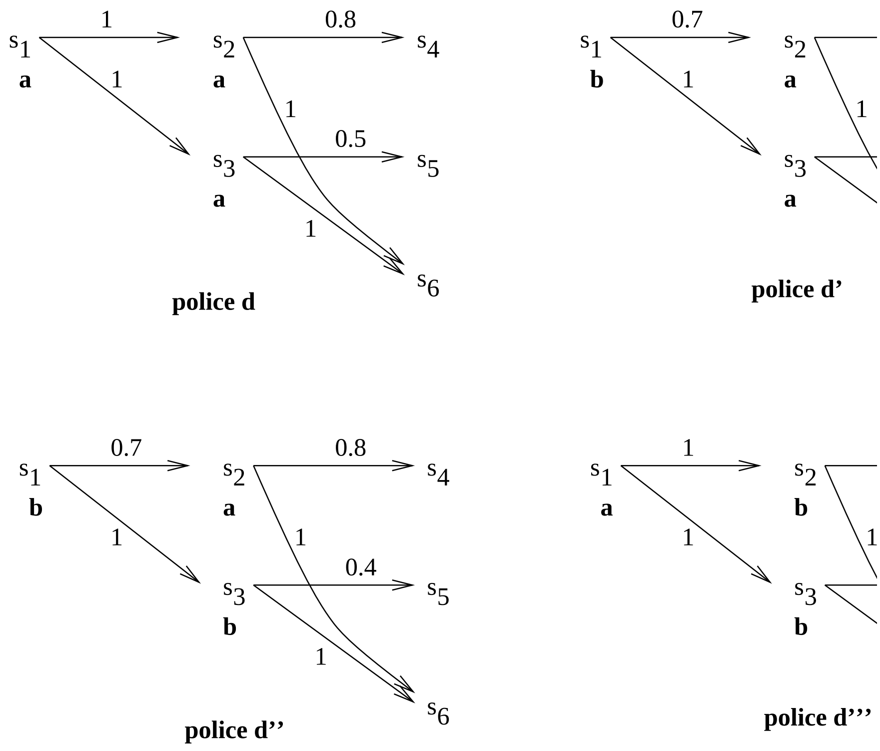

FIG. 2.2 – *Quatre stratégies*

exemple montre qu'une sous-stratégie d'une stratégie optimale (qui n'est pas calculée par induction arrière) peut être sous-optimale, ce qui est une conséquence de l'utilisation de l'opérateur idempotent min pour le calcul de la possibilité d'une trajectoire.

Algorithme 2.1 : CALCUL RÉCURSIF DE STRATÉGIES OPTIMALES DANS LE CAS DE PRÉFÉRENCES BINAIRES

% Cet algorithme est basé sur des méthodes de programmation dynamique ;
% Il évalue par induction arrière alternativement la "valeur" de chaque état puis l'action optimale en cet état ;
Données : d, la stratégie optimale,
G, l'ensemble des états buts,
N, le nombre d'étapes,
S_t, les ensembles d'états possibles pour chaque étape,
$A_{s,t}$ les actions disponibles pour chaque état possible de chaque étape.

début
- % Initialisations ;
- $d \leftarrow \emptyset$;
- **pour** $s \in S_{N+1}$ **faire**
  - **si** $s \in But$ **alors**
    - $\mathsf{N}(But_{N+1}(s)) \leftarrow 1$; $\Pi(But_{N+1}(s)) \leftarrow 1$
  - **sinon**
    - $\mathsf{N}(But_{N+1}(s)) \leftarrow 0$; $\Pi(But_{N+1}(s)) \leftarrow 0$
- % Calcul récursif ;
- **pour** $t \leftarrow N$ *jusqu'à 1* **faire**
  - **pour** $s \in S_t$ **faire**
    - **pour** $a \in A_{s,t}$ **faire**
      - % Calcul de $\mathsf{N}(But_t(s,a))$ et $\Pi(But_t(s,a))$;
      - $\mathsf{N}(But_t(s,a)) \leftarrow \min_{s' \in S_{t+1}} \max\{n(\pi_t(s'|s,a)), \mathsf{N}(But_{t+1}(s'))\}$;
      - $\Pi(But_t(s,a)) \leftarrow \max_{s' \in S_{t+1}} \min\{\pi_t(s'|s,a), \Pi(But_{t+1}(s'))\}$
    - % a^* maximise $[\mathsf{N}(But_t(s,a)), \Pi(But_t(s,a))]$;
    - $d_t(s) \leftarrow a^*$;
    - $\mathsf{N}(But_t(s)) \leftarrow \mathsf{N}(But_t(s,a^*))$;
    - $\Pi(But_t(s)) \leftarrow \Pi(But_t(s,a^*))$
  - $d \leftarrow d \cup d_t$

fin

Grâce à ces résultats on peut proposer un algorithme de calcul de stratégies optimales qui est une variante des méthodes de programmation dynamique de Bellman [Bel57] (algorithme 2.1). Il calcule les stratégies optimales récursivement (de la dernière étape à la première).

Algorithme 2.2 : VERSION AMÉLIORÉE

% Cet algorithme diffère du précédent dans la partie calcul de N et Π ;
% Il permet d'éviter certains calculs superflus ;
Données : % En plus des variables de l'algorithme précédent ;
    n et π, deux variables temporaires de stockage de N et Π,
    n_{opt}, π_{opt} et a_{opt}, valeurs optimales de N et Π pour s et t fixés ainsi qu'une action optimale correspondante.

début
- % Initialisations ;
- $d \leftarrow \emptyset$;
- **pour** $s \in S_{N+1}$ **faire**
  - **si** $s \in But$ **alors**
    - $\mathsf{N}(But_{N+1}(s)) \leftarrow 1$; $\Pi(But_{N+1}(s)) \leftarrow 1$
  - **sinon**
    - $\mathsf{N}(But_{N+1}(s)) \leftarrow 0$; $\Pi(But_{N+1}(s)) \leftarrow 0$
- % Calcul récursif ;
- **pour** $t \leftarrow N$ *jusqu'à 1* **faire**
  - **pour** $s \in S_t$ **faire**
    - $n_{opt} \leftarrow 0$; $\pi_{opt} \leftarrow 0$;
    - **pour** $a \in A_{s,t}$ **faire**
      - % Calcul de N et Π ;
      - $n \leftarrow 1$; $\pi \leftarrow 0$;
      - **pour** $s' \in S_{t+1}$ **faire**
        - **si** $\pi_{opt} < 1$ **alors**
          - $\pi \leftarrow \max\{\pi, \min\{\pi_t(s'|s,a), \Pi(But_{t+1}(s'))\}\}$;
          - **si** $\pi > \pi_{opt}$ **alors**
            - $\pi_{opt} \leftarrow \pi$; $a_{opt} \leftarrow a$;
            - % Dans le cas contraire, π_{opt} ne peut être amélioré et $\Pi(But_t(s,a))$ est inutile ;
        - $n \leftarrow \min_{s' \in S_{t+1}} \max\{n(\pi_t(s'|s,a)), \mathsf{N}(But_{t+1}(s'))\}$;
        - **si** $n < n_{opt}$ **alors** Sortie de la boucle (*) % Dans ce cas a est sous-optimale puisque n ne peut que décroître ;
      - **si** $n > n_{opt}$ **alors** $n_{opt} \leftarrow n$; $a_{opt} \leftarrow a$
    - $\Pi(But_t(s)) \leftarrow \pi_{opt}$; $\mathsf{N}(But_t(s)) \leftarrow n_{opt}$; $d_t(s) \leftarrow a_{opt}$
  - $d \leftarrow d \cup d_t$

fin

L'utilisation des opérateurs min et max à la place de la multiplication et du produit dans les PDM classiques permet d'éviter des calculs inutiles et de simplifier l'algorithme. En effet, une action peut être détectée comme sous-optimale avant la fin du calcul de $\mathsf{N}(But_t(s,a))$ et $\Pi(But_t(s,a))$ (*). Ceci conduit à la version améliorée de l'algorithme (algorithme 2.2).
La complexité de cet algorithme peut être facilement calculée :elle est en $O(N.|S|^2.|A|)$ comme celle de l'algorithme de programmation dynamique [Put94].

2.3 Décision séquentielle possibiliste avec observabilité totale et préférences flexibles

2.3.1 L'ensemble flou des trajectoires satisfaisantes

Jusqu'ici nous n'avons distingué que deux types d'états finaux : ceux qui satisfont les préférences de l'agent et les autres. Cette dichotomie est insuffisante pour modéliser la plupart des problèmes de décision réels dans lesquels un agent exprime des préférences (ou de l'indifférence) entre les états éventuels du système. C'est pourquoi nous allons affecter aux états du système un degré d'utilité qualitatif (ordinal), reflétant les préférences de l'agent, comme exposé dans le chapitre précédent.
Dans de nombreux problèmes l'agent n'exprime de préférences qu'entre les états finaux du système (à l'étape $N+1$). La fonction d'utilité associée est interprétée comme un *degré de satisfaction* $u(s) = \mu_B(s)$, où μ_B est une application de S_{N+1} dans une échelle ordinale L (éventuellement l'intervalle $[0,1]$). $\mu_B(s)$ exprime à quel point s est un état final satisfaisant pour l'agent. μ_B n'est pas forcément normalisée : il se peut qu'aucun état final du système ne soit pleinement satisfaisant.
Plus généralement, des degrés de satisfaction peuvent aussi être assignés aux états intermédiaires du système. On peut également affecter de tels degrés aux actions et pas seulement aux états : de tels degrés peuvent être considérés comme des “anti-coûts” (ordinaux), avec la convention que plus $\mu(a)$ est élevé, moins l'action a est coûteuse ($\mu(a) = 1_L$ signifie que a est gratuite et $\mu(a) = 0_L$ signifie qu'elle est si coûteuse qu'aucune stratégie admissible ne peut la contenir). Pour simplifier l'exposé des méthodes proposées ici nous ne prendrons pas en compte les coûts d'actions, cependant leur intégration aux modèles que nous allons montrer est facile.

L'utilité globale d'une trajectoire est définie à partir des utilités des états qu'elle contient, en les agrégeant par le min :

Définition 2.3.1 *Utilité qualitative d'une trajectoire :*
Soient les fonctions d'utilité qualitative $u_t : S_t \to L$, pour chaque étape t. L'utilité globale d'une trajectoire τ est définie par :

$$u(\tau) = \min_{i \in 1,\ldots,N+1} u_t(s_i)$$

Notons que dans le cas où $L = [0,1]$ n'importe quelle norme triangulaire peut être utilisée à la place du min pour agréger les utilités de chaque étape.
La fonction d'utilité qualitative u sur les trajectoires permet de définir un ensemble flou de bonnes trajectoires dont la fonction d'appartenance est $\mu_{BT}(\tau) = u(\tau)$. Pour la même raison que pour μ_B, μ_{BT} n'est pas forcément normalisée : il se peut qu'il existe une étape dont aucun état n'est pleinement satisfaisant.

2.3.2 Stratégies optimales au sens de l'utilité qualitative pessimiste

A partir des distributions de possibilité et des fonctions d'utilité qualitatives définies sur les trajectoires nous pouvons maintenant exprimer les fonctions d'utilité qualitative pessimiste et optimiste (décrites dans le chapitre précédent) d'une stratégie.
Commençons par l'utilité pessimiste :

Définition 2.3.2 *Utilité qualitative pessimiste associée à une stratégie $d_{t \to N}$:*

$$v_*(d_{t \to N}|s_t) = \min_{\tau \in Traj_{t \to N+1}} \max\{n(\pi(\tau|s_t, d_{t \to N})), u(\tau)\}$$

$d_{t\to N}$ *est* v_**-optimale ssi elle maximise* v_*.

Intuitivement, cette définition suggère que l'utilité qualitative pessimiste d'une stratégie est d'autant plus élevée que toutes les trajectoires les plus plausibles qu'elle induit ont une forte utilité. Ce critère est une généralisation du critère pessimiste proposé par Dubois et Prade [DP95b] à l'évaluation d'une stratégie (et non simplement d'une action). Notons que comme dans la section précédente, des stratégies peuvent être v_*-optimales et toutefois contenir des sous-stratégies qui ne le sont pas.
Deux cas particuliers peuvent être soulignés. Premièrement, si les possibilités de transition $\pi(\tau|s_t, d_{t\to N})$ sont binaires (appartiennent à $\{0_L, 1_L\}$), on reconnaît l'habituel critère de Wald qui assigne à une stratégie l'utilité de la pire trajectoire possible, i.e. du pire état qui peut être rencontré au cours de l'horizon du problème en appliquant cette stratégie. Deuxièmement, si c'est la fonction d'utilité u qui est binaire, l'ensemble des bonnes trajectoires devient un ensemble classique et $v_*(d_{t\to N}|s_t)$ n'est autre que $\mathsf{N}(But_t(s, d_{t\to N}))$ définie dans la section précédente.

Comme précédemment on montre qu'une stratégie v_*-optimale peut être calculée récursivement.

Définition 2.3.3 *Définition récursive d'une stratégie optimale (pessimiste).*

- $\forall s \in S_{N+1}, v_*^{N+1}(s) = u_{N+1}(s)$,
- $\forall s \in S_t, \forall a \in A_{s,t}, v_*^t(s,a) = \min_{s' \in S_{t+1}} \max\{n(\pi_t(s'|s,a)), v_*^{t+1}(s')\}$,
- $\forall s \in S_t, v_*^t(s) = \max_{a \in A_{s,t}} v_*^t(s,a)$.

De plus, on écrit que a est au moins aussi bonne que a' (au sens pessimiste) dans l'état $s \in S_t$, noté $a \geq_s^{Pes} a'$ si et seulement si $v_*^t(s,a) \geq v_*^t(s,a')$. On définit $A_{s,t}^{Pes}$ comme l'ensemble des actions optimales en s, i.e. celles qui maximisent $v_*^t(s,a)$. Bien sûr, $\forall s \in S_t, v_*^t(s) = v_*^t(s,a^*)$ où a^* est une action de $A_{s,t}^{Pes}$.
En fait, $v_*^t(s)$ est la mesure de nécessité de l'événement flou “il existe une stratégie qui à partir de l'état s conduit à une trajectoire optimale”. En d'autres termes,

Théorème 2.3.1 $v_*^t(s) = \max_{d \in D_t} v_*^t(s, d_t(s))$

La preuve de ce théorème, très similaire à celle du théorème 2.2.2 page 116 est omise.

Corollaire 2.3.1 *Toute stratégie calculée récursivement comme dans la définition 2.3.3 est* v_**-optimale.*

Ainsi, comme dans le cas de préférences binaires sur les états finaux, des stratégies maximisant la nécessité d'atteindre un état but peuvent être calculées par induction arrière, affectant à chaque état une action $a \in A_{s,t}^{Pes}$.

2.3.3 Stratégies optimales au sens de l'utilité qualitative optimiste

On peut également classer les stratégies suivant leur utilité qualitative optimiste. Si on adopte un point de vue optimiste, une stratégie est d'autant meilleure que que le degré d'intersection de l'ensemble flou des trajectoires possibles et de celui des bonnes trajectoires est élevé.

Définition 2.3.4 *Utilité qualitative optimiste d'une stratégie* $d_{t\to N}$.

$$v^*(d_{t\to N}|s_t) = \max_{\tau \in Traj_{t\to N+1}} \min\{\pi(\tau|s_t, d_{t\to N}), u(\tau)\}$$

$d_{t\to N}$ *est* v^**-optimale ssi elle maximise* v^*.

Cette fois, une stratégie est d'autant meilleure que l'ensemble flou des trajectoires possibles qu'elle définit et qui ont une forte utilité est non-vide. Si les possibilités de transition sont binaires, on retrouve un critère du type $maximax$: une stratégie est d'autant meilleure qu'elle permet de rencontrer (en suivant une des trajectoires possibles qu'elle définit) au moins un état satisfaisant. Si ce sont les utilités qui deviennent binaires, on retrouve le critère optimiste exposé dans la section précédente.
Tout comme pour v_*, on montre qu'une stratégie v^*-optimale peut être calculée récursivement.

Définition 2.3.5 *Définition récursive d'une stratégie optimale (optimiste).*

- $\forall s \in S_{N+1}, v^*_{N+1}(s) = u_{N+1}(s)$,
- $\forall s \in S_t, \forall a \in A_{s,t}, v^*_t(s,a) = \max_{s' \in S_{t+1}} \min\{\pi_t(s'|s,a), v^*_{t+1}(s')\}$,
- $\forall s \in S_t, v^*_t(s) = \max_{a \in A_{s,t}} v^*_t(s,a)$.

$\geq_s^{Opt}$ est définie de manière semblable à $\geq_s^{Pes}$ et $A_{s,t}^{Opt}$ est l'ensemble des actions $a \in A_{s,t}$ maximisant $v^*_t(s,a)$. Comme précédemment, on obtient aisément que $v^*_t(s) = v^*_t(s,a^*)$ où $a^* \in A_{s,t}^{Opt}$.

En fait, $v^*_t(s)$ est la possibilité de l'événement flou “il existe une stratégie définissant une bonne trajectoire à partir de s” :

Théorème 2.3.2 $v^*_t(s) = \max_{d \in D_t} v^*_t(s, d_t(s))$

Corollaire 2.3.2 *Toute stratégie calculée par induction arrière en suivant la définition 2.3.5 est v^*-optimale*

2.3.4 Buts flexibles ou non-flexibles?

Jusqu'ici, dans ce chapitre nous avons fait une distinction entre les cas où les préférences d'un agent, exprimées à l'étape $N+1$ étaient flexibles ou non. Nous allons maintenant voir que le cas où les préférences sont flexibles peut se réduire au précédent, en supposant que les degrés d'utilité affectés aux états de S_{N+1} sont en fait reliés à des possibilités ou des nécessités d'atteindre un état but à l'étape $N+2$. Ceci conduit à ajouter une étape de décision entre les instants $N+1$ et $N+2$, instant auquel seules des préférences binaires sont maintenant exprimées et une seule action est applicable.
Cependant, les possibilités de transition entre les étapes $N+1$ et $N+2$ sont *différentes*, selon que l'on souhaite calculer l'utilité pessimiste ou optimiste d'une stratégie.

Supposons que nous ayons une fonction d'utilité qualitative u sur S_{N+1}. Définissons le nouveau problème comportant une étape supplémentaire, $N+2$:
- $S_{N+2} = \{but, \overline{but}\}$,
- $But_{N+2} = \{but\}$,
- $\forall s \in S_{N+1}, A_{s,N+1} = \{a^*\}$,
- $\forall s \in S_{N+1}, \pi_{N+1}(but|s,a^*) = 1_L, \pi_{N+1}(\overline{but}|s,a^*) = n(u(s))$.

On peut montrer la proposition suivante :

Proposition 2.3.1 $\forall s \in S_{N+1}, u(s) = \mathsf{N}(But_{N+1}(s))$.

Preuve :

Par définition de $\mathsf{N}(But_{N+1}(s))$ on a :

$\mathsf{N}(But_{N+1}(s)) = \max_{a \in A_{s,N+1}} \min_{s' \in S_{N+2} - But_{N+2}} n(\pi_{N+1}(s'|s,a))$.

Donc, $\mathsf{N}(But_{N+1}(s)) = n(n(u(s))) = u(s)$.

De la même façon, en étendant l'horizon du problème initial par :
- $S_{N+2} = \{but, \overline{but}\}$,
- $But_{N+2} = \{but\}$,
- $\forall s \in S_{N+1}, A_{s,N+1} = \{a^*\}$,
- $\forall s \in S_{N+1}, \pi_{N+1}(but|s,a^*) = u(s), \pi_{N+1}(\overline{but}|s,a^*) = 1_L$.

On peut montrer que :

Proposition 2.3.2 $\forall s \in S_{N+1}, u(s) = \Pi(But_{N+1}(s))$.

Preuve :

Par définition de $\Pi(But_{N+1}(s))$ on a :

$\mathsf{N}(But_{N+1}(s)) = \max_{a \in A_{s,N+1}} \max_{s' \in But_{N+2}} \pi_{N+1}(s'|s,a)$. Donc, $\Pi(But_{N+1}(s)) = u(s)$.

Ainsi, $u(s)$ peut être obtenue, soit comme degré de nécessité soit comme degré de possibilité en considérant une étape supplémentaire et une seule action disponible à l'étape $N+1$.
Il est important de noter que les actions disponibles à l'étape $N+1$ sont différentes et les utilités pessimistes et optimistes calculées pour une action a à une étape $t < N+1$ ne sont pas des degrés de nécessité et de possibilité duaux. En effet, $\Pi(But_t(s)) < 1$ n'implique pas forcément $\mathsf{N}(But_t(s)) = 0$.

2.3.5 Raffiner le critère pessimiste par un critère optimiste

En général on choisira le critère pessimiste v_* pour discriminer les stratégies, car celui-ci est prudent alors que le critère optimiste v^* peut conduire au choix de stratégies plus risquées. Cependant, n'utiliser que v_* pour classer les stratégies peut se révéler peu discriminant. En conséquence on peut songer à utiliser le critère v^* pour raffiner v_*, d'une manière similaire à celle utilisée dans la section 2.2 [FLS98].

Définition 2.3.6 *Raffiner le critère pessimiste.*

1. $d_{t\to N} >_s d'_{t\to N}$ *ssi l'une de ces deux conditions est vérifiée :*
   - $v^t_*(s, d_{t\to N}) > v^t_*(s, d'_{t\to N})$ *ou*
   - $v^t_*(s, d_{t\to N}) = v^t_*(s, d'_{t\to N})$ *et* $v^*_t(s, d_{t\to N}) > v^*_t(s, d'_{t\to N})$.
2. $d_{t\to N}$ *est optimale ssi il n'existe aucune stratégie* $d'_{t\to N}$ *telle que* $d'_{t\to N} >_s d_{t\to N}$.

Il est alors facile de montrer que $d_{t\to N}$ est optimale ssi elle est v_*-optimale et maximise v^* parmi l'ensemble des stratégies v_*-optimales.

Puisque $v_*^t(s, d_{t\to N})$ et $v_t^*(s, d_{t\to N})$ sont respectivement la nécessité et la possibilité d'un ensemble flou, en général ils *ne vérifient pas* $v_*^t(s, d_{t\to N}) > 0_L \Rightarrow v_t^*(s, d_{t\to N}) = 1_L$, mais seulement la relation plus faible : $v_t^*(s, d_{t\to N}) \geq v_*^t(s, d_{t\to N})$. En conséquence, des actions optimales suivant v_* ne sont plus forcément optimales suivant v^*.

