Un Modèle pour la Prise de Décision Multi-agent sous Incertitude Stricte

THÈSE

soutenue publiquement le 14 12 2009

pour l'obtention du

Doctorat de l'Université d'Artois

Spécialité Informatique

par

Ramzi BEN LARBI

Composition du jury

Rapporteurs :	Didier Dubois, Directeur de Recherche CNRS à l'IRIT Toulouse Abel-Illah Mouaddib, Professeur à l'Université de Caen-Basse Normandie
Examinateurs :	Bruno Beaufils, maître de conférences à l'Université des Sciences et Technologies de Lille Salem Benferhat, Professeur à l'Université d'Artois Sébastien Konieczny, Chargé de Recherche CNRS au CRIL Lens (co-directeur de thèse) René Mandiau, Professeur à l'Université de Valenciennes et du Hainaut-Cambrésis Pierre Marquis, Professeur à l'Université d'Artois (co-directeur de thèse) Nicolas Maudet, maître de conférences à l'Université Paris-Dauphine

Centre de Recherche en Informatique de Lens – CNRS UMR 8188
Université d'Artois, rue Jean Souvraz, S.P. 18 F-62307, Lens Cedex France
Secrétariat : Tél.: +33 (0)3 21 79 17 23 – Fax : +33 (0)3 21 79 17 70
http://www.cril.fr

Résumé

Le contexte informationnel dans lequel évolue un agent possède une importance extrême quand celui-ci élabore son comportement futur. Un agent rationnel doit en effet baser ses choix sur les informations qu'il possède pour choisir ses actions. Or, dans les applications réelles, l'information disponible à l'agent est souvent rare et peu précise. De multiples modèles ont été élaborés dans les différents cadres d'application de l'intelligence artificielle afin de caractériser une décision rationnelle dans chacun des contextes informationnels possibles. Les travaux présentés dans cette thèse concernent l'élaboration d'un modèle permettant à un agent de prendre des décisions rationnelles dans un contexte informationnel très pauvre. La seule information dont dispose un agent à propos du résultat de ses actions est la donnée de l'ensemble de résultats de chacune d'entre elles. En particulier, aucune information sur la conséquence la plus susceptible de se produire n'est disponible. L'agent est supposé égoïste (au sens où seul compte pour lui son propre intérêt) et autonome. Il évolue de plus dans un environnement où il coexiste avec d'autres agents (qui sont aussi égoïstes et autonomes). Les actions d'un agent influent sur les autres agents. La démarche entreprise pour élaborer le modèle est la suivante. D'abord, nous caractérisons les critères de décision rationnels d'un agent seul dans le contexte informatif étudié. Ensuite, nous étendons ces critères de décision individuelle au cas multi-agent en nous appuyant sur la théorie des jeux qui est le meilleur cadre pour exprimer les interactions entre agents rationnels et en particulier les possibilités de coordination entre les agents. Enfin, le domaine de la planification est un excellent cadre pour représenter et exprimer les concepts du modèle.

Abstract

The informative context in which an agent evolves is extremely important when she elaborates her future behaviour. A rational agent must base her choices on the available information. In realistic applications, the information is often rare and imprecise. Many models have been introduced to caracterize rational decision in each possible informative context. This thesis is about the elaboration of a model that allows an agent to make rational decisions in an extremely poor informative context. The only information that is available to an agent about her actions' consequences is the result set of each of her actions. No information about which consequence of any action will eventually happen is available. The agent is supposed to be selfish (which means that her own interest is her only concern) and autonomous. She evolves in an environment in which she coexists with other agents (that are as selfish and autonomous as her). An agent action may inflence those of other agents. We used the following approach to build our model. First, we caracterized the rational decision criteria for an agent to use in the context of complete ignorance. Then we extended these criteria, by using game theory concepts, to a multiagent environment. Finally, the planning framework is an excellent framework to represent the introduced concepts.

Table des matières

Introduction

Les êtres humains se trompent. Ils tombent, ratent la cible, perdent leur argent en bourse. Ils ont conscience que cela peut arriver. Cela ne les paralyse pourtant pas en général. Ils évaluent leurs chances de réussite, prennent une décision puis agissent. On dit alors communément, quand ils se trompent, qu'ils ont fait un mauvais calcul. La décision est en effet très souvent prise alors qu'ils ne possèdent pas toute l'information nécessaire pour déterminer complètement les conséquences de leurs choix. Ils se contentent de l'information qu'ils possèdent. Ils s'y adaptent et la considèrent (de manière consciente ou pas) comme un élément du problème qui se pose à eux.

Ce fait que l'on retrouve dans les gestes les plus anodins d'un être humain est particulièrement intéressant pour le domaine de l'intelligence artificielle (IA). Un des objectifs de l'IA est de reproduire sur machine des comportements considérés comme "intelligents". La faculté de décider est sans doute le trait le plus spécifique à ce que l'on entend par être intelligent. Un agent (humain ou artificiel) met en oeuvre un comportement "intelligent" lorsque, en fonction de ses objectifs (buts, préférences, intentions), il raisonne sur les informations dont il dispose (typiquement de natures variées : connaissances et croyances sur le "monde extérieur", les autres agents, perceptions diverses de ce monde extérieur, règles de décision, arguments, etc.) pour décider, agir au mieux en exploitant les actions dont il dispose. On voit apparaître dans cette description un certain nombre de concepts clés (croyances, préférences, actions, etc.) et de processus d'exploitation des informations correspondantes (raisonnement, décision). En particulier, les connaissances sont ici vues comme synonymes d'informations, munies de mécanismes d'exploitation de celles-ci et pas dans le sens plus technique de croyances certaines. C'est la même acception de "connaissances" qui est à l'oeuvre dans "représentation des connaissances".

La représentation des connaissances constitue le pilier de l'approche dite symbolique de l'intelligence artificielle, s'appuyant sur le postulat mécaniste de Simon et Newell : un système physique est capable d'un comportement "intelligent" si et seulement si c'est un système physique de symboles, i.e., une machine qui produit au cours du temps un assemblage évolutif de structures symboliques. Cette approche ne préjuge en rien des moyens à mettre en oeuvre pour aboutir à un comportement vu comme "intelligent". En fait, dans "représentation des connaissances", le terme "représentation" doit être pris dans un sens assez large. En effet, il s'agit tout autant de définir des modèles pour les processus d'exploitation des concepts d'intérêt (et évidemment ces concepts eux-mêmes) que des langages permettant la représentation proprement dite de ces concepts.

Cette séparation, pas toujours claire entre modélisation et représentation, peut s'expliquer par la prépondérance des approches logiques en représentation des connaissances. En effet, dans toute logique, le point de départ est un langage, donc un formalisme de représentation ; le modèle considéré pour les concepts représentés existe mais est souvent moins apparent. Par exemple, en logique propositionnelle épistémique mono-agent, on s'intéresse à la représentation des connaissances propositionnelles d'un agent. La logique prépondérante pour cela est S5. Dans cette logique, on traite en fait des connaissances propositionnelles d'un agent idéalement rationnel (i.e. ayant des capacités de déduction non limitées), qui ne connaît que des propositions vraies (et a donc des connaissances cohérentes) et a des capacités parfaites d'introspection positive et négative (quand il sait, il sait ce qu'il sait et quand il ne sait pas, il sait aussi qu'il ne sait pas). Que les connaissances dans S5 doivent vérifier ces principes apparaît dans son axiomatique (et dans sa sémantique) mais évidemment pas directement dans le langage correspondant. On peut donc parfaitement utiliser le langage de S5 (i.e. un langage propositionnel classique enrichi par une modalité de connaissance) sans se conformer nécessairement aux principes de la logique pour ce qui est de l'utilisation des "connaissances".

En règle générale, il n'existe pas de modèle unique et consensuel pour les différents processus d'exploitation des connaissances ainsi que pour les concepts sous-jacents. De la même façon, une fois un modèle de concept fixé, il n'existe pas forcément un langage de représentation idéal qui lui correspondrait. L'absence d'un modèle absolu pour la prise de décision est bien connue. On sait bien en effet qu'il n'y a pas forcément un concept de "meilleure décision" lorsque l'état du "monde" dans lequel la décision est prise et/ou les conséquences des décisions disponibles ne sont pas complètement connus (décision dans l'incertain) ou encore lorsque plusieurs agents évaluent séparément les conséquences des décisions (décision de groupe) ou agissent de concert ou pas (là, on entre dans le domaine de la théorie des jeux). Dans un cadre extrêmement simplifié comportant des actions déterministes (chacune conduit de manière sûre à une unique conséquence) et un état du monde parfaitement connu à chaque instant, on pourrait valider le choix de l'agent en comparant les conséquences des actions à ses buts. Ce n'est plus aussi direct dans les contextes informationnels précédents. Le modèle choisi est alors primordial. Celui-ci doit prendre en compte le mieux possible la quantité ainsi que la qualité de l'information disponible. Faire varier ces deux propriétés donne une multitude de modèles possibles.

Pour ce qui est des concepts, la pluralité des modèles est aussi la règle. Considérons par exemple la notion de préférences. Nombreux sont les modèles possibles (fonctions d'utilité sur un ensemble de conséquences, préordre total ou non, etc.). Pour un même concept (par exemple préordre), diverses représentations sont usuellement possibles (CP-net, bases de formules dans diverses logiques pondérées, etc.). Il en va de même pour les actions (ontiques, que l'on peut modéliser simplement par exemple comme des applications entre états du "monde") ; ainsi, une action peut être représentée de façon indépendante des états (en utilisant différents formalismes comme le calcul des situations, la logique dynamique, les formalismes à la STRIPS, etc.) ou pas (i.e. comme un graphe étiqueté, type abstrait de données qui peut à son tour être réalisé par des structures de données différentes, matrice booléenne, liste d'adjacence, etc).

Se posent donc deux questions clés : qu'est-ce qu'un bon modèle (pour un concept donné, pour un processus d'exploitation d'informations donné) ? Qu'est-ce qu'un bon langage de représentation de concept (une fois le modèle fixé) ? Répondre à ces questions nécessite d'analyser modèles et langages de représentation selon différents critères et de faire un choix sur la base de ceux-ci.

Pour les modèles, les critères de rationalité et d'adéquation à l'information disponible semblent primordiaux. La rationalité peut s'envisager selon différents points de vue, en particulier descriptif (il s'agit de s'inspirer des comportements humains observés) ou normatif (ce qu'un humain devrait faire ou concevoir). L'expression de la rationalité passe souvent par l'explicitation de postulats à respecter. Par exemple, si l'on estime qu'une préférence doit être transitive (ce qui n'est pas toujours vrai chez l'humain), il faut adopter un modèle mathématique qui l'impose. Pour prendre un exemple de processus d'exploitation d'informations, si l'on observe que les propositions qui ne sont impliquées dans aucune contradiction sont inférées, il faut choisir un modèle de raisonnement para-consistant qui l'assure. Concilier les deux formes de rationalité n'est pas forcément facile (voir le paradoxe d'Allais par exemple qui remet en cause le modèle de l'utilité espérée en décision dans l'incertain).

L'adéquation à l'information disponible exprime simplement le fait qu'il faut considérer et utiliser des modèles qui vont permettre d'exploiter autant que possible toute l'information disponible et surtout ne vont pas nécessiter d'informations autres que celles que l'agent possède pour pouvoir être mis en oeuvre. Par exemple, le modèle (dominant) de l'utilité espérée en décision dans l'incertain offre un certain nombre de garanties (en particulier en terme de rationalité) mais il est totalement inadapté à la modélisation de scénarios de prise de décision où l'information disponible est pauvre (par exemple, un ensemble d'actions ayant plusieurs conséquences possibles et un ordre sur celles-ci exprimant la préférence de l'agent).

Les recherches actuelles en représentation des connaissances visent à définir de nouveaux modèles pour lesquels les modèles existants sont inadaptés (que ce soit pour des raisons de rationalité ou d'adéquation à l'information disponible) et de nouveaux langages offrant de meilleurs compromis que ceux qui existent déjà compte tenu des différents critères mentionnés. Si on examine les différents modèles existants du point de vue des contextes informatifs auxquels ils font référence, on observe qu'entre le cas de certitude dans lequel l'agent connaît parfaitement le résultat de ses actions et le cas où il n'en connaît rien existent de nombreuses situations intermédiaires où l'agent possède une certaine information quoiqu'incomplète. Dans toutes ces situations, on suppose que l'agent connaît au moins, pour chacune de ses actions, l'ensemble de ses résultats possibles. C'est essentiellement la quantité et la qualité de l'information qui différencie les différents contextes. Ainsi, du plus précis au moins précis, nous avons donc à propos des résultats une distribution de probabilités, une distribution imprécise, plusieurs distributions de probabilités, une information qualitative telle une distribution de possibilités ou une simple partition des états en états plus ou moins crédibles. A chaque contexte informatif correspond un modèle qui permet d'en exploiter les concepts. Il s'agit essentiellement dans chaque cas de se représenter de la manière la plus rationnelle possible (la plus cohérente avec les buts de l'agent) les attentes de celui-ci associées aux différentes alternatives qui s'offrent à lui. Le modèle le plus prisé dans des applications économiques notamment est celui de l'utilité espérée, sans doute parce qu'il se rapproche du cas de certitude au travers de son concept de l'équivalent certain, i.e., associer à un ensemble de résultats possibles pondérés par une distribution de probabilités une utilité numérique qui permet de les comparer à une issue certaine. Pourtant ce modèle nécessite une grande quantité d'information ainsi que du temps et des capacités de calcul et de stockage étendues. L'interprétation qui sous-tend le plus souvent son utilisation est l'interprétation fréquentiste qui requiert de connaître le résultat d'un grand nombre d'expériences. Les travaux d'axiomatisation de l'utilité espérée de von Neumann et Morgenstern ont permis d'en adapter l'esprit à des contextes informatifs plus pauvres. Ainsi, dans le cas où l'agent ne possède pas de distribution de probabilités et en se basant sur une interprétation subjective des probabilités, Savage a montré que l'agent se comporte encore comme un maximisateur d'une certaine utilité espérée même s'il n'en a pas forcément conscience. La même direction permit de retrouver l'esprit de l'utilité espérée jusque dans le cas où l'information n'est plus que qualitative comme le montrent les travaux de Dubois. Toutefois, en appauvrissant encore le contexte informationnel, cette approche atteint ses limites. Il existe en effet des contextes dans lequels une information aussi structurée qu'une distribution de probabilités ou de possibilités, ou même une partition sur les états du monde selon qu'ils sont plus ou moins plausibles, n'est pas disponible pour l'agent. On peut se poser la question de l'utilité d'étudier ce genre de contextes extrêmement pauvres dans lesquels il semble difficile d'imaginer un critère rationnel de décision. En fait, la même motivation qui a poussé à introduire le modèle de l'utilité espérée est encore présente. Il s'agit de sortir d'un cadre idéalisé pour s'adapter aux conditions d'une décision réelle. Or, ces modèles exigent encore des structures d'information particulières. La question est alors de savoir que faire dans le cas d'une information minimale. L'approche axiomatique utilisée est particulièrement adaptée dans des contextes informationnels pauvres. Il s'agit de déterminer les conditions adaptées à l'information disponible que doit satisfaire un critère de décision rationnel. Plusieurs travaux, notamment ceux d'Arrow et Hurwicz, ont montré qu'on peut dépasser le critère prudent intuitif qui consiste à comparer les actions selon la pire conséquence possible de chacune vers des critères qui utilisent de l'information disponible et laissée de côté par ce critère. Les axiomes de rationalité retenus induisent le plus souvent l'existence de points focaux dans l'ensemble de résultat possibles d'un choix. Comparer les choix disponibles revient alors à comparer ces (ensembles de) points focaux.

Dans le paysage précédemment dépeint, cette thèse est axée sur les cadres les plus pauvres en information à la fois en théorie de la décision et en théorie des jeux.

Nous nous intéressons d'abord au cadre de l'ignorance en théorie de la décision. C'est un cadre très pauvre en information dans lequel aucune information autre que l'ensemble des conséquences possibles d'un choix n'est disponible au moment de la prise de décision. L'approche que nous adoptons, pour définir les critères de décision adéquats, est une approche axiomatique. C'est l'approche utilisée en général en théorie de la décision. Le point essentiel une fois que l'on a défini le cadre informationnel dans lequel évolue l'agent est de garantir que son comportement sera rationnel. La démarche axiomatique remplit ce rôle. Il s'agit de déterminer une famille d'axiomes qui traduisent ce que l'on entend intuitivement par un comportement rationnel. Les critères de décision à utiliser par l'agent sont alors ceux qui satisfont cette famille d'axiomes. Si l'on adhère aux axiomes, on doit adhérer au caractère rationnel des critères. Dans le cadre particulier de l'ignorance, nous introduisons des axiomes simples et de justification aisée qui caractérisent des critères opérationnels en ce sens qu'ils dépassent la tentation de se comporter de manière de plus en plus prudente à mesure que l'agent manque d'information.

Certains axiomes que nous introduisons partent d'une situation type dans laquelle l'agent peut intuitivement formuler une comparaison entre deux actions. On peut en effet avancer que si l'agent préfère toutes les conséquences d'une action a à toutes celles d'une action b alors il peut sans doute comparer l'action a à l'action b. L'axiome stipule alors que l'ajout d'une conséquence commune aux deux actions ne doit pas inverser cette préférence. L'information de l'ajout de la conséquence commune n'est en effet pas accompagnée d'une information sur la plausibilité de réalisation de cette dernière.

Nous montrons qu'alors les critères caractérisés par nos axiomes partagent le fait de comparer les ensembles de résultats en tenant surtout compte des leurs éléments extrémaux par rapport à la préférence d'un agent. Ces critères ne tiennent en particulier compte que de la comparaison entre les conséquences et pas de leur valeur intrinsèque. En particulier, ils sont opérationnels dans le cas où l'agent ne peut formuler que des préférences ordinales sur les conséquences de ses actions. C'est une hypothèse cohérente avec un contexte informationnel pauvre dans lequel l'agent ne possède pas d'information précise sur son environnement, soit parce qu'elle ne lui pas accessible, soit parce qu'il ne peut pas, pour des raisons de capacité de stockage limitées ou de contraintes de rapidité de la prise de décision.

Nous nous intéressons ensuite au cadre de la planification multi-agent pour y illustrer l'utilisation des concepts de la théorie de la décision en considérant donc la planification comme un choix entre plans disponibles. C'est une hypothèse cohérente avec le cadre de la planification classique dans lequel l'agent planificateur ne peut observer le résultat de l'exécution de son plan. Elle est de plus particulièrement adaptée au cadre informationnel pauvre que nous étudions. Plus précisément, nous considérons le cas d'agents autonomes (ayant possiblement des buts différents et même antagonistes) qui essayent de choisir des plans leurs permettant d'atteindre leurs buts respectifs. Nous donnons un modèle permettant de formaliser puis de traiter le cas où les agents considérés possèdent des préférences dichotomiques (l'agent est soit complètement satisfait soit complètement insatisfait). En ce sens, nous étendons les concepts éprouvés de la planification classique au cadre multi-agent.

A ce modèle, on peut ajouter la possibilité que les agents coopèrent autour d'un plan "collectif" s'ils y voient chacun leur intérêt. C'est une possibilité qui peut changer grandement la manière dont l'agent classe les possibilités qui s'offrent à lui. Typiquement, des plans qui, si l'agent était seul, ne lui permettraient pas d'atteindre son but, peuvent se révéler de "bons" plans une fois mélangés aux plans des autres agents. Nous proposons alors une procédure qui permet aux agents concernés de prendre en compte cette possibilité dans l'évaluation de leurs plans disponibles. Rappelons que le choix entre plans se fait en amont de l'exécution des plans. Les agents doivent donc identifier les possibilités de coordination offertes par chacune des exécutions simultanées d'un de leur plans avec les plans des autres agents en se basant notamment sur la rationalité des autres agents présents qui induit que ceux-là accepteront de se coordonner dès lors qu'ils y voient leur intérêt (i.e., la possibilité d'atteindre leurs buts).

Cette possibilité est surtout pertinente quand les préférences des agents sont graduées (i.e., non pas seulement dichotomiques). C'est une supposition minimale qui rend notamment les résultats obtenus applicables dans le cas où les préférences des agents sont exprimées au moyen d'échelles numériques. Dans le cas de préférences graduées, les possibilités de coordination sont plus nombreuses et appellent une évaluation plus fine des plans disponibles. Nous étendons la procédure précédemment introduite à ce cas. Nous montrons en particulier que la comparaison entre les différentes possibilités de coordination peut s'appuyer dans le contexte informationnel considéré sur les mêmes critères qui servent à comparer les ensembles de conséquences possibles des plans.

Enfin, la dernière partie de la thèse aborde le domaine de la théorie des jeux. On y définit principalement un nouveau type de jeu, les jeux qualitatifs à résultats multiples, dont l'introduction a un double intérêt au regard des parties précédentes.

D'abord, cela a un intérêt propre de combler le manque constaté dans la théorie des jeux en ce qui concerne le cadre le plus pauvre en information qu'est le cadre où la seule information disponible est l'ensemble des conséquences des actions. Les modèles présents dans la littérature de la théorie des jeux sous incertitude supposent toujours que nous pouvons isoler plusieurs situations de jeu délimitées par ce que l'on définit comme types des joueurs dans les jeux bayésiens. On obtient alors une situation de jeux "classique" pour chaque combinaison de types des joueurs. Cela reste vrai y compris dans le cas de jeux étudiant un contexte informationnel pauvre dans lequel une distribution de probabilités sur les types des autres joueurs (disponible dans les jeux bayésiens) n'est plus accessible (on obtient alors ce que l'on nomme un jeu pré-bayésien). Nous montrons que ce modèle ne couvre pas le cas que nous étudions et que les jeux qualitatifs à résultats multiples y sont, eux, adaptés.

Ensuite, ce type de jeux permet de généraliser le premier modèle que nous proposons pour la planification multi-agent. Il permet aussi de justifier le modèle en l'appuyant sur l'approche axiomatique élaborée dans le cadre d'un seul agent. Les résultats obtenus sont donc à la fois soutenus par le fait que l'évaluation d'un agent pour ses plans s'appuie sur des critères de décision dont nous avons montrés qu'ils sont les plus rationnels dans le cadre informationnel étudié et que la solution du jeu possède la légitimité des concepts de solutions largement étudiés et utilisés dans la théorie des jeux que sont par exemple les équilibres de Nash.

Première partie

Etat de l'art

Chapitre 1

Préliminaires

On ne peut aborder le sujet de la décision sans utiliser le concept de préférences d'un agent. On a donc besoin, pour définir les concepts et les formaliser, d'un modèle de ces préférences.

L'outil formel principal pour exprimer les préférences d'un agent est celui de *relation binaire*. Une relation binaire $\Re$ sur un ensemble S est complètement définie par la donnée d'un ensemble de couples d'éléments de S. L'appartenance d'un couple (a, b) d'éléments de S à $\Re$ signifie alors " a est en relation $\Re$ avec b". Cette manière de présenter les relations permet d'utiliser les outils de la théorie des ensembles sur les relations. Par exemple $(a, b) \in \Re_1 \cup \Re_2$ signifiera "a est en relation $\Re_1$ avec b ou a est en relation $\Re_2$ avec b". On utilisera toutefois la notation ensembliste $(a, b) \in \Re$ que quand elle sera utile pour privilégier la notation $a\Re b$ plus légère.

- Une relation $\Re$ est dite *réflexive* si pour tout élément a de l'ensemble S on a $a\Re a$.
- Une relation $\Re$ est dite *antisymétrique* si pour tous éléments a et b de l'ensemble S, $a\Re b$ et $b\Re a$ implique $b = a$ (par exemple la relation $\geq$ sur les réels est antisymétrique).
- Une relation $\Re$ est dite *symétrique* si pour tous éléments a et b de l'ensemble S, $a\Re b$ implique $b\Re a$ (par exemple la relation = sur les réels est symétrique).
- Une relation $\Re$ est dite *transitive* si pour tous éléments a, b et c de l'ensemble S, $a\Re b$ et $b\Re c$ implique $a\Re c$ (par exemple la relation > sur les réels est transitive).

La transitivité induit une sorte de cohérence dans les préférences. Avec une relation $\Re$ non transitive, on peut avoir $a\Re b$ et $b\Re c$ mais a et c incomparables.

- Une relation $\Re$ est dite *totale* si pour tous éléments a et b de l'ensemble S, soit $a\Re b$ soit $b\Re a$ (soit les deux à la fois) (par exemple la relation > sur les réels n'est pas totale).

Le fait que les préférences du décideur soient modélisées par une relation totale signifie que, s'il est questionné sur un couple d'alternatives, il saura toujours en préférer l'une à l'autre (ou être indifférent entre les deux). Pourtant, pour choisir entre deux alternatives A et B, est-il vraiment utile que l'agent dise (ou sache) s'il préfère A à C ou B à C ? De fait les agents humains ne font pas appel lors de chaque situation de choix à toutes les préférences qu'ils ont à propos des éléments du monde qui les entoure. Par ailleurs, l'exigence de totalité peut conduire à faire des comparaisons saugrenues. Que signifie pour un agent quand il a à choisir une voiture de préférer une banane à un melon ?

- La structure de la préférence la moins exigeante communément utilisée est celle de *préordre*, parfois aussi appelé *quasi-ordre*, qui est une relation réflexive et transitive. Quand un préordre est total, c'est un ordre faible (ou *weak order*).

  Si on demande qu'un préordre soit de plus antisymétrique, alors on obtient une relation d'*ordre*.
- La partie stricte d'un ordre $\leq$ est généralement notée $<$ et définie par $x < y$ si et seulement si $(x \leq y)$ et $\neg(y \leq x)$. C'est une relation irréflexive (un élément n'est jamais en relation avec lui même) et transitive. Remarquons que dans ce cas la relation est aussi antisymétrique.

En effet, ces deux conditions impliquent l'antisymétrie. En fait elles impliquent une propriété d'asymétrie :

$$\text{Si } x\Re y \text{ alors } \neg(y\Re x)$$

En termes de préférences de l'agent et dans l'optique d'une prise de décision, le préordre est une structure de préférence assez grossière. Une relation d'ordre est de ce point de vue plus forte car elle impose une structure "plus fine".

- Etant donnée une relation $\Re$, sa *relation stricte* associée $\wp$ est définie comme le sous-ensemble des couples $(a, b) \in \Re$ tels que $a\Re b$ et $\neg(a\Re b)$.
- Etant donnée une relation $\Re$, sa *relation d'indifférence* associée $\Im$ est définie comme le sous-ensemble des couples $(a, b) \in \Re$ tels que $a\Re b$ et $a\Re b$.

On peut définir de la même manière deux relations associées à un préordre. Soit $\leq$ un préordre sur un ensemble χ, i.e., $\leq$ est une relation réflexive et transitive. Le préordre $<$ et la relation d'équivalence $\sim$ associés sont respectivement définis par :

- x $\sim$ y ssi x $\leq$y et y $\leq$ x.
- x $<$y ssi x $\leq$y et y $\nleq$ x.

Etant données n relations $\leq_i, i \in [1, n]$ sur χ, la *relation lexicographique* lex($\leq_1, \cdots, \leq_n$) est définie par

- x $\leq_{lex(\leq_1,\cdots,\leq_n)}$ y ssi x $<_{lex(\leq_1,\cdots,\leq_n)}$ y ou x $\sim_{lex(\leq_1,\cdots,\leq_n)}$ y avec :
  - x $<_{lex(\leq_1,\cdots,\leq_n)}$ y ssi $\exists j \in [1, n]$ s.t. $\forall i < j$ x $\sim_i$ y et x $<_j$y.
  - x $\sim_{lex(\leq_1,\cdots,\leq_n)}$ y ssi $\forall i \in [1, n]$ x $\sim_i$ y.

Nous notons $\subseteq$ la relation d'inclusion entre ensembles et $\subset$ sa partie stricte. Pour un ensemble donné χ, nous notons 2^χ l'ensemble des parties de χ.

Soient $x = \langle x_1, \cdots, x_n\rangle$ et $y = \langle y_1, \cdots, y_n\rangle$ deux vecteurs dans χ^n. y *Pareto domine* x ssi $\forall i \in [1, n], x_i \leq y_i$ et $\exists i \in [1, n], x_i < y_i$. Le *front de Pareto* d'un ensemble $E \subseteq \chi^n$ est le sous-ensemble de E formé par ses points non Pareto-dominés, noté $\mathsf{pf}(E)$, i.e., $\mathsf{pf}(E) = \{x \in E \mid \nexists y \in E$ y Pareto domine $x\}$.

Nous notons $\min(\chi, \leq)$ l'ensemble des minima de χ et $\max(\chi, \leq)$ l'ensemble des maxima de χ par rapport à la relation $\leq$, i.e., $x \in \min(\chi, \leq)$ ssi $\nexists y \in \chi$ t.q. $y < x$, et $x \in \max(\chi, \leq)$ ssi $\nexists y \in \chi$ t.q. $x < y$.

Soient $x = \langle x_1, \cdots, x_n\rangle$ et $y = \langle y_1, \cdots, y_n\rangle$ deux vecteurs dans χ^n, et $\leq$ une relation sur χ, alors $x \leq^i y$ est une notation pour $x_i \leq y_i$, et x_{-i} est une notation pour $\langle x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n\rangle$.

On considèrera quelquefois l'ordre suivant sur $I\!R^2$ (l'ordre produit de la relation habituelle sur les réels dont on peut facilement vérifier que c'est un ordre partiel) : $\forall(x_1, y_1) \succeq (x_2, y_2)$ ssi $x_1 \geqslant x_2$ et $y_1 \geqslant y_2$.

Soient A et B deux ensembles. L'ensemble des applications de A dans B est noté B^A.

Soit A un ensemble. Une relation $\leq$ sur 2^A est préadditive si pour chaque sous-ensemble E,F et G de A on a

$$E \cap (F \cup G) = \emptyset \Rightarrow (F \leq G \Leftrightarrow (F \cup E \leq G \cup E))$$

Chapitre 2

Théorie de la décision sous incertitude

2.1 Théorie de la décision

La théorie de la décision est un domaine qui offre une approche mathématique rigoureuse de la prise de décision (d'une "bonne décision") dans les situations d'incertitude. La définition du dictionnaire de philosophie précise ce que l'on entend par une situation d'incertitude. On y trouve à l'entrée *théorie de la décision*

> "la théorie des choix faits quand chaque option est associée à un risque ou une attente de gain ou de perte."

Ces termes familiers restant à définir de manière rigoureuse, on comprend intuitivement qu'ils évoquent une situation dans laquelle le décideur ne maîtrise pas totalement toutes les contingences ayant un effet sur l'issue de ses actions. Par cette définition même transparaît le caractère pluri-disciplinaire de cette théorie. On peut l'approcher différemment selon les outils avec lesquels on modélise l'agent (philosophie, psychologie, sociologie, mathématiques) ou le domaine d'application où l'on se place (économie, théorie du vote, intelligence artificielle).

La théorie traite la notion de bonne décision de deux manières différentes (et complémentaires). Une grande partie de littérature existante adopte un point du vue normatif où le but est de caractériser ce que doit être une bonne décision. On se base alors sur une idéalisation du cadre. Le décideur est supposé avoir une information parfaite et une capacité infinie de calcul. Il est surtout supposé être rationnel. Et c'est essentiellement le sens que l'on va donner à cette notion de rationalité qui définira les "bonnes" décisions à prendre. Cette approche trouve son débouché naturel dans le domaine de l'aide à la décision. Les logiciels et méthodes développés viennent en renfort au client (non idéal) pour lui préscrire les décisions à prendre. Un deuxième point de vue, celui-là descriptif, est nécessaire et ce pour deux raisons. D'abord cette approche sert à alimenter la première ambition en définissant, parmi les comportements observés ceux que l'on estime être rationnels. Ce jugement est souvent établi sur des situations particulières. L'approche normative prend alors le relais pour imaginer ce qu'implique cette rationalité dans les différentes autres conditions imaginables, la cohérence (le fait que les choix futurs de l'agent ne viennent pas contredire ceux déjà effectués) étant le minimum de rationalité que l'on peut exiger. D'un autre côté, si le décideur n'est pas seul dans son environnement, ce qui n'est pas rare, une décision prescrite par la théorie doit prendre en compte l'évolution du système, donc la prédire au mieux. Or le système est formé par des agents non idéaux. D'où l'utilité de pouvoir décrire ce que les décideurs font dans la réalité.

Le concept central de cette théorie est celui d'un problème de décision. Il capture une situation où un agent se trouve face à des choix parmi des actions qui peuvent causer dans le monde qui l'entoure des conséquences qu'il désire ou redoute. Les actions sont introduites comme les changements que l'agent peut apporter aux aspects du monde qu'il peut directement contrôler. Les conséquences représentent tout ce qui a de la valeur (positive ou non) pour l'agent après l'exécution d'une action. L'incertitude signifie que l'agent ne peut pas tout contrôler par ses actions, c'est-à-dire que quand une action est exécutée, plusieurs conséquences sont envisageables. L'agent ne sait pas laquelle va survenir.

Traditionnellement, cette incertitude porte sur une notion d'état du monde. Un état du monde est ce que l'agent saurait de son environnement et de sa position dans celui-ci s'il en avait une information complète et parfaite. L'incertitude peut venir de la méconnaissance des circonstances dans lesquelles une action s'exécute (qui peuvent comporter des éléments totalement incontrôlables tels que le temps qu'il fera ou d'autres plus prévisibles tels que ce que d'autres agents feront). Si l'agent pouvait connaître parfaitement l'état du monde, il pourrait savoir quelle conséquence aurait chaque action. Les états représentent donc tout ce que l'agent ne peut contrôler. Chaque couple formé par une action et un état (a, s) fixe alors une conséquence qui est le résultat de l'exécution de l'action a dans l'état s. Cette relation suggère, dans le cas où les actions et les états sont finis, d'utiliser la représentation matricielle suivante (table 2.1), appelée *table de décision* pour représenter les liens entre états, actions et conséquences. Le résultat de l'exécution de l'action a_i dans l'état s_j y est noté $a_i(s_j)$. Cette représentation rappelle celle du résultat d'une application et de fait, une action est traditionnellement présentée comme une application associant chaque état à un autre (résultat de son exécution dans cet état) ou à une conséquence.

	s_1	$s_2 \cdots$	s_n
a_1	$a_1(s_1)$	$a_1(s_2) \cdots$	$a_1(s_n)$
a_2	$a_2(s_1)$	$a_2(s_2) \cdots$	$a_2(s_n)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_m	$a_m(s_1)$	$a_m(s_2) \cdots$	$a_m(s_n)$

TABLE 2.1 – TABLE DE DÉCISION

Voici un exemple classique ([Savage 1954]) qui permet d'illustrer les notions introduites :

Mazen se prépare une omelette pour diner. Il prend dans le frigidaire une boîte d'oeufs que sa mère a achetée. Il y reste 3 oeufs. Il en casse deux dans un bol. Le problème suivant se présente alors à lui. Il doute de la fraîcheur du troisième oeuf. L'ensemble des états de la nature est constitué de deux états : oeuf frais et oeuf avarié.

Dans cet exemple, et selon la représentation précédente, les actions dont dispose Mazen sont *Casser dans le bol*, *Casser dans un autre bol* et *Jeter*, les états du monde sont *oeuf frais* et *oeuf avarié*. La donnée d'un couple (action,état) détermine un unique résultat représenté dans la table 2.2.

	oeuf frais	oeuf avarié
Casser dans le bol	omelette à 3 oeufs	pas d'omelette 2 oeufs gachés
Casser dans un autre bol	omelette à 3 oeufs un bol de plus à laver	omelette à 2 oeufs
Jeter	omelette à 2 oeufs un oeuf gaché	omelette à 2 oeufs

TABLE 2.2 – OMELETTE

La décision est le fait d'asseoir un choix sur une base logique et rationnelle [Chernoff 1954]. Cette base est ce qui définit le "critère de décision". Un critère de décision représente les préférences de l'agent, son expression la plus simple est la suivante "choisir l'action dont on préfère les conséquences". Rappelons ici que les préférences sur les conséquences sont une donnée du problème, au moins sur le plan où nous nous plaçons. Dans l'exemple précédent, nous supposons que l'agent sait parfaitement qu'il préfère une omelette contenant 3 oeufs à une omelette contenant 2 oeufs et laver un bol à en laver deux. Les préférences à déterminer par l'application du critère sont celles que l'agent a (ou devrait avoir) sur les actions. Afin d'utiliser la matrice précédente pour analyser une décision, on a besoin à la fois de l'évaluation des résultats qui y figurent et d'une information sur l'état du monde qui est le plus susceptible de se produire.

La manière la plus répandue de présenter la valeur des résultats est de leur assigner une utilité. Ainsi, dans l'exemple précédent, les résultats décrits peuvent être représentés numériquement comme à la table 2.3.

	oeuf frais	oeuf avarié
Casser dans le bol	4	0
Casser dans un autre bol	2	3
Jeter	1	3

TABLE 2.3 – OMELETTE NUMÉRIQUE

Il y est par exemple spécifié que Mazen préfère avoir une omelette contenant 2 oeufs à pas d'omelette du tout. Qu'exprime une telle représentation ? Jusqu'à présent tout ce que l'on a exigé est qu'elle soit compatible avec la relation de préférence, c'est-à-dire que si l'on préfère un résultat a à un résultat b alors la représentation numérique devra être telle que le nombre associé à a soit plus grand que celui associé à b.

Définition 2.1 *Soit U une fonction de X, un ensemble d'alternatives (ou conséquences), dans $\mathbb{R}$. On dit que U représente la préférence $\succeq$ si et seulement si quels que soient $x, x' \in X$ on a*

$$U(x) \geq U(x') \Leftrightarrow x \succeq x'$$

La fonction U est dite alors fonction d'utilité, et la valeur $U(x)$ est l'utilité du résultat x.

Il se peut d'après cette définition que la fonction d'utilité reflète ou non une notion d'intensité de préférence. Dans le deuxième cas, l'échelle de représentation est ordinale (elle exprime simplement comment sont rangées les conséquences). On verra dans la suite que les deux cas permettent une prise de décision mais qu'ils requièrent chacun un type différent de critères de décision. Une grande partie des travaux de la théorie de la décision est quasi-exclusivement consacrée à des problèmes qui peuvent être exprimés au travers de matrices numériques : des "matrices d'utilité". Dans beaucoup de problèmes réels, nous avons pourtant des informations qui sont loin d'être aussi précises que des valeurs numériques, peut être au mieux exprimées par une relation de préférence incomplète.

Une fois acquise la représentation des préférences de l'agent sur les conséquences, la caractérisation à utiliser dépend de la deuxième information disponible : celle que l'on a sur les mondes les plus susceptibles de se produire, et de sa nature quantitative ou non. Le premier travail marquant fût celui d'Abraham Wald's dans [Wald 1950] où il généralisait la littérature existante à son époque vers un cadre de décision tel que nous le connaissons aujourd'hui. Cet article préfigurait en grande partie ce qui allait devenir l'univers moderne de la théorie de la décision. Le terme "théorie de la décision" en lui même fût probablement utilisé pour la première fois dans les années 50 notamment par E. L. Lehmann dans [Lehmann 1959].

La procédure connue comme celle de la valeur espérée fût initiée au 17ème siècle et invoquée notamment par Blaise Pascal dans ses Pensées, publiées en 1670. L'idée derrière la valeur espérée est que, quand on a à sa disposition un ensemble d'actions, chacune pouvant induire plus d'une conséquence possible, la procédure rationnelle est la suivante : d'abord identifier toutes les conséquences possibles, déterminer leurs valeurs et leurs probabilités respectives avant de multiplier les deux pour obtenir la valeur espérée. L'action choisie est alors celle ayant la valeur espérée globale la plus élevée. En 1738, Daniel Bernoulli a fait évoluer ce concept dans son article "Exposition of a New Theory on the Measurement of Risk". Il commença par en montrer les limites au moyen du paradoxe de St. Petersburg.

> Soit le jeu suivant : on lance en l'air une pièce de monnaie. Si face apparaît, la banque paie 2 euros au joueur, et on arrête le jeu. Sinon, on relance la pièce. Si face apparaît, la banque paie 4 euros, et on arrête le jeu. Sinon, on relance la pièce. Si face apparaît, la banque paie 8 euros au joueur, et ainsi de suite. Donc, si face apparaît pour la première fois au n-ième lancer, la banque paie 2^n euros au joueur. Quelle est la mise initiale pour que le jeu soit équitable, c'est-à-dire pour que ni la banque ni le joueur ne soient avantagés par ce jeu ?

Ses critiques se portent sur le pouvoir normatifs de la théorie. En effet dans le cas décrit particulier de ce paradoxe, le comportement des agents confrontés à cette situation n'est pas celui prescrit par le modèle. Et c'est dans la solution qu'il propose à cette faiblesse qu'il définit une fonction d'utilité et remplace dans la procédure le calcul de la valeur financière espérée par celui de l'utilité espérée.

L'avènement de la probabilité subjective , i.e., une probabilité reflétant essentiellement la croyance de l'agent dans le fait qu'un évènement se produira plutôt qu'un autre, dans les travaux de Frank Ramsey [Ramsey 1931], Bruno de Finetti [de Finetti 1937], Leonard Savage [Savage 1954] entre autres, permit d'étendre le domaine d'application de l'utilité espérée à des situations dans lesquelles le décideur ne dispose que de probabilités subjectives. Il était alors largement admis parmi les économistes que les personnes physiques se comportent comme des décideurs rationnels et que par conséquent la théorie de l'utilité espérée permet de prévoir les comportements humains. Les travaux de Maurice Allais [Allais 1953] et Daniel Ellsberg [Ellsberg 1961] ont largement jeté le doute sur cette vision. Daniel Kahneman et Amos Tversky [Kahneman & Tversky 1979] adoptèrent une démarche empirique. Ils montrèrent dans une vision descriptive que, dans un comportement réel, les pertes comptent plus que les gains, les gens sont plus attentifs aux changements qui peuvent se produire dans leurs utilités qu'aux changements dans leurs environnement, et leurs utilités subjectives s'en trouvent biaisées.

Enfin, parallèlement à ces approches quantitatives, des chercheurs comme Shackle [Shackle 1955] et Hurwicz [Hurwicz 1951] ou encore Maskin [Maskin 1979] se sont intéressés aux types de critères à adopter dans le cas où l'information disponible est qualitative. Ce genre d'approche semble ituitivement utile dans le cadre d'applications réelles dans lesquelles les informations sont rares, imprécises ou difficiles à mesurer.

2.1.1 Attentes

Le premier pilier de la décision est sans doute l'attente que l'on a par rapport à la décision que l'on va prendre. Notre préférence sur les actions disponibles dépend de ce que les actions vont nous rapporter si on les exécute. Dans le cas déterministe, c'est-à-dire si chaque action mène de manière certaine à un unique résultat après son exécution, notre attente pour cette action est la valeur que l'on accorde à son résultat. Une grande partie des problèmes de décision ne sont pas déterministes. L'exécution d'une action peut mener à un ensemble de résultats possibles. Il semble évident que cette nouvelle caractéristique ne devrait pas changer notre manière de décider. Nous avons encore ici les mêmes éléments de base constitutifs d'un problème de décision : on est face à des alternatives dont chacune a des conséquences sur notre satisfaction. Cette vision est particulièrement claire dans la définition que donnent [Luce & Raiffa 1957] de la décision sous incertitude :"étant donné un ensemble d'actes possibles, choisir un (ou tous) de ceux qui maximisent (ou minimisent) un index donné [...] , bien souvent le coeur du problème est de trouver l'index approprié." L'index évoqué renvoie clairement à la notion d'attente. Il s'agit encore d'évaluer nos attentes pour chaque action puis de choisir l'action aux attentes les plus élevées. Ce qui change ici est que les attentes ne sont pas attachées à la valeur qu'on donne à un certain résultat mais à la valeur que l'agent donne à un ensemble de résultats.

Nos attentes, attX, d'un nombre inconnu X sont usuellement définies comme la moyenne pondérée des valeurs dont nous pensons que X peut les prendre, dans laquelle les poids sont nos probabilités que X soit l'un de ces nombres. Les attentes sont classiquement définies en termes de probabilités.

Par exemple si on sait que X peut prendre un nombre fini de valeurs $x_1, \cdots, x_n$, on peut définir nos attentes comme suit :

$$attX = \sum_{i=1}^{n} p_i x_i \text{ où } p_i = p(X = x_i)$$

Nous voyons bien que cette définition des attentes en termes de probabilités repose sur deux hypothèses. D'abord, les probabilités en question doivent être disponibles. Ce qui est loin d'être acquis dans tous les problèmes de décision. D'un autre côté, le fait qu'on puisse multiplier des probabilités p_i par des valeurs x_i de X, appelée hypothèse de commensurabilité, écarte du domaine de définition des attentes celles attachées à des préférences qualitatives (sans intensité de préférence) ou des préférences incomplètes (contenant potentiellement des incomparabilités).

Si l'on veut donner une définition générale des attentes, il est possible, en prenant le chemin inverse de la définition précédente, de voir les attentes comme le concept initial. Pour illustrer ce fait, on peut définir les probabilités commes les attentes dans le cas de nombres X particulièrement simples appelés "indicateurs". L'indicateur d'une hypothèse H est une constante, I_H, qui vaut 1 si H est vraie et 0 si H est fausse. Dans ce cas les probabilités sont des attentes d'indicateurs :

$$p(H) = att(I_H)$$

Déjà au 18ème siècle, Thomas Bayes définissait les probabilités en termes d'attentes :

> " La probabilité d'un évènement quelconque est le ratio entre la valeur à laquelle une attente dépendant de la réalisation de l'évènement devrait être calculée, et la valeur de la chose attendue à sa réalisation."

Il est donc nécessaire de se donner un moyen de calculer nos attentes pour chaque type de problèmes.

2.1.2 Qualité de l'information

Afin de calculer (évaluer) les attentes du décideur, les matrices d'utilité sont combinées avec différents types d'information à propos des états du monde. On peut voir les différents cas dans lesquels peut se trouver un agent, en fonction du type et de la précision de l'information dont il dispose, comme un spectre à une extrémité duquel est celui où le décideur sait quel état de la nature adviendra. A l'autre limite du spectre de l'information disponible se trouve le cas où le décideur n'a strictement aucune information. Entre les deux et dans les autres parties du spectre, se trouvent des situations que l'on nomme parfois risque ou incertitude (et ignorance).

Afin de classifier les parties du spectre, on peut se baser sur le travail de Knight [Knight 1921], qui écrivit

> "le terme 'risque', comme grossièrement utilisé dans le dicours quotidien et dans les discussions économiques, couvre en réalité deux choses qui, fonctionellement au moins, dans leurs relations causales au phénomène de l'organisation économique, sont catégoriquement différentes."

En d'autres termes, le mot risque est invariablement utilisé pour décrire deux points du spectre qui demandent chacun une procédure particulière de décision. Il est en effet souvent utilisé à la fois pour décrire des quantités mesurables ou non mesurables. Les quantités évoquées se rapportent à l'incomplétude de l'information. Il proposa alors de réserver le terme "incertitude" aux cas non quantifiables. Le "risque" ne couvrant plus alors qu'une évaluation quantifiable de l'information incomplète.

Ces distinctions sont faites pour différencier le cas probabiliste des autres. Knight reste un peu vague sur ce qui est susceptible d'être mesuré. De plus les frontières entre les classes distinguées sont floues si on considère la multiplicité des représentations possibles de la "non-certitude". La distinction se base surtout sur le caractère quantitatif de cette représentation. Par exemple, [Lehmann 1996] considère qu'une représentation du manque de connaissance par des probabilités imprécises rentre dans le cas qualitatif.

Une autre classification est celle présentée par Luce et Raiffa [Luce & Raiffa 1957] :

" on dira que l'on est dans le royaume de la décision sous :

- (a) **Certitude** si chaque action est connue comme menant invariablement à un résultat spécifique.
- (b) **Risque** si chaque action mène à un ensemble de résultats spécifiques, chaque résultat se produisant avec une probabilité connue. Les probabilités sont supposées connues du décideur. Par exemple, une action peut mener au résultat risqué suivant : un gain de 10 si une pièce non truquée tombe sur face, et une perte de 5 si elle tombe sur pile. Bien sûr, la certitude est un cas dégénéré du risque où les probabilités sont 0 et 1.
- (c) **Incertitude** si une ou plusieurs actions ont pour conséquence un ensemble de résultats spécifiques, mais que les probabilités de ces résultats sont complètement inconnues ou ne sont même pas porteuses de sens". ([Luce & Raiffa 1957], p. 13)

Luce et Raiffa illustrent le fait que passer de (a) à (b) puis (c) se fait au prix d'une perte d'information en utilisant un changement d'acteur dans l'exemple de l'omelette de Mazen. Si l'on suppose que celui qui doit préparer l'omelette est un "fermier scientifique", alors on peut supposer que l'on est dans une situation de risque puisqu'il peut utiliser une certaine distribution de probabilités acquise par exemple lors de précédentes expériences d'omelettes. Si maintenant l'acteur est un simple citadin "qui n'a absolument aucune idée d'une telle distribution", alors on est dans un cas d'incertitude.

La classification se fait sur la nature de l'information disponible, mais alors que Knight reste vague sur le sujet de l'information en parlant de "quantités" mesurables, Luce et Raiffa dans l'énumération précédente évoquent les conséquences de l'action comme objets. Dans la représentation utilisée, cela revient à une information sur l'occurrence des états du monde. Dans ces deux classifications, le terme d'incertitude est réservé au cas où des probabilités ne sont pas disponibles. Luce et Raiffa semblent de plus le réserver au cas où aucune information probabiliste n'est disponible. Il existe toutefois une grande variété de cas intermédiaires où une information existe bien qu'elle ne soit pas une distribution de probabilités. Ainsi en est il du cas où l'agent considère simplement que certains états du mondes sont plus plausibles que les autres, voire le cas où il a à sa disposition un ensemble de distributions de probabilités dont il sait que l'une d'elles décrit la probabilité des états. Et beaucoup, peut être la plupart, des problèmes de décision sont entre les catégories du risque et de l'incertitude, comme définies par Luce et Raiffa. Prenons, en l'occurence, la situation d'un agent qui décide de ne pas prendre un parapluie en sortant. Il ne connaissait certainement pas la probabilité qu'il pleuve, ce qui fait que ce n'était pas une décision sous risque. D'un autre côté, la probabilité de la pluie ne lui était pas complètement inconnue. Compte tenu de la saison et les bulletins météo qu'il a pu consulter, il savait que la probabilité serait au dessus de 5% et moins que 99%. Il est commun d'utiliser le terme d'incertitude pour couvrir aussi de telles situations de connaissance partielle des probabilités. L'incertitude plus stricte évoquée par Luce et Raiffa est nommée "ignorance". (cf. [Alexander 1975]).

La classification que nous adopterons dans la suite considére que le terme incertitude couvre tous les cas ne correspondant pas au risque (qui, lui, correspond à une distribution unique et parfaitement connue de probabilités). On réserve le terme d'incertitude stricte (ou complète ou ignorance) à la sous-classe de ces situations dans laquelle aucune information (probabiliste ou non) n'est disponible. Nous obtenons donc le spectre donné à la figure 2.1 :

Dans la suite, nous allons parcourir le spectre pour examiner, dans chacune de ses parties, la possibilité pour l'agent de prendre une décision rationnelle. Nous commencerons par la partie risque du spectre. Dans cette situation, les notions importantes pour la décision sont celle de probabilité et d'utilité espérée. Nous examinerons ensuite les situations d'incertitude en commençant par celle où l'agent essaye de revenir vers la situation précédente au moyen des probabilités subjectives. Puis nous arriverons à celle où, ne possédant pas cette information, on peut quand même prendre une décision basée sur des critères qualitatifs. Manquer même de ce type d'information mène aux situations d'incertitude stricte dont l'appauvrissement nous met dans une situation d'ignorance, l'autre limite du spectre.

2.2 Décision face au risque

Dans cette section, nous examinons la situation dans laquelle un décideur doit prendre une décision dans le risque. Nous présentons la notion de probabilité utilisée ainsi que ses différentes interprétations. Ces interprétations permettent de bien déterminer le cadre dans lequel cette notion peut être utilisée. Cela fait, on pourra mieux cerner l'intuition derrière cette notion, ce qui sera utile à évaluer les attentes par rapport aux actions entreprises dans des cadres où les probabilités ne pourront plus être utilisées.

2.2.1 Probabilités

On lance un dé, et on s'intéresse au nombre qui apparaît sur la face supérieure du dé. Cette expérience est une expérience aléatoire : son résultat dépend du hasard. L'ensemble des résultats possibles de cette expérience aléatoire s'appelle l'univers des possibles. Si le dé comporte 6 faces, l'univers des possibles est $\{1, 2, 3, 4, 5, 6\}$.

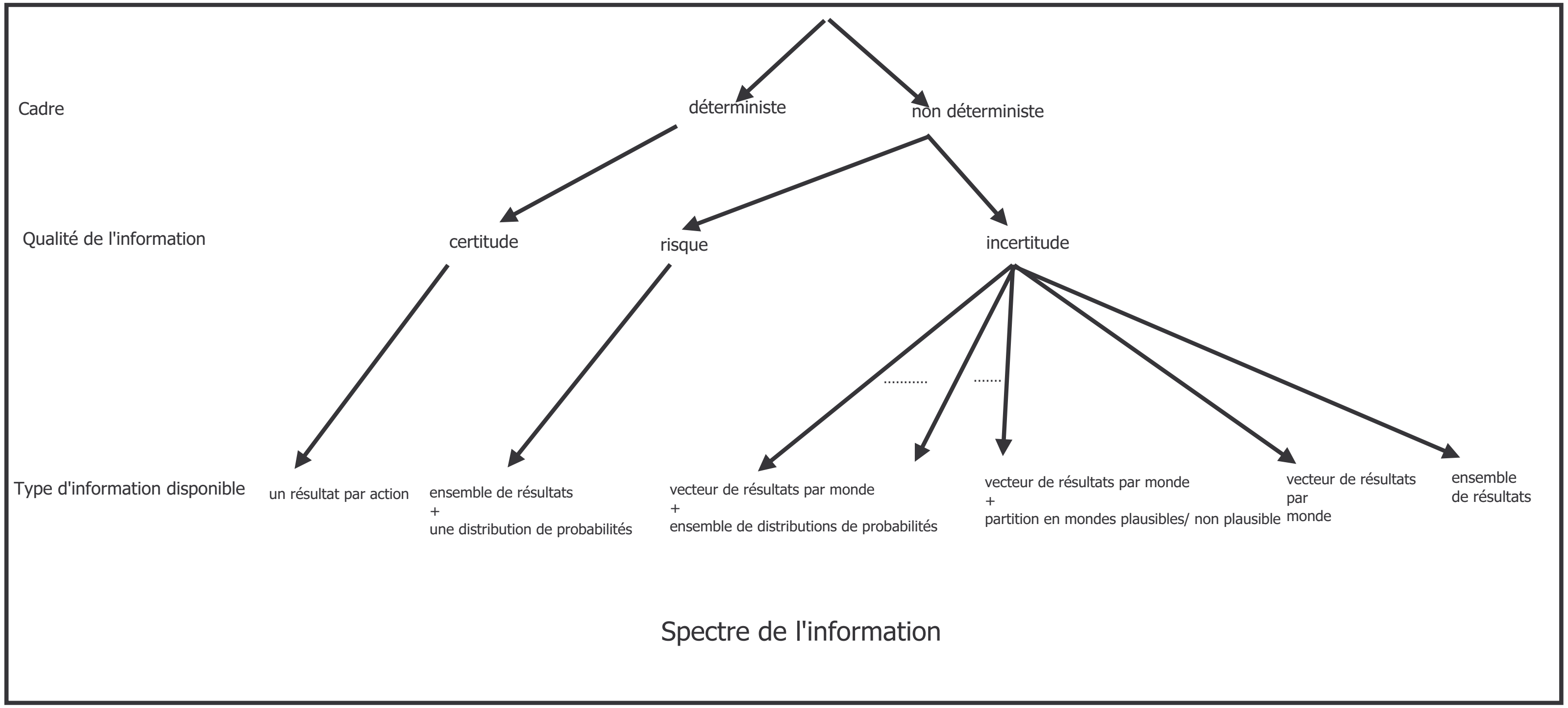
Intéressons-nous à des évènements de l'expérience aléatoire, c'est-à-dire à des faits qui peuvent se produire. Par exemple, on peut choisir les évènements A="on obtient un 6" et B="on obtient un nombre pair". On répète l'expérience plusieurs fois, et on étudie la fréquence de réalisation de l'évènement A. Quand le nombre de tirages augmente, la fréquence de réalisation de A tend à se stabiliser autour d'un nombre limite, compris entre 0 et 1. Ce nombre limite signifie intuitivement la chance qu'a l'évènement A de se produire lorsqu'on réalise une expérience : on l'appelle probabilité de A, et on le note $p(A)$. Dans notre exemple, on a bien sûr $p(A) = 1/6$ et $p(B) = 1/2$ si le dé n'est pas pipé.