Exemple 2.3.1 *Soit* $T = 1$, $S_1 = \{s_0\}$, $S_2 = \{s_1, s_2, s_3\}$, $A_{s_0,1} = \{a, b, c\}$ *avec* $v_1(s_0) = 1$ *et*

$$\left|\begin{array}{ccc|c} \pi(s_1|s_0,a)=1 & \pi(s_2|s_0,b)=0 & \pi(s_3|s_0,a)=0,2 & v_2(s_1)=1 \\ \pi(s_1|s_0,b)=0,7 & \pi(s_2|s_0,b)=1 & \pi(s_3|s_0,b)=0,1 & v_2(s_2)=0,6 \\ \pi(s_1|s_0,c)=1 & \pi(s_2|s_0,b)=0 & \pi(s_3|s_0,b)=0 & v_2(s_3)=0,2 \end{array}\right|$$

Les stratégies d, d' *et* d'' *associant respectivement* a,b *et* c *à* s_0 *ont pour utilité pessimiste et optimiste :*

$$\left|\begin{array}{ll} v_*^1(d)=0,2 & v_1^*(d)=1 \\ v_*^1(d')=0,6 & v_1^*(d')=0,7 \\ v_*^1(d'')=0,6 & v_1^*(d'')=0,6 \end{array}\right|$$

Ainsi, d' *et* d'' *sont* v_**-optimales alors que* d *est* v^**-optimale.* d' *est optimale suivant le critère de choix raffiné.*

Maintenant montrons que l'on peut obtenir par calcul récursif des stratégies optimales suivant le critère raffiné. Nous gardons les mêmes définitions que précédemment pour le calcul récursif de $v_*^t(s, a)$ et $v_*^t(s)$ mais nous utilisons maintenant $\hat{v}_t^*$ à la place de v_t^* :

Définition 2.3.7 *Calcul récursif du critère optimiste raffinant le critère pessimiste.*

– $\forall s \in S_{N+1}$, $\hat{v}_{N+1}^*(s) = v_{N+1}(s)$

– $\forall s \in S_t$, $\forall a \in A_{s,t}$, $\hat{v}_t^*(s,a) = Sup_{s' \in S_{t+1}} min(\pi_t(s'|s,a), \hat{v}_{t+1}^*(s'))$

– $a >_s a'$ *ssi* $v_*^t(s,a) > v_*^t(s,a')$ *ou* ($v_*^t(s,a) = v_*^t(s,a')$ *et* $\hat{v}_t^*(s,a) > \hat{v}_t^*(s,a')$)

– $A_{s,t}^* = \{a \in A_{s,t}$ *tels qu'il n'existe pas* $a' \in A_{s,t}$ *telle que* $a' >_s a\}$

– $\forall s \in S_t$, $\hat{v}_t^*(s) = \min(u_t(s), \hat{v}_t^*(s,a^*))$ *pour* $a^* \in A_{s,t}^*$*, arbitraire.*

Ainsi, $\hat{v}_t^*(s)$ représente la possibilité de l'événement flou "atteindre un état but à partir de s en appliquant une stratégie v_*-optimale". Notons que pour tout t et s on a les inégalités suivantes : $v_t^*(s) \geq \hat{v}_t^*(s) \geq v_*^t(s)$. En général, ces inégalités sont strictes. En particulier, $\hat{v}_t^*$ ne doit pas être confondu avec v_t^* : alors que $v_t^*(s)$ représente la possibilité de l'événement flou "atteindre un état but en appliquant une stratégie optimale suivant le critère purement optimiste", $\hat{v}_t^*(s)$ représente la possibilité de l'événement flou "atteindre un état but en appliquant une stratégie optimale suivant le critère raffiné", i.e. v_*-optimale et parmi celles-ci, maximisant v^*.

Proposition 2.3.3 $\hat{v}_t^*(s) = Sup_{d \in D_t | d\ est\ v_*-optimale} v_t^*(s, d_t(s))$

Corollaire 2.3.3 *Toute stratégie calculée récursivement par induction arrière suivant la définition 2.3.7 est optimale pour le critère raffiné.*

Grâce à ces propriétés il est possible de définir encore un algorithme dans le style "programmation dynamique", permettant de calculer une stratégie optimale suivant le critère raffiné. Encore une fois, comme précédemment, il est possible de le simplifier afin d'éviter certains calculs superflus.

Algorithme 2.3 : CALCUL RÉCURSIF DE STRATÉGIES OPTIMALES SUIVANT LE CRITÈRE RAFFINÉ

Données : d, la stratégie optimale,
N, le nombre d'étapes,
S_t, les ensembles d'états possibles pour chaque étape,
$A_{s,t}$ les actions disponibles pour chaque état possible de chaque étape

début
- % Initialisations ;
- $d \leftarrow \emptyset$;
- **pour** $s \in S_{N+1}$ **faire**
  - $\hat{v}^*_{N+1}(s) = u(s)$;
  - $v^{N+1}_*(s) = u(s)$
- % Calcul récursif ;
- **pour** $t \leftarrow N$ *jusqu'à 1* **faire**
  - **pour** $s \in S_t$ **faire**
    - **pour** $a \in A_{s,t}$ **faire**
      - % Calcul de $v^t_*(s,a)$;
      - % Calcul de $\hat{v}^*_t(s,a)$
    - % a^* maximise $v^t_*(s,a)$ puis $\hat{v}^*_t(s,a)$;
    - $d_t(s) \leftarrow a^*$;
    - $v^t_*(s) \leftarrow v^t_*(s,a^*)$;
    - $\hat{v}^*_t(s) \leftarrow v^*_t(s,a^*)$
  - % Ajout à la stratégie optimale ;
  - $d \leftarrow d \cup d_t$

fin

Une solution alternative pour raffiner le critère pessimiste pourrait consister en le remplacement de l'opérateur min pour calculer la possibilité d'une trajectoire par un opérateur de *minimum lexicographique*. Dans ce cas il s'agirait de conserver pour chaque trajectoire, non pas la possibilité de transition du "chaînon le plus faible", mais toutes les possibilités de transition, rangées par ordre croissant. Alors, deux trajectoires sont comparées d'abord selon leur élément le plus faible, puis le suivant, etc...

Des méthodes de programmation dynamique ont été proposées pour résoudre ce type de problème (dont on peut montrer qu'il ressemble à des problèmes classiques de programmation dynamique de type "plus court chemin"), entre autres par Fortemps [For97].

Une autre méthode de raffinement, également applicable dans le cas de décisions non séquentielles, est suggérée dans [DP95c]. Cette méthode consiste à départager les actions v_*-optimales en utilisant les ordres lexicographiques ou *discrimin* ([FLS93], voir aussi la section 1.2.1 page 11) sur les vecteurs dont les coordonnées sont les termes $\max\{n(\pi(s)), \mu(d(s))\}$ (dans le cas non séquentiel), pour $s \in S$.

2.3.6 Exemple

Un robot doit se déplacer dans une salle représentée par la Figure 2.3. Il entre par la case en haut à gauche (de coordonnées $(1, 1)$) et son objectif est de se rendre “en bas à droite”. Son but est entièrement satisfait si il réussit à rejoindre la case $(3, 3)$ et partiellement si il rejoint l’une des deux cases adjacentes.

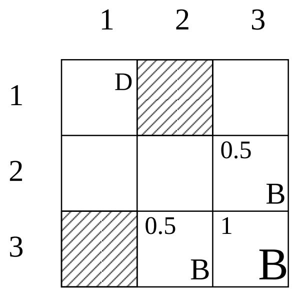

FIG. 2.3 – *Espace d’états et fonction d’utilité.*

Les actions disponibles sont (H)aut, (B)as, (G), (D)roite et (R)ester en place. Si le robot choisit de (R)ester en place, il est *certain* de demeurer sur la même case. S’il choisit H, B, G ou D, il est tout à fait possible qu’il se retrouve dans la case choisie si elle est libre mais il est également possible qu’il se retrouve dans une des cases voisines. La figure 2.4 montre les possibilités de transition pour l’action D, les autres se déterminent aisément, par symétrie.

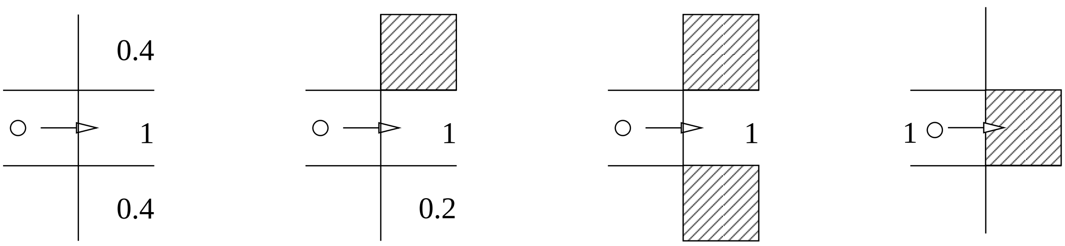

FIG. 2.4 – *Possibilités de transition pour l’action “aller à droite”.*

Supposons maintenant que l’horizon du problème soit 5, ce qui signifie que les buts sont proposés pour l’étape 6. La Figure 2.3 représente la fonction d’utilité μ, donc elle représente également v_*^6 : $v_*^6(s_{33}) = 1$, $v_*^6(s_{32}) = v_*^6(s_{23}) = 0,5$ et $v_*^6(s) = 0$ pour les autres états.
Calculons maintenant les actions optimales pour les états de S_5, i.e. la stratégie optimale (pessimiste) à l’étape 5. Pour chaque action a et chaque état s, nous avons $u_*^5(s, a) = min_{s' \in S_6} max(1 - \pi(s'|s, a), u_*^6(s'))$ (π ne dépend pas de l’étape courante) ainsi que $u_*^5(s) = max_{a \in \{T,D,L,R,S\}} u_*^5(s, a)$. La Figure 2.5 représente les utilités pessimistes de chaque état de S_5 (lorsque celle-ci n’est pas nulle) ainsi qu’une action optimale pour chacune de ces étapes. Notons que chaque état ne comporte qu’une seule action optimale associée, excepté pour l’état s_{33} pour lequel les actions D et B sont également optimales.
Il nous faut maintenant itérer le processus pour obtenir une stratégie optimale. La Figure 2.6 représente un tel processus itéré cinq fois. Notons qu’après quatre itérations l’utilité pessimiste associée à chaque état ainsi que les actions optimales correspondantes ne changent plus. De manière similaire on pourrait calculer des stratégies optimales suivant le critère optimiste ou suivant le critère raffiné.

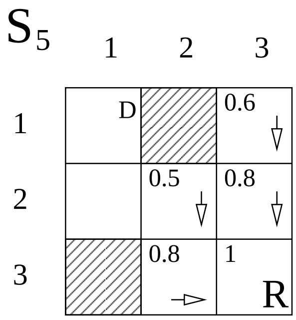

FIG. 2.5 – *Actions optimales à l'étape 5.*

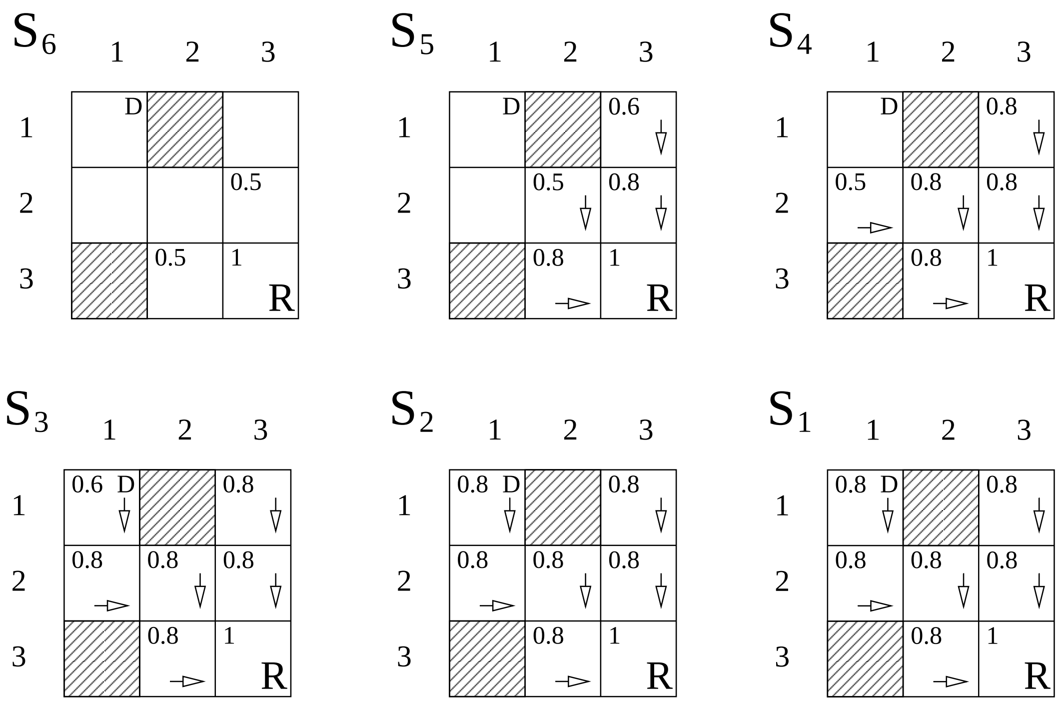

FIG. 2.6 – *Calcul récursif d'une stratégie pessimiste optimale.*

2.4 Conclusion

La principale contribution de ce chapitre consiste en une extension de la théorie de la décision qualitative possibiliste à la décision séquentielle. Deux généralisations ont été proposées (la seconde plus générale que la première), correspondant dans notre modèle aux quadruplets $< N, \pi, B$, maximiser (N,Π) $>$ et $< N, \pi, \min$, maximiser v_* puis $\hat{v}^* >$. Nous avons proposé des algorithmes, dans le style de la programmation dynamique, s'appliquant à la résolution de ces problèmes.

Une méthode alternative pour la décision séquentielle possibiliste à été explorée par Da Costa et col. [DGLM97a], [DGLM97b], [Da 98]. Ces travaux utilisent également des distributions de possibilité pour représenter les effets des actions. Ils se placent dans un cadre non observable et en conséquence, plutôt que de calculer des stratégies, ils calculent des plans inconditionnels (ou séquences d'actions), maximisant la possibilité ou la nécessité d'atteindre un état but. Ils utilisent une représentation des

actions possibilistes “à la STRIPS” qui évite une énumération complète des états, comme nous le faisons ici et proposent un algorithme de calcul de plans optimaux, plus dans le style “IA et planification” que dans le style “programmation dynamique”.
Dans le chapitre suivant nous proposerons également une méthode de représentation logique des problèmes de décision sous incertitude possibiliste et des algorithmes de résolution associés, pour le cas d’une seule étape de décision.

Plusieurs extensions de la programmation dynamique en environnement “flou” ont été proposées. Une revue de ces extensions peut être trouvée dans [EK98]. Le premier article sur ce thème est celui de Bellman et Zadeh [BZ70] dans lequel une fonction d’utilité est utilisée (et les utilités intermédiaires sont agrégées par le $minimum$) et les transitions sont soit déterministes soit stochastiques. Ces deux approches correspondent aux quadruplets $< N, D, \min, u >$ et $< N, pr, \min, \bar{u} >$ respectivement. La seconde approche est en fait mi-qualitative mi-quantitative : le critère de choix correspond à maximiser l’espérance mathématique d’une fonction d’utilité qualitative. Au contraire, notre approche est entièrement qualitative (aussi bien du côté des préférences que de l’incertitude).
D’autres travaux importants utilisent la théorie des possibilités pour la décision séquentielle, dont ceux de Kacprzyk [Kac83, Kac98] et Baldwin et Pilsworth [BP82]. Kacprzyk [Kac83, Kac98] propose de construire une stratégie maximisant l’utilité optimiste v^* grâce à un algorithme de type *branch and bound*. Il propose également des extensions de sa méthode aux cas où l’horizon du problème est infini, voire flou. L’approche alternative de Baldwin et Pilsworth [BP82], basée sur des méthodes de programmation dynamique, maximise également le critère optimiste mais la recherche est effectuée sur un ensemble flou d’états et la maximisation s’opère sur un ensemble flou de décisions, ce qui, comme le note Kacprzyk, rend l’approche d’une complexité non raisonnable.

Ce chapitre ne propose qu’un ensemble de résultats préliminaires quoique importants. Trois directions principales peuvent être envisagées pour la suite :

– Considérer la contribution des états et actions intermédiaires pour l’évaluation globale d’une stratégie : la manière la plus naturelle de le faire dans un cadre purement qualitatif consisterait à les combiner par un $minimum$. Pour v_* cela pourrait donner :

$$v_*^t(s,a) = \min\{v(s), v(a), \min_{s' \in S_{t+1}} \max\{n(\pi(s'|s,a)), v_*^{t+1}(s')\}\},$$

si on considère que les utilités sont des degrés de satisfaction ou de priorité.

Néanmoins, revenant à l’expression de base de l’utilité pessimiste d’une stratégie δ :

$$v_*(\delta) = \min_{\tau \in TRAJ} \max\{n(\pi(\tau)), \mu(\tau)\},$$

Toute lattitude est laissée pour le calcul effectif de l’utilité $\mu(\tau)$ d’une trajectoire (il en est de même pour $\pi(\tau)$. Cela signifie que l’on peut, par exemple, considérer que l’utilité d’une trajectoire est fonction d’une somme de récompenses récoltées au long de son parcours, mais que seul l’ordre entre les sommes associées à chaque treajectoire compte pour la définition de μ (afin de conserver une échelle ordinale). Dans ce cas, on obtient une classe de problèmes généralisant celui de la recherche de plus court chemin dans un arbre (on retrouve exactement ce problème si les transitions entre états sont déterministes).

– Se placer en environnement partiellement observable et poursuivre les pistes proposées dans ce chapitre pour définir une contrepartie qualitative des POMDP et des méthodes de résolution adaptées. Les méthodes envisagées pourraient être d'une complexité moindre que celles des POMDP classiques, grâce aux propriétés des opérateurs $minimum$ et $maximum$, permettant d'obtenir un espace de connaissances *fini*, à partir d'un ensemble d'états *fini*.

– Enfin, des théories qualitatives de la décision sous incertitude plus générales que celle de l'utilité possibiliste pourraient être étendues à la décision séquentielle. En particulier, la théorie basée sur l'intégrale de Sugeno pourrait être étudiée dans ce cadre.

Nous pourrions également nous intéresser à des méthodes de représentation structurée pour les problèmes de décision qualitative séquentielle : trouver une contrepartie qualitative des réseaux de décision de Zhang [Zha94] ou utiliser des méthodes de représentation logique (à la STRIPS, par exemple, comme dans [Da 98]. Ce dernier problème de représentation logique (et de résolution) de problèmes de décision sous incertitude qualitative va être abordé dans le chapitre suivant, dans le cas non séquentiel.

Chapitre 3

Une approche logique de la décision qualitative dans l'incertain basée sur les ATMS

3.1 Introduction

Nous avons exposé dans le dernier chapitre de la partie bibliographique de cette thèse un certain nombre de représentations de problèmes qualitatifs de décision sous incertitude.
Nous avons décrit de manière détaillée l'approche de Tan et Pearl [TP94] basée sur les préférences ceteris paribus de Doyle et Wellman [DW94], ainsi que celles de Boutilier [Bou94] et de Lang [Lan96].

Dans les théories de la décision sous incertitude que nous avons montrées dans les chapitres 1 et 2 de la première partie et dans le chapitre 1 de la seconde, les préférences d'un agent sont exprimées par l'intermédiaire d'une fonction d'utilité (à valeurs réelles ou dans une échelle simplement ordonnée) alors que ses connaissances sont exprimées par l'intermédiaire d'une mesure de l'incertitude : mesure de probabilité, possibilité, etc...
Cependant, il serait intéressant de proposer une expression plus compacte des préférences et des connaissances d'un agent, compatible avec une théorie "évoluée" de la décision sous incertitude, plus riche en tous cas que les théories qui sous-tendent les approches logiques décrites dans le chapitre 3 de la partie I.
Cette expression pourrait être, et c'est l'approche que nous allons proposer, sous la forme de formules logiques (propositionnelles) valuées, dont nous pourrons déduire les fonctions d'utilité et d'incertitude correspondantes. Dans ce chapitre, nous supposerons ainsi que les préférences et les connaissances d'un agent seront exprimées par l'intermédiaire de propositions valuées.

Dans la prochaine section nous allons montrer deux approches syntaxiques basées sur la *logique possibiliste*, l'une pessimiste et l'autre optimiste, pour résoudre des problèmes de décision sous incertitude. Dans un premier temps nous considérerons que les préférences et l'incertitude sont binaires (conséquences buts/non buts, états du monde possibles/impossibles), puis nous envisagerons le cas de préférences et de connaissances incertaines *graduelles*. Les connaissances et les préférences seront exprimées dans des bases stratifiées de formules logiques propositionnelles. Nous montrerons alors, en reprenant nos résultats exposés dans [DPS97b], que les deux approches syntaxiques sont respectivement en accord avec les deux fonctions d'utilité qualitative possibilistes, pessimiste et optimiste.

Ensuite, reprenant nos travaux présentés dans [DLPS98a], nous donnerons de brefs rappels sur le cadre des *Assumption-based Truth Maintenance Systems* (ATMS) de de Kleer [de 86] et nous montrerons comment un problème de décision qualitative possibiliste peut être codé dans ce cadre. Nous décrirons alors une procédure appelée MPL (pour *Modèles Préférés et Littéraux*) proposée par Castell et col. [CCCL96] et basée sur la procédure de Davis et Putnam [DP60] et nous montrerons comment elle peut être utilisée pour le calcul de décisions optimales, pessimistes ou optimistes.

Le cadre original basé sur les ATMS que nous proposons pour la décision qualitative possibiliste, peut être généralisé aux calculs de décisions optimales vis-à-vis d'autres critères de décision non-classiques. Dans la section suivante nous exposerons le lien entre ATMS et décision dans l'incertain basée sur les fonctions de croyance et de plausibilité. Ces travaux ont été présentés dans [LS97]. Nous décrirons également un cadre général basé sur les ATMS proposé dans [Sab98], permettant d'intégrer la représentation de problèmes qualitatifs de décision sous incertitude, dans lequel outre les hypothèses classiques, sont inclus des symboles de préférence et de décision.

3.2 Décision qualitative sous incertitude et bases stratifiées

3.2.1 Notations

Dans ce chapitre, les lettres majuscules $(K, D, P, H, \ldots)$ représentent des ensembles de propositions (éventuellement littérales). Pour tout ensemble de formules A, $A^\wedge$ représente la formule conjonction des propositions de A et $A^\vee$ la formule disjonction des propositions de A. Si $H = \{l_i\}$ est un ensemble de littéraux, $\sim H = \{\neg l_i, l_i \in H\}$ est l'ensemble constitué de la négation de ces littéraux.

3.2.2 Décision dans le cas binaire

Voyons comment on peut décrire un problème de décision sous incertitude dans un cadre logique.
Nous utiliserons, dans cette section et dans la suivante, un vocabulaire contenant des variables propositionnelles de deux types : variables de décision et variables d'état. Les variables de décision sont contrôlables, ce qui signifie que l'agent-décideur peut fixer leur valeur. Prendre une décision revient pour l'agent à fixer la valeur de vérité (vraie ou fausse) de toutes les variables contrôlables ou d'une partie d'entre elles seulement. La valeur de vérité des variables d'états, au contraire, est fixée par *la nature* et l'agent n'en a aucun contrôle, même s'il peut exprimer des préférences sur ces valeurs. L'ensemble des variables de décision est noté D.

K est une base de connaissances (de la logique propositionnelle) décrivant ce que l'agent *sait* du monde, ainsi que les contraintes liant éventuellement les valeurs de vérité des variables de décision. P est une autre base, de préférences, décrivant les états-buts de l'agent, i.e. ceux satisfaisant ses préférences. K et P sont supposées finies, tout comme le langage $\mathcal{L}$ considéré.
Supposons dans un premier temps que K et P sont des bases de la logique propositionnelle classique, représentant des connaissances certaines et des préférences *tout ou rien*. Un agent dont les connaissances certaines sont modélisées par K et les préférences *absolues* par P va chercher à rendre "vraies" toutes les formules de P (i.e. à satisfaire ses préférences absolues) en agissant sur les variables de décision qui contrôlent les modèles de K et P. Une *bonne* décision (d'un point de vue pessimiste) sera

une conjonction de littéraux, $d^\wedge$ ($d \subseteq D$), entraînant la satisfaction de toutes les formules de P lorsque les formules de K sont supposées vraies. Ainsi, une bonne décision $d^\wedge$ doit satisfaire :

$$K^\wedge \wedge d^\wedge \vdash P^\wedge. \tag{3.1}$$

De plus, $K^\wedge \wedge d^\wedge$ doit être cohérente, car sinon 3.1 est trivialement satisfaite.

Dans ce chapitre nous supposerons implicitement que le résultat de l'action représentée par la décision $d^\wedge$ ne modifie pas le contenu de la base K qui contiendra, par exemple, des connaissances *génériques*. Il est doit être clair que cette hypothèse n'est pas toujours vérifiée. Considérons par exemple une base de connaissances *factuelles* décrivant que dans une pièce, la porte ou la fenêtre est ouverte. Supposons que la décision de fermer la porte soit prise. On se retrouve devant un problème de *mise à jour* des connaissances, dans lequel l'action de fermer la porte ne doit pas déboucher sur la conclusion que la fenêtre est ouverte...
Dans le cas général, donc, on ne devrait pas utiliser $K^\wedge \wedge d^\wedge$, mais $K^\wedge \circ d^\wedge$, où $\circ$ est un opérateur de *mise à jour* et $K^\wedge \circ d^\wedge$ est le résultat de la mise à jour. Dans ce cas, la cohérence de $K^\wedge \circ d^\wedge$ devrait être remplacée par la cohérence de $d^\wedge$ avec la partie *factuelle* de K seulement, lorsque K contient également des connaissances *génériques*.
L'étude de l'utilisation d'opérateurs de mise à jour est laissée en dehors de cette thèse : notre seul propos ici est de faire le lien entre des approches qualitatives de la décision dans l'incertain et représentation logique. Nous nous contenterons d'une représentation peu raffinée (logique propositionnelle classique), mais nous mènerons le lien jusqu'à définir des méthodes de résolution opérationnelles.

D'un point de vue optimiste, on peut se contenter de décisions qui sont simplement cohérentes avec les connaissances et l'obtention des buts, soit telles que :

$$K^\wedge \wedge d^\wedge \wedge P^\wedge \neq \bot. \tag{3.2}$$

Ce point de vue est optimiste dans le sens qu'il suppose que les préférences sont satisfaites dès lors qu'elles sont cohérentes avec la décision prise et les connaissances, i.e. dès que leurs négations ne peuvent être prouvées.

Ces deux modes de décision présentent une certaine ressemblance avec les deux modes de diagnostic basés sur les modèles (voir par exemple Hamscher et col. [HCd92]) : *diagnostic abductif* et *diagnostic basé sur la cohérence*.
Il est alors tentant de coder l'expression logique d'un problème de décision sous incertitude dans le cadre des *Systèmes de maintien de la cohérence basés sur les hypothèses* (dont l'abréviation anglaise est ATMS) de de Kleer [de 86]. Pour mieux s'adapter au cadre des ATMS et utiliser les outils de résolution correspondants, nous devrons changer le codage des décisions (comme nous le verrons dans la section suivante), de conjonctions de littéraux en ensembles de littéraux positifs.

3.2.3 Décision dans le cas stratifié

Lorsqu'un problème de décision sous incertitude est posé sous forme logique, les propositions de la base de connaissances peuvent être entachées d'incertitude et les préférences peuvent ne pas être d'égale priorité. Dans la théorie de l'utilité espérée, l'incertitude est modélisée par une distribution de probabilité sur les états possibles du monde, les préférences par une fonction d'utilité sur les conséquences et l'agent cherche une décision maximisant l'espérance mathématique de la fonction d'utilité.

Nous allons enrichir notre cadre logique en affectant des degrés de *certitude* aux formules de la base de connaissance et des niveaux de priorité aux formules de la base de préférences. Nous obtenons ainsi deux bases stratifiées, aptes à modéliser des connaissances et des préférences graduelles. Dubois Lang et Prade [DLP94] ont montré que la sémantique d'une base de la logique possibiliste (i.e. une base stratifiée dont les formules sont rassemblées en strates d'égal niveau de certitude ou de préférence) peut être codée par une distribution de possibilité.
Dans ce paragraphe, nous allons d'abord voir comment un problème de décision sous incertitude peut être exprimé à l'aide de bases de connaissances et de préférences stratifiées, puis nous montrerons que la sémantique correspondant à cette expression logique peut être représentée par les fonctions d'utilité qualitative possibiliste de Dubois et Prade [DP95b].

Dans tout ce chapitre nous supposerons encore que les degrés de certitude et de priorité sont commensurables et appartiennent à la même échelle simplement ordonnée, L, finie comme le langage $\mathcal{L}$. Le plus grand élément de L sera en général noté $\mathbf{1}$ et le plus petit, $\mathbf{0}$, à moins que nous mentionnions explicitement d'autres notations. Les connaissances et les préférences sont rassemblées dans deux bases distinctes. La base de connaissances est $K = \{(\phi_i, \alpha_i)\}$ où $\alpha_i \in L$ ($\alpha_i > \mathbf{0}$) est un degré de certitude et les ϕ_i sont des formules de $\mathcal{L}$ dans lesquelles des littéraux de décision peuvent apparaître. La base de préférences est $P = \{(\psi_i, \beta_i)\}$, où $\beta_i \in L$ ($\beta_i > \mathbf{0}$) est un degré de priorité et les ψ_i sont également des formules de $\mathcal{L}$ dans lesquelles des littéraux de décision peuvent apparaître.

Avant de poursuivre sur les aspects "techniques" de la décision sous incertitude avec des bases stratifiées, posons-nous la question du "sens" donné aux différents niveaux de préférence ou de certitude attachés aux propositions de ces bases. Clairement, il revient à l'agent d'exprimer l'ordre de *préférence* entre les propositions exprimant ses souhaits. Les différents niveaux de *certitude*, eux, peuvent être révélés par un agent unique classant les différentes propositions par ordre de plausibilité. Les connaissances peuvent aussi être issues de plusieurs sources. Dans ce cas, les sources peuvent avoir des niveaux de fiabilité (éventuellement différents) et les connaissances doivent être ordonnées selon ces différents niveaux de fiabilité. D'un autre côté, les sources peuvent être d'égale fiabilité, mais peuvent exprimer chacune un ordre de "plausibilité" sur les différentes propositions. Dans ce cas, il nous faut supposer qu'il existe un accord entre les différentes sources sur une signification commune des différents niveaux de certitude, afin de pouvoir fusionner les ordres de ces sources. Par ailleurs, on peut tenir compte des niveaux de spécificité des différentes connaissances exprimées (comme dans les approches de Tan et Pearl ou Boutilier, décrites dans le chapitre 3 de la partie I) et utiliser par exemple la méthode System Z de Pearl [Pea90] pour ordonner les connaissances (voir aussi (Benferhat, Dubois et Prade [BDP92])).

Revenons à l'expression de problèmes de décision sous incertitude avec des bases stratifiées. Soit K_α (resp. P_β) l'ensemble des formules de degré de certitude (resp. de priorité) supérieur ou égal à α (resp. β). Notons que seules les connaissances ou les préférences de degré de certitude ou de préférence strictement supérieur à $\mathbf{0}$ sont à prendre en compte puisque $K_{\mathbf{0}} = P_{\mathbf{0}} = \mathcal{L}$. Dans la suite on utilisera les notations $K_{\bar{\alpha}}$ et $P_{\bar{\beta}}$ (avec $\alpha < \mathbf{1}$ et $\beta < \mathbf{1}$) pour représenter les ensembles de formules de degré de certitude *strictement* supérieur à α ou de priorité *strictement* supérieur à β. En particulier, $K_{\bar{\mathbf{0}}} = K^*$ et $P_{\bar{\mathbf{0}}} = P^*$ où K^* et P^* sont respectivement les ensembles de toutes les formules de K ou de P, sans les degrés qui y sont attachés. Remarquons enfin que puisque l'échelle L est finie, $K_{\bar{\alpha}} = K_{\alpha'}$ et $P_{\bar{\beta}} = P_{\beta'}$, où α' est le degré de L immédiatement supérieur à α et β' le degré immédiatement supérieur à β.

Prendre une décision revient à choisir un sous-ensemble d de l'ensemble des décisions disponibles, $D = \{l_i\}$, dans lequel les l_i sont des variables spéciales du langage $\mathcal{L}$. La représentation logique de cette décision est la conjonction $d^\wedge$ des variables de d. Si l'ensemble des variables de décision est D, les variables qui sont hors de D sont appelées *variables d'état*.
Notre objectif est d'ordonner les décisions, et pour ce faire nous utiliserons une fonction d'utilité *syntaxique* $U : \mathcal{P}(D) \to S$ telle que $d \preceq d' \Leftrightarrow U(d) \leq U(d')$. Dans la suite nous proposerons deux fonctions d'utilité syntaxiques différentes : Une fonction U_* qui est en accord avec une vision prudente (ou pessimiste) du problème de décision sous incertitude et une fonction aventureuse (ou optimiste), U^*.

Dans le premier cas (prudent), nous cherchons une décision d (lorsqu'elle existe) telle que :

$$K_\alpha^\wedge \wedge d^\wedge \vdash P_\beta^\wedge \tag{3.3}$$

avec α élevé et β faible, c'est-à-dire que nous exigeons de la décision d, qu'avec la partie la plus certaine de K elle induise la satisfaction de toutes les préférences de P, y compris les moins prioritaires. Nous supposons bien entendu implicitement que d est cohérente avec K_α, pour les α satisfaisant 3.3. Un moyen d'assurer cette cohérence est de supposer que $K^* \cup d$ est cohérent, soit que les toutes les connaissances un tant soit peu certaines sont cohérentes entre elles et autorisent la décision d. Par convention, nous attribuerons une utilité $\mathbf{0}$ à toutes les décisions qui ne sont pas cohérentes avec K^*, ce qui éliminera les décisions non autorisées. Par ailleurs, notons que les β satisfaisant 3.3 sont forcément strictement supérieurs à $\mathbf{0}$ puisque $P_{\mathbf{0}}^\wedge = \mathcal{L}$ est incohérent.

Idéalement, d, avec uniquement la partie la plus certaine de K ($K_{\mathbf{1}}$) doit permettre de satisfaire toutes les contraintes (ou préférences) de P, de priorité non nulle ($P_{\overline{\mathbf{0}}}$). Les pires décisions sont celles qui, avec toutes les connaissances (même les moins certaines) ($K_{\overline{\mathbf{0}}}$), ne parviennent même pas à inférer les préférences les plus prioritaires ($P_{\mathbf{1}}$). De telles décisions doivent avoir $\mathbf{0}$ pour utilité.
Posons n, la fonction de renversement de L : si $L = \mathbf{0} = \alpha_0 < \ldots < \alpha_i < \ldots < \alpha_n = \mathbf{1}$ alors $n(\alpha_i) = \alpha_{n-i}$. Supposons que d soit telle que $K_\alpha^\wedge \wedge d^\wedge \vdash P_\beta^\wedge$, avec $\beta < n(\alpha)$. Alors nous devons avoir $K_\alpha^\wedge \wedge d^\wedge \vdash P_{\overline{n(\alpha)}}^\wedge$ parce que $P_{\overline{n(\alpha)}} \subseteq P_\beta$, donc $P_\beta \vdash P_{\overline{n(\alpha)}}$.
Réciproquement, si $\beta > n(\alpha)$ alors $n^{-1}(\beta) < \alpha$, donc puisque $K_\alpha \subseteq K_{n^{-1}(\beta)}$ et $P_{\overline{\beta}} \subseteq P_\beta$ et $K_\alpha^\wedge \wedge d^\wedge \vdash P_\beta^\wedge$, nous obtenons $K_{n^{-1}(\beta)}^\wedge \wedge d^\wedge \vdash P_{\overline{\beta}}^\wedge$. En posant $\beta = n(\alpha)$, nous obtenons $K_\alpha^\wedge \wedge d^\wedge \vdash P_{\overline{n(\alpha)}}^\wedge$.
Donc on peut supposer que $\beta = n(\alpha)$ dans le problème de maximisation qui devient : maximiser α tel que $K_\alpha^\wedge \wedge d^\wedge \vdash P_{\overline{n(\alpha)}}^\wedge$, avec $\alpha > \mathbf{0}$.

Finalement, l'utilité prudente d'une décision d, définie au niveau syntaxique, prend la forme suivante :

Définition 3.2.1

$$U_*(d) = \max_{\alpha / K_\alpha^\wedge \wedge d^\wedge \vdash P_{\overline{n(\alpha)}}^\wedge,\ K_\alpha^\wedge \wedge d^\wedge \neq \perp} \alpha.$$

et $U_*(d) = \mathbf{0}$ *si* $\{\alpha > \mathbf{0}, K_\alpha^\wedge \wedge d^\wedge \vdash P_{\overline{n(\alpha)}}^\wedge$ *et* $K_\alpha^\wedge \wedge d^\wedge \neq \perp\} = \emptyset$.

Si on considère maintenant le cas optimiste, on cherche une décision d telle que :

$$K_\alpha^\wedge \wedge d^\wedge \wedge P_\beta^\wedge \neq \perp \tag{3.4}$$

avec α et β aussi petits que possible (α et β doivent être strictement supérieurs à $\mathbf{0}$). En quelque sorte, on cherche une décision d qui soit cohérente avec les connaissances les plus certaines et avec la satisfaction des préférences les plus prioritaires. Du point de vue syntaxique, l'utilité d'une décision d est définie par :

Définition 3.2.2

$$U^*(d) = \max_{K^{\wedge}_{\overline{\alpha}} \wedge d^{\wedge} \wedge P^{\wedge}_{\overline{\alpha}} \neq \bot} n(\alpha),$$

et $U^*(d) = \mathbf{0}$ *si* $\{\alpha < \mathbf{1}\ , K^{\wedge}_{\overline{\alpha}} \wedge d^{\wedge} \wedge P^{\wedge}_{\overline{\alpha}} \neq \bot\} = \emptyset$.

Notons que $U^*(d) = \mathbf{1}$ ssi $K^{\wedge}_{\overline{\mathbf{0}}} \wedge d^{\wedge} \wedge P^{\wedge}_{\overline{\mathbf{0}}} \neq \bot$, c'est-à-dire si la décision est cohérente avec l'ensemble des connaissances de degré de certitude non nul et des préférences de priorité non nulle. U^* représente un point de vue optimiste dans la mesure où $U^*(d) = \mathbf{1}$ dès que d est cohérente avec K et P (on se place d'office dans le meilleur cas).

3.2.4 Sémantique possibiliste de l'expression en logique stratifiée des problèmes de décision sous incertitude

Voyons maintenant quelle est la sémantique de cette expression logique des problèmes de décision sous incertitude. Nous allons interpréter les α_i attachés aux strates de K comme les degrés de nécessité des formules des strates correspondantes de $K \cup d$. A partir de cela on peut calculer une distribution de possibilité π_{K_d} sur Ω, l'ensemble des interprétations de $\mathcal{L}$, exprimant la sémantique de K (voir par exemple [DLP94]) :
$\forall \omega \in \Omega$, $\pi_{K_d}(\omega) = \min_{(\phi_i,\alpha_i) \in K / \omega \models \neg\phi_i} n(\alpha_i)$ si $\omega \models d^{\wedge}$ et $\pi_{K_d}(\omega) = \mathbf{1}$ si $\{\phi_i / \omega \models \neg\phi_i\} = \emptyset$ et $\pi_{K_d}(\omega) = \mathbf{0}$ si $\omega \not\models d^{\wedge}$.

La distribution de possibilité π_{K_d} ordonne les interprétations suivant leur niveau de possibilité, induit par les degrés de certitude des formules de K. Cette sémantique est en accord avec l'idée qu'une interprétation ω est d'autant moins possible qu'elle viole des formules de degré de certitude élevé. Notons que puisque $K^{\wedge}_{\overline{\mathbf{0}}} \wedge d^{\wedge}$ est cohérent, π_{K_d} est normalisée, c'est-à-dire qu'il existe une interprétation ω telle que $\pi(\omega) = \mathbf{1}$.

En interprétant les β_i attachés aux différentes strates de P comme les degrés de priorité des formules de P, on peut construire de la même manière une fonction d'utilité μ sur Ω (ω est d'autant plus "satisfaisante" qu'elle viole moins de contraintes à forte priorité) :

$$\mu(\omega) = \min_{(\psi_j,\beta_j) \in P, \omega \models \neg\psi_j} n(\beta_j).$$

et $\mu(\omega) = \mathbf{1}$ si $\{\psi_j / \omega \models \neg\psi_j\} = \emptyset$.

Comme nous l'avons montré dans [DPS97b], les deux fonctions d'utilité U_* et U^* que nous venons de définir "syntaxiquement" peuvent s'exprimer en fonction de la distribution de possibilité π_{K_d} et de la fonction d'utilité μ :

Théorème 3.2.1 *Expression sémantique de* $U_*(d)$.

$$U_*(d) = \max_{\alpha / K^{\wedge}_{\overline{\alpha}} \wedge d^{\wedge} \vdash P^{\wedge}_{\overline{n(\alpha)}}} \alpha = \min_{\omega \in \Omega} \max(n(\pi_{K_d}(\omega)), \mu(\omega)).$$

Preuve : (pour une échelle L finie)

- $\forall \omega / \omega \models d^{\wedge}$, $(\omega \models K_\alpha \cup d) \Leftrightarrow (\forall (\phi_i, \alpha_i) \in K, \alpha_i \geq \alpha \Rightarrow \omega \models \phi_i)$
$\Leftrightarrow (\forall (\phi_i, \alpha_i) \in K, \omega \models \neg\phi_i \Rightarrow \alpha_i < \alpha) \Leftrightarrow (\min_{(\phi_i,\alpha_i) \in K / \omega \models \neg \phi_i} n(\alpha_i) > n(\alpha))$
$\Leftrightarrow \pi_{K_d}(\omega) > n(\alpha)$.

- De la même manière on peut montrer que $(\omega \models P_{\overline{n(\alpha)}} \Leftrightarrow (\mu(\omega) \geq \alpha)$

- Utilisons ces résultats dans la suite :
$\forall \alpha / \alpha > \mathbb{0}$, $(K_\alpha^{\wedge} \wedge d^{\wedge} \vdash P_{\overline{n(\alpha)}}^{\wedge}) \Leftrightarrow (\forall \omega, \omega \models K_\alpha \cup d \Rightarrow \omega \models P_{\overline{n(\alpha)}})$
$\Leftrightarrow (\forall \omega, \pi_{K_d}(\omega) > n(\alpha) \Rightarrow \mu(\omega) \geq \alpha) \Leftrightarrow (\forall \omega, n(\pi_{K_d}(\omega)) < \alpha \Rightarrow \mu(\omega) \geq \alpha))$
$\Leftrightarrow (\forall \omega, \max(n(\pi_{K_d}(\omega)), \mu(\omega)) \geq \alpha) \Leftrightarrow (\min_{\omega \in \Omega} \max(n(\pi_{K_d}(\omega)), \mu(\omega)) \geq \alpha)$.

- Ainsi, nous avons montré que

$\forall \alpha > \mathbb{0}, (K_\alpha^{\wedge} \wedge d^{\wedge} \vdash P_{\overline{n(\alpha)}}^{\wedge}) \Leftrightarrow (\min_{\omega \in \Omega} \max(n(\pi_{K_d}(\omega)), \mu(\omega)) \geq \alpha)$.

- Il devient alors facile de montrer que
$\min_{\omega \in \Omega} \max(n(\pi_{K_d}(\omega)), \mu(\omega)) \geq \max_{\alpha / K_\alpha^{\wedge} \wedge d^{\wedge} \vdash P_{\overline{n(\alpha)}}^{\wedge}} \alpha$, en passant à la limite.

- L'inégalité réciproque peut être montrée *par l'absurde* en supposant que

$\min_{\omega \in \Omega} \max(n(\pi_{K_d}(\omega)), \mu(\omega)) = \beta > \max_{\alpha / K_\alpha^{\wedge} \wedge d^{\wedge} \vdash P_{\overline{n(\alpha)}}^{\wedge}} \alpha$.
Alors, puisque $\min_{\omega \in \Omega} \max(n(\pi_{K_d}(\omega)), \mu(\omega)) \geq \beta \Rightarrow K_\beta^{\wedge} \wedge d^{\wedge} \vdash P_{\overline{n(\beta)}}^{\wedge}$, nous obtenons une contradiction avec l'hypothèse de départ, ce qui prouve le résultat.

Ce résultat est proche d'un résultat prouvé par Prade [Pra82], exprimant la nécessité d'un ensemble flou en fonction d'alpha-coupes, dans le cas d'une échelle continue [0,1], en remarquant que l'expression sémantique contenue dans le théorème 3.2.1 exprime la nécessité d'un événement flou.

Un théorème similaire peut être montré dans le cas optimiste :

Théorème 3.2.2 *Expression sémantique de $U^*(d)$.*

$$U^*(d) = \max_{\alpha / K_{\overline{\alpha}}^{\wedge} \wedge d^{\wedge} \wedge P_{\overline{\alpha}}^{\wedge} \neq \bot} n(\alpha) = \max_{\omega \in \Omega} \min(\pi_{K_d}(\omega), \mu(\omega)).$$

Preuve : (pour une échelle finie)

- $\forall \alpha / K_{\overline{\alpha}}^{\wedge} \wedge d^{\wedge} \wedge P_{\overline{\alpha}}^{\wedge} \neq \bot, \exists \omega^* / \omega^* \models K_{\overline{\alpha}} \cup d \cup P_{\overline{\alpha}}$

- $\omega^* \models K_{\overline{\alpha}} \cup d \Leftrightarrow \forall (\phi_i, \alpha_i) \in K, \alpha_i > \alpha \Rightarrow \omega^* \models \phi_i$
$\Leftrightarrow \forall (\phi_i, \alpha_i) \in K, \omega^* \models \neg\phi_i \Rightarrow \alpha_i \leq \alpha \Leftrightarrow \min_{i, \omega^* \models \neg\phi_i} n(\alpha_i) \geq n(\alpha)$
$\Leftrightarrow \pi_{K_d}(\omega^*) \geq n(\alpha)$

- de la même manière on peut montrer que $\omega^* \models P_{\overline{\alpha}} \Leftrightarrow \mu(\omega^*) \geq n(\alpha)$.

Ainsi, on a montré que $\exists \omega^* / \omega^* \models K_{\overline{\alpha}} \cup d \cup P_{\overline{\alpha}} \Leftrightarrow \min(\pi_{K_d}(\omega^*), \mu(\omega^*)) \geq n(\alpha)$, c'est à dire, $\forall \alpha / K_{\overline{\alpha}}^{\wedge} \wedge d^{\wedge} \wedge P_{\overline{\alpha}}^{\wedge} \neq \bot$, $\max_{\omega \in \Omega} \min(\pi_{K_d}(\omega), \mu(\omega)) \geq n(\alpha)$

- par passage à la limite, on montre $\max_{\alpha / K_{\overline{\alpha}}^{\wedge} \wedge d^{\wedge} \wedge P_{\overline{\alpha}}^{\wedge} \neq \bot} n(\alpha) \leq \max_{\omega \in \Omega} \min(\pi_{K_d}(\omega), \mu(\omega))$

- L'inégalité réciproque peut être montrée par l'absurde de la même manière que dans la preuve du théorème 3.2.1 en supposant que $\max_{\alpha / K_{\overline{\alpha}}^{\wedge} \wedge d^{\wedge} \wedge P_{\overline{\alpha}}^{\wedge} \neq \bot} n(\alpha) < \max_{\omega \in \Omega} \min(\pi_{K_d}(\omega), \mu(\omega))$

Comme on pouvait s'y attendre, les expressions sémantiques des fonctions d'utilité syntaxiques U_* et U^* sont exactement les fonctions d'utilité qualitative possibilistes v_* et v^* telle que nous les avons décrites en détails dans le chapitre 1 de cette partie.

Nous avons donc maintenant à notre disposition un moyen d'expression logique des problèmes de décision sous incertitude possibiliste. Dans la prochaine section de ce chapitre nous allons proposer des méthodes et algorithmes de résolution pour ces nouveaux problèmes logiques, qui ne requièrent pas de repasser à une expression sémantique et de résoudre un problème d'optimisation. Nous avons proposé ces méthodes et algorithmes basés sur une procédure d'évaluation sémantique de type Davis et Putman décrite dans [CCCL96], dans [DLPS98a] et [DLPS98b].