Définition 2.2 (Loi de probabilité) *E est un ensemble d'évènements sur lequel se déroule une épreuve aléatoire. E est considéré comme l'évènement "un évènement quelconque se produit". $\emptyset$ est considéré comme l'évènement "aucun évènement dans E ne se produit". Un sous-ensemble $A \subseteq E$ est considéré comme l'évènement "un des évènements dans A se produit".*

Une loi de probabilité est une application p qui, à un évènement, fait correspondre un nombre réel compris entre 0 et 1.

Cette loi de probabilité vérifie les propriétés suivantes :

– $p(E) = 1$. *La probabilité de l'évènement E a une probabilité égale à 1.*
– $p(\emptyset) = 0$ *La probabilité de l'évènement impossible $\emptyset$ a une probabilité égale à 0.*


Cadre
déterministe
non déterministe
Qualité de l'information
certitude
risque
incertitude
....
....
Type d'information disponible
un résultat par action
ensemble de résultats
+
une distribution de probabilités
vecteur de résultats par monde
+
ensemble de distributions de probabilités
vecteur de résultats par monde
+
partition en mondes plausibles/ non plausible
vecteur de résultats
par
monde
ensemble
de résultats
Spectre de l'information

– *Si A et B sont deux sous ensembles disjoints ($A \cap B = \emptyset$) de E alors* $p(A \cup B) = p(A) + p(B)$.

La signification des probabilités a été un objet de débat depuis sa première formulation à la fin du 17ème siècle. Cette signification a une importance particulière dans le cadre de la décision en intelligence artificielle. Les probabilités vont être utilisées, comme déjà évoqué, pour estimer la valeur qu'on donne à une action dont l'issue est incertaine. Une interprétation différente influe forcément sur cette estimation ou plus clairement sur le champ d'application de la notion d'utilité associée.

Le débat se situe principalement entre les supporters de deux interprétations : les subjectivistes et les fréquentistes. Les fréquentistes généralisent l'intuition présente dans l'exemple introductif de cette section. Ils soutiennent que la probabilité d'un évènement est la fréquence avec laquelle il se produit sur un nombre suffisant de répétitions. Les subjectivistes, quant à eux, arguent que cette probabilité est le degré auquel un agent va y croire. Ce degré peut être mesuré à la vue de son intérêt à parier dessus.

Interprétation fréquentiste

Au vingtième siècle, des techniques statistiques fréquentistes développées se sont montrées utiles dans les domaines de la biologie et des sciences sociales. Leur application dans ces domaines n'a pas d'interprétation subjective évidente. Le statisticien et généticien R.A. Fisher, dont l'apport des travaux statistiques est dominant, était parmi les plus grands critiques de l'approche subjectiviste. L'intérêt est ici de déterminer la probabilité d'un certain évènement dans une certaine expérience. Par exemple, l'expérience peut être le lancer de deux dés, et nous serions intéressés par la probabilité de l'évènement "la somme des nombres obtenus est 6". On suppose que l'on peut répéter cette expérience autant de fois qu'on le désire avec exactement les mêmes conditions de départ. Dans cet exemple, on suppose que l'on peut lancer les dés "de la même manière" à chaque fois.

Voici le résultat de 50 de ces expériences identiques :

[4, 10, 6, 7, 5, 10, 4, 6, 5, 6, 11, 11, 3, 3, 6, 7, 10, 10, 4 , 4, 7 , 8, 8, 7, 7, 4, 10, 11, 3, 8, 6, 10, 9, 4 , 8 , 4, 3, 8, 7, 3, 7, 5, 4, 11, 9, 5, 5, 5, 8, 5].

Afin d'approcher la probabilité de l'évènement précédent, on compte le nombre de fois où la somme est égale à 6 et on le divise par le nombre total d'expériences. En d'autres termes, la probabilité de l'évènement est (approximativement) égale à sa fréquence pendant les expériences réalisées.

$$p(\text{la somme des nombres obtenus est } 6) = 5/50 = 0.1$$

La fréquence relative observée est seulement une approximation de la probabilité réelle d'un évènement. Toutefois, si nous étions capable de refaire cette expérience autant de fois qu'on le désire, la fréquence de l'évènement convergerait de plus en plus vers sa probabilité.

Cette interprétation se base donc sur l'hypothèse importante que l'expérience peut être répétée à l'identique. Cela la rend notamment très populaire surtout pour des situations économiques. A contrario, elle implique la nécessité d'une autre approche comme remède à l'inadaptation de l'inteprétation statistique dans des cas de décision impliquant des contingences qui par nature ne se produisent qu'une fois.

Interprétation subjective

L'interprétation fréquentiste qui vient d'être présentée est utile et adaptée dans des situations de choix qui peuvent se répéter dans des conditions identiques. Mais nous voulons aussi traiter des situations de choix dont on sait qu'elles se présentent une fois (ou rarement).

L'interprétation subjective vise à répondre à ce besoin. Les subjectivistes, dont Leonard J. Savage [Savage 1954] est sans doute celui qui fait triompher cette interprétation, soutiennent que la probabilité d'un évènement est le degré auquel un agent y croit, degré que l'on peut mesurer comme la quantité d'argent qu'il est prêt à parier sur sa réalisation. Prenons par exemple le cas du bug de l'an 2000 et supposons que l'on soit intérressé par le fait qu'il se produise. L'ensemble E des évènements contient deux éléments "le bug se produit" et "le bug ne se produit pas". On pourra vérifier le résultat une et unique fois. toute simulation à moins d'être globale ne saurait être considérée comme une reproduction de l'évènement. Et une simulation globale est extrêmement difficile sinon impossible. Dans ce cas, comment doit-on envisager les probabilités ? On assigne dans l'exemple précédent un nombre (une probabilité) qui reflète sa croyance personnelle dans le fait que le bug se réalisera. Si on est sûr que le bug se produira, on assignera une probabilité de 1 à l'évènement "le bug se produit". Si on en doute fortement et qu'on pense que c'est impossible. On assignera une probabilité de 0. Si on ne sait pas, on assignera un nombre entre 0 et 1 qui sera d'autant plus proche de 0 ou 1 selon qu'on soit plus enclin à penser que le bug se produira ou non.

Dans son livre "The foundations of statistics" [Savage 1954], Savage met en avant un ensemble d'axiomes sur les préférences des agents au sujet des actions (selon ses termes des actes). Le respect de ces axiomes par ses préférences conditionne selon Savage la rationalité de l'agent.

La probabilité subjective reflète une opinion sur la réalisabilité d'un évènement. La probabilité que l'on assigne à sa propre réussite dans un examen peut être différente de la probabilité qu'assignera un professeur à ce même évènement. Les probabilités sont personnelles et dépendent des personnes. Cette probabilité doit en même temps satisfaire les conditions mathématiques d'une probabilité "bien formée". Ces deux exigences peuvent être difficiles à concilier.

Discussion

Le principal reproche des fréquentistes aux probabilités subjectives est qu'elles reposent sur le concept de degré de croyance qui n'est pas une valeur empirique en soi. Ce point de vue n'est pas partagé par exemple par les économistes qui ont pour objet principal d'observation et d'expérimentation le comportement d'agents qui, comme le montre la proposition de Savage, traduit une certaine distribution de degrés de croyances ou de probabilités subjectives.

En probabilités bayésiennes, le compromis se fait entre les deux thèses autour de la notion de probabilités a priori. C'est la description de ce qui est connu par l'agent (son état de croyance) au sujet d'une variable (ici les mondes possibles) en l'absence de preuves (des données statistiques). Le but y est d'estimer les probabilités objectives à partir de données statistiques, et ce en utilisant des probabilités a priori pour améliorer les estimations. Parmi les auteurs adhérants à cette vision l'on peut citer John Maynard Keynes [Keynes 1921], Harold Jeffreys [Jeffreys 1939]. [Nagel 1939], [Kyburg & Smokler 1964] donnent plus d'information sur l'historique du débat.

Les deux approches se complètent. La fréquentiste étant en général préférable lorsque les informations sont abondantes et d'un faible coût de collecte. La subjectiviste l'est dans le cas où elles sont rares et difficiles à rassembler. Dans l'optique d'un problème de décision, l'approche statistique nous sert à justifier/ guider la décision de l'agent quand on dispose d'une information probabiliste suffisante. On est alors dans le cadre du risque. L'approche subjectiviste nous permet d'étendre la méthode de décision dans le risque au cas d'expériences uniques ou dans lesquelles les observations ne sont pas suffisantes. Elle affirme en effet que, dans ces cas, un agent rationnel se comporte comme s'il possédait une distribution de probabilités.

2.2.2 Utilité espérée

Dans le cadre général (et traditionnel), on suppose que le décideur connaît l'ensemble des états du monde. Concernant le degré d'incertitude et suivant la taxonomie adoptée précédemment, il existe deux situations ; celle où l'agent est capable d'assigner à chaque état potentiel une probabilité précise (qui sera traité ici), et celle où il n'en est pas capable (qui sera traitée dans les sections suivantes). Pour traiter la situation de décision (d'un seul agent) face à une situation de risque, on peut supposer acquise l'existence d'une distribution de probabilités sur les états du monde, qu'elle soit donnée ou formulée par l'agent. La décision face au risque commence là et suppose que les probabilités sont une donnée de base. L'agent est supposé connaître le résultat de l'exécution d'une action dans chacun des états du monde. Il connaît alors la probabilité d'obtenir chacun des résultats possibles. Sous ces hypothèses, une action a peut être identifiée à une liste (finie) de résultats, avec une probabilité associée à chacune. Ainsi représentée, elle est appellée lotterie et notée comme suit :

$$a = \{p_1 : x_1; p_2 : x_2; ...; p_j : x_j; ...; p_n : x_n\}.$$

Le décideur reçoit le résultat x_j avec la probabilité p_j pour $j = 1, 2, ...n$

Afin d'être capable de faire des choix rationnels, le décideur doit expliciter ses préférences sur les actions et donc sur les lotteries. Le modèle classique de décision face au risque est celui de l'utilité espérée (EU). Dans ce modèle une lotterie est préférée à une autre si son utilité espérée est plus élevée.

Définition 2.3 (Utilité espérée) *Considérons une lotterie a qui donne p_i chances d'obtenir un résultat x_i $(i = 1, \cdots, n)$. L'utilité espérée de cette lotterie est définie par $EU(a) = \sum p_i u(x_i)$, où $u(x_i)$ représente l'utilité d'avoir x_i.*

L'utilité espérée fût d'abord introduite par [Bernoulli 1954] en 1738 comme résolution du paradoxe de st. Petersbourg (cf section 2.1).

Il nous faut calculer le gain moyen du joueur au cours d'une partie : ce doit être la mise initiale pour que le jeu soit équitable. Si face intervient dès le premier lancer, on gagne 2 euros. La probabilité pour que cela arrive est $1/2$, ce qui donne une espérance pour ce coup de $1/2 \times 2 = 1$. Si face intervient pour la première fois au 2ème lancer, ce qui se produit avec une probabilité de $1/2 \times 1/2 = 1/4$, le gain est de 4 euros, ce qui fait une espérance de gain de 1 euro pour ce coup. Plus généralement, si face apparaît pour la première fois au n-ième lancer, ce qui se produit avec une probabilité de $1/2^n$, le gain est de 2^n euros, d'où une espérance de 1 euro pour ce coup. Maintenant, l'espérance totale s'obtient en sommant l'espérance de tous les cas possibles. On somme une infinité de termes qui valent tous 1 : la somme est bien sûr infinie. Il faudrait donc miser une infinité d'euros pour que le jeu soit équitable. Aucun joueur rationnel ne devrait donc participer à ce jeu qui ne sera pas équitable car sa fortune est finie. Pourtant ce n'est pas le comportement observé.

Alors que Bernouilli suppose l'existence de fait de la fonction d'utilité, les travaux de von Neumann et Morgenstern [von Neumann & Morgenstern 1947] conduisirent à une approche axiomatique pour fournir un ensemble de conditions nécessaires et suffisantes à la validité de l'utilité espérée. Ce sont des axiomes dont l'objet est de spécifier quelle doit être une préférence rationnelle sur les lotteries. L'axiomatisation en question suppose qu'il existe une relation binaire $\leq$ sur l'ensemble X des lotteries. Le premier axiome est alors

axiome (A1) *$\leq$ est transitive et totale.*

Cet axiome n'est pas particulier au cas de la décision dans le risque. Le deuxième axiome est donné par :

axiome (A2) *Soient $x, y, z \in X$. Si $z < x < y$ alors $\exists \alpha \in [0,1], z < (\alpha * x + (1 - \alpha) * y)$ et $\exists \beta \in [0,1], (\beta * x + (1 - \beta) * y) < z$.*

[Mongin 1997] note qu'en présence du premier "cet axiome donne la possibilité de représenter la donnée qualitative [i.e., la préférence que l'on a à propos des éléments de X] par une certaine fonction numérique $u(x)$ même si cette dernière reste non spécifiée."

Le troisième axiome est souvent présenté comme le plus important :

axiome (A3) $\forall x, y, z \in X, \forall \alpha$ *tel que* $0 < \alpha < 1, x \leq y$ *si et seulement si* $\alpha * x + (1-\alpha) * z \leq \alpha * y + (1-\alpha) * z$.

En d'autres termes, si on préfère y à x alors on préfere toute lotterie du type $\{\alpha : y; 1-\alpha : z\}$ à une lotterie du type $\{\alpha : x; 1-\alpha : z\}$.

Quand ces axiomes sont satisfaits, le théorème de l'utilité espérée peut s'appliquer.

Théorème 2.1 *Soient* $\{p_1 : x_1; \cdots; p_k : x_k\}$ *une lotterie* x, *et soit* $\{p_1^{'} : x_1^{'}; \cdots; p_k^{'} : x_k^{'}\}$ *une lotterie* $x^{'}$. *Si les axiomes (A1),(A2) et (A3) sont vérifiés alors*

$$\exists u \quad \textit{une fonction d'utilité telle que} \quad x \leq y \quad \textit{si et seulement si} \quad \sum p_i * u(x_i) \leq \sum p_i^{'} u(x_i^{'})$$

Cette fonction u *est unique à une transformation linéaire positive près.*

Il n'est pas supposé que u existe de manière consciente chez le décideur. On dit seulement que les conditions imposées sur les préférences induisent que le décideur se comporte comme un maximisateur d'utilité espérée.

L'axiome (A3) a été souvent remis en question. [Savage 1954] avance l'argument suivant en sa faveur. Soit le choix entre $\{p : x_1, (1-p) : x_3\}$ et $\{p : x_2, (1-p) : x_3\}$. Supposons alors que le décideur reçoit x_1 ou x_2 si une pièce truquée retombe sur face et x_3 si elle retombe sur pile. Si la pièce retombe sur pile, peu importe le choix du décideur. Si le résultat est face, il doit choisir x_1 s'il le préfère à x_2 et il choisira x_2 autrement. Ainsi cette logique ramène le choix entre x_1 et x_2 au choix entre $\{p : x_1, (1-p) : x_3\}$ et $\{p : x_2, (1-p) : x_3\}$. Cet argument tend à justifier l'axiome d'un point de vue normatif. En d'autres termes, cet argument montre que le décideur doit agir comme s'il était dans ce cas.

La validité de la théorie de l'utilité espérée fût remise en doute par Allais au moyen de l'exemple suivant [Mongin 1997].

Exemple 2.1 On soumet à un groupe de personne le questionnaire suivant :
- Question 1 : quelle lotterie choisiriez-vous entre x1= recevoir 100 millions avec la probabilité 1, et y1= recevoir 500 millions avec la probabilité 0.10, 100 millions avec la probabilité 0.89, et rien avec la probabilité 0.01.
- Question 2 : quelle lotterie choisiriez-vous entre x1= recevoir 100 millions avec la probabilité 0.11 et rien avec la probabilité 0.89, et y1= recevoir 500 millions avec la probabilité 0.10 et rien avec la probabilité 0.90.

Allais trouva que la majorité des réponse étaient x1 à la question 1 et y2 à la question 2, et argumenta que ces deux lotteries pouvaient être en effet choisies pour de bonnes raisons. Mais ce choix viole la théorie de l'utilité espérée, puisqu'il n'existe pas de fonction U qui satisferait à la fois : $U(100) > 10/100U(500) + 89/100U(100) + 1/100U(0)$, et $11/100U(100) + 89/100U(0) < 10/100U(500) + 90/100U(0)$.

2.3 Décision face à l'incertitude : cadre subjectif

Savage a entrepris en 1954 [Savage 1954] d'étendre le puissant outil de choix qu'est l'utilité espérée du cadre du risque à celui de l'incertitude. La nécessité d'une théorie spécifique vient du fait que dans le cadre de l'incertitude, les attentes de l'agent ne peuvent plus être quantifiées au moyen de probabilités comme c'était le cas dans le cadre du risque. Savage identifia pour ce faire un certain nombre de propriétés que l'on souhaite voir vérifiées par la fonction de choix rationnel entre les alternatives dans le cadre d'incertitude étudié. Il montra qu'alors cette fonction de choix est identifiable à l'application de la maximisation d'une utilité espérée dont les probabilités sur lesquelles elle se base sont cette fois-ci subjectives. Cette interprétation fût par ailleurs d'abord introduite dans [Ramsey 1931], puis développée de manière indépendante dans [de Finetti 1937].

2.3.1 Cadre formel

Une action est modélisée comme une application $a : S \to \mathcal{C}$. Suivant l'approche de Savage, n'importe quelle application dans l'ensemble $\mathcal{C}^{\mathcal{S}}$ est considérée comme une action possible. Un évènement est un sous-ensemble d'états, compris de manière disjonctive. Dire qu'un évènement E se réalise signifie que l'un des états dans E est le "vrai" état. Parmi les actions dans $\mathcal{C}^{\mathcal{S}}$ se trouvent les actions constantes telles que : $\exists c \in \mathcal{C}$, $\forall s \in S, a(s) = c$. Elles sont notées a_c. Soit $E \subseteq S$ un évènement, a et b deux actions, et notons par aEb l'action telle que : $aEb(s) = a(s)$ si $s \in E$, et $b(s)$ si $s \notin E$. Cette définition est l'expression de l'action comme application dont le domaine de départ est l'ensemble des états. On définit une action particulière qui a les mêmes conséquences que a sur une partie des états et les mêmes conséquences que b sur le reste des états. Le conditionnement suggéré n'est qu'apparent. Il n'implique pas une observation supplémentaire. Cette notation permet de comparer de manière aisée le résultat de deux actions sur un sous-ensemble d'états. Plus généralement, la notation $a_1 E_1 a_2 E_2, \cdots, E_{n-1} a_n$, où $\{E_1, \cdots, E_{n-1}, E_n\}$ est une partition de S, dénotent l'action dont le résultat est $a_i(s)$ si $s \in E_i (i = 1, \cdots, k-1)$ et $a_n(s)$ si $s \notin \bigcup_{i=1}^{n-1} E_i$.

Exemple 2.2 Soient les actions a : "contourner la ville" et b : "traverser la ville". E est l'évènement : "présence d'un bouchon". S repésente des descriptions de l'état du réseau routier, et $\mathcal{C}$ représente une échelle de temps pour le temps passé par l'agent qui conduit vers son travail. aEb signifie alors : "contourner la ville s'il existe des bouchons, et traverser la ville sinon".

L'approche de Savage est alors la suivante. Supposons qu'un décideur ait une relation de préférence $\trianglerighteq$ sur les actions. Savage énonce un ensemble d'axiomes traduisant une prise de décision rationnelle dans le cadre étudié puis explicite l'effet de la satisfaction de cet ensemble sur $\trianglerighteq$.

2.3.2 Axiomes

L'axiome central est celui du "principe de la chose sûre" (*sure thing principle*). Cet axiome se base sur la même intuition que l'axiome (A3) rencontré dans le cadre de la décision face au risque (voir section 2.2). Il exige que la préférence relative entre deux actions ne dépende pas des états où les actions ont les mêmes conséquences. En d'autres termes la préférence entre aEh et bEh ne dépend pas du choix de l'action h :

axiome (Principe de la chose sûre) *Soit E un évènement et a, b, h, h' des actions, $aEh \triangleright bEh$ ssi $aEh' \triangleright bEh'$.*

	E_1	E_2	E_3
a_1	100	0	0
a_2	0	100	0

TABLE 2.4 – PRINCIPE DE LA CHOSE SÛRE (1)

Considérons pour l'illustrer l'exemple de la table 2.4 :

Préférer a_1 à a_2 peut être interprété comme venant d'une croyance que E_1 est plus probable que E_2. Comme E_3 entraînera certainement une conséquence de 0, cet évènement ne doit pas influer sur les préférences de l'agent sur a_1 et a_2. Ainsi, le décideur devrait faire le même choix dans la situation suivante (table 2.5) dans laquelle on remplace E_3 par $E_3^{'}$.

	E_1	E_2	$E_3^{'}$
a_1	100	0	100
a_2	0	100	100

TABLE 2.5 – PRINCIPE DE LA CHOSE SÛRE (2)

Savage adjoint au principe de la chose sûre un ensemble d'axiomes, ce qui donne la liste suivante.

axiome (S1) *La relation de préférence entre les actions, $\trianglerighteq$, est un préordre.*

axiome (S2) *Soit E un évènement et a, b, h, h' des actions, $aEh \triangleright bEh$ ssi $aEh' \triangleright bEh'$.*

axiome (S3) $\forall E \subseteq S$, *E non vide, $x, y \in \mathcal{C}$. $(a_x \trianglerighteq a_y)E$ ssi $x > y$.*

axiome (S4) $\forall x, y, x', y' \in \mathcal{C}$ *t.q.* $x > y, x' > y', xEy \triangleright xEy$ *ssi* $x'Ey' \triangleright x'Ey'$.

axiome (S5) *$\mathcal{C}$ contient au moins deux éléments x, y tels que $a_x \triangleright a_y$ (où $x > y$).*

axiome (S6) *Pour toutes actions a, b avec $a \triangleright b$ et tout $c \in \mathcal{C}$, il existe une partition $S_1, \cdots, S_n$ de S telle que $\forall i = 1, \cdots, n, cS_ia \triangleright b$ et $a \triangleright cS_ib$.*

L'axiome (S1) semble incontournable dans le cadre de la théorie de l'utilité. Si les actions sont ordonnés par rapport à l'utilité espérée, alors la préférence sur les actes sera transitive, réflexive et complète. Ce que cet axiome implique, si $\mathcal{C}$ et S sont finis, est qu'il existe une échelle totalement ordonnée, disons L, qui peut servir à évaluer la valeur des actions. La classe d'équivalence d'une action a est alors le niveau d'utilité qualitatif de a.

Il semble raisonnable d'identifier l'ensemble des actions constantes $\{a_x, x \in \mathcal{C}\}$ à $\mathcal{C}$. La préférence sur $\mathcal{C}$ peut être induite de $\trianglerighteq$ comme suit :

Définition 2.4 *La relation de préférence $\geq$ sur $\mathcal{C}$ est telle que $\forall x, y \in \mathcal{C}$, $x \geq y$ ssi $a_x \trianglerighteq a_y$.*

Cette définition est consistante seulement si les préférences entre les actions constantes ne sont pas changées par le conditionnement. C'est le troisième axiome (S3). C'est une propriété d'extension que présente ici Savage. Les préférences sur les actions ne devraient pas contredire les préférences sur les conséquences.

Définition 2.5 *Une action a Pareto-domine une autre action b si et seulement si $\forall s \in S, a(s) > b(s)$ ou $a(s) = b(s)$, ce qui noté aPb.*

Clairement, la Pareto dominance devrait induire une préférence faible sur les actions.

Définition 2.6 *a est dit faiblement préféré b, conditionné sur E, que l'on note $(a \trianglerighteq b)E$, ssi $\forall h, aEh \trianglerighteq bEh$.*

Et effectivement sous $S1, S2$, et $S3$, pour deux actions a et b, aPb implique $a \triangleright b$.

L'axiome (S4) sert à ce que la préférence sur les actions induise une relation de probabilité parmi les évènements au moyen d'un isomorphisme entre les évènements et les actions ne produisant que deux conséquences possibles (qui sont toujours enviseageables d'après l'hypothèse introduite plus haut). Par cet axiome, on obtient un préordre sur les évènements.

Ensuite Savage suppose que l'ensemble $\mathcal{C}$ n'est pas trivial par l'axiome (S5). Finalement, l'axiome (S6) lui permet de dériver l'existence (et l'unicité) d'une mesure de probabilité numérique sur S qui peut représenter la relation obtenue sur les évènements.

Considérons l'expérience suivante $(e_1, c_1; \cdots; e_n, c_n)$ qui offre c_i si e_i se produit. L'utilité espérée subjective de cette expérience est donnée par $\sum p(e_i) * u(c_i)$, où $u(c_i)$ est une fonction d'utilité standard comme dans le modèle (EU) et p est une mesure de probabilité subjective qui obéit aux axiomes des probabilités standard. Ainsi la théorie de l'utilté espérée subjective est une extension naturelle de (EU) au cadre de l'incertitude.

La principale motivation qu'affiche Savage est de donner un argument en faveur de son interprétation des probabilités. Un second type de motivation porte sur l'utilisation du modèle d'utilité espérée comme théorie descriptive du comportement humain en situation d'incertitude. En effet, on peut alors considérer les axiomes essentiellement comme les conséquences observables du comportement d'un agent maximisateur d'utilité espérée. Ils fournissent donc un moyen de mettre à l'épreuve la validité descriptive de la théorie. L'axiomatisation est dans ce cas un outil d'évaluation descriptive. Enfin, on peut concevoir une motivation normative : l'axiomatisation nous donne les moyens d'évaluer la validité normative du modèle d'utilité espérée. L'axiomatisation est un moyen de répondre à la question de savoir si la maximisation de l'utilité est le bon guide à suivre en situation d'incertitude.

2.4 Cadre de l'incertitude qualitative

L'approche présentée dans la section précédente se base sur un cadre probabiliste. Ce cadre est pertinent si l'incertitude considérée peut être décrite par une distribution de probabilités. Le rangement des actions se fait alors en utilisant l'utilité espérée des actions. Cette approche est particulièrement adaptée aux sciences économiques au sein desquelles elle fût développée avant d'être axiomatisée par [Savage 1954].

Les chercheurs en intelligence artificielle et spécialement dans le domaine de planification sous incertitude, qui présente des hypothèses proches, ont repris à leur compte cette approche. Dans ce cadre, un agent doit former un plan avec des actions dans un monde sur lequel il a une information incomplète. Si l'on suppose que son incertitude est décrite par une distribution de probabilités alors la notion d'utilité espérée semble toute indiquée pour guider son choix.

Toutefois, des travaux comme ceux de Brafman et Tennenholtz [Brafman & Tennenholtz 1996] et de Doyle et Thomason [Doyle & Thomason 1999] présentent des points clés d'une approche parallèle à la précédente. Cette approche nommée théorie de la décision qualitative garde encore la même structure générale d'un problème de décision comme décrite à la section 2.1. Elle s'affranchit toutefois de la notion d'utilité espérée (numérique) comme fondement de la rationalité de l'agent.

Dans le cadre de cette théorie, l'incertitude n'est plus supposée décrite par une distribution de probabilité mais par une relation de préférence ordinale. On doit alors aller chercher ailleurs la justification de la rationalité des choix. Par exemple, dans les problèmes de planification de mouvements de robots, la qualité de l'information disponible (recueillie par des capteurs physiques) conjuguée à la rapidité avec laquelle une décision doit être prise (sans plus grande investigation) font rapidement apparaître la théorie de l'utilité espérée comme trop exigeante. Dans un autre type de problèmes qui peuvent être rencontrés par des agents logiciels d'aide à la décision, le décideur humain, source de l'information de l'agent logiciel, n'est souvent pas capable d'exprimer son incertitude en termes de probabilités précises ou même de quantifier ses préférences [Boutilier 1994], ce qui met en défaut une hypothèse importante du modèle (EU) (voir section 2.2.2) qui est celle de la commensurabilité entre probabilité et utilité (c'est-à-dire qu'on peut les représenter sur la même échelle).

Sur un autre plan, considérons l'interprétation statistique des probabilités. La notion d'utilité espérée basée sur cette interprétation est beaucoup plus adaptée à des situations expérimentales susceptibles de se répéter (et un grand nombre de fois) avec des conditions identiques à chaque fois. Elle est de ce fait beaucoup moins adaptée à des situations de planification de déplacement d'une navette dans l'espace par exemple où la multiplicité des facteurs influant sur le résultat des actions rend chaque expérience différente. L'interprétation subjective n'a, quant à elle, pas beaucoup de sens. On peut être raisonnablement réticent à parier sur le résultat d'une décision technique surtout si la navette est habitée. Dans ce cas, une modélisation qualitative du problème, particulièrement adaptée à un environnement changeant et partiellement observable semble toute indiquée [Sabbadin 2000].

Le terme de "cadre qualitatif" a été, dans ce qui précède, laissé vague. Les arguments précédents faisaient tous appel à l'intuition qu'on peut avoir de ce qu'un tel cadre doit être. Cette intuition correspond à plusieurs interprétations, chacune menant à une théorie différente. Ces théories remplissent un spectre qui va en s'éloignant de la théorie quantitative (probabiliste) de référence. Les modèles faisant partie du spectre s'en écartent et s'en différencient en relâchant les hypothèses qui la sous-tendent, par exemple en supposant que les probabilités existent mais qu'en on a une connaissance partielle comme chez [Lang 1996], [Bacchus & Grove 1996]. Une autre vision est celle de considérer que c'est le caractère précis, mathématique des outils du modèle (EU) qui l'empêche d'appréhender des préférences et des raisonnements qualitatifs. On introduit alors pour y remédier des relâchements à cette caractéristique. On peut le faire sur les probabilités elles-mêmes au moyen de la notion de probabilités vagues [Pearl 1996]. On peut aussi le faire en en enrichissant l'utilité espérée d'un concept qualitatif : la négligeabilité [Lehmann 1996].

A la limite opposée du spectre de la qualité de l'information disponible se trouve la vision du qualitatif comme le fait de se priver/ s'affranchir de tout moyen de représentation ou de raisonnement quantitatif dans le processus de décision. S'imposent alors des approches faisant appel à la notion de plausibilité [Boutilier 1994] ou à la théorie des possibilités, pendant qualitatif de la théorie des probabilités [Dubois & Prade 1988]. Toujours dans cette partie du spectre (essayant de ne pas utiliser de notion quantitative) se trouvent deux types d'approches qui ont pour but de garder une cohérence avec la théorie de la décision classique. Une hypothèse faite dans le cadre (EU) est celle de commensurabilité (de manière triviale elle signifie que l'on peut multiplier des probabilités par des utilités. On la retrouve dans la notion d'équivalent certain d'un évènement incertain : on peut évaluer une collection d'utilités pondérées par des probabilités comme une utilité). Quand on essaye de rester près de la théorie de Savage dans un cadre qualitatif, on doit faire un choix entre garder cette hypothèse, qui devient l'existence d'une échelle commune à la fois à l'utilité et à l'incertitude sur les états, ou l'écarter. La garder mène aux travaux de [Fargier & Sabbadin 2003a, Fargier & Sabbadin 2005]. Ne pas la garder mène aux travaux tels que [Dubois *et al.* 2000] qui supposent seulement l'existence d'un ordre partiel sur les états qui en exprime l'incertitude.

La théorie de Savage présentée dans la section 2.3 recherchait une justification de la rationalité de choix se basant sur une information qualitative (probabilités). Il utilise pour cela des axiomes qui capturent la notion de rationalité dans ce cadre. Si on "appauvrit" l'information disponible, il est tentant de garder une approche similaire pour caractériser un comportement de choix rationnel.

Dans [Dubois *et al.* 2003], les auteurs proposent un axiome qui prend en compte l'hypothèse de non-commensurabilité. Cet axiome se base sur la notion d'équivalence ordinale entre deux paires d'actions qui requiert que, pour chaque état de la nature, la préférence du décideur sur la première paire soit la même que celle sur la deuxième.

Définition 2.7 (Equivalence ordinale) *Soit S l'ensemble des états du monde. On note $\leq_e$ la relation de préférence du décideur sur ces états. Deux paires d'actions (a, b) et (a', b') sont ordinalement équivalentes ($(a, b) \equiv (a', b')$) si et seulement si $\forall s \in S, a(s) \leq_e b(s) \Leftrightarrow a'(s) \leq_e b'(s)$.*

L'axiome s'exprime alors de la manière suivante :

axiome (Invariance Ordinale) *Soit A un problème de décision. On note $\leq$ la relation de préférence du décideur sur les actions. $\forall a, a', b, b' \in A$, $(a, b) \equiv (a', b') \Rightarrow (a \leq b \Leftrightarrow a' \leq b')$*

Cet axiome signifie que, quand un décideur a à comparer deux actions, les conséquences d'une action ne sont pas importantes par elles-mêmes pour son choix. Seules comptent ici les positions relatives de ces conséquences par rapport à celles de l'autre action, et seulement celles de l'autre action. On pourrait alors changer la valeur de ces conséquences sans altérer le choix du décideur si on n'en change pas les positions relatives. La notion d'intensité de préférence n'est donc pas pertinente dans ce cadre. Cet axiome peut être vu comme "une version forte des postulats clés de la théorie de Savage, (S2) et (S4)". Le lien entre l'axiomatisation développée et la théorie de Savage ne s'arrête pas aux axiomes. [Dubois *et al.* 2003] suivent aussi son cheminement en partant de la supposition que le décideur possède une relation de préférence sur les actions et en reconstruisant alors à l'aide des axiomes les motivations qui l'on poussé à choisir cette relation. [Dubois *et al.* 2003] construisent alors en se basant sur les notions d'acte constant et d'acte conditionnel respectivement, deux relations de préférence. La première est une relation de préférence sur les conséquences :

Définition 2.8 (1) *Soit $\mathcal{A}$ un ensemble d'actions et $\mathcal{C}$ l'ensemble de conséquences associé. Soient $x, y \in \mathcal{C}$. Soient alors $a_x, a_y \in \mathcal{A}$ les actions constantes respectives (l'action a_x a pour seul résultat x et l'action a_y a pour seul résultat y). Alors la relation de préférence $\leq_c$ sur les conséquences est définie à partir de la relation de préférence $\leq$ sur les actions par :*

$$x \leq_c y \text{ si et seulement si } a_x \leq a_y$$

La deuxième relation exprime la confiance que possède un agent dans le fait qu'un certain évènement se réalisera.

Définition 2.9 (2) *Soit $\mathcal{A}$ un ensemble d'actions et $\mathcal{C}$ l'ensemble de conséquences associé. Soient A et B deux évènements. Alors la relation de confiance $\leq_C$ sur les évènements est définie par : $A \leq_C B$ si et seulement si $\exists x, y \in \mathcal{C}$ tels que $xAy \leq xBy$.*

On peut maintenant présenter le théorème principal de cette approche :

Théorème 2.2 *Si le décideur possède une relation de préférence sur les actions $\leq$ qui est complète sur les actions constantes, réflexive et satisfait l'axiome (S5) et l'axiome d'Invariance Ordinale alors il existe une relation de préférence $\leq_c$ $sur \mathcal{C}$ définie par (1) et une relation réflexive, préadditive et non triviale $\leq_C$ sur S définie par (2) telles que $\forall a, b \in \mathcal{A}$, $a \leq b$ si et seulement si $\{s \in S, a(s) \leq_c b(s)\} \leq_C \{s \in S, b(s) \leq_c a(s)\}$.*

Notons que les hypothèses faites sur $\leq$ dans le théorèmes sont plus faibles que dans la théorie de Savage. En effet $\leq$ n'est pas supposée transitive.

L'intérêt principal de ce théorème est qu'il montre que même si on abondonne l'hypothèse de commensurabilté entre la mesure des préférences du décideur et sa mesure de l'incertitude sur les résultats des actions, on obtient quand même un comportement guidé par à la fois par les attentes qu'il a (ses préférences sur les conséquences) et sur l'information qu'il détient sur la réalisabilité des différents évènements (ici la relation de confiance). On peut donc supposer que même quand on appauvrit encore l'information dont dispose le décideur, ce seront encore ces deux aspects qui forgeront sa décision.

2.5 Cadre de l'incertitude stricte

Dans cette section, on arrive à la partie du spectre la plus pauvre en information. On a jusqu'ici passé en revue les situations où l'agent possède une information probabiliste. On a vu qu'alors la théorie de l'utilité espérée (subjective ou non) pouvait être utilisée comme fondement d'un critère de décision. On a vu que même dans le cas où l'information que possède l'agent est plus qualitative on pouvait encore déterminer des critères raisonnables de décision, y compris en revenant, moyennant quelques adaptations, vers la théorie de l'utilité espérée. Dans cette section, on va examiner le cas encore plus pauvre en information où le décideur ne connaît d'une action que ses résultats possibles dans les différents états du monde. Il ne peut même pas former d'appréciation qualititative lui permettant de penser qu'un état se réalisera plutôt qu'un autre. Cela peut arriver quand l'agent est privé de toute information sur l'occurrence des états du monde dans un cadre totalement nouveau pour lui ou dans un cadre où la caractère critique de la décision à prendre l'empêche d'avancer toute hypothèse subjective.

2.5.1 Cadre formel

On suppose que l'agent dispose de l'ensemble Ω d'états du monde possibles. Soit $\mathcal{A}$ un ensemble d'actions et $\mathcal{C}$ un ensemble de conséquences (résultats possibles des actions de A). $\mathcal{C}$ contient aussi bien des conséquences "sûres" que des lotteries sur ces conséquences possibles. L'agent possède une relation de préférence $\geq$ sur $\mathcal{C}$. Selon le besoin des axiomes, on pourra demander que cette relation vérifie les axiomes de von Neumann et Morgenstern (présentés à la section 2.2.2) permettant de définir une fonction d'utilité sur $\mathcal{C}$. Une action (et plus généralement une décision) est définie comme une aplication $a : \Omega \mapsto \mathcal{C}$. Un problème de décision A est un ensemble de telles applications qui partagent un même ensemble d'états. Résoudre un problème de décision consiste à lui associer un sous-ensemble non vide d'éléments optimaux $\hat{A}$.

2.5.2 Axiomes

Quand seulement l'ensemble des états du monde est connu, on ne peut mettre plus de poids sur un état plutôt qu'un autre. Par hypothèse, ce type d'information n'est pas disponible. Il est toutefois possible d'accorder plus d'importance à certaines conséquences d'une action. Ainsi un critère qui est souvent utilisé met l'accent sur la pire conséquence d'une action. L'idée est que, comme les attentes du décideur ne peuvent plus être calculées comme précédement, on doit utiliser les informations disponibles dans le cadre pour les évaluer. Ici les informations sont d'abord les conséquences d'une action. Le critère peut alors être le suivant :

1. Pour chaque action, chercher sa conséquence la moins préférée.
2. Comparer les actions comme si pour chacune, c'était cette "pire" conséquence qui se produisait.

Ce critère est peut être le plus sûr que l'on puisse envisager car l'agent cherche, en l'utilisant, à éviter une conséquence catastrophique avant tout. Il n'utilise pas toute l'information disponible. En particulier, dans ce cadre on peut, de la même manière qu'on isole dans ce critère les conséquences d'une seule action (ce qui représente une ligne dans la table de décision), isoler les conséquences de toutes les actions pour un état donné (ce qui correspond à une colonne de la table de décision). On peut aller plus loin et vouloir que dans le critère choisi on puisse intégrer ce fait. On peut demander par exemple que si une action a a une conséquence préférée à celle d'une autre action b pour chaque état du monde alors le critère suggère de préférer a à b.

Concrétiser cette exigence ainsi que d'autres que l'on peut avoir s'est principalement fait dans la littérature en adoptant une démarche axiomatique. Il s'agit de délimiter les exigences (de rationalité) impliquées par le cadre sous forme d'un ensemble d'axiomes pour ensuite caractériser la famille des critères qui satisfont cet ensemble. Les critères ainsi caractérisés tirent leur légitimité (en termes de rationalité) de celle des axiomes. [Maskin 1979] passe en revue les axiomes présentés dans la littérature ainsi que les théorèmes qu'ils permettent de caractériser. On peut réunir les axiomes présentés en fonction de l'aspect de la décision qu'ils traitent.

La première famille d'axiomes est la suivante :

axiome (1) $\forall A_1, A_2 \subseteq \mathcal{A}$, *Si* $A_1 \subseteq A_2$ *et* $A_1 \cap \hat{A}_2 \neq \emptyset$ *alors* $\hat{A}_1 = A_1 \cap \hat{A}_2$.

axiome (2) $\forall A_1, A_2 \subseteq \mathcal{A}$, $a \in A_1 \subseteq A_2 \Rightarrow [a \in \hat{A}_2 \Rightarrow a \in \hat{A}_1]$.

axiome (3) $\forall A_1, A_2 \subseteq \mathcal{A}$, $[a, a_1 \in \hat{A}_1$ *et* $A_1 \subseteq A_2] \Rightarrow [a \in \hat{A}_2 \Rightarrow a_1 \in \hat{A}_2]$.

axiome (4) $\forall A_1, A_2 \subseteq \mathcal{A}$, $a \in A_1 \subseteq A_2 \Rightarrow [a \in \hat{A}_1, a \notin \hat{A}_2 \Rightarrow \hat{A}_2 \setminus A_1 \neq \emptyset]$.

Le premier axiome est l'axiome (A) dans [Arrow & Hurwicz 1977]. Il stipule que si des actions sont désignées comme optimales dans un certain ensemble d'actions et si ultérieurement cet ensemble est réduit mais qu'il contient encore ces actions alors elle restent optimales dans ce nouvel ensemble (et aucune action ne "devient" optimale). Les deux axiomes suivants sont introduits par [Sen 1970], le deuxième est aussi utilisé dans [Milnor 1954].

Dans [Maskin 1979], l'auteur fait remarquer que la satisfaction simultanée des axiomes (2) et (3) équivaut à la satisfaction du premier. La combination des axiomes (2) et (4), plus faible que (1) est utilisée dans [Luce & Raiffa 1957] et dans [Chernoff 1954].

Les axiomes suivants partagent la motivation de tirer les conclusions d'une situation de domination (état par état) comme celle évoquée plus haut.

axiome (5)

$\forall A \subseteq \mathcal{A}$, $a, a_1 \in A$, *si* $a_1 \in \hat{A}$ *et* $a(\omega) \geq a_1(\omega)\ \forall \omega \in \Omega$, *alors* $a \in \hat{A}$.

axiome (6)
$\forall A \subseteq \mathcal{A}$, $a \in \hat{A}, a_1 \in A$, si $a(\omega) > a_1(\omega)\ \forall \omega \in \Omega$, alors $a_1 \notin \hat{A}$.

axiome (7)
$\forall A \subseteq \mathcal{A}$, $a, a_1 \in A$, si $a(\omega) > a_1(\omega)\ \forall \omega \in \Omega$ et $\exists \omega_0 \in \Omega$ tel que $a(\omega_0) > a_1(\omega_0)$, alors $a_1 \notin \hat{A}$.

Ces axiomes visent à faire que la comparaison entre deux actions s'effectue état par état. Cela se rapproche fortement de la préférence faible de Savage (voir section 2.3).

L'axiome suivant spécifie que la manière dont on nomme (range) les différents états du monde n'influe pas sur les décisions de l'agent.

axiome (8) *Supposons qu'il existe une bijection $h : \Omega^1 \to \Omega$. Pour un problème A, définissons $A^1 \subseteq \mathcal{A}$ par*

$$A^1 = \{a^1 | a^1 = a \times h, a \in A\}.$$

Alors $a \in \hat{A}$ ssi $a \times h \in \hat{A}^1$.

L'axiome suivant, présenté dans [Arrow & Hurwicz 1977], rend compte d'un autre type d'indépendance de la décision. Il stipule en effet que diviser un état en plusieurs autres états sans changer les conséquences des actions ne devrait pas changer la décision de l'agent. [Arrow & Hurwicz 1977] définissent d'abord la notion de dérivation d'un problème vers un autre par élimination d'états répétés. Un problème dérive d'un autre si leurs ensembles d'actions sont en bijection par h et que leurs ensembles d'états le sont aussi modulo la relation d'équivalence selon laquelle deux états ω et ω' sont équivalents si chaque action a la même conséquence dans ω que dans ω^1.

axiome (9)
$\forall A, A^1 \subseteq \mathcal{A}$, si A^1 dérive de A via h alors $h(d) \in \hat{A^1}$ ssi $d \in \hat{A}$.

[Maskin 1979] présente un axiome qui stipule qu'ajouter au problème des états dans lesquels toutes les actions mènent à la même conséquence ne change pas la décision de l'agent. Cet axiome rappelle celui de la chose sûre dans la théorie de Savage (section 2.3).

axiome (10) *Notons, pour un problème A, $\Omega(A)$ l'ensemble des mondes possibles qui lui sont associés. Soient $A, A^1 \subseteq \mathcal{A}$ t.q. $\Omega(A^1) \supseteq \Omega(A)$ et une surjection $g : A \to A^1$ t.q. $\forall d \in A, \forall \omega \in \Omega(A)$, $g(d)(\omega) = d(\omega)$. Alors, si pour $d, d^1 \in \hat{(A^1)}, g(d)(\omega) \sim g(d^1)(\omega) \forall \omega \in \omega(A^1) \setminus \omega(A), g(d) \in \hat{A}^1$ ssi $g(d^1) \in \hat{A}$.*

D'autres axiomes s'intéressent à la structure de l'ensemble des conséquences et à l'effet sur la préférence de l'agent de changements effectués sur ces dernières. Les axiomes présentés précédement ne dépendent pas des hypothèses faites sur la topologie de l'ensemble des conséquences. Or cette topologie peut influer de deux manières sur la décision. D'un côté, la propriété de finitude, dénombrabilité ou densité par exemple de cet ensemble permet de donner (ou non) un sens à des propositions du type "une conséquence est beaucoup plus préférée qu'une autre". De l'autre, elle permet de rendre plus ou moins riche la relation de préférence que l'agent pour les conséquences, par exemple une fonction d'utilité $u : \mathcal{C} \to I\!R$. Les deux axiomes suivants introduits par [Milnor 1954] illustrent ce point .

axiome (11) *Considérons une suite (A_i) de problèmes de décision et A un problème. Supposons que pour tout $i, \Omega(A_i) = \Omega(A)$ et $|A_i| = |A| = n$. Ecrivons alors $A = \{a_1, \cdots, a_n\}$ et $A^i = \{a_1^i, \cdots, a_n^i\}$. Alors, si $\forall j, \forall \omega, lim_{i\to\infty} a_j^i(\omega) = a_j(\omega)$, on a $a_j^i \in \hat{A}_i \forall i$ implique $a_j \in \hat{A}$.*

axiome (12) *$\forall A \subseteq \mathcal{A}$, $\forall a_1, a_2, a \in A$, s'il existe une fonction d'utilité sur $\mathcal{C}$ telle que $u \circ a = 1/2 * u \circ a_1 + 1/2 * u \circ a_2$, alors $[a_1, a_2 \in \hat{A}] \Rightarrow a \in \hat{A}$.*

L'axiome (11) transcrit l'idée qu'un petit changement dans l'ensemble de conséquences d'une action ne doit pas modifier dramatiquement la préférence de l'agent pour elle.

Le deuxième axiome (12) stipule que l'ensemble optimal est convexe. Sa justification tient dans le sens que l'on donne à l'ensemble optimal. L'agent peut indifférement choisir n'importe quelle action qui s'y trouve.

Enfin, un certain nombre d'axiomes ont été proposés qui traduisent chacun une disposition de l'agent face à l'incertitude. En voici un exemple.

axiome (13)
$\exists A \subseteq \mathcal{A}$, $a \in \hat{A}, a' \in A \setminus \hat{A}$ t.q. $max_{\omega\in\Omega(A)} a(\omega) < max_{\omega\in\Omega(A)} a'(\omega)$.

Cet axiome stipule que l'agent ne se comporte pas comme s'il était sûr que la meilleure conséquence surviendrait toujours.

Ces axiomes sont combinés les uns aux autres pour caractériser des critères de décision qui alors en tirereont leur légitimité. Bien que chacun ait sa propre justification, [Milnor 1954] a montré qu'on arrivait assez vite à une situation d'incohérence si on veut tous les réunir.

Un théorème important dans ce cadre est celui démontré dans [Arrow & Hurwicz 1977].

Théorème 2.3 *Un critère de décision satisfait les axiomes (1, 5, 8, 9) si et seulement si il existe un préordre total $\succcurlyeq$ sur l'ensemble des paires de conséquences (m, M) avec $M \geqslant m$ tel que les conditions (1) et (2) sont vérifiées :*

*(1) Si $m_1 \geqslant m_2$ et $M_1 \geqslant M_2$
alors $(m_1, M_1) \succcurlyeq (m_2, M_2)$.*

(2) $\forall A \subseteq \mathcal{A}$ avec $A \neq \emptyset$, $\hat{A} = \{a | (min_{\omega\in\Omega(A)} a(\omega), max_{\omega\in\Omega(A)} a((\omega)) \succcurlyeq (min_{\omega\in\Omega(A)} a'(\omega), max_{\omega\in\Omega(A)} a'(\omega)), \forall a' \in A\}$.

Ce théorème caractérise la famille de critères dont la décison induite ne se base que sur les éléments extrémaux des ensembles de conséquences des actions. Remarquons que la comparaison dans le deuxième point du théorème ne se fait pas état par état. C'est bien à chaque fois les éléments minimaux et maximaux de l'ensemble de toutes les conséquences (tous états confondus) qui sont utilisés.

Cette famille contient nombre de critères dont certains ont été caractérisés de manière unique. Ainsi en est-il du critère introduit au début de cette section, le critère min que l'on peut formaliser de la manière suivante :

Définition 2.10 (critère min) *$\forall A \in \mathcal{A}$, soit $a \in A$. $a \in \hat{A}$ ssi $min_{\omega\in\Omega(A)} a(\omega) \geq min_{\omega\in\Omega(A)} a'(\omega), \forall a' \in A$.*

Ce critère de "prudence" est caractérisé par le théorème suivant, établi par [Milnor 1954] :

Théorème 2.4 *Un critère satisfait les axiomes (2,3,6,8,9,10) si et seulement si c'est le critère min.*

Notons que d'après ce qui précède le fait que min satisfait (2,3,6,8,9,10) implique qu'il satisfait (1,5,8,9) et qu'il fait donc partie de la famille caractérisée par le théorème 2.1, ce qui peut se voir par ailleurs en vérifiant qu'il remplit bien les conditions du théorème. [Milnor 1954] caractérise aussi le critère d'Hurwicz [Hurwicz 1951] qui compare entre deux axiomes les sommes pondérées de l'utilité des éléments extrémaux de leurs ensembles de conséquences respectifs.

Définition 2.11 (critère de Hurwicz) *Soit* $\alpha \in [0,1]$. $\forall A \in \mathcal{A}$*, soit* $a \in A$. $a \in \hat{A}$ *ssi* $(\alpha * min_{\omega \in \Omega(A)} a(\omega)) + ((1-\alpha) * max_{\omega \in \Omega(A)} a(\omega)) \geq (\alpha * min_{\omega \in \Omega(A)} a'(\omega) + ((1-\alpha) * max_{\omega \in \Omega(A)} a'(\omega)), \forall a' \in A.$

Ce critère est encore dans la famille caractérisée par le théorème précédent. D'autres travaux ont cherché à caractériser des critères en dehors de cette famille, c'est-à-dire qui ne prennent pas seulement en compte les éléments extrémaux des conséquences d'une action. Ainsi [Milnor 1954] caractérise "Le principe de la raison insuffisante". En l'absence d'information privilégiant, pour une action, un résultat à un autre, les travaux jusqu'ici cités considéraient que les seuls résultats à prendre en compte sont les éléments extrémaux de l'ensemble des résultats. Le principe de la raison insuffisante contient une interprétation différente du manque d'information. Il avance l'idée que si l'on n'a aucune information donnant plus de poids à un résultat qu'à un autre, on peut supposer que tous ont la même chance de se produire. En termes de probabilité, les résultats sont équiprobables.

Définition 2.12 (Principe de la raison insuffisante) *Soient* $A \subseteq \mathcal{A}$ *et* u *une fonction d'utilité* $u : \mathcal{C} \to I\!R$. *Pour* $a \in A$, $a \in \hat{A}$ *ssi* $\sum_{\omega \in \Omega(A)} u(a(\omega)) \geqslant \sum_{\omega \in \Omega(A)} u(a'(\omega)) \forall a' \in A$.

Théorème 2.5 *Un critère satisfait les axiomes (2,3,5,6,8,9) ssi c'est le principe de la raison insuffisante.*

Pourtant, Arrow et Hurwicz réfutent dans leur article [Arrow & Hurwicz 1977] la légitimité du critère de raison insuffisante. Ils avancent pour ce faire l'objection suivante. Supposons que l'agent doive estimer ses attentes par rapport à une certaine action. D'après le principe de la raison insuffisante, il doit le faire en attribuant un poids identique à chacune des conséquences de l'action. D'après le cadre celà signifie qu'il attribue une probabilité identique à chaque état possible du monde. Supposons alors que cet ensemble se divise en deux états s_1 et s_2. Alors $p(s_1) = p(s_2)$. Supposons maintenant que sans rien changer de l'action elle même, on divise s_2 en deux états $s_2^{'}$ et $s_2^{''}$ (si par exemple s_2 est l'état "il fait jour" alors les deux sous-états peuvent être "c'est le matin" et "c'est l'après-midi"). Dans ce cas le principe de la raison insuffisante pousse l'agent à attribuer une probabilité identique à chacun des états du monde qui sont maintenant s_1, $s_2^{'}$ et $s_2^{''}$: $p(s_1) = p(s_2^{'}) = p(s_2^{''})$ (E). Mais alors la probabilité de s_1 serait strictement inférieure à celle de l'union des deux autres mondes c'est-à-dire s_2. En effet $p(s_2) = p(s_2^{'}) + p(s_2^{''})$ et (E) entraîne $p(s_1) < s_2^{'} + s_2^{''}$. Contradiction.

Remarquons, comme le fait d'ailleurs [Arrow & Hurwicz 1977], que pour conclure à une contradiction on a fait appel à la propriété d'addivité des probabilités. Cet argument n'est plus valable dans le cas de mesure non additives de l'attente d'un agent. C'est le cas par exemple dans le cadre de l'approche de Shackle (voir par exemple [Shackle 1955]).

Un autre exemple de critères qui ne satisfont pas le théorème d'Arrow/Hurwicz est celui développé par [Bhattacharyya 2008]. C'est le critère de la médiane.

Définition 2.13 *Soit une action* $a \in \mathcal{A}$*, on suppose que* $|R(a)| = n$ *et on note* $R(a) = \{c_1, \cdots, c_n\}$. *med(a) est défini par* $\{c_{1+(n-1/2)}\}$ *si n est impair et* $\{c_{n/2}, c_{1+(n/2)}\}$ *sinon. Un agent utilise un critère basé sur la médiane si et seulement si* $\forall A \subseteq \mathcal{A}$ *et* $\forall a, b \in A$, *aIb si* $|med(a)| = |med(b)|$ *et qu'il existe une bijection* $h : med(a) \to med(b)$ *telle que* $\forall c \in med(a), c \sim h(c)$.