3.3 Représentation logique et calcul de solutions optimales pour des problèmes de décision sous incertitude possibiliste

Nous allons présenter dans cette section une méthode de calcul de décisions possibilistes optimales, proposée dans [DLPS98b]. Nous décrirons des algorithmes basés sur l'utilisation de la procédure MPL (pour *Modèles préférés et littéraux*), initialement donnée dans [CCCL96].
La procédure MPL est donnée en annexe A page 159 (pour une description complète, voir [CCCL96]) et nous montrerons comment calculer les décisions cohérentes avec K et P en utilisant un seul appel à MPL, alors qu'en effectuant deux appels à cette même procédure on peut calculer le "label" de P.
Avant de passer à la procédure MPL nous allons donner quelques notions sur le cadre des ATMS et son application à la décision dans l'incertain qualitatif.

3.3.1 ATMS et décisions

Les ATMS ont été introduits par de Kleer [de 86]. Considérons un ensemble $\mathcal{S}$ de symboles propositionnels, séparé en deux sous-ensembles : $\mathcal{H}$ un ensemble de *symboles hypothèses* et $\mathcal{NH}$, un ensemble de *symboles non-hypothèses*. Un ensemble d'hypothèses est appelé *environnement*. Un environnement E est dit *incohérent* pour une base de clauses K si et seulement si $K^\wedge \wedge E^\wedge \vdash \perp$. Un environnement est *cohérent* si et seulement si il n'est pas incohérent.

Définition 3.3.1 *Un "nogood" est un environnement incohérent minimal pour l'inclusion ensembliste (i.e. E est un nogood ssi $K^\wedge \wedge E^\wedge \vdash \perp$ et $\nexists E' \subset E / K^\wedge \wedge E'^\wedge \vdash \perp$).*

Définition 3.3.2 *Le* label *d'une formule ψ, noté $label_K(\psi)$, est l'ensemble des environnements E_i cohérents, minimaux pour l'inclusion ensembliste tels que $K^\wedge \wedge E_i^\wedge \vdash \psi$*[1].

Exemple 3.3.1 *Soit $K = \{A\ b \to c,\ B \to b,\ c \to\}$. $\mathcal{S} = \{A, B, b, c\}$. $\mathcal{H} = \{A, B\}$. $\{\{A, B\}\}$ est l'ensemble des nogoods de K. $label_K(b)$=$\{\{B\}\}$, $label_K(c)$=$\{\}$, $label_K(A)$=$\{\{A\}\}$, $label_K(B)$=$\{\{B\}\}$.*

Les hypothèses de l'ATMS peuvent être utiles dans le calcul de décisions optimales. Les actions seront modélisées par des ensembles de littéraux positifs de D, cohérents avec K. Ainsi, D jouera le rôle de l'ensemble d'hypothèses $\mathcal{H}$ d'un ATMS. Lorsqu'une action (ou décision) d contient un ensemble de littéraux positifs, d ne doit pas être interprétée comme une séquence d'actions, mais comme une

1. Ces définitions diffèrent légèrement de celles proposées par de Kleer. La notion de label, par exemple, était originellement définie par rapport à un littéral.

action unique, consistant en l'affectation de valeurs de vérité positives aux littéraux de d. L'ensemble des décisions qui satisfont la propriété 3.1 et dont aucun sous ensemble strict ne satisfait 3.1 (lorsque les décisions sont vues en termes d'ensembles) peut être vu comme une extension de la notion de label d'un littéral au sens de de Kleer [de 86] au label d'une conjonction de formules, $label_K(P^\wedge)$.

Exemple 3.3.2 *Les actions disponibles sont d'acheter zéro, un ou deux objets. L'ensemble de décision est* $D = \{Zero, Un, Deux\}$ *et* K *doit contenir au moins les contraintes exprimant le caractère exclusif des différentes actions :* $\{Zero \vee Un \vee Deux, \neg Zero \vee \neg Un, \neg Zero \vee \neg Deux, \neg Un \vee \neg Deux\} \subseteq K$.
Dans cet exemple, les décisions cohérentes avec les contraintes sont : $d_1 = \{Zero\}$, $d_2 = \{Un\}$, $d_3 = \{Deux\}$.

Nous allons traduire un problème de décision qualitative pessimiste binaire dans le langage des ATMS. Pour cela, on définit l'ensemble des variables hypothèses : $\mathcal{H} = D$. Supposons que K soit la base de connaissances du problème de décision, exprimée sous une forme normale conjonctive et considérons la base de préférences P sous la forme d'une formule $P^\wedge$. Une décision d est un sous-ensemble de $\mathcal{H} = D$. Pour toute décision d telle que $K^\wedge \wedge d^\wedge \vdash P^\wedge$ et $K^\wedge \wedge d^\wedge \neq \perp$ il existe au moins un élément E de $label_K(P^\wedge)$, sous ensemble de $\mathcal{H}$ tel que $E \subseteq d$.

Proposition 3.3.1

$$K^\wedge \wedge d^\wedge \vdash P^\wedge \; et \; K^\wedge \wedge d^\wedge \neq \perp \; si \; et \; seulement \; si \; \exists E \in label_K(P^\wedge) \; t.q. \; E \subseteq d$$

Preuve : immédiate en considérant la définition d'un label.

Une bonne action (au sens pessimiste) pour un problème de décision sous incertitude binaire est donc un sur-ensemble cohérent d'un élément de $label_K(P^\wedge)$. Dans la suite nous nous intéresserons seulement aux décisions minimales pour l'inclusion ensembliste, i.e. les éléments de $label_K(P^\wedge)$.
Soit $K^\wedge = \phi_1 \wedge \phi_2 \ldots \wedge \phi_n$ et $P^\wedge = \psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_m$. Trouver toutes les décisions d maximisant α tel que : $K^\wedge_\alpha \wedge d^\wedge \vdash P^\wedge_{\overline{n(\alpha)}}$ et $K^\wedge_\alpha \wedge d^\wedge \neq \perp$ est équivalent à trouver $label_{K_\alpha}(P_{\overline{n(\alpha)}}) \neq \{\}$ maximisant α.

3.3.2 Calcul de décisions optimales, optimistes et pessimistes, à l'aide de la procédure MPL

Nous utilisons la procédure MPL introduite dans [CCCL96] pour le calcul de labels. L'objectif de cette procédure est le suivant : Étant donné une formule ϕ, exprimée sous forme normale conjonctive (FNC) et comportant deux types distincts de littéraux, MPL calcule une projection de ϕ sur l'ensemble des littéraux d'un des deux types. Cette projection est une formule sous forme normale disjonctive (FND) qui est une représentation minimale de l'ensemble des modèles de ϕ, restreints à ce type de littéral. On peut montrer que les nogoods d'une base de connaissances (au sens des ATMS) peuvent être calculés facilement par cette procédure.

L'algorithme MPL est basé sur une procédure énumérative du type Davis et Putnam [DP60]. Le problème du calcul de la satisfaisabilité d'un ensemble de clauses (SAT) est un problème NP-complet. Parmi les algorithmes complets qui résolvent le problème SAT, certains parmi les plus efficaces, utilisent un algorithme du type Davis et Putnam [DP60], [DLL62].
Tous les éléments de base d'un ATMS : labels, nogoods, etc. peuvent être calculés par l'algorithme MPL et ce sans utiliser la phase de minimisation présente dans les algorithmes de de Kleer [de 86]

puisque les modèles calculés par MPL sont minimaux.

Nous décrivons sommairement la procédure MPL, ainsi que son application au calcul des différents éléments d'un ATMS dans l'annexe A page 159. Pour une description plus détaillée, on se réferera à [CCCL96] ou [CCCL98].

Décisions optimistes

L'application de la procédure MPL au calcul de décisions optimistes est immédiate. Soient K et P les bases (mises sous forme normale conjonctive) représentant les connaissances et les préférences d'un agent. Pour toute décision d telle que $K^{\wedge} \wedge d^{\wedge} \wedge P^{\wedge} \neq \bot$ il existe au moins un élément E de $MPL(K \cup P, \{\}, D)$ tel que $E \subseteq d$.

Proposition 3.3.2

$$K^{\wedge} \wedge d^{\wedge} \wedge P^{\wedge} \neq \bot \equiv \exists E \in MPL(K \cup P, \{\}, D)\ t.q.\ E \subseteq d$$

La preuve est immédiate en partant de la définition des modèles D-restreints minimaux (voir la définition A.1.3 page 159).

Une bonne décision (optimiste) est donc un sur-ensemble cohérent d'un élément de $MPL(K \cup P, \{\}, D)$. Dans la suite nous nous limiterons, parmi les bonnes décisions, à celles qui sont minimales pour l'inclusion ensembliste.
Soit $K^{\wedge} = \phi_1 \wedge \phi_2 \ldots \wedge \phi_n$ et $P^{\wedge} = \psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_m$. Trouver d maximisant $U^*(d) = n(\alpha)$ tel que : $K^{\wedge}_{\underline{\alpha}} \wedge d^{\wedge} \wedge P^{\wedge}_{\underline{\alpha}} \neq \bot$ (voir définition 3.2.2) est équivalent à trouver $MPL(K_{\overline{\alpha}} \cup P_{\overline{\alpha}}, \{\}, D) \neq \{\}$ minimisant α.
Cette méthode à une restriction : $K^{\wedge} \wedge P^{\wedge}$ doit être sous forme normale conjonctive (FNC), donc $P^{\wedge}$ doit être sous FNC, ce qui signifie que P ne doit contenir que des clauses.

L'algorithme suivant calcule les meilleures décisions, étant données K et P, en accord avec le critère de l'utilité qualitative optimiste :

Algorithme 3.1 : DÉCISIONS_OPTIMISTES

% algorithme simple calculant les décisions optimistes optimales ;
Données : K base de connaissances stratifiée, P base de préférences graduelles, D l'ensemble des symboles de décision.
Résultat : - α^* tel que $n(\alpha^*)$ est l'utilité de la meilleure décision optimiste,
- $\mathcal{D}$ l'ensemble des décisions optimistes optimales.

début
  $\alpha \leftarrow 0$ % a priori on utilise toutes les connaissances ;
  $\mathcal{D} \leftarrow MPL(K_{\overline{\alpha}} \cup P_{\overline{\alpha}}, \{\}, D)$;
  tant que $\mathcal{D} = \{\}$ *et* $\alpha < \mathbf{1}$ **faire**
    $Inc(\alpha)$ % On élimine la strate la moins certaine de K et la moins prioritaire de P ;
    $\mathcal{D} \leftarrow MPL(K_{\overline{\alpha}} \cup P_{\overline{\alpha}}, \{\}, D)$;
  retourner $< n(\alpha), \mathcal{D} >$;
fin

Décisions pessimistes

Les nogoods et labels d'un ATMS en utilisant un ou deux appels à la procédure MPL (voir l'annexe A page 159). L'avantage principal de l'utilisation de la procédure MPL réside dans le fait qu'elle permet de calculer le label d'un littéral sans avoir à calculer celui de tous les autres, comme dans la technique de de Kleer. De plus, un ATMS basé sur MPL peut être appliqué à n'importe quel ensemble de clauses (représenté par une formule sous FNC) et permet de calculer de la même manière le label d'un littéral, d'une disjonction ou d'une conjonction de littéraux. Le calcul d'un label présuppose le calcul des nogoods afin d'enlever du label les environnements incohérents. Le calcul des nogoods et des labels s'effectue de la même manière, en partant de $K^\wedge$ pour les nogoods et de $K^\wedge \wedge \neg\phi$ pour le calcul du label de ϕ.

Algorithme 3.2 : DÉCISIONS_PESSIMISTES

% algorithme basé sur la procédure MPL ;
% Soit $P_{(\beta)} = P_\beta - P_{\overline{\beta}}$. $P_{(\beta)}$ est l'ensemble des préférences de priorité β;
% $\underline{\alpha}$ représente le niveau (de certitude ou de priorité);
% de la prochaine strate non vide sous α ;
Données : K la base de connaissances,
P la base de préférences,
D l'ensemble des symboles de décision.
Résultat : $\alpha*$ l'utilité de la meilleure décision pessimiste,
$\mathcal{D}$ l'ensemble des meilleures décisions pessimistes.
début
$\alpha \leftarrow \mathbf{1}$ % on commence par la strate la plus certaine ;
$S \leftarrow \{\}$;
tant que $S = \{\}$ *et* $\alpha > \mathbf{0}$ **faire**
% on commence par calculer les nogoods de K_α ;
$NG' \leftarrow MPL(K_\alpha, \{\}, \sim D)$ % premier appel à MPL ;
$NG \leftarrow\sim MPL(\neg NG', \{\}, D)$ % NG contient les nogoods de K_α ;
$S' \leftarrow MPL(K_\alpha \cup \sim P_{\mathbf{1}}, \{\}, \sim D)$ % ;
$S \leftarrow\sim MPL(\neg S', \sim NG, D)$ % S contient le label de $P_{\mathbf{1}}$;
$\beta \leftarrow \mathbf{1}$;
tant que $\beta > n(\alpha)\ et\ S \neq \{\}$ **faire**
$S' \leftarrow S' \wedge MPL(K_\alpha \cup \sim P_{(\beta)}, \{\}, \sim D)$;
% S contient le label de P_β ;
$S \leftarrow MPL(\neg S', \sim NG, D)$;
$Dec(\beta)$;
si $S = \{\}$ **alors** $\alpha \leftarrow max(\underline{\alpha}, n(\beta))$;
retourner $< \alpha, S >$;
fin

Nous sommes maintenant à même de décrire un algorithme pour le calcul de décisions pessimistes. Nous devons calculer d'abord les nogoods, puis le label. $K^\wedge \wedge (\sim P)^\wedge$ doit être sous forme normale conjonctive, donc $P^\vee$ doit être sous forme normale disjonctive. Puisque P est sous FNC, la procédure n'acceptera comme bases de préférences que des clauses ou des phrases (les deux types de formule sont à la fois sous forme normale conjonctive et disjonctive). Pour calculer le label de formules sous FNC il nous faut utiliser une particularité de la procédure MPL : le label d'une conjonction $\psi \wedge \phi$

peut être calculé à partir des labels de ψ et ϕ (proposition A.3.3). Cette propriété permet également d'interrompre le calcul du label d'une conjonction de formules dès qu'un label "intermédiaire" est trouvé vide.
Illustrons le fonctionnement de cet algorithme sur un exemple :

Exemple 3.3.3 *K et P contiennent chacune 5 niveaux (les deux échelles sont commensurables). En premier lieu, on ne prend en compte que la strate la plus certaine de K et on calcule successivement les labels de P_4, P_3, P_2 et finalement de P_1 qui est le premier que l'on trouve vide (voir Figure 3.1). Maintenant α prend la valeur* $\max(\underline{\alpha}, n(\beta)) = \max(3, 3) = 3$. *on considère alors K_3 en plus de K_5*

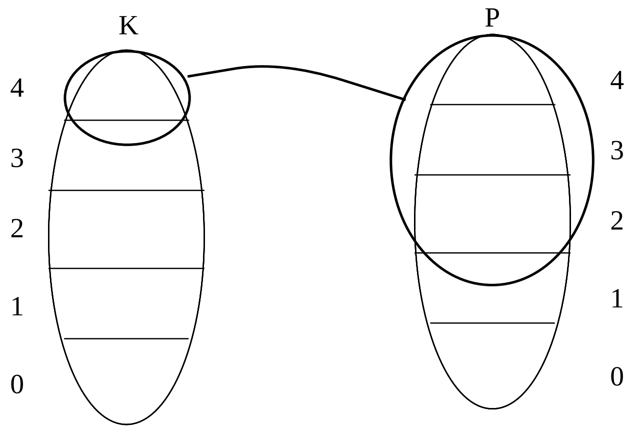

FIG. 3.1 – *Exemple*

et on calcule les labels. Une fois de plus les labels de P_4, P_3 et P_2, sont trouvés non-vides (voir Figure 3.2), la valeur de β devient 1, ce qui n'est pas strictement supérieur à $n(\alpha) = 1$. C'est pourquoi, α reste inchangé, et puisque $S \neq \{\}$, l'algorithme retourne α et S.

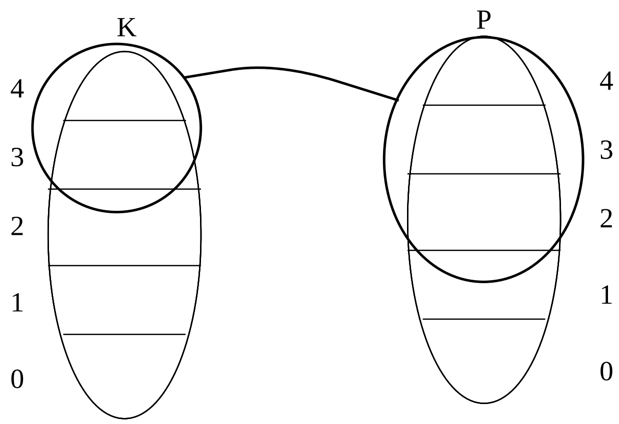

FIG. 3.2 – *Exemple (suite)*

Remarques sur l'implémentation

On doit remarquer que l'algorithme n'utilise pas l'hypothèse que $K_{\mathcal{D}}^{\wedge} \wedge d^{\wedge}$ est cohérent : pour chaque valeur α, les nogoods de K_α sont calculés et les décisions retournées par l'algorithme avec l'utilité α

sont garanties cohérentes avec K_α. En supposant (comme nous l'avons fait plus tôt dans ce chapitre) que $K^\wedge_{\mathbf{D}} \wedge d^\wedge$ est cohérent on pourrait simplifier l'algorithme, puisque le calcul des nogoods deviendrait inutile. Qui plus est, lorsqu'une strate de connaissances supplémentaire est prise en compte il devient inutile de recalculer les labels des strates de préférences trouvées non vides auparavant, puisqu'en ajoutant des connaissances "cohérentes" on ne peut rendre ces labels vides. Donc dans la seconde boucle, β ne doit pas forcément être réinitialisé à **1** .
Néanmoins l'algorithme que nous avons proposé ici, quoique plus complexe, a l'avantage de pouvoir fonctionner avec des bases de connaissances et de préférences partiellement incohérentes.

L'aspect "anytime" de l'algorithme MPL peut être remarqué ici. Si l'on arrête l'algorithme avant sa fin normale, on peut obtenir éventuellement un sous-ensemble de l'ensemble des décisions optimales. Ceci peut être utile, par exemple, si l'on ne cherche qu'une décision optimale et/ou son utilité et non pas *toutes* les décisions optimales.

Exemple

Reprenons l'exemple de Savage que nous allons exprimer en logique. Rappelons le tableau résumant le problème de décision :

	État du monde	
Action	oeuf sain (s)	oeuf pourri (p)
Casser l'oeuf dans le bol (CB)	omelette à 6 oeufs (6O)	pas d'omelette, 5 oeufs gâchés (G)
Le casser dans une tasse (CT)	omelette à 6 oeufs une tasse à laver (6T)	omelette à 5 oeufs une tasse à laver (5T)
Le jeter (J)	omelette à 5 oeufs un oeuf gâché (5G)	omelette à 5 oeufs (5O)

Nous allons exprimer les connaissances et les préférences de l'agent sous la forme de propositions logiques valuées. Les connaissances "certaines" représentant le résultat de chaque action lorsque l'état du monde est connu sont faciles à modéliser par des propositions logiques. Par exemple, savoir que casser l'oeuf dans le bol (CB) donne une omelette à six oeufs (6O) si celui-ci est sain (s) se traduit par : $CB \wedge s \rightarrow 6O$. La seule connaissance incertaine est celle concernant l'état de fraîcheur de l'oeuf : sain (s) ou pourri (p).

Les préférences sont modélisées par des propositions auxquelles sont attachés des degrés de priorité. En termes logiques, si $L = \{0, 1, 2, 3, 4, 5\}$, la base de préférence est $P = \{(\neg G, 5)$, $(\neg OG, 4)$, $(\neg T \vee \neg 5O, 3)$, $(\neg 5O, 2)$, $(\neg T, 1)\}$, où T signifie "tasse à laver" et OG signifie "un oeuf gâché". On rappelle que la fonction d'utilité qualitative sémantique associée, μ, est définie par $\mu(\omega) = \min\{n(\beta_j) t.q. (\psi_j, \beta_j) \in P$ et $\omega \models \neg\psi_j\}$. On vérifie que l'interprétation syntaxique ci-dessus correspond à la fonction d'utilité sémantique $\mu(6O) = 5$, $\mu(6T) = 4$, $\mu(5O) = 3$, $\mu(5T) = 2$, $\mu(5G) = 1$, $\mu(G) = 0$, donnée dans le chapitre 1 de la partie II.

Dans cet exemple, la base de préférences est relativement réduite. La base de connaissances K est plus importante puisqu'elle comprend la description de toutes les actions dans tous les états du monde, les connaissances sur l'état réel du monde et les contraintes entre les variables de décision :

// Base de connaissances K

`Strate` 5
// les décisions CB, CT et J sont mutuellement exclusives ;
$CB \vee CT \vee J$;
$\neg CB \vee \neg CT$;
$\neg CB \vee \neg J$;
$\neg CT \vee \neg J$;

// on obtient une omelette à six oeufs ssi l'oeuf est sain et
// si on le casse dans l'omelette ou dans la tasse ;
$s \wedge CB \rightarrow 6O$;
$s \wedge CT \rightarrow 6O$;
$6O \rightarrow s$;
$\neg 6O \vee \neg J$;

// on obtient une omelette à cinq oeufs ssi l'oeuf est pourri et
// si on le casse dans la tasse, ou si on le jette ;
$J \rightarrow 5O$;
$p \wedge CT \rightarrow 5O$;
$5O \rightarrow J \vee p$;
$\neg 5O \vee \neg CO$;

// on à une tasse à laver ssi on y casse l'oeuf ;
$CT \rightarrow T$;
$T \rightarrow CT$;

// Un oeuf est gâché ssi on jette un oeuf qui est sain ;
$s \wedge J \rightarrow OG$;
$OG \rightarrow s$;
$OG \rightarrow J$;

// l'omelette est gâchée ssi on y casse un oeuf pourri ;
$p \wedge CB \rightarrow G$;
$G \rightarrow CB$;
$G \rightarrow p$;

// un oeuf est soit sain soit pourri ;
$\neg s \vee \neg p$;
$s \vee p$;

`Strate` 2
// dans cet exemple on est relativement convaincu que l'oeuf est sain ;
s ;

Faisons tourner les algorithmes de recherche de décisions optimales optimiste et pessimiste sur cet exemple :

- **Décision optimiste**
  Première étape : α = $\mathbf{0}$= 0.
  Le calcul de MPL($K_{\overline{\mathbf{0}}} \cup P_{\overline{\mathbf{0}}}$,$\{\}$,D) retourne la solution $\{CB\}$.
  L'utilité optimiste de cette solution est $n(\alpha) = \mathbf{1}$.

- **Décision pessimiste**
  Première étape : α = **1** = 5.
  Calcul des nogoods : $NG = \{\{CB, CT\}, \{CB, J\}, \{CT, J\}\}$.
  Calcul du label de $\neg G \wedge \neg OG \wedge (\neg 5O \vee \neg T) \wedge \neg 5O \wedge \neg T$ ($\overline{n(\alpha)} = 1$):
  Label$_{K_\alpha}(\neg G) = \{\{CT\}, \{J\}\}$. $\beta = 5$.
  Label$_{K_\alpha}(\neg G \wedge \neg OG) = \{\{CT\}\}$. $\beta = 4$.
  Label$_{K_\alpha}(\neg G \wedge \neg OG \wedge (\neg 5O \vee \neg T) = \{\}$. $\beta = 3$.
  Stop. On essaie avec $\alpha = n(\beta) = 2$ (la nouvelle valeur de α est $max(\underline{\alpha}, n(\beta)) = 2$ puisque $K_{(4)}$ et $K_{(3)}$ sont vides).
  Deuxième étape : $\alpha = 2$.
  Calcul des nogoods : $NG = \{\{CB, CT\}, \{CB, J\}, \{CT, J\}\}$.
  Calcul du label de $\neg G \wedge \neg OG$ ($n(\alpha) = 4$):
  Label$_K(\neg G) = \{\{\}\}$. $\beta = 5$.
  Label$_K(\neg G \wedge \neg OG) = \{\{CT\}, \{CB\}\}$. $\beta = 4$.
  Stop. $\overline{n(\alpha)} > 3 = \beta$.
  Les décisions pessimistes optimales sachant que $N(s) = 2$ sont CB et CT, d'utilité pessimiste $\alpha = 2$.

Remarques

- Les deux bases représentant les connaissances et les préférences d'un agent et en particulier la base de connaissances de cet exemple peuvent exprimer plus que ce qui est réellement nécessaire pour résoudre le problème de décision. Cependant l'agent est en général incapable de distinguer, parmi ses connaissances, lesquelles vont réellement être utiles à la résolution du problème de celles qui ne le seront pas. En fait, l'agent ne se préoccupera que d'exprimer suffisamment de connaissances pour calculer des décisions optimales, d'autant plus que la redondance des informations n'est pas un problème pour les procédures de décision que nous avons définies.

- Nous avons vu, dans cette section, que l'on pouvait coder un problème de décision sous incertitude possibiliste dans le langage des ATMS, et nous avons proposé des algorithmes de résolution, faisant appel à des procédures spécifiques appliquées aux ATMS. Dans la section suivante, nous allons approfondir le lien entre abduction et décision pessimiste, satisfaisabilité et décision optimiste.

3.4 Extension du lien entre décision et abduction

Nous allons exposer en premier lieu une généralisation du cadre des ATMS dans laquelle les symboles distingués ne représenterons pas seulement les hypothèses (donc la connaissance incertaine), mais aussi des contraintes (ou préférences et des décisions applicables. Cette généralisation proposée dans [Sab98] peut être spécialisée, en "valuant" les symboles distingués, à la représentation et au calcul de décisions possibilistes pessimistes, mais aussi de décisions optimales au sens des fonctions de croyance de Shafer [Sha76], comme nous l'avons montré dans [LS97].

3.4.1 D-ATMS : une généralisation du cadre des ATMS à la représentation de problèmes de décision sous incertitude

Jusqu'ici nous avons exposé une méthode de représentation et de résolution de problèmes de décision sous incertitude possibiliste basée sur les ATMS (pour le cas pessimiste). La méthode de calcul proposée, basée sur la procédure MPL est incrémentale. L'ensemble des symboles hypothèses est constitué des symboles de décision (D), la base de connaissances est réduite aux connaissances certaines au départ, puis augmentée au fur et à mesure. Dans une boucle intérieure, on cherche les labels des préférences, en "incorporant" celles-ci une à une.
La méthode de calcul proposée prend bien entendu appui sur les propriétés spécifiques de la procédure MPL. Cependant, nous allons voir que l'on peut facilement représenter un problème de décision sous incertitude possibiliste dans le cadre général des ATMS (classiques), en considérant un ensemble d'hypothèses, constitué non seulement des symboles de décision, mais également de symboles représentant l'incertitude et les préférences. Alors, à partir d'un calcul de label classique, on obtient en quelque sorte une "stratégie", i.e. une fonction qui associe à chaque état du monde les décisions permettant de satisfaire certaines des préférences. Nous verrons dans les deux paragraphes suivants qu'à partir de ces labels ou stratégies et en attachant des valeurs aux symboles représentant l'incertitude et les préférences, on peut calculer des décisions optimales, au sens de l'utilité qualitative pessimiste ou au sens de l'utilité basée sur les fonctions de croyance, proposée par Jaffray et Wakker [JW94] et décrite dans le chapitre 2 de la partie I.

Un D-ATMS est un ATMS *orienté vers la décision*. La différence principale avec le cadre classique est que l'ensemble d'hypothèses est partagé entre les symboles hypothèses, les symboles de décision, et les symboles buts.

Définition 3.4.1 *Un D-ATMS est défini par le quintuplet:* $< \mathcal{L}, K, H, D, B >$, *où:*
- $\mathcal{L}$ *est le langage,*
- K *est l'ensemble de formules de* $\mathcal{L}$, *exprimant les connaissances de l'agent,*
- H *est un ensemble de symboles hypothèses* [2],
- D *est un ensemble de symboles de décision,*
- B *est un ensemble de symboles, constituant le but à atteindre.*

Le langage peut comprendre des symboles qui ne sont pas dans $H \cup D \cup B$.

Exemple 3.4.1 *Le parapluie.*
Cet exemple initialement proposé par Boutilier a été repris dans [DP95b]. Il s'agit de déterminer, sachant qu'il risque de pleuvoir, s'il vaut mieux prendre un parapluie ou non.
H *contient deux symboles,* $PLUIE$ *et* $NONPLUIE$, *déterminant s'il va pleuvoir ou non,* D *contient deux symbole représentant les deux actions "prendre un parapluie" et "ne pas prendre de parapluie",* $PARA$ *et* $NONPARA$,*. Enfin,* B *contient deux symboles* B_1 *et* B_2, *correspondant aux deux objectifs de l'agent : ne pas être mouillé, et garder les mains libres en ne portant pas de parapluie. La base*

2. H est différent de l'ensemble des symboles hypothèses des ATMS classiques : l'ensemble des symboles hypothèses d'un D-ATMS est constitué de $H \cup D \cup B$.

de connaissances K comprend des formules modélisant les conséquences des actions :
$PARA \rightarrow \neg mouille$ *: si on prend un parapluie on n'est pas mouillé,*
$PLUIE \wedge NONPARA \rightarrow mouille$ *: s'il pleut et qu'on est sans parapluie, on est mouillé,*
$\neg PLUIE \rightarrow \neg mouille$ *: s'il ne pleut pas, on n'est pas mouillé,*

Mais aussi des formules modélisant les préférences de l'agent :
$\neg mouille \rightarrow B_1$ *: notre premier but est de ne pas être mouillé,*
$NONPARA \rightarrow B_2$ *: notre second but est de ne pas porter de parapluie,*
$B_1 B_2 \rightarrow BUT$ *: BUT est un symbole supplémentaire dont on calculera le label afin de déterminer les actions optimales.*

Et enfin des formules exprimant les contraintes entre variables de décision ou variables incertaines :
$\neg PLUIE \leftrightarrow NONPLUIE$*,*
$\neg PARA \leftrightarrow NONPARA$*.*

En effectuant un calcul de label classique sur le symbole BUT, avec pour ensemble d'hypothèses $\{PLUIE, NONPLUIE, B_1, B_2, PARA, NONPARA\}$, on obtient :
$label_K(BUT) = \{\{B_1, B_2\}, \{B_2, PARA\}, \{NONPARA, B_1\}, \{NONPLUIE, B_2\}, \{NONPLUIE, NONPA$

A l'examen de $label_K(BUT)$ on obtient des informations sur les effets des actions, en terme de satisfaction des préférences. Chaque environnement du label constitue un *moyen* de satisfaire le BUT. Par exemple, l'environnement $\{B_2, PARA\}$ suggère que le fait de prendre un parapluie permet de satisfaire B_1 (rester au sec), le fait de ne pas en prendre permet de satisfaire B_2 (environnement $\{NONPARA, B_1\}$) et permet de satisfaire toutes les préférences *s'il ne pleut pas* (environnement $\{NONPLUIE, NONPARA\}$)[3].

Plus généralement, tout environnement E est constitué d'un ensemble de symboles représentant les hypothèses (E^H), d'un ensemble de symboles de décision (E^D) et d'un ensemble de symboles de préférences, ou buts (E^B). Chaque environnement E peut alors être interprété intuitivement par : "Si l'action (cohérente avec K) que je choisis est un sur-ensemble de E^D, si l'ensemble des hypothèses satisfaites par le monde réel contient E^H alors toutes les préférences autres que celles contenues dans E^B sont satisfaites" (en effet : les préférences contenues dans E^B sont nécessaires pour atteindre le but).

Maintenant que nous avons défini le cadre des D-ATMS, nous allons voir dans les deux paragraphes qui suivent comment on peut, en partant de l'expression de $label_K(BUT)$ et en attachant des valeurs aux symboles de préférence et d'incertitude, calculer l'utilité des actions disponibles. Comme nous l'avons montré dans [Sab98], suivant que l'on attache des degrés de priorité aux préférences et des degrés de certitude aux hypothèses ou des utilités additives et des degrés de probabilité, on pourra calculer les utilités des actions au sens de l'utilité qualitative possibiliste pessimiste ou au sens de l'utilité basée sur les fonctions de croyance.

3. La méthode suppose que la *seule* manière d'obtenir BUT est de prouver tous ses sous-buts $B_1, B_2, \ldots$ Ainsi, on est certain que si un sous-but est absent d'un des environnement de $label_K(BUT)$ c'est que celui-ci est satisfait. Cette méthode à l'avantage d'intégrer en un seul label toutes les informations nécessaires au calcul de décisions optimales, néanmoins si on cherche seulement les actions permettant d'obtenir un sous-but donné, on pourra préférer utiliser des symboles de données et non d'hypothèses pour les sous-buts. Un ATMS classique permettra d'obtenir les labels (actions et hypothèses) de chaque donnée (dont les sous-buts).

3.4.2 D-ATMS et utilité qualitative pessimiste

Dans l'expression logique des problèmes de décision sous incertitude qualitative possibiliste que nous avons proposée, les bases de connaissances et de préférences K et P sont stratifiées, i.e. les formules sont rassemblées en strates de niveau de certitude ou de priorité donné. Faire entrer un tel problème dans le cadre des D-ATMS est aisé. Il suffit de :

- Rajouter un ensemble de symboles hypothèses $\{H_i\}$, un par strate de connaissances de certitude donnée.
- Rajouter un ensemble de symboles $\{B_i\}$ exprimant chacun la satisfaction de toutes les préférences d'un niveau de priorité donné.

Ensuite, nous pouvons regrouper les deux bases stratifiées K et P en une nouvelle base, *non stratifiée* K'. Pour ce faire, nous remplaçons chaque clause ϕ de la strate de K de niveau i par la clause $H_i \to \phi$. Le problème de la base de préférences est un peu plus complexe. A chaque strate de la base de préférences correspond un symbole B_j. B_j exprime la satisfaction de toutes les préférences $\{\psi_j^1, \psi_j^2, \ldots\}$ de niveau de priorité j. La formule que l'on doit rajouter à K' pour chaque strate de préférences est donc $\psi_j^1 \wedge \psi_j^2 \wedge \ldots \to B_j$. Le problème est que les formules ajoutées ne sont plus des clauses lorsque les strates de préférences comportent plus d'une clause, ce qui pose des problèmes pour certaines implémentations d'ATMS. Ce problème peut néanmoins être contourné en utilisant des symboles intermédiaires, non-hypothèses, b_j^i, un pour chaque clause ψ_j^i de la strate de priorité j. Dans ce cas, K' est obtenue en ajoutant les clauses $\psi_j^i \to b_j^i$, et pour chaque strate j, la clause $b_j^1 \wedge b_j^2 \wedge \ldots \to B_j$. Enfin, dans tous les cas, la clause $B_1 \wedge B_2 \wedge \ldots \to BUT$ est ajoutée pour signifier que le but est atteint si toutes les préférences sont satisfaites.

Dans l'exemple de l'omelette considéré dans sa forme logique, la nouvelle base K', obtenue à partir de K et P est :
// Base de connaissances K'

Connaisances certaines
// les décisions CB, CT et J sont mutuellement exclusives ;
$\neg H_1 \vee CB \vee CT \vee J$;
$\neg H_1 \vee \neg CB \vee \neg CT$;
$\neg H_1 \vee \neg CB \vee \neg J$;
$\neg H_1 \vee \neg CT \vee \neg J$;
// on obtient une omelette à six oeufs ssi l'oeuf est sain et
// si on le casse dans l'omelette ou dans la tasse ;
$H_1 \vee s \wedge CB \to 6O$;
$H_1 \vee s \wedge CT \to 6O$;
...
Connaissances incertaines
$H_2 \to s$;
$H_3 \to p$;

Préférences
$B_5 \vee G$;
$B_4 \vee OG$;

$B_3 \vee 5O$; $B_3 \vee T$;
$B_2 \vee 5O$;
$B_1 \vee T$;
$B_1 \wedge B_2 \wedge B_3 \wedge B_4 \wedge B_5 \rightarrow BUT$;

Environnements de $label_{K'}(BUT)$	Utilité pessimiste	Décisions optimales
$\{B_1, B_2, B_3, B_4, B_5\}$	0	CT, CO, J
$\{H_1, H_2, B_1, B_2, B_4\}$	$min(N(s), 1)$	CT, CO, J
$\{H_1, H_3, B_1, B_2, B_3, B_5\}$	0	CT, CO, J
$\{H_1, B_1, B_2, B_3, CT\}$	2	CT
$\{H_1, H_2, B_1, CT\}$	$min(N(s), 4)$	CT
$\{H_1, B_2, B_4, J\}$	1	J
$\{H_1, H_3, B_2, J\}$	$min(N(p), 3)$	J
$\{H_1, B_5, CO\}$	0	CO
$\{H_1, H_2, CO\}$	$N(g)$	CO

L'utilité pessimiste d'une décision peut être calculée à partir du "meilleur chemin" qu'elle "ouvre" vers le but, où un tel "chemin" est un environnement E dont la partie E^D est contenue dans la décision vue comme un ensemble de littéraux. Si par exemple on choisit la décision CO, les "chemins" possibles vers le but sont $\{B_1, B_2, B_3, B_4, B_5\}$, $\{H_1, H_2, B_1, B_2, B_4\}$, $\{H_1, H_3, B_1, B_2, B_3, B_5\}$, $\{H_1, B_5, CO\}$ et $\{H_1, H_2, CO\}$. L'utilité d'un chemin ou environnement est fonction de l'hypothèse la moins certaine et de la préférence la plus prioritaire qu'il contient. Par exemple, l'utilité de $\{H_1, H_2, B_1, CT\}$ dépend de l'hypothèse H_2 de degré de certitude $N(s)$ dont il a besoin pour prouver BUT, ainsi que de la préférence B_1 de priorité 1, que la décision CT ne peut satisfaire à elle seule. Ainsi, l'utilité de cet environnement est (en supposant que l'action CT soit effectuée) $min(N(s), n(1)) = min(N(s), 4)$.
Remarquons qu'avec notre mode de calcul, l'environnement $\{B_1, B_2, B_3, B_4, B_5\}$ est toujours présent dans le label de BUT. Cet environnement est un moyen *artificiel* d'atteindre le but puisqu'il signale que celui-ci est atteint si on suppose que toutes les préférences sont satisfaites. L'utilité correspondant à cet environnement est toujours minimale puisqu'il contient la préférence de priorité maximale, B_5.
Chaque décision de l'exemple de l'omelette peut être évaluée par l'utilité du meilleur chemin qu'elle ouvre :

- $u_*(BIO) = max(0, min(N(g), 1), 0, 0, N(g))$,
- $u_*(TA) = max(0, min(N(g), 1), 0, 1, min(N(r), 3))$,
- $u_*(BAC) = max(0, min(N(g), 1), 0, 2, min(N(g), 4))$.

Bien entendu, ces utilités sont les mêmes que celles obtenues syntaxiquement par la méthode exposée plus tôt dans ce chapitre et que celles obtenues sémantiquement comme exposé dans le chapitre 1 de cette partie, même si elles sont obtenues ici sous forme $maxmin$ et non $minmax$.

3.4.3 D-ATMS et utilité basée sur les fonctions de croyance

Pourquoi n'existe-t-il pas de méthode de représentation en logique propositionnelle valuée des problèmes de décision dans l'incertain dont la sémantique soit en accord avec le critère de l'utilité espérée ?

Il n'est pas possible en général de déduire une distribution de probabilité unique sur l'ensemble des interprétations à partir de connaissances exprimées sous la forme d'un ensemble de formules propositionnelles, mais valuées par des degrés de probabilité :

Exemple 3.4.2 *Soient les deux variables propositionnelles a et b. La connaissance d'un agent est représentée par la formule $\varphi = \neg a \wedge b$: il sait que soit $\neg a$ est faux, soit b est vrai. De plus, il sait que cette connaissance est entachée d'incertitude, il connait la probabilité que φ soit vraie : $P(\varphi) = 0,5$. Cette connaissance ne permet pas de déterminer une distribution de probabilité p unique sur l'ensemble des interprétations sur $\{a, b\}$: tout ce que l'agent peut déduire, c'est que $p(a \wedge b) + p(\neg a \wedge \neg b) + p(\neg a \wedge b) = 0,6$, et que $p(aa \wedge \neg b) = 0,4$.*

On se retrouve dans un cas où les probabilités sont affectés à des ensembles d'interprétation, et non à des singletons. Nous verrons également dans la seconde partie de cette thèse que dans ce cas on peut construire des méthodes logiques se représentation dont la sémantique est en accord avec les critères généralisant l'utilité espérée, proposés par Jaffray et Wakker [JW94].
Les fonctions de croyance [Dem67] [Sha76] (voir aussi le chapitre 2 de la partie I généralisent les fonctions de probabilité : la distribution de probabilité sur l'ensemble S est remplacée par une *fonction de masse* sur l'ensemble des parties de S.

Définition 3.4.2 *m est une fonction de masse ssi :*
$m : 2^S \to [0, 1]$ *vérifie :*
$(i) m(\emptyset) = 0$
$(ii) \sum_{A \subseteq S} m(A) = 1$

On définit alors une *fonction de croyance* :

Définition 3.4.3 *$Bel : 2^S \to [0, 1]$ est une fonction de croyance ssi il existe une fonction de masse m telle que : $\forall A \subseteq S, Bel(A) = \sum_{B \subseteq A} m(B)$*

[JW94] ont défini une contrepartie de l'utilité espérée au sens de von Neumann et Morgenstern, basée sur les fonctions de croyance:

Définition 3.4.4 *Utilité d'une action f :*

$$u_{Bel}(f) = \sum_{A \subseteq X} m(f^{-1}(A)) \cdot inf_{x \in A}(\mu(x))$$

où m est une fonction de croyance sur S, et μ est une fonction d'utilité sur X.

Dans le cas où X est un ensemble fini, il existe un nombre fini de sous-ensembles de X : $\{X_0, \ldots, X_q\}$, tels que $X_q \subseteq \ldots \subseteq X_0 = X$, et autant de nombres réels, $\{\mu_0, \ldots, \mu_q\}$ tels que $\mu_0 < \ldots < \mu_q$ et $\forall i \in 0 \ldots q-1, \forall x \in X_i - X_{i+1}, \mu(x) = \mu_i$ et $\forall x \in X_q, \mu(x) = \mu_q$.
A partir d'une telle décomposition, [LS97] ont montré que u_{Bel} pouvait être réécrite sous la forme :

Proposition 3.4.1

$$u_{Bel}(f) = \sum_{i \in 1 \ldots q} Bel(f^{-1}(X_i)) \cdot (\mu_i - \mu_{i-1}) + \mu_0$$

Preuve :

$u_*(R) = \sum_{A \subseteq X} m(A|R) \cdot \inf_{x \in A} \mu(x) \Leftrightarrow u_*(R)$
$= \sum_{i \in 0 \ldots n-1} (\sum_{A \subseteq G_i, A \not\subseteq G_{i+1}} m(A|R)) \cdot \mu_i + (\sum_{A \subseteq G_n} m(A|R)) \cdot \mu_n.$

Remarquons que $\sum_{A\subseteq G_i, A\not\subseteq G_{i+1}} m(A|R) = (\sum_{A\subseteq G_i} m(A|R)) - (\sum_{A\subseteq G_{i+1}} m(A|R))$
$= Bel(G_i|R) - Bel(G_{i+1}|R)$,
on obtient : $u_*(R) = \sum_{i\in 0\ldots n-1}(Bel(G_i|R) - Bel(G_{i+1}|R))\cdot \mu_i + Bel(G_n|R)\cdot \mu_n$
ce qui s'écrit finalement :

$$u_{Bel}(R) = \sum_{i\in 1\ldots n} Bel(G_i|R)\cdot(\mu_i - \mu_{i-1}) + \mu_0$$

Dans le cas où les buts sont d'une certaine forme, dite *graduelle* [4], [LS97] ont montré que l'utilité u_{Bel} d'une décision d peut être calculée à partir de $label_{\mathcal{K}'}(BUT)$ par la procédure suivante :

1) Calculer $label_{\mathcal{K}'}(BUT)$,

2) éliminer de $label_{\mathcal{K}'}(BUT)$ tous les environnements E tels que $d \not\vdash E^D$,

3) Pour chaque but B_i, calculer $Bel(B_i|d)$.

On utilisera l'une des méthodes proposées par [Pro89] et [KM95] pour calculer la "probabilité" de cet ensemble d'environnements à partir des probabilités des hypothèses. Ces méthodes passent par la mise sous une forme particulière, dite forme "mutuellement exclusive ou indépendante" (MEI), de la formule exprimant le label, à partir de laquelle on peut facilement calculer la "probabilité" du label [5]. Les probabilités calculées dans le point 3) correspondent aux $Bel(f^{-1}(X_i))$ de la Proposition 3.4.1. A partir d'elles on peut évaluer directement l'utilité de la décision f.

Exemple 3.4.3 *(tiré de [LS97])*
Une voiture doit passer un contrôle technique. Sa carrosserie, ses freins et son silencieux sont susceptibles d'être considérés comme défaillants. Leurs probabilités respectives de ne pas entraver l'obtention du certificat sont: $P(\neg Ab(c)) = 0.7; P(\neg Ab(f)) = 0.4; P(\neg Ab(s)) = 0.3$.
Les coûts de réparation respectifs sont: $Cout(rc)$ *= 10000 FF ;* $Cout(rf)$ *= 2500 FF ;* $Cout(rs)$ *= 500 FF. Le prix d'une voiture neuve est:* $Cout(nc) = 50000$ *FF. Quelle est la meilleure décision au sens de l'utilité espérée basée sur les fonctions de croyance?*
Acheter une nouvelle voiture (nv) *permet d'être sûr de passer le contrôle technique, et réparer un élément le rend apte à passer le contrôle (ok).* $\mathcal{K} = \{nv \to controle; rc \to okC; rf \to okF; rs \to okS;$
$\neg Ab(c) \to okC; \neg Ab(f) \to okF; \neg Ab(s) \to okS;$
$okC \wedge okF \wedge okS \to controle\}$
Alors, $\mathcal{D} = \{rc, rf, rs, nv\}$ *(il y a* $2^4 = 16$ *décisions possibles),* $\mathcal{H} = \{Ab(c), Ab(f), Ab(s)\}$ *(il y a huit états de panne possibles).*
Les coûts des différentes réparations sont pris en compte pour l'élaboration des buts graduels:
$B_0 \leftarrow \top, \mu_0 = -75000$ *(coût estimé d'être à pieds...).*
$B_1 \leftarrow controle, \mu_1 = -Cout(nv) - Cout(rc) - Cout(rf) - Cout(rs) = -63000$*;*
$B_2 \leftarrow controle \wedge \neg nv, \mu_2 = -Cout(rc) - Cout(rf) - Cout(rs) = -13000$*;*

4. $\{B_0,\ldots,B_n\}$ sont des buts sous forme graduelle ssi $B_n \vdash \ldots \vdash B_0$ et $\mu(B_i) > \mu(\wedge_{j\in 0\ldots i-1} B_j)$. La satisfaction d'un but est plus "utile" que la satisfaction de tous les buts de plus faible priorité.

5. La forme MEI est une forme normale disjonctive, $E_1 \vee \ldots \vee E_n$, où les E_i sont soit indépendants (n'ont pas de symboles hypothèse en commun), soit exclusifs (deux environnements contiennent un même littéral, mais positif chez l'un, et négatif chez l'autre). On calcule alors facilement la probabilité d'une telle formule, car si E_i et E_j sont MEI, $P(E_i \vee E_j) = P(E_i) + P(E_j)$.