On peut remarquer que ce critère reprend une idée sous-jacente au théorème d'Arrow/Hurwicz qui est celle de l'existence de points focaux. L'idée est que face à une situation où l'utilité espérée n'est plus d'aucun secours pour évaluer les attentes de l'agent, celui-ci se tourne vers certains points de l'ensemble des conséquences d'une action. C'est un comportement largement observé chez des agents humains. Une expérience classique est la suivante. On informe deux personnes ne se connaissant pas qu'elles ont rendez-vous un jour donné à un certain endroit. On ne leur indique pas l'heure du rendez-vous. On observe alors souvent que les personnes soumises à cette expérience ont tendance à privilégier, sans se concerter, certaines heures particulières du jour telles que midi.

L'agent considère ainsi que son attente est "portée" par ce point focal. Une question qui peut se poser est pourquoi ce point ? Pourquoi le critère précédent se base-t-il sur la médiane et pas sur par exemple la moyenne ? [Stiglitz 2005] donne l'argument suivant en faveur de la médiane.

> "Consider the following thought experiment : If you could choose which country to live in but would be assigned an income randomly from within that country's income distribution, would you choose the 1 country with the highest GDP per capita ? No. More relevant to that decision is median income" [1]."

Dans cette situation, le résultat à propos duquel l'agent est incertain est sa place dans l'échelle des salaires du futur pays. En se basant sur le médian, il sait qu'il y a autant de possibilités de "positions" (symbolisant des citoyens du pays gagnant effectivement cette valeur) au dessus qu'en dessous, ce qui n'est pas du tout garanti par la moyenne qui peut à cause d'une mauvaise répartition des richesses être au dessus ou en dessous du médian.

Dans le cadre de l'incertitude stricte, les critères caractérisés ont en commun d'utiliser le moins d'information possible et d'être ainsi particulièrement adaptés à un cadre informativement faible. Pourtant leurs caractérisations se basent sur la représentation de l'action qui est elle même une information. Les axiomes de dominance présentés plus haut par exemple comparent les axiomes état par état. La question qui se pose alors est "peut-on trouver des critères de décision rationnels dans le cas encore plus informativement pauvre où l'on ne possède même pas cette information ?". La section suivante passe en revue la littérature qui s'y est intéressée.

2.6 Cadre de l'ignorance

2.6.1 Introduction

Tous les modèles décrits jusqu'ici supposaient que l'agent se représentait une action comme une application reliant l'ensemble des mondes à l'ensemble des conséquences. Une action donne une conséquence donnée dans un monde donné (voir [Arrow & Hurwicz 1977] et [Luce & Raiffa 1957]).

Une interprétation pourrait en être l'introduction d'un deuxième agent symbolisant la nature, à la manière de la théorie des jeux, selon le scénario suivant. L'agent effectue l'action a, suite à quoi la nature décide de l'état du monde s, ce qui va donner la conséquence c. Remarquons que les actions sont considérées dans cette interprétation comme intrinsèquement déterministes, l'incertitude étant transportée sur le coup joué par la nature. Et cette interprétation suppose un niveau minimum d'information. En effet, en isolant un coup joué par la nature, on pourrait alors connaître de manière sûre les conséquences de chaque action. C'est là une information dont on peut se servir pour comparer deux actions. Par exemple, si une action a une meilleure conséquence qu'une autre quelque soit l'état considéré, alors on peut intuitivement la préférer car on sait qu'on aura si on l'exécute une meilleure conséquence quoiqu'il arrive. Dans le cas général comparer deux actions revient alors à comparer des manières d'ordonner

1. Supposons que vous soyez un immigrant qui ait à choisir entre plusieurs pays. Choisiriez vous le pays à la moyenne de revenu la plus élevée ou celui au revenu médian le plus élevé ?

des vecteurs d'utilité comme évoqué dans [Nehring & Puppe 1996]. Par exemple, pour comparer deux actions a_1 et a_2, les états du monde possibles étant s_1 et s_2. On compare le vecteur $(a_1(s_1), a_1(s_2)$ au vecteur $(a_2(s_1), a_2(s_2))$. Le résultat de cette comparaison dépend de la représentation choisie des états du monde. Et quand il n'y a pas de subdivision unique et évidente de l'évolution du monde en états, l'argument précédent de dominance état par état n'est clairement plus valable.

Par ailleurs, il arrive souvent que cette modélisation se fasse au prix de l'efficacité du processus de décision. L'ensemble des mondes envisageables peut en effet être arbitrairement grand. Si une action a une certaine conséquence pour une certaine année, on peut aussi représenter cela en posant qu'elle a cette conséquence pour chacune des semaines, ou jours, ou heures qui la composent. On peut donc changer la description des mondes pour en obtenir de plus en plus et ainsi alourdir artificiellement ne serait-ce que la tâche de représenter l'application entre mondes et conséquences. Enfin, dans le contexte de choisir la personne qui partagera notre vie par exemple, la segmentation des mondes possibles semble clairement saugrenue [Gravel *et al.* 2008].

Une autre approche ne comporte pas ces inconvénients. Dans cette approche, une action est directement reliée à un ensemble de conséquences possibles. Si on exécute l'action a, une et une seule des conséquences de l'ensemble associé sera réalisée. Trouver un critère de décision se basant sur cette approche signifie que d'un côté on évite l'écueil d'une représentation trop lourde et que de l'autre, ce critère sera valable dans un domaine plus large que l'approche précedente. En effet, un telle représentation utilise moins d'information que la précédente. Pour le voir examinons la relation entre les deux représentations. On pourrait revenir vers la précédente pour un problème donné sous cette forme en "recréant" des mondes correspondants chacun à un élément du produit cartésien des sous-ensembles de conséquences présents avec l'interprétation suivante : dans un état ainsi obtenu, une action aura comme conséquence l'élément du vecteur du produit cartésien défini qui est associé à ce monde. Par contre le chemin inverse n'est pas toujours possible. Un problème de décision impliquant une représentation explicite des mondes ne correspond pas forcément en termes d'information au modèle action-conséquence intuitif. L'intuition conduirait en effet à associer directement une action à l'ensemble d'arrivée de son application associée. L'exemple suivant montre que l'opération n'est pas aussi directe :

Exemple 2.3 Considérons le problème de décision suivant formé par deux actions a_1 et a_2.

	s_1	s_2
a_1	α_1	α_2
a_2	β_1	β_2

Soit le problème de décision formé par les actions a_1 et a_2 mais dans lequel on sait seulement que a_1 peut avoir comme conséquences α_1 ou α_2 et que a_2 peut avoir comme conséquences β_1 ou β_2. Si maintenant on reprend le chemin inverse de manière intuitive en recréant un monde possible pour chaque combinaison de conséquences des deux actions, on obtient un problème différent :

	s_1	s_2	s_3	s_4
a_1	α_1	α_2	α_1	α_2
a_2	β_1	β_2	β_2	β_1

On n'a en effet aucun moyen de savoir que dans le problème original, α_1 ne se produit pas quand β_2 se produit.

Dans le modèle impliquant une description explicite des états du mondes (que l'on appellera monde-conséquence), il s'agit de comparer des vecteurs d'utilité (ici (α_1, β_1) et (α_2, β_2)) alors que dans le modèle associant directement une action à son ensemble de conséquences (modèle action-conséquence) les vecteurs (α_2, β_1) et (α_1, β_2) sont aussi permis. L'information contenue dans le premier modèle peut être vue comme l'impossibilité de certaines colonnes de la matrice de décision. On peut l'interpréter d'une autre manière. Même si l'agent ne connaît rien de la distribution de probabilités de s_1 et s_2, il peut du moins inférer de cette matrice que si cette distribution p existe, alors $p(\alpha_1) = p(\beta_1)$. Dans les deux interprétations, Le modèle action- conséquence se révèle moins informatif. Dans ce modèle, l'appartenance d'une conséquence à l'ensemble associé à une action donné signifie seulement qu'il existe au moins un état du monde dans lequel l'action conduit à cette action.

Ces différences entre les deux approches, déjà évoquées dans la section précédente, impliquent des différences dans les critères de décision à étudier. L'approche monde- conséquence permet d'utiliser des comparaisons qui se basent sur le nombre d'états dans lesquels se réalise une conséquence. Or, comme le souligne [Barbera *et al.* 2001], le cadre de l'ignorance incite à éviter des critères attachant de l'importance à cette information car ce nombre peut varier moyennant une subdivision arbitraire des états. Il s'en suit que dans cette approche, comparer deux actions peut revenir à comparer leurs ensembles de conséquences possibles. Cela la rapproche alors des travaux effectués par [Barbera *et al.* 1984] sur l'extension d'ordres sur les élements d'un ensemble vers un ordre sur ses sous-ensembles (finis). Un élément est interprété comme une conséquence et un sous-ensemble comme un ensemble de conséquences associé à une action. Cette opération trouve aussi une interprétation dans le champ du choix social (cf. [Klemisch-Ahlert 1993]), chaque sous-ensemble est alors un ensemble d'alternatives (ou menu) parmi lesquelles doit choisir l'agent. Ordonner les sous-ensembles revient à trouver le meilleur menu pour l'agent.

Ce modèle semble particulièrement adapté au cadre de l'ignorance complète. Comme nous l'avons précédement exposé, l'ignorance complète est à l'extrémité du spectre d'information que l'on peut avoir sur le problème de décision. Il est alors cohérent de caractériser un raisonnement se basant sur le minimum d'information possible. Il s'agit d'abord, sous certaines hypothşes sur l'ensemble des éléments, de délimiter un ensemble d'axiomes qui sont autant de "conditions minimales" à un raisonnement rationnel dans le cadre désiré. Reste alors à caractériser un critère ou une classe de critères permettant de classer les sous-ensembles.

2.6.2 Cadre formel

$\mathcal{A}$ est défini comme un ensemble de sous-ensembles de $\mathcal{C}$ (qui est interprété comme l'ensemble de conséquences). On identifie donc une action a à son ensemble de conséquences. Nous ferons référence à l'ensemble de conséquences de l'action a par $R(a)$. Soit $\geq \subseteq \mathcal{C} \times \mathcal{C}$ une relation binaire sur $\mathcal{C}$ qui représente la relation de préférence de l'agent sur les conséquences. $\geq$ est dans la plupart des travaux considérés dans ce cadre une relation d'ordre sauf quand $\mathcal{C}$ est supposé non dénombrable auquel cas $\geq$ est supposée être seulement un préordre.

Un critère de décision est une fonction de choix associant à chaque ensemble non vide A de $\mathcal{A}$ un sous-ensemble $\hat{A}$ de A (formellement, c'est une application). La donnée d'un critère de décision permet de résoudre n'importe quel problème de décision A : les éléments optimaux sont ceux dans $\hat{A}$. Afin de "ranger" les éléments de $\mathcal{A}$, on utilisera une relation $\trianglerighteq \subseteq \mathcal{A} \times \mathcal{A}$. On se servira de cette relation pour représenter le critère de décision recherché avec l'interprétation suivante. Si on utilise un critère $\trianglerighteq$ alors pour un ensemble d'actions A, l'ensemble $\hat{A}$ est l'ensemble des éléments de A non dominés pour la relation binaire $\trianglerighteq$.

Définition 2.14 *Soit A un problème de décision. Soit $\trianglerighteq \subseteq \mathcal{A} \times \mathcal{A}$. Soit a une action dans A. Alors $a \in \hat{A}$ si et seulement si $\nexists a' \in A$ telle que $a' \triangleright a$.*

Parmi les axiomes et résultats présentés, certains supposent de plus que l'ensemble des conséquences $\mathcal{C}$ soit un espace de Haussdorf arc-connexe et séparable. Un espace topologique X est arc-connexe si et seulement si $\forall x, y \in X$, il existe un chemin continu $\omega : [0,1] \mapsto X$ tel que $\omega(0) = x$ et $\omega(1) = y$. Un espace topologique X est séparable si et seulement si il contient un sous-ensemble dense dénombrable. Un ensemble topologique est dit de Haussdorff si et seulement pour tout couple d'éléments distincts de X, les deux éléments qui le composent possèdent chacun un voisinage disjoint.

Dans ce cas on peut représenter la relation de préférence sur les conséquences par une fonction d'utilité, i.e., une fonction continue $u : \mathcal{C} \mapsto I\!R$ telle que $\forall c, c' \in \mathcal{C}\ c \geq c'$ si et seulement si $u(c) \geq u(c')$. Cela permet de définir la pire conséquence et la meilleure conséquences d'un ensemble de la manière suivante. Pour $a \in \mathcal{A}$, $U_M(a) = max_{x \in R(a)} u(x)$ et $U_m(a) = min_{x \in R(a)} u(x)$.

2.6.3 Axiomes

Les différents travaux qui se sont intéressés à la décision dans le cadre de l'ignorance complète ont présenté des systèmes d'axiomes dont la satisfaction permet de caractériser une famille de critères de décision, parfois réduite à un élément. Dans cette section, nous présentons ces axiomes en les regroupant par nature et motivation.

Le premier axiome, commun à la plupart des travaux, est l'axiome d'extension. Il n'est cependant pas toujours évoqué explicitement car il souvent impliqué par d'autres axiomes.

axiome (Extension (E)) $\forall x, y \in \mathcal{C}$, $x \geq y \Leftrightarrow \{x\} \trianglerighteq \{y\}$.

La justification de cet axiome est assez directe. On demande que la relation obtenue sur les ensembles de conséquences ne contredise pas les préférences du décideur sur les conséquences elles-mêmes. On identifie pour cela de manière naturelle une conséquence à l'action dont elle est le seul élément de l'ensemble associé.

Les axiomes qui suivent partagent la même forme. Ils comprennent une situation initiale constituée par deux actions ayant une certaine configuration l'une par rapport à l'autre. Ils tirent ensuite une conclusion sur leur préférences relatives après l'ajout de conséquences supplémentaires. Cet ajout peut avoir deux interprétations, soit les actions dans la situation initiale sont différentes de celles dans la situation finale, et dans ce cas les axiomes ont pour sens que les préférences de l'agent sur certaines actions induisent des préférences sur d'autres actions. Soit le décideur a en premier lieu une vue partielle des conséquences des actions. Il reçoit alors un complément d'information sur les mêmes actions. L'axiome indique donc l'effet de cette information complémentaire sur la préférence entre les actions disponibles. Dans la suite nous ferons donc indifférement appel à l'une ou l'autre de ces interprétations.

axiome (Indépendance faible (I)) *$\forall a \in \mathcal{A}$ et $\forall x, y \in \mathcal{C}$ t.q. $\{x, y\} \cap R(a) = \emptyset$. Soient $a', b' \in \mathcal{A}$ telles que $R(a') = R(a) \cup \{x\}$ et $R(b') = R(a) \cup \{y\}$. Si $x \geq y$ alors $a' \trianglerighteq b'$.*

L'axiome (I) est présenté par [Puppe 1995] comme une condition "d'indépendence du contexte". La préférence qu'on a entre deux alternatives reste valable quel que soit le contexte dans lequel on la teste. Si on remplace dans l'ensemble d'alternatives a une alternative x par une autre y telle que l'on préfère x à y, alors on préfère a au nouvel ensemble obtenu. Cet axiome illustre le rapport complémentaire entre l'approche normative et l'approche descriptive de la théorie de la décision. On y part d'une situation particulière dans laquelle on exprime le jugement qui nous semble légitime (au moyen par exemple d'observations de comportements humains). Toutefois, malgré cette conclusion bénigne, la famille de critères de décision caractérisée est si large qu'elle admet des critères contre-intuitifs si l'on considère un décideur maximisateur d'utilité. Elle contient par exemple le critère de cardinal qui séléctionne les actions qui ont le plus grand nombre de conséquences.

Définition 2.15 (cardinal) *Considérons deux actions $a, b \in A$. $a \trianglerighteq_{card} b$ ssi $|R(a)| \geq |R(b)|$.*

Ce critère conduit à préférer une action a menant à deux conséquences c_1, c_2 à une autre b menant à une seule conséquence c_3 et ce même si c_3 est largement préférée à c_1 et à c_2. D'après l'interprétation adoptée, une fois les actions exécutées, l'action a aura pour conséquence c_1 ou c_2. Dans les deux cas, a mène donc à un résultat moins préféré que celui de b. Cet exemple éclaire le fait que la plupart des axiomes présentés dans la littérature ne suffisent pas par eux-mêmes pour délimiter une famille de critères raisonnables pour la décision sous ignorance complète. C'est la conjonction de plusieurs axiomes qui délimite souvent une famille satisfaisante.

L'axiome suivant, présenté dans [Puppe 1995], ressemble fort à l'axiome (I).

axiome (J) *$\forall a \in \mathcal{A}$ et $\forall x, y \in \mathcal{C}$, Soient $a', a'' \in \mathcal{A}$ telles que $R(a') = R(a) \cup \{x\}$ et $R(a'') = R(a) \cup \{y\}$. $x \geq y \Rightarrow a' \trianglerighteq a''$.*

La différence entre (I) et (J) réside dans la condition $\{x, y\} \cap R(a) = \emptyset$ qui n'est plus demandée dans (J). Cela rend (J) plus fort que (I) puisque pour vérifier (I) on ne doit comparer que des ensembles "de même taille", pour (J) on doit en vérifier plus.

axiome (Independance Forte (IF)) *$\forall a, b \in \mathcal{A}$, $\forall x \in \mathcal{C} \setminus (R(a) \cup R(b))$, Soient $a', b' \in \mathcal{A}$ telles que $R(a') = R(a) \cup \{x\}$ et $R(b') = R(b) \cup \{x\}$. Si $a \trianglerighteq b$ alors $a' \trianglerighteq b'$.*

L'axiome (IF) énonce que l'ajout d'une conséquence commune aux ensembles de conséquences respectifs de deux actions ne doit pas changer la préférence du décideur entre ces deux actions. C'est un axiome largement utilisé dans la littérature à la fois de la décision sous ignorance complète (voir e.g. [Bossert 1997]) et de la liberté de choix (voir e.g. [Bossert *et al.* 1994], [Klemisch-Ahlert 1993]), ce qui le fait paraitre comme un axiome naturel. En réalité, cet axiome est strictement plus fort que (I). Pour le montrer, [Nehring & Puppe 1996] montre d'abord que (IF) implique (I), puis exhibe un exemple de critère satisfaisant (I) mais non (IF). C'est la cas par exemple du critère de Hurwicz (voir [Hurwicz 1951]).

Les tentatives de caractérisation utilisent souvent d'autres versions de cet axiome qui apportent différentes nuances dans les hypothèses de départ ou des conclusions qui sont tirées. L'axiome (IND), introduit dans [Kannai & Peleg 1984], en est un xemple.

axiome (Independance (IND)) *$\forall a, b \in \mathcal{A}$, $\forall x \in \mathcal{C} \setminus (R(a) \cup R(b))$, Soient $a', b' \in \mathcal{A}$ telles que $R(a') = R(a) \cup \{x\}$ et $R(b') = R(b) \cup \{x\}$. Si $a \triangleright b$ alors $a' \trianglerighteq b'$.*

(IND) requiert que l'ajout d'une conséquence commune ne change pas la préférence d'un décideur à propos de deux actions si initialement cette préférence est stricte. Plus précisément, cet ajout peut rendre les deux actions obtenues équivalentes mais pas inverser la préférence. Il est aisé de voir que (IF) implique (IND).

axiome (Independance Stricte (IS)) *$\forall a, b \in \mathcal{A}$, $\forall x \in \mathcal{C} \setminus (R(a) \cup R(b))$, Soient $a', b' \in \mathcal{A}$ telles que $R(a') = R(a) \cup \{x\}$ et $R(b') = R(b) \cup \{x\}$. Si $a \triangleright b$ alors $a' \triangleright b'$.*

L'axiome d'Indépendance stricte (IS) [Barbera & Pattanaik 1984] est une autre version où les préférences sont strictes dans les deux phases initiale et finale. En d'autres termes, si un décideur a une préférence stricte entre deux actions, alors l'ajout d'une conséquence commune non seulement ne change pas la préférence mais la garde en plus stricte. On part ici de la même hypothèse que dans (IND) et on en renforce la conclusion. On peut voir alors que (IS) implique (IND). (IS) a pour ambition de caractériser une famille de critères plus fins que ceux caractérisés par les autres axiomes car on peut voir intuitivement que les seconds peuvent induire que deux actions vérifiant la même condition initiale, c'est-à-dire une préférence stricte, peuvent se retrouver dans la même classe d'équivalence après l'ajout d'une conséquence commune alors que ce n'est clairement pas le cas pour les premiers.

axiome (Independance Etendue (IE)) $\forall a, b \in \mathcal{A}$, $\forall C \subseteq \mathcal{C} \setminus (R(a) \cup R(b))$, *Soient* $a', b' \in \mathcal{A}$ *telles que* $R(a') = R(a) \cup C$ *et* $R(b') = R(b) \cup C$. *Si* $a \rhd b$ *alors* $a' \trianglerighteq b'$.

L'axiome d'Indépendance Etendue (IE) [Barbera *et al.* 1984] généralise ces précédents axiomes et énonce plus directement leur motivation commune. Les deux premiers stipulent que l'ajout d'une conséquence commune n'influe pas dramatiquement sur la préférence du décideur. Le dernier (IS) va plus loin et stipule qu'en fait cet ajout n'influe pas du tout dessus. (IE) stipule directement que l'ajout de n'importe quel nombre de conséquences commune ne change pas de manière dramatique les préférences. Tous ces axiomes véhiculent donc la même idée. Dans la comparaison de deux actions disponibles, seules comptent les conséquences non communes. On peut faire comme si les conséquences communes n'existaient pas.

L'axiome d'Indépendance restreinte (IR) ([Barbera *et al.* 2004]) partage cette motivation mais restreint son application au cas d'actions dont l'ensemble de conséquences est un couple.

- Indépendance Restreinte (IR) : $\forall x, y, z, w, v \in \mathcal{C}$ telles que $x \geq y$ et $z \geq w$.
  - Si $\{x, y\} \rhd \{z, w\}$ et $x \neq v$ et $v > z$ alors $\{max(v, x), y\} \trianglerighteq \{v, w\}$.
  - Si $\{x, y\} \rhd \{z, w\}$ et $w \neq v$ et $y > v$ alors $\{x, v\} \trianglerighteq \{z, min(v, w)\}$.

Selon la position de la conséquence commune ajouté, (IR) tire alors une conclusion sur la préférence entre deux autres actions dont les ensembles associés comportent encore deux conséquences. A travers cet axiome, on aperçoit la famille de critères qui est le plus souvent visée par les axiomes de la littérature. Ce sont les critères qui se focalisent sur la pire et la meilleure conséquence de l'ensemble associé à une action. L'idée est que ces éléments sont les seuls points dont la prise en compte reste compatible avec le cadre de l'ignorance complète.

Nous avons vu dans les chapitres précédents que la décision à propos d'actions futures se base dans le cas général sur une évaluation des attentes associées à son exécution. Dans le cadre de l'ignorance complète, aucune information probabiliste n'est disponible. Dans ce cas, ne sachant pas laquelle des conséquences se réalisera, tout ce que l'on peut dire de l'attente liée à l'action est qu'elle n'est pas pire que la moins préférée des conséquences et pas meilleure que la plus préférée des conséquences. En particulier, si on suppose que l'on est dans un cas où les probabilités des conséquences ont un sens mais ne sont pas disponibles, dire que l'on ne connaît pas la distribution de probabilités des conséquences peut être interprété comme "toutes les distributions sont admissibles". On peut alors, pour n'importe quel point entre le minimum et le maximum de l'ensemble des conséquences, trouver une distribution telle que ce point égale l'utilité espérée induite par cette distribution. On ne peut donc pas baser le choix sur l'un de ces points.

axiome (Independance Intermédiaire (II)) $\forall a, b \in \mathcal{A}$, $\forall x, y \in \mathcal{C}$ *telles que* $x > z > y \forall z \in R(a) \cup R(b)$, *Soient* $a', b' \in \mathcal{A}$ *telles que* $R(a') = R(a) \cup \{x, y\}$ *et* $R(b') = R(b) \cup \{x, y\}$. *Si* $a \trianglerighteq b$ *alors* $a' \trianglerighteq b'$.

L'axiome Indépendance Intermédiaire (II) [Barbera *et al.* 2004], abondonne cette focalisation sur les conséquences extrémales. Il garde la même idée initiale, à savoir que l'ajout de deux conséquences communes n'influe pas sur le choix entre deux actions, mais en contraignant les actions à être "de part et d'autre" des conséquences des deux actions, il montre clairement que les conséquences extrémales ne définissent pas la préférence (car les ensembles de conséquences extrémales associées aux actions finalement obtenues sont identiques).

Enfin l'axiome suivant introduit dans [Gravel *et al.* 2008] sert dans un ensemble d'axiomes qui abondonne complètement l'idée des points focaux.

axiome (Indépendance Forte restreinte) *Soient $a, b \in \mathcal{A}$ tel que $|R(a)| = |R(b)|$ et $C \subseteq \mathcal{C}$. $a \trianglerighteq b \Rightarrow R(a) \cup C \trianglerighteq R(b) \cup C$.*

Cet axiome requiert que la préférence entre deux ensembles de conséquences de même cardinal est indépendant des conséquences communes entre ces deux ensembles. C'est une version plus forte de l'axiome (RI) .

Les axiomes qui suivent traitent de situations, plus particulières que celles dont partent les axiomes précédents, desquelles elles dégagent des conclusions sur les préférences de l'agent.

Le premier axiome, Principe de Gärdenfors (PG), a été présenté par [Kannai & Peleg 1984] dans le cadre fini et est appelé Dominance dans [Barbera *et al.* 2004].

axiome (Dominance/ Principe de Gärdenfors (PG)) *$\forall a \in \mathcal{A}$, $x \in \mathcal{C}$. Soit $a' \in \mathcal{A}$ telle que $R(a') = R(a) \cup \{x\}$,*
- *Si $(x > y \; \forall y \in R(a))$ alors $a' \triangleright a$.*
- *Si $(y > x \; \forall y \in R(a))$ alors $a \triangleright a'$.*

L'axiome (PG) stipule que si on ajoute à l'ensemble de conséquences d'une action a une conséquence strictement préférée à toutes les autres conséquences, alors on accroît, de manière stricte, la préférence de l'agent pour cette action.

axiome (Dominance Simple/ Simple Monotonie (DS)) *$\forall x, y \in \mathcal{C}$, Si $x > y$ alors $\{x\} > \{x, y\} > \{y\}$.*

[Bossert *et al.* 2000] présente l'axiome de dominance simple (DS) qui est une restriction de (PG) au cas d'actions ayant initialement une seule conséquence. [Bossert 2001] note que si on suppose que la relation de préférence de l'agent sur les conséquences est reflexive et antisymétrique et que la relation visée sur les actions est réflexive et quasi-transitive, alors (DS) implique (E).

[Barbera *et al.* 2004] montre le résultat suivant qui fait d'ailleurs le lien entre l'axiome (DS) et les axiomes (IND) et (IE). Il montre que sous réserve que (DS) soit vérifié, requérir une condition sur l'ajout d'un ensemble de conséquences communes revient à la requérir sur l'ajout d'une seule conséquence commune.

Théorème 2.6

Soit R un ordre sur $\mathcal{C}$ et $\trianglerighteq$ un critère qui soit une relation binaire reflexive et transitive (un pré-ordre) sur $\mathcal{A}$ qui satisfait la Dominance Simple. $\trianglerighteq$ satisfait Indépendance si et seulement si $\trianglerighteq$ satisfait l'axiome d'indépendance étendue.

Le troisième axiome Monotonie de Gärdenfors (MG) [Barbera *et al.* 2004] joue pour les deux précédents le même rôle de généralisation que (IE) par rapport aux axiomes de la même famille.

axiome (Monotonie de Gärdenfors (MG)) *$\forall a, b \in \mathcal{A}$ t.q. $R(a) \cap R(b) = \emptyset$. Considérons $c \in \mathcal{A}$ telle que $R(c) = R(a) \cup R(b)$. Si $x > y \;\; \forall x \in R(a) \;$ et $\; \forall y \in R(b)$, alors $a \trianglerighteq c \trianglerighteq b$.*

Enfin l'axiome (D) [Barbera *et al.* 2004] traduit, en se restreignant au cas de deux conséquences par action que ce sont les deux éléments extrémaux de l'ensemble des conséquences qui déterminent la préférence de l'agent pour une action.

axiome (D) $\forall x, y, x', y' \in \mathcal{C}$, *si* $x > x'$ *et* $y > y'$ *alors* $\{x, y\} \rhd \{x', y'\}$.

L'axiome (C") présenté par [Nehring & Puppe 1996] repose sur la relation de préférence de l'agent. Il stipule que l'ajout d'une conséquence équivalente par cette relation à l'une des conséquences déjà connues d'une action ne change pas la préférence que l'agent a pour cette action. On peut remarquer que cet axiome est porteur de sens seulement si la relation de préférence sur les conséquences est au plus un préordre. Si cette relation est un ordre, le résultat devient trivial.

axiome (C") $\forall a \in \mathcal{A}$ *et* $\forall y \in \mathcal{C}$, *si* $\exists x \in \mathcal{C}$ *t.q.* $y \sim x$ *alors, si on considère* $b \in \mathcal{A}$ *t. q.* $R(b) = R(a) \cup \{y\}$, $a \diamond b$.

Les axiomes suivants essayent de modéliser l'attitude du décideur face à l'incertitude. En d'autres termes, le décideur préfère-t-il qu'il y ait plus ou moins d'incertitude sur le résultat d'une action ? En ce sens ces axiomes sont plus descriptifs que les précédents.

axiome (Recherche d'Incertitude (RI)) $\forall a, b \in \mathcal{A}$ *t.q.* $R(a) \cap R(b) = \emptyset, \forall x \in \mathcal{C} \setminus (R(a) \cup R(b))$, *si* $y > x > z$ $\forall y \in R(a)$ *et* $\forall z \in R(b)$. *Considérons* $c \in \mathcal{C}$ *t.q.* $R(c) = R(a) \cup R(b)$ *et* $d \in \mathcal{C}$ *t.q.* $R(d) = R(a) \cup R(b) \cup \{x\}$. *On a alors* $c \rhd \{x\}$ *et* $d \rhd \{x\}$.

Le premier axiome de Recherche d'Incertitude (RI) traduit une attitude assez inattendue. Le décideur préfèrera une situation risquée à une situation sûre. La recherche délibérée de l'incertitude peut se rencontrer par exemple chez des entrepreneurs qui préfèrent se lancer dans un projet à l'issue incertaine à la solution de garder pour eux le capital de départ. Un autre exemple en est le comportement observé en situation désespérée ou le statu quo a une conséquence (certaine) si indésirable à l'agent que celui-ci s'engagera dans une action incertaine même si son issue favorable tient du miracle.

axiome (Aversion à l'Incertitude (AI)) $\forall a, b \in \mathcal{A}$ *t.q.* $R(a) \cap R(b) = \emptyset, \forall x \in \mathcal{C} \setminus (R(a) \cup R(b))$, *si* $y > x > z$ $\forall y \in R(a)$ *et* $\forall z \in R(b)$. *Considérons* $c \in \mathcal{C}$ *t.q.* $R(c) = R(a) \cup R(b)$ *et* $d \in \mathcal{C}$ *t.q.* $R(d) = R(a) \cup R(b) \cup \{x\}$. *On a alors* $\{x\} \rhd c$ *et* $\{x\} \rhd d$.

L'axiome d'Aversion à l'Incertitude (AI) décrit une situation plus familière en cas de choix importants pour l'agent. Le décideur y est averse à l'incertitude, c'est-à-dire qu'il préfère être le plus sûr possible du résultat, ce qui se traduit par le moins de conséquences possibles.

axiome (Neutralité face à l'Incertitude (NI)) $\forall a, b \in \mathcal{A}$ *t.q.* $R(a) \cap R(b) = \emptyset, \forall x \in \mathcal{C} \setminus (R(a) \cup R(b))$, *si* $a > x > b$ $\forall a \in R(a)$ *et* $\forall b \in R(b)$. *Considérons* $c \in \mathcal{C}$ *t.q.* $R(c) = R(a) \cup R(b)$ *et* $d \in \mathcal{C}$ *t.q.* $R(d) = R(a) \cup R(b) \cup \{x\}$. *On a alors* $c \diamond \{x\}$ *et* $d \diamond \{x\}$.

L'axiome de Neutralité face à l'Incertitude (NI) décrit le comportement non couvert par les deux autres et qui est tout aussi imaginable ; le décideur y est indifférent à l'incertitude. L'explication de son introduction est aussi descriptive. L'attitude la plus souvent observée dans le cadre économique est celle d'un agent neutre à l'incertitude si elle est limitée et averse à celle-ci si elle est grande.

axiome (Simple Aversion à l'Incertitude (SAI)) $\forall x, y, z \in \mathcal{C}$, *si* $x > y > z$ *alors* $\{y\} > \{x, z\}$.

axiome (Simple Recherche d'Incertitude (SRI)) *$\forall x, y, z \in \mathcal{C}$, si $x > y > z$ alors $\{x, z\} > \{y\}$.*

Les deux axiomes de Simple Recherche d'Incertitude (SRI) et Simple Aversion à l'Incertitude (SAI) sont des restrictions des attitudes précédentes au cas de conséquences ayant une ou deux conséquences.

axiome (Monotonie (M)) *$\forall a, b \in \mathcal{A}$, si $R(a) \supseteq R(b)$ alors $a \rhd b$.*

L'axiome de Monotonie (M) est à rapprocher des premiers axiomes précédement présentés. Il décrit en effet, de manière plus large que ces derniers, une certaine attitude du décideur face à l'incertitude. Il dit en substance qu'enlever des alternatives dans un ensemble le rend plus attractif car moins incertain. Cette axiome est vérifié par des critères qui génèrent des décisions incompatibles avec le comportement d'un décideur maximiseur d'utilité. Ainsi est il satisfait par le critère de cardinal précédement présenté. Ce critère guidera par exemple un décideur, à qui on proposerait de choisir entre recevoir 1000 euros avec certitude et recevoir 5 euros ou 10 euros de manière équiprobable, à préférer le second choix. On peut toutefois noter, comme le fait Nehring, que si on associe cet axiome à d'autres (tels que (C)+(I)) qui caractérisent des critères ne prenant en compte que les éléments extrémaux des ensembles traités conduit à la caractérisation unique du bien connu critère optimiste max.

Définition 2.16 (max) *Considérons deux actions a_1, a_2.*

$$a_1 \trianglerighteq_{max} a_2 \;\; ssi \;\; maxR(a_1) \geq maxR(a_2).$$

axiome (Simple Monotonie par le Haut) *$\forall x, y, z \in \mathcal{C}$, si $x > y > z$ alors $\{x, z\} > \{y, z\}$.*

axiome (Simple Monotonie par le Bas) *$\forall x, y, z \in \mathcal{C}$, si $x > y > z$ alors $\{x, z\} > \{y, z\}$.*

axiome (Consistance monotone) *Soit $a, b \in \mathcal{A}$ et $a' \in \mathcal{A}$ telle que $R(a') = R(a) \cup R(b)$. Si $a \trianglerighteq b$ alors $a' \trianglerighteq b$.*

axiome (Robustesse) *Soit $a, b, c \in \mathcal{A}$ et $b' \in \mathcal{A}$ telle que $R(b') = R(b) \cup R(c)$. Si $a \trianglerighteq b$ et $b \trianglerighteq c$ alors $a \trianglerighteq b'$.*

Ces axiomes donnent d'autres versions de l'axiome de monotonie. Les axiomes suivants ne rentrent pas particulièrement dans l'une des familles précédentes.

axiome (Continuité (C)) *Soient $a, b \in \mathcal{A}$ et $(a_n)_{n \in I\!N}$ une suite dans $\mathcal{A}$ convergente vers a,*
Si $\forall n \in I\!N, a_n \trianglerighteq b$ alors $a \trianglerighteq b$. De même si $\forall n \in I\!N, b \trianglerighteq a_n$ alors $b \trianglerighteq a$.

Cet axiome d'aspect purement mathématique exprime en fait, sous des conditions topologiques imposées au domaine de définition des alternatives, qu'un changement minime de l'ensemble des alternatives ne devrait pas causer de changement majeur dans le jugement du décideur.

L'axiome suivant, plus normatif, stipule qu'un simple changement de nom des alternatives ne devrait pas changer le choix du décideur. Seule importe la position relative des alternatives, c'est-à-dire leur impact sur le décideur.

axiome (Neutralité (NT)) *Soient* $a, b \in \mathcal{A}$ *et une bijection* $g : R(a) \cup R(b) \to \mathcal{C}$*. Soient alors* $a', b' \in \mathcal{A}$ *telles que* $R(a') = \{c' \in \mathcal{C} \mid \exists c \in R(a) \text{ t.q. } g(c) = c'\}$ *et* $R(b') = \{c' \in \mathcal{C} \mid \exists c \in R(b) \text{ t.q. } g(c) = c'\}$.
$[x \geq y \Leftrightarrow g(x) \geq g(y) \;\; et \;\; y \geq x \Leftrightarrow g(y) \geq g(x) \;\; \forall x \in R(a), y \in R(b)]$
$\Rightarrow \;\; [a \trianglerighteq b \Leftrightarrow a' \trianglerighteq b' \;\; et \;\; b \trianglerighteq a \Leftrightarrow b' \trianglerighteq a']$.

Ces axiomes combinés permettent de caractériser les critères de décision qui les vérifient. La première série de résultats montre les limites d'associations d'axiomes présentés précédement. Le théorème suivant est peut être le plus connu.

Théorème 2.7 *Supposons que* $|\mathcal{C}| \geqslant 6$ *et R une relation d'ordre sur* $\mathcal{C}$*. Il n'existe aucun ordre* $\trianglerighteq$ *sur* $\mathcal{A}$ *satisfaisant les axiomes Dominance et Indépendance.*

Ce théorème fût établi par [Kannai & Peleg 1984]. Il y est dit que l'on ne peut concilier deux exigences raisonnables qui trouvent chacune sa justification de son côté comme on l'a vu plus haut. [Nehring & Puppe 1996] explicite la nature de ce résultat. Il montre en effet, moyennant l'adjonction de l'axiome (C) dans le cas d'un ensemble de conséquences ayant une structure d'espace de Haussdorf, que la satisfaction simultannée de (IS) et du premier point de (PG) caractérise un unique critère de décision différent de celui caractérisé par la satisfaction simultanée de (SI) et du deuxième point de (PG). Les deux théorèmes suivants montrent d'autres résultats d'impossibilité [Barbera *et al.* 2004].

Théorème 2.8 *Supposons que* $|\mathcal{C}| \geqslant 3$ *et R une relation d'ordre sur* $\mathcal{C}$*. Il n'existe aucune relation binaire* $\trianglerighteq$ *sur* $\mathcal{A}$ *satisfaisant les axiomes Simple Dominance et Indépendance Stricte.*

Théorème 2.9 *Supposons que* $|\mathcal{C}| \geqslant 4$ *et R une relation d'ordre sur* $\mathcal{C}$*. Il n'existe aucun ordre* $\trianglerighteq$ *sur* $\mathcal{A}$ *satisfaisant les axiomes Dominance, Indépendance et Neutralité.*

Les théorèmes de la série suivante énoncent un résultat commun à la quasi-totalité des travaux qui se sont intéressés à la décision dans le cadre non probabiliste, y compris [Hurwicz 1951] qui montre un résultat similaire. Ce résultat constitue une marche intermédiaire vers des résultats qui définissent les familles de critères de décision caractérisées. Ces résultats explicitent les points de l'ensemble de conséquences qui peuvent potentiellement compter pour un critère compatible avec les axiomes. Il éclairent par là le fait que les axiomes de rationalité communément utilisés impliquent l'existence de points focaux dans les ensembles de conséquences.

Théorème 2.10 *Soit R un ordre sur* $\mathcal{C}$ *et* $\trianglerighteq$ *un critère qui soit une relation binaire réflexive et transitive sur* $\mathcal{A}$*. Si* $\trianglerighteq$ *satisfait les axiomes Dominance (GP) et Indépendance (IND) alors* $\forall a \in \mathcal{A}$*,* $a \diamond \{minR(a), maxR(a)\}$.

Ce premier théorème a été montré par [Kannai & Peleg 1984]. Il stipule que si un critère de décision se basant sur un préordre (les meilleures actions sont celles préférées par le préordre) satisfait les axiomes (GP) et (IF) alors toute action est équivalente par ce critère à l'action dont l'ensemble de conséquences comporte seulement le couple formé par les pire et meilleure conséquences.

Théorème 2.11 *Soit R un ordre sur* $\mathcal{C}$ *et* $\trianglerighteq$ *un critère qui soit une relation binaire réflexive et quasi-transitive sur* $\mathcal{A}$*. Si* $\trianglerighteq$ *satisfait les axiomes Dominance (GP) et Indépendance (IND) alors* $\forall a \in \mathcal{A}$*,* $a \diamond \{minR(a), maxR(a)\}$.

Théorème 2.12 *Soit R un ordre sur $\mathcal{C}$ et $\trianglerighteq$ un critère qui soit une relation binaire réflexive et transitive (un pré-ordre) sur $\mathcal{A}$. Si $\trianglerighteq$ satisfait les axiomes Dominance Simple et Indépendance alors $\forall a \in \mathcal{A}$, $a \diamond \{minR(a), maxR(a)\}$.*

Les deux théorèmes suivants présentent d'autres versions du premier qui obtiennent la même conclusion en affaiblissant les prémisses. Le premier d'entre eux, montré par [Barbera *et al.* 2004], le fait en stipulant que l'on peut affaiblir l'hypothèse faite sur le critère visé en exigeant seulement que le critère soit basé sur une relation réflexive et quasi-transitive. L'autre, prouvé dans [Barbera *et al.* 2004, Bossert *et al.* 2000] le fait en affaiblissant l'axiome (GP) vers l'axiome (DS).

Pour présenter le théorème qui va suivre, on aura besoin de quelques axiomes particuliers présentés dans [Barbera *et al.* 2004]. Ils sont particuliers car ils s'appliquent seulement quand l'ensemble des conséquences d'une action a une forme particulière.

axiome (Extension de l'Equivalence aux cardinaux Pairs) $\forall a \in \mathcal{A}$ *tel que* $|R(a)|$ *est pair,* $\forall x, y \in \mathcal{C}$,

$$[A \cup \{x\} \diamond \{x\} \;\; et \;\; A \cup \{y\} \diamond \{y\}] \Rightarrow A \cup \{x, y\} \diamond \{x, y\}$$

axiome (Neutralité pour les cardinaux Impairs) $\forall a, b \in \mathcal{A}$ *tels que* $|R(a)|$ *et* $|R(b)|$ *sont impairs. Pour toute application injective* $\phi : R(a) \cup R(b) \to \mathcal{C}$,

$$[x \geq y \Leftrightarrow \phi(x) \geq \phi(y) \;\; et \;\; y \geq x \Leftrightarrow \phi(y) \geq \phi(x)] \forall x \in R(a) \;\; et \;\; y \in R(b)$$

$$\Rightarrow (a \trianglerighteq b \Leftrightarrow \phi(a) \trianglerighteq \phi(b) \;\; et \;\; b \trianglerighteq a \Leftrightarrow \phi(b) \trianglerighteq \phi(a))$$

axiome (Dualité pour les Cardinaux Impairs) $\forall a, b \in \mathcal{A}$ *tels que* $|R(a)|$ *et* $|R(b)|$ *sont impairs. Pour toute application injective* $\phi : R(a) \cup R(b) \to \mathcal{C}$ $(R(a) \cup R(b))$,

$$[x \geq y \Leftrightarrow \phi(x) \geq \phi(y) \;\; et \;\; y \geq x \Leftrightarrow \phi(y) \geq \phi(x)] \forall x \in R(a) \;\; et \;\; y \in R(b)$$

$$\Rightarrow (a \trianglerighteq b \Leftrightarrow \phi(a) \trianglerighteq \phi(b) \;\; et \;\; b \trianglerighteq a \Leftrightarrow \phi(b) \trianglerighteq \phi(a))$$

Théorème 2.13 *Supposons que $|\mathcal{C}| \geqslant 6$. Soit R une relation d'ordre sur $\mathcal{C}$ et $\trianglerighteq$ un préordre total sur $\mathcal{A}$ satsfaisant l'axiome d'Extension. $\trianglerighteq$ satisfait les axiomes d'Indépendance Intermédiaire, d'Extension de l'Equivalence aux cardinaux Pairs, de Neutralité pour les cardinaux Impairs et de la Dualité pour les Cardinaux Impairs si et seulement si $\forall a \in \mathcal{A}$, $a \diamond medR(a)$.*

Ce dernier théorème, montré par [Nitzan & Pattanaik 1984], tourne autour de l'axiome (II) et concrétise l'intuition évoquée lors de son introduction. Moyennant plusieurs autres axiomes, ce théorème montre que (II) donne un rôle important à un point focal différent en général des éléments extrémaux des ensembles de conséquences. Cet élément est le médian de l'ensemble des conséquences.

Dans la suite, on présente trois résultats de caractérisation qui spécifient la forme des critères qui satisfont un ensemble d'axiomes donné. Ces trois résultats vont donc plus loin que les précédents. Et de fait, ils en font intervenir certains dans leurs preuves.

Théorème 2.14 *Soit $\mathcal{A}$ un ensemble d'actions et soit $\mathcal{C}$ l'ensemble des conséquences associées. On suppose que $\mathcal{C}$ est un espace de Haussdorf. Un critère de décision $\trianglerighteq$ sur les problèmes $A \subseteq \mathcal{A}$ satisfait les axiomes (C) et (I) si et seulement si il existe un pré-ordre $\succcurlyeq$ sur l'ensemble des paires de conséquences (m, M) avec $M \geqslant m$ tel que les conditions (1) et (2) soient vérifiées :*

(1) Si $m_1 \geq m_2$ alors $(m_1, M) \succcurlyeq (m_2, M)$.

(2) Si $M_1 \geq M_2$ alors $(m, M_1) \succcurlyeq (m, M_2)$.

Et tel que la relation $\trianglerighteq$ est définie par $\forall a, b \in \mathcal{A}$ $a \trianglerighteq b$ si et seulement si $(U_M(a), U_m(a)) \succcurlyeq (U_M(b), U_m(b))$.

Le premier, présenté dans [Nehring & Puppe 1996], s'applique dans le cadre particulier d'un ensemble de conséquences $\mathcal{C}$ qui est un espace de Haussdorf. Cette structure est riche (et restrictive). Les auteurs justifient l'étude spécifique d'une telle topologie par le rôle central joué par les préférences continues dans le champ économique. On peut voir que ces deux théorèmes caractérisent dans ce cas l'ensemble des critères basé sur un préordre entre les paires de conséquences. La paire retenue pour chaque action étant formée par le pire et le meilleur élément de son ensemble de conséquences. La famille caractérisée dans le second théorème est incluse dans celle retenue par le premier.

Théorème 2.15 *Soit $\mathcal{A}$ un ensemble d'actions et soit $\mathcal{C}$ l'ensemble des conséquences associées. On suppose que $\mathcal{C}$ est un espace de Haussdorf. Un critère $\trianglerighteq$ sur les problèmes $A \subseteq \mathcal{A}$ satisfait les axiomes (C) et (IF) si et seulement si il existe un préordre $\succcurlyeq$ sur l'ensemble des paires de conséquences (m, M) avec $M \geqslant m$ tel que les conditions (1), (2), (3) et (4) soient vérifiées :*

(1) Si $m_1 \geq m_2$ alors $(m_1, M) \succcurlyeq (m_2, M)$.

(2) Si $M_1 \geq M_2$ alors $(m, M_1) \succcurlyeq (m, M_2)$.

(3) Si $(m, M) \succcurlyeq (m', M')$ alors $(max(m, t), M) \succcurlyeq (max(m', t), M')$.

(4) Si $(m, M) \succcurlyeq (m', M')$ alors $(m, min(M, t)) \succcurlyeq (m', min(M', t))$.

Et tel que la relation $\trianglerighteq$ est définie par $\forall a, b \in \mathcal{A}$ $a \trianglerighteq b$ si et seulement si $(U_M(a), U_m(a)) \succcurlyeq (U_M(b), U_m(b))$.

Le dernier théorème, quant à lui, s'intéresse au cas général pour trouver un résultat ayant à première vue la même allure. Il énonce qu'un critère de décision qui satisfait les axiomes (DS) et (IND) dépend d'abord seulement des éléments extrémaux des ensembles de conséquences et ensuite répond à une condition particulière. Pour spécifier cette forme, les auteurs se concentrent sur les actions ayant un ensemble de conséquences qui se réduit à deux conséquences. Cela est justifié par le résultat présenté plus haut qui montre dans ce cas l'équivalence entre une action et celle dont l'ensemble de conséquences compte deux éléments : les éléments extrémaux de l'ensemble associé à la première action. La condition est alors que le critère satisfasse les axiomes (DS) et (RI) sur ces actions particulières. C'est le sens de l'expression "basé sur le maximin" dans le résultat.

Théorème 2.16 *Soit $\mathcal{A}$ un ensemble d'actions et soit $\mathcal{C}$ l'ensemble des conséquences associées. Soit $\trianglerighteq$ un critère telle que la relation de préférence associée soit une relation d'ordre sur $\mathcal{A}$. $\trianglerighteq$ satisfait les axiomes Dominance Simple et Indépendance si et seulement si $\trianglerighteq$ est basé sur le maximin.*

Même si cette condition fait apparaître le résultat comme similaire au précédent. Elle ne caractérise pas la même famille de critères comme le montre l'exemple du critère min qui est dans la première sans appartenir à la deuxième.

Définition 2.17 (min) *Considérons deux actions a_1, a_2.*

$$a_1 \trianglerighteq_{min} a_2 \;\; ssi \;\; minR(a_1) \geq minR(a_2).$$

Le critère min est habituellement connu sous le nom de critère de Wald [Wald 1950].

Les résultats suivants s'attachent à caractériser non pas une famille mais un unique critère à chaque fois. Avant de les présenter, introduisons d'abord les critères en question.

Définition 2.18 (minmax - maxmin) *Considérons deux actions* a_1, a_2.

$$a_1 \trianglerighteq_{minmax} a_2 \text{ ssi } a_1 \trianglerighteq_{lex(\trianglerighteq_{min}, \trianglerighteq_{max})} a_2.$$

$$a_1 \trianglerighteq_{maxmin} a_2 \text{ ssi } a_1 \trianglerighteq_{lex(\trianglerighteq_{max}, \trianglerighteq_{min})} a_2.$$

On a alors les deux théorèmes suivants :

Théorème 2.17 *Soit R un ordre sur* $\mathcal{C}$ *et* $\trianglerighteq$ *un critère qui est une relation binaire réflexive, transitive et complète (un préordre complet) sur* $\mathcal{A}$. $\trianglerighteq$ *satisfait la Dominance Simple, la Simple Aversion à l'Incertitude, la Simple Monotonie par le Haut, la Consistance Monotone et la Robustesse si et seulement si* $\trianglerighteq = \trianglerighteq_{minmax}$.

Théorème 2.18

Soit R un ordre sur $\mathcal{C}$ *et* $\trianglerighteq$ *un critère qui est une relation binaire réflexive, transitive et complète (un préordre complet) sur* $\mathcal{A}$. $\trianglerighteq$ *satisfait la Dominance Simple, la Simple Attrait à l'Incertitude, la Simple Monotonie par le Bas, la Consitance Monotone et la Robustesse si et seulement si* $\trianglerighteq = \trianglerighteq_{maxmin}$.

Enfin nous présentons ici une charactérisation axiomatique d'un critère qui traduit une vision différente de l'ignorance. Dans [Gravel *et al.* 2008], les auteurs présentent en effet un critère particulier qui donne la même importance à chacune des conséquences. L'interprétation qui le sous-tend est la suivante. Un décideur qui ignore le mécanisme par lequel les conséquences sont choisies et qui est seulement capable d'identifier l'ensemble des conséquences d'une action, n'a aucune raison légitime de penser qu'une action se réalisera plutôt qu'une autre. On reconnaît dans cet argument le principe de la raison insuffisante.

axiome (Moyennisation (M)) $\forall c_1, c_2 \in \mathcal{A}$ *tel que* $R(c_1) \cap R(c_2) = \emptyset$. *Soit* $c \in \mathcal{C}$ *telle que* $R(c) = R(c_1) \cup R(c_2)$, *on a* $c_1 \trianglerighteq c_2 \Rightarrow c_1 \trianglerighteq c \Rightarrow c \trianglerighteq c_2 \Rightarrow c_1 \trianglerighteq c_2$.

Cet axiome requiert que quand on ajoute à une action des conséquences meilleures que les siennes, on augmente son attractivité pour le décideur, et inversement si les conséquences sont pires que les siennes.

axiome (Atténuation (At)) $\forall c_1, c_2, c_3 \in \mathcal{C}$ *tel que* $c_1 \diamond c_2$, $R(c_1) \cap R(c_3) = R(c_2) \cap R(c_3) = \emptyset$ *et* $|R(c_1)| > |R(c_2)|$. *Soient* $\forall c_{13}, c_{23} \in \mathcal{C}$ *telles que* $R(c_{13}) = R(c_1) \cup R(c_3)$ *et* $R(c_{23}) = R(c_2) \cup R(c_3)$. *On a* $c_3 \triangleright c_1 \Rightarrow c_{13} \triangleright c_{23}$ *et* $c_1 \triangleright c_3 \Rightarrow c_{23} \triangleright c_{13}$.

Cet axiome traduit une aversion à l'incertitude un peu particulière. Il spécifie qu'un agent peut être indifférent entre deux actions dont l'ensemble de résultats de l'une est plus incertain (contient plus d'éléments) que celui de l'autre. Pourtant si on rajoute à ses ensembles un ensemble de conséquences commun, alors la balance penchera d'un côté ou de l'autre. Si cet ensemble est préféré aux deux autres, c'est l'ensemble obtenu le plus sûr qui sera préféré et inversement. Notons que cet axiome donne un rôle important aux conséquences communes entre deux actions, ce qui n'est pas le cas de la plupart des axiomes jusqu'ici présentés.

axiome (Combinaison de Paires (CP)) *$\forall a, b, c, d \in \mathcal{C}$ distinctes. Si $\exists e, f \in \mathcal{C}$ (non nécessairement distinctes) telles que $\{a, b\} \diamond \{e\}$ et $\{c, d\} \diamond \{f\}$ alors $\{a, b, c, d\} \diamond \{e, f\}$.*

Cet axiome requiert une cohérence dans les équivalences entre les ensembles de conséquences.

axiome (Equivalent Certain (EC)) *$\forall C \subseteq \mathcal{C}$, $\exists a \in \mathcal{C}$ tel que $\{a\} \diamond C$.*

Cet axiome présente une condition forte sur l'ensemble des conséquences $\mathcal{C}$. Il est à rapprocher des hypothèses faites pour la définition de l'utilité espérée. On peut en effet voir l'utilité espérée comme une certaine utilité équivalente à un ensemble d'utilités auquel on adjoint une certaine distribution de probabilités. On voit donc intuitivement que cet axiome s'écarte de l'interprétation, précédemment présentée, qui voyait l'ignorance comme "toute distribution de probabilité est admissible".

axiome (Richesse (R)) *$\forall a, b \in \mathcal{A}$. Si $\exists c_1, c_2 \in \mathcal{C}$ telles que $R(a) \cup \{c_1\} \trianglerighteq b$ et $b \trianglerighteq R(a) \cup \{c_1\}$ alors $\exists c \in \mathcal{C}$ telle que $R(a) \cup \{c\} \diamond b$.*

axiome (Archimedien (Ar)) *Soit $(c_i)_{i=1,2,\cdots}$ une suite d'éléments de $\mathcal{C}$. Soient alors deux conséquences distinctes a et b telles que $\{a\}$, $\{b\}$ et $\forall i = 1, 2, \cdots, \{c_i, a\} \diamond \{b, c_{i+1}\}$. Alors si $\exists x, y \in \mathcal{C}$ telles que $\forall i = 1, 2, \cdots, \{x\} \triangleright \{c_i\} \triangleright \{y\}$ alors la suite $(c_i)_{i=1,2,\cdots}$ est finie.*

Ces deux théorèmes décrivent une structure assez contraignante de l'ensemble de conséquences considéré. En particulier, l'axiome (R) demande que l'on puisse toujours rajouter des éléments à l'ensemble de conséquences (bien que cela soit conditionné par l'existence d'une situation particulière sur les préférences). Enfin l'axiome (Ar) est un axiome souvent utilisé dans la théorie de la mesure. Il est trivialement vérifié si l'ensemble des conséquences est fini.