$B_3 \leftarrow controle \wedge \neg nv \wedge \neg rc, \mu_3 = -Cout(rf) - Cout(rs) = -3000$;
$B_4 \leftarrow controle \wedge \neg nv \wedge \neg rc \wedge \neg rf, \mu_4 = -Cout(rs) = -500$;
$B_5 \leftarrow controle \wedge \neg nv \wedge \neg rc \wedge \neg rf \wedge \neg rs, \mu_5 = 0$;

Après introduction du symbole BUT, et de la connaissance $\wedge B_i \vdash BUT$, on calcule label(BUT) =
$\{\{B_1, B_2, B_3, B_4, B_5\}, \{\{\}, \neg Ab(c), \neg Ab(f), \neg Ab(s)\}$,
$\{$**nv**$, B_2, B_3, B_4, B_5\}, \{$**rc**$, \neg Ab(s), \neg Ab(f), B_3, B_4, B_5\}$,
$\{$**rf**$, \neg Ab(c), \neg Ab(s), B_4, B_5\}, \{$**rs**$, \neg Ab(c), \neg Ab(f), B_5\}$,
$\{$**rc**, **rf**$, \neg Ab(s), B_3, B_4, B_5\}, \{$**rc**, **rs**$, \neg Ab(f), B_3, B_4, B_5\}$,
$\{$**rs**, **rf**$, \neg Ab(c), B_4, B_5\}, \{$**rc**, **rf**, **rs**$, B_3, B_4, B_5\}\}$.
A partir de là, on peut évaluer pour chaque sous but B_i et chaque décision d, le degré de croyance $Bel(B_i|d)$ (voir [LS97] pour plus de détails. La méthode utilisée est basée sur la procédure proposée par [CCCL96]).
Les utilités des différentes décisions disponibles peuvent être calculées:
$u_*(\{\}) = -75000 + 12000 \cdot Bel(B_1|\{\}) + 50000 \cdot Bel(B_2|\{\})$
$+10000 \cdot Bel(B_3|\{\}) + 2500 \cdot Bel(B_4|\{\}) + 500 \cdot Bel(B_5|\{\}) = -75000 + 0.084 \cdot 75000 = -68700$
$u_*(\{$**rb, rc, rm**$\}) = -75000 + 12000 + 50000 = -13000$
$u_*(\{\mathbf{nc}\}) = -75000 + 12000 = -63000$
$u_*(\{\mathbf{rc}, \mathbf{rb}\}) = -75000 + 0.3 \cdot (12000 + 50000) = -56400$
$u_*(\{\mathbf{rc}\}) = -75000 + 0.12 \cdot (12000 + 50000) = -67560 \ldots$

3.5 Conclusion

Dans ce chapitre nous avons proposé un mode de représentation de problèmes de décision sous incertitude et de calcul de décisions optimales, basé sur la logique propositionnelle.
En particulier, nous avons fait le lien entre les théories possibilistes optimistes et pessimistes présentées dans le chapitre 1 de la partie II et les problèmes d'abduction et de cohérence dans des bases stratifiées. Profitant d'une collaboration avec un autre doctorant de l'équipe, ayant proposé une implémentation particulière d'un ATMS, nous avons proposé et codé des méthodes spécifiques de résolution de problèmes de décision sous incertitude possibiliste, représentés logiquement.

Un lien entre représentation logique, abduction et décision sous incertitude basée sur les fonctions de croyance a également été mis en évidence. Un lien semblable existe entre cohérence et décision basée sur des fonctions de plausibilité.
Enfin, nous avons montré que ces deux types de décision sous incertitude (possibiliste ou basé sur les fonctions de croyance) peuvent être intégrés dans une même approche logique dans le cadre des ATMS (pour les cas pessimistes) utilisant des symboles hypothèses étendus.

La représentation utilisée, basée sur la logique propositionnelle (éventuellement valuée) permet un compromis entre une représentation concise de problèmes de décision sous incertitude, et des méthodes de calcul de décisions efficaces (basées sur une implémentation particulière d'un ATMS utilisant une procédure de type Davis et Putnam [CCCL96]). D'autres représentations structurées de problèmes de décision comportant à la fois de l'incertitude et des préférences ont été proposées récemment: basées sur les CSP (par exemple, [FLS96b], [BBGP97]), où sur les diagrammes d'influence (par exemple [Zha94]). Les diagrammes d'influence permettent de représenter graphiquement de manière concise les dépendances entre les variables du problème. Le calcul de décisions optimales se fait toujours de manière classique (bayésienne), grâce à des tables de probabilités conditionnelles, résumées par les arcs d'un diagramme d'influence.

Le travail décrit dans ce chapitre devra être poursuivi dans plusieurs directions :

– Utiliser le modèle de représentation logique basé sur les ATMS pour recouvrir d'autres théories non-classiques. La théorie proposée par [BT96] par exemple, doit pouvoir supporter très bien la modélisation logique que nous avons proposé, puisqu'elle peut être vue comme un cas particulier de la théorie de l'utilité qualitative pessimiste de [DP95b].

– Etendre notre modèle à la décision séquentielle. Dans ce cas, il s'agit non plus de déterminer une action optimale, mais une séquence d'actions. Des méthodes basées sur une contrepartie possibiliste de la programmation dynamique qui ont été proposées par [FLS96a] et ont été décrites dans le chapitre précédent. Une extension du modèle décrit dans ce chapitre, dans laquelle le temps est explicitement représenté, pourrait permettre de représenter et de résoudre de tels problèmes de décision séquentielle, en multipliant néanmoins le nombre de variables par le nombre d'étapes du problème de décision séquentielle.

Conclusion

Une étude complète des processus décisionnels doit inclure trois composantes (Geffner [Gef98]) : étude des modèles de décision (axiomatique), des langages de représentation et des algorithmes de résolution.
Nous avons proposé une étude d'une théorie ordinale de la décision basée principalement sur la théorie des possibilités. Notre étude a abordé, de façon sans doute incomplète, ces différents points. Dans cette conclusion nous allons dresser le bilan de notre contribution, puis nous verrons comment notre étude doit être poursuivie.

Bilan

La première contribution de cette thèse consiste en l'axiomatisation d'une famille de critères de décision dans l'incertain, basés sur l'intégrale de Sugeno. Ces critères de décision font appel à des mesures d'incertitude et de préférence ordinales. Les mesures d'incertitude, en particulier, sont monotones pour l'inclusion.

Les principales différences entre ces critères de décision et le critère de l'utilité espérée, outre leur forme plus *qualitative*, sont les suivantes :

- La notion de compensation entre les différentes conséquences possibles d'une action est rejetée. Le critère de l'utilité espérée évalue les actions en fonction de la moyenne (pondérée) des utilités des conséquences possibles, introduisant ainsi l'hypothèse que le problème de décision sera répété et que l'action préférée sera celle qui donne le meilleur résultat "en moyenne". Utiliser l'intégrale de Sugeno, au contraire, suggère que l'utilité d'une action reflète soit l'utilité de l'une de ses conséquences possibles, soit la plausibilité d'un événement conduisant à l'une de ces conséquences.

  Si on introduit une hypothèse supplémentaire de pessimisme ou d'optimisme de l'agent, ceci conduit à l'utilisation de l'un des critères d'utilité qualitative possibiliste de [DP95b].

- Le "principe de la chose certaine" de Savage, stipulant que modifier les conséquences identiques de deux actions ne modifie pas l'ordre de préférence entre ces deux actions, n'est en général pas respecté par les critères de décision basés sur l'intégrale de Sugeno. Les critères possibilistes vérifient toutefois une version affaiblie de ce principe, souffrant d'un éventuel "effet de noyade" des préférences, lorsque des conséquences extrêmement graves ou bénéfiques peuvent plausiblement être atteintes. Notons que lorsqu'on utilise les critères de décision possibilistes on ne peut jamais renverser une préférence *stricte* entre deux actions.

Si les critères de décision basés sur l'utilité espérée et sur l'intégrale de Sugeno diffèrent en plusieurs aspects, ils conservent toutefois un point commun important : l'hypothèse de commensurabilité entre les niveaux d'incertitude et de préférence. Cette hypothèse découle directement du fait que l'on souhaite construire un préordre complet entre les actions : vouloir comparer actions binaires et actions constantes revient à supposer l'existence d'une échelle commune pour mesurer l'incertitude et les préférences.

La deuxième contribution de cette thèse est une extension de l'utilisation des critères de décision possibilistes à la décision séquentielle en environnement totalement observable.
Nous avons défini une version *possibiliste* des processus décisionnels markoviens. Dans cette version, le principe d'optimalité de Bellman [Bel57], exprimant que "toute sous police d'une police optimale est à son tour optimale" ne tient plus. Nous avons cependant proposé des algorithmes de type "programmation dynamique", calculant par induction arrière des polices optimales au sens des critères possibilistes. Ceci est possible car les processus décisionnels markoviens possibilistes vérifient une certaine forme du principe de Bellman : toute sous police entre les étapes t et N (N est l'horizon du problème) d'une police optimale entre les étapes 1 et N (calculée récursivement), est optimale.

Enfin, nous avons proposé des méthodes de représentation logique de problèmes de décision sous incertitude. L'intérêt principal de ces méthodes est qu'elles permettent de représenter de manière concise et simple les préférences graduelles et les connaissances plus ou moins certaines d'un agent, tout en restant, au niveau de leurs sémantiques, en accord avec les théories possibilistes de la décision dans l'incertain. Pour ce faire, nous avons étendu le langage des ATMS par l'adjonction de symboles de préférence et de décision.
Nous plaçant ensuite dans le cadre, soit de la logique possibiliste, soit de la logique probabiliste, nous avons proposé des algorithmes de résolution pour les problèmes de décision sous incertitude à base de fonctions de croyance ou de possibilités.

Perspectives

Nous avons présenté un ensemble de résultats dans le domaine de la décision "qualitative" dans l'incertain. Certains de ces résultats sont tournés vers la modélisation (axiomatisation de critères de décision), d'autres vers le calcul effectif (algorithmes de résolution de problèmes de décision séquentiels ou exprimés en logique). Ces résultats ont été appliqués, par exemple, à un problème d'ordonnancement dans l'industrie vinicole...

Depuis quelques années la planification sous incertitude utilisant le modèle des *Processus Décisionnels Markoviens Partiellement Observables* (PDMPO) connaît un intérêt croissant, comme le prouve le prochain workshop de l'*American Association on Artificial Intelligence* qui va se tenir a Stanford à la fin de cette année. Or le modèle des PDMPO pose un double problème de complexité spatiale et temporelle, que tentent de résoudre la plupart des travaux récents.

En effet, le fait de remplacer l'ensemble (fini) des états d'un PDM par un ensemble (infini) d'états de connaissance (distributions de probabilités sur les états) rend les PDMPO terriblement complexes à résoudre. C'est pourquoi les travaux concernant les PDMPO se tournent donc soit vers des algorithmes de résolution approchée (anytime), soit vers des représentations structurées qui en diminuent la complexité spatiale (Dearden et Boutilier [DB94]).

Nos travaux sur une théorie qualitative de la décision dans l'incertain ouvrent une nouvelle voie vers le traitement de la planification sous incertitude, permettant de limiter les problèmes de complexité inhérents aux PDMPO classiques. Cependant, avant d'être utilisables dans le domaine de la planification, les résultats présentés dans cette thèse doivent être étendus selon trois directions :

– prise en compte des observation au niveau même de la définition des utilités qualitatives possibilistes,

– étude des PDMPO possibilistes et proposition d'algorithmes de résolution,

– représentation structurée des PDMPO, afin d'en réduire la complexité spatiale.

Caractérisations axiomatiques des préférences conditionnelles

Nous avons proposé des caractérisations axiomatiques pour un certain nombre de critères de décision. Ce faisant, nous avons justifié un certain nombre de mesures d'incertitude monotones (possibilités, nécessités, mesures décomposables...). Cependant, nous avons laissé de côté les notions de conditionnement dans ces mesures. Or, observer change les connaissances d'un agent, donc modifie sa mesure de l'incertitude. La manière dont sa mesure d'incertitude est modifiée par une observation dépend de la notion de conditionnement que l'agent utilise. Or, la notion de conditionnement, en théorie des possibilités comme plus généralement pour beaucoup de “mesures monotones”, n'a pas de définition aussi universellement acceptée que celle du conditionnement bayésien. Plusieurs définitions coexistent, non-équivalentes et pas forcement bien “interprétables”, c'est pourquoi une axiomatisation du conditionnement en termes de préférences “conditionnelles” entre actions devra être proposée.

Décision séquentielle en environnement partiellement observable

La notion de conditionnement, couplée à la notion d'utilité qualitative trouve une application directe en décision séquentielle lorsque l'on souhaite prendre en compte l'observabilité partielle de la plupart des problèmes réels. Nous avons proposé des solutions pour les problèmes complètement observables et des solutions ont été données dans le cas d'observabilité nulle par Da Costa [Da 98]. Cependant, le cas plus général des problèmes de décision qualitative en environnement *partiellement* observable n'a pas encore été traité. Afin de le traiter et de proposer des contreparties possibilistes des processus décisionnels markoviens partiellement observables (PDMPO), il nous faudra exploiter les résultats sur le problème du conditionnement (pour définir la notion d'utilité qualitative, conditionnée par une observation).

Représentation structurée des problèmes de décision sous incertitude qualitative et application à la planification

Nous avons montré qu'il est possible de représenter en logique (possibiliste) des problèmes de décision sous incertitude qualitative. Il ne s'agit là que d'un premier pas vers la définition de langages plus évolués, dont la sémantique devra continuer à s'appuyer sur les résultats théoriques concernant l'utilité qualitative possibiliste.
De tels langages de représentation pourraient s'appuyer sur :

– les langages classiquement utilisés pour la planification, de type STRIPS ou autres, qui ont déjà été utilisés dans le cadre de la planification sous incertitude en environnement non observable : Kushmerick, Hanks et Weld [KHW95] ont, par exemple proposé un planificateur de type

STRIPS, probabiliste dont on pourrait s'inspirer (des premiers pas dans ce sens ont été réalisés par Guerre [Gue98]).

– Les réseaux de décision bayésiens proposés par Zhang [Zha94] qui étendent les diagrammes d'influence (Pearl [Pea88]) peuvent être utilisés pour représenter des PDMPO de manière compacte. Or, Benferhat et col. [BDGP98] ont défini une contrepartie possibiliste des diagrammes d'influence et mis en évidence leur lien avec la logique possibiliste. Appliquer les PDMPO possibilistes à la planification nécessiterait sans doute, afin de diminuer la complexité spatiale du problème, d'effectuer un travail analogue à celui de Zhang en adaptant les diagrammes d'influence possibilistes à la décision qualitative dans l'incertain.

Annexe A

La procédure MPL

Nous présentons ici la procédure MPL introduite dans [CCCL96] que nous utilisons pour le calcul de labels. L'objectif de cette procédure est le suivant : Étant donnée une formule ϕ, exprimée sous forme normale conjonctive (FNC) et comportant deux types distincts de littéraux, MPL calcule une projection de ϕ sur l'ensemble des littéraux d'un des deux types. Cette projection est une formule sous forme normale disjonctive (FND) qui est une représentation minimale de l'ensemble des modèles de ϕ, restreints à ce type de littéral. On peut montrer que les nogoods d'une base de connaissances (au sens des ATMS) peuvent être calculés facilement par cette procédure.
Les définitions qui suivent sont légèrement différentes de celles de [CCCL96], mais le principe reste le même. Cette nouvelle formulation qui facilite la lecture des preuves des théorèmes peut être retrouvée dans [CCCL98].

A.1 Définitions

On définit tout d'abord les notions d'impliquant et d'impliqué :

Définition A.1.1 (impliquants,impliqués) *soit ϕ une formule propositionnelle. Soient E et F deux ensembles cohérents de littéraux.*

– *La conjonction de littéraux (ou* phrase*) $E^\wedge$ est un* impliquant *de ϕ ssi $E^\wedge \vdash \phi$. $E^\wedge$ est un* impliquant premier *de ϕ ssi $E^\wedge \vdash \phi$ et $\nexists F^\wedge$, impliquant de ϕ tel que $F \subset E$.*

– *La disjonction de littéraux (ou* clause*) $E^\vee$ est un* impliqué *de ϕ ssi $\phi \vdash E^\vee$. $E^\vee$ est un* impliqué premier *de ϕ ssi $\phi \vdash E^\vee$ et $\nexists F^\vee$, impliqué de ϕ tel que $F \subset E$.*

Dans la suite, H représente un ensemble cohérent de littéraux distingués du langage $\mathcal{L}$ sur lequel seront projetées les formules de $\mathcal{L}$.

Définition A.1.2 (restriction) *Soit M un ensemble de littéraux correspondant à une interprétation. On appelle* restriction *de M à H l'ensemble $H \cap M$, noté $R_H(M)$.*

Si l'on considère M comme une conjonction de littéraux, $R_H(M)$ est une phrase dans laquelle seuls des littéraux de H apparaissent. On peut voir $R_H(M)$ comme une interprétation partielle.
La définition suivante concerne la notion de modèle *H-restreint* :

Définition A.1.3 (modèles restreints) *Soit ϕ une formule. Soit $I \subseteq H$. I est un modèle H-restreint de ϕ ssi $\exists M$, modèle de ϕ tel que $I = R_H(M)$.*

Le résultats suivants ont été montrés dans Castell et col. [CCCL96][CCCL98].

Proposition A.1.1 *I est un modèle H-restreint de ϕ ssi la phrase $(I \cup \sim (H - I))^\wedge$ est cohérente avec ϕ.*

Preuve :

> Puisque H est cohérent, si $(I \cup \sim (H - I))$ était incohérent, cela signifierait qu'il existe un littéral appartenant à la fois à I et à $H - I$.

Maintenant, $(I \cup \sim (H - I))^\wedge$ cohérent avec ϕ est équivalent à l'existence d'un modèle commun [1].

Théorème A.1.1 *I est un modèle H-restreint de ϕ ssi $\exists E^\wedge$, impliquant de ϕ tel que $H \cap E = I$.*

Preuve :

> Pour prouver la première implication il suffit de construire $E^\wedge$ en prenant une phrase bâtie sur un modèle commun de $I^\wedge$ et ϕ. Réciproquement, tout impliquant peut être étendu pour obtenir un modèle.

En fin de compte, un modèle H-restreint de ϕ peut être obtenu en prenant un modèle quelconque de ϕ et en "masquant" les littéraux de ce modèle qui ne sont pas dans H. Notons que la phrase obtenue n'est plus forcément un impliquant de ϕ.

Définition A.1.4 *Soit ϕ une formule. Soit $I \subseteq H$. I est un H-impliquant de ϕ ssi pour toute interprétation M telle que $I \subseteq M$, M est un modèle de ϕ.*

Un H-impliquant de ϕ est donc un modèle H-restreint qui implique également ϕ. On peut montrer que les H-impliquants de ϕ minimaux pour l'inclusion sont exactement les impliquants premiers de ϕ qui sont inclus dans H.

En fait, les modèles H-restreints minimaux d'une formule ϕ exprimée sous forme normale disjonctive peuvent être facilement calculés :

Théorème A.1.2 *L'ensemble des modèles H-restreints minimaux d'une FND $\phi = E_1^\wedge \vee E_2^\wedge \vee \ldots \vee E_n^\wedge$ est l'ensemble des éléments minimaux pour l'inclusion de $R_H(E_1), R_H(E_2), \ldots, R_H(E_n)$*

Preuve :

> On peut montrer facilement que les $R_H(E_i)$ sont des modèles H-restreints de ϕ. Supposons l'existence de E, un modèle H-restreint minimal de ϕ qui ne se trouve pas dans l'ensemble des éléments minimaux de $R_H(E_1), R_H(E_2), \ldots, R_H(E_n)$. Alors on peut trouver M, modèle de ϕ, tel que $R_H(M) = E$, et $\forall i, R_H(E) \neq R_H(E_i)$. Or, puisqu'il est impossible de trouver $R_H(E_i) \subset E$ (minimalité de E), il existe donc un littéral $l \in R_H(E_i)$ et $l \notin E$ pour tout i. On a $l \notin M$, parce que $l \in H$ et $E = R_H(M) = M \cap H$. Donc, M falsifie les $R_H(E_i)$ et donc ϕ, ce qui contredit l'hypothèse de départ.

1. Si $\mathcal{D}$ est l'ensemble des symboles impliqués dans H, $I \cup \sim (H - I))$ est exactement un $\mathcal{D}$-modèle préféré de ϕ dans le sens de [CCCL96].

Exemple A.1.1 *Soit $K = \{\neg A \vee c, \neg A \vee B, \neg B \vee \neg c \vee A\}$ et $H = \{\neg A, \neg B\}$. La forme normale disjonctive de K est $(\neg B \wedge \neg A) \vee (\neg A \wedge \neg c) \vee (A \wedge B \wedge c)$. Les modèles H-restreints sont $\{\neg A, \neg B\}$, $\{\neg A\}$ et $\{\}$. Le seul H-impliquant de K est $\{\neg A, \neg B\}$.*

On peut maintenant définir une fonction MP_H (resp. PI_H) qui à toute formule ϕ du langage associe une autre formule faite de littéraux de H, sous forme normale disjonctive et telle que les impliquants premiers de $MP_H(\phi)$ (resp. $PI_H(\phi)$) (i.e. les éléments de la disjonction) sont exactement les modèles H-restreints (resp. les H-impliquants) minimaux pour l'inclusion de ϕ.
ϕ, $MP_H(\phi)$ et $PI_H(\phi)$ sont reliés par les propriétés suivantes :

Propriété A.1.1 $\forall \phi, PI_H(\phi) \vdash \phi \vdash MP_H(\phi)$

Preuve :

Si $E^\wedge$ est un élément de la disjonction $PI_H(\phi)$, c'est un impliquant premier de ϕ, ce qui prouve le premier $\vdash$. Maintenant, tout modèle M de ϕ définit un modèle H-restreint de $R_H(M)$. Alors, il existe $E^\wedge$, élément de la disjonction $MP_H(\phi)$, tel que $E \subseteq R_H(M)$, donc le deuxième $\vdash$ est prouvé.

Théorème A.1.3 *Soit $E \subseteq H$. $\forall \phi, \phi \vdash E^\vee$ ssi $MP_H(\phi) \vdash E^\vee$.*

Preuve :

($\Leftarrow$) Soit M, modèle de ϕ. $R_H(M)$ est un modèle H-restreint de ϕ, donc il contient un élément de $MP_H(\phi)$. A fortiori, M contient un élément de $MP_H(\phi)$, donc M est un modèle de $MP_H(\phi)$. Puisque $MP_H(\phi) \vdash E^\vee$, M satisfait $E^\vee$.

($\Rightarrow$) Supposons $\phi \vdash E^\vee$. Supposons maintenant que l'on n'a pas $MP_H(\phi) \vdash E^\vee$. Alors, il existe $E_i^\wedge$ dans $MP_H(\phi)$ tel que $E_i \cap E = \emptyset$. Par définition, $\exists M$, modèle de ϕ tel que $E_i = R_H(M)$, donc $E_i \subseteq M$. Puisque $E \subseteq H$, $E \cap M \subseteq E \cap P \cap M = E \cap E_i = \emptyset$, donc M ne satisfait pas $E^\vee$, ce qui est contradictoire.

Théorème A.1.4 *Soit $E \subseteq H$. $\forall \phi, E^\wedge \vdash \phi$ ssi $E^\wedge \vdash PI_H(\phi)$.*

Preuve :

($\Leftarrow$) est évident.

($\Rightarrow$) Comme $E^\wedge \vdash \phi$, il contient un impliquant premier $F^\wedge$ de ϕ. Puisque $E \subseteq H$, $F^\wedge$ appartient à $PI_H(\phi)$, et donc $E^\wedge \vdash PI_H(\phi)$.

Corollaire A.1.1 *Si tous les impliquants premiers de ϕ sont dans H alors $MP_H(\phi) \equiv PI_H(\phi)$*

Ce corollaire est très important puisque c'est lui qui nous permet de calculer facilement les labels et les nogoods d'un ATMS en utilisant la fonction MP_H, lorsque H est l'ensemble des hypothèses de l'ATMS.

A.2 Principe de l'algorithme MPL

L'algorithme MPL est basé sur une procédure énumérative du type Davis et Putnam [DP60]. Le problème du calcul de la satisfaisabilité d'un ensemble de clauses (SAT) est un problème NP-complet. Parmi les algorithmes complets qui résolvent le problème SAT, certains parmi les plus efficaces, utilisent un algorithme du type Davis et Putnam [DP60], [DLL62].
Un algorithme de Davis et Putnam énumère les interprétations d'une base de connaissances K jusqu'à ce qu'il en trouve un modèle (si la base est cohérente) s'il y en a un (sinon, la base est incohérente). En opérant de la sorte, la recherche d'un modèle de la base de connaissances K est étroitement reliée à la recherche d'une forme normale disjonctive de K, puisque l'on obtient facilement les modèles de K depuis sa FND.
La différence principale entre l'algorithme de Davis et Putnam et l'algorithme MPL est que ce dernier ne se contente pas de chercher un modèle de K, mais en cherche tous les modèles H-restreints.

L'algorithme de Davis et Putnam construit un arbre de recherche binaire qui bifurque à chaque noeud sur la valeur de vérité d'un littéral. Si on définit un préordre entre les ensembles de littéraux, on peut ordonner les interprétations I en se basant sur le préordre entre leurs restrictions à H, $R_H(I)$: à chaque noeud on falsifie un littéral de H. Si M est la première interprétation trouvée satisfaisant la base K, $R_H(M)$ est un modèle H-restreint minimal (pour la relation d'ordre).
Le problème est maintenant d'éliminer les modèles H-restreints de K qui ne sont pas minimaux. L'idée utilisée dans l'algorithme est d'ajouter une clause que les modèles H-restreints non-minimaux violent dès que l'on a trouvé un modèle H-restreint minimal. Pour ceci, dès que l'on a trouvé un modèle M (le premier trouvé est forcément minimal), on ajoute à K la clause $C = \vee_{l \in R_H(M)} \neg l$. Toute interprétation I telle que $R_H(M) \subset R_H(I)$ falsifie C.

Les deux algorithmes suivants implémentent la procédure de Davis et Putnam modifiée. L'ensemble des clauses est divisé en deux : la base initiale K et une base de clauses ajoutées au cours du calcul, KA. Cet algorithme, appelé avec K, $KA = \{\}$ et H comme paramètres, calcule exactement $MP_H(K^\wedge)$.

Algorithme A.1 : MPL(K,KA,H)

% cette fonction retourne les modèles H-restreints minimaux de K ;
Données : K un ensemble de clauses ;
  KA ensemble de clauses ajoutées ;
  H ensemble de littéraux cohérents ;
Résultat : M ensemble de modèles H-restreints minimaux
début
  $M \leftarrow \{\}$ % M est une variable locale ;
  MODEL_PREF_LIT(K,KA,$\{\}$,M,H);
  retourner M;
fin

L'algorithme A.2 est une version modifiée de l'algorithme de Davis et Putnam. Il produit un arbre de recherche binaire qui est développé (les variables de H sont instanciées successivement) jusqu'à ce qu'un modèle de K soit trouvé. L'algorithme de Davis et Putnam "normal" s'arrête à ce point, mais puisque nous voulons trouver *tous* les modèles H-restreints minimaux de K, on doit poursuivre la recherche.

Algorithme A.2 : MODEL_PREF_LIT(K, var KA, IP, var M,H)

Données : K un ensemble de clauses ;
KA ensemble de clauses ajoutées ;
IP ensemble cohérent de littéraux ;
H ensemble cohérent de littéraux ;
Résultat : M un ensemble de modèles H-restreints minimaux
début
  SIMPLIFY(K,IP) ;
  % $\rightarrow$ est la clause vide ;
  si $(K)_{IP} = \{\}$ *et* $\rightarrow \notin (KA)_{IP}$ **alors**
    % on a trouvé un impliquant de K ;
    $M \leftarrow M \cup \{IP \cap H\}$;
    $KA \leftarrow KA \cup \{\vee_{l \in IP \cap H)} \neg l\}$;
  sinon
    si $\rightarrow \notin (K \cup KA)_{IP}$ **alors**
      % on cherche un littéral de H dans K ;
      $l \leftarrow$ CHOOSE_LITERAL(K,IP,H) ;
      si $l \neq$ *"rien"* **alors**
        % on commence par la négation du littéral ;
        MODEL_PREF_LIT(K,KA,$IP \cup \{\neg l\}$,M,H);
        MODEL_PREF_LIT(K,KA,$IP \cup \{l\}$,M, H);
      sinon
        % on effectue un test de cohérence ;
        si $(K \cup KA)_{IP}$ *cohérent* **alors**
          $M \leftarrow M \cup \{IP \cap H\}$;
          $KA \leftarrow KA \cup \{\vee_{l \in IP \cap H} \neg l\}$;
fin

Dans la fonction SIMPLIFY on peut utliser les techniques classiques des algorithmes de type Davis et Putnam pour la simplification de bases de clauses (propagation des clauses unitaires...). La fonction CHOOSE_LITTERAL retourne un littéral de H auquel on n'a pas encore affecté de valeur, s'il en existe un et sinon, le mot "rien". Pour choisir le prochain littéral à instancier, on se peut se servir de méthodes initiées dans la communauté SAT.

Exemple A.2.1 *Soit $K = \{B \vee c, \neg c \vee A \vee B\}$. Soit $H = \{A, B\}$. La figure A.1 montre l'arbre binaire développé par notre algorithme. Le fonctionnement peut être séparé en deux étapes : en premier lieu on utilise l'algorithme de Davis et Putnam décrit précédemment, afin d'affecter une valeur à tous les littéraux de H impliqués dans la formule. Ensuite, en s'appuyant sur le théorème A.1.1 on cherche simplement à savoir si la formule résultante est cohérente. Ce test de cohérence est effectué par un appel à une procédure de Davis et Putnam classique. Si la formule est cohérente, on a trouvé un modèle H-restreint minimal.*

La restriction à H du premier modèle trouvé est $\{B\}$. On ajoute donc à la base la clause $\neg B$. On

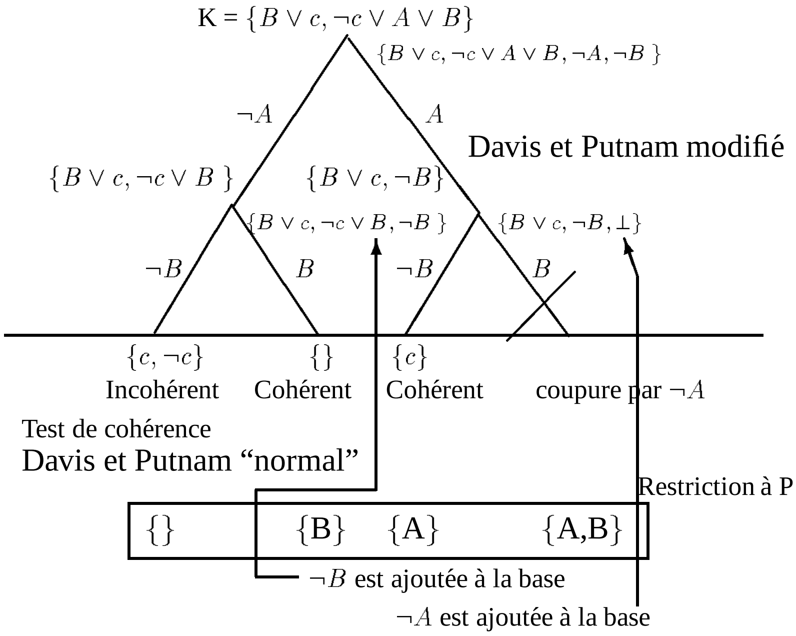

FIG. A.1 – *Arbre de Davis et Putnam modifié*

poursuit, on ajoute la clause $\neg A$, ce qui permet de couper la dernière branche. Les deux modèles H-restreints minimaux sont donc $\{A\}$ et $\{B\}$.

A.3 Application aux ATMS

Tous les éléments de base d'un ATMS : labels, nogoods, etc. peuvent être calculés par l'algorithme MPL et ce sans utiliser la phase de minimisation présente dans les algorithmes de de Kleer [de 86] puisque les modèles calculés par MPL sont minimaux.

Propriété A.3.1 *L'ensemble des nogoods d'une base de connaissances K par rapport à l'ensemble d'hypothèses H est exactement l'ensemble des impliquants premiers de* $MP_H(\neg MP_{\sim H}(K^{\wedge}))$.

Preuve :

E est un nogood ssi $E \subseteq \mathcal{H}$ est minimal et tel que $K^{\wedge} \wedge E^{\wedge} \vdash \bot$, ou $K^{\wedge} \vdash (\sim E)^{\vee}$.
$K^{\wedge} \vdash (\sim E)^{\vee}$ ssi $MP_{\sim H}(K^{\wedge}) \vdash (\sim E)^{\vee}$ (théorème A.1.3),
$MP_{\sim H}(K^{\wedge}) \vdash (\sim E)^{\vee} \equiv E^{\wedge} \vdash \neg MP_{\sim H}(K^{\wedge})$
$E^{\wedge} \vdash \neg MP_{\sim H}(K^{\wedge})$ ssi $E^{\wedge} \vdash PI_H(\neg MP_{\sim H}(K^{\wedge}))$ (théorème A.1.4).
Ou, de manière équivalente, $E^{\wedge} \vdash MP_H(\neg MP_{\sim H}(K^{\wedge}))$ (corollaire A.1.1 puisque les littéraux de $\neg MP_{\sim H}(K^{\wedge})$ sont dans H). Finalement, un nogood est exactement un impliquant premier de $MP_H(\neg MP_{\sim H}(K^{\wedge}))$.

Propriété A.3.2 *Soit K un ensemble de clauses et un ensemble d'hypothèses H . Soit ϕ une formule. Le label de ϕ contient exactement l'ensemble des impliquants premiers de* $MP_H(\neg MP_{\sim H}(K^{\wedge} \wedge \neg\phi))$ *qui ne sont pas parmi les nogoods de K.*

La preuve est similaire à la précédente puisque l'on cherche les nogoods de $K \cup \neg\phi$. On doit ôter les nogoods du label puisque toute formule est conséquence logique d'un environnement incohérent. Ceci peut être fait (après un calcul préalable des nogoods) en initialisant la base de clauses ajoutées (KA) avec l'ensemble des nogoods, à l'appel de la procédure MPL.

Propriété A.3.3 *Soit K un ensemble de clauses. Soient ϕ et ψ deux formules. Le label de $\phi \wedge \psi$ est exactement l'ensemble des impliquants premiers de $MP_H(\neg MP_{\sim H}(K \wedge \neg\phi) \wedge \neg MP_{\sim H}(K \wedge \neg\psi))$ (duquel on enlève les nogoods de K).*

Exemple A.3.1 *(suite)*

$K = \{\neg A \vee \neg b \vee c, \neg B \vee b, \neg c\}$
$MP_{\sim H}(K) = \neg A \vee \neg B$ *(*$\{\{\neg A\}, \{\neg B\}\}$*),*
$\neg MP_{\sim H}(K) = A \wedge B$ *(*$\{\{A\}, \{B\}\}$*),*
$MP_H(\neg MP_{\sim H}(K)) = A \wedge B$ *(*$\{\{A, B\}\}$*).*
Donc, l'ensemble des nogoods est $\{\{$*A,B*$\}\}$.

A.4 Interface du prototype

Ce prototype, implémentant le modèle de représentation logique de problèmes de décision dans l'incertain a été réalisé, dans le cadre de sa thèse, par D. Le Berre. Les deux figures suivantes représentent les interfaces de saisie des bases de connaissances et de préférences, ainsi que le résultat d'un calcul de décisions pessimistes optimales.

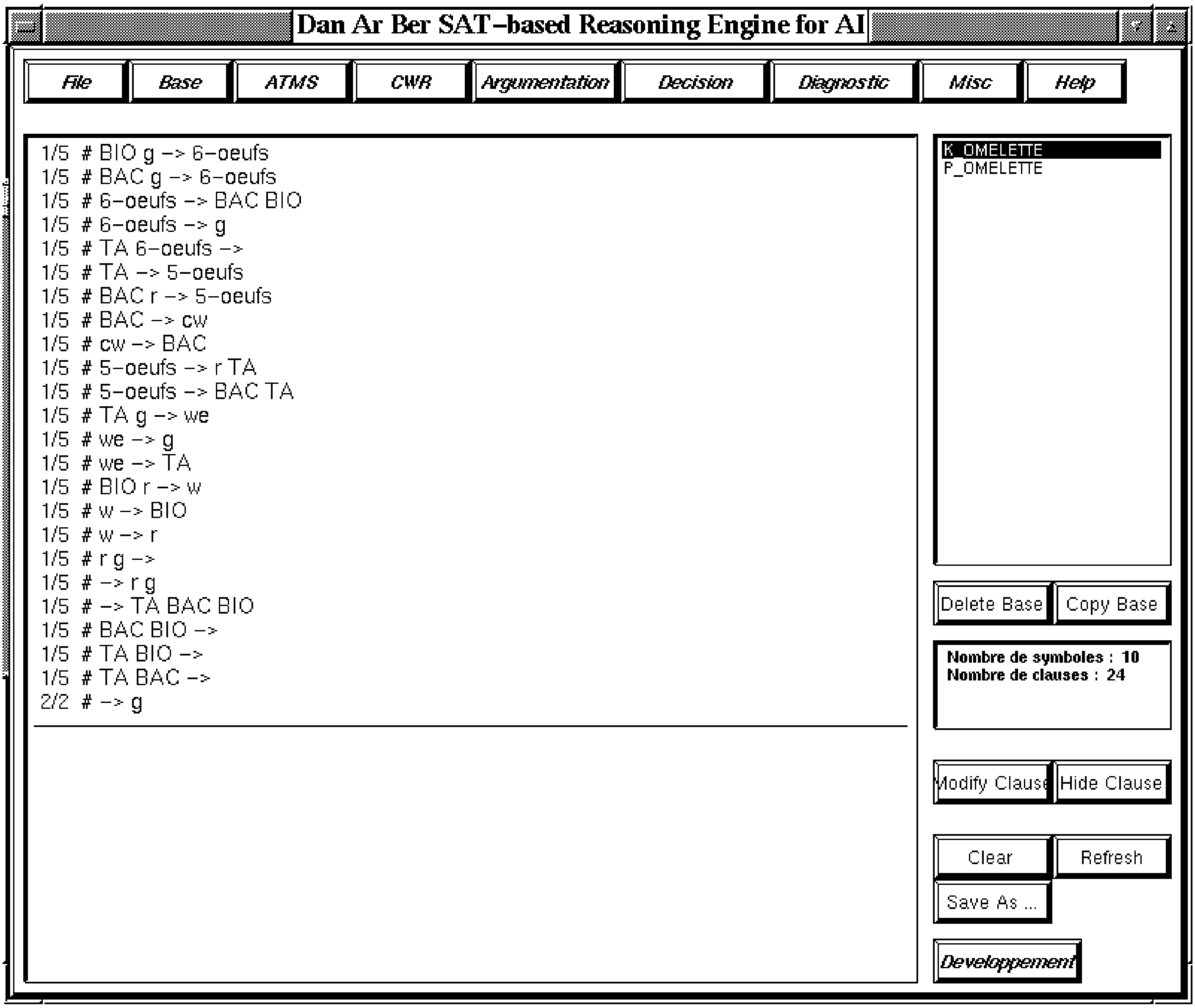

FIG. A.2 – *Interface de saisie de la base de connaissances.*

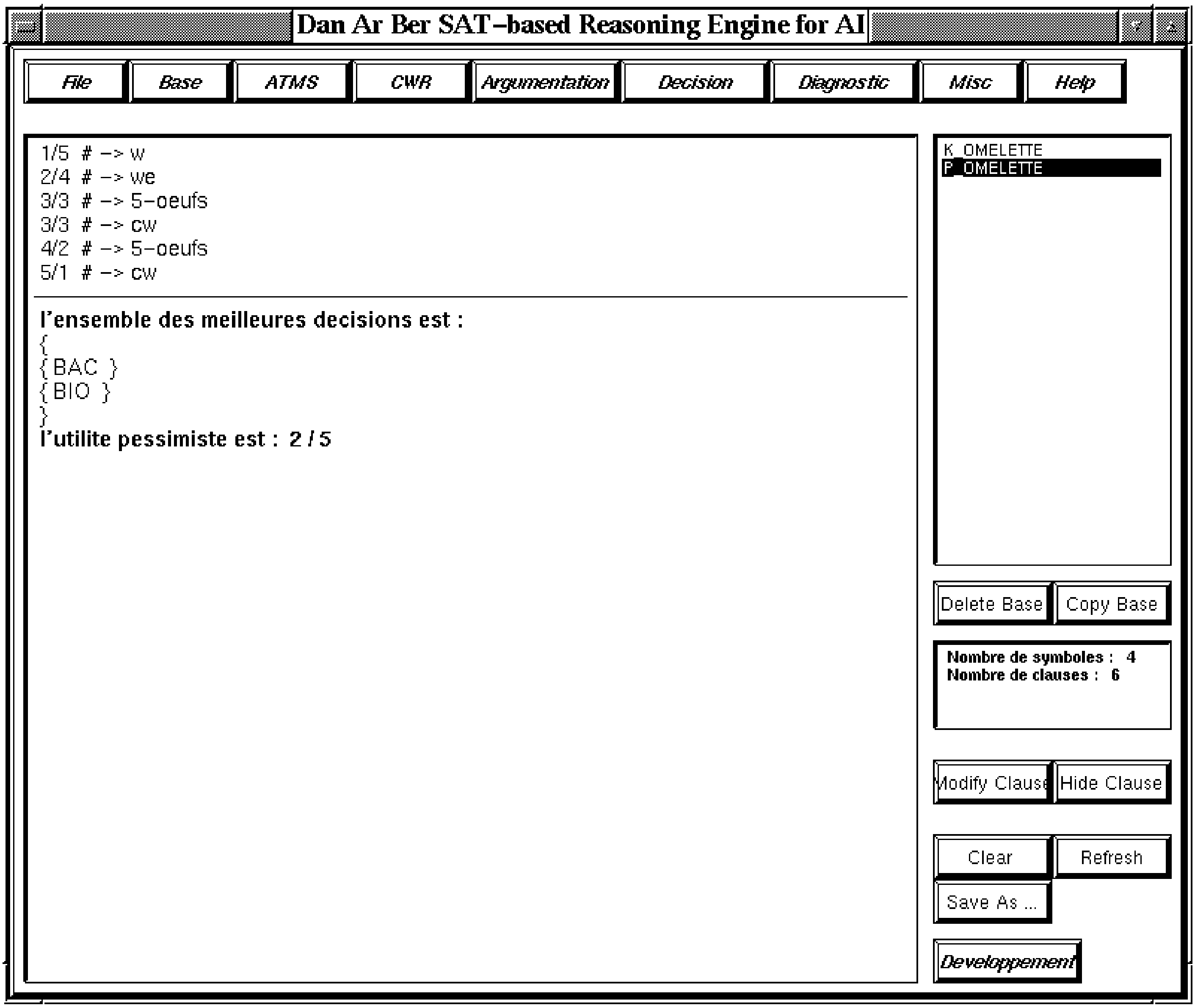

FIG. A.3 – *Base de préférences et décisions optimales.*

Annexe B

Décision dans l’incertain et contraintes flexibles

Bibliographie

[AA63] Ansccombe (A.) et Aumann (R. J.). – A definition of subjective probability. *Annals of Mathematical Statistics*, vol. 34, Mars 1963, pp. 199–205.

[Aki95] Akian (M.). – *Theory of cost measures: convergence of decision variables*. – Rapport technique n° 2611, Rocquencourt, France, INRIA, Juillet 1995.

[All53] Allais (M.). – Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine. *Econometrica*, vol. 21, 1953, pp. 503–546.

[AQV94] Akian (M.), Quadrat (J. P.) et Viot (M.). – Bellman processes. *In : Proc. 11th Int. Conf. on Analysis and Optimization of Discrete Events Systems*.

[Arr51] Arrow (K.). – *social Choice and Individual values*. – Wiley, New York, 1951.

[BBGP97] Boutilier (C.), Brafman (R.), Geib (C.) et Poole (D.). – A constraint-based approach to preference elicitation and decision making. *In : Proc. AAAI 97 Spring Symposium Series (Qualitative Preferences in Deliberation and Practical Reasoning)*. – Standford University, CA, 24-26 Mars 1997.

[BDGP98] Benferhat (S.), Dubois (D.), Garcia (L.) et Prade (H.). – Directed possibilistic graphs and possibilistic logic. *In : Proc. 7th Inter. Conf on Information Processing and Management of Uncertainty in knowledge based systems*, pp. 1470–1477. – Paris, France, 6-10 Juillet 1998.

[BDP92] Benferhat (S.), Dubois (D.) et Prade (H.). – Representing default rules in possibilistic logic. *In : Proc. 3rd Inter. Conf. on Principles of Knowledge Representation and Reasonning (KR'92)*, éd. par Nebel (B.), Rich (C.) et Swartout (W.), pp. 673–684. – Cambridge MA, 25-29 Oct. 1992.

[Bel57] Bellman (R. E.). – *Dynamic Programming*. – Princeton University Press, Princeton N.J., 1957.

[Ber96] Bernhard (P.). – A separation theorem for expected value and feared value discrete time control. *European Series in Applied and Industrial Mathematics : Contrôle, Optimisation et Calcul des Variations*, vol. 1, Juillet 1996, pp. 191–206.

[BG96] Bonet (B.) et Geffner (H.). – Arguing for decisions: A qualitative model of decision making. *In : Proc. 12th Con. on Uncertainty in Artificial Intelligence (UAI'96)*, éd. par E. Horwitz (F. Jensen eds.), pp. 98–105. – Portland, Oregon, 31 juillet-4 août 1996.

[Bou94] Boutilier (C.). – Toward a logic for qualitative decision theory. *In : Proc. 4th Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'94)*, éd. par J. Doyle, E. Sandewall (P. Torasso eds.), pp. 75–86. – Bonn, Allemagne, 24-27 mai 1994.

[BP82] Baldwin (J. F.) et Pilsworth (B. W.). – Dynamic programming for fuzzy systems with fuzzy environments. *J. Math. Analysis and Applications*, no85, 1982, pp. 1–23.

[Bra97] Brafman (R. I.). – A heuristic variable grid solution method for pomdps. *In : Proc. of 13th Nat. Conf. on Artificial Intelligence (AAAI'97)*, pp. 727–733.

[BT96] Brafman (R. I.) et Tennenholtz (M.). – On the foundations of qualitative decision theory. *In : Proc. 12th National Conf. on Artificial Intelligence*, pp. 1291–1296. – Portland, Oregon, 4-8 Août 1996.

[BT97] Brafman (R. I.) et Tennenholtz (M.). – On the axiomatization of qualitative decision criteria. *In : Proc. 13th National Conf. on Artificial Intelligence*, pp. 76–81.

[BZ70] Bellman (R. E.) et Zadeh (L. A.). – Decision making in a fuzzy environment. *Management Science*, vol. B, n° 17, 1970, pp. 141–164.

[CCCL96] Castell (T.), Cayrol (M.), Cayrol (C.) et Le Berre (D.). – Using the Davis and Putnam procedure for an efficient computation of preferred models. *In : Proc. 12th European Conf. on Artificial Intelligence (ECAI'96)*, pp. 350–354. – Budapest, Hongrie, 11-16 Août 1996.

[CCCL98] Castell (T.), Cayrol (M.), Cayrol (C.) et Le Berre (D.). – Modèles p-restreints. application à l'inférence propositionnelle. *In : Actes du 11eme congrès Reconnaissance des Formes et Intelligence Artificielle (RFIA'98)*, pp. 205–214. – Clermont-Ferrand, 20-22 Janvier 1998.