Ces axiomes réunis déterminent une famille de critères de décision : les critères d'utilité espérée uniforme définis par

Définition 2.19 (UEU) *Un critère de décision est dit d'utilité espérée uniforme s'il existe une application $u : \mathcal{C} \to I\!R$ telle que $\forall a, b \in \mathcal{A}$,*

$$a \trianglerighteq b \text{ si et seulement si } \sum_{c \in R(a)} (u(c)/|R(a)|) \geq \sum_{c \in R(b)} (u(c)/|R(b)|)(1)$$

Un critère satisfaisant cette définition peut être "pensé comme résultant de comparaisons des utilités espérées des conséquences [des actions] pour une certaine fonction d'utilité sous l'hypothèse que le décideur assigne à chaque conséquence [d'une action] une égale probabilité d'occurrence"
[Gravel *et al.* 2008].

Cette famille est caractérisée au moyen du théorème suivant.

Théorème 2.19 *Soit $\trianglerighteq$ une relation réflexive, transitive et totale sur $\mathcal{C}$ qui satisfait les axiomes (EC, R, Ar). Alors $\trianglerighteq$ satisfait (M, IRF, At, PC) si et seulement si c'est un critère d'utilité espérée uniforme. De plus, la fonction u dans la défintion (UEU) est unique à une transformation affine positive près.*

2.7 Conclusion

Dans ce chapitre, nous avons passé en revue l'état de l'art de la théorie de la décision sous incertitude en prenant comme parti pris de s'éloigner progressivement du cadre probabiliste en appauvrissant à chaque fois l'information disponible au décideur et en exposant alors les critères de décision rationnels, dans le cadre obtenu, sur lesquels il peut s'appuyer. Le fil conducteur entre ces différents cadres est que le décideur se sert toujours de l'information qui lui est disponible (distribution de probabilités sur les états du monde, informations qualitatives sur les états du monde, préférences sur les conséquences des actions) pour formuler ses attentes par rapports aux différentes actions qu'il peut choisir. La plupart des travaux qui se sont intéressés à caractériser ces critères de décision ont adopté une démarche axiomatique traduisant les exigences que l'on peut avoir à propos d'un critère de décision pour qu'il soit rationnel dans un cadre donné. Le critère tire alors sa légitimité de celle des axiomes qu'il satisfait.

Nous avons vu que même dans des cadres extrêmement pauvres en information comme celui de l'incertitude stricte ou de l'ignorance, le décideur peut utiliser des critères moins prudents que le critère min qui juge une action par sa pire conséquence. Ces critères reviennent la plupart du temps à faire émerger des points particuliers, les points focaux, dans les ensembles de conséquences associées aux actions.

Chapitre 3

Planification sous incertitude

3.1 Introduction

Le but des algorithmes de planification est de "trouver une suite d'actions que doit exécuter un certain agent pour atteindre ses buts" ([Blythe 1999]). La suite d'actions en question est ce qui forme ce que nous appellons un plan. La notion de plan a été décrite d'une manière générale dans [Schank & Abelson 1975] comme " l'ensemble des choix d'une personne quand elle entreprend de réaliser un but". Deux variables importantes pour rendre la définition précédente opérationnelle sont les manières dont on modélise chacun des concepts : action et but. Ces deux variables font partie de ce qui est plus généralement un modèle pour le problème de planification. Celui-ci englobe tous les aspects influant sur la décision finale de l'agent. [Munson 1971] introduit la notion de modèle de la manière suivante : "dans toute formulation d'un problème d'intelligence artificielle, il existe une structure d'information dans l'ordinateur qui constitue un modèle pour le domaine du problème". Un modèle est donc ce qui existe entre l'agent (l'ordinateur) et l'environnement dans lequel il agit.

Spécifier seulement une modélisation des buts est inefficace car la cohérence du modèle impose qu'à celle-là corresponde une modélisation à la fois des états du monde dans lequel évolue l'agent et de l'état de satisfaction de l'agent. La détermination du modèle ou plutôt le choix d'un modèle dépend de plusieurs impératifs au premier rang desquels se trouve la nécessité de se rapprocher le plus possible de la réalité du monde décrit. Plusieurs modèles peuvent correspondre à un même problème. Un modèle ne doit avant tout pas laisser de côté un aspect du monde qui est pertinent. Chacun de ses aspects doit se trouver représenté par un des concepts du modèle. D'un autre côté, la finalité du modèle est opérationnelle en ce sens que l'agent doit pouvoir le manipuler pour trouver une solution, soit en l'occurence un plan, à son problème. Il s'en suit que, parmi les modèles couvrants les différents aspects du monde à traiter, le meilleur pour l'agent est sans doute le plus simple à manipuler. Une façon de comparer deux modèles est de comparer leurs manipulations dans les algorithmes de planification. Cette manipulation passe forcément par le choix d'un langage de représentation.

Wittgenstein [Wittgenstein 1986] soutient qu'il existe une relation très étroite entre la pensée et le langage. La pensée, affirme-t-il, n'est rien d'autre qu'une proposition linguistique douée de sens. Le langage prend d'autant plus d'importance lorsqu'il s'agit de déléguer complètement le raisonnement à une machine (c'est bien le sens de l'automatisation du raisonnement qui est une des grandes caractéristiques de l'intelligence artificielle). Un langage permet en effet de ramener le lien entre les concepts à un schéma syntaxique autrement plus aisé à manipuler de manière automatique que des concepts abstraits. Le choix d'un langage qui permet à l'agent de raisonner sur les concepts qui constituent le modèle est crucial car un mauvais choix de langage peut nous faire perdre tout le bénéfice de l'attention portée au choix du modèle. Deux critères permettent de juger de l'opprtunité du choix d'un langage de représenta-

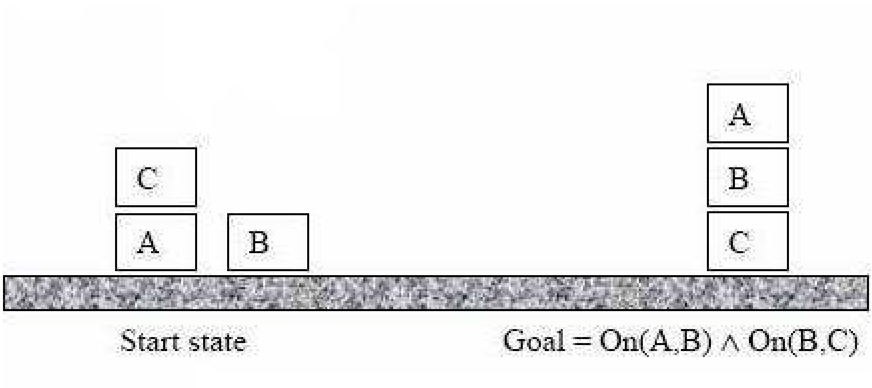


FIGURE 3.1 – Monde des blocs

tion. Il s'agit d'un côté d'exprimer le plus précisément les concepts du modèle. C'est ce que l'on appelle l'expressivité du langage. De l'autre, on cherche à réduire la complexité du problème de planification une fois exprimé dans ce langage.

Une fois que le monde et les actions ont été modélisés, et que l'agent est capable de raisonner sur eux, sa tâche est de résoudre le problème de planification. Dans le cadre de la planification classique (le plus simple), un tel problème est constitué de trois parties :

- des propriétés initiales qui décrivent l'état dans lequel le monde est avant de l'exécution d'un plan.
- les buts à atteindre .
- les actions disponibles dans le but de modifier les relations entre les objets de l'environnement.

Résoudre un tel problème consiste à générer un ensemble ordonné d'actions (un plan) permettant à un agent de faire évoluer le monde de l'état initial vers un état dans lequel les propriétés de ses buts sont vérifiées. Plusieurs méthodes ont été proposées pour générer des plans.

Avant d'aller plus loin dans la discussion, et pour rendre celle-ci moins abstraite, présentons d'abord un exemple typique de problème de planification qui devra être soluble par tout modèle de planification : le monde des blocs. La version standard peut être trouvée dans [Nilsson 1982].

Exemple 3.1 Dans cet exemple, plusieurs blocs sont posés sur une table et un robot agit sur eux au moyen d'un bras mobile.

- les blocs peuvent être soit posés sur la table soit posés sur un autre cube soit tenus par le bras du robot. Une dernière caractéristique à déterminer est s'ils sont ou non libres (un cube est libre si aucun autre n'est posé dessus).
- Le bras peut être vide ou non-vide.
- Le robot peut exécuter les actions suivantes :
  - prendre un bloc sur la table.
  - poser un bloc sur la table.
  - dépiler un bloc (le prendre au dessus d'un autre bloc).
  - empiler un bloc (le poser au dessus d'un autre bloc).
- Le but de l'agent est une configuration particulière des cubes.

Nous nous concentrerons dans ce chapitre sur la présentation des modèles de planification existants. La littérature concernant les aspects de représentation ainsi que les aspects de génération de plans ne sera pas abordée.

3.2 Planification classique

3.2.1 Modèle : concepts

L'intérêt s'est d'abord historiquement porté sur un modèle simple convenant à monde simplifié comme celui de l'exemple 3.1.

Reprenons d'abord la définition d'un problème de planification présentée dans l'introduction.

Définition 3.1 *Un problème de planification est défini par la donnée d'un triplet <S,A,G> où :*
- *S est l'ensemble des états possibles du monde.*
- *A est l'ensemble des actions disponibles pour l'agent.*
- *G est l'ensemble des buts de l'agent.*

Dans le cadre de la planification classique, l'élément de base du modèle est la notion d'état du monde. Il décrit en totalité tout ce qui est pertinent pour l'agent dans le monde qui l'entoure. On suppose que les états du monde sont connus à l'agent et que de plus celui-ci connaît parfaitement l'état dans lequel il se trouve à tout moment. Une action est définie par ses effets sur le monde dans lequel se trouve l'agent. Vue la manière dont on a défini un état du monde, les changements occasionnés par une action définissent un autre état du monde (ou plusieurs). En planification classique, les actions sont déterministes. Cela signifie que l'agent connaît exactement l'état du monde que produit une action sachant l'état dans lequel elle est exécutée. On peut alors définir formellement une action comme une application de l'ensemble des états du monde dans lui même. Reste à définir le concept de but. Un but est défini comme un état (ou plusieurs états) que l'agent préfère à tous les autres. En planification classique, la relation de préférence évoquée est une relation dichotomique. Il n'y a pas de gradation dans cette relation. Un état est soit un but soit il ne l'est pas. On dit aussi que l'agent a une structure binaire de buts.

Le cadre de la planification classique a été l'objet de nombreux travaux car étant assez restrictif et bien délimité, il permet de se concentrer sur l'objectif d'obtenir des méthodes de résolution efficaces. Le choix d'une méthode de résolution peut être considéré comme faisant partie du modèle lui même. Il s'agit en effet de choisir un modèle pour le processus de planification. Ce choix dépend de la manière dont on voit les concepts de but, d'action, etc. Il dépend donc du modèle que l'on vient d'évoquer. Il est par ailleurs très lié au langage de représentation retenu.

Il existe deux principales manières de considérer le processus de planification. Dans la première, on considère que les plans possibles sont une donnée du problème, quitte à ce que ce soient tous les plans que l'on peut obtenir à partir des actions disponibles. Ils s'agit alors de choisir un plan dans cet ensemble. On compare les plans dans leur globalité. Dans le cas de la planification classique, on voit tout de suite un critère permettant de discriminer les plans. Il suffit en effet de regarder l'état du monde après exécution d'un plan. Cet état est parfaitement connu d'après les hypothèses du modèle. L'agent préfèrera alors un plan qui mène à un état but à un plan qui mène à un état non-but.

La deuxième est de considérer que seules sont données les actions disponibles à l'agent. Le processus consiste alors à construire un plan de manière incrémentale en rajoutant successivement des actions à la suite déjà choisie. L'agent se trouve donc face à des points de choix multiples alors que dans l'approche précédente le choix est unique. Ces choix peuvent être pris de manière définitive avant l'exécution du plan. En effet, comme nous l'avons signalé plus haut, le modèle de la planification classique suppose que les actions sont déterministes. L'agent "sait" donc dans quel état sera le monde après l'exécution de d'une action. Il prend alors le choix suivant exactement comme s'il se tenait dans l'état correspondant.

Les deux approches ne sont pas si éloignées l'une de l'autre. En effet, choisir une action à un moment de la construction d'un plan peut être vu comme le fait de réduire l'espace de choix parmi les plans disponibles (aux plans commençant par la suite d'actions choisies jusque-là). Ce qui détermine le choix de l'une ou de l'autre de ces approches, c'est avant tout le modèle qui peut par exemple permettre l'observation des résultats des actions (ce qui n'est pas le cas en planification classique) et de ce fait la réparation de plans (pouvoir corriger des choix qui se sont révélés désavantageux par l'ajout d'une action appropriée). Dans cet exemple, il semble clair que le planificateur devrait privilégier l'approche de construire ses plans à partir des actions. Par ailleurs, les aspects de représentation et de complexité peuvent aussi peser aussi sur ce choix.

Après avoir ainsi défini le modèle, on doit pour le rendre opérationnel choisir un langage de représentation approprié. La dynamique d'utilisation du modèle consiste alors en une méthode de résolution qui fait interagir les concepts du modèle à travers leurs représentation dans le langage choisi.

3.2.2 Généralisations du cadre

Dès que l'on s'écarte du cadre idéalisé de la planification classique en changeant l'hypothèse que l'on y fait, les modèles précédents montrent leurs limites. Par exemple, quand des logiciels de planification sont utilisés pour produire des suites d'actions dans le monde réel, comme des protocoles de traitement médicaux ou des déplacements de robots, ils doivent prendre en compte le fait que leur environnement ne correspond pas à ce monde parfait sur plusieurs points au moins. Le premier est que les actions peuvent avoir plusieurs effets possibles, à cause de la présence d'évènements exogènes. Il se peut qu'une action entreprise ne réussisse simplement pas, i.e., ne donne pas une conséquence "normale". Il peut aussi arriver que deux plans menant à un état but aient des coûts différents, l'agent voulant alors intuitivement choisir celui de coût minimum. De telles considérations poussent à rechercher une autre théorie sur laquelle s'appuyer pour bâtir un modèle plus pertinent.

L'essor de la décision bayésienne et des approches markoviennes font de la théorie de la décision le cadre idéal pour développer des techniques de planification sous information incomplète. Toutefois, la problématique de planification n'épouse pas parfaitement celle de la théorie de la décision. On peut voir la planification comme une décision. Mais d'autres défis se présentent à l'agent. La manière dont les actions prennent place par rapport au temps d'exécution et entre elles détermine aussi des approches différentes du problème de planification. Dans STRIPS par exemple, les actions sont instantanées et deux actions ne peuvent se produire en même temps. Toutefois, un examen des préconditions et effets des actions peut révéler que certaines actions sont indépendantes les unes des autres (l'ordre de leur exécution n'influe pas sur le résultat final) ou que l'agent planificateur préfère retarder le plus possible le choix de la sous-suite qu'elles devront former. Dans le but de prendre en compte ces cas, il existe des planificateurs non linéaires (e.g. [Sacerdoti 1977], [Tate 1977], [Wilkins 1988]). Par exemple, dans [Sacerdoti 1975], l'auteur décrit une structure d'information particulière : un réseau procédural, structure qui sera étendue et raffinée dans [Drummond 1985]. Cette structure représente un plan comme un préordre sur les actions par rapport au temps. Le plan est alors non linéaire. Les planificateurs non linéaires permettent d'envisager le cas où les actions se déroulent de manière parallèle. Se pose alors la question de savoir si l'effet de deux actions parallèles est simplement l'union de leurs effets prises chacune à part ou si un nouvel effet naît de cette simultanéité. Des extensions aux travaux précédement cités permettent de traiter ce cas pour des actions indépendantes. Toutes les suites qu'on peut former avec de telles actions donnent le même résultat que la suite où elle s'exécutent simultanément.

Une telle représentation des buts a permis le développement de techniques de planification heuristique. Mais elle ne permet pas d'incorporer la notion de satisfaction partielle. Par exemple, un agent dont le but est de collecter 10 euros peut vouloir encoder dans son raisonnement sur les plans qu'il préfère avoir 9 euros à ne rien avoir du tout. D'un autre côté, la représentation binaire des buts ne permet pas de prendre en compte pour l'évaluation d'un plan le coût de son exécution qui peut être un paramètre important pour un agent réel aux ressources limitées. Certains travaux ont essayé d'y parvenir en étendant le cadre de la planification classique, soit en prenant en compte le coût des actions [Haddawy & Hanks 1998], soit en intégrant le facteur temps dans le processus de planification

[Smith & Weld 1999], ou encore en gérant des ressources limitées [Koehler 1998]. [Wellman & Doyle 2000] notent que ces approches, qui consistent finalement à combiner de manière *ad hoc* plusieurs paramètres influant sur la satisfaction de l'agent,

> "[$\cdots$] fournissent des performances raisonnables dans des systèmes particuliers de planification [mais] manquent typiquement d'une quelconque signification et ainsi ne fournissent ni une base pour évaluer leur cohérence et leur caractère approprié pour d'autres domaines de problèmes, ni une justification pour les opérations d'inférence et choix exécutés par les architectures de planification qui les sous-tendent."

La théorie de la décision procure un cadre théorique riche permettant la construction de modèles d'utilité qui capturent précisément la notion de satisfaction partielle et de balance entre coût et profit. Voir un problème de planification comme un problème de décision ne pose pas de difficulté. Suivant le cadre exposé à la section 2.1, le choix d'un plan est l'objet de la décision. La conséquence de cette décision est l'état auquel aboutit l'exécution du plan choisi. L'utilité d'un plan est alors la satisfaction qu'il tire du monde obtenu. Cette interprétation correspond au cadre de la planification classique. Le fait de représenter la satisfaction de l'agent par une fonction d'utilité permet d'envisager des buts d'importances différentes pour l'agent. De plus, la manière dont on définit l'utilité permet d'incorporer facilement le coût d'un plan dans la satisfaction liée à son exécution. [Wellman & Doyle 2000] proposent une approche permettant la construction modulaire d'une fonction d'utilité pour l'agent comme la composition de fonctions de préférence (portant sur différentes composantes de sa satisfaction). La fonction d'utilité choisie doit incorporer (comme l'indique sa définition dans le cadre de la théorie de la décision) tous les aspects auquels l'agent attache une importance. Deux modalités d'utilisation de la notion d'utilité s'offrent alors à l'agent. La première permet de rester dans l'esprit de la planification comme génération de plans par composition d'opérateurs (actions). Il s'agit dans cette approche de guider le processus de planification en limitant la construction de plans au moyen de contraintes d'utilité comme le suggèrent [Feldman & Sproull 1977]. La deuxième approche consiste à considérer la planification comme le fait de comparer les plans possibles et de choisir celui induisant la plus grande utilité. [Wellman 1990] propose aussi d'introduire une notion de dominance entre classes de plans au moyen d'une mesure d'utilité qualitative.

Par ailleurs, on se trouve souvent dans des environnements où l'information que possède l'agent est incomplète. Une des raisons les plus communes est le manque de ressources de stockage ou de calcul qui permettraient à l'agent de tirer profit de ces informations. Mais il se peut aussi que ces informations lui soient volontairement cachées par d'autres agents autour de lui. [Dean & Kanazawa 1989] appliquent des techniques de la théorie de la décision à des problèmes de planification. La plupart utilisent une variante de modèle d'utilité utilisé dans les processus de Markov [Howard 1960]. Ce modèle comporte une fonction de récompense associée à certains états et une fonction de coût associée aux actions. La valeur d'un plan est la valeur associée à un état à laquelle on retranche le coût des actions du plan.

3.3 Modèles pour la planification sous incertitude

En planification classique, on suppose que l'environnement ne recèle pas d'évènements exogènes. Evaluer un plan est alors une opération directe. Selon nos hypothèses, on calcule l'utilité de la suite d'actions qui est l'utilité du monde dans lequel on arrive après exécution ou la somme des utilités des actions. Si maintenant on relaxe cette hypothèse pour traiter par exemple la planification des déplacements d'un robot physique ([Krithivasan *et al.* 1999]), il nous faut encore une fois nous baser sur la théorie de la décision. Elle offre en effet les modèles permettant de prendre en compte l'information incomplète dans la détermination de la qualité d'un plan. Le modèle utilisé dépend de la source de l'imperfection de l'information à laquelle se trouve confronté l'agent. D'abord les actions peuvent avoir des effets aléatoires.

L'exemple le plus frappant est celui d'une action "lancer un dé". Un autre exemple plus pratique pourrait être dans un système d'organisation de voyage l'action "prendre un train" dans le cas ou on suppose que l'on a 85% de chances d'arriver à destination à l'heure, 10% pourcent de chances d'arriver en retard et 5% de chances d'avoir un accident. Un moyen de traiter l'incertitude dans le monde est de construire des plans qui incluent en même temps des actions d'observation et d'autres actions dont l'exécution dépend du résultat de ces observations. Par exemple si on désire acheter une voiture, on peut planifier de demander à un mécanicien d'examiner une voiture particulière et de l'acheter seulement si son rapport indique qu'elle est en bon état. Quand les résultats des actions peuvent être observés, un plan peut être construit qui inclut des actions d'observation. Cela peut donner lieu soit à une solution où on alterne des phases de planification et d'observation, soit une solution dont le plan résultat comporte des branches conditionnelles. Dans ce dernier cas, on doit spécifier toutes les branches possibles pour tous les résultats d'une action. La plupart des systèmes distinguent entre les variables observables et non observables du domaine. Un plan STRIPS classique ne suppose aucune observabilité puisque la suite d'actions est exécutée de manière aveugle. Grâce à son expressivité, le calcul de situations permet de gérer une situation initiale incomplète. Toutefois, si quelques propriétés sont laissées indéfinies, cela signifie que plusieurs états peuvent satisfaire la description de ce qui est connu, et le planificateur cherche un plan qui les satisfait tous. Cela est bien sûr impossible si un plan reste une suite d'actions. Par exemple, si on ne sait pas exactement où se trouve un bloc, plusieurs alternatives doivent être considérées pour manipuler ce bloc, ce qui mène à plusieurs plans différents.

Dans les deux cas evoqués précédement, une imperfection de l'information peut avoir pour source la défaillance ou le résultat partiel des actions d'observation. Cette source peut être assimilée à la précédente si on considère que le succès et la défaillance sont deux résultats possibles d'une action particulière qu'est l'action d'observation. Elle en diffère pourtant en ceci. Si l'imperfection de l'information porte sur un plan futur dans le premier cas, elle porte sur l'histoire de l'exécution dans le deuxième. Une dernière source d'imperfection est la présence d'autres agents agissant dans l'environnement. Dans ce cas, et même si les actions qui forment un plan sont déterministes, le fait que les actions d'un autre agent peuvent venir s'intercaler de manière incontrôlée entre deux actions d'un plan fait que le résultat de celui-ci devient inconnu.

Parmi les approches qui étendent la planification classique, [Wellman 1990] étend le formalisme STRIPS pour la planification sous incertitude. [Boutilier *et al.* 1999] étend aussi le cadre de la planification classique en y utilisant le biais suivant. Il nous faut considérer l'interaction avec les autres agents comme complètement aléatoire ou en d'autres termes assimilables aux actions d'une nature sans but personnel. Si maintenant on veut prendre en compte le fait que les agents ont des buts et sont proactifs (agissent pour atteindre leurs buts), on doit accorder sa place à la coordination entre agents dans la détermination du devenir de l'environnement. Ce cadre fera l'objet de la prochaine section.

Planification robuste

Le problème de planification robuste (voir e.g. [Cimatti & Roveri 1999, Cimatti *et al.* 2004], [Palacios & Geffner 2006]) consiste à générer des plans dans le cas d'une incertitude sur l'état initial du monde ainsi que sur l'effet des actions. L'agent ne peut généralement pas non plus observer ce résultat pendant l'exécution du plan. La motivation est alors d'atteindre un état but dans ces conditions. En d'autres termes, le plan doit réussir quel que soit l'état initial et quels que soient les résultats des différentes actions entreprises. Ce modèle est utile car aucune observation n'est possible.

Définition 3.2 ***(planification robuste)***

- *Une* action non déterministe α *sur un ensemble fini et non vide* S *d'états est une application de* S *dans* $2^S \setminus \{\emptyset\}$.

– *Un* plan non déterministe π *sur un ensemble* A *d'actions non déterministes est une suite finie d'éléments de* A*.*
– *Une* trajectoire *pour un plan non déterministe* $\pi = \alpha_1. \cdots . \alpha_n$ *étant donné un état initial* $s_0 \in S$ *est une suite d'états* $s_0, \ldots, s_{n+1}$ *telle que pour tout* $i \in 0 \ldots n$*,* $s_{i+1} \in \alpha_i(s_i)$*.*
– *Un plan non-déterministe* $\pi = \alpha_1. \cdots . \alpha_n$ *sur* A *est* robuste *pour un but* $G \subseteq S$ *étant donné un état initial* $s_0 \in S$ *si et seulement si pour chaque trajectoire* $s_O, \ldots, s_{n+1}$ *pour* π*,* $s_{n+1} \in G$*.*

Il existe plusieurs modèles servant à traiter les problèmes de planification robuste. Ainsi, [Palacios & Geffner 2007] proposent un moyen de faire correspondre à chaque problème de planification robuste un problème équivalent de planification classique. [Hyafil & Bacchus 2003] présentent quant à eux un algorithme pour les problèmes de planification probabiliste. Il s'agit de problèmes dans lesquelles l'agent possède en plus des hypothèses habituelles une distribution de probabilités sur les résultats possibles des actions mais toujours aucune observation possible de ces résultats durant l'exécution.

Processus de décision de Markov

Cette description des processus de Markov suit l'approche de [Littman *et al.* 1995].

Un processus de Markov (ou MDP) M est un tuple $< S, A, F, R >$ où S est un ensemble fini d'états du système. A est un ensemble fini d'actions. $F : A \times S \rightarrow P(S)$ est la fonction de transition entre états, liant une action et un état à une distribution de probabilités sur S pour les états résultants possibles. La probabilité d'atteindre un état s en exécutant l'action a est notée $F(a; s; s_0)$. $R : S \times A \rightarrow R$ est la fonction de récompense. $R(s; a)$ est la récompense que le système reçoit s'il exécute l'action a dans l'état s. Une politique pour un MDP est une application $Y : S \rightarrow A$ qui séléctionne une action pour chaque état. Etant donnée une politique, nous pouvons définir une fonction de valeur à horizon fini $V(Y, n) : S \rightarrow R$, où $V(Y, n)(s)$ est la valeur espérée de l'application de la politique Y pour n pas en commençant dans l'état s. Elle est définie de manière inductive par $V(Y, 0)(s) = R(s; Y(s))$ et

$$V(Y, m)(s) = R(s; Y(s)) + \sum_{u \in S} F(Y(s); s; u)V(Y, m-1)(u)$$

Une solution à un MDP est une politique qui maximise sa valeur espérée. L'élaboration d'un modèle des processus de Markov comme dans [Lang *et al.* 2000] permet de les utiliser comme procédure de choix dans la planification. Ainsi, [Littman *et al.* 1995], proposent ces processus comme un modèle à la planification sous incertitude. [Boutilier *et al.* 1999] présentent un modèle général pour la spécification de problèmes de planification.

3.4 Planification multi-agent

3.4.1 Introduction

Une situation de planification multi-agent est celle par exemple d'un conducteur qui doit prendre une route avec sa voiture [Wood 1983]. Il a en même temps à obéir aux règles de circulation (rouler à droite, s'arrêter au feu rouge, etc) et à prendre en compte les autres véhicules et éviter d'entrer en collision avec eux. Dans les sections précédentes, la présence des autres agents planificateurs (ou du moins actifs) a été évoquée comme une source possible de l'incertitude à laquelle se trouve confronté l'agent planificateur. Les méthodes de planification proposées la traitent comme toute autre source d'information. En planification robuste par exemple, l'agent forme un plan qui lui assurera d'atteindre son but quelles que soient les contingences. Pour cela il adopte une attitude prudente dans laquelle il considère que les pires conséquences sont possibles et c'est cela à quoi l'agent se prépare en supposant que ce sont celles-là qui se produiront. Pourtant, dans les problèmes de la vie réelle, on a affaire à des agents multiples qui ont leurs propres buts.

Cela introduit deux différences importantes avec le cas précédent. L'influence de la présence de l'autre (ou d'autres) agent(s) est d'abord prévisible par le fait que sa rationalité supposée et ce que l'on sait de ses (leurs) propres buts, connaissances et actions limitent le nombre d'options. Ensuite, les effets de ces actions ne sont pas seulement négatifs. Ils peuvent aussi être positifs. Ces deux points réunis rendent appréciable, et même nécessaire dans certains cas, d'essayer de trouver une manière de contrôler cette interaction entre plans. Ce contrôle peut se réduire simplement à une anticipation des plans de l'autre. Il peut aller plus loin quand c'est possible au moyen d'un protocole de coordination avec l'autre agent. En effet, si les agents ne prennent pas en compte les dépendances entre leurs actions, ils peuvent se retrouver dans une situation de conflit au moment de l'exécution de leurs plans. Afin de résoudre leur dépendances, les agents doivent coordonner leurs efforts. Dans la littérature multi-agent, plusieurs définitions de la coordination sont données. Une définition claire et concise en a été donnée dans [Malone & Crowston 1994] :

Définition 3.3 *Le coordination est l'acte de gérer les interdépendances entre les activités.*

[Jennings 1996] en donne une autre qui est plus spécifique au domaine multi-agent :

> "La participation à toute situation sociale devrait être à la fois contraignante, en ce que les agents doivent y contribuer, et enrichissante, en ce que la participation procure des ressources et des opportunités qui sinon seraient indisponibles. La coordination, le processus par lequel un agent raisonne à propos de ses actions locales et les actions (anticipées) des autres agents afin de s'assurer que la communauté agit d'une manière cohérente, est la clé de la réalisation de son objectif."

Il est donc clair que la planification multi-agent traite autant de la planification que de la coordination. Pour résumer, l'équation suivante résume le problème de planification multi-agent :

$$\text{Planification multi-agent} = \text{planification} + \text{coordination}$$

Dans la suite de ce chapitre, on s'attachera à développer ces deux point importants. On commencera d'abord par présenter les différents aspects de la coordination multi-agents. On décrira ensuite les différentes approches à la planification utilisées et on verra que le choix de l'une de ces approches implique une procédure particulière de coordination.

3.4.2 Coordination

Nous avons vu que la coordination est ce qui permet de différencier la planification multi-agent du cadre général de planification dans l'incertain. Afin d'étudier le concept de coordination, dans le but de déterminer quel mécanisme de coordination est approprié pour un problème particulier, plusieurs points sont à examiner. On peut d'abord se poser la question de savoir la raison qui crée le besoin de coordination. [Jennings 1996] et [Nwana *et al.* 1996] en montrent cinq ; l'évitement de l'anarchie, la gestion de contraintes globales, l'utilisation de capacités particulières d'autres agents, l'existence d'une interaction entre les actions des agents et l'efficacité globale. On voit alors clairement que ces raisons reviennent d'une façon ou d'une autre à prendre en compte le fait que les actions des agents ont des effets sur les autres agents. L'anarchie ne serait pas génante si elle n'influait pas sur la capacité des agents à atteindre leurs buts. De même si les ressources des agents sont séparées ou infinies, il n'y a pas à se soucier des contraintes des autres agents. Les capacités des agents représentent une interaction qui peut être positive pour eux. Si des agents peuvent accomplir des tâches à la place d'autres c'est parce que cela leur permet d'atteindre leur but ou du moins améliorer leur utilité en leur fournissant un plan moins coûteux. La quatrième raison présentée vise elle l'existence d'une interaction négative entre les actions (comme des actions en rendant d'autres impossibles). Enfin la dernière met en lumière l'opportunité d'éviter par exemple les doublons dans les actions des agents.

La question qui se pose ensuite est de savoir la forme que doit prendre le processus de coordination. Deux approches à la coordination multi-agent coexistent. La coordination explicite implique des agents qui raisonnent au sujet de leurs interactions et leurs négociations. Un problème avec la coordination explicite est qu'elle peut être extrêmement coûteuse en temps, ce qui peut la rendre inutilisable en pratique. Avec la seconde approche, la coordination implicite, les agents suivent des règles locales de comportement qui assurent qu'ils peuvent coopérer sans avoir à se soucier des interférences avec les autres.

Un autre aspect de la coordination est le moment auquel elle se produit. Dans la littérature, on peut distinguer de ce point de vue deux approches dominantes à la coordination multi-agent. Dans la première approche, la coordination entre agents est établie après la complétion de planification individuelle. Il est supposé que les agents travaillent indépendamment sur une partie du problème. Ensuite, dans une phase de coordination, les conflits possibles entre les plans individuels sont résolus et des interactions positives entre eux sont exploitées en échangeant et en révisant des parties des plans individuels. La deuxième approche traite la coordination et la planification comme deux processus entremêlés où les agents échangent continuellement des informations pour arriver à une solution jointe. Par exemple dans GPGP ([Durfee & Lesser 1991, Decker *et al.* 2000], [Durfee & Lesser 1991]) la planification et la coordination sont considérées comme deux étapes d'un processus itératif avec des plans de différents niveaux d'abstraction qui sont échangés entre les agents. La différence avec la première approche est que les interactions positives (négatives) entre les plans individuels sont exploitées avant que chaque agent produise un plan complet. Cette approche devrait alors être considérée comme une coordination pendant la planification où les agents sont supposés coopératifs à la fois pour ce qui est de l'échange d'information mais aussi dans leur volonté de réviser des plans déjà calculés.

Dans ces deux approches, il est nécessaire que les agents soient d'accord pour partager l'information et, dans le cas d'un conflit, de réviser leurs plans. Ainsi, ces deux approches sont moins appropriées si les agents ne sont pas coopératifs. Dans de tels cas, les agents doivent posséder une totale autonomie de planification. Cela implique que la seule possibilité pour les agents de se coordonner est de le faire avant que les plans ne commencent à être construits. Dans [ter Mors *et al.* 2004] une méthode de coordination avant planification est étudiée. Celle-ci ajoute un ensemble minimal de contraintes sur les sous-buts qui doivent être réalisés. Ces contraintes servent à s'assurer une solution coordonnée en planifiant chacun de son côté.

Une autre méthode étudiée est l'introduction des normes sociales [Shoham & Tennenholtz 1995]. Une norme sociale est une convention généralement acceptée que chaque agent suit. De telles normes restreignent le comportement des agents. Elles peuvent être utilisées pour réduire les coûts de communication ainsi que les délais de planification et de coordination. Il en est ainsi des travaux de [Foulser *et al.* 1992]. Un exemple typique de norme sociale dans le monde réel est celui du code de la route : parce que chacun conduit sur le côté droit de la route, virtuellement aucune coordination avec les voitures venant dans l'autre sens n'est nécessaire. Généralement, les solutions basées sur les normes ne sont pas optimales, mais sont rapides à mettre en oeuvre.

3.4.3 Planification distribuée

Nous allons aborder dans la suite les méthodes de planification dans le contexte multi-agent. La caractéristique première qui nous servira à les distinguer est celle de l'architecture adoptée. Il s'agit principalement de savoir si les agents sont coopératifs ou non. Dans le premier cas, on peut toujours se ramener au cas où les agents partagent un but commun global. Cette configuration où des agents essayent de résoudre ensemble un problème de planification est aussi appelée planification distribuée. Le terme de planification distribuée est toutefois ambigu car il n'explicite pas ce qui est distribué. En effet, les plans peuvent être construits de manière centralisée puis distribués aux agents, ou bien chaque agent peut construire localement son propre plan puis le coordonner de manière distribuée. Dans le premier cas, seule l'exécution du plan est distribuée. En revanche, dans le second, la synthèse de plans, le processus de coordination ainsi que l'exécution sont réalisés de manière complètement distribuée.

Nous présentons d'abord les mécanismes de coordination centralisée de plans, puis les trois grands mécanismes de coordination décentralisée qui sont la planification partiellement globale, la synchronisation distribuée de plans et la fusion incrémentale de plans.

Architecture centralisée

Une première approche requiert un agent central qui planifie pour tous les agents à la fois. Cet agent centralise l'ensemble des plans des agents du système et résout les conflits potentiels entre leurs activités en introduisant des actions de synchronisation. Soit l'agent coordinateur planifie pour l'ensemble des agents et, dans ce cas, il doit décomposer le plan global en sous-plans synchronisés pouvant être exécutés par les agents, soit chaque agent planifie localement et, dans ce cas, le rôle de l'agent coordinateur se limite à la synchronisation des plans reçus. Dans ce dernier cas, l'agent coordinateur doit examiner les plans partiels des agents pour en identifier les interactions. Il utilise alors un mécanisme de résolution de conflits pour rendre les plans cohérents et les réunir pour former un plan global. Il existe plusieurs mécanismes de résolution de conflits.[Georgeff 1983] présente un mécanisme où l'agent coordinateur insère des commandes explicites de coordination dans le plan global multi-agent.

Dans [Thorndyke *et al.* 1981], les auteurs présentent une solution adaptée au problème de contrôle de trafic aérien au moyen de la planification. Les avions qui prévoient un conflit entre leurs plans de vols respectifs décident entre eux de désigner un agent coordinateur qui reçoit les plans des autres et essaye de modifer son propre plan pour éviter le conflit.

Architecture décentralisée

Au delà même du problème posé par la complexité de la tâche de l'agent central, l'approche centralisée présentée dans la section précédente souffre de son inadaptation à de nombreuses situations de planification. On se trouve en effet souvent en présence d'agents qui n'ont pas le même but comme par exemple deux robots qui se déplacent dans la même pièce. Sans forcément avoir l'intention de nuire, les buts des deux agents peuvent temporairement se contredire. On voit dans l'exemple précédent des robots que le problème vient du fait que les actions des agents interfèrent les unes avec les autres. Si ce n'était pas le cas, un agent pourrait simplement ignorer les autres, réaliser ses buts comme s'il était seul. Enfin les agents peuvent avoir un intérêt à se cacher une partie de l'information qu'ils possèdent. Par exemple, une entreprise peut vouloir garder son bilan hors de portée de ses concurrents. Un modèle basé sur un agent central qui "connaît" tout de l'environnement ne peut clairement pas capturer cet aspect de l'interaction.

Planification individuelle En principe, n'importe quelle technique de planification peut être utilisée et différents agents peuvent utiliser différentes techniques. Il y a quelques approches qui entremêlent la planification et la coordination des plans. Dans le cadre Partial Global Planning (PGP) [Durfee & Lesser 1991], et ses extensions, Generalized PGP [Decker & Lesser 1992], chaque agent possède une conception partielle des plans des autres. Dans cette méthode, la coordination s'opère de la manière suivante. Si un agent A informe un agent B d'une partie de son propre plan, B fusionne cette information dans son propre plan partiel global. L'agent B peut alors essayer d'améliorer le plan global en éliminant par exemple les redondances qu'il observe. Un tel plan amélioré est montré aux autres agents, qui peuvent soit accepter soit rejeter ce plan ou le modifier. Ce processus prend place en parallèle de l'exécution de la première partie du plan local. Une revue des approches semblables à PGP est donnée dans [Lesser 1998]. [Clement & Barrett 2003] ont amélioré ce cadre de PGP en séparant l'algorithme de planification de la coordination en utilisant une approche plus modulaire appelée activité partagée (SHAC).

Allocation de tâches Les méthodes centralisées mentionnées plus haut traitent usuellement aussi de l'allocation des tâches aux agents. Il existe, cependant, plusieurs autres méthodes pour établir une telle allocation d'une manière plus distribuée, qui donnent aux agents un plus grand degré d'autonomie, par exemple, via des protocoles complexes d'allocations [Shehory & Kraus 1998] ou des enchères et simulations de marché. Une enchère est un moyen d'être sûr qu'une tâche est allouée à l'agent qui y attache le plus d'importance. [Vickrey 1961] présente un protocole d'enchère dans lequel chaque agent peut faire une seule enchère cachée, et une tâche est assignée au plus grand enchérisseur au prix de la seconde plus grande enchère. Ce protocole d'enchère a la propriété souhaitable qu'il pousse les agents à enchérir la vraie valeur qu'il attachent au produit de l'enchère. Des simulations de marché peuvent être aussi utilisées pour distribuer de larges quantités de ressources entre des agents [Walsh & Wellman 1999], [Walsh & Wellman 1998]. Par exemple, [Huberman & Clearwater 1995] montrent comment la monnaie et les coûts peuvent être transformés en moyen de coordination. Ces méthodes sont utilisées pour la coordination de plans déjà construits par les agents. Dans le contexte d'un environnement orienté valeur, des approches basées sur la théorie des jeux (où les agents raisonnent à propos du coût de leur prise de décision) deviennent plus importantes. Un exemple en est le travail de [Sandholm & Lesser 1997] pour le problème de routage de véhicules.

Planification hiérarchique Une approche qui considère à la fois des conflits et des relations positives est proposée par [von Martial 1992]. Celui-ci présente des plans de manière hiérarchique, et le niveau le plus haut doit être échangé entre les agents pour déterminer de telles relations. Si possible, les relations sont résolues et exploitées à ce niveau d'abstraction. Sinon, un raffinement des plans est effectué, et le processus est répété. Pour chaque type spécifique de relation entre les plans, une solution différente est présentée. Les relations entre les plans d'agents autonomes sont catégorisées. Les aspects principaux sont les relations positives/négatives, les ressources consommables/ non consommables.

Synchronisation distribuée de plans Les travaux basés sur la synchronisation voient les activités des agents comme des processus concurrents qui doivent être synchronisés. Dans ce cadre, [Melliti *et al.* 2005] proposent un algorithme distribué qui permet de réaliser une telle coordination. Ces travaux s'appuient sur une représentation du monde par états de manière similaire au formalisme STRIPS. La production des plans est réalisée par les agents en fonction de leurs buts locaux. Chaque agent est alors responsable de la synchronisation de ses plans avec les autres agents. Pour qualifier un plan local validé dans un contexte multi-agent, i.e., exempt de conflits, [Melliti *et al.* 2005] introduit la notion de plan structuré. Tout d'abord, l'agent élabore un plan de manière indépendante sans tenir compte des plans produits par les autres agents. Puis, il essaie de synchroniser son plan en le diffusant aux autres agents.

Fusion de plans [Tsamardinos *et al.* 2000] a développé un algorithme de fusion de plans qui traite d'actions qui ont une durée dans le temps. Les auteurs construisent un réseau temporel simple pour spécifier les conflits (temporels) entre les plans. Basé sur cette spécification, un ensemble de contraintes est dérivé qui peut être résolu par un solveur de contraintes. La solution spécifie les relations temporelles entre les actions dans le plan résultat de la fusion. Une autre approche pour fusionner un ensemble de plans dans un plan global traite de problèmes surgissant des actions redondantes ou en conflit en utilisant des méthodes de recherche A^* et une méthode basée sur une heuristique intelligente : [Ephrati & Rosenschein 1993] ont montré que, en divisant la tâche de construction de sous-plans entre plusieurs agents, on peut réduire la complexité globale de l'algorithme de fusion. Dans d'autres travaux sur la fusion de plans, [Ephrati *et al.* 1995] [Rosenschein 1995] proposent un algorithme distribué en temps polynomial pour améliorer le bien-être social (i.e., la somme des bien-être des agents). Au travers d'un processus d'agrégation de contraintes de groupe, les agents construisent de manière incrémentale un plan global amélioré en votant autour d'actions jointes. Ils proposent même des algorithmes qui traitent des agents insincères et pour entremêler planification, coordination et exécution.

3.4.4 Agents "égoïstes" dans un contexte multi-agent

Il est important pour aborder cette partie de la théorie de la planification de bien spécifier ce que l'on entend par le terme d'agent égoïste. D'abord, cela ne veut pas dire qu'un tel agent veuille causer du tort aux autres agents. On vise plutôt par ce terme un agent qui possède sa description propre des états du monde qu'il préfère et dont les actions sont motivées seulement par l'obtention de ces états.

Dans cette section, nous allons passer en revue les modèles proposés lorsque le système comporte des agents égoïstes. La différence avec les modèles présentés dans la section précédente est que les agents ne partagent plus leurs buts. On peut toutefois toujours considérer si on voit le groupe d'agents dans sa globalité que le processus de planification est distribué. Certains des modèles utilisés précédement peuvent donc être étendus pour traiter le cadre présent.

Les systèmes de fusion de plans [Tsamardinos *et al.* 2000], [de Weerdt *et al.* 2003] peuvent être étendus au cas des agents égoïstes afin de coordonner leurs plans. Dans ces systèmes, chaque agent contruit son propre plan, sans échanger d'information avec les autres agents. Quand tous les plans ont été calculés, peu d'information est échangée pour détecter les interactions possibles. Clairement, ces approches sont statiques et ne peuvent plus être utilisées pour demander l'aide d'autres agents durant le processus de planification. Une autre voie de recherche concerne le raisonnement pendant la construction des plans multi-agents. Des exemples d'une telle recherche sont les travaux de [Cohen & Levesque 1991] sur les intentions jointes et l'approche des plans partagés de [Grosz *et al.* 1999]. Ces deux approches se concentrent sur un comportement coopératif. Les modèles d'allocation de tâches sont aussi applicables pour des agents égoïstes. Par exemple, les mécanismes de marché [Walsh & Wellman 1999], ou l'utilisation d'extensions du protocole des réseau de contrats [Smith 1980]. [Bai & Zhang 2008] considère quant à lui que la structure d'équipe est la mieux indiquée pour que les agents coopèrent. En d'autres termes, les agents se regroupent en équipes pour exécuter des tâches précises. L'effort à fournir est alors celui de choisir un procédé rationnel qui permet aux agents de choisir les équipes dans lequelles ils vont s'engager. Toutefois, l'agent peut avoir besoin de changer d'équipe car dans un système dynamique et par définition concurrentiel, des paramètres de l'environnement, dont le plus évident est celui des ressources disponibles, peuvent changer. Cela signifie du point de vue des agents que le caractère commun ou conflictuel de leurs buts peut changer. Cette possibilité supplémentaire d'entrer et de sortir des équipes implique de rechercher un mécanisme qui permet aux agents, sinon de contraindre, au moins d'anticiper l'engagement et surtout le désengagement des autres agents. Ce mécanisme doit s'adapter au fait que la nature des collaborations entres agents (au sein d'une équipe) peuvent être de durée variable. Les agents peuvent en effet coopérer de manière ponctuelle ou alors de manière plus stable et durable. Un exemple de mécanisme basé sur les contrats se trouve dans [Rathod & des Jardins 2004] dans lequel des "chefs d'équipe" recrutent des agents pour leur équipe en leur proposant des contrats contenant essentiellement un partage des gains espérés liés à la formation de l'équipe (et à l'exécution de la tâche visée). Ces contrats peuvent être de longue durée et dans ce cas ils sont basés sur une pénalité de rupture censée retenir les agents. Ils peuvent aussi être de court terme et porter sur une tâche précise.

Une autre partie des travaux abandonne cette idée et traite le cas d'un système multi-agent dans lequel la coopération se décide au terme d'un processus de négociation [Florea 1999]. La négociation se base sur une balance entre les gain et les pertes associées à la coopération. La négociation est guidée chez chaque agent par un ensemble de règles d'inférence. Celle-ci incorporent notamment un profil des autres agents qui code leur comportement durant les précedentes interactions. Les agents se construisent alors plusieurs types d'agents qui diffèrent selon des critères tels que la tenue d'engagements passés, le désir d'atteindre un gain maximum ou la volonté affichée d'établir de bonnes relations de coopération. C'est un modèle que l'on retrouve par exemple dans les applications de commerce pair à pair sur internet où des systèmes de réputation permettent d'avoir une idée sur le sérieux de la personne avec qui on est sur le point d'effectuer une transaction. Le système de réputation n'apporte toutefois pas assez de garanties dans le cas de systèmes critiques tels que celui de la gestion de traffic aérien traité dans [Geert *et al.* 2007]. Dans ce travail, les auteurs présentent un système de négociation au sein d'un marché. Il s'agit d'un système monétaire dans lequel chaque agent peut émettre de la monnaie qu'il signe au moment où il la transmet. L'inconvénient de ce genre de système est que certains agents peuvent abuser de leur situation dominante dans le marché pour imposer des accords injustes. Les auteurs montrent qu'il est possible de dépasser ce problème en dotant le système de stratégies dominantes consistant à éviter les contre-venants. L'idée est de réintroduire une notion de confiance mais cette fois-ci sur la monnaie émise par un agent.

3.5 Conclusion

Dans ce chapitre, nous avons passé en revue les modèles existants pour la résolution d'un problème de planification. La variété des modèles vise surtout à prendre en compte l'incertitude à laquelle le planificateur doit faire face dans des situations réalistes. Une grande partie des modèles sont basés sur l'idée d'introduire dans le modèle idéalisé de la planification classique des concepts permettant la prise en compte de l'incertitude. La différence se situe alors dans les hypothèses du cadre auquel ces modèles seront adaptés. Y a-t-il possibilité d'observer le résultat des actions ? Peut-on réparer un plan après le début de son exécution ? etc. La nature de l'incertitude guide aussi le choix du modèle. Par exemple, des planificateur markoviens sont adaptés à une information probabiliste. Par ailleurs, le fait que cette incertitude soit due à l'interaction avec d'autres agents ouvre de nouvelles possibilités devant le planificateur. Il ne s'agit alors plus seulement de prévoir l'évolution future de l'environnement et de s'y adapter. L'agent peut de plus profiter des synergies positives qui peuvent naître de l'interaction avec d'autres agents rationnels, y compris en communicant explicitement avec eux.

Chapitre 4

Théorie des jeux sous information incomplète

4.1 Introduction à la théorie des jeux

4.1.1 Un point historique

Voici comment John von Neuman et Oskar Morgenstern présentent la théorie des jeux :

> Un grand nombre de choses sont couvertes par cet intitulé, de la roulette aux échecs, du baccarat au bridge. Et après tout, tout évènement, étant données les conditions externes et les participants à une situation (pourvu que ces derniers agissent de leur propre chef) peut être vu comme un jeu de stratégie si chacun tient compte des effets qu'il a sur les participants. [von Neumann 1928]

Cette citation de von Neumann et Morgenstern définit bien le cadre de la théorie des jeux, et ce à double titre. D'abord, les auteurs y dépeignent clairement la grande variété de situations et d'évènements qui peuvent être modélisés en utilisant cette théorie. Ensuite, cette citation, et plus largement le livre dont elle est issue, marque le moment où la théorie des jeux fût acceptée comme une nouvelle discipline. Bien sûr il y eut des précurseurs comme Cournot. Il apporta au 19ème siècle une formalisation mathématique à l'étude de phénomènes économiques. Mais le point historique de la citation présentée est important car toute la théorie développée plus tard se rapporte au formalisme et aux hypothèses posées par ce travail. Cela ne se fît toutefois pas rapidement et sans heurt. Aussi bien pour les travaux de Cournot et de von Neuman-Morgenstern, leur formalisation mathématique et les hypothèses particulières adoptées ont fait que les économistes, public privilégié pour l'application d'une théorie des interactions, mirent quelques temps à s'y adapter, les comprendre et en faire ce qui est aujourd'hui une partie classique de la théorie économique.

Voici comment von Neumann arriva à la citation introductive. Le mathématicien Ernst Zermelo s'est intéressé au problème suivant : comment peut-on modéliser mathématiquement le jeu d'échecs ? A partir de la théorie des ensembles, Zermelo a montré qu'il existait une stratégie optimale pour chacun des joueurs, autrement dit qu'on pouvait, moyennant une capacité de calcul suffisante, trouver une stratégie unique déterminée selon toutes les possibilités choisies par l'adversaire. Von Neumann reprend alors le problème à son compte en le généralisant bien au delà du jeu d'échec de la manière suivante : comment peut-on caractériser la rationalité mathématiquement ou en d'autres termes trouver un algorithme permettant de déterminer une action en fonction de celles des autres ? Sa rencontre avec Morgenstern et à travers lui de la théorie économique permit de cristalliser la question centrale de la théorie des jeux : caractériser les fondements de la rationalité dans une rencontre entre acteurs rationnels. En effet, l'éco-

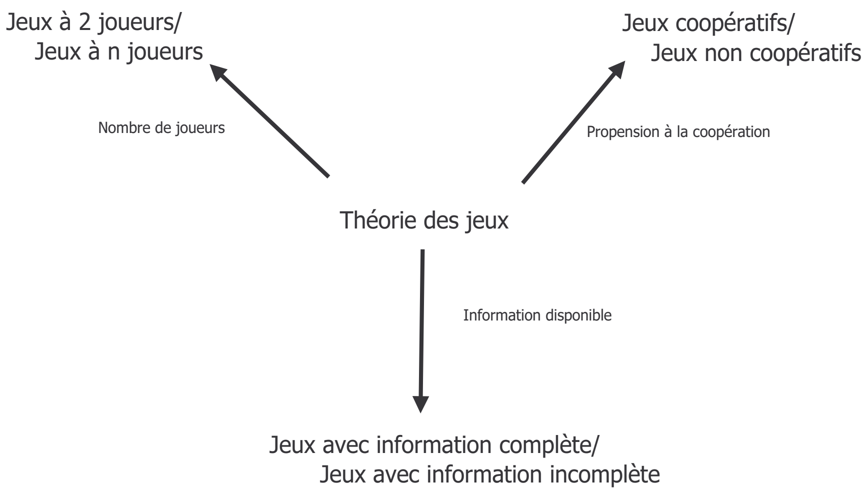


FIGURE 4.1 – Taxonomie des jeux

nomie théorique n'avait jusque là jamais été capable de décrire les interactions directes entre individus. Il fallait supposer quelque chose d'abstrait comme "un marché" pour en parler. Là, grâce à cette nouvelle théorie, on allait enfin pouvoir découvrir cette dimension fondamentale de la rationalité interactive.

4.1.2 Caractéristiques d'un jeu

Plusieurs modèles existent pour décrire la situation de jeu. Parler d'un jeu pour en exprimer les caractéristiques n'a pas de sens si on ne précise pas dans quel modèle précis nous nous plaçons. On part dans chacun de ces modèles du concept d'interaction. Suivant ce concept, un jeu est formé avant tout par des joueurs. Chacun d'eux a un intérêt propre qu'il cherche à défendre [2]. On arrive à un modèle en précisant les conditions de l'interaction entre les joueurs.

Commençons par une taxonomie des modèles de jeux existants 4.1.

Un modèle de jeu pourra être par exemple celui de jeux à deux joueurs non coopératifs sous information complète. Une fois que le modèle est défini, restent à donner les caractéristiques du jeu permettant de développer un raisonnement à propos de son issue possible. On convient généralement qu'on y parvient en répondant à la liste suivante de questions :

– Quoi ? (actions,stratégies)
– Combien ? (que rapporte chaque issue à chaque joueur)
– Comment ? (déroulement d'un jeu)

2. Nous pouvons l'illustrer dans le domaine de l'économie en modélisant un marché donné comme un jeu. Pour le producteur, son intérêt sera de maximiser son profit face à des concurrents en choisissant, par exemple, le meilleur prix de vente. Le consommateur cherchera quant à lui à acquérir le bien qui l'intéresse au prix le plus bas après un marchandage avec le vendeur.

Les joueurs peuvent effectuer des actions qui vont influer sur l'environnement qu'ils partagent. Du résultat de l'ensemble des actions de tous dépend l'utilité de chacun.

Quoi ? En fait le choix des joueurs se fait entre des stratégies. Une stratégie peut être envisagée comme "une collection de décisions pour toute situation future concevable du jeu"[von Neumann 1928]. En d'autres termes, le joueur doit prévoir une action préférée pour chaque combinaison possible des actions des autres. Le choix d'une stratégie par joueur définit ce que l'on appelle un profil de stratégies.

Combien ? D'une manière générale, quelle que soit l'unité dans laquelle il est formulé et même s'il n'en vérifie pas les hypothèses basiques, le résultat du jeu pour un joueur est souvent appelé un gain (sous-entendu monétaire). Deux aspects de ces gains permettent de définir chacun des types différents de jeux. D'abord il y a la relation entre les gains de joueurs. Historiquement, von Neuman et Morgenstern étudièrent des jeux à somme nulle (et à deux joueurs). Ceux-là se définissent de la manière suivante. Le résultat associé à un profil donné de stratégies est tel que ce qui constitue un gain positif d'un joueur constitue une perte pour l'autre. Ensuite la nature même du gain est importante. Nous avons précédemment vu qu'étant donné le lien historique entre la théorie des jeux et le domaine économique, le gain des joueurs est le plus souvent représenté par une fonction d'utilité telle que définie à la section 2.2.2. Il est facile de voir que cette hypothèse, bien que commode, n'est pas nécessaire. Il est en effet possible de considérer un gain qui est sous l'une des formes plus ou moins quantifiées et précises utilisées dans la théorie de la décision.

Comment ? Les agents suivent un certain nombre de règles en exécutant les actions prescrites par leurs stratégies. Chaque joueur est supposé agir rationnellement. La rationalité visée ici implique que l'agent essaye de maximiser son utilité quel que soit ce que les autres feront. La situation est donc la suivante. Chaque profil de stratégies mène à un résultat pour chaque joueur (un vecteur de résultats). C'est pour essayer d'améliorer leur résultat personnel (la ième composante du vecteur) que les agents raisonnent ainsi : "mon action seule ne suffit pas à définir le résultat, les actions des autres influent aussi. Or mon action influe aussi sur le résultat des autres. Dans l'établissement de leur stratégie, ils doivent donc la prendre en compte de la même manière que moi les leurs. Je dois trouver la stratégie qui incitera/poussera/obligera les autres joueurs à choisir celles de leurs stratégies qui me donnent le meilleur résultat possible". Les joueurs doivent donc se faire une idée aussi précise que possible des actions choisies par les autres. Pour cela, deux hypothèses importantes sont faites dans le raisonnement précédent :

- chaque joueur s'efforce de prendre les meilleures décisions pour lui-même et sait que les autres joueurs font de même.
- chacun sait que l'autre sait et ainsi de suite ad infinitum.

4.1.3 Représentation d'un jeu

Regardons maintenant comment on peut représenter une situation de jeu et les notions s'y rapportant que nous venons de présenter. Pour représenter un jeu, deux possibilités existent.

Forme stratégique On peut représenter le jeu sous forme normale (aussi appelée forme stratégique), c'est-à-dire sous la forme d'un tableau contenant les gains des joueurs en fonction de leurs stratégies respectives. Un jeu sous forme stratégique est une collection de stratégies décrivant les actions de chaque joueur dans toutes les situations concevables du jeu, ainsi que les gains que chacun obtient lorsque les stratégies de tous les joueurs sont connues.

Définition 4.1 *Un jeu est la donnée de* $G = (N, (S^i)_{i \in N}; (g^i)_{i \in N})$ *où*
- *N est un ensemble non vide appelé ensemble des joueurs,*
- *pour chaque joueur i dans N,* S^i *est un ensemble non vide appelé ensemble de stratégies du joueur i.*
- *pour chaque joueur i dans N,* g^i *est une application de* $\prod_{j \in N} S^j$ *dans* $\mathbb{R}$ *appelée fonction de gain du joueur i.*

Chaque joueur i doit choisir une stratégie s^i dans S^i. il n'est pas nécessaire que les choix d'actions soient simultanés, il suffit en fait qu'au moment de choisir, chaque joueur ne soit pas au courant des choix éventuellement déjà exécutés par les autres joueurs. C'est le cas par exemple lorsque chaque joueur écrit l'action qu'il choisit dans une enveloppe cachetée, et qu'ensuite toutes les enveloppes sont récoltées et ouvertes par un arbitre. A la fin du jeu, si chaque joueur j a choisi l'action s^j de S^j , le gain d'un joueur i est $g^i(s)$, où $s = (s^j)_{j \in N}$. Chaque joueur cherche uniquement à maximiser son propre gain. Tous les joueurs connaissent G.

Une stratégie pure est une action, ou un plan d'actions, qui est choisie par chaque joueur avec certitude. Cette notion a été étendue à celle de stratégie mixte définie comme une distribution de probabilités sur l'ensemble des stratégies pures.

Définition 4.2 *Soit* G *un jeu,* i *un joueur et* $s^i_1, \cdots s^i_m \in S^i$ *m stratégies pures possibles, une stratégie mixte du joueur* i *est la donnée d'une distribution* $(p_1, p_2, \ldots p_m)$ *telle que* $\sum_{i=1}^m p_i = 1$ *où* p_i *est la probabilité de jouer la ième stratégie.*

Exemple 4.1 Il y a deux joueurs : $N = \{1, 2\}$. Le joueur 1 doit choisir une ligne Haut ou Bas, on a : $S^1 = \{H, B\}$. Le joueur 2 doit choisir une colonne Gauche ou Droite, on a : $S^2 = \{G, D\}$. Les gains, c'est-à-dire les fonctions g^1 et g^2, sont donnés par la table 4.1 :

	G	D
H	(1, 1)	(3,0)
B	(0, 3)	(0, 0)

TABLE 4.1 – EXEMPLE DE JEU

Les cases de cette matrice représentent les éléments de $S^1 \times S^2$. Dans chaque case, on trouve un couple de réels : la première composante donne le gain du joueur 1, la seconde celui du joueur 2. Par exemple, dans la case (H, D) se trouve le couple $(3, 0)$: cela signifie que $g^1(H, D) = 3$ et que $g^2(H, D) = 0$. Cette matrice représente donc de façon pratique les fonctions de gains. Le joueur 1 choisira la ligne, le joueur 2 choisira la colonne.

On peut s'interroger sur l'intuition qui sous-tend le concept de stratégie mixte. Pourquoi un joueur aurait-il recours à un mécanisme aléatoire, qu'il choisit, auquel il remet son pouvoir de décision alors qu'il pourrait choisir directement ? Dans [Binmore 1999], l'auteur reprend à son compte ce questionnement et propose un argument qui en même temps un bon exemple d'interprétation d'une stratégie mixte. Il écrit

> Il peut sembler bizarre qu'un joueur rationnel doive choisir d'introduire le hasard de cette manière. Mais considérons le bluff au poker. Il est évident qu'un joueur rationnel doit incorporer le bluff dans son jeu. Supposons, par exemple, que Mary ne bluffe jamais. Quand elle mise beaucoup, chacun saura, en conséquence, qu'elle doit avoir une bonne mise. Dès

lors les opposants se hâteront tous de passer, sauf en ces rares occasions où quelqu'un pense qu'il a une main encore meilleure. Mary gagnera donc gros très rarement quand elle tire une bonne main et elle perdra de temps en temps. Pour avoir une chance de gagner un pot assez important lorsqu'elle a une bonne main, Mary doit parfois miser haut quand elle a une mauvaise main. En d'autres termes, elle doit bluffer à l'occasion. Évidemment le bluff ne lui sera d'aucune utilité si les autres joueurs peuvent anticiper le moment où elle bluffe. Elle a besoin de rendre son comportement de bluff imprévisible, ce qu'elle peut se garantir en déléguant le choix du moment de bluffer à un mécanisme aléatoire soigneusement choisi. Même Mary ne sera pas capable de déterminer à l'avance l'instant où elle bluffera. Un tel stratagème laissera donc certainement dans le doute ses opposants.

Forme extensive La forme extensive ou forme arborescente fût développée par [Kuhn 1953]. Il s'agit de représenter le jeu sous la forme d'un arbre. Le jeu se déroule de manière séquentielle. Partant de la racine, chaque noeud de l'arbre représente un point de choix pour l'un des joueurs. Le choix s'effectue entre les arcs partant de ce noeud symbolisant chacun une action à entreprendre. Ce choix constitue un coup du jeu, qui mène à un noeud fils. Ce noeud est alors un point de choix pour le joueur suivant. Les gains que chaque joueur peut réaliser après avoir suivi un des chemins possibles au sein de l'arbre sont donnés aux feuilles. Cette description d'un jeu en forme extensive suppose que les ensembles d'actions sont finis.

Définition 4.3 *Un jeu sous forme extensive est défini par :*

- N *un ensemble fini de joueurs.*
- C *un ensemble de points de choix.*
- H *un ensemble de séquences distinctes de points de choix (ensemble d'historiques complets).*
- J *une fonction qui assigne un joueur à chaque point de choix.*
- *pour chaque joueur i dans* N*,* g^i *est une application de* H *dans* $\mathbb{R}$ *appelée fonction de gain du joueur i.*

Dans un jeu sous forme extensive, une stratégie pure est une collection de règles décrivant les choix d'un joueur dans chacun des points de choix en fonction de son information. Une stratégie mixte pour un joueur i est une distribution de probabilités sur les stratégies pures de ce joueur.

On peut associer un jeu en forme stratégique à tout jeu en forme extensive en combinant toutes les stratégies possibles et en évaluant les gains correspondants. Pour l'illustrer reprenons l'exemple de jeux représenté ci-dessus sous forme stratégique. La représentation sous forme extensive en est donnée à la figure 4.2.

Exemple 4.2 Les feuilles de l'arbre représentent les résultats possibles du jeu. Dans chaque feuille, on trouve un couple de réels : la première composante donne le gain du joueur 1, la seconde celui du joueur 2. Par exemple, dans la feuille obtenu lorsque au premier point de choix le joueur 1 choisit H puis qu'au point de choix suivant le joueur 2 choisit D se trouve le couple $(3, 0)$: cela signifie que $g^1(H, D) = 3$ et que $g^2(H, D) = 0$.

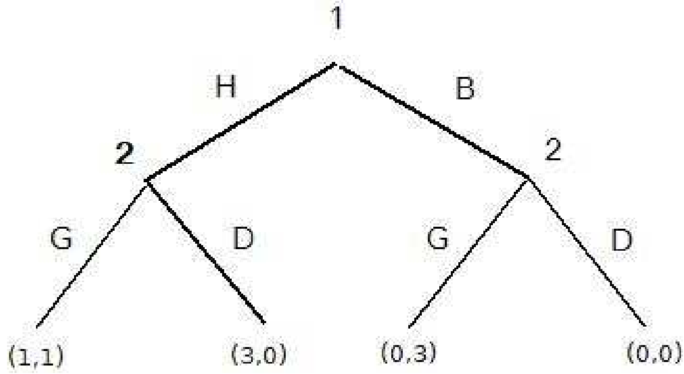


FIGURE 4.2 – Jeu sous forme extensive

4.2 Propension à la coopération

4.2.1 Jeux non coopératifs

Introduction

On convient généralement de distinguer deux grandes familles de jeux : les jeux coopératifs et les jeux non coopératifs.

Lorsque que les joueurs n'ont pas la possibilité de former des coalitions, le jeu est non coopératif. Par définition, dans un jeu non coopératif on spécifie toutes les options stratégiques offertes aux joueurs, alors que les contrats qui sous-tendent les coalitions dans un jeu coopératif ne sont pas décrits.

La théorie des jeux non coopératifs s'est avérée très utile pour l'analyse des luttes d'influence entre groupes au sein d'une institution (par exemple une entreprise), des négociations internationales entre gouvernements ou encore de la concurrence électorale entre partis politiques. Elle permet une description détaillée des mécanismes incitatifs qui guident le comportement des joueurs sur la base de leurs seuls objectifs personnels.

La coopération peut survenir dans des jeux non coopératifs et ce quand les joueurs pensent que c'est dans leur intérêt propre.

Le dilemme des prisonniers Pour commencer, on considère le jeu qui est sans doute le plus célèbre, à savoir le dilemme du prisonnier.

Le premier jeu ayant une structure de dilemme des prisonniers est un jeu proposé par deux mathématiciens, Melvin Dresher er Merrill Flood, en 1950, dans le but de vérifier la solidité du concept d'équilibre de Nash. Le nom du jeu fût donné par Albert Tucker, un autre mathématicien, qui le popularisa.

Voici comment se présente classiquement un tel jeu. On suppose que deux suspects sont interrogés séparément par la police pour une action délictueuse grave. La police ne dispose pas d'éléments de preuve suffisants pour obtenir la condamnation des prévenus pour l'acte dont ils sont accusés. L'aveu d'au moins l'un des deux est donc indispensable. La police propose à chaque accusé d'avouer, dans quel cas il sera relâché. S'il n'avoue pas mais que l'autre le fait, il écope d'une peine de prison de 15 ans. Si les deux avouent, ils peuvent espérer bénéficier de circonstances atténuantes et recevoir une peine de 8 ans chacun. Enfin, si aucun des deux n'avoue, ils seront condamnés à 1 an de prison. La matrice des gains a donc la forme suivante (table 4.2). Ici , il est plus aisé de représenter la sentence consécutive à un profil de stratégies comme des gains négatifs.

	Avoue	N'avoue pas
Avoue	(-8, -8)	(0,-15)
N'avoue pas	(-15, 0)	(-1,-1)

TABLE 4.2 – DILEMME DES PRISONNIERS

Ce qui fait qu'un jeu est un dilemme des prisonniers est la relation suivante entre l'évaluation pour chaque joueur (ici le joueur 1) des profils de stratégies :

$$(Avoue, N'avouepas) \succ (N'avouepas, N'avouepas) \succ (Avoue, Avoue) \succ (N'avouepas, Avoue)$$

Vers un concept de solution

On donne au fait de trouver une solution à un jeu la même signification rencontrée précédemment dans le cadre de la théorie de la décision. Il s'agit de trouver une manière d'analyser le jeu et d'en prédire l'issue. L'issue d'un jeu est un choix de stratégie pour chaque joueur, i.e., un profil de stratégies. Le terme solution d'un jeu désigne donc aussi bien la procédure qu'on applique à tout jeu que l'on rencontre que l'issue du jeu particulier que l'on traite.

Dominance Considérons le cas simplifié où le joueur dispose de seulement deux stratégies s_1 et s_2. Supposons alors que, quel que soit le choix de l'adversaire, le résultat de la stratégie s_1 est meilleur que celui de la stratégie s_2. On dit que la stratégie s_1 domine la stratégie s_2. Comme les joueurs sont supposés rationnels, ils font des choix censés mener à un résultat qu'ils préfèrent, étant donné ce que leurs adversaires font. Un joueur rationnel ne choisit donc pas une stratégie dominée, c'est-à-dire une stratégie qui lui rapporte systématiquement (quelle que soit la stratégie choisie par son adversaire) moins qu'une autre à sa disposition. On peut donc conclure que les profils de stratégies contenant une stratégie dominée ne feront pas partie de la solution du jeu. Si dans un jeu donné tous les joueurs disposent d'une stratégie dominante (qui lui rapporte systématiquement plus que toute autre stratégie à sa disposition) et qu'ils choisissent effectivement cette stratégie, le résultat du jeu est appelé équilibre en stratégies dominantes.

Reprenons l'exemple du dilemme des prisonniers (table 4.2) pour illustrer cette notion et son utilisation possible. On remarque immédiatement qu'avouer constitue une stratégie qui conduit toujours à une peine moins lourde que ne pas avouer, et ce quel que soit le choix effectué par l'autre. Dès lors, il semble naturel de penser que chacun des prévenus va choisir cette stratégie dans l'intention de réduire sa peine. Le résultat est qu'ils vont tous les deux être condamnés à 8 ans de prison, ce qui constitue malgré tout une condamnation assez lourde. Dans ce jeu, la stratégie $Avouer$ est "optimale" puisque, chaque joueur peut prendre sa décision sans avoir besoin de se faire une idée de ce que va faire l'autre. Le couple $(Avouer, Avouer)$ est un équilibre en stratégies dominantes.

Cet exemple fait apparaître un concept plus faible que celui de stratégie dominante, à savoir celui de stratégie dominée. Formellement, on dit qu'une stratégie $s_i \in S_i$ est strictement dominée pour le joueur i s'il existe une autre stratégie $s \in S_i$ telle que :

$$u_i(s, s_{-i}) > u_i(s_i, s_{-i}) \forall s_{-i} \in S_{-i}.$$

On dit que la stratégie s_i est faiblement dominée si dans l'inégalité précédente on remplace $>$ par $\geqslant$, l'inégalité étant stricte pour au moins une combinaison s_{-i}, $u_i(s_i, s_{-i}) < u_i(s, s_{-i})$.

Bien entendu, s'il existe une stratégie dominante pour un joueur, toutes les autres sont dominées et peuvent donc être éliminées. On suit en cela le processus appelé de dominance successive. Il s'agit de choisir une stratégie dominée et de l'éliminer du jeu. On reprend cette opération ensuite sur le nouveau jeu obtenu tant que l'on trouve des stratégies dominées. L'emploi du processus de dominance successive ne conduit pas nécessairement à une solution unique, ce qui laisse donc subsister une indétermination concernant l'issue du jeu. L'ordre d'élimination des stratégies dominées n'a pas d'importance. Les stratégies qui subsistent après élimination sont les mêmes quelle que soit la séquence suivie. Il n'en va plus de même lorsque l'on élimine des stratégies faiblement dominées.

L'application successive de l'élimination par dominance stricte semble donc fournir une solution satisfaisante quand elle conduit a une solution unique. Il arrive pourtant que la solution trouvée ne corresponde pas au comportement observé des joueurs réels (ce qui signifie qu'ils ne satisfont pas toujours à l'hypothèse de rationalité). Le jeu suivant (table 4.3) montre un tel cas :

	G	D
H	(8,10)	(-100,9)
B	(7,6)	(6,5)

TABLE 4.3 – L'ÉLIMINATION DE STRATÉGIES DOMINÉES POUR DES JOUEURS RÉELS

Pour le joueur 1, D est strictement dominée. Après élimination de cette stratégie, B est strictement dominé pour le joueur 2. On peut alors conclure que le résultat du jeu sera le profil (H, G). Toutefois, s'il à la moindre incertitude quant au comportement de 2, le joueur 1 pourrait se retrouver dans une situation catastrophique : s'il joue H tandis que 2 se trompe et joue D, 1 réalise une perte énorme. Certains individus prudents préfèrent jouer B qui leur garantit dans tous les cas un gain significatif et une perte relative assez faible par rapport à H.

Equilibre de Nash John Nash introduit une notion de solution (voir par exemple [J.Nash 1951]) qui permet d'étendre les concepts introduits par von Neumann au cas de jeux qui ne sont plus à somme nulle (ainsi que des jeux à plus de deux joueurs).

Définition 4.4 *Soit G un jeu où N est un ensemble de n joueurs. On dit qu'un profil de stratégies s est un équilibre de Nash si l'inégalité suivante est satisfaite pour chaque joueur*

$$\forall i \in [1, n], \forall s_i^{'} \in S_i, u_i(s_i^{'}, s_{-i}) \leqslant u_i(s).$$

En d'autres termes, si le joueur i anticipe correctement que les autres participants au jeu vont choisir les stratégies associées à s_{-i}, il maximise son gain en choisissant au sein de l'ensemble S^i la stratégie s_i. Cette propriété de stabilité étant satisfaite pour chaque joueur, la solution décrite par l'équilibre de Nash capte ainsi l'idée intuitive de ce que l'on entend communément par "équilibre".

Exemple 4.3 Dans le jeu du dilemme des prisonniers, on a

$$u_1(Avoue, Avoue) = 1, u_1(N'avouepas, Avoue) = 0$$

$$u_2(Avoue, Avoue) = 1, u_2(Avoue, N'avouepas) = 0.$$

Ainsi
$u_1(Avoue, Avoue) > u_1(N'avouepas, Avoue)$ et $u_2(Avoue, Avoue) > u_2(Avoue, N'avouepas)$.

D'après la définition précédente, le profil de stratégies $(Avoue, Avoue)$ est un équilibre de Nash. On peut aisément vérifier que c'est le seul.

Voyons maintenant comment on peut justifier le recours à l'équilibre de Nash comme solution possible d'un jeu. Une première réponse est de voir ce concept comme un ensemble de conditions raisonnables sur le comportement des participants au jeu : chaque joueur choisit la meilleure stratégie pour lui-même compte tenu des croyances qu'il a sur les stratégies qui vont être choisies par les autres. Étant donnée sa capacité à se mettre à la place des autres et à reproduire leurs raisonnements, chaque joueur est à même de formuler des anticipations correctes. Dès que le jeu est terminé et que chacun en observe les résultats, il n'éprouve aucun regret car il n'aurait pas pu augmenter son gain en faisant un choix différent.

Une seconde réponse consiste à imaginer que les joueurs se réunissent avant de jouer et discutent du choix d'une solution possible. Si les joueurs arrivent à un accord sur la façon de jouer et si chaque joueur est personnellement convaincu que les autres participants au jeu vont effectivement se comporter selon les règles de l'accord, alors il est "optimal" pour lui d'en suivre les recommandations car toute déviation unilatérale conduirait à un résultat qui lui serait moins favorable. Bien qu'il ne soit pas a priori contraignant, l'accord s'impose donc à chaque joueur considéré individuellement parce qu'il repose sur l'intérêt de chacun. Dans cette interprétation, l'interdépendance stratégique prend forme lors du processus qui permet la sélection d'un accord.

Équilibres corrélés On définit une stratégie corrélée (notion introduite par [Aumann 1974]) comme une variable aléatoire dont les valeurs sont des profils de stratégies (pures). La corrélation est une forme plus générale d'introduction du hasard dans le choix de stratégies que les stratégies mixtes. Dans les deux cas, les joueurs basent leurs choix sur l'observation d'un évènement aléatoire. Pour les stratégies mixtes, les observations sont indépendantes, alors que cela n'est pas exigé dans le cas des stratégies corrélées. Il se peut en effet que les joueurs décident leurs stratégies en se basant sur la même variable aléatoire.

Si on considère le jeu du dilemme des prisonniers (table 4.2), les joueurs peuvent parvenir à un gain espéré supérieur à ce qu'ils auraient par ailleurs en basant leur choix sur le résultat d'une même expérience aléatoire. Ils obtiennent un gain de $(-\frac{9}{2}, -\frac{9}{2})$ si tous deux observent simultanément le lancer d'une pièce non truquée, puis jouent $(Avoue, Avoue)$ ou $(N'avouepas, N'avouepas)$ en fonction du résultat. Considérons pour le montrer le gain espéré du joueur 1 pour cette stratégie que l'on nommera s (le gain du joueur 2 est calculé de manière symétrique). Si le résultat du lancement de la pièce est Pile, il va jouer $Avoue$ et il sait que le joueur 2 jouera $Avoue$. Si le résultat est Face, il jouera $N'avouepas$ et saura que le joueur 2 fera de même. Son gain espéré est alors, puisque la pièce est supposée non truqué :

$$g(s) = \frac{1}{2} * (-8) + \frac{1}{2} * (-1) = -\frac{9}{2}$$

Dans le cas général, on peut utiliser l'interprétation suivante donnée par [Aumann 1987] pour définir formellement la notion d'équilibre corrélé. On suppose qu'une certaine entité (disons la chance) choisit un profil de stratégies et en fait la suggestion aux joueurs. On représente cette suggestion par $f = (f_1, \cdots, f_n)$. Si chacun des joueurs i suit la suggestion f_i alors la stratégie corrélée est réalisée.

Définition 4.5 *Un équilibre corrélé dans G est un profil de stratégies tel que*

$$E_i(f) \geqslant E_i(f_{-i}, s) \forall i \in [1, n] \;\; et \;\; s \in S^i,$$

où $E_i(f)$ est le gain espéré pour le profil de stratégies f.

Cette formulation se rapproche de celle d'un équilibre de Nash en stratégies mixtes. Et de fait [Aumann 1987] introduit la notion d'équilibre corrélé comme un moyen de répondre à certaines objections sur la rationalité de l'équilibre de Nash.

4.2.2 Jeux coopératifs

Introduction

Un jeu est coopératif lorsque des joueurs peuvent passer entre eux des accords qui les lient de manière contraignante (par exemple, sous la forme d'un contrat qui prévoit une sanction légale dans le cas du non-respect de l'accord). On dit alors qu'ils forment une coalition dont les membres agissent de concert. La spécification d'un jeu coopératif, aussi appelé jeu de coalition, porte sur les gains que chaque groupe potentiel de joueurs peut atteindre en coopérant et sur la manière dont les coalitions se forment.

Marchandage La théorie des négociations ou du marchandage entre deux agents s'inspire des travaux pionniers de Nash [Nash 1950, J.Nash 1951]. Voici comment Nash formule un problème de négociation entre deux joueurs dans [Nash 1950] (dans la suite nous nous intéresserons aussi à ce cadre). Deux agents (individus, firmes, pays, $\cdots$) ont la possibilité de collaborer afin d'atteindre des situations mutuellement favorables. Une telle situation est typiquement le partage d'une certaine somme d'argent ou le choix d'un profil d'actions ou tout autre choix des agents que tous connaissent et sur lequel tous sont d'accord. Notons que dans ce cadre particulier, les agents poursuivent encore leur intérêt personnel avant tout.

Définition 4.6 *Un problème de marchandage est défini par la donnée de deux joueurs, notés 1 et 2, X un ensemble de points qui correspond à l'ensemble des accords réalisables et $D \in X$ un point représentant la situation de désaccord.*

On peut de plus doter chaque joueur $i = 1, 2$ d'une fonction d'utilité $u_i : X \rightarrow I\!R$ qui associe à chaque accord $x \in X$ une utilité $u_i(x) \in I\!R$. Ainsi, $u_i(x)$ représente le niveau de satisfaction du joueur $i = 1, 2$ si l'accord $x \in X$ est sélectionné. On note $U = \{(u_1(x), u_2(x)) | x \in X\}$ l'ensemble des couples d'utilités en cas d'accord et $d = (u_1(D), u_2(D))$ (ou plus simplement d = (d1, d2)) le couple d'utilités en cas de désaccord. Un jeu de négociation ou de marchandage G est une paire (U, d) qui satisfait les trois conditions suivantes :

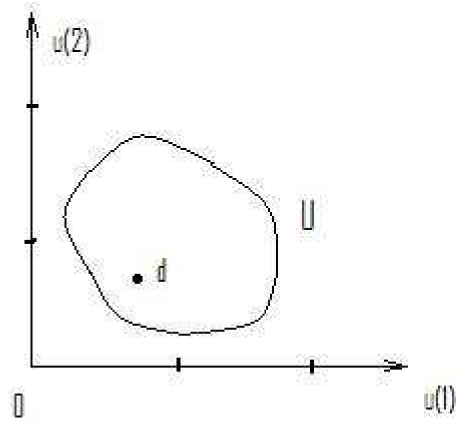
- $d \in U$,
- L'ensemble U est fermé et convexe.

La première condition requiert que le point de désaccord appartient à l'ensemble des utilités atteignables par les deux agents. La deuxième condition garantit la non-vacuité de frontière efficiente au sens de Pareto (c'est l'ensemble des points non dominés au sens de Pareto) et, d'autre part, signifie que pour chaque paire d'éléments de U, toute combinaison convexe de ces deux éléments appartient également à U.

La situation de désaccord joue une rôle important dans le problème de marchandage. Si jamais les agents ne se mettaient finalement pas d'accord, c'est cette situation là qui se produira. son existence influe sur le comportement des agents car comme chaque agent cherche son intérêt personnel, il n'a justement aucun intérêt à se lier d'un accord qui lui rapporte moins que ce point.

Lorsque les deux agents négocient pour s'accorder sur le choix de l'une de ces situations, ou lorsqu'un arbitre impartial doit choisir l'issue de la négociation (ce qui rejoint l'interprétation classique que fait par exemple [Harsanyi 1962] [Harsanyi 1968] de l'analyse stratégique), l'allocation des gains que chacun des deux agents estime pouvoir obtenir de la négociation ne doit dépendre que des deux données suivantes :

- l'ensemble des couples de gains que les deux agents peuvent réaliser au cours de la négociation.
- le couple des gains que chaque agent obtiendra dans la situation où la négociation ne se termine pas par un accord.



Concepts de solution

Étant donnés un ensemble U de couples d'utilités réalisables et un point de désaccord d, le problème de négociation consiste donc à déterminer quelle solution sera le résultat "plausible" de la négociation entre deux joueurs parfaitement rationnels. Dans ce but, la notion de solution au problème de marchandage s'énonce ainsi.

Définition 4.7 *Une solution est une fonction de l'ensemble des problèmes de marchandages $\mathcal{G}$ dans $I\!R^2$, $f : \mathcal{G} \Rightarrow I\!R^2$ telle que $\forall (U, d) \in \mathcal{G}$, $f(U, d) \in U$.*

Certains points particuliers de l'ensemble X servent à construire les solutions de manière géométrique. Le premier point est le point idéal :

Définition 4.8 (Point idéal) $i(U) = (i_1(U), i_2(U))$ *avec*

$$i_1(U) = sup\{x \in I\!R | \exists y \in I\!R \ \text{t.q.} \ (x, y) \in U\}$$

$$i_2(U) = sup\{y \in I\!R | \exists x \in I\!R \ \text{t.q.} \ (x, y) \in U\}$$

Ce point nous permet de définir une notion utile : la fonction de meilleur réponse.

Définition 4.9 *Soit la fonction $g_s(x)$ définie sur l'ensemble des $x \leqslant s_1(U)$ par*

$$g_s(x) = y \ \text{ si } \ (x, y) \ \text{ Pareto domine } \ (U, d),$$

$$= s_2(S) \ \text{ s'il n'existe pas un tel } \ y$$

$g_s(x)$ est le maximum que peut obtenir le joueur 2 si le joueur 1 obtient au moins x.

Un autre point utile est celui de point d'attentes minimales :

Définition 4.10 $m(U) = (m_1(U), m_2(U))$ *avec*

$$m_1(U) = inf\{x \in I\!R | \exists y \in I\!R \ \text{t.q.} \ (x, y) \in U\}$$

$$m_2(U) = inf\{y \in I\!R | \exists x \in I\!R \ \text{t.q.} \ (x, y) \in U\}$$

Il existe une analogie avec la théorie de la décision dans la manière dont ont été caractérisées les solutions. Il s'agit encore d'exhiber un ensemble d'axiomes décrivant un comportement rationnel et de désigner comme solutions acceptables les fonctions satisfaisant cet ensemble d'axiomes. Commençons par exposer les axiomes retenus dans la littérature.

axiome (Pareto optimalité (PO)) $\forall (U,d) \in \mathcal{G}$, $\nexists y \in U$ *tel que* $y > f(U,d)$ *(au sens du préordre défini sur* $I\!R^2$*) et* $y \notin f(U,d)$.

Cet axiome signifie que si la solution considère un certain point comme une issue possible du marchandage et si un deuxième point domine ce dernier au sens de Pareto (en d'autres termes composante par composante) alors la solution doit aussi retenir le deuxième. Que le deuxième point domine le premier signifie qu'il satisfait plus les deux agents participant au marchandage ce qui justifie son acceptation étant donné que chacun des agents recherche à maximiser sa satisfaction. Cet axiome est commun à une majorité de solutions présentées dans la nature. Une autre version imposant une condition moins restrictive à la définition d'un point Pareto optimal est la suivante :

axiome (Pareto optimalité faible POF) $\forall (U,d) \in \mathcal{G}$, $\nexists y \in U$ *tel que* $y \geqslant f(U,d)$ *(au sens du préordre défini sur* $I\!R^2$*)et* $y \notin f(U,d)$.

L'axiome suivant spécifie que les deux agents jouent un rôle symétrique dans le marchandage. Il rejoint en cela l'hypothèse classique que tout ce qui constitue la satisfaction d'un agent est déjà encodé dans sa fonction d'utilité.

axiome (Symétrie (S)) *Soit* $T : I\!R^2 \to I\!R^2$ *tel que* $T(x_1, x_2) = (x_2, x_1)$ *et* $\forall (U,d) \in \mathcal{G}$, $f(U,d) = T(f(U,d))$.

axiome (Indépendance à une transformation affine de l'utilité (ITAU)) *Soit* $A = (A_1, A_2) : I\!R^2 \to I\!R^2$ *tel que les applications* A_1 *et* A_2 *sont de la forme* $c_i x + d_i$ *où* c_i *et* d_i *sont des constantes positives. Alors pour une telle transformation,* $f(A(U), A(d)) = A(f(U,d))$.

Un autre axiome a été proposé dans [Roth 1977] qui se veut la description d'un pendant individuel à la rationalité collective décrite par l'axiome (PO)

axiome (Rationalité Individuelle (RI)) $f(U,d) > d$.

Voici maintenant une autre famille d'axiomes qui se ressemblent beaucoup.

axiome ***Invariance de la solution par rapport aux alternatives non pertinentes (INV)*** *Si* (U,d) *et* (V,d) *sont tels que* $U \subseteq V$ *et* $f(V,d) \in U$ *alors* $f(V,d) = f(U,d)$.

axiome ***Invariance de la solution par rapport aux alternatives non pertinentes autres que le point de désaccord et le point idéal (INVDI)*** *Si* (U,d) *et* (V,d) *sont tels que* $i(U) = i(V)$, $U \subseteq V$ *et* $f(V,d) \in U$ *alors* $f(V,d) = f(U,d)$.

axiome ***Invariance de la solution par rapport aux alternatives non pertinentes autres que le point idéal (INVI)*** *Si* (U,d) *et* (V,d') *sont tels que* $i(U) = i(V)$, $U \subset V$ *et* $f(V,d') \in U$ *alors* $f(V,d') = f(U,d)$.

axiome ***Invariance de la solution par rapport aux alternatives non pertinentes autres que le point des attentes minimales (INVM)*** *Si* (U, d) *et* (V, d') *sont tels que* $m(U) = m(V)$*,* $U \subseteq V$ *et* $f(V, d') \in U$ *alors* $f(V, d') = f(U, d)$*.*

Le premier axiome a été présenté par Nash [Nash 1950]. Cet axiome signifie qu'enlever des points qui font pas partie de la solution ne change pas la dite solution. Dans ce travail, Nash démontre le théorème suivant :

Théorème 4.1 *Il existe une unique solution qui vérifie les axiomes (PO),(S),(ITAU) et (INV).*

Nash présente la solution ainsi caractérisé de manière géométrique. Pour un problème de marchandage (U, d), on considère pour chaque point $(x_1, x_2) \in U$ le produit (aire du rectangle) $(x_1 - d1) * (x_2 - d2)$. Alors $f(U, d)$ est le point de U qui maximise ce produit. Il est intéressant de noter que [Roth 1977] a montré que l'on pouvait remplacer l'axiome (PO) par l'axiome (Ri) dans le théorème précédent sans altérer la solution caractérisée.

L'axiome (INV) a souvent été critiqué (voir [Kalai & Smorodinsky 1975, Luce & Raiffa 1957]). Les axiomes qui le suivent dans la liste précédente ont été introduits pour pallier cela en apportant des changements dans les prémisses de l'axiome. Cet axiome peut en effet paraître trop fort en ce qu'il écarte peut être des points qui influent sur le jugement des agents et ce même s'ils n'apparaissent pas dans la solution. Cela est le cas pour le point idéal que les axiomes (INVDI) et (INVI) se proposent de prendre en compte, mais aussi du point d'attentes minimales qui est utilisé par l'axiome (INVM). [Roth 1977] contient une discussion sur l'effet d'exiger la satisfaction de ces axiomes par les solutions. Il montre en particulier une solution satisfaisant les axiomes (PO),(S),(ITAU) et (INVM) au moyen du théorème suivant.

Théorème 4.2 *Il existe une unique solution satisfaisant les axiomes (PO),(S),(ITAU) et (INVM).*

C'est la fonction M définie par

$$\text{M(U,d)=}\begin{cases} \text{m(U) si m(U)} \in P(U) \\ \text{N(U,m(U)) sinon} \end{cases}$$

où N est la solution de Nash décrite plus haut.

[Kalai & Smorodinsky 1975] propose un axiome assez différent pour remplacer (INV).

axiome (Monotonie (M)) *Si* (U_1, d) *et* (U_2, d) *sont deux problèmes de marchandage tels que* $b_l(U_1) = b_l(U_2)$ *et* $g_{s_1} \leqslant g_{s_2}$ *alors* $f_2(U_1, d) \leqslant f_2(U_2, d)$

Cet axiome spécifie que, si on fixe l'utilité de l'agent 1, le maximum d'utilité que peut atteindre l'agent 2 est augmenté, alors l'utilité de l'agent 2 au point solution devrait aussi augmenter. L'introduction de cet axiome permet à [Kalai & Smorodinsky 1975] de caractériser une solution particulière au moyen du théorème suivant.

Théorème 4.3 *Il existe une unique solution qui vérifie l'axiome (M). Cette solution a l'interprétation géométrique suivante. Pour un problème de marchandage* (U, d)*, on considère la ligne* $L(d, b(U))$ *reliant* d *au point* $b(U)$*. La solution* $f(U, d)$ *est alors le point maximal d'intersection entre* U *et* $L(d, b(U))$*.*

Cette solution est aussi celle pointée par [Luce & Raiffa 1957]. [Roth 1977] note qu'elle satisfait aussi les axiomes (PO),(S) et (ITAU) mais pas (INV).

Une autre solution qui forme selon [Thomson 1994] avec les deux précédentes "trois solutions [qui] jouent un rôle central dans la théorie comme elle apparaît aujourd'hui" est la solution égalitaire (E) présentée par [Kalai 1977].

Définition 4.11 ***La solution égalitaire (E)*** *est définie comme le point maximal de l'ensemble de marchandage d'égales coordonnées, i.e.,* $E_1(U, d) - d_1 = E_2(U, d) - d_2$.

Cette solution se distingue des deux autres principalement par le fait qu'elle requiert des comparaisons interpersonnelles d'utilité. en d'autres termes, les utilités des agents sont supposées être représentées sur la même échelle. Elle est caractérisée par le théorème suivant :

Théorème 4.4 *Il existe une unique solution qui satisfait les axiomes (POF),(S), et (SM). C'est la fonction égalitaire E.*

Toutes les solutions précédentes ont au moins un point en commun. Elles supposent toutes que l'ensemble de marchandage U est convexe. [Conley & Wilkie 1991] expose les raisons qui poussent à assouplir cette hypothèse. Une hypothèse qui permet cet assouplissement est alors de considérer que l'ensemble de marchandage U a une structure d'étoile [3]. Dans le champ économique, cela a l'interprétation d'une disponibilité de l'utilité. On peut alors s'affranchir de l'interprétation probabiliste qui vient avec le cadre de von Neumann et Morgenstern. Plusieurs travaux ont caractérisé des solutions sous cette hypothèse comme [Herrero 1989, Kaneko 1980, Anant *et al.* 1990] ou [Conley & Wilkie 1991] qui étendit les solutions de Kalai et égalitaires à ce nouveau cadre. Notons par ailleurs pour illustrer les différences entre les deux cadres que [Conley & Wilkie 1991] montre que sous l'hypothèse d'un ensemble U en étoile, il n'existe pas de solution satisfaisant à la fois les axiomes (PO) et (S).

Allant toujours plus loin dans la relaxation des hypothèses de départ, certains travaux se sont intéressés a rechercher des solutions ordinales au problème de marchandage, i.e., des solutions qui sont invariantes par n'importe quelle transformation monotone de l'utilité des agents. Cette condition est l'axiome d'invariance ordinale (IO).

Conclusion Cette section présente un panorama des solutions existantes au problème de marchandage. Si l'on voulait définir une solution sur un domaine nouveau (sous de nouvelles hypothèses), la démarche serait de conserver les motivations de ces différentes solutions car chacune exprime une intuition de ce que doit être un comportement rationnel.

4.3 Jeux sous incertitude

4.3.1 Jeux bayésiens

Les jeux que l'on a examinés jusqu'à maintenant partagent tous une même caractéristique. Ce sont des jeux avec information complète. Cela signifie que chaque joueur connaît, au moment de faire son choix, ses stratégies, les stratégies des autres joueurs, les gains résultants de ces stratégies pour lui même ainsi que pour les autres joueurs, etc.

Les situations caractérisées par un manque d'information peuvent être modélisées dans le cadre des jeux à information incomplète. Un jeu avec information incomplète peut être défini comme un jeu dans lequel un des joueurs au moins manque d'une certaine information pertinente sur ses adversaires.

3. Soient $x, y \in U$, $z \in [x, y] \Rightarrow z \in U$.

Une information pertinente est une information qui peut influer sur son choix et donc sur l'issue du jeu. Par exemple, du fait de l'incertitude sur les gains (cartes de l'adversaire cachées), le poker est un jeu à information incomplète.

Le modèle des jeux bayésiens offre une représentation aisée des jeux sous information incomplète. Dans ce modèle, l'information inconnue pour un joueur est représentée, comme si elle dépendait du hasard, par une distribution de probabilités décrivant la vraisemblance de ses valeurs possibles. Harsanyi proposa dans [Harsanyi 1968] le cadre général suivant pour ces jeux. Prenons un jeu avec deux joueurs, les notions présentées se généralisant de manière évidente au cas de n joueurs. Supposons que le joueur 1 ne possède pas une information pertinente sur le joueur 2, comme par exemple les préférences du joueur 2. Notons les valeurs possibles de cette valeur inconnue par $\theta_1, \theta_2, \cdots, \theta_n$. Le joueur 2 peut de la même manière manquer d'une information pertinente sur le joueur 1. Notons alors les valeurs possibles de cette information par $\psi_1, \psi_2, \cdots, \psi_p$. Par exemple, dans une enchère à mises secrètes chaque joueur connaît la valeur qu'il donne à l'objet mis en vente mais ne connaît pas la valeur que les autres joueurs donnent à celui-ci.

Un élément de $\theta_1, \theta_2, \cdots, \theta_n$ est habituellement appelé le "type" du joueur 2. Un type de joueur décrit toute information qui n'est pas une connaissance commune des joueurs. C'est ce que l'on appelle aussi son information privée. Un joueur a donc plusieurs types, un pour chaque valeur que peut prendre son information privée. En fonction de la "vraie" valeur de la variable θ_i, le joueur 2 agira potentiellement d'une manière différente (il sera alors de "type" différent). De manière similaire, une valeur dans $\psi_1, \psi_2, \cdots, \psi_p$ est le type du joueur 1. Une paire de ces variables (ψ_j, θ_i) est un profil de types.

Ce jeu avec information incomplète peut maintenant être réduit à un jeu sous information imparfaite [4] (ce qui est une manière d'interpréter la valeur réelle de l'information privée comme une action du jeu et donc de revenir en terrain connu) en introduisant un troisième joueur, la nature, et en imaginant que la nature agit la première et choisit une paire ψ_j, θ_i. Ce jeu est avec information imparfaite parce que les joueurs ignorent ce choix. Le joueur 2 est seulement mis au courant de θ_i. Symétriquement, le joueur 1 est mis au courant de ψ_j. La gain d'un joueur dépend donc à la fois des actions et des types des autres joueurs. Dans la suite, nous supposons que les ensembles de types et de stratégies sont finis. Il n'est pas suffisant pour définir une stratégie de spécifier une action pour le joueur, on doit spécifier une action pour chaque type.

La philosophie des jeux bayésiens est la suivante : on évite de travailler avec un modèle sous-spécifié en disant qu'à chaque fois qu'une information est connue par une partie des joueurs seulement, cette information est modélisée par une variable (aléatoire) dans le modèle. Le modèle peut inclure des distributions de probabilités qui décrivent les croyances des différents joueurs sur les variables aléatoires. Le modèle ne contient pas les valeurs réelles de ces variables.

La construction du modèle des jeux bayésiens commence par l'affirmation suivante. Toute incertitude dans un jeu peut être modélisée comme une incertitude sur les gains des joueurs. Par exemple, l'incertitude sur la faisabilité d'une stratégie d'un joueur peut être convertie en une incertitude sur ses gains, en spécifiant un gain extrêmement négatif pour le joueur s'il essaye d'utiliser cette stratégie quand elle est supposée infaisable. [Harsanyi 1968] insiste sur le fait que pour analyser le jeu l'on doit prendre, le point de vue d'un observateur qui connait seulement l'information commune à tous les joueurs du jeu. Le type réel de chaque joueur, étant une information privée, doit être traité comme inconnu au cours de l'analyse du jeu. Même si nous prenons le point de vue d'un joueur en particulier qui connait donc son type, nous devons analyser les problèmes de décision qui se posent aux autres joueurs qui ne connaissent pas ce type.

Pour résoudre de tels jeux, les notions présentées pour le cadre de l'information complète sont reprises et adaptées au nouveau cadre plus pauvre en information. Ainsi le concept de dominance peut-

4. Un jeu est avec information imparfaite si un joueur ne peut distinguer entre plusieurs choix de l'un de ses adversaires

il être étendu au cadre des jeux avec information incomplète. Dans un jeu avec information complète, une stratégie s est dominée par une stratégie s' si elle donne un gain plus faible que celui de s' pour toute stratégie de l'adversaire. Dans un jeu avec information incomplète, un joueur ignore le type de son adversaire. Une stratégie s est alors dite être dominée par une stratégie s' si elle donne un gain espéré $Eu_P(s, s{-}i)$ plus bas que celui de s pour toute liste de stratégies ; une liste contient une stratégie pour chaque type.

$$Eu_P(s, s{-}i) = \sum_{i=1}^{i=n} p_{\theta_i} * u_{\theta_i}(s, s{-}i)$$

Dans le même sens nous pouvons définir le concept de meilleure réponse et ensuite d'équilibre de Nash, qui devient maintenant équilibre de Bayes-Nash. Pour un profil donné des stratégies des autres joueurs s_{-i}, une meilleure réponse du joueur i est une stratégie s qui maximise $Eu_P(s, s{-}i)$.

Le joueur 2 choisit une stratégie pour maximiser son gain espéré étant donné ce qu'il croit être la distribution de probabilités sur les types du joueur 1. Le joueur 1 de type j choisit une stratégie pour maximiser son gain espéré étant donné ce qu'il croit être la distribution de probabilités sur les types du joueur 2. Si nous pouvons trouver des stratégies pour chaque type de chaque joueur qui sont les meilleures réponses les unes aux autres alors nous avons un équilibre de Bayes-Nash (version bayésienne de l'équilibre de Nash).

Il existe plusieurs interprétations différentes sur lesquelles s'appuyer pour établir comment les types des joueurs peuvent être générés dans n'importe quel jeu bayésien. On peut supposer que les types sont déterminés par une loterie initiale, d'après une distribution de probabilités a priori P dans le jeu, après quoi chaque joueur est informé de son type mais reste incertain du type des autres joueurs. [Harsanyi 1968] suggéra aussi que l'on peut à la place imaginer l'existence d'un agent différent pour chaque type de chaque participant au jeu, et alors la loterie initiale détermine lequel de ces joueurs participe effectivement au jeu. Cette dernière interprétation possède l'inconvénient de supposer l'existence de plusieurs joueurs fictifs qui n'apparaissent pas réellement dans le jeu. [Harsanyi 1968] suggéra cette interprétation parce qu'elle a l'avantage de vider de son sens toute question sur l'utilité espérée du joueur avant de connaître son type. Pour Harsanyi, le type représente ce que le joueur sait au début du jeu, et donc le calcul de son utilité espérée avant ce type n'apporte rien décisif au joueur.

Une interprétation de la notion de type d'un joueur s'inscrit dans une interprétation plus large des phases du jeu. Le déroulement d'un jeu apporte de fait de plus en plus d'information à un joueur. Cette information modifie largement sa vision de ses intérêts. Le moment dans lequel les joueurs prennent la décision de choisir une stratégie est donc important pour la détermination de l'issue du jeu. On peut diviser les jeux sur une base chronologique en trois phases différentes. Dans la première, les joueurs ne connaissent rien du jeu et ne sont donc pas capables de déterminer leur comportement face aux différentes éventualités d'actions des autres joueurs. En reprenant la notion de type on dit que dans cette phase les joueurs ne connaissent pas leurs types respectifs. C'est la phase ex ante qu'évoque Harsanyi plus haut. La phase suivante commence au moment où les joueurs acquièrent l'information leur permettant de visualiser les différentes situations face auxquelles ils devront faire un choix. Dans ce cas on peut raisonnablement penser que chaque joueur saura prédire comment il se comportera dans ces situations. Ainsi que nous l'avons introduit plus haut, chaque joueur connaît donc son type dans cette phase. Il ne connaît toutefois pas les types des autres joueurs. Il doit se préparer à affronter chacun de ces types. C'est la phase in interim du jeu (voir [Holmstrom & Myerson 1983]). La dernière phase d'importance pour les jeux sous information incomplète est la phase ex post dans laquelle un joueur prend connaissances des actions choisies par les autres joueurs. C'est la phase où toute incertitude disparaît.

Cette phase du jeu permet de définir une autre sorte d'équilibre. C'est l'équilibre ex post : un profil de stratégies, tel qu'aucun joueur n'a une déviation profitable, et ce quel que soit le profil de types des joueurs. C'est une condition forte et de ce fait rarement vérifiée.

Restent à spécifier les distributions de probabilités dont se serviront les joueurs pour calculer leurs gains espérés. Les joueurs sont supposés partager une estimation a priori de la probabilité de toute paire (ψ_j, θ_i).

L'exemple 4.4, illustre les concepts précédents sur une version modifiée du jeu classique du dilemme des prisonniers.

Exemple 4.4 Supposons que les préférences du joueur 2 soient inconnues (du joueur 1) ; il peut être soit un joueur difficile ou accommodant. Un joueur difficile possède les préférences habituelles d'un joueur du dilemme du prisonnier classique, mais un joueur accommodant préfère ne pas avouer à avouer. Supposons maintenant que les préférences du joueur 1 dépendent du fait que le joueur 2 soit difficile ou accommodant. Quand il est face à un joueur difficile, il possède les préférences habituelles d'un joueur du dilemme du prisonniers classique. Face à un joueur accommodant, il préfère ne pas avouer. Alors le jeu peut être représenté par deux matrices de gain, une pour chaque type du joueur 2 (table 4.4). Le joueur 1 ne sait pas quelle est la matrice réelle.

	D	C
D	0,0	7,-2
C	-2,7	5,5

	D	C
D	-2,-2	5,0
C	0,5	7,7

TABLE 4.4 – Dilemme des prisonniers avec information incomplète

Un joueur 2 difficile a une stratégie dominante D alors qu'un joueur 2 accommodant possède une stratégie dominante C. Le joueur 1 ne sait pas s'il se trouve dans la première ou la seconde table décrivant le jeu (une pour chaque type du joueur 2). Cela signifie qu'il ne sait pas s'il sera confronté à la stratégie D ou C. S'il est relativement confiant que ce sera la première, il voudra jouer D, mais dans le cas contraire, il voudra jouer C.

Supposons qu'il pense qu'il y a une probabilité de 0.9 qu'il soit face à un joueur 2 difficile, c'est-à-dire qu'il y a 0.9 de chances que le joueur 2 joue D et 0.1 qu'il joue C. Jouer C mène alors à un gain espéré de $(0.9 * 0) + (0.1 * 5) = 0.5$, et jouer D à un gain espéré de $[0.9 * (-2)] + (0.1 * 7) = -1.1$. Il est clair qu'il vaut mieux jouer D dans ce cas.

4.3.2 Jeux sous incertitude stricte : jeux prébayésiens

Dans les jeux précédents avec information incomplète, on a vu que l'idée pour les modéliser est de représenter l'information manquante par une variable aléatoire dont les joueurs partagent la distribution de probabilité régissant ses valeurs (les types des joueurs). Cette dernière information partagée permet alors de définir les gains espérés des joueurs et de revenir aux notions classiques de solution en les étendant au nouveau cadre.

La disponibilité de cette information a été remise en question dans certains cas ou l'on peut encore modéliser la situation par un jeu avec information incomplète. La justification est celle utilisée pour introduire les cadres faibles en information dans la théorie de la décision. La raison de cette analogie est évidente et a été utilisée dans la partie de planification multi-agent : on peut considérer que chacun des agents doit résoudre un problème de décision particulier. Pour traiter ce genre de situation, un type particulier de jeux a été introduit : les jeux avec incertitude stricte.

Dans les jeux bayésiens, les joueurs utilisent une distribution (a priori) commune pour estimer le gain espéré de leurs stratégies. Les jeux prébayésiens [Ashlagi *et al.* 2006, Ashlagi *et al.* 2007] sont des jeux particuliers où des joueurs bayésiens (au sens où ils ont le même modèle du jeu) manquent de cette information qu'est la distribution de probabilité régissant la variable aléatoire qui modélise le paramètre manquant du jeu. Du point de vue de la représentation, on peut voir les jeux prébayésiens comme des jeux représentés (sous forme normale) par un ensemble de matrices, une pour chaque profil de types des joueurs.

Les jeux prébayésiens sont aussi connus sous l'appellation de jeux informatifs, jeux sans information probabiliste, jeux sous information incomplète avec incertitude stricte et jeux sans distribution de probabilités avec information incomplète (voir [Aghassi & Bertsimas 2006, Holzman & Monderer 2004, Holzman *et al.* 2004, Hyafil & Boutilier 2004]). Le cadre formel est le même que celui des jeux bayésiens. La seule différence est un manque d'information. Si les joueurs d'un jeu prébayésien partagent une distribution de probabilités a priori, on obtient un jeu bayésien. Cette situation mène à une notion assez prudente de solution. Comme nous l'avons vu plus tôt, les notions de dominance et de meilleure réponse reposent classiquement sur celle de gain espérée qui elle-même repose sur la distribution de probabilité régissant les types des joueurs. Sans ces données, [Ashlagi *et al.* 2007] définit seulement le concept de solution MINMAX qui est facilement acceptable et justifiable mais qui peut n'être d'aucun secours à des joueurs devant prendre des décisions critiques et notamment vitales comme le montre l'exemple suivant.

Exemple 4.5 Imaginons que nous soyons dans la savane africaine. Un impala est en train de se reposer dans des buissons quand il aperçoit un lion se déplaçant dans sa direction. Le lion ne semble pas l'avoir vu mais cela changera probablement à très court terme s'il ne bouge pas. L'impala est face à un choix entre s'enfuir maintenant et risquer d'attirer l'attention du lion, ou se cacher plus longtemps en espérant que le lion ne ne le voie pas mais en risquant alors d'avoir moins de chances pour s'échapper si finalement le lion le découvre. L'impala ne sait pas s'il est face à un lion rapide ou lent.

Nous pouvons représenter la situation par le jeu suivant de la table 4.5. Dans la matrice de gauche on peut voir la situation stratégique d'un impala qui est face à un lion rapide (et dans ce cas en s'élançant l'impala ne peut échapper au lion). Dans la matrice de droite se trouve décrite la situation stratégique de l'impala face à un lion de type lent (qu'il pourrait distancer). Il semble évident (pour des raison de temps et de capacité de raisonnement) que l'impala ne peut estimer la distribution de probabilité qui décrit les chances que le lion soit de l'un ou l'autre types.

Lent	A	NA	Rapide	A	NA
bondir	1,0	1,0	bondir	0,1	1,0
se cacher	0,1	1,0	se cacher	0,1	1,0

TABLE 4.5 – Dilemme de l'impala

Dans l'exemple 4.5 le concept de solution MINMAX n'est d'aucun secours pour l'impala puisque les deux stratégies qui s'offrent à lui peuvent toutes deux mener à une même satisfaction minimale de 0.

Equilibre de niveaux de sécurité

Le fait que le joueur manque de l'information qu'est la distribution de probabilité sur les états du monde l'empêche de calculer son gain espéré selon le modèle de von Neumann/Morgenstern. Il ne peut donc pas s'en servir pour comparer ses stratégies. Les joueurs sont réduits au raisonnement prudent suivant. Chaque joueur maximise son gain au pire cas sur les états du monde en supposant que ses adversaires ne changent pas leur stratégies. Les équilibres de niveaux de sécurité ont été définis par

[Aghassi & Bertsimas 2006]. [Ashlagi *et al.* 2007] montre que ce concept d'équilibre prudent a l'avantage d'exister dans chaque jeu prébayésien concave. La définition d'un tel équilibre passe par la définition par une fonction de meilleure réponse en niveaux de sécurité.