[Cha96] Chateauneuf (A.). – Decomposable measures, distorted probabilities and concave capacities. *Math. Soc. Sci.*, no31, 1996, pp. 19–37.

[Che88] Cheng (H. T.). – *Algorithm for Partially Observable Markov Decision Processes*. – Canada, Thèse de PhD, University of British Columbia, 1988.

[Cho54] Choquet (G.). – Theory of capacities. *Annales de l'Institut Fourier*, no5, 1953-1954, pp. 131–295. – (Grenoble).

[Chr92] Chrisman (L.). – Reinforcement learning with perceptual aliasing: The perceptual distinctions approach. *In : Proc. 10th Nat. Conf. on Artificial Intelligence (AAAI'92)*. pp. 183–188. – San Jose, CA, 1992.

[CJ80] Cohen (M.) et Jaffray (J-Y). – Rational behavior under complete ignorance. *Econometrica*, vol. 48, n° 5, 1980, pp. 1281–1299.

[CKL94] Cassandra (A. R.), Kaelbling (L. P.) et Littman (M. L.). – Acting optimally in partially observable stochastic domains. *In : Proc. 11th Nat. Conf. on Artificial Intelligence (AAAI'94)*. pp. 1023–1028. – Seattle, WA, 31 Juillet-4août 1994.

[Da 98] Da Costa Pereira (C.). – *Planification d'actions en environnement incertain : une approche fondée sur la théorie des possibilités*. – Toulouse, France, Thèse de PhD, Université Paul Sabatier, Mai 1998.

[dB92] de Campos (L. L.) et Bolanos (M. J.). – Characterization and comparison of sugeno and choquet integrals. *Fuzzy Sets and Systems*, no52, 1992, pp. 61–67.

[DB94] Dearden (R.) et Boutilier (C.). – Using abstraction for decision-theoretic planning with time constraints. *In : Proc. 11th Nat. Conf. on Artificial Intelligence (AAAI'94)*. – Seattle, WA, 31 Juil.-4 août 1994.

[de 86] de Kleer (J.). – An Assumption-based TMS et Extending the ATMS . *Artificial Intelligence*, vol. 28, 1986, pp. 127–162 and 163–196.

[Dem67] Dempster (A. P.). – Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, no38, 1967, pp. 325–339.

[dF37] de Finetti (B.). – La prévision : ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré*, no7, 1937, pp. 1–68.

[DFL+96] Dubois (D.), Fargier (H.), Lang (J.), Prade (H.) et Sabbadin (R.). – Qualitative decision theory and multistage decision making: A possibilistic approach. *In : Proc. European Workshop on Fuzzy Decision Analysis for Management, Planning and Optimization*, éd. par Felix (R.), pp. 1–6. – Dortmund, Germany, 21-22 Mai 1996.

[DFP96] Dubois (D.), Fargier (H.) et Prade (H.). – Possibility theory in constraint satisfaction problems. *Applied Intelligence*, no6, 1996, pp. 287–309.

[DFP97] Dubois (D.), Fargier (H.) et Prade (H.). – Decision-making under ordinal preferences and uncertainty. *In : Actes 13eme conf. on Uncertainty in Artificial Intelligence (UAI'97)*, éd. par D. (Geiger) et P. (Shenoy P.). pp. 157–164. – Providence, R. I., 1997.

[DFPR96] Dubois (D.), Fodor (J. C.), Prade (H.) et Roubens (M.). – Aggregation of decomposable measures with application to utility theory. *Theory and Decision*, no41, 1996, pp. 59–95.

[DGLM97a] Da Costa Pereira (C.), Garcia (F.), Lang (J.) et Martin-Clouaire (R.). – Planning with graded nondeterministic actions: a possibilistic approach. *Int. Journal of Intelligent Systems*, no12, 1997, pp. 935–962.

[DGLM97b] Da Costa Pereira (C.), Garcia (F.), Lang (J.) et Martin-Clouaire (R.). – *Recent Advances in AI Planning*, chap. Possibilistic planning: representation and complexity, pp. 143–155. – Springer Verlag, S. Steel and R. Alami, 1997, *Lecture Notes in Artificial Intelligence*.

[DGPZ98] Dubois (D.), Godo (L.), Prade (H.) et Zapico (A.). – Possibilistic representation of qualitative utility: an improved characterization. *In : Proc. of the Int. Conf. on Information Processing and Management of Uncertainty* (IPMU'98). – Paris, Juillet 1998.

[DLL62] Davis (H.), Logemann (G.) et Loveland (D.). – A machine program for theorem proving. *Commun. ACM*, no5, 1962, pp. 394–397.

[dLM91] de Campos (L. L.), Lamata (M. T.) et Moral (S.). – A unified approach to define fuzzy integrals. *Fuzzy Sets and Systems*, no39, 1991, pp. 75–90.

[DLP94] Dubois (D.), Lang (J.) et Prade (H.). – Automated reasoning using possibilistic logic: Semantics, belief revision and variable certainty weights. *IEEE Trans. on Knowledge and Data Engineering*, vol. 1, n° 6, 1994, pp. 64–69.

[DLPS98a] Dubois (D.), Le Berre (D.), Prade (H.) et Sabbadin (R.). – Logical representation and computation of optimal decisions in a qualitative setting. *In : Proc. AAAI'98*. pp. 588–593. – Madison, Wisconsin, 26-30 Juillet 1998.

[DLPS98b] Dubois (D.), Le Berre (D.), Prade (H.) et Sabbadin (R.). – Using possibilistic logic for modeling qualitative decision: ATMS-based algorithms. *Fundamenta Informaticae*, 1998. – Accepté pour publication.

[DP60] Davis (H.) et Putnam (L.). – A computing procedure for quantification theory. *Journal of the ACM*, no7, 1960, pp. 201–215.

[DP80] Dubois (D.) et Prade (H.). – *Fuzzy Sets and Systems: Theory and application*. – New York, N. Y., Academic Press, 1980.

[DP82] Dubois (D.) et Prade (H.). – A class of fuzzy measures based on triangular norms. *Int. J. of General Systems*, no8, 1982, pp. 225–233.

[DP88] Dubois (D.) et Prade (H.). – *Possibility theory*. – Plenum Press, 1988.

[DP95a] Dubois (D.) et Prade (H.). – Possibilistic mixtures and their applications to qualitative utility theory. part ii: decision under incomplete knowledge. *In : Proc. of conf. Foundations and Applications of Possibility Theory (FAPT'95)*, éd. par de Cooman (G.), Ruan (D.) et Kerre (E. E.). pp. 256–266. – World Scientific.

[DP95b] Dubois (D.) et Prade (H.). – Possibility theory as a basis for qualitative decision theory. *In : Proc. 14th Inter. Joint Conf. on Artificial Intelligence (IJCAI'95)*, pp. 1925–1930. – Montreal, Canada, 20-25 août 1995.

[DP95c] Dubois (D.) et Prade (H.). – Towards possibilistic decision theory. *In : IJCAI'95 workshop-Fuzzy logic in Artificial Intelligence*, éd. par T. P. Martin (A. L. Ralescu). pp. 240–251. – Springer.

[DPS97a] Dubois (D.), Prade (H.) et Sabbadin (R.). – Decision under qualitative uncertainty with Sugeno integrals-an axiomatic approach. *In : Actes du 7eme Congrès Mondial IFSA'97*. pp. 441–446. – Prague, 1997.

[DPS97b] Dubois (D.), Prade (H.) et Sabbadin (R.). – A possibilistic logic machinery for qualitative decision. *In : Proc. AAAI 97 Spring Symposium Series (Qualitative Preferences in Deliberation and Practical Reasoning)*, pp. 47–54. – Standford University, CA, 24-26 Mars 1997.

[DPS98] Dubois (D.), Prade (H.) et Sabbadin (R.). – Qualitative decision theory with Sugeno integrals. *In : Actes de la 14eme Conf. Uncertainty in Artificial Intelligence (UAI'98)*, éd. par Cooper (G. F.) et Moral (S.). pp. 121–128. – Madison, WI, 24-26 Juillet 1998.

[DSW91] Doyle (J.), Shoham (Y.) et Wellman (M. P.). – A logic of relative desire (preliminary report). *In : Proc. 6eme International Symposium on methodologies for Intelligent Systems*, pp. 16–31.

[Dub86] Dubois (D.). – Belief structures, possibility theory and decomposable confidence measures on finite sets. *Computers and Artificial Intelligence*, no5, 1986, pp. 404–416.

[Dup96] Dupin de Saint-Cyr (F.). – *Gestion de l'évolutif et de l'incertain en logiques pondérées*. – Toulouse, Thèse de PhD, Université Paul Sabatier, 1996.

[DW94] Doyle (J.) et Wellman (M. P.). – Representing preferences as *ceteris paribus* comparatives. *In : AAAI Spring Symposium on decision theoretic planning*, pp. 69–75.

[EK98] Esogbue (O.) et Kacprzyk (J.). – *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, chap. Fuzzy Dynamic Programming, pp. 281–307. – Kluwer Academic Publishers, R. Slowinski, 1998.

[Ell61] Ellsberg (D.). – Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, no75, 1961, pp. 643–669.

[Far94] Fargier (H.). – *Problèmes de satisfaction de contraintes flexibles, application à l'ordonnancement de production*. – Thèse de PhD, Université Paul Sabatier, Toulouse, 1994.

[FLS93] Fargier (H.), Lang (J.) et Schiex (T.). – Selecting preferred solutions in fuzzy constraint satisfaction problems. *In : Proc. 1st European Congress on Fuzzy and Intelligent Technologies (EUFIT'93)*. pp. 1128–1134. – Aachen, 1993.

[FLS96a] Fargier (H.), Lang (J.) et Sabbadin (R.). – Towards qualitative approaches to multi-stage decision making. *In : Proc. 6th International Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'96)*, pp. 31–36. – Granada, Spain, 1-5 juillet 1996.

[FLS96b] Fargier (H.), Lang (J.) et Schiex (T.). – Mixed constraint satisfaction: a framework for decision problems under uncertainty. *In : Proc. 13th National Conf. on Artificial Intelligence (AAAI'96)*, pp. 175–180.

[FLS98] Fargier (H.), Lang (J.) et Sabbadin (R.). – Towards qualitative approaches to multi-stage decision making. *International Journal of Approximate Reasoning*, 1998. – Version révisée et rallongée (36 pages).

[FN71] Fikes (R. E.) et Nilsson (N. J.). – Partial constraint satisfaction. *Artificial Intelligence*, vol. 2, 1971, pp. 189–208.

[For97] Fortemps (P.). – *Fuzzy Sets for Modelling and Handling Imprecision and Flexibility*. – Thèse de PhD, Faculté Polytechnique de Mons (Belgique), 1997.

[Gef92] Geffner (H.). – *Default reasoning: Causal and conditional theories*. – MIT Press, 1992.

[Gef98] Geffner (H.). – Markov decision processes for planning. *In : ECAI'98 workshop : decision theory meets AI*, pp. 5–9. – Brighton, UK, 1998.

[Gil87] Gilboa (I.). – Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, no16, 1987, pp. 65–88.

[GKP97] Grant (S.), Kajii (A.) et Polak (B.). – Weakening the sure-thing principle: Decomposable choice under uncertainty. *In : Workshop on Decision Theory*. – Chantilly, France, Juin 1997.

[GP91] Goldszmidt (M.) et Pearl (J.). – System z^+: A formalisation for reasoning with variable strength defaults. *In : Proc. National Conf. of Artificial Intelligence (AAAI'91)*, pp. 399–404. – Anaheim, CA, 1991.

[Gue98] Guere (E.). – *Etude et développement d'un planificateur possibiliste/non-déterministe*. – Rapport technique, Toulouse, France, Laboratoire d'Analyse et d'Architecture des Systèmes, 1998.

[Hau97] Hauskrecht (M.). – Incremental methods for computing bounds in partially observable markov decision processes. *In : 13th Nat. Conf. on Artificial Intelligence (AAAI'97)*, pp. 734–739.

[HCd92] Hamscher (W.), Console (L.) et de Kleer (eds.) (J.). – *Readings in Model-Based Diagnosis*. – San Mateo, CA, Morgan Kaufmann, 1992.

[Hur51] Hurwicz (L.). – *Optimality Criteria for Decision Making Under Ignorance*. – Cowles Comission Discussion Paper, Statistics, 1951.

[IIT89] Inuiguishi (M.), Ichihashi (H.) et Tanaka (H.). – Possibilistic linear programming with multiattribute value functions. *ORSA Journal on Computing*, no1, 1989, pp. 146–158.

[Jaf89] Jaffray (J.-Y.). – Linear utility theory for belief functions. *Operations Research Letters*, no8, 1989, pp. 107–112.

[JW94] Jaffray (J.-Y.) et Wakker (P.). – Decision making with belief functions : Compatibility and incompatibility with the sure-thing principle. *Journal of Risk and Uncertainty*, no8, 1994, p. 1994. – Kluwer Academic Publishers.

[Kac83] Kacprzyk (J.). – *Multi-stage Decision Making under Fuzzyness*. – Verlag TÜV Rheinland, 1983.

[Kac98] Kacprzyk (J.). – *Multistage Fuzzy Control*. – Wiley, 1998.

[Kas93] Kast (R.). – *La théorie de la décision*. – Paris, La Découverte, 1993.

[KHW95] Kushmerick (N.), Hanks (S.) et Weld (D. W.). – An algorithm for probabilistic planning. *Artificial Intelligence*, 1995, pp. 239–286.

[KM95] Kohlas (J.) et Monney (P.A.). – *A Mathematical Theory of Hints. An Approach to the Dempster-Shafer Theory of Evidence*, chap. chapter 5 - Probabilistic Assumption-based Reasonning. – Springer, 1995, *Lecture Notes in Economics and Mathematical Systems*.

[KPS59] Kraft (C. H.), Pratt (J. W.) et Seidenberg (A.). – Intuitive probability on finite sets. *Annals of Mathematical Statistics*, no30, 1959, pp. 408–419.

[Lan96] Lang (J.). – Conditional desires and utilities-an alternative logical approach to qualitative decision theory. *In : Proc. 12th European Conference on Artificial Intelligence (ECAI'96)*, éd. par Wahlster (W.), pp. 318–322. – Budapest, Hongrie, 11-16 Août 1996.

[Leh96] Lehmann (D.). – Generalized qualitative probability: Savage revisited. *In : Proc. 12th Conf. Uncertainty in Artificial Intelligence* UAI'98. pp. 381–388. – Portland, Oregon, 1996.

[Lov91] Lovejoy (W. S.). – A survey of algorithmic methods for partially observed markov decision processes. *Annals of Operation Research*, vol. 1, n° 28, 1991, pp. 47–65.

[LR57] Luce (R.) et Raiffa (H.). – *Games and Decisions*. – New York, Wiley and Sons, 1957.

[LS97] Le Berre (D.) et Sabbadin (R.). – Decision-theoretic diagnosis and repair: representational and computational issues. *In : Proc. 8th International Workshop on Principles of Diagnosis (DX'97)*, pp. 141–145. – Le Mont-Saint-Michel, France, 15-18 Septembre 1997.

[Mak93] Makinson (D.). – Five faces of minimality. *Studia logica*, no52, 1993, pp. 339–379.

[Mar97] Marichal (J. L.). – *On Sugeno integral as an aggregation function*. – Rapport technique, Liège, Belgique, GEMME, Faculté d'Economie, de Gestion et de Sciences Sociales, 1997.

[Mas87] Maslov (V.). – *Méthodes opératorielles*. – Moscou, MIR, 1987.

[MH69] Mac Carthy (J.) et Hayes (P. J.). – Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, vol. 4, 1969, pp. 463–502. – B. Meltzer et D. Mitchie eds.

[Mil54] Milnor (J. W.). – *Games against nature*, pp. 49–60. – Thrall, Coombs and Davis, 1954.

[Mon82] Monahan (G. E.). – A survey of partially observable markov decision processes. *Management Science*, vol. 1, n° 28, 1982, pp. 1–16.

[Mou88] Moulin (H.). – *Axioms of Cooperative Decision Making*. – New York, Wiley, 1988.

[Pea88] Pearl (J.). – *Probabilistic Reasoning in Intelligent Systems*. – San Mateo, CA: Morgan Kaufmann, 1988.

[Pea90] Pearl (J.). – System z: A natural ordering of defaults with tractable applications to default reasonning. *In : Proc. 3rd Conf. on Theoretical Aspects of Reasonning About Knowledge (TARK'90)*. pp. 121–135. – Morgan Kaufmann.

[Pea93] Pearl (J.). – From conditionnal oughts to qualitative decision theory. *In : Proc. of the 9th Conf. on Uncertainty in Artificial Intelligence (UAI'93)*, éd. par Heckerman (D.) et Mamdani (A.). pp. 12–20. – San Mateo, CA, 1993.

[Pin91] Pinkas (G.). – Propositional nonmonotonic reasoning and inconsistency in symetric neural networks. *In : Proc. of the 12th International Joint Conf. on Artificial Intelligence (IJCAI'91)*. pp. 525–530. – Sydney, Australia, 1991.

[PR95] Parr (R.) et Russel (S.). – Approximating optimal policies for partially observable stoc. *In : Int. Joint Conf. on Artificial Intelligence (IJCAI'95)*, pp. 1088–1094. – Montreal, Canada, 1995.

[Pra82] Prade (H.). – *Fuzzy Sets and Possibility Theory: Recent Developments*, chap. Model semantics and fuzzy set theory, pp. 232–246. – R. R. Yager. Pergamon Press, 1982.

[Pro89] Provan (G. M.). – An analysis of atms-based techniques for computing dempster-shafer belief functions. *In : Proc. of the 11th Inter. Joint Conf. on Artificial Intelligence (IJ-CAI'89)*, pp. 1115–1120.

[PT87] Papadimitriou (C. H.) et Tsitsiklis (J. N.). – The complexity of markov decision processes. *Mathematics of Operations Research*, vol. 3, n° 12, 1987, pp. 441–450.

[Put87] Puterman (M. L.). – *Encyclopedia of Physical Science and Technology*, chap. Dynamic Programming, pp. 438–463. – Academic Press, 1987.

[Put94] Puterman (M. L.). – *Markov Decision Processes*. – New York, John Wiley and Sons, 1994.

[Roy85] Roy (B.). – *Méthodologie multicritère d'aide à la décision*. – Paris, Economica, 1985.

[RS96] Ralescu (D.) et Sugeno (M.). – Fuzzy integral representation. *Fuzzy Sets and Systems*, no84, 1996, pp. 127–133.

[Sab98] Sabbadin (R.). – Decision as abduction. *In : Proc. 13th European Conf. on Artificial Intelligence (ECAI'98)*, pp. 600–604. – Brighton, UK, 1998.

[Sav51] Savage (L. J.). – The theory of statistical decision. *Journal of the American Statistical Association*, no46, 1951, pp. 55–67.

[Sav54] Savage (L. J.). – *The Foundations of Statistics*. – New York, J. Wiley and Sons, 1954.

[Sch89] Schmeidler (D.). – Subjective probability and expected utility without additivity. *Econometrica*, no57, 1989, pp. 571–587.

[SDGP95] Sabbadin (R.), Dubois (D.), Grenier (P.) et Prade (H.). – Un exemple de problème de satisfaction de contraintes flexibles dans l'industrie vinicole. *In : Actes des rencontres francophones sur la Logique Floue et ses applications (LFA'95)*. pp. 265–272. – Cépadu'es Editions.

[SDGP98] Sabbadin (R.), Dubois (D.), Grenier (P.) et Prade (H.). – A fuzzy constraint satisfaction problem in the wine industry. *Inter. Journal of Intelligent and Fuzzy Systems*, no37, 1998. – à paraître.

[Sha76] Shafer (G.). – *A Mathematical Theory of Evidence*. – Princeton, NJ, Princeton University Press, 1976.

[Spo88] Spohn (W.). – Ordinal conditional functions: A dynamic theory of epistemic states. *In : Causation in Decision, Belief Change and Statistics*, éd. par Harper (W. L.) et Skyrms (B.), pp. 105–134. – Reidel, Dordrecht, 1988.

[SS73] Smallwood (R. D.) et Sondik (E. J.). – The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, no21, 1973, pp. 1071–1088.

[SS83] Schweitzer (B.) et Sklar (A.). – *Probabilistic Metric Spaces*. – Amsterdam, North-Holland, 1983.

[Sug77] Sugeno (M.). – *Fuzzy Automata and Decision Processes*, chap. Fuzzy measures and fuzzy integrals - A survey, pp. 89–102. – North-Holland, Amsterdam, M. M. Gupta, G. N. Saridis et B. R. Gaines éditeurs, 1977.

[SW92] Sarin (R.) et Wakker (P.). – A simple axiomatization of nonadditive expected utility. *Econometrica*, vol. 60, n° 6, 1992, pp. 1255–1272.

[TP94] Tan (S. W.) et Pearl (J.). – Qualitative decision theory. *In : Proc. 11th Nat. Conf. on Artificial Intelligence (AAAI'94)*, pp. 928–933. – Seattle, WA, 31 juillet-4 août 1994.

[van94] van der Torre L. – Violated obligations in a defeasible deontic logic. *In : 11th European Conf. on Artificial Intelligence*, pp. 371–375.

[van97] van der Torre L. – *Reasoning about Obligations*. – Thèse de PhD, Erasmus University Rotterdam, 27 Fév. 1997 1997.

[Vin89] Vincke (Ph.). – *L'Aide multicritère à la décision*. – Paris, Ellipses, 1989.

[vNM44] von Neumann (J.) et Morgenstern (O.). – *Theory of Games and Economic Behavior*. – Princeton University Press, 1944.

[Wal50] Wald (A.). – *Statistical Decision Functions*. – New York, Wiley and Sons, 1950.

[WD92] Watkins (C. J.) et Dayan (P.). – Q-learning. *Machine Learning*, vol. 3, n° 8, 1992, pp. 279–292.

[Web84] Weber (S.). – $\perp$-decomposable measures and integrals for archimedean t-conorms. *J. of Math. Annal. Appl.*, no101, 1984, pp. 114–138.

[Wha84] Whalen (T.). – Decision making under uncertaintywith various assumptions about available information. *IEEE Transactions on Systems, Man and Cybernetics*, no14, 1984, pp. 888–900.

[Wil95] Wilson (N.). – An order of magnitude calculus. *In : Proc. 10th Conf on Uncertainty in Artificial Intelligence UAI'95)*. pp. 548–555. – Morgan Kaufmann.

[WYBB91] Wong (S. K. M.), Yao (Y. Y.), Bollmann (P.) et Bürger (H. C.). – Axiomatization of qualitative belief structure. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 21, n° 4, July-August 1991, pp. 726–734.

[WYL93] Wong (S. K. M.), Yao (Y. Y.) et Lingras (P.). – Comparative beliefs and their measurements. *Int. Journal of General Systems*, no22, 1993, pp. 69–89.

[Yag79] Yager (R. R.). – Possibilistic decision making. *IEEE Trans. on Systems, Man and Cybernetics*, no9, 1979, pp. 388–392.

[Zad65] Zadeh (L. A.). – Fuzzy sets. *Int. Journal on Information and Control*, no8, 1965, pp. 338–353.

[Zad75] Zadeh (L. A.). – *Fuzzy Sets and their Applications to Cognitive and Decision Processes*, chap. Calculus of fuzzy restrictions, pp. 1–39. – Academic Press, New York, L. A. Zadeh and K. S Fu and K. Tanaka and M. Shimura, 1975.

[Zad78] Zadeh (L. A.). – Fuzzy sets as a basis for theory of possibility. *Fuzzy Sets and Systems*, no1, 1978, pp. 3–28.

[Zha94] Zhang (N. L.). – *A Computational Theory of Decision Networks*. – Rapport technique n° HKUST-CS94-3, Dept. of Computer Science, Hong Kong University of Science and Technology, 1994.