Définition 4.12 *On dit qu'une stratégie x_i est optimale pour le type t_i étant donné un profil de stratégies x_{-i} pour les autres joueurs si la valeur maximale du coût au pire cas (défini comme le minimum de l'utilité possiblement atteinte en utilisant x_i par rapport à toutes les valeurs possibles de l'information privée) est atteinte en x_i . Une stratégie du joueur i, x_i, est une meilleur réponse en niveaux de sécurité à x_{-i} si pour chaque type t_i, x_i est optimal pour t_i étant donné x_{-i}.*

Cela permet de définir l'équilibre de manière classique.

Définition 4.13 *Un profil de stratégies $x = (x_i)_{i \in N}$ est appelé un équilibre en niveaux de sécurité si pour chaque joueur i, x_i est une meilleure réponse en niveaux de sécurité à x_{-i}.*

On s'écarte donc de la notion d'équilibre de Bayes-Nash. La formulation de l'équilibre ressemble pourtant à celle d'un équilibre de Nash dans un jeu sous information complète. Et de fait, un équilibre en niveaux de sécurité dans un jeux prébayésien contenant un unique état du monde possible (qui est de manière évidente un jeu avec information complète) est un équilibre de Nash pour ce jeu. C'est une propriété appréciable dans une démarche de généralisation des concepts de solution classiques à des cadres de plus en plus pauvres en information.

Équilibre ex-post

La solution la plus utilisée pour les jeux prébayésiens est celle de l'équilibre ex-post. Examinons cette notions sur l'exemple suivant :

Exemple 4.6 On considère deux joueurs 1 et 2. On suppose que le joueur 1 possède un seul type mais que le joueur 2 est soit du type A soit du type B.

	s_2	$s_2^{'}$
s_1	(5, 0)	(3,1)
$s_1^{'}$	(0, 0)	(4,2)

TABLE 4.6 – 2 EST DE TYPE A

	s_2	$s_2^{'}$
s_1	(2, 2)	(0,0)
$s_1^{'}$	(3, 3)	(5,2)

TABLE 4.7 – 2 EST DE TYPE B

Dans ce jeu, on peut voir que le joueur 2 aura $s_2^{'}$ comme stratégie dominante s'il est de type A et s_2 comme stratégie dominante s'il est de type B. Tenant compte de cela, le joueur 1 aura intérêt à jouer la stratégie $s_1^{'}$ dans les deux cas. Le profil de stratégies comprenant $s_1^{'}$ pour le joueur 1 et pour le joueur 2 $s_2^{'}$ s'il est de type A et s_2 s'il est de type B est un équilibre ex-post.

L'exemple suivant montre que l'existence d'un équilibre ex-post, condition forte, est rarement vérifiée.

	s_2	$s_2^{'}$
s_1	(5, 2)	(3,0)
$s_1^{'}$	(0, 0)	(2,2)

TABLE 4.8 – 2 EST DE TYPE A

	s_2	$s_2^{'}$
s_1	(2, 2)	(0,0)
$s_1^{'}$	(3, 0)	(5,2)

TABLE 4.9 – 2 EST DE TYPE B

Exemple 4.7 On considère deux joueurs 1 et 2. On suppose que le joueur 1 possède un seul type mais que le joueur 2 est soit du type A soit du type B.

Dans ce jeu, on peut voir que le joueur 1 a s_1 comme stratégie strictement dominante s'il est de type A et $s_1^{'}$ comme stratégie dominante s'il est de type B. Il jouera donc une stratégie différente selon le type du joueur 2. Or le joueur 1 ne connaît pas le type du joueur 2. Il n'existe donc pas d'équilibre ex-post.

Dans le premier exemple, si les joueurs suivent les recommandations de l'équilibre ex-post alors ils seront dans une situation qu'ils ne pourront regretter. En effet, les joueurs ne connaissent que leur propre type avant de choisir leur stratégie. Supposons alors que le joueur 2 soit de type A. Il jouera $s_2^{'}$. Le joueur 1 jouera lui la stratégie $s_1^{'}$. Le profil obtenu est un équilibre dans la matrice correspondante au profil de types des joueurs. Il en va de même si le joueur 2 est de type B. Maintenant dans le deuxième exemple, au moment de jouer le joueur 1 ne connaît pas le type du joueur 2. Et selon ce type, il a intérêt à jouer une stratégie différente. Il n'existe donc pas pour lui de stratégie gagnante quel que soit le type révélé dans la phase ex-post de l'autre joueur. A fortiori il n'existe donc pas de profil de stratégies qui soit un équilibre ex-post.

Médiateur

Afin de répondre au problème de rareté des équilibres ex-post qui sont pourtant les équilibres les plus sûrs (et profitables) pour les joueurs, [Ashlagi *et al.* 2007] proposent d'introduire la notion de médiateur. Un cas spécial de médiateur est décrit dans [Kalai & Rosenthal 1996].

Un médiateur est une entité sur laquelle les joueurs peuvent compter et à qui ils délèguent leurs actions. Précisons qu'un médiateur se borne à ce rôle de délégué et ne peut en aucun cas obliger un joueur à adopter une stratégie quelconque. De plus le joueur est libre de ne pas faire appel au médiateur durant le jeu. De tels médiateurs ont été étudiés dans le cas de jeu à information complète dans [Monderer & Tennenholtz 2006], [Rozenfeld & Tennenholtz 2007]. Le gain dans le jeu initial d'un équilibre dans le jeu avec médiateur est appelé un équilibre médié.

Afin d'illustrer le concept de médiateur, [Monderer & Tennenholtz 2006] considèrent le cas du dilemme des prisonniers.

	D	C
D	(-8, -8)	(0,-15)
C	(-15, 0)	(-1,-1)

Dans cet exemple, rappelons que le seul équilibre se produit quand les deux joueurs choisissent de trahir (profil (D, D)). Toutefois, on voit facilement que cet équilibre n'est optimal en terme de gains pour aucun des joueurs. Considérons maintenant un médiateur (en qui les joueurs peuvent avoir confiance). Supposons que si les deux joueurs font tous deux appels au médiateur, alors celui-ci exécutera la stratégie C au nom de chaque joueur. Si en revanche seul l'un des deux joueurs fait appel à ses services, alors il jouera la stratégie D au nom de ce joueur.

	D	C	M
D	(-8,-8)	(0,-15)	(-8,-8)
C	(-15, 0)	(-1,-1)	(-15,0)
M	(-1,-1)	(0,-15)	(-8,-8)

TABLE 4.10 – DILEMME DES PRISONNIERS AVEC MÉDIATEUR

Analysons maintenant ce nouveau jeu. Il existe un équilibre qui assure un vecteur de gain $(4, 4)$. Dans cet équilibre, les deux joueurs font appel au médiateur. Notons pour finir que cet équilibre possède la propriété appréciable d'être stable par rapport aux déviations par des coalitions de joueurs.

[Ashlagi *et al.* 2007] donnent l'exemple suivant pour illustrer l'intérêt de l'introduction d'un médiateur dans un jeu prébayésien. La particularité de cet exemple est qu'il ne présente pas d'équilibre ex-post.

A	a	b
a	(5,2)	(3,0)
b	(0,0)	(2,2)

B	a	b
a	(2, 2)	(0,0)
b	(3, 0)	(5,2)

TABLE 4.11 – JEUX AVEC INFORMATION INCOMPLÈTE SANS MÉDIATEUR

On doit spécifier le comportement du médiateur une fois introduit. La différence est qu'ici le joueur doit aussi transmettre son type au médiateur au moment où il fait appel à lui (information qui est, rappelons-le, privée). Supposons alors que si les deux joueurs font appel au médiateur, le médiateur appliquera le profil (a, a) si le joueur 2 est de type A et (b, b) dans l'autre cas. Si seulement le joueur 1 fait appel au médiateur, le médiateur appliquera la stratégie a en son nom. Si seulement le joueur 2 fait appel au médiateur, alors celui-ci appliquera a en son nom s'il transmet le type B et b dans le cas restant.

L'introduction du médiateur génère alors le jeu 4.12.

A	m -A	m- B	a	b
m	(5,2)	(2,2)	(5,2)	(3,0)
a	(3,0)	(5,2)	(5,2)	(3,0)
b	(2,2)	(0,0)	(0,0)	(2,2)

B	m -A	m- B	a	b
m	(2,2)	(5,2)	(2,2)	(0,0)
a	(0,0)	(2,2)	(2,2)	(0,0)
b	(5,2)	(3,0)	(3,0)	(5,2)

TABLE 4.12 – JEUX AVEC INFORMATION INCOMPLÈTE SANS MÉDIATEUR

Ce nouveau jeu a la particularité de contenir un équilibre ex-post. Il est réalisé si les deux joueurs font appel au médiateur et que le joueur 2 indique son type.

Minimisation de regret

Les équilibres de Minimax-regret ont été définis dans le cadre des jeux prébayésiens par [Hyafil & Boutilier 2004] avec l'argument (commun aux approches précédemment présentées) qu'en l'absence d'information probabiliste, on doit comparer les stratégies en utilisant un critère qualitatif (qui ne repose pas sur une telle information). La minimisation de regret en elle-même est un critère qui a été utilisé à la fois dans le domaine de décision sous incertitude stricte [Borodin & El-Yaniv 1998] que celui de la prise de décision sous information imprécise [Claus & Boutilier 1997]. Intuitivement, pour un choix de stratégies et un profil de types donnés pour les adversaires, le regret d'un joueur i de type θ_i par rapport au choix d'une stratégie s_i est la perte que ressent i en jouant cette stratégie au lieu d'une stratégie optimale.

Voici la définition de la fonction de meilleure réponse en minmax regret.

Définition 4.14 *On dit que (y_i) est une stratégie de minimax regret pour t_i étant donné s_{-i} si la valeur minimale de $MR(x_i, t_i, s_{-i})$ est atteint en y_i. Une stratégie s_i est une meilleure réponse en minimax regret pour s_{-i} si pour chaque type t_i, s_i est une stratégie minimax regret à t_i étant donné s_{-i}.*

L'équilibre se définit alors de la manière suivante :

Définition 4.15 *s est un équilibre de minimax regret si pour chaque joueur i s_i est une meilleure réponse en minimax regret à s_{-i}.*

[Hyafil & Boutilier 2004] expose le critère de minimisation de regret dans le cas où l'incertitude porte sur l'utilité des joueurs. Cela peut sembler être un cas particulier par rapport à ceux déjà présentés, mais il s'y ramène en fait. Pour cela on peut se rappeler le point de vue de Harsanyi [Harsanyi 1968] dans sa définition des jeux bayésiens. Il considère en effet qu'on doit pouvoir toujours ramener l'incertitude dans le jeu à une incertitude sur les utilités des joueurs. L'existence des équilibres de minimisation de regret a été mise en évidence par [Hyafil & Boutilier 2004] dans le cas de stratégies mixtes. Malgré ce résultat et le caractère intuitif du critère, celui-ci est difficile à mettre en oeuvre à cause de la complexité des espaces de décision et d'utilité.

4.4 Conclusion

Dans ce chapitre, nous avons exposé les concepts de base de la théorie des jeux. Nous avons adopté une démarche similaire à celle entreprise dans le chapitre sur la théorie de la décision. Nous avons en effet examiné les différents concepts de solutions disponibles à mesure que l'on prive l'agent joueur de l'information qu'il possède sur le résultat de ses actions dans le jeu et sur les autres agents qui y participent. Ainsi nous avons parcouru successivement les champs des jeu sous certitude, les jeux sous information incomplète (où les joueurs possèdent quand même une information probabiliste sur les différents profils de types des joueurs) et la théorie des jeux sous incertitude stricte ou jeux prébayésiens.

Deuxième partie

Contribution

Chapitre 5

Contribution à la théorie de la décision sous incertitude

5.1 Introduction

Nous avons présenté plusieurs modèles pour la prise de décision dans le cas où l'agent n'est pas certain du résultat de ses actions. En résumé, on peut établir les correspondances suivantes :

Cadre Probabiliste Ensemble d'état+ probabilités

Cadre Qualitatif Ensemble d'état+ niveaux de plausibilité

Incertitude Stricte Ensemble d'état

Ignorance Pas d'ensemble d'état

Un modèle de décision est avant tout un modèle qui entreprend de caractériser les meilleures actions en tenant compte de l'information disponible au moment de la prise de décision. Cette affirmation reste valable que l'information disponible soit complète ou pas. En particulier, si l'information est incomplète, la décision doit s'appuyer sur une théorie permettant d'estimer les attentes de l'agent. Alors que les modèles basés sur la théorie de l'utilité espérée sont intéressants de part le fait que les actions qu'ils choisissent satisfont des exigences qui semblent naturelles, ils ne sont d'aucune aide pour l'agent quand aucune distribution de probabilités sur les états du monde n'est disponible. Cela est la motivation essentielle à l'introduction des trois autres familles de modèles. Avec à chaque fois ce même argument de non disponibilité de l'information pour passer d'un modèle à un autre qui a besoin de moins d'hypothèses pour caractériser une décision rationnelle. Notons par ailleurs qu'en même temps que l'on se débarrasse du formalisme probabiliste puis de la notion même d'état du monde, on soulage l'agent de la charge de représenter toute cette information qui peut être arbitrairement lourde.

Dans cette partie, nous nous intéressons au le cas le moins informatif qui est celui de la décision sous ignorance complète. Nous suivons pour ce faire une démarche axiomatique en introduisant d'abord deux axiomes naturels que les critères de décision devront satisfaire pour être considérés comme rationnels dans le cadre étudié. Nous montrons alors que les critères d'optimalité retenus prennent en compte dans leur choix uniquement les éléments extrémaux des ensembles de conséquences associés aux actions. Nous comparons ensuite les axiomes introduits à ceux déjà présents dans la littérature (et présentés dans la section 2.6.3).

La suite de ce chapitre est organisée comme suit. Dans la section 5.2 nous commençons par quelques préliminaires formels. Dans la section 5.3 nous présentons les deux axiomes nouveaux que nous introduisons. Dans la section 5.4 nous présentons un théorème qui caractérise la famille de critères induite par la satisfaction de ces deux critères ainsi qu'un deuxième théorème qui raffine le premier dans le sens où il s'attache au cas de préférences strictes. Nous repassons en revue à la section 5.5 quelques exemples de critères intéressants et examinons leur appartenance à la famille que nous caractérisons. Nous comparons ensuite à la section 5.6 nos axiomes à ceux présentés dans les travaux existants en explicitant plus précisément les relation logiques qui les lient avant de conclure à la section 5.7.

5.2 Cadre formel

Rappelons d'abord le cadre formel dans lequel nous pourrons exprimer les notions résultats et preuves. Ce cadre est celui de la décision sous ignorance complète présenté dans la section 2.6.

Définition 5.1 *Un cadre de décision est déterminé par $(\mathcal{A}, \mathcal{C}, \mathcal{R}, \geq)$, où l'on considère $\mathcal{A}$ un ensemble d'actions, $\mathcal{C}$ un ensemble de conséquences, R une application de $\mathcal{A}$ vers $2^{\mathcal{C}} \setminus \{\emptyset\}$ qui associe à chaque action $a \in \mathcal{A}$ un sous-ensemble non vide $R(a)$ de $\mathcal{C}$ décrivant les conséquences possibles de a. Nous supposons que $\mathcal{C}$ est ordonné au moyen d'une relation d'ordre total $\geq$, i.e., une relation réflexive, anti-symétrique et transitive. $\geq$ exprime les préférences du décideur sur $\mathcal{C}$.*

Un problème de décision est un sous-ensemble non vide A de $\mathcal{A}$.

Définition 5.2 *Un critère de décision dans le cadre de décision $(\mathcal{A}, \mathcal{C}, \mathcal{R}, \geq)$ est une fonction de choix associant à chaque ensemble non vide A de $\mathcal{A}$ un sous-ensemble $\hat{A}$ de A (formellement, c'est une application).*

La donnée d'un critère de décision dans le cadre $(\mathcal{A}, \mathcal{C}, \mathcal{R}, \geq)$ permet de résoudre n'importe quel problème de décision A : les éléments optimaux sont ceux dans $\hat{A}$.

Une conséquence d'une action $a \in A$ ne sera pas forcément associée à un nombre réel. La seule hypothèse exigée sur l'ensemble de conséquences $\mathcal{C}$ est que l'agent peut comparer n'importe quelles alternatives $c, c^{'} \in \mathcal{C}$ au moyen de $\geq$.

Si $a_1, a_2 \in A$ et $a_1 \in \hat{A}$, nous disons que a_1 est préféré ou indifférent à a_2 par rapport à A et le symbolisons par $a_1 \trianglerighteq a_2[A]$.

Si $a_1, a_2 \in A$, et $a_1, a_2 \in \hat{A}$ ou $a_1, a_2 \in A \setminus \hat{A}$, nous disons que a_1 et a_2 sont équivalents par rapport à A, ce qui est symbolisé par $a_1 \diamond a_2[A]$.

Similairement, si $a_1, a_2 \in A$, $a_1 \in \hat{A}$ et $a_2 \in A \setminus \hat{A}$, a_1 est dit être préféré à a_2 par rapport A ; la relation est symbolisée par $a_1 \triangleright a_2[A]$.

En particulier, pour des problèmes de décision contenant seulement les actions a_1, a_2 et noté par $[a_1, a_2]$, nous simplifions les notations comme suit :
$a_1 \diamond a_2$ si $a_1 \diamond a_2[a_1, a_2]$,
$a_1 \trianglerighteq a_2$ si $a_1 \trianglerighteq a_2[a_1, a_2]$,
$a_1 \triangleright a_2$ si $a_1 \triangleright a_2[a_1, a_2]$.

Nous notons $minR(a)$ le minimum de $R(a)$ et $maxR(a)$ le maximum de $R(a)$ par rapport à la relation de préférence sur les conséquences, i.e., $minR(a) = c$ ssi $\nexists c^{'} \in R(a)$ t.q. $c > c^{'}$, et $maxR(a) = c$ ssi $\nexists c^{'} \in R(a)$ t.q. $c^{'} > c$.

5.3 Propriétés des cadres et des critères de décision

Nous faisons les hypothèses suivantes sur les problèmes de décision étudiés :

Hypothèse 1 (A) *$\forall A \subseteq \mathcal{A}$, si $\mathcal{A} \neq \emptyset$ alors $\hat{A} \neq \emptyset$.*

Nous considérons seulement des problèmes qui possèdent un ensemble optimal non vide.

Hypothèse 2 (B) *$\forall a \in \mathcal{A}$, $minR(a)$ et $maxR(a)$ existent.*

Pour chaque action, l'ensemble des conséquences est noethérien pour $\geq$ et pour $\leq$. Cela signifie que nous supposons que pour chaque action, l'ensemble de conséquences possibles contient toujours un élément qui est minimal et un élément qui est maximal par rapport aux préférences de l'agent. Cette hypothèse nous évite de traiter des cas induisant des chaînes infinies de comparaisons. Ces deux hypothèses délimitent l'ensemble des problèmes étudiés. elles écartent les cas dégénérés dans lesquels l'ensemble optimal ne peut être dérivé.

Hypothèse 3 (C) *$\forall C' \subseteq \mathcal{C}$, si $C' \neq \emptyset$, $\exists a \in \mathcal{A}$ t.q. $R(a) = C'$.*

Cette hypothèse stipule que pour chaque ensemble de conséquences on peut envisager une action dont les conséquences possibles sont exactement celles contenues dans cet ensemble. Pour chaque sous-ensemble $C' = \{c_1, c_2, \ldots\}$ de $\mathcal{C}$, on note $[c_1, c_2, \ldots]$ une action a de $\mathcal{A}$ t.q. $R(a) = C'$. Elle exprime une sorte de propriété de densité de l'ensemble des conséquences possibles. Cette hypothèse est similaire à celle sous jacente à l'utilisation des probabilités.

Un critère d'optimalité est une règle définissant $\hat{A}$ pour chaque $A \in \mathcal{A}$. Nous allons maintenant présenter les propriétés que nous voudrions voir vérifiées par les critères et nous déduirons ensuite la forme des critères retenus.

axiome (α) *Si $A_1 \subseteq A_2$ et $A_1 \cap \hat{A}_2$ est non vide, alors $\hat{A}_1 = A_1 \cap \hat{A}_2$.*

En d'autres termes, si une action est supposée optimale dans un ensemble donné d'alternatives et si par la suite l'ensemble des actions disponibles est réduit mais que cette action est encore disponible, alors elle est encore optimale dans le nouvel ensemble (et aucune action ne "devient" optimale). C'est une propriété que l'on retrouve souvent dans les axiomatisations des critères de décision sous incertitutde. Elle est nécéssaire pour pouvoir obtenir que le choix de l'agent soit constitué par des comparaisons deux à deux des actions, ce qui le processus de séléction le plus naturelle.

axiome (β) *Soit A un problème de décision, $a, a' \in A$ telles que $\forall c \in R(a), \forall c' \in R(a'), c \geqslant c'$. Considérons $\forall c'' \in C$ et $a_1, a'_1 \in A$ t.q. $R(a_1) = R(a) \cup \{c''\}$ et $R(a'_1) = R(a') \cup \{c''\}$. On a alors $a_1 \trianglerighteq a'_1$.*

Si deux actions sont telles qu'elles partagent une conséquence commune et que toutes les autres conséquences de la première sont préférées ou indifférentes à toutes les conséquences de la suivante alors la première action est préférée ou indifférente à la deuxième. La rationalité de cet axiome est la suivante. Supposons que le décideur est confronté au choix entre deux actions telles que les conséquences de la première soient préférées à toutes celles de la seconde (ce qui la rend intuitivement préférée). Plus tard, il est informé que ces actions partagent une conséquence commune qui ne faisait pas partie de sa vision initiale (qui était donc partielle). Alors cette information ne devrait pas changer dramatiquement sa préférence. Cet axiome est cohérent avec l'hypothèse d'ignorance complète. Celle-ci stipule que l'agent ne dispose pas d'une distribution de probabilités sur les résultats possibles d'une action. Voyons ce que cela signifie sur l'exemple suivant. Supposons que l'agent doit choisir entre deux actions a et b

(d'après l'axiome précédent, le choix de l'agent reviendra toujours finalement à des choix entre des couples d'actions). Supposons pour pouvoir aisément les comparer que les résultats des actions sont des réels. L'action a implique un résultat dans l'ensemble $\{4, 5\}$ (sans que l'on dispose d'une distribution de probabilités sur l'occurence éventuelle de ces résultats). Sous la même hypothèse, l'action b implique un résultat dans l'ensemble $\{1, 2\}$. On peut observer qu'en l'état, on peut raisonnablement supposer que l'agent préfèrera l'action a car si on exécute simultannément les deux actions et quelque soit le résultat obtenu pour chacune, la conséquence de a sera préférée à celle de b. Maintenant si on ajoute aux ensembles de résultats une même conséquence (par exemple 6) pour obtenir respectivement les actions a' et b', la préférence de l'agent entre les deux action n'est plus aussi intuitive. On peut dire cependant que cette transformation n'apporte pas de raison particulière pour inverser complètement la première préférence et affirmer que b serait strictement préférée à a. L'axiome s'appuie sur cette idée pour poser que a' sera préférée ou indifférente à b'. L'idée derrière l'introduction de cet axiome est aussi opérationnelle. On y part en effet d'une situation simple dans laquelle on peut ituitivement conclure et on induit alors une relation de préférence sur une situation plus complexe. On peut alors espérer au moyen de cet axiome conclure dans toutes les situations possibles en partant des préférences de l'agent sur les conséquences (que l'on identifie aux actions ayant une seule conséquence possible).

Le lemme suivant donne une conséquence intéressante de cet axiome ($\boldsymbol{\beta}$). Il stipule que, quand on considère une action et son ensemble associé de conséquences, il est suffisant pour décider de son optimalité de considérer seulement les éléments extrémaux de cet ensemble, i.e., le couple formé par la pire conséquence et la meilleure conséquence (leur existence est assurée par l'hypothèse (B)).

Lemme 5.1 *Supposons un critère d'optimalité satisfaisant l'axiome (β).*
Soit $a \in \mathcal{A}$. $a \diamond [minR(a), maxR(a)]$.

Preuve
Soit $a \in \mathcal{A}$. Définissons a_1, a_n, a'_1, a'_n par $R(a_1) = \{minR(a)\}$, $R(a_n) = \{maxR(a)\}$, $R(a'_1) = R(a) \setminus \{minR(a)\}$, $R(a'_n) = R(a) \setminus \{maxR(a)\}$. Alors nous avons $\forall c \in R(a_1), \forall c' \in R(a'_n), c' \geqslant c$. Puisque $R([minR(a), maxR(a)]) = R(a_1) \cup \{maxR(a)\}$ et $R(a) = R(a'_n) \cup \{maxR(a)\}$, l'axiome ($\boldsymbol{\beta}$) donne $a \trianglerighteq [minR(a), maxR(a)]$. D'un autre côté, nous avons $\forall c \in R(a_n), \forall c' \in R(a'_1), c \geqslant c'$. Puisque $R([minR(a), maxR(a)]) = R(a_n) \cup \{minR(a)\}$ et$R(a) = R(a'_1) \cup \{minR(a)\}$, l'axiome ($\boldsymbol{\beta}$) donne $[minR(a), maxR(a)] \trianglerighteq a$. Finalement, nous pouvons conclure $a \diamond [minR(a), maxR(a)]$. □

Le corollaire suivant est obtenu comme une conséquence directe du lemme 5.1 :

Corollaire 5.1 *Soient $a, a^{'} \in A$. Si $R(a) = R(a^{'})$ alors $a \diamond a^{'}$.*

Preuve
Par le lemme précédent, il vient directement que :

$$a \diamond [minR(a), maxR(a)] \diamond a'.$$

□

Ce corollaire montre que la seule information pertinente pour la comparaison d'actions est la donnée de leurs ensembles de conséquences. Ainsi, deux actions ayant le même ensemble de conséquences ne peuvent être distinguées par le critère retenu. Cela implique que les actions peuvent être identifiées à leurs ensembles de conséquences.

Finalement, le lemme suivant stipule essentiellement que la comparaison entre deux actions dépend seulement de ces actions (et de leurs conséquences), et non des autres actions disponibles.

Lemme 5.2 *Soit un critère d'optimalité qui satisfait l'axiome* $(\boldsymbol{\alpha})$*. Alors nous avons :*

1. *Si* $a_1 \diamond a_2$*, alors* $a_1 \diamond a_2[A]$ *pour tout* A *t.q.* $a_1, a_2 \in A$*.*
2. *Si* $a_1 \trianglerighteq a_2[A]$ *pour un certain* A*, alors* $a_1 \trianglerighteq a_2$*.*
3. *Si* $a_1 \rhd a_2[A]$ *pour un certain* A*, alors* $a_1 \rhd a_2$*.*

Preuve
Prouvons d'abord le point 1. Si a_1, a_2 ne sont pas équivalents par rapport à A, supposons sans perte de généralité $a_1 \in \hat{A}$, $a_2 \in A \setminus \hat{A}$. Puisque $\{a_1, a_2\}$ intersecte $\hat{A}$, l'ensemble optimal pour $\{a_1, a_2\}$ consiste en un élément a_1 par l'axiome $(\boldsymbol{\alpha})$. Cela contredit l'hypothèse. Les points 2 et 3 suivent par des applications de l'axiome $(\boldsymbol{\alpha})$. □

5.4 Caractérisation des critères d'optimalité

Dans cette section nous énonçons et prouvons un théorème qui permet de caractériser la famille de critères d'optimalité qui satisfont à la fois l'axiome $(\boldsymbol{\alpha})$ et l'axiome $(\boldsymbol{\beta})$.

Théorème 5.1 *Sous les hypothèses (A), (B) et (C), un critère d'optimalité satisfait les axiomes* $(\boldsymbol{\alpha})$ *et*$(\boldsymbol{\beta})$ *si et seulement si il existe un pré-ordre total* $\succcurlyeq$ *sur l'ensemble des paires de conséquences* (m, M) *avec* $M \geq m$ *satisfaisant :*

si $m_1 \geq m_2$ *and* $M_1 \geq M_2$ *alors* $(m_1, M_1) \succcurlyeq (m_2, M_2)$*,*

tel que les ensembles optimaux sont déterminés par $\forall A \in \mathcal{A}$ *(1)*

$\hat{A} = \{a | (minR(a), maxR(a)) \succcurlyeq (minR(a'), maxR(a')), \forall a' \in A\}$*. (2)*

Preuve
Afin de prouver la nécessité, nous supposons que les axiomes $(\boldsymbol{\alpha})$ et $(\boldsymbol{\beta})$ sont vérifiées par le critère et nous voulons prouver qu'il existe un pré-ordre total qui satisfait (1) et (2). Dans ce but nous définissons une relation (que nous allons montrer être un pré-ordre total) et nous prouvons qu'elle convient.

Nous supposons donc l'existence d'un critère d'optimalité qui satisfait les axiomes $(\boldsymbol{\alpha})$ et $(\boldsymbol{\beta})$.

Définissons $\succcurlyeq$ sur les paires de conséquences (m, M) comme suit :

$$(m_1, M_1) \succcurlyeq (m_2, M_2) \text{ ssi } [m_1, M_1] \trianglerighteq [m_2, M_2].$$

Prouvons d'abord que la relation $\succcurlyeq$ est un préordre total, i.e., une relation qui est réflexive, transitive et totale.

Pour la réflexivité, nous devons prouver que pour chaque paire (m, M) nous avons

$$(m, M) \succcurlyeq (m, M).$$

Par définition de la relation $\succcurlyeq$, ceci revient à montrer que si l'on considère deux actions a_1 et a_2 telles que

$$R(a_1) = R(a_2) = \{m, M\}$$

alors

$$a_1 \trianglerighteq a_2$$

.

D'après le corollaire 5.1, si $R(a_1) = R(a_2)$ nous avons

$$a_1 \diamond a_2.$$

Soit A le problème de décision dont les actions sont a_1 et a_2. Puisque d'après l'hypothèse (A)

$$\hat{A} \neq \emptyset$$

alors par définition de $\diamond$ et puisque $A = \{a_1, a_2\}$, nous avons

$$\hat{A} = A$$

Il s'en suit que

$$a_1 \in \hat{A}$$

donc par définition de $\trianglerighteq$

$$a_1 \trianglerighteq a_2.$$

Pour prouver la transitivité, supposons

$$(m_1, M_1) \succcurlyeq (m_2, M_2)$$

et

$$(m_2, M_2) \succcurlyeq (m_3, M_3).$$

Soit A le problème de décision dont les actions sont $[m_1, M_1], [m_2, M_2], [m_3, M_3]$. D'après l'hypothèse (A), nous avons $\hat{A} \neq \emptyset$.

Supposons que $[m_3, M_3]$ est la seule action optimale, i.e. $\hat{A} = \{[m_3, M_3]\}$, alors

$$[m_3, M_3] \triangleright [m_2, M_2][A].$$

Par le lemme 1, nous avons que

$$[m_3, M_3] \triangleright [m_2, M_2]$$

alors

$$[m_3, M_3] = \hat{A}'$$

où $A' = \{[m_2, M_2], [m_3, M_3]\}$. Mais par hypothèse

$$(m_2, M_2) \succcurlyeq (m_3, M_3)$$

ce qui donne par définition $[m_2, M_2] \trianglerighteq [m_3, M_3]$ et alors $[m_2, M_2] \in \hat{A}'$. Contradiction.

Supposons maintenant que

$$[m_2, M_2] \in \hat{A} \text{ mais } [m_1, M_1] \notin \hat{A}.$$

Nous avons par hypothèse

$$(m_1, M_1) \succcurlyeq (m_2, M_2).$$

Ce qui par défintion donne $[m_1, M_1] \trianglerighteq [m_2, M_2]$ et alors $[m_2, M_2]\hat{A}$ implique $[m_1, M_1]\hat{A}$. Contradiction.

En conséquence, $[m_3, M_3]$ ne peut pas être le seul élément dans $\hat{A}$ et si $[m_2, M_2]$ est optimal alors il en est de même pour $[m_1, M_1]$.

Nous devons conclure que

$$[m_1, M_1] \in \hat{A},$$

qui donne

$$[m_1, M_1] \trianglerighteq [m_3, M_3][A]$$

Alors par le lemme 5.2

$$[m_1, M_1] \trianglerighteq [m_3, M_3]$$

Donc par définition de $\succcurlyeq$,

$$(m_1, M_1) \succcurlyeq (m_3, M_3).$$

La relation $\succcurlyeq$ est alors transitive.

D'après l'hypothèse (A), puisque chaque problème de décision $\{[m_1, M_1], [m_2, M_2]\}$ contient au moins un élément optimal, nous avons $(m_1, M_1) \succcurlyeq (m_2, M_2)$ ou $(m_2, M_2) \succcurlyeq (m_1, M_1)$, ce qui signifie que la relation $\succcurlyeq$ est totale.

Afin de prouver que ce préordre satisfait la condition (1) du théorème, supposons maintenant que $m_1 \geqslant m_2, M_1 \geqslant M_2$ et montrons que $(m_1, M_1) \succcurlyeq (m_2, M_2)$.

Si $m_1 \geq m_2$ alors l'axiome ($\boldsymbol{\beta}$) nous permet de conclure en ajoutant M_2, que

$$[m_1, M_2] \trianglerighteq [m_2, M_2]$$

De plus, puisque $M_1 \geq M_2$, l'axiome ($\boldsymbol{\beta}$) nous permet de conclure en ajoutant m_1, que

$$[m_1, M_1] \trianglerighteq [m_1, M_2]$$

Donc par définition du préordre

$$(m_1, M_1) \succcurlyeq (m_1, M_2) \text{ et } (m_1, M_2) \succcurlyeq (m_2, M_2).$$

La transitivité du préordre implique que

$$(m_1, M_1) \succcurlyeq (m_2, M_2).$$

Prouvons maintenant que le deuxième point du théorème est aussi satisfait.

Supposons que $a \in \hat{A}$. Donc $a \trianglerighteq a'[A]$ pour chaque $a' \in A$, et donc par le lemme 5.2 nous avons $a \trianglerighteq a'$ pour chaque $a' \in A$.

Par le lemme 5.1 nous avons

$$[minR(a), maxR(a)] \diamond a$$

et
$$[minR(a'), maxR(a')] \diamond a'.$$

Considérons maintenant le problème

$$A_1 = \{a, a', [minR(a), maxR(a)], [minR(a'), maxR(a')]\}.$$

Par le lemme 5.2 (point 1) nous avons

$$[minR(a), maxR(a)] \diamond a[A_1] \text{ et } [minR(a'), maxR(a')] \diamond a'[A_1].$$

Si
$$a' \rhd a[A_1](E)$$

alors par le lemme 5.2 (point 3)
$$a' \rhd a,$$

ce qui est faux par hypothèse.

Donc $\hat{A}_1$ est dans l'un des cas suivants :

– $a \in \hat{A}'$ et $a' \in \hat{A}'$.
– $a \in \hat{A}'$ et $a' \notin \hat{A}'$.
– $a \notin \hat{A}'$ et $a' \notin \hat{A}'$. Mais dans ce cas , (E) implique $[minR(a), maxR(a)] \notin \hat{A}'$ et $[minR(a'), maxR(a')] \notin \hat{A}'$. Donc $\hat{A}' = \emptyset$ ce qui est impossible.

Nous pouvons voir que dans tous les cas possibles

$$a \in \hat{A}_1$$

et donc, puisque $[minR(a), maxR(a)] \diamond a[A_1]$, nous pouvons déduire que

$$[minR(a), maxR(a)] \in \hat{A}_1.$$

Et donc par définition de $\unrhd$,

$$[minR(a), maxR(a)] \unrhd [minR(a'), maxR(a')][A_1]$$

et par le lemme 5.2 (point 2)

$$[minR(a), maxR(a)] \unrhd [minR(a'), maxR(a')]$$

Et donc par la définition du préordre $\succcurlyeq$ nous avons

$$(minR(a), maxR(a)) \succcurlyeq (minR(a'), maxR(a')).$$

Supposons maintenant $a \in A \setminus \hat{A}$. D'après l'hypothèse (A), il existe un élément a' dans $\hat{A}$ et nous avons donc $a' \rhd a[A]$. En conséquence par le lemme 5.2 (point 3), nous avons aussi $a' \rhd a$.

Par le lemme 5.1 nous avons

$$[minR(a), maxR(a)] \diamond a \text{ et } [minR(a'), maxR(a')] \diamond a'.$$

Nous pouvons conclure par un argument analogue :

$$(minR(a'), maxR(a')) \succ (minR(a), maxR(a)) \text{ pour un certain } a' \in A.$$

Nous allons maintenant prouver la réciproque. Supposons que nous avons un préordre total sur les paires d'éléments de $\mathcal{C}$ (i.e. $\succcurlyeq \subseteq (\mathcal{C} \times \mathcal{C}) \times (\mathcal{C} \times \mathcal{C})$) qui satisfait les conditions (1) et (2) du théorème. Vérifions la satisfaction de l'axiome $(\boldsymbol{\alpha})$.

Supposons que nous avons $A_1 \subseteq A_2$ et $A_1 \cap \hat{A}_2$ n'est pas vide. Soit $a \in A_1 \cap \hat{A}_2$. Puisque $a \in \hat{A}_2$ par la condition (2) du théorème nous avons

$$\forall a' \in A_2, (minR(a), maxR(a)) \succcurlyeq (minR(a'), maxR(a')).$$

Puisque $A_1 \subseteq A_2$, $a \in A_1 \cap \hat{A}_2$ implique $a \in \hat{A}_1$, nous avons

$$(minR(a), maxR(a)) \succcurlyeq (minR(a'), maxR(a')), \forall a' \in A_1.$$

et donc $a \in \hat{A}_1$. Nous avons donc

$$A_1 \cap \hat{A}_2 \subseteq \hat{A}_1.$$

Soit maintenant $a \in \hat{A}_1$. Par la condition (2) nous avons

$$(minR(a), maxR(a)) \succcurlyeq (minR(a'), maxR(a')), \forall a' \in A_1$$

et en particulier pour les éléments de $A_1 \cap \hat{A}_2$.

Soit $a'' \in A_1 \cap \hat{A}_2$. $a'' \in \hat{A}_2$ implique par la définition de $\succcurlyeq$ que

$$(minR(a''), maxR(a'')) \succcurlyeq (minR(a'''), maxR(a''')), \forall a''' \in A_2,$$

ce qui implique par la transitivité de $\succcurlyeq$ que

$$(minR(a), maxR(a)) \succcurlyeq (minR(a'''), maxR(a''')), \forall a''' \in A_2,$$

ou en d'autres termes que $a \in \hat{A}_2$. Nous pouvons conclure que

$$a \in A_1 \cap \hat{A}_2.$$

Nous avons donc

$$\hat{A}_1 \subseteq A_1 \cap \hat{A}_2.$$

Finalement nous pouvons conclure

$$\hat{A}_1 = A_1 \cap \hat{A}_2.$$

Vérifions maintenant l'axiome $(\boldsymbol{\beta})$.

Soient $a, a' \in A$ tels que

$$\forall c \in R(a), \forall c' \in R(a'), c \geq c'.$$

Et soit $c'' \in C$. Soient a_1, a'_1 tels que

$$R(a_1) = R(a) \cup \{c''\} \text{ et } R(a'_1) = R(a') \cup \{c''\}.$$

Montrons que $a_1 \trianglerighteq a_1^{'}$.

- Si $c'' \geq maxR(a)$, alors $minR(a_1) \geq minR(a'_1)$ et $maxR(a_1) = maxR(a'_1)$.
- Si $minR(a') \geq c''$, alors $minR(a_1) = minR(a'_1)$ et $maxR(a_1) \geq maxR(a'_1)$.
- Si $maxR(a) \geq c'' \geq minR(a')$, alors $min(c'', minR(a)) \geq minR(a')$ et $maxR(a) \geq max(maxR(a'), c'')$. d'où $minR(a_1) \geq minR(a'_1)$ et $maxR(a_1) \geq maxR(a'_1)$.

Dans chacun de ces trois cas, nous avons $minR(a_1) \geq minR(a'_1)$ et $maxR(a_1) \geq maxR(a'_1)$, ce qui implique par la condition (1) que

$$(minR(a_1), maxR(a_1)) \succcurlyeq (minR(a'_1), maxR(a'_1)).$$

□

Ce théorème montre que tout critère satisfaisant les axiomes (α) et (β) sélectionne dans un problème A les éléments non dominés par rapport à un préordre total sur les actions qui est lié à un préordre $\succcurlyeq$ sur les paires de conséquences par la propriété qu'une action a est préférée ou indifférente à une autre action a' si et seulement si $(minR(a), maxR(a)) \succcurlyeq (minR(a'), maxR(a'))$.

Assez clairement, le théorème ne caractérise pas un seul critère mais une famille de critères possibles. Cela est dû à la condition sur les conséquences extrémales du théorème qui est seulement une implication (et non une équivalence) , ce qui permet de multiples raffinements possibles.

La force du théorème est qu'il utilise des exigences relativement simples pour caractériser une famille de critères qui dépendent seulement des éléments des ensembles de conséquences qui sont extrémaux par rapport à la relation de préférence de l'agent.

C'était déjà le cas dans le théorème prouvé dans [Arrow & Hurwicz 1977]. Ce que nous prouvons c'est que leur résultat peut être étendu au cadre plus pauvre en information que nous considérons. Bien sûr cela requiert des axiomes différents.

La proposition qui suit nous permet de prouver l'équivalence entre l'axiome (β) et un deuxième axiome qui pourrait sembler plus fort à prime abord :

Property $(\boldsymbol{\beta}^*)$ *Soit A un problème de décision, $a, a' \in A$ telles que $\forall c \in R(a), \forall c' \in R(a'), c \geqslant c'$. Considérons $C \in \mathcal{C}$ et $a_1, a'_1 \in A$ t.q. $R(a_1) = R(a) \cup C$ et $R(a'_1) = R(a') \cup C$. Alors $a_1 \trianglerighteq a'_1$.*

Par rapport au scénario présenté pour justifier l'axiome $(\boldsymbol{\beta})$, cet axiome stipule que la découverte par l'agent d'un nombre quelconque de conséquences communes aux deux actions ne doit pas changer de manière dramatique ses préférences.

Proposition 5.1 *Supposons qu'un critère de décision satisfait l'axiome $(\boldsymbol{\alpha})$. Alors Il satisfait l'axiome $(\boldsymbol{\beta})$ si et seulement s'il satisfait l'axiome $(\boldsymbol{\beta}^*)$*

Preuve
De manière évidente, si un critère satisfait l'axiome $(\boldsymbol{\beta}^*)$ alors il satisfait l'axiome $(\boldsymbol{\beta})$. Montrons alors le sens restant de la proposition. Soit A un problème de décision, $a, a' \in A$ telles que $\forall c \in R(a), \forall c' \in R(a'), c \geqslant c'$. Considérons $C \in \mathcal{C}$ et $a_1, a'_1 \in A$ t.q. $R(a_1) = R(a) \cup C$ et $R(a'_1) = R(a') \cup C$. Considérons les 4 points suivants $m = min(minR(a), minC)$, $m' = min(minR(a'), minC)$, $M = max(maxR(a), maxC)$ et $M' = max(maxR(a'), maxC)$. Le lemme 5.1 nous donne alors $a_1 \diamond [m, M]$ et $a_1^{'} \diamond [m', M']$. On voit clairement d'après leur définition que l'on a toujours $m \geq m'$ et $M \geq m'$ ce qui, puisque l'on suppose $(\boldsymbol{\alpha})$ et $(\boldsymbol{\beta})$ vérifiés, induit par le théorème 5.1 que Alors $a_1 \trianglerighteq a'_1$.

□

Cette proposition éclaire une autre interprétation de l'axiome $(\boldsymbol{\beta})$. Elle stipule que quand on a à comparer deux actions a et a' par un critère d'optimalité satisfaisant l'axiome $(\boldsymbol{\beta})$, si en éliminant leurs conséquences communes, toutes les conséquences restantes de a sont préférées à toutes celles de a', alors on devrait préférer a à a'.

Notons enfin que l'axiome $(\boldsymbol{\beta}^*)$ pourrait très bien remplacer l'axiome $(\boldsymbol{\beta})$ dans le théorème (c'est une conséquence directe du lemme 5.1). Toutefois, il est plus intérressant d'énoncer le résultat en utilisant l'axiome $(\boldsymbol{\beta})$ qui est moins exigeant.

Considérons maintenant l'axiome suivant :

axiome $(\boldsymbol{\beta'})$ *Soit A un problème de décision, $a, a' \in A$ t.q. $\forall c \in R(a), \forall c' \in R(a'), c > c'$. Considérons alors $c'' \in C$ and $a_1, a'_1 \in A$ t.q. $R(a_1) = R(a) \cup \{c''\}$ et $R(a_1^{'}) = R(a') \cup \{c''\}$ on a alors $a_1 \triangleright a'_1$.*

Cet axiome est similaire à $(\boldsymbol{\beta})$ à la différence qu'à part les conséquences communes, toutes les conséquences de la première action sont strictement préférées à toutes celles de la seconde action.

Nous pouvons obtenir un théorème similaire au théorème 5.1 basé sur l'axiome $(\boldsymbol{\beta'})$ qui conduit à imposer une condition supplémentaire sur le préordre défini dans le théorème 5.1 :

Théorème 5.2 *Sous les hypothèses (A), (B) et (C), un critère d'optimalité satisfait les axiomes* $(\boldsymbol{\alpha})$, $(\boldsymbol{\beta'})$ *et* $(\boldsymbol{\beta})$ *si et seulement si il existe un pré-ordre total* $\succcurlyeq$ *sur l'ensemble des paires de conséquences* (m, M) *avec* $M \geqslant m$ *vérifiant les conditions :*

(1) Si $m_1 \geq m_2$ *et* $M_1 \geq M_2$ *alors* $(m_1, M_1) \succcurlyeq (m_2, M_2)$*.*

(2) Si $m_1 \geq m_2$ *et* $M_1 \geq M_2$ *et* $(m_1 > m_2$ *or* $M_1 > M_2)$ *alors* $(m_1, M_1) \succ (m_2, M_2)$*.*

tel que les ensembles optimaux sont définis par
$\hat{A} = \{a | (minR(a), maxR(a)) \succcurlyeq (minR(a'), maxR(a')), \forall a' \in A\}$.

Preuve
Vu la preuve du théorème 5.1, nous avons juste à prouver que l'ajout de la condition (2) dans le théorème est équivalent à l'ajout de l'axiome $(\boldsymbol{\beta'})$.

Comme pour le théorème 5.1, définissons :

$$(m_1, M_1) \succcurlyeq (m_2, M_2) \text{ si et seulement si } [m_1, M_1] \trianglerighteq [m_2, M_2]$$

Nous définissons maintenant de manière naturelle la relation stricte associée :

$$(m_1, M_1) \succ (m_2, M_2) \text{ si et seulement si } [m_1, M_1] \triangleright [m_2, M_2]$$

Vérifions d'abord que l'ajout de l'axiome $(\boldsymbol{\beta'})$ implique la condition (2). Il y a deux cas possibles : $m_1 > m_2$ ou $M_1 > M_2$.

- Soient m_1, m_2, M_1, M_2 tels que $m_1 > m_2$ et $M_1 \geq M_2$. Puisque $m_1 > m_2$, l'axiome $(\boldsymbol{\beta'})$ nous permet de conclure en ajoutant M_2 que $[m_1, M_2] \triangleright [m_2, M_2]$. De plus $M_1 \geq M_2$ et l'axiome $(\boldsymbol{\beta})$ nous permet de conclure en ajoutant m_1 que $[m_1, M_1] \trianglerighteq [m_1, M_2]$. Ainsi nous avons finalement par transitivité que $[m_1, M_1] \triangleright [m_2, M_2]$ et par définition $(m_1, M_1) \succ (m_2, M_2)$.
- Soient m_1, m_2, M_1, M_2 tels que $m_1 \geq m_2$ et $M_1 > M_2$. Puisque $m_1 \geq m_2$, l'axiome $(\boldsymbol{\beta})$ nous permet de conclure en ajoutant M_2 que $[m_1, M_2] \trianglerighteq [m_2, M_2]$. De plus $M_1 > M_2$ et l'axiome $(\boldsymbol{\beta'})$ nous permet de conclure en ajoutant m_1 que $[m_1, M_1] \triangleright [m_1, M_2]$. Ainsi nous avons finalement par transitivité $[m_1, M_1] \triangleright [m_2, M_2]$ et par définition $(m_1, M_1) \succ (m_2, M_2)$.

Cela termine la preuve du sens directe du théorème.

Vérifions maintenant l'axiome $(\boldsymbol{\beta'})$. Soient a, a' telles que $\forall c \in R(a), c' \in R(a'), c > c'$. Et soit $c'' \in C$. Soit a_1, a'_1 telles que $R(a_1) = R(a) \cup \{c''\}$ et $R(a'_1) = R(a') \cup \{c''\}$:

- Si $c'' \geq maxR(a)$, alors $minR(a_1) > minR(a'_1)$ et $maxR(a_1) = maxR(a'_1)$.
- Si $minR(a') \geq c''$, alors $minR(a_1) = minR(a'_1)$ et $maxR(a_1) > maxR(a'_1)$.
- Si $maxR(a) > c'' > minR(a')$, alors $min(c'', minR(a_1)) > minR(a'_1)$ et $maxR(a_1) > max(maxR(a'_1), c'')$.

Nous pouvons voir que dans chaque cas, en ajoutant une conséquence commune on se retrouve encore dans une situation où la condition (2) s'applique. Nous pouvons alors conclure que $a_1 \triangleright a'_1$. $\square$

Ajouter l'axiome ($\boldsymbol{\beta'}$) restreint l'ensemble des critères retenus. Le théoreme 5.1 implique seulement des comparaisons par rapport à $\geq$ sur les conséquences extrémales pour caractériser les actions optimales. Le théoreme 5.2 introduit des comparaisons strictes dans la condition (2) et exclut ainsi d'autres actions de l'ensemble optimal. Ainsi, une action a telle qu'il existe une autre action a' avec $minR(a') \geq minR(a)$ et $maxR(a') \geq maxR(a)$ peut être dans l'ensemble optimal pour un critère admis par le théorème 5.1, même si $minR(a') > minR(a)$ ou $maxR(a') > maxR(a)$, cela n'est pas possible pour un critère admis par le théorème 5.2.

5.5 Exemples de critères d'optimalité

Dans cette section nous passons en revue quelques critères présentés à la section 2.6 ainsi que quelques autres critères et examinons la question de savoir s'ils satisfont les axiomes précédemment présentés. C'est une manière d'expliciter la nature des critères caractérisés par les théorèmes que nous avons démontrés. Pour faciliter la présentation, nous allons donner pour chaque critère son expression pour un problème constenant seulement deux actions. La définition de l'ensemble optimal dans le cas général s'en déduit de manière directe.

Examinons d'abord le critère le plus prudent. C'est le critère Min appelé aussi critère de Wald [Wald 1950] :

Définition 5.3 (Min) *Soient deux actions a_1, a_2.*

$$a_1 \trianglerighteq_{min} a_2 \ \textit{ ssi } \ minR(a_1) \geq minR(a_2).$$

On peut tout de suite lui associer le critère dual exprimant une sorte d'état d'esprit optimiste de l'agent qui l'utilise.

Définition 5.4 (Max) *Soient deux actions a_1, a_2.*

$$a_1 \trianglerighteq_{max} a_2 \ \textit{ ssi } \ maxR(a_1) \geq maxR(a_2).$$

Il est facile de vérifier que ces deux critères définissent deux préordres totaux vérifiant la condition (1) du théorème 5.1 et qu'ils satisfont les axiomes ($\boldsymbol{\alpha}$) et ($\boldsymbol{\beta}$).

Proposition 5.2 *Les critères Min et Max satisfont les axiomes* ($\boldsymbol{\alpha}$) *et* ($\boldsymbol{\beta}$).

Preuve
Montrons d'abord que le critère Min satisfait l'axiome ($\boldsymbol{\alpha}$). Considérons pour cela un problème de décision $A \subseteq \mathcal{A}$. Soient une action $a \in \hat{A}$ et un problème $A_1 \subseteq A$ un problème tel que $a \in A_1$.

Si le critère utilisé est le critère Min, alors $a \in \hat{A}$ implique $a \trianglerighteq_{min} a', \forall a' \in A$ ou par la définition du Min $minR(a) \geq minR(a'), \forall a' \in A$. Mais alors on a en particulier $minR(a) \geq minR(a'), \forall a' \in A_1$ ce qui est la définition de $a \in \hat{A}_1$. Donc $a \in \hat{A}_1$. Ainsi le critère Min satisfait l'axiome ($\boldsymbol{\alpha}$).

Un argument dual montre que le critère Max satisfait aussi l'axiome ($\boldsymbol{\alpha}$).

Montrons maintenant que le critère Min satisfait l'axiome ($\boldsymbol{\beta}$). Considérons pour cela deux actions $a, a' \in \mathcal{A}$ telles que $\forall c \in R(a), c' \in R(a')$ on a $c \geq c'$. On a alors $minR(a) \geq minR(a')$. Soit maintenant $c'' \in \mathcal{C}$ et $a_1, a_1' \in \mathcal{A}$ telles que $R(a_1) = R(a) \cup \{c''\}$ et $R(a_1') = R(a') \cup \{c''\}$. Alors on peut se trouver dans l'un des cas suivants :

– Si $c'' \geq maxR(a)$, alors $minR(a_1) \geq minR(a'_1)$.
– Si $minR(a') \geq c''$, alors $minR(a_1) = minR(a'_1)$.
– Si $maxR(a) \geq c'' \geq minR(a')$, alors $min(c'', minR(a)) \geq minR(a')$. D'où $minR(a_1) \geq minR(a'_1)$.

Ainsi dans ces trois cas, $minR(a_1) \geq minR(a'_1)$ ce qui par définition du critère Min donne $a_1 \trianglerighteq_{min} a'_1$. Donc le critère Min satisfait l'axiome $(\boldsymbol{\beta})$. Un argument dual montre que le critère Max satisfait aussi l'axiome $(\boldsymbol{\beta})$. □

En fait nous pouvons caractériser exactement chacun de ces deux critères en ajoutant à l'axiome $(\boldsymbol{\beta})$ l'un des axiomes suivants *Monotonie* (M) et une forme simplifiée d'*Aversion à l'Incertitude* (m) :

(m) $\forall a, b \in A, R(a) \subseteq R(b) \Rightarrow a \trianglerighteq b$.

(M) $\forall a, b \in A, R(b) \subseteq R(a) \Rightarrow a \trianglerighteq b$.

Le premier axiome exprime une aversion du décideur pour la dispersion des conséquences d'une action avec la motivation que moins les conséquences associées à une action sont nombreuses, plus le décideur est sûr de ses attentes. Le deuxième axiome exprime une attitude duale et a déjà été présenté plus haut.

Nous avons alors le résultat suivant qui montre que l'axiome (m) (respectivement (M)) une fois joint aux axiomes $(\boldsymbol{\alpha})$ et $(\boldsymbol{\beta})$ raffine la famille caractérisée par ces derniers en la réduisant au seul critère Min (respectivement Max) :

Proposition 5.3 *Un critère d'optimalité satisfait (α), (β) et (M) si et seulement si c'est le critère Min. Un critère d'optimalité satisfait (α), (β) et (m) si et seulement si c'est le critère Max.*

Preuve

– De manière évidente, $\trianglerighteq_{max}$ satisfait (β) et (M). Montrons alors le sens réciproque. Soit $\trianglerighteq$ un pré-ordre sur $\mathcal{A}$ satisfaisant les axiomes $(\beta + M)$. Pour $a \in \mathcal{A}$ nous avons par (M) que $a \trianglerighteq [maxR(a)]$. D'un autre côté, puisque $\forall c \in R(a), maxR(a) \geq c$, en ajoutant $maxR(a)$ on peut conclure $[maxR(a)] \trianglerighteq a$. Ainsi nous avons que $[maxR(a)] \diamond a$ (F).
Soient $a, b \in A$ t.q. $a \trianglerighteq_{max} b$, i.e., $maxR(a) \geq maxR(b)$. Nous avons donc $[maxR(a)] \geq [maxR(b)]$ ce qui par (F) conduit à $a \trianglerighteq b$.
Soient maintenant $a, b \in A$ t.q. $a \trianglerighteq b$. Supposons alors que $maxR(b) > maxR(a)$. Par (F) nous avons $[maxR(a)] \geq [maxR(b)]$. Contradiction.
– De manière évidente, $\trianglerighteq_{min}$ satisfait les axiomes (β) et (m). Montrons maintenant le sens réciproque. Soit $\trianglerighteq$ un pré-ordre sur $\mathcal{A}$ satisfaisant les axiomes $(\beta + m)$. Pour $a \in \mathcal{A}$ nous avons par (m) que $[minR(a)] \trianglerighteq a$. Par ailleurs, comme on a $\forall c \in R(a), c \geq minR(a)$, en ajoutant $minR(a)$ nous avons $a \trianglerighteq [minR(a)]$. Ainsi, $[minR(a)] \diamond a$ (E1).
Soient $a, b \in A$ t.q. $a \trianglerighteq_{min} b$, i.e., $minR(a) \geq minR(b)$. Nous avons alors $[minR(a)] \geq [minR(b)]$ ce qui par (E1) donne $a \trianglerighteq b$.
Soient maintenant $a, b \in A$ t.q. $a \trianglerighteq b$. Supposons alors que $minR(b) > minR(a)$. Par (E1) nous avons $[minR(a)] \geq [minR(b)]$. Contradiction.

□

Il est aisé de voir que ni le critère Min ni le critère Max ne satisfont l'axiome $(\boldsymbol{\beta'})$. En effet, considérons deux actions a_1, a_2 telles que $minR(a_1) \geq minR(a_2)$ et $maxR(a_1) > maxR(a_2)$. Nous avons $a_1 \diamond_{min} a_2$. Cela exclut le critère Min de la famille caractérisée par le théorème 5.2 puisque la condition (2) dans ce théorème impliquerait dans ce cas que $a_1 \rhd_{min} a_2$. Un argument similaire exclut le critère Max.

Ainsi, les deux critères précédents ne font pas partie de la famille caractérisée par le théorème 5.2. Ils satisfont le théorème 5.1, mais ils le font de manière triviale parce qu'ils prennent en compte seulement l'une des conséquences extrémales là où le théorème 5.1 suggère que le décideur devrait utiliser un critère qui tient compte des deux conséquences extrémales à la fois.

Voici deux critères parmi ceux présentés à la section 5.5 qui tirent parti de cette possibilité.

Définition 5.5 (Minmax - Maxmin) *Considérons deux actions a_1, a_2.*

$$a_1 \unrhd_{minmax} a_2 \ \ \textit{ssi} \ \ a_1 \unrhd_{lex(\unrhd_{min}, \unrhd_{max})} a_2.$$

$$a_1 \unrhd_{minmax} a_2 \ \ \textit{ssi} \ \ a_1 \unrhd_{lex(\unrhd_{max}, \unrhd_{min})} a_2.$$

Le critère Minmax est un raffinement du critère Min, prenant en compte le Max quand les conséquences minimales des deux actions sont équivalentes.

De manière duale, le critère Max peut être raffiné vers un critère lexicographique Maxmin, i.e., raffinant le Max par la comparaison des conséquences minimales.

On a alors le résultat suivant :

Proposition 5.4 *Les critères Minmax et Maxmin satisfont les axiomes* $(\boldsymbol{\alpha})$, $(\boldsymbol{\beta})$ *et* $(\boldsymbol{\beta'})$.

Preuve
Nous devons prouver que les conditions (1), (2) du théorème 5.2 ainsi que la condition définissant les ensembles optimaux dans ce même théorème sont satisfaites.

Définissons un préordre total $\succcurlyeq$ sur les paires de conséquences basé sur le critère Minmax au travers de la condition de sélection des ensembles optimaux du théorème 5.2.

Considérons deux actions a_1, a_2 telles que $minR(a_1) \geq minR(a_2)$ et $maxR(a_1) \geq maxR(a_2)$. La définition de la relation lexicographique implique alors que $a_1 \unrhd_{lex(\unrhd_{min}, \unrhd_{max})} a_2$ ce qui signfie $a_1 \unrhd_{minmax} a_2$ et par définition $(minR(a_1), maxR(a_1)) \succcurlyeq (minR(a_2), maxR(a_2))$. Cela prouve la condition (1).

Afin de prouver la condition (2), considérons deux actions a_1, a_2 telles que ($minR(a_1) \geq minR(a_2)$ et $maxR(a_1) > maxR(a_2)$) ou ($minR(a_1) > minR(a_2)$ et $maxR(a_1) \geq maxR(a_2)$). La définition de la relation lexicographique implique alors que $a_1 \rhd_{lex(\unrhd_{min}, \unrhd_{max})} a_2$ ce qui signifie $a_1 \rhd a_2$ et par définition $(minR(a_1), maxR(a_1)) \succ (minR(a_2), maxR(a_2))$. Cela prouve la condition (2). □

Le critère suivant min' illustre une autre manière de raffiner le critère Min. Une interprétation en est la situation où un agent quand même prudent ne veut pas être trop pessimiste. Il est disposé à croire que la pire conséquence ne va pas forcément se produire et écarte donc les pires conséquences des actions qu'il a à comparer. Toutefois il ne s'avance pas trop et compare les ensembles de conséquences restants en se basant sur le critère Min. En résumé, il compare les actions sur la base du deuxième moins bon élément de leurs ensembles de conséquences. Le critère $min2$ est un compromis entre le critère précédent et le critère Min. Il raffine le Min lexicographiquement comme le fait le critère Minmax mais utilise de manière plus pessimiste le critère min' à la place du critère Max.

Définition 5.6 (min') *Soit A un problème de décision et* $a, b, a', b' \in A$ *tels que* $R(a') = R(a) \setminus \{minR(a)\}$ *et* $R(b') = R(b) \setminus \{minR(b)\}$*. Alors* $a \trianglerighteq_{min1} b$ *ssi* $a' \trianglerighteq_{min} b'$

Définition 5.7 (min2) *Soient A un problème de décision et* $a, b \in A$*. Alors* $a \trianglerighteq_{min'} b$ *ssi* $a \trianglerighteq_{lex(min,min1)} b$*.*

Comme on peut s'y attendre d'après le théorème 5.1, ces deux critères prenant en compte d'autres éléments que les éléments extrémaux des ensembles de conséquences ne satisfont pas l'axiome ($\boldsymbol{\beta}$).

Proposition 5.5 *Les critères* min' *et* $min2$ *ne satisfont pas l'axiome* ($\boldsymbol{\beta}$)*.*

Preuve
Considérons deux actions $a, a' \in \mathcal{A}$ et des conséquences $c_1, c_2, c_3 \in \mathcal{C}$ telles que $c_1 < c_2 < c_3$, $R(a) = c_1, c_2, c_3$ et $R(a') = c_1$. On a alors $\forall c \in R(a), c' \in R(a'), c \geq c'$. Soient maintenant les actions $a_1, a_1^{'} \in \mathcal{A}$ telles que $R(a_1) = R(a) \cup \{c^*\}$ et $R(a_1^{'}) = R(a') \cup \{c^*\}$ où $c_2 < c^* < c_3$. Alors nous avons $a_1^{'} \triangleright_{min'} a'$ et $a_1^{'} \triangleright_{min1} a'$ alors que l'axiome ($\boldsymbol{\beta}$) prescrit l'opposé. □

Remarquons que cet argument montre aussi que min' et $min2$ ne satisfont pas l'axiome ($\boldsymbol{\beta'}$).

Toujours dans la même visée de raffiner le critère Min se trouve le critère de comparaison lexicographique leximin qui va plus loin que le critère $min2$. On peut l'interpréter comme suit. On compare deux actions sur la base des premières conséquences différentes dans leurs ensembles respectifs de conséquences (en partant de la pire conséquence pour chacune). En voici une définition formelle :

Définition 5.8 (leximin) *Soit un problème de décision A et* $a, b \in A$*. Supposons* $|R(a)| = n$ *et* $|R(b)| = p$ *tel que sans perte de généralité* $p \geq n$*. Indexons les éléments de* $R(a)$ *comme suit :* $c_1, c_2, \cdots, c_n$ *et supposons que nous avons* $c_n \geq \cdots \geq c_1$*. Similairement, indexons les éléments de* $R(p)$ *comme suit :* $c_1^{'}, c_2^{'}, \cdots, c_p^{'}$ *et supposons que nous avons* $c_p^{'} \geqslant \cdots \geq c_1^{'}$*.*

Le préordre Leximin $\trianglerighteq_{leximin}$ *est défini par* $a \trianglerighteq_{leximin} b$ *ssi* $a \triangleright_{leximin} b$ *ou* $a \diamond_{leximin} b$ *avec :*

- $a \trianglerighteq_{leximin} b$ *ssi* ($\exists j \in [1, n]$ *t.q.* $\forall i < j$ $c_i = c_i^{'}$ *et* $c_j > c_j^{'}$) *ou* ($p > n$ *et* $\forall i \in [1, n]$, $c_i = c_i^{'}$)*.*
- $a \diamond_{leximin} b$ *ssi* $p = n$ *et* $\forall i \in [1, n]$, $c_i = c_i^{'}$*.*

Le même contre exemple de la preuve de la proposition 5.5 montre aussi que le critère leximin ne satisfait pas non plus l'axiome ($\boldsymbol{\beta}$) (ni l'axiome ($\boldsymbol{\beta'}$)).

Proposition 5.6 *Le critère* $leximin$ *ne satisfait pas l'axiome* ($\boldsymbol{\beta}$)*.*

Un autre critère bien connu (en fait c'est une famille de critères) qui satisfait nos trois axiomes est le critère d'Hurwicz [Hurwicz 1951] :

Définition 5.9 (Hurwicz) *Soit* $\mathcal{C} \subseteq I\!R$*. Considérons deux actions* a_1, a_2 *et* $\alpha \in [0, 1]$*.* $a_1 \trianglerighteq_\alpha a_2$ *ssi* $(\alpha * minR(a_1) + (1 - \alpha) * maxR(a_1)) \geq (\alpha * minR(a_2) + (1 - \alpha) * maxR(a_2))$*.*

Nous voyons que pour la bonne définition de ce critère, le terme $\alpha * minR(a_1)$ doit notamment avoir un sens. Cela implique une structure particulière de l'ensemble des conséquences $\mathcal{C}$. Typiquement ce critère est utilisé quand $\mathcal{C} \subseteq I\!R$. Les théorèmes 5.1 et 5.2 n'utilisaient pas d'autres hypothèses sur l'ensemble des conséquences que les hypothèses (B) et (C). Il s'applique donc particulièrement au cas précédent. Il est donc pertinent de tester si le critère de Hurwicz satisfait bien les axiomes que nous avons introduits. Le résultat suivant montre que c'est le cas quand $\alpha \neq 0$ et $\alpha \neq 1$.

Proposition 5.7 *Supposons $\mathcal{C} \subseteq I\!R$. $\forall \alpha \in]0, 1[$, le critère de Hurwicz satisfait les axiomes ($\boldsymbol{\alpha}$), ($\boldsymbol{\beta}$) et ($\boldsymbol{\beta'}$).*

Preuve
Soit $\alpha \in]0, 1[$. Nous devons prouver que les conditions (1), (2) du théorème 5.2 ainsi que la condition définissant les ensembles optimaux dans ce même théorème sont satisfaites.

Définissons un préordre total $\succcurlyeq$ sur les paires de conséquences basé sur le critère de Hurwicz au travers de la condition de sélection des ensembles optimaux du théorème 5.2.

Considérons deux actions a_1, a_2 telles que $minR(a_1) \geq minR(a_2)$ et $maxR(a_1) \geq maxR(a_2)$. On a alors $(\alpha * minR(a_1) + (1-\alpha) * maxR(a_1)) \geq (\alpha * minR(a_2) + (1-\alpha) * maxR(a_2))$. La définition du critère de Hurwicz implique alors que $a_1 \trianglerighteq_\alpha a_2$ et par définition du préordre $\succcurlyeq$, $(minR(a_1), maxR(a_1)) \succcurlyeq (minR(a_2), maxR(a_2))$. Cela prouve la condition (1).

Afin de prouver la condition (2), considérons deux actions a_1, a_2 telles que $(minR(a_1) \geq minR(a_2)$ et $maxR(a_1) > maxR(a_2))$ ou $(minR(a_1) > minR(a_2)$ et $maxR(a_1) \geq maxR(a_2))$. Puisque $\alpha \in]0, 1[$, $(\alpha * minR(a_1) + (1-\alpha) * maxR(a_1)) > (\alpha * minR(a_2) + (1-\alpha) * maxR(a_2))$. La définition du critère de Hurwicz implique alors que $a_1 \triangleright_\alpha a_2$ et par définition du préordre $\succcurlyeq$, $(minR(a_1), maxR(a_1)) \succ (minR(a_2), maxR(a_2))$. Cela prouve la condition (2).
$\square$

Remarquons que la proposition 5.7 écarte les critères obtenus pour $\alpha = 0$ et $\alpha = 1$. La raison en est simple. Dans ces cas les critères sont ceux du Min et du Max dont nous avons précédement montré qu'ils satisfont l'axiome ($\boldsymbol{\beta}$) mais pas l'axiome ($\boldsymbol{\beta'}$).

Un autre cas particulier de critère est celui obtenu pour $\alpha = 1/2$. C'est le critère de la moyenne qui satisfait donc d'après le résultat précédent les axiomes (α), (β) et (β').

Définition 5.10 (Moyenne) *Soit A un problème de décision et $a, b \in A$. Alors $a \trianglerighteq_{av} b$ ssi $(minR(a) + maxR(a))/2 \geq (minR(b) + maxR(b))/2$.*

Proposition 5.8 *Supposons $\mathcal{C} \subseteq I\!R$. Le critère de la Moyenne satisfait les axiomes ($\boldsymbol{\alpha}$), ($\boldsymbol{\beta}$) et ($\boldsymbol{\beta'}$).*

Preuve
Prouvons que les conditions (1), (2) du théorème 5.2 ainsi que la condition définissant les ensembles optimaux dans ce même théorème sont satisfaites.

Soit un préordre total $\succcurlyeq$ sur les paires de conséquences basé sur le critère de la Moyenne au travers de la condition de sélection des ensembles optimaux du théorème 5.2.

Considérons deux actions a_1, a_2 telles que $minR(a_1) \geq minR(a_2)$ et $maxR(a_1) \geq maxR(a_2)$. On a alors $(*minR(a_1) + maxR(a_1))/2 \geq (minR(a_2) + maxR(a_2))/2$. La définition du critère implique alors que $a_1 \trianglerighteq_{moy} a_2$ et par définition du préordre $\succcurlyeq$, $(minR(a_1), maxR(a_1)) \succcurlyeq (minR(a_2), maxR(a_2))$. Cela prouve la condition (1).

Afin de prouver la condition (2), considérons deux actions a_1, a_2 telles que $(minR(a_1) \geq minR(a_2)$ et $maxR(a_1) > maxR(a_2))$ ou $(minR(a_1) > minR(a_2)$ et $maxR(a_1) \geq$

$maxR(a_2)$). On a alors $(minR(a_1) + maxR(a_1))/2 > (minR(a_2) + maxR(a_2))/2$. La définition du critère implique alors que $a_1 \rhd_{moy} a_2$ et par définition du préordre $\succcurlyeq$, $(minR(a_1), maxR(a_1)) \succ (minR(a_2), maxR(a_2))$. Cela prouve la condition (2). □

Dans la section 2.5, le critère de la médiane a été introduit comme une manière d'expliquer un comportement observé d' agents "réels". La formulation présentée réduisait son champ d'application au cas d'un ensemble de conséquences inclus dans $I\!R$. On peut en imaginer une variante "ordinale" qui permet d'étendre son utilisation à un ensemble de conséquences quelconque. Notons que cette définition nécessite que l'ensemble de conséquences associé à une action soit de cardinal fini.

Définition 5.11 (médiane) *Pour une action $a \in \mathcal{A}$, définissons l'action $m(a)$ par son ensemble de conséquences comme suit. Supposons $R(a) = \{c_1, \cdots, c_n\}$. Si n est pair alors $R(m(a)) = c_{n+1/2}$ et sinon $R(m(a)) = \{c_{n/2}, c_{1+(n/2)}\}$.*
Soient $a, b \in \mathcal{A}$, alors $a \trianglerighteq_{med} b$ si et seulement si $m(a) \trianglerighteq_{minmax} m(b)$.

Proposition 5.9 *Le critère $mdian$ ne satisfait pas l'axiome ($\boldsymbol{\beta}$).*

Preuve
Considérons deux actions $a, a' \in \mathcal{A}$ et des conséquences $c_1, c_2 \in \mathcal{C}$ telles que $c_1 < c_2$, $R(a) = \{c_1, c_2\}$ et $R(a') = \{c_2\}$. On a alors $\forall c \in R(a), c' \in R(a'), c \leq c'$. Soient maintenant les actions $a_1, a_1^{'} \in \mathcal{A}$ telles que $R(a_1) = R(a) \cup \{c^*\}$ et $R(a_1^{'}) = R(a') \cup \{c^*\}$ où $c^* < c_1$. Alors nous avons $R(m(a_1)) = c_1$ et $R(m(a_1^{'})) = \{c^*, c_2\}$. Or cela donne $m(a_1) \rhd_{minmax} m(a_1^{'})$ qui implique $a_1 \rhd_{med} a_1^{'}$ alors que l'axiome ($\boldsymbol{\beta}$) prescrit l'opposé. □

Evoquons enfin un critère présenté à la section 2.5 qui malgré son aspect contre-intuitif satisfait certains des axiomes utilisés dans la littérature.

Définition 5.12 (Cardinal) *Soient A un problème de décision et $a, b \in A$. Alors $a \trianglerighteq_{card} b$ ssi $|R(a)| \geq |R(b)|$.*

Nous pouvons aisément prouver que cet axiome ne satisfait pas les axiomes ($\boldsymbol{\beta}$) et ($\boldsymbol{\beta'}$).

Proposition 5.10 *Le critère $Cardinal$ ne satisfait pas l'axiome ($\boldsymbol{\beta}$).*

Preuve
Considérons deux actions $a, a' \in \mathcal{A}$ et des conséquences $c_1, c_2, c_3 \in \mathcal{C}$ telles que $c_1 < c_2 < c_3$, $R(a) = \{c_1, c_2\}$ et $R(a') = \{c_3\}$. On a alors $\forall c \in R(a), c' \in R(a'), c \leq c'$. Soient maintenant les actions $a_1, a_1^{'} \in \mathcal{A}$ telles que $R(a_1) = R(a) \cup \{c^*\}$ et $R(a_1^{'}) = R(a') \cup \{c^*\}$ où c^* quelconque. Alors nous avons $|R(a_1)| = 3$ et $\left|R(m(a_1^{'}))\right| = 2$. Or cela donne $a_1 \rhd_{med} a_1^{'}$ alors que l'axiome ($\boldsymbol{\beta}$) prescrit l'opposé. □

5.6 Comparaison avec d'autres axiomes

Dans cette section, nous comparons les axiomes α, β et β' à d'autres axiomes de la littérature.

Le premier axiome considéré est le plus faible axiome d'indépendance que nous avons présenté à la section 2.6.3. C'est l'axiome (I) suivant :

(I) $\forall a \in A$, Soient $x, y \in \mathcal{C}$ t.q $x \notin R(a)$ et $y \notin R(a)$, et soit $a' \in A$ t.q. $R(a') = R(a) \cup \{x\}$ et $a'' \in A$ t.q. $R(a'') = R(a) \cup \{y\}$. Alors $x \geq y \Rightarrow a' \trianglerighteq a''$.

Proposition 5.11 *$(\beta + \alpha) \Rightarrow (I)$ et $(I) \not\Rightarrow (\beta + \alpha)$.*

Preuve

Montrons d'abord que $(\beta + \alpha) \Rightarrow (I)$.

Soient $a \in A, x, y \in \mathcal{C} \setminus R(a)$ t.q. $x \geq y$. Soient alors $a_x, a_y \in A$ t.q. $R(a_x) = R(a) \cup \{x\}$ et $R(a_y) = R(a) \cup \{y\}$. Nous avons $y \geq minR(a_y)$ et $minR(a) \geq minR(a_y)$. Soit $t \in R(a_x)$. Donc $t = x$ ou $t \in R(a)$. Dans le premier cas, nous avons par définition de x et y, $x \geq y$ et donc $x \geq minR(a_y)$. Dans le dernier cas, puisque $t \in R(a), t \geq minR(a)$ et donc $t \geq minR(a_y)$. Nous avons alors $\forall t \in R(a_x)$, $t \geq minR(a_y)$ et spécialement $minR(a_x) \geq minR(a_y)$. Un argument similaire donne $maxR(a_x) \geq maxR(a_y)$. L'axiome β nous permet de conclure, puisque $minR(a_x) \geq minR(a_y)$, $R([minR(a_y), maxR(a_y)]) = \{minR(a_y)\} \cup \{maxR(a_y)\}$ et $R([minR(a_x), maxR(a_y)]) = \{minR(a_x)\} \cup \{maxR(a_y)\}$, que $[minR(a_x), maxR(a_y)] \trianglerighteq [minR(a_y), maxR(a_y)]$. Un argument similaire donne $[minR(a_x), maxR(a_x)] \trianglerighteq [minR(a_x), maxR(a_y)]$. Nous pouvons alors conclure par transitivité du préordre que $[minR(a_x), maxR(a_x)] \trianglerighteq [minR(a_y), maxR(a_y)]$. Finalement le lemme 5.1 donne $a_x \trianglerighteq a_y$. Nous avons donc prouvé que le préordre satisfait l'axiome (I).

Pour montrer que $(I) \not\Rightarrow (\beta + \alpha)$, il nous suffit de trouver un exemple de critère qui satisfait l'axiome (I) sans satisfaire l'axiome $(\boldsymbol{\beta})$. C'est le cas pour le critère leximin.

- leximin satisfait (I).
  Soient $a \in A, x, y \in \mathcal{C} \setminus R(a)$ t.q. $x \geq y$. Soient alors $a_x, a_y \in A$ t.q. $R(a_x) = R(a) \cup \{x\}$ et $R(a_y) = R(a) \cup \{y\}$. Supposons $R(a) = \{c_i, i \in |R(a)|\}$, $R(a_x) = \{c_{xi}, i \in |R(a)| + 1\}$ et $R(a_y) = \{c_{y_i}, i \in |R(a)| + 1\}$.
  - Si $x = y$ alors $a_x = a_y$ et $a_x \diamond_{lex} a_y$.
  - Si $x > y$ alors soit $m = max\{i | y > c_i\}$. Nous avons alors $\forall i \leq m, c_{xi} = c_{y_i}$. Donc $y = c_{y_{i+1}}$. Considérons alors c_{xi+1}. Nous avons l'un des deux cas suivants :
    - $c_{xi+1} = x$ et donc $c_{xi+1} > y$.
    - $c_{xi+1} < x$ et donc $c_{xi+1} \in R(a)$ ce qui signifie $c_{xi+1} \neq y$ et par définition de m $c_{xi+1} > y$.

    Donc dans les deux cas $c_{xi+1} > c_{y_{i+1}}$. Nous pouvons conclure par la définition du critère leximin que $a_x \triangleright_{lex} a_y$.

  Ainsi si $x \geq y$ alors $a_x \trianglerighteq_{lex} a_y$ ce qui termine la preuve.
- leximin ne satisfait pas l'axiome $(\boldsymbol{\beta})$.
  Soient $x, y, z, t \in \mathcal{C}$, $x \geq y \geq z \geq t$. Soient A un problème de décision, $a, a', a_1, a_1' \in A$ telles que $R(a) = \{x, z, t\}$, $R(a') = \{t\}$. $R(a_1) = R(a) \cup \{y\}$ et $R(a_1') = R(a') \cup \{y\}$. Donc $\forall c \in R(a), c' \in R(a'), c \geq c'$, mais $a_1' \triangleright_{leximin} a_1$ là où (β) prescrirait $a_1 \trianglerighteq a'_1$.

$\square$

Cette proposition montre que l'axiome $(\boldsymbol{\beta})$ (en présence de l'axiome (α)) est strictement plus fort que l'axiome (I). Cela suggère qu'il intéressant de comparer $(\boldsymbol{\beta})$ à des axiomes plus forts. Un exemple présenté à la section 2.6.3 est l'axiome (IF) suivant :

IS $\forall a, b \in A$ t.q. $a \trianglerighteq b$, soit x $\in \mathcal{C} \setminus (R(a) \cup R(b))$ et soit $a', b' \in A$ t.q. $R(a') = R(a) \cup \{x\}$ et $R(b') = R(b) \cup \{x\}$. Alors $a' \trianglerighteq b'$.

Le résultat suivant montre que les axiomes $(\boldsymbol{\beta})$ et (IF) sont logiquement indépendants :

Proposition 5.12 $(SI) \nRightarrow (\beta)$ *et* $(\beta) \nRightarrow (SI)$.

Preuve

- $\trianglerighteq_{card}$ satisfait (IF) mais pas (β). Soient $a, b \in A$ t.q. $a \trianglerighteq_{card} b$ et $x \in \mathcal{C} \setminus (R(a) \cup R(b))$. $a \trianglerighteq_{card} b$ signifie que $|R(a)| \geq |R(b)|$. Il est alors direct que $|R(a) \cup \{x\}| \geq |R(b) \cup \{x\}|$, ce qui permet de conclure.
  Soient $a, b \in A$ t.q. $R(a) = \{y\}$ et $R(a) = \{z, t\}$ et $y > z > t$. Soit alors $x \in \mathcal{C} \setminus (R(a) \cup R(b))$. Soient $a_1, b_1 \in A$ t.q. $R(a_1) = R(a) \cup \{x\}$ et $R(b_1) = R(b) \cup \{x\}$. Alors $b_1 \triangleright_{card} a_1$ là où (β) prescrit l'opposé.
- Minmax satisfait (β) et mais pas (IF).
  Ce que nous avons à prouver c'est que Minmax satisfait (IF). Soient $a, a' \in \mathcal{A}$ t.q. $minR(a) > minR(a')$ et $maxR(a') > maxR(a)$. Clairement $a \triangleright_{minmax} a'$. Soient alors $c \in \mathcal{C}$ t.q. $minR(a') \geq c$ et $a_1, a_1^{'} \in \mathcal{A}$ t.q. $R(a_1) = R(a) \cup \{c\}$ et $R(a_1^{'}) = R(a') \cup \{c\}$. Donc nous avons $minR(a_1) = minR(a_1^{'})$ et $maxR(a_1^{'}) > maxR(a_1)$ ce qui signifie $a_1^{'} \triangleright_{minmax} a_1$ là où (IF) prescrit le contraire.

$\square$

Un autre représentant des axiomes d'indépendance est l'axiome (IND).

(IND) $\forall a, b \in A$, soit $x \in \mathcal{C} \setminus (R(a) \cup R(b))$ et soit $a' \in A$ t.q. $R(a') = R(a) \cup \{x\}$ et $b' \in A$ t.q. $R(b') = R(b) \cup \{x\}$. Si $a \triangleright b$ alors $a' \trianglerighteq b'$.

Le résultat suivant montre que (β) est logiquement indépendant de (IND) :

Proposition 5.13 $(IND) \nRightarrow (\beta)$ *et* $(\beta) \nRightarrow (IND)$.

Comparer (β) à (IF) et (IND) a du sens car chacun de ces axiomes stipule en quelque sorte que les conséquences communes des actions ne doivent pas compter pour la caractérisation des actions optimales. Et prouver que (β) est logiquement indépendant de (IF) et de (IND), est intéressant puisque (β) semble plus faible au premier abord. Si de plus on prend en compte le fait que (IF) implique (I) et (IND), (β) apparaît comme un bon compromis entre ces axiomes. Ainsi (β) exclut quelques critères qui ne semblent pas correspondre au cadre de l'ignorance complète mais qui sont tout de même retenus par ces axiomes (le critère Cardinal est l'un de ceux là). Remplacer (β) par l'un de ces axiomes dans le théorème 5.1 ne conduit donc pas à la même famille de critères.

La proposition 2.6 indique que l'axiome (IE) doit avoir la même relation logique avec (β) que l'axiome (IF).

Pour ce qui est de l'axiome (IS), il semble plus naturel de le comparer avec l'axiome (β'). Le résultat suivant montre que ces deux axiomes sont logiquement indépendants :

Proposition 5.14 *$(IS) \nRightarrow (\beta')$ et $(\beta') \nRightarrow (IS)$.*

Examinons maintenant la relation entre (PG) et (β) et (β') :

(PG) $\forall a \in A$, soit $x \in \mathcal{C}$ et soit $a' \in A$ t.q. $R(a') = R(a) \cup \{x\}$:
- Si $x > c, \forall c \in R(a)$ alors $a' \rhd a$,
- Si $c > x, \forall c \in R(a)$ alors $a \rhd a'$.

Proposition 5.15 *1. $(PG) \nRightarrow (\beta)$ et $(\beta) \nRightarrow (PG)$.*
2. $(PG) \nRightarrow (\beta')$ et $(\beta') \nRightarrow (PG)$.
3. $(\beta + \alpha) \Rightarrow [(PG) \Leftrightarrow (\beta')]$.

Preuve
Montrons d'abord le point (1).

Pour montrer que $(\beta) \nRightarrow (PG)$, il nous suffit d'exhiber un critère satisfaisant (β) qui ne satisfait pas (PG). Le critère Min répond à cette description. Nous avons en effet montré que Min satisfait (β). Montrons qu'il ne satisfait pas (PG). Considérons pour cela une action $a \in \mathcal{A}$, une conséquence $x \in \mathcal{C}$ telle que $\forall c \in R(a), x > c$. Soit $a' \in \mathcal{A}$ telle que $R(a') = R(a) \cup \{x\}$. On a alors clairement $a \diamond_{min} a'$ alors que (PG) prescrit que $a' \rhd a$.

Montrons maintenant que $(PG) \nRightarrow (\beta)$. On a déjà montré que leximin ne satisfait pas (β). Tout ce que nous avons à prouver est qu'il satisfait (PG) :
Soit $a \in A$ et $x \in \mathcal{C}$. Soit alors $a' \in A$ t.q. $R(a') = R(a) \cup \{x\}$,

- Si $x > c, \forall c \in R(a)$ alors $|R(a')| > |R(a)|$ et $\forall c \in R(a), \exists c' \in R(a')$ t.q. $c = c'$. Par définition du pré-ordre nous avons alors $a' \rhd_{leximin} a$.
- Si $c > x, \forall c \in R(a)$ alors en particulier $minR(a) > x$ ce qui entraîne que $a \rhd_{leximin} a'$.

Montrons maintenant le point (2) de la proposition. Pour cela montrons en premier que $(\beta') \nRightarrow (PG)$ au moyen du critère $min2$.

Montrons d'abord que $min2$ satisfait (β').

Soient $a, a' \in A$ t.q. $\forall c \in R(a), c' \in R(a'), c > c'$. Soit $c'' \in C$, considérons $R(a_1) = R(a) \cup \{c''\}$ et $R(a_1') = R(a') \cup \{c''\}$. Nous pouvons observer que nous avons l'un des cas suivants :
- $c'' > minR(a_1')$. $minR(a_1) = min(c'', minR(a))$. Donc $minR(a_1) > minR(a_1')$ d'où on peut conclure $a_1 \rhd a_1'$.
- $c'' = minR(a_1')$. Donc $minR(a_1) = minR(a_1')$ et comparer ces deux actions revient à comparer deux éléments de $R(a)$ et $R(a')$ (ou à évaluer le cas dégénéré où R(a') est un singleton), ce qui donne clairement $a_1 \rhd a_1'$.

Ainsi dans les deux cas $a_1 \rhd a_1'$ ce qui signifie que (β') est satisfait.

Afin de montrer que min' ne satisfait pas (PG), examinons l'exemple suivant :

Soit $a \in A$ telle que $R(a) = \{c_1, c_2, c_3\}$ où les éléments sont strictement ordonnés, et soit $x \in \mathcal{C}$ tel que $x > c_3$. Alors si $a' \in A$ est telle que $R(a') = R(a) \cup \{x\}$, nous pouvons aisément conclure que $a \diamond_{min'} a'$ là où (PG) prescrit $a >_{min'} a'$.

Montrons maintenant que $(PG) \nRightarrow (\beta')$. Utilisons pour cela le critère de la médiane défini plutôt. Montrons d'abord que ce critère ne satisfait pas (β'). Considérons trois conséquences

$c_1, c_2, c_3 \in \mathcal{C}$ telles que $c_1 < c_2 < c_3$ et deux actions $a, a' \in \mathcal{A}$ telles que $R(a) = \{c_3\}$ et $R(a') = \{c_1, c_2\}$. Soient alors $c^{''} \in \mathcal{C}$ et $a_1, a_1^{'} \in \mathcal{A}$ telles que $R(a_1) = R(a) \cup \{c^{''}\}$ et $R(a_1^{'}) = R(a') \cup \{c^{''}\}$. On suppose de plus $c_1 > c''$. On a alors $R(m(a_1)) = \{c^{''}, c_3\}$ et $R(m(a_1^{'})) = \{c_1\}$. On a alors clairement $m(a_1^{'}) \rhd_{minmax} m(a_1)$ ce qui par définition du critère de la médiane donne $a_1^{'} \rhd_{med} a_1$ alors que (β') prescrit l'inverse.

Il nous reste à vérifier que le critère de la médiane satisfait bien (PG). Soit alors $a \in \mathcal{A}$. Supposons que $|R(a)|$ = n et notons $R(a) = \{c_1, \cdots, c_n\}$. Soient alors $x \in \mathcal{C}$ et $a' \in A$ telle que $R(a') = R(a) \cup \{x\}$:

- Si $x > c, \forall c \in R(a)$ alors si n est impair alors $R(m(a)) = c_{n+1/2}$ et $R(m(a')) = \{c_{n+1/2}, c_{1+(n+1/2)}\}$. On en conclut aisément $m(a') \rhd_{minmax} m(a)$ ce qui par définition du critère de la médiane donne $a' \rhd_{med} a$ comme le prescrit (PG). Si maintenant n est pair alors $R(m(a')) = c_{1+(n/2)}$ et $R(m(a)) = \{c_{n/2}, c_{1+(n/2)}\}$. On en conclut aisément $m(a') \rhd_{minmax} m(a)$ ce qui par définition du critère de la médiane donne $a' \rhd_{med} a$ comme le prescrit (PG).
- Si $c > x, \forall c \in R(a)$, un argument similaire au précédent nous donne le résultat recherché.

Ainsi dans les deux cas, le résultat prescrit par (PG) est vérifié. Donc le critère de la médiane satisfait l'axiome (PG).

Montrons maintenant le point (3) de la proposition. Supposons d'abord que (β') est satisfait.

Soient $a \in A$ et soit $x \in \mathcal{C}$. Soit $a' \in A$ telle que $R(a') = R(a) \cup \{x\}$.

Par le lemme 5.1 nous avons $a \diamond [minR(a), maxR(a)]$ et $a' \diamond [minR(a'), maxR(a')]$.

- Si $x > c, \forall c \in R(a)$ alors $minR(a) = minR(a')$ and $maxR(a') > maxR(a)$. Donc (β') implique $[minR(a'), maxR(a')] \rhd [minR(a), maxR(a)]$ puisque $maxR(a') > maxR(a)$ et nous pouvons conclure en ajoutant $minR(a)$. On conclut finalement $a' \rhd a$.
- Si $c > x, \forall c \in R(a)$ alors $minR(a) > minR(a')$ et $maxR(a') = maxR(a)$. Donc (β') implique $[minR(a), maxR(a)] \rhd [minR(a'), maxR(a')]$ puisque $minR(a) > minR(a')$ et nous pouvons conclure en ajoutant $maxR(a)$. Nous concluons enfin $a \rhd a'$.

Donc (PG) est satisfait.

Supposons maintenant que (PG) est satisfait.

Soient A un problème de décision, $a, a', a_1, a'_1 \in A$ telles que $\forall c \in R(a), \forall c' \in R(a'), c > c'$.

Soit $c'' \in C$, si $R(a_1) = R(a) \cup \{c''\}$ et $R(a_1^{'}) = R(a') \cup \{c''\}$, alors nous sommes dans l'une des configurations suivantes :

- $c^{''} \geq maxR(a)$. Dans ce cas $minR(a_1) > minR(a'_1)$ et $maxR(a_1) = maxR(a'_1)$. L'axiome (PG) implique alors que $[minR(a_1), maxR(a_1^{'})] \rhd [minR(a_1), minR(a_1^{'}), maxR(a_1^{'})]$. Ce qui signifie $[minR(a_1), maxR(a_1)] \rhd [minR(a_1), minR(a_1^{'}), maxR(a_1^{'})]$ car $maxR(a_1) = maxR(a_1^{'})$. Par le lemme 5.1 $[minR(a_1^{'}), maxR(a_1^{'})] \diamond [minR(a_1), minR(a_1^{'}), maxR(a_1^{'})]$. Nous avons finalement $[minR(a_1), maxR(a_1)] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$ ce qui par le lemme 5.1 signifie $a_1 \rhd a_1^{'}$.
- $minR(a) \geq c^{''}$. Dans ce cas $minR(a_1) = minR(a'_1)$ et $maxR(a_1) > maxR(a'_1)$. L'axiome (PG) implique alors que $[minR(a_1), maxR(a_1^{'}), maxR(a_1)] > [minR(a_1), maxR(a_1^{'})]$. Ce qui signifie $[minR(a_1), maxR(a_1), maxR(a_1^{'})] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$ puisque $minR(a_1) = minR(a_1^{'})$. Le lemme 5.1 donne $[minR(a_1), maxR(a_1)] \diamond [minR(a_1), maxR(a_1^{'}), maxR(a_1)]$. Nous avons finalement $[minR(a_1), maxR(a_1)] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$ ce qui signifie par le lemme 5.1 que

$a_1 \rhd a_1^{'}$.

– $maxR(a) > c^{''} > minR(a')$. Dans ce cas $minR(a_1) > minR(a_1^{'})$ et $maxR(a_1) > maxR(a_1^{'})$. En fonction des positions relatives de $maxR(a_1^{'})$ et $minR(a_1)$, nous pouvons voir que nous avons l'un des sous-cas suivants :
  – $maxR(a_1^{'}) > minR(a_1)$. L'axiome (PG) implique alors que $[minR(a_1), maxR(a_1^{'})] \rhd [minR(a_1^{'}), minR(a_1), maxR(a_1^{'})]$. Le lemme 5.1 donne $[minR(a_1), maxR(a_1^{'})] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$. Par ailleurs, l'axiome (PG) implique alors que $[minR(a_1), maxR(a_1^{'}), maxR(a_1)] \rhd [minR(a_1), maxR(a_1^{'})]$ et par le lemme 5.1 $[minR(a_1), maxR(a_1)] \rhd [minR(a_1), maxR(a_1^{'})]$. La transitivité du préordre implique que nous avons alors $[minR(a_1), maxR(a_1)] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$. Par le lemme 5.1, nous avons $a_1 \rhd a_1^{'}$.
  – $maxR(a_1) > minR(a_1^{'})$. L'axiome (PG) implique alors que $[minR(a_1), maxR(a_1^{'})] \rhd [minR(a_1), maxR(a_1^{'}), minR(a_1^{'})]$. L'axiome (PG) implique aussi que $[minR(a_1), maxR(a_1^{'}), minR(a_1^{'})] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$. Alors par transitivité $[minR(a_1), maxR(a_1^{'})] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$. Un argument similaire donne $[minR(a_1), maxR(a_1)] \rhd [minR(a_1), maxR(a_1^{'})]$. Par la transitivité du pré-ordre nous avons alors $[minR(a_1), maxR(a_1)] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$ et le lemme 5.1 nous permet de conclure que $a_1 \rhd a_1^{'}$.
  – $maxR(a_1^{'}) = minR(a_1)$. L'axiome (PG) implique alors que $[minR(a_1), maxR(a_1^{'})] \rhd [minR(a_1^{'}), minR(a_1), maxR(a_1)]$. Et $[minR(a_1), maxR(a_1), minR(a_1^{'})] = [minR(a_1^{'}), maxR(a_1^{'}), maxR(a_1)]$. (PG) implique alors que $[minR(a_1^{'}), maxR(a_1^{'}), maxR(a_1)] \rhd [minR(a_1^{'}), maxR(a_1^{'})]$. Par le lemme 5.1, nous avons $a_1 \rhd a_1^{'}$.

Ainsi, dans chacun des cas nous avons $a_1 \rhd a_1^{'}$. Nous pouvons donc conclure que (β') est satisfait.

□

L'axiome (β') a été introduit comme une contre-partie stricte de (β). On a montré que la satisfaction simultanée de ces axiomes conduit au raffinement de la famille de critères caractérisée par le théorème 5.1. Un raffinement similaire a été effectué dans [Nehring & Puppe 1996] par l'introduction de l'axiome (PG). On y montre que (PG) raffine, quand l'ensemble des conséquences $\mathcal{C}$ est un espace de Haussdorf, la famille de critères caractérisée par la satisfaction simultanée des axiomes (C) et (I). Il est donc intéressant de s'interroger sur la relation logique entre (PG) et (β').

On peut noter que, même si ces deux axiomes sont différents (point (2) de la proposition 5.15), en présence des axiomes (α) et (β), ils deviennent équivalents.

5.7 Conclusion

Nous avons proposé dans ce chapitre une caractérisation de critères de décision pour la prise de décision sous ignorance complète. Nous avons montré que les critères qui satisfont les axiomes (α) et (β) sont exactement ceux qui prennent en compte les conséquences minimales et maximales des actions.

Notre résultat de caractérisation est proche de celui de [Arrow & Hurwicz 1977]. Les auteurs y caractérisent le même ensemble de critères mais dans le cadre de l'incertitude stricte dans lequel l'agent peut énumérer les états du monde. Ainsi notre travail peut être considéré comme une extension de [Arrow & Hurwicz 1977] pour la prise de décision sous ignorance complète. Et de fait nous partageons avec eux l'axiome $(\boldsymbol{\alpha})$ qui ne dépend pas de la manière dont on représente une action. Remarquons aussi que les deux axiomes utilisés dans leurs théorèmes et qui mettent en évidence l'indépendance de la décision rationnelle vis à vis de la manière dont on nomme les états du monde ou du nombre d'états dans lesquels survient une même conséquence n'ont plus de sens dans le cadre de l'ignorance. Enfin, certains axiomes d'indépendance utilisés dans des travaux ayant pour cadre l'incertitude stricte traitent dans leurs prémisses des situations basées sur les conséquences des actions comparées. Cela n'est plus le cas dans la littérature portant sur le cadre d'ignorance complète où les prémisses des axiomes d'indépendance évoquent des situations de préférences sur les actions elles-mêmes. Notre axiome (β) revient vers cette intuition en partant d'une situation particulière des conséquences de deux actions pour arriver à une conclusion sur la préférence de l'agent entre ces deux actions.

Par ailleurs, dans le cadre d'ignorance complète, nous avons présenté un autre résultat de caractérisation décrit au théorème 2.16. La famille de critères qui y est délimitée semble au premier abord similaire à la nôtre. Toutefois, la proposition suivante montre qu'il n'en est rien. En effet, ni le critère Min ni le critère minmax n'appartiennent à cette famille.

Proposition 5.16

$\trianglerighteq_{min}$ ne satisfait pas (DS).

$\trianglerighteq_{minmax}$ ne satisfait pas (IND).

Preuve

– Soient $x, y \in \mathcal{C}$ telles que $x > y$, alors nous avons clairement $[y] \diamond_{min} [x, y]$ ce qui n'est pas la conclusion prescrite par l'axiome (DS).
– Soient $x, y, z, w, v \in \mathcal{C}$ telles que $x > y, z > w, y > w, z > x, w > v$ alors clairement $[x, y] \triangleright_{minmax} [z, w]$ and $[v, w, z] \triangleright_{minmax} [v, y, x]$ alors que (IND) prescrit une conclusion opposée.

□

Le travail le plus proche du nôtre dans le cadre de l'ignorance complète est celui de [Nehring & Puppe 1996], qui caractérise la même famillle de critères, avec des axiomes différents. Mais leur caractérisation requiert un cadre continu. L'ensemble des conséquences est ainsi supposé être un espace de Haussdorf. C'est une structure riche (et restrictive). Les auteurs justifient l'étude spécifique de cette topologie par le rôle central joué par les préférences continues dans le domaine économique. Ils avancent par ailleurs la justification suivante

> "$\cdots$ we believe that an analysis of the continuous case can also contribute to our understanding of problems which originally arose in a finite framework"

Mais il semble difficile d'exporter leur caractérisation au cas discret (ou fini), comme dans notre cadre. Le fait qu'un critère soit caractérisé par un ensemble d'axiomes incluant l'axiome (C) ne saurait notre point de vue être un argument suffisant pour le retenir dans le cas discret car la preuve de sa caractérisation repose fortement sur la propriété d'arc-connection de l'ensemble des conséquences. En particulier, comme expliqué à la section précédente, les axiomes (I), (IF) et (PG) qu'ils utilisent ne sont pas suffisants pour prouver le résultat dans le cas discret.

Chapitre 6

Extension de la planification classique au cadre multi-agent : une approche par la théorie des jeux

6.1 Introduction

Depuis l'émergence du problème de planification en informatique, la majeure partie des travaux a porté sur l'efficacité et la complexité des algorithmes de planification. Assez peu d'attention s'est portée sur le modèle qui le plus souvent était celui de la planification classique. Ce cadre commun offrait en effet des hypothèses simples et restrictives qui permettaient de se concentrer sur la tâche de trouver des méthodes de plus en plus rapides et efficaces pour planifier. Parmi les hypothèses standard de planification classique figurent le fait que l'agent connaît l'état initial du monde, chaque action possible est déterministe et son résulat peut être parfaitement prédit quel que soit l'état où elle est exécutée, les buts sont binaires (i.e. un état du monde est soit complètement satisfaisant soit complètement insatisfaisant), et le monde est statique dans le sens où la seule manière de le modifier est d'exécuter l'une des actions de l'agent (ainsi, non seulement il n'y a pas d'évènement exogène mais aussi le monde n'a pas de dynamique intrinsèque).

Plus tard dans l'histoire de la discipline est venue une phase où l'on s'est posé la question de l'adéquation entre le modèle et le monde réel, et ce notamment avec l'essor de la robotique dans laquelle les robots ont à se mouvoir dans un monde réel et à interagir avec lui. Dans ces expériences, l'efficacité d'une méthode de planification ne tenait pas seulement à la complexité algorithmique, mais aussi au fait que les concepts du modèle épousent les éléments de l'environnement pertinents pour l'agent. L'ajustement du modèle s'est souvent fait en généralisant le cadre de la planification classique. Nous avons vu au chapitre 3 que ces généralisations du cadre consistaient à relâcher certaines des hypothèses précédemment citées. Cette approche a entre autres l'avantage d'isoler l'effet de chaque hypothèse relâchée. Parmi ces hypothèses, la plus questionnée est celle de la connaissance qu'a l'agent de son état initial et du résultat de chacune de ses actions. Dans une application réaliste impliquant de faire se mouvoir un robot sur une planète lointaine, présumer que l'on connaît parfaitement l'état du monde après chaque action semble en effet une hypothèse très forte. Nous avons vu que cette incertitude sur les actions pouvait avoir plusieurs origines appelant chacune une représentation et une résolution adaptée du problème. Parmi ces sources existe celle de contingences incontrôlables, i.e., d'évènements exogènes qui peuvent influer sur le résultat de l'exécution mais que l'agent ne peut contrôler. Dans ce cas, l'agent doit s'appuyer sur les informations qu'il possède sur les résultats possibles pour former sa décision. Il existe dans la classe des problèmes induite par cette hypothèse une sous-classe, particulièrement intéressante pour le problème de planification, qui est le sujet de cette partie. C'est le cas où la présence d'autres agents planificateurs introduit de l'incertitude dans l'environnement de l'agent. Du point de vue d'un agent, la prise en compte la plus "naïve" de l'incertitude sur ses actions est de considérer que les contingences sur l'exécution du plan sont le fait de la nature prise dans le sens qu'on lui attribue en théorie des jeux. Dans cette théorie, la nature est un joueur (agent) comme les autres possédant des actions qu'il peut exécuter, sauf qu'il n'a aucun intérêt propre. Il ne poursuit aucun but particulier. Les actions sont modélisées comme choisies selon une loi sur laquelle on a plus ou moins d'information. Supposons qu'à la place elles soient dues à la présence d'autres agents possédant une utilité personnelle ou du moins une relation de préférence sur les états du monde. Ils poursuivent donc par leurs actions des buts propres. Supposons aussi qu'ils aient une faculté de raisonnement capable de prendre en compte les autres agents autour d'eux. Plus précisément, ils sont aptes à concevoir les choix des autres agents et à les anticiper. Dans cette situation, une incertitude vient de l'interaction stratégique entre les plans des agents présents au moment de l'exécution. Ainsi chaque agent ignore les plans qui vont être exécutés par les autres. Il doit choisir le meilleur plan à exécuter en tenant compte du fait que l'autre (ou les autres agents) fait (font) de même. L'agent choisit alors en fonction du résultat escompté de l'exécution simultanée de son plan et de ceux des autres. Il se comporte comme un joueur au sens de la théorie des jeux.

L'incertitude peut donc avoir plusieurs causes :
- actions intrinsèquement non déterministes
- choix des autres agents

Si nous avons à prendre des actions ponctuelles comme en théorie de la décision, écarter ces deux causes suffit à connaître parfaitement le résultat d'une action. En planification, on est dans un cas quelque peu différent. Dans ce cadre, l'agent exécute (ou du moins choisit) une suite d'actions. Considérons-en deux successives a et b. Avant que ne se finisse l'action a et ne commence l'action b, une action c du plan d'un autre agent peut venir s'intercaler entre les deux. L'état du monde au moment de l'exécution de b s'en trouve changé. Ce n'est plus en effet celui après exécution de a mais après exécution de la suite $a.c$. Si l'on reste dans le cadre de la planification classique, sans observation, cela revient à dire que l'agent n'est plus sûr du résultat de son action a sur lequel il s'était basé pour choisir l'action b. Cette incertitude introduite sur un point du plan induit une incertitude sur le résultat de l'exécution simultanée de deux plans, pourtant chacun linéaire et formé par des actions déterministes.

Cette spécificité de la planification par rapport à la théorie de la décision appelle un examen particulier. L'agent se trouve bien dans une situation de jeu telle que définie dans l'introduction du chapitre 4. Mais c'est un jeu particulier car à chaque profil de plan correspond un ensemble de résultats possibles. Partir du cadre de la planification classique est un bon moyen d'isoler l'effet de cette incertitude particulière sur la décision de l'agent. Se concentrer sur des plans linéaires peut être justifié dans ce cadre lorsque l'agent en charge de l'exécution (qui peut être différent de l'agent qui construit le plan) ne peut observer l'environnement et ainsi ne peut adapter son plan aux événements extérieurs (i.e. aux actions des autres agents), ou lorsqu'il peut observer l'environnement mais ne peut replanifier dynamiquement à cause d'un manque de ressources calculatoires ou la présence de contraintes temps réel (c'est le cas par exemple d'agents autonomes et mobiles comme des drones volant à grande vitesse, ou des infobots - agents logiciels - devant agir sur des marchés hautement volatiles).

Une autre question à se poser quand un agent fait face à un choix dont chaque alternative implique plusieurs conséquences possibles est, et c'est le point central de la théorie de la décision, de circonscrire l'information dont dispose l'agent à propos de ces différentes conséquences. Dans le modèle que nous proposons, et pour rester en cohérence avec le but affiché de coller le plus possible à un environnement réel dans lequel les agents ont souvent très peu d'information sur les conséquences de leurs actions, nous supposons que la seule information disponible est l'ensemble des mélanges possibles. Cela signifie, puisque de tels plans sont linéaires, que l'agent connaît les résultats de plans possibles consécutifs au choix d'un plan par chacun des agents. L'agent se trouve donc dans une situation tout à fait similaire à celle du choix dans l'ignorance présenté dans le chapitre 1.

Par ailleurs, une des grandes caractéristiques de la planification impliquant plusieurs agents agissant dans un même environnement est la possibilité qui leur est offerte de se coordonner. Le résultat de cette coordination est un plan commun (dans notre cas, linéaire). Grâce à la coordination, l'incertitude causée par l'interaction est dissipée. Toutefois, il n'est pas toujours dans l'intérêt d'un agent de se coordonner. Les agents que nous considérons cherchent avant toujours leurs propres satisfactions qui sont conditionnées par le fait d'atteindre leurs buts. Si les buts des agents ne sont pas compatibles (si par exemple tout état du monde satisfait au plus un agent) alors aucun état ne pourra faire l'objet d'un accord.

6.2 Préférences dichotomiques

6.2.1 Introduction

Exemple 6.1 Deux agents, un peintre et un électricien, agissent dans une même pièce. L'ampoule doit être changée (ce qui est le but de l'électricien) et le plafond doit être peint (ce qui est le but du peintre). L'électricien a une nouvelle ampoule et le peintre a le matériel nécessaire à la peinture. Il y a une seule échelle dans la pièce (l'échelle est donc une ressource critique). De plus, le peintre a besoin de lumière pour peindre. L'électricien possède trois actions Prendre-Echelle-Electricien (PEE), Changer-Ampoule (CA), Reposer-Echelle-Electricien (REE), et le peintre trois actions : Prendre-Echelle-Peintre (PEP), Peindre (P), Reposer-Echelle-Peintre (REP). Peindre réussit seulement si Changer-Ampoule a déjà été exécuté. Prendre-Echelle-Peindre réussit seulement si l'échelle est disponible (i.e., elle a été reposée auparavant). Les interactions suivantes peuvent être facilement envisagées :

- si le peintre prend l'échelle en premier, il ne sera pas capable d'atteindre son but (l'ampoule doit être changée avant) ; s'il ne repose pas l'échelle, l'électricien ne sera pas capable d'atteindre son but.
- si l'électricien prend l'échelle en premier, il sera capable d'atteindre son but ; alors, le peintre sera capable d'atteindre son but si et seulement si l'électricien repose l'échelle. En conséquence, si les deux agents peuvent se coordonner pour exécuter le plan joint PEE.CA.REE.PEP.P, alors les deux agents seront satisfaits.

Dans cette partie, nous étendons le cadre de la planification classique à un cadre de planification multi-agent, i.e., nous considérons un groupe d'agents. Chaque agent possède ses propres actions et buts. Les agents agissent dans un environnement commun. Nous suggérons de traiter ce cas en utilisant des concepts de théorie des jeux qui permettront à l'agent de construire un diagnostic stratégique exprimant ses chances d'atteindre ses buts étant données les interactions possibles avec les autres agents. Nous supposons que l'agent connaît les buts de chaque agent du groupe, ainsi que les plans que chaque agent peut proposer. C'est un hypothèse commune en théorie des jeux.

Dans ce chapitre, nous allons aborder les questions suivantes. Pour chaque agent du groupe, quels sont ses "meilleurs" plans ? Est-ce qu'un plan donné requiert une coordination afin d'être exécuté d'une manière satisfaisante pour les deux agents ? En nous concentrant principalement sur le cas de deux agents et en considérant seulement des buts binaires, nous montrons comment un jeu peut être associé à n'importe quel problème de planification multi-agent ; en conséquence, les "meilleurs" plans pour un agent rationnel peuvent être caractérisés en utilisant des notions de théorie des jeux, spécialement l'équilibre de Nash. Nous identifions aussi les scénarios pour lesquels une coopération entre agents est opportune et montrons comment plusieurs informations stratégiques peuvent être dérivées du jeu sous forme stratégique.

6.2.2 Un cadre formel pour la planification multi-agent

On considère un groupe d'agents N= $\{1, 2, \cdots, k\}$, où chaque agent est identifié par un entier. Soit S un ensemble fini non vide d'états abstraits. Notons s_0 l'état initial, supposé être l'état actuel du monde. s_0 est connu par chacun des agents de N. Chaque agent est associé à un ensemble d'actions :

Définition 6.1 (action) *Une action α est une application de S dans S. L'ensemble des actions de l'agent i est noté A^i.*

Dans la suite, une action sera notée par une lettre grecque ($\alpha, \beta, \cdots$). La définition précédente impose que les actions soient déterministes et toujours exécutables. Cette dernière hypothèse n'est pas excessive. En effet, si l'on veut modéliser le fait qu'une action n'est pas exécutable si l'état du monde est s, on peut typiquement la représenter par une action qui ne change pas l'état du monde dans cet état, i.e. $\alpha(s) = s$, ou qui conduit à un état "puits", i.e., $\alpha(s) = s_\perp$, avec $s_\perp$ un état qui a la pire évaluation par rapport aux buts de l'agent et tel que $\beta(s_\perp) = s_\perp$ pour toute action β. A partir de son ensemble d'actions, chaque agent peut construire des plans :

Définition 6.2 ***(plan)*** *Soit A un ensemble d'actions. Un plan p sur A est une suite (finie mais possiblement vide) d'actions de A, i.e., $p = \alpha_1.\alpha_2.\cdots.\alpha_n$, où chaque $\alpha_i \in A$. Sémantiquement, c'est une application de S dans S, définie à partir de la composition de ses actions, i.e., pour toute action $\alpha \in S$, p(s) = s si $p = \epsilon$ (la séquence vide), et $p(s) = \alpha_n(\cdots(\alpha_1(s))\cdots)$ autrement. L'ensemble des plans sur A est noté A^*.*

Soit un plan $p = \alpha_1.\alpha_2.\cdots.\alpha_n$. Un sous-plan de p est une sous-suite de ses actions, i.e., $p' = \alpha_1'.\cdots.\alpha_m'$ est un sous-plan de p si et seulement si il existe une application strictement croissante t de $\{1, \cdots, m\}$ dans $\{1, \cdots, n\}$ telle que $\forall q \in \{1, \cdots, m\}$, $\alpha_a' = \alpha_{t(q)}$. Soit un autre plan $p' = \beta_1.\cdots.\beta_k$, $p.p'$ dénote la concaténation de p et p', i.e., $p.p'$= $\alpha_1.\cdots.\alpha_n.\beta_1.\cdots.\beta_k$. Enfin, si p est un plan sur A, $A(p)$ dénote le sous-ensemble de A formé par les actions de p.

Les buts d'un agent sont exprimés d'une manière qualitative au moyen d'une relation de préférence (un pré-ordre) sur l'ensemble des états : $\preceq_G{}^i \subseteq S \times S$. Ainsi, pour $s, s' \in S, s \preceq_G{}^i s'$ signifie que pour chaque agent i, l'état s' est au moins aussi préféré que s. Quand S est fini, chaque préordre $\preceq_G{}^i$ sur S peut être représenté par une fonction réelle G^i telle que pour tout $s, s' \in S$, $s \preceq_G{}^i s'$ si et seulement si $G^i(s) \leq G^i(s')$. Dans la suite, nous allons souvent nous concentrer sur le cas binaire dans lequel les états sont divisés entre états buts et états non-buts :

Définition 6.3 ***(préférences dichotomiques)*** *On dit qu'un agent i a des préférences dichotomiques $G^i \subseteq S$ si et seulement si sa relation de préférence $\preceq_G{}^i$ est telle que $s \preceq_G{}^i s'$ si et seulement si $s' \in G^i$ ou $s, s' \notin G^i$. Nous utiliserons la notation $G^i(s) = 1$ si $s \in G^i$ et $G^i(s) = 0$ si $s \notin G^i$.*

Assez naturellement, toute relation de préférence sur les états induit une relation de préférence sur les plans :

Définition 6.4 ***(préférence sur les plans)*** *Soit i un agent, A^* un ensemble de plans, s_0 un état et $\preceq_G{}^i$ une relation de préférence sur les états. La relation de préférence $\leq^i$ sur A^* est définie comme suit : pour tout $p, p' \in A^*$, $p \leq^i p'$ si et seulement si $p(s_0) \preceq_G{}^i p'(s_0)$.*

La qualité d'un plan est donnée par la qualité de l'état atteint ; comme en planification classique, des critères additionnels (e.g., le coût du plan) peuvent être utilisés pour discriminer les meilleurs plans ainsi définis. Dans plusieurs cas, il est raisonnable de supposer que seulement un sous-ensemble Π^i de

A^{i^*} est envisagé par l'agent i ; en particulier, à cause de capacités de calcul limitées, les plans dont la longueur excède un seuil donné peuvent être éliminés. Toutefois, cela a du sens de considérer que Π^i est clos pour les sous-plans, i.e., quand un plan p appartient à Π^i, alors tout sous-plan de p appartient aussi à Π^i. En particulier, le plan vide ϵ appartient toujours à Π^i. Nous sommes maintenant prêts à définir la notion de représentation d'un agent et celle de problème de planification multi-agent :

Définition 6.5 ***(représentation d'un agent)*** *Chaque agent $i \in N$ est caractérisé par un triplet $\mathscr{A}^i = \langle A^i, \Pi^i, \preceq_G{}^i \rangle$ formé par un ensemble d'actions A^i, un ensemble de plans $\Pi^i \subseteq A^i$ et une relation de préférence $\preceq_G{}^i$.*

Définition 6.6 ***(problème de planification multi-agent)*** *Un problème de planification multi-agent (MAPP) pour un ensemble N d'agents est un triplet $\langle S, s_0, \{\mathscr{A}^i | i \in N\} \rangle$ formé par un ensemble d'états S, un état initial $s_0 \in S$ et un ensemble de représentations d'agents $\mathscr{A}_i$. Un MAPP avec préférences dichotomiques est tel que chacun des agents possède une structure de préférences dichotomiques.*

Lorsque chaque agent a choisi un plan, la suite d'évènements correspondant à leur exécution jointe est l'un de leurs mélanges, sauf si une coordination est réalisée. Nous notons $\oplus$ l'application de $A^* \times A^*$ dans 2^{A^*} qui associe à chaque paire de plans p_i et p_j, l'ensemble contenant leurs mélanges :

Définition 6.7 ***(mélange, ensemble de mélanges)*** *Soit $p_i = \alpha_1^i. \ldots .\alpha_n^i \in A^i, p_j = \alpha_1^j. \ldots .\alpha_p^j \in A^j$. Alors $p_i \oplus p_j$ est l'ensemble de plans p qui sont des permutations de $p_i.p_j$ pour lesquel p_i et p_j sont des sous-plans. Chaque p est appelé un* mélange *de p_i et p_j et $p_i \oplus p_j$ est appelé l'* ensemble de mélanges *de p_i et p_j.*

Exemple 6.2 Reprenons l'exemple 6.1 avec p_1 le plan de l'éléctricien : PEe.CA et p_2 le plan du peintre : PEp.P. Alors $p_1 \oplus p_2 = \{$PEe.CA.PEp.P, PEe.PEp.CA.P, PEe.PEp.P.CA, PEp. PEe.P.CA, PEp.P.PEe.CA, PEp.PEe. CA.P$\}$.

Observons que $\oplus$ est une fonction permutative (i.e., commutative et associative). Il s'en suit que les définitions précédentes de mélange et d'ensemble de mélanges peuvent être facilement étendues au cas où $n > 2$. Observons aussi que ϵ (la suite vide) est un élément neutre pour $\oplus$. Dans le cas déterministe avec un seul agent, évaluer un plan est une tâche facile. L'état prédit est le résultat de l'exécution du plan. Caractériser un meilleur plan est aussi facile pour l'agent considéré : le plan est d'autant meilleur que l'état atteint l'est.

Dans le cas non déterministe, l'agent doit considérer tous les états possiblement atteints et agréger leurs scores afin d'évaluer le plan (plusieurs fonctions d'agrégation peuvent être utilisées, e.g. min (critère de Wald [Wald 1950]) pour traduire le comportement d'un agent pessimiste, ou un critère d'utilité espérée quand les scores sont quantitatifs et les actions non déterministes sont données par des ensembles de distributions de probabilité).

Dans le cas multi-agent (quoique déterministe), la situation est similaire à celle du cas non-déterministe à un agent dans le sens où chaque agent doit considérer tous les états possiblement atteints pour évaluer ses plans. La différence principale vient de la nature de l'incertitude : dans notre cadre, l'incertitude vient de l'interaction avec les plans fournis par les autres agents. En conséquence, chaque agent doit exploiter le fait qu'il connaît les représentations des autres agents (il connaît les buts des agents ainsi que leurs plans) afin de déduire quel est son "meilleur" plan. Il diffère en cela du cas non déterministe où l'incertitude vient de l'impossibilité de prédire précisement le résultat de certaines actions, comme "tirer à pile ou face". Dans plusieurs cas, une telle impossibilité résulte d'évènements extérieurs (sur lesquels notre connaissance est imparfaite), qui ne peuvent être totalement observés ou prédits et qui ont un certain effet sur le monde. Par exemple, dans le cas de planification des mouvements d'un robot, l'effet normal de l'action "avancer(1m)" est d'avancer le robot d'un mètre ; toutefois, il se peut que cet effet normal ne se produise pas : si le sol est mouillé (et que cela ne puisse pas être observé), un effet exceptionnel de "avancer(1m) sera d'avancer le robot de 0.5 mètre, seulement. Toutefois, dans la section 6.2.5, nous expliquerons comment la planification robuste, qui traite le problème de trouver un plan robuste dans un cadre non déterministe, peut être exprimée dans notre cadre.

Exemple 6.3 Si le peintre dans l'exemple 6.1 propose le plan p = PEp.P.REp , il est seulement assuré que les actions de p seront exécutés dans l'ordre désiré. Alors qu'il connaît la représentation de l'électricien, il ne sait pas quel plan l'électricien va proposer (en effet, l'ensemble des plans possibles n'est en général pas un singleton). Même si cet ensemble est un singleton, le peintre ignore encore l'ordre d'exécution, i.e., comment son plan va s'intercaler avec celui de l'électricien. Supposons que l'électricien propose le plan p' = PEe.CA.REe, le plan joint qui va être finalement exécuté peut être n'importe quel plan de $p \oplus p'$. L'incertitude résultante disparait dès que les deux agents se coordonnent pour exécuter un plan commun comme p''= PEp.P.REp.PEe.CA.PEe.

Dans notre approche, une tâche capitale pour chaque agent est celle d'évaluer l'interaction de ses plans avec ceux des autres agents. Formellement, cela requiert l'évaluation de chaque ensemble de mélanges. Considérons un de ces mélanges. Savoir que c'est ce mélange qui s'exécutera signifie que l'agent connaît non seulement l'ordre d'exécution de ses actions (son plan) mais aussi l'ordre dans lequel les actions de l'autre agent viennent s'intercaler entre les siennes. On est encore sous les hypothèses de la planification classique. Cela veut dire qu'il connaît le résultat de l'exécution de chacune des actions de la suite d'actions formée par ses actions et celles de l'autre agent étant donné qu'il peut maintenant déterminer avec certitude l'état du monde dans lequel elle s'exécute. Il connaît alors l'état du monde dans lequel va le conduire l'exécution du mélange. Dans le cadre de la planification classique, son évaluation du mélange est celle de l'état du monde auquel il aboutit. Quand chacun des agents choisit un plan, le résultat obtenu est celui de l'exécution de l'un des mélanges. Ce que sait un agent, c'est que sa satisfaction appartiendra à l'ensemble formé par les satisfactions résultats des mélanges. Il doit pouvoir comparer les satisfactions générées par le choix de deux plans. La théorie des jeux préconise que l'agent doit aussi tenir compte des satisfactions des autres agents. Nous proposons pour l'y aider une vue compacte de l'évaluation des mélanges pour tous les agents du groupe. A cette fin, nous associons à chaque ensemble de mélanges son profil de satisfaction (SP).

Un SP est la représentation graphique des points résultats possibles de l'exécution d'un profil de plans des agents. Chaque point traduit, pour un résultat, la satisfaction des agents si le mélange correspondant est réalisée. L'idée est de considérer les échelles de préférences des agents comme les axes d'un référentiel dans lequel on place les points résultats. La satisfaction d'un agent est la projection sur l'axe de l'agent du profil de satisfaction. Dans le cas binaire elle peut être $\{0\}$, $\{1\}$, ou $\{0, 1\}$. Remarquons que cette représentation ne nécessite pas que les échelles des agents soient identiques ni même comparables.

Expliquons comment construire un profil de satisfaction pour un groupe de deux agents ayant des buts binaires. Étant donnée un couple de plans, $p_i \in \Pi^i$ et $p_j \in \Pi^j$, chaque mélange de l'ensemble de mélanges $p_i \oplus p_j$ est un plan construit à partir des actions des deux agents ; l'exécution d'un tel plan conduit à un état spécifique qui est plus ou moins satisfaisant pour chaque agent. L'évaluation d'un plan dépend de l'état résultant de son exécution. On peut représenter l'évaluation d'un ensemble de mélanges par les agents en utilisant une représentation sur 2 axes (chaque axe exprime la satisfaction de l'agent correspondant) qui associe un point de coordonnées (x,y) à un mélange p ssi $G^i(p(s_0)) = x$ et $G^j(p(s_0)) = y$. Notons qu'une telle représentation peut être facilement généralisée au cas de n agents.

Définition 6.8 ***(profil de satisfaction)*** *Soit un MAPP avec buts binaires pour un ensemble $N = \{1, \ldots, m\}$ d'agents, avec un état intial s_0. Un* profil de satisfaction (SP) *pour l'ensemble de mélanges $p_1 \oplus p_2 \oplus \ldots \oplus p_k$ ($p_i \in \Pi^i$ pour $i \in \{1, \ldots k\}$) est un ensemble $SP(p_1 \oplus p_2 \oplus \ldots \oplus p_k)$ de vecteurs $(x_1, \ldots, x_k)$ vérifiant $(x_1, \ldots, x_k) \in SP(p_1 \oplus p_2 \oplus \ldots \oplus p_k)$ si et seulement si $\exists p \in p_1 \oplus p_2 \oplus \ldots \oplus p_k$ tel que $\forall i \in \{1, \ldots k\}$, $G^i(p(s_0)) = x_i$.*

Quand nous considérons seulement deux agents ayant des buts binaires, l'ensemble des profils de satisfaction possibles est décrit dans la figure 6.1.

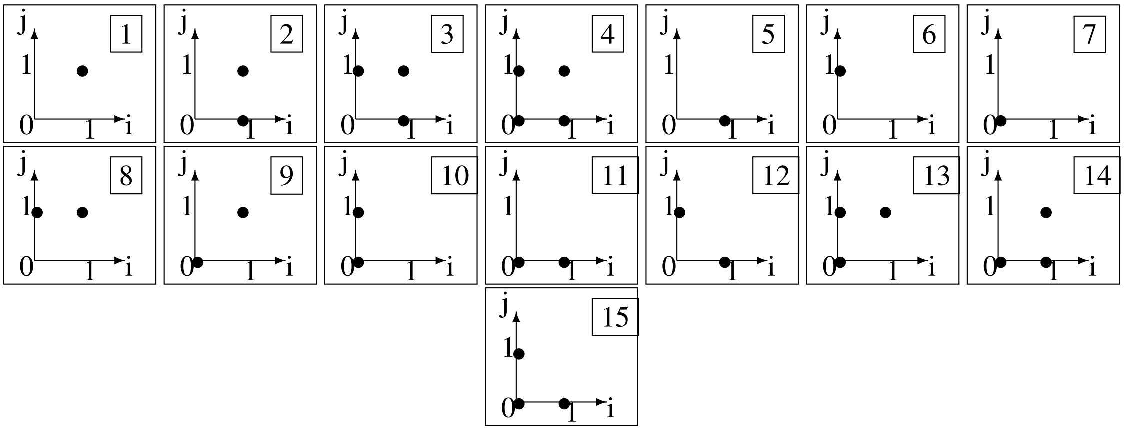


FIGURE 6.1 – SPs possibles dans le cas binaire avec deux agents

Un SP traduit la situation stratégique d'un agent. En plus d'énumérer les résultats possibles d'un choix de plan, il permet de connaître l'attrait des autres joueurs pour chacun de ses résultats. Si chaque agent peut évaluer de son point de vue l'ensemble de résultats, i.e., comparer les SPs entre eux, alors on se trouve dans les conditions d'application des raisonnements de la théorie des jeux. La possibilité de se coordonner sur un résultat particulier introduit un deuxième élément à prendre en compte. L'existence de points intéressants pour tous les agents qui peuvent émerger comme les résultats "réels" d'un SP peut améliorer l'appréciation de ce SP par l'agent. Par exemple, un point de coordination fortement apprécié de la part d'un agent peut contrebalancer un point du SP fortement craint par l'agent (qui ne ferait donc pas partie de la possibilité de coordination). On peut, en examinant selon ces critères les SPs précédents, tirer plusieurs conclusions stratégiques sur la situation d'un agent. Ces conclusions portent sur le fait qu'un agent soit ou pas dans une meilleure situation pour un SP que pour un autre. La situation est meilleure quand, une fois le profil de plan choisi, l'agent est plus enclin à penser qu'il pourra atteindre un état but. Ainsi, dans le schéma précédent, quelques SPs représentent des situations bien meilleures pour un agent donné que d'autres. Clairement, SP 2 dans lequel les mélanges conduisent seulement à des états que l'agent i évalue à 1, est plus intéressant pour lui que SP 10, dans lequel les mélanges conduisent toujours à des états non buts pour cet agent. De plus, considérons SP 3 : pour chacun des deux agents, au moins l'un des mélanges conduit à un mauvais état (i.e., un état non but), et au moins l'un des mélanges conduit à un état but. Ce SP montre aussi l'existence d'au moins un mélange gagnant-gagnant (évalué comme le vecteur (1,1)). Dans un tel cas, si les deux agents sont rationnels (i.e., ils agissent pour changer le monde vers un état but), alors il devraient se coordonner pour exécuter un tel mélange. En effet, la coordination est un moyen d'éliminer l'incertitude. Si les deux agents i et j proposent deux plans $p_i \in A^{i*}$ et $p_j \in A^{j*}$ de manière indépendante, ils courent le risque que l'exécution jointe de $p_i \oplus p_j$ conduise à un état évalué à (0,1) ou (1,0), auquel cas, l'un des agents sera insatisfait. A l'inverse, s'ils se coordonnent et proposent conjointement un plan correspondant à un mélange gagnant-gagnant, ils auront la garantie d'être tous deux satisfaits. Dans une situation correspondant au SP 3, les deux agents ont intérêt à offrir (et accepter) une coordination. En l'absence de plus d'information (comme une distribution de probabilité sur l'ensemble des mélanges), cela a un sens de classer les SPs sur une échelle ordinale. Prenons pour cela le point de vue de l'agent i et montrons comment les SPs peuvent être rassemblés et ordonnés :

- **Toujours Satisfait** SP 1,2,5. Pour ces SPs, l'agent i est assuré d'atteindre ses buts même si l'agent j n'accepte aucune coordination. C'est le cas le plus favorable pour i. L'incertitude sur les mélanges n'est pas significative pour l'agent. Qu'importe en effet le mélange, il sait qu'il sera satisfait.
- **Intérêt Mutuel** SP 3,4,9,13,14. Pour chacun de ces SPs, une certaine exécution jointe est bénéfique et d'autres non (pour les deux agents), mais ils partagent tous le vecteur (1,1), ce qui signifie que si les deux agents se coordonnent, il peuvent tous deux atteindre leurs buts. Le vecteur (1,1) émerge comme un résultat plus probable que les autres. Toutefois la coordination peut échouer. Cette situation n'est pas aussi intéressante pour l'agent i que la précédente, ne serait-ce que parce qu'elle est moins sûre.
- **Dépendance** SP 8,11. Pour ces SPs, l'évaluation des mélanges par l'autre agent est constante. Cela signifie que, a priori, il n'y a aucune raison pour l'autre agent d'accepter une coordination afin d'aider l'agent i à atteindre son but. Le vecteur (1,1) est encore un résultat probable. L'autre agent n'a pas de raison de refuser de se coordonner sur ce point. La situation est moins favorable à l'agent i que la précédente étant donné que l'autre agent n'a pas non plus d'incitation à se coordonner.
- **Antagonisme** SP 12,15. Ces SPs reflètent des situations plus problématiques étant donné que les intérêts des deux agents sont clairement distincts. Cela signifie que si l'un est satisfait, alors l'autre ne l'est pas (i.e. la coordination (1,1) n'est pas une option). Dans de tels cas, l'agent i peut juste espérer que l'exécution jointe lui sera favorable sans qu'il n'y ait coordination.
- **Toujours Insatisfait** SP 6,7,10. Quelle que soit la suite des évènements, l'agent i sera insatisfait (aucune exécution jointe ne permet à l'agent d'atteindre son but). De tels SPs sont clairement les pires pour l'agent i.

Notre thèse est que, en l'absence d'information supplémentaire, une telle classification est la plus rationnelle. Par conséquent, nous considérons que chaque agent possède les préférences suivantes sur les classes d'ensembles de mélanges :

Toujours Satisfait> Intérêt Mutuel > Dépendance > Antagonisme> Toujours Insatisfait

X>Y signifie que les SPs de la classe X sont strictement préférés à ceux de la classe Y. Tous les SPs d'une même classe sont indifférents. On peut facilement représenter un tel préordre total en utilisant des nombres. Ainsi, nous écrivons $e_i(p_i \oplus p_j) = 4$ si et seulement si $SP(p_i \oplus p_j) \in$ **Toujours Satisfait(i)**, $\cdots$, $e_i(p_i \oplus p_j) = 0$ si et seulement si $SP(p_i \oplus p_j) \in$ **Toujours Insatisfait(i)** (voir Table6.1).

Classe	Evaluation
Toujours Satisfait	4
Intérêt Mutuel	3
Dépendance	2
Antagonisme	1
Toujours Insatisfait	0

TABLE 6.1 – Evaluation des SPs

De telles évaluations $e_i(p_i \oplus p_j)$ peuvent grossièrement être vues comme des utilités, mais elles ne dépendent pas seulement des buts de l'agent i. Notons aussi que les nombres utilisés importent peu, seul l'ordre compte (notre cadre n'est pas quantitatif). Notons finalement que, alors que les définitions à venir vont utiliser des évaluations $e_i(p_i \oplus p_j)$ et $e_j(p_i \oplus p_j)$, elles ont encore du sens quand d'autres évaluations sont utilisées. Ainsi, si quelqu'un est en désaccord avec l'échelle proposée, les définitions suivantes s'appliquent toujours (tant que l'on utilise une évaluation qui permet de comparer tous les couples de plans).

6.2.3 Résolution du jeu et génération de diagnostic stratégique

A partir de la construction précédente, nous sommes maintenant capables d'associer à chaque ensemble de mélanges une évaluation pour chaque agent. Ceci nous permet de modéliser l'interaction entre les plans des agents comme un jeu (non coopératif) sous forme stratégique. En faisant cela, on peut utiliser deux concepts de solutions pour ces jeux : ceux de niveau de sécurité et d'équilibre de Nash. En effet, à chaque MAPP à buts binaires pour un ensemble de deux agents N=$\{i, j\}$, on peut associer un jeu sous forme stratégique, défini par l'ensemble N de joueurs, l'ensemble de stratégies pour chaque joueur (les ensembles Π^i et Π^j de plans dans notre cas), et par une fonction d'évaluation pour chaque joueur qui associe une évaluation à chaque profil de stratégies (les évaluations $e_i(p_i \oplus p_j)$ et $e_j(p_i \oplus p_j)$ pour chaque ensemble de mélanges $p_i \oplus p_j$ dans notre cas).

Exemple 6.4 Considérons le MAPP suivant : $\langle S, s_0, \{\mathscr{A}^i \mid i \in 1, 2\}\rangle$. $\mathscr{A}^1$ = $\langle A^1, \Pi^1 = \{p_1, p_1'\}, \preceq_G^1\rangle$. $\mathscr{A}^2$ = $\langle A^2, \Pi^2 = \{p_2, p_2'\}, \preceq_G^2\rangle$. Supposons qu'il en résulte le SP de la figure 6.2 :

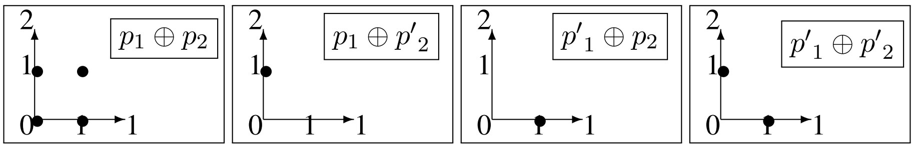


FIGURE 6.2 – Exemple de SPs

On peut maintenant associer un MAPP avec le jeu suivant sous forme stratégique de la Table 6.2.

	p_2	p_2'
p_1	(3,3)	(0,4)
p_1'	(4,0)	(1,1)

TABLE 6.2 – Jeu associé

Une première analyse qu'un agent peut faire est basée sur la notion de niveau de sécurité de ses plans.

Définition 6.9 ***(niveau de sécurité d'un plan)*** *Etant donné un MAPP avec buts binaires pour $N = \{1, 2\}$, le* niveau de sécurité *d'un plan p_i d'un agent i ($i \in N$) face à un ensemble Π^j de plans de l'agent j ($j \neq i$) est défini comme l'évaluation minimum de l'ensemble de mélanges entre le plan p_i et un plan du joueur j, i.e.,*

$$S_{\Pi^j}(p_i) = \min_{p_j \in \Pi^j} e_i(p_i \oplus p_j).$$

A partir des niveaux de sécurité des plans d'un agent on peut définir le niveau de sécurité de l'agent :

Définition 6.10 ***(niveau de sécurité d'un agent)*** *Étant donné un MAPP avec buts binaires $N = \{1, 2\}$, le* niveau de sécurité *de l'agent i face à l'ensemble Π^j de plans de l'agent j est le plus grand niveau de sécurité des plans de l'agent i, i.e.,*

$$S_{\Pi^j}(i) = \max_{p_i \in \Pi^i} S_{\Pi^j}(p_i).$$

Une solution au jeu associé à un MAPP donné peut être définie comme un couple de plans $\langle p_1 \in \Pi^1, p_2 \in \Pi^2 \rangle$ telle que p_1 (resp. p_2) maximise le niveau de sécurité de l'agent 1 (resp. 2) face à Π^2 (resp. Π^1). Une telle solution a du sens dans notre cadre étant donné qu'elle peut être vue comme une analyse au pire cas de l'interaction stratégique. En effet, les SPs sont des ensembles de vecteurs de satisfaction possibles, et comme la classification des SPs que nous avons fournie repose sur une analyse au pire cas, il semble raisonnable d'utiliser les niveaux de sécurité pour comparer des mélanges. Toutefois, les niveaux de sécurité ne prennent pas en compte toutes les opportunités offertes aux agents. Une notion de solution beaucoup plus largement acceptée est basée sur la notion d'équilibre de Nash.

Définition 6.11 ***(équilibre de Nash)*** *Étant donné un MAPP avec buts binaires pour $N = \{i, j\}$, un couple de plans $\langle p_i \in \Pi^i, p_j \in \Pi^j \rangle$ est un équilibre de Nash si aucun des agents ne peut avoir une meilleure évaluation en choisissant un autre plan, i.e., $\langle p_i, p_j \rangle$ est un équilibre de Nash si et seulement si $\nexists p \in \Pi^i$ s.t $e_i(p \oplus p_j) > e_i(p_i \oplus p_j)$ et $\nexists p \in \Pi^j$ s.t. $e_j(p_i \oplus p) > e_j(p_i \oplus p_j)$.*

Exemple 6.5 Revenons au jeu donné à la Table 6.2. Considérons le couple $\langle p'_1, p'_2 \rangle$. L'agent 1 n'a aucun intérêt à dévier seul de ce couple. En effet, $\langle p_1, p'_2 \rangle$ le conduit à une situation moins favorable ($e_1(p_1 \oplus p'_2) < e_1(p'_1 \oplus p'_2)$). De même, $\langle p'_1, p_2 \rangle$ est clairement moins favorable à l'agent 2 que $\langle p'_1, p'_2 \rangle$. Ainsi, on peut conclure que $\langle p'_1, p'_2 \rangle$ est un équilibre de Nash. Il est facile de vérifier que c'est le seul du jeu.

Comme dans le cas général en théorie des jeux, il se peut dans notre cadre qu'un jeu n'ait pas d'équilibre de Nash en stratégie pure [J.Nash 1951], ou qu'il y en ait plusieurs. Quand il y a plusieurs équilibres de Nash, d'autres critères, comme la Pareto optimalité [5], sont souvent utilisés pour les différentier. Les propositions suivantes donnent deux conditions suffisantes à l'existence d'un équilibre de Nash.

Proposition 6.1 *Considérons un MAPP avec buts binaires et deux agents 1 et 2 tel que $G^1 = G^2$. Alors le jeu associé exhibe un équilibre de Nash.*

Preuve
$G^1 = G^2$ implique que pour tout SP s du jeu ; $e_1(s) = e_2(s)$. Cela signifie que les seuls couples d'évaluation possibles sont $(0, 0)$,$(3, 3)$ et $(4, 4)$. S'il existe un SP ayant pour couple de satisfaction $(4, 4)$ alors il est facile de voir qu'il répond à la définition d'un équilibre de Nash car 4 est la plus haute satisfaction possible dans le cadre. Si un tel SP n'existe pas et qu'il existe un SP ayant pour couple de satisfaction $(3, 3)$ alors 3 est la plus haute satisfaction possible dans le jeu et le SP est un équilibre de Nash. Si un tel SP n'existe pas non plus, alors tous les SPs ont pour évaluation $(0, 0$. Chacun est évidemment un équilibre de Nash. □

5. Un vecteur Pareto domine un autre si chacune des composantes du premier est supérieure ou égale à la composante correspondante du second

Autrement dit, si les deux agents partagent les mêmes buts et s'il existe un plan formé sur l'ensemble de leurs actions qui permette d'y parvenir, alors le modèle présenté retient ce plan comme solution.

Notons que, dans notre cadre, le dilemme des prisonniers peut aussi être atteint.

Exemple 6.6 Considérons encore une fois l' exemple 6.4. Le jeu associé (table 6.2) exhibe une situation de dilemme des prisonniers. $\langle p'_1, p'_2 \rangle$ est un équilibre de Nash. Le couple $\langle p_1, p_2 \rangle$ qui est plus profitable que $\langle p'_1, p'_2 \rangle$ pour les deux agents n'est pas un équilibre de Nash (chaque agent est tenté d'utiliser un autre plan).

Au delà de la notion de solution, chacun des deux agents i et j considérés dans le MAPP peut dériver des informations stratégiques à partir du jeu sous forme stratégique associé.

Concentrons-nous maintenant sur les notions de plans robustes, d'effets de synergie, et d'indépendance. Un plan p_i pour l'agent i est *robuste* par rapport à l'agent j si et seulement si son exécution jointe avec n'importe quel plan de l'agent j lui assure d'atteindre son but. Dans le jeu sous forme stratégique, un tel plan correspond à une ligne (ou colonne) pour laquelle toutes les évaluations pour cet agent sont égales à 4 : $\forall p_j \in \Pi^j, e_i(p_i \oplus p_j) = 4$. Assez clairement, un tel plan maximise le niveau de sécurité de l'agent i. Si un plan robuste existe pour un agent i, alors aucune coordination n'est nécessaire avec l'agent j. L'existence d'une synergie entre deux agents peut assez facilement être déduite du jeu sous forme stratégique.

Définition 6.12 ***(effet synergique)*** *Soient deux agents i et j.Un effet synergique pour les agents i et j est possible si et seulement si il existe* $p_i \in \Pi^i$ *et* $p_j \in \Pi^j$ *tel que* $e_i(p_i \oplus p_j) > max_{p \in \Pi^i} e_i(p)$ *et* $e_j(p_i \oplus p_j) > max_{p \in \Pi^j} e_j(p)$.

Assez clairement, aucun effet synergique n'est possible quand au moins l'un des agents possède un plan robuste. La proposition suivante donne une condition suffisante pour assurer qu'un couple de plans $\langle p_1, p_2 \rangle$ exhibant un effet synergique pour les deux agents 1 et 2 soit aussi une solution du jeu :

Une notion d'indépendance entre agents, reflétant le fait qu'il n'y a pas d'interaction entre leurs plans, peut aussi être facilement dérivée du jeu sous forme stratégique.

Définition 6.13 (indépendace) *Soient deux agents i et j. les deux agents sont indépendants si et seulement si* $\forall p_i \in \Pi^i, \forall p_j \in \Pi^j, e_i(p_i \oplus p_j) = e_i(p_i \oplus \epsilon)$ *et* $e_j(p_i \oplus p_j) = e_j(\epsilon \oplus p_j)$.

6.2.4 Exemple : le pont

On considère deux agents 1 et 2. L'agent 1 est en position a et l'agent 2 est en position b. Afin d'aller de a à c, l'agent 1 doit traverser le pont, qui doit être ouvert auparavant (action B^1). Il en va de même pour 2 (action B^2). Si le pont est ouvert pour 1, il est fermé pour 2 et inversement. Chaque agent a une action C^i lui permettant de traverser le pont (mais requiert que le pont soit ouvert pour réussir), e.g., C^1 change la position de l'agent 1 de a à c. L'agent 1 a une action supplémentaire J^1 qui lui permet de sauter par dessus le pont. En utilisant cette action, l'agent 1 n'a pas besoin d'ouvrir le pont. Cette action conduit directement l'agent 1 à la position c.

Clairement, l'agent 1 doit exécuter le plan $B^1.C^1$ ou le plan J^1 afin d'atteindre la position désirée ; l'agent 2 doit exécuter $B^2.C^2$. Étant donné qu'un agent ne peut ouvrir le pont pour un autre, les plans qui ne contiennent pas l'un de ces sous-plans ne peuvent conduire à un état but. Si chaque agent était tout seul dans cet environnement, le problème de planification serait facilement résolu étant donné qu'un agent serait alors sûr d'atteindre son but ($B^i.C^i$ permet d'atteindre le but de i). Ce n'est plus la même histoire

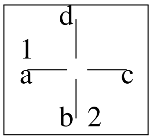


FIGURE 6.3 – Traverser le pont

lorsque les deux agents agissent conjointement. En effet, dans ce cas, une coordination est nécessaire : si l'exécution jointe est $B^1.B^2.C^1.C^2$ l'agent 1 ne pourra pas traverser le pont et ne pourra pas atteindre son but. Représentons le jeu sous forme stratégique associé à ce problème de planification multi-agent. Nous restreignons la longueur des plans examinés à deux actions (observons que les plans de longueur supérieure à 2 sont inutiles). Ce jeu peut être simplifié en supprimant les plans nuls (un plan nul est un plan qui conduit à une satisfaction de 0 quel que soit le mélange dans lequel il est impliqué). Toutefois, nous gardons le plan ϵ dans la version simplifiée, même lorsque c'est un plan nul (voir Table 6.3).

	ϵ	B^2C^2
ϵ	$0,0$	$0,4$
J^1	$4,0$	$4,4$
B^1C^1	$4,0$	$3,3$
B^1J^1	$4,0$	$4,2$
C^1J^1	$4,0$	$4,4$
J^1B^1	$4,0$	$4,2$
J^1C^1	$4,0$	$4,4$

TABLE 6.3 – Jeu sous forme stratégique (simplifié)

Avec n'importe lequel des plans $J^1.B^1$, $J^1.C^1$, J^1, $B^1.J^1$ ou $C^1.J^1$, l'agent 1 a un niveau de sécurité de 4. Comme J^1 est un sous-plan de tous ces plans, les autres plans incluent des actions inutiles. L'agent 1 va probablement choisir le plan J^1. Pour l'agent 2, le seul plan dont le niveau de sécurité est non nul est $B^2.C^2$ (tous les autres plans sont des plans nuls, ils ne peuvent donc conduire à un état but). Ainsi, dans cette situation, le résultat probable du jeu sera le couple de plans $\langle J^1, B^2.C^2\rangle$ qui est évaluée à 4 par chaque agent, ce qui signifie que les deux agents vont sûrement atteindre leurs buts et que cette situation stratégique ne requiert aucune coordination. L'agent 1 peut aussi choisir $B^1.J^1$ au lieu de J^1. Ces deux plans sont pareillement évalués par l'agent 1. Cependant, avec $B^1.J^1$ l'agent 1 peut s'assurer que le plan de l'agent 2 obtiendra une plus faible évaluation (2 au lieu de 4) face à J^1. Si l'agent 1 choisit $B^1.J^1$, il exhibe un comportement agressif par rapport à l'agent 2. On ne développera pas ce point dans la suite, mais il est intéressant d'observer que de telles attitudes peuvent être modélisées dans notre cadre. Les équilibres de Nash de ce MAPP correspondent ici exactement aux solutions obtenues en utilisant la notion de niveau de sécurité.

6.2.5 Relation avec des modèles existants

Nous allons montrer dans ce paragraphe que plusieurs cadres dans lesquels on considère l'interaction entre agents peuvent être facilement vus comme des cas particuliers du nôtre.

– **Planification robuste** En planification robuste (see e.g. [Cimatti & Roveri 1999]), le but est de déterminer si une suite d'actions (i.e., un plan) est robuste, i.e., s'il permet d'atteindre le but pour toutes les contingences possibles. Rappelons en la définition :

Définition 6.14 ***(planification robuste)***
– *Une* action non déterministe α *sur un ensemble fini et non vide S d'états est une application de S dans $2^S \setminus \{\emptyset\}$.*
– *Un* plan non déterministe π *sur un ensemble A d'actions non déterministes est une suite finie d'éléments de A.*
– *Une* trajectoire *pour un plan non déterministe $\pi = \alpha_1.\cdots.\alpha_n$ étant donné un état initial $s_0 \in S$ est une suite d'états $s_0, \ldots, s_{n+1}$ telle que pour tout $i \in 0 \ldots n$, $s_{i+1} \in \alpha_i(s_i)$.*
– *Un plan non déterministe $\pi = \alpha_1.\cdots.\alpha_n$ sur A est* robuste *pour un but $G \subseteq S$ étant donné un état initial $s_0 \in S$ si et seulement si pour chaque trajectoire $s_0, \ldots, s_{n+1}$ pour π, $s_{n+1} \in G$.*

Le problème de la planification robuste peut être facilement exprimé dans notre cadre. Chaque trajectoire du plan non déterministe considéré π peut être vue comme le résultat d'un mélange avec le plan d'un second agent qui joue le rôle de la nature ; considérons la première action α de π et supposons qu'elle possède au plus k effets ; dans ce cas, le plan du second agent va débuter par le sous-plan $\alpha'_1, ..., \alpha'_k$, où α'_j est l'action vide si α n'a pas été exécutée lorsque α'_j est rencontrée (information que l'on mémorise dans les états via un fluent supplémentaire), et produit le $k$ème effet de α sinon ; il reste essentiellement à répéter ce traitement pour les actions suivantes de π en mettant à jour le plan du second agent par concaténation avec les sous-plans produits à chaque étape.
– **bowling** Dans [Bowling & Veloso 2003], les auteurs décrivent un modèle proche du notre. Dans ce travail, les choix des agents se font entre des politiques. Les politiques sont évaluées au niveau du groupe par rapport à chaque agent et les "meilleures" sont caractérisés comme des équilibres de Nash, comme c'est le cas dans notre travail. Cette approche est néanmoins différente de la nôtre par de nombreux aspects :
  – le cadre formel considéré est celui de la planification sous incertitude et observabilité totale et non celui de la planification classique. Des actions non déterministes sont considérées et un ensemble d'états initiaux possibles (et non un seul état) est connu par chaque agent. Les politiques sont des applications associant des actions à des états et non des plans linéaires (suites d'actions), et la qualité d'un plan n'est pas de nature binaire par essence (à l'inverse de ce qui se passe dans le cadre classique).
  – les politiques au niveau du groupe font partie de l'entrée du problème mais les politiques au niveau des agents ne le sont pas (alors que les plans possibles au niveau du groupe sont caractérisés comme des mélanges de plans au niveau des agents dans notre travail).
  – enfin, aucune notion de diagnostic stratégique n'est abordée (en particulier, le besoin de coordination ne peut être déduit de l'entrée considérée).
– **Jeux booléens** Les jeux booléens (voir e.g. [P. Harrenstein & Witteveen 2001] [Dunne & van der Hoek 2004] traitent le cas d'agents contrôlant un ensemble de variables (binaires) propositionnelles. Plus précisément, ce sont des jeux où les utilités des agents sont binaires et les buts sont spécifiés par des formules propositionnelles.

Définition 6.15 ***(jeu booléen)*** *Un* jeu booléen *est un quadruplet $\langle A, V, \Pi, \Phi \rangle$ où $A = \{1 \cdots n\}$ est un ensemble d'agents, V est un ensemble de variables propositionnelles (variables de décision), $\Pi : A \to 2^V$ une fonction d'assignation qui induit une partition $\{\pi_1, \cdots, \pi_n\}$ de V où π_i est l'ensemble de variables contrôlées par l'agent i, $\Phi = \{\phi_1 \cdots \phi_n\}$ un ensemble de formules propositionnelles.*

Pour un joueur $i \in A$, une stratégie est une instanciation des variables qu'il contrôle (i.e., une application de $\Pi(i) = \pi_i$ dans $\{0, 1\}$). Un profil de stratégies P consiste en l'instanciation de toutes les variables considérées et peut être vu comme une application de V dans $\{0,1\}$). Un agent i est satisfait par un profil de stratégies P si et seulement si P est un modèle de Φ_i. On peut exprimer ce cadre dans le nôtre en associant à chaque variable $v \in V$ une action v^+ qui affecte la variable v à 1. A chaque jeu booléen $G = \langle A, V, \Pi, \Phi \rangle$ nous associons un MAPP $\langle S, s_0, \{\mathscr{A}^i \mid i \in 1 \cdots n\} \rangle$ où P est l'ensemble de toutes les affectations de V, s_0 est l'affectation telle que $s_0(v) = 0$ pour tout $v \in V$. Pour chaque agent i, $A^i = \{v^+ \mid v \in \pi_i\}$, Π^i est le sous-ensemble de plans de A^{i^*} tels que chaque action possède au plus une seule occurrence dans chaque plan et G^i est l'ensemble des modèles de ϕ_i.

6.3 Préférences non dichotomiques : satisfaction graduée

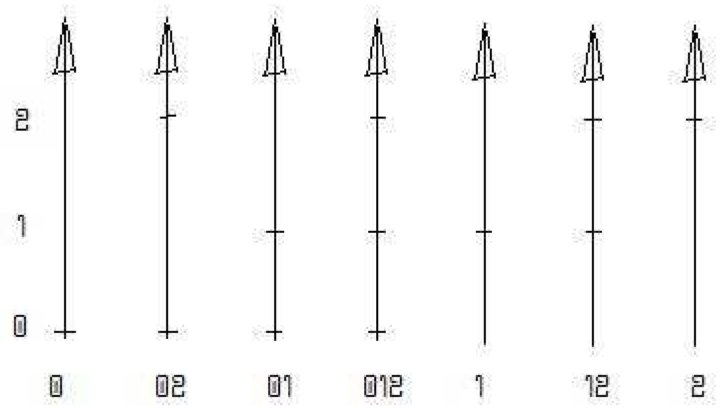
6.3.1 Introduction

Nous nous sommes jusqu'à maintenant conformés à l'hypothèse de préférences binaires. Cela nous a permis de passer en revue un à un tous les cas de profils de satisfaction possibles entre deux agents et d'en saisir de manière aisée et intuitive les caractéristiques permettant de les classer. Si on veut relâcher cette hypothèse, la question clé est : ces caractéristiques restent-elles valables dans le cas où les préférences des agents ne sont plus binaires mais graduées ? C'est une configuration que l'on rencontre souvent dans des applications réelles. Par exemple, un agent dont le but est de réunir 10 euros sera sans doute plus satisfait s'il ne réunit que 9 euros que s'il ne réunit pas du tout d'argent. On peut s'attendre dans de telles situations impliquant de plus plusieurs agents à ce que les cas d'interactions soient plus variés et appellent des critères différents. La théorie de la décision offre les outils nécessaires pour estimer la satisfaction graduée des agents. Moyennant cette remarque, le modèle peut être étendu naturellement à une échelle de préférences plus riche. Pour illustrer cela, gardons la même approche consistant à partir de cas simples en considérant d'abord un cadre dans lequel les échelles de préférences des agents comportent un seul échelon de plus que précédemment. Cette gradation élémentaire de la satisfaction de l'agent permet déjà de constater les paramètres supplémentaires qui interviennent dans la décision de l'agent.

Ce qui apparaît en premier est que le nombre de projections à classer est forcément plus grand. Ensuite, alors que dans le cas de préférences dichotomique, les agents pouvaient raisonnablement se mettre d'accord seulement sur $(1,1)$, si on a plus de deux niveaux de satisfaction, il faut de plus se figurer les points sur lesquels ils vont se mettre d'accord.

6.3.2 Evaluation des ensembles de résultats

Examinons d'abord les configurations possibles. Supposons que l'échelle de préférences d'un agent comporte trois niveaux que nous notons $0, 1, 2$. Le nombre de profils de satisfaction possibles pour deux agents est alors de $2^9 - 1 = 511$ profils (ce sont toutes les parties possibles d'un ensemble à 9 élément sauf l'ensemble vide car le choix des plans aura toujours au moins une conséquence même si c'est "il ne se passe rien"). Si nous augmentons le nombre d'agents ou les niveaux de préférence, ce nombre croît de manière rapide. Cela n'est pas un facteur rebutant de complexité car dans un problème pratique, un agent n'aura pas à classer tous les profils imaginables mais uniquement ceux à sa disposition. De plus, nombre de ces profils sont indiscernables pour la satisfaction immédiate de l'agent. En effet, l'agent les compare suivant leurs projections sur son axe. Étant donné que l'échelle contient trois niveaux, on peut s'attendre à l'une des projections suivantes :



Dans ce schéma, chaque flèche représente une projection possible de l'ensemble de résultats sur l'échelle de préférences de l'agent. Dans la suite on notera chaque projection comme le nombre formé par les préférences (0,1 ou 2) qui la forment. Par exemple 0 représentera la première flèche en partant de la gauche.

Dans le cas général de n niveaux de satisfaction, il y a $2^n - 1$ classes. L'agent doit d'abord ordonner ces classes. La première exigence que l'on peut avoir est que les préférences sur les ensembles de satisfaction soient cohérentes avec les satisfactions "simples" de l'agent. On peut donc poser quand $n = 3$

$$2 \succ 1 \succ 0$$

. Ensuite par extension

$$2 \succ 01$$

car toutes les conséquences dans un profil 12 sont préférées à toutes celles dans 01. Le même argument donne

$$01 \succ 0 \;\text{ et }\; 1 \succ 01 \;\text{ et }\; 2 \succ 12$$

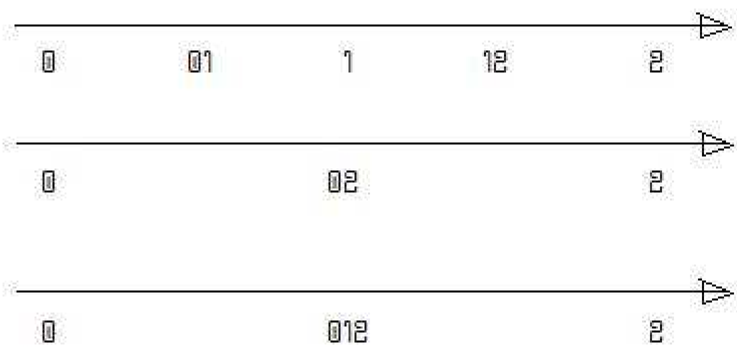
On obtient donc

$$2 \succ 21 \succ 1 \succ 01 \succ 0$$

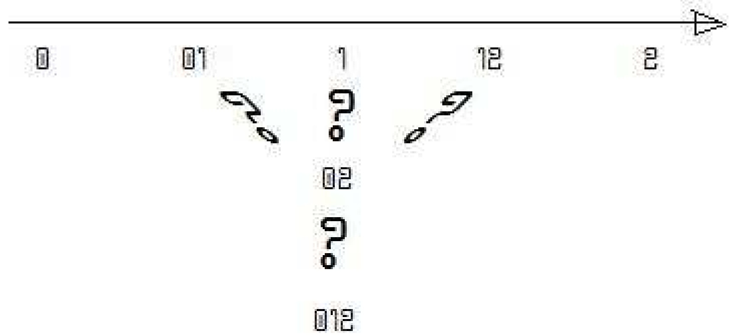
Par ailleurs, le même argument nous permet de conclure :

$$2 \succ 02 \succ 0 \;\text{ et }\; 2 \succ 210 \succ 0$$

En rassemblant ces conclusions, nous obtenons :



Dans ce schéma, chaque flèche représente une échelle de préférence qu'on peut attribuer à un agent à propos des projections possibles. On obtient 3 échelles car on peut par exemple comparer 012 à 2 et 2 à 1 mais pas 012 à 1. On pourrait tout à fait s'arrêter ici. On n'a utilisé jusqu'ici que des arguments très similaires à ceux invoqués pour l'ordonnancement des classes de résultats dans le cas binaire. On peut voir que ces argument permettent à l'agent d'avoir des préférences cohérentes (sans contradiction) même si elles sont incomplètes. On obtient en effet un préordre partiel sur les projections. On peut le représenter dans le schéma suivant qui regroupe les échelles suivantes en une.



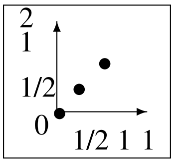
Cette forme de relation n'interdit pas la prise de décision. Par exemple, [Aumann 1962] montre comment construire une fonction d'utilité sans faire appel à l'hypothèse de complétude de la relation de préférence généralement requise. Cette possibilité induit que l'on peut encore utiliser des notions de théories des jeux comme celui de l'équilibre de Nash pour guider la décision de l'agent. Le problème est que l'on ne peut espérer obtenir des décisions fines. Certains cas posent problème car ils marquent la différence avec le cas binaire ; les projections de deux classes à comparer peuvent se chevaucher et se contenir. La question qui subsiste est alors celle du choix que devra faire l'agent quand il sera face à des situations qu'il ne peut comparer. En d'autres termes il s'agit de placer sur une même échelle 02 et 012. Pour aller plus loin et pouvoir obtenir une décision plus efficace, on peut suivre l'approche définie au chapitre 5. On est en effet dans le cadre d'une décision sous ignorance. La décision consiste à choisir la meilleure de deux projections ne sachant pas lequel des éléments des projections seront les résultat réels et n'ayant aucune information sur les chances d'un élément en particulier de se réaliser. Les critères produits par la théorie de la décision sous ignorance sont donc ici utilisables. Dans le chapitre 5, nous avons montré que les critères pertinents sont ceux se basant essentiellement sur une comparaison des éléments extrémaux des résultats possibles d'un choix. Cela revient ici à comparer les valeurs extrémales des projections des différents SPs sur l'axe de l'agent.

Dans cette section, nous avons ébauché une échelle dont peut se servir un agent pour classer les différents SPs qui représentent la situation dans laquelle il se trouve quand chacun a choisi un plan. Cette échelle prend en compte la projection d'un SP sur son échelle de satisfaction. En d'autres termes, il prend en compte les retombées possibles du choix d'un profil de plan sur sa propre satisfaction. D'après les hypothèses adoptées dans le modèle, les agents peuvent de plus se coordonner au sujet d'un SP particulier pour en choisir un point particulier. Cette possibilité peut altérer l'évaluation qu'a un agent d'un SP et doit donc être prise en compte dans la détermination de ses attentes. Dans le cas de préférences dichotomiques par exemple, les agents savent que s'ils se coordonnent dans le profil suivant, ils le feront sur une exécution associée au point $(1, 1)$. Nous dégagerons dans la section suivante les principes servant à guider la prise en compte de cette possibilité. Le premier principe reste similaire à celui utilisé dans le cas binaire, à savoir que les agents se mettront d'accord seulement si cela ne contredit pas leur intérêt propre. Toutefois, de la même manière que l'introduction d'échelons supplémentaires dans la préférence de l'agent appelle à une évaluation plus fine des SPs, le fait que les agents ont le choix pour se coordonner entre plusieurs point du SP requiert une évaluation plus fine des opportunités de coordination.

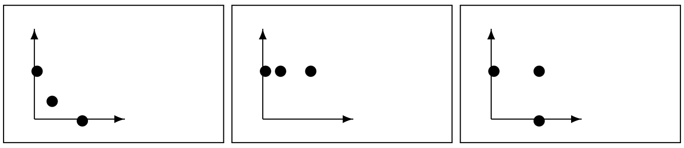
6.3.3 Évaluation des opportunités de coordination

L'objet de cette section est de proposer une approche qui prenne en compte, dans la comparaison, des profils de satisfaction la coordination possible autour d'un mélange particulier.

Dans l'exemple suivant, les points $(1/2, 1/2)$ et $(1, 1)$ peuvent être des points de coordination pour les deux agents. Supposons en effet que les deux agents choisissent deux plans qui induisent ce SP. S'ils ne se coordonnent pas, alors n'importe quel point du SP peut être le mélange atteint réel. D'après les hypothèses retenues, les agents n'en savent pas plus. Ils n'ont pas de raison de penser qu'un point est plus probable qu'un autre. Il se peut tout à fait que se soit le point $(0, 0)$ qui se réalise. La possibilité de coordination leur offre l'occasion d'éviter cette issue. Dans le cas binaire, le point de coordination était tout indiqué (le point $(1, 1)$). Que devient ce point dans le cas où l'échelle de préférence d'un agent comporte plus de deux échelons ? Dans l'exemple présent, le point $(1, 1)$ est meilleur pour les deux agents individuellement. Il semble donc plus intéressant pour les deux agents à la fois. L'intérêt personnel des agents reste donc un principe de base dans le choix d'un point de coordination.



Chaque agent doit donc utiliser un critère pour décider lequel des trois profils suivants est meilleur de ce point de vue. Nous n'en donnons ici que l'allure.



Sur ce point, le troisième profil semble meilleur car le point de coordination possible est meilleur. Qu'en est-t-il du deuxième profil ?

Nous ne présageons pas de la procédure de coordination qui sera utilisée afin que notre modèle soit le plus général possible. Quand les préférences sont dichotomiques, le problème ne se pose pas car s'il y a coordination, elle se fait sur un point tout indiqué. L'agent doit ici évaluer que serait un résultat de coordination "rationnel", c'est-à-dire un résultat que n'importe quelle procédure pourrait admettre. Chaque procédure particulière doit choisir un point dans ce résultat.

Dans le cas que nous traitons ici, nous pouvons visualiser les différentes configuration de points de coordinations possibles. Nous représentons dans les schémas suivants les portions des SPs possibles qui contiennent des points de coordination.

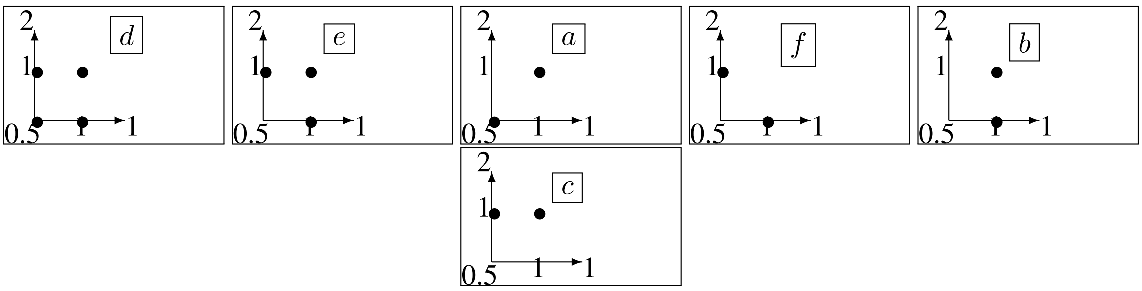


FIGURE 6.4 – Opportunités de coordination dans le cas ternaire

Dans le cas a on peut se dire que si deux agents sont prêts à se coordonner sur le point $(1/2, 1/2)$ alors pourquoi ne le feraient-ils pas sur une trajectoire $(1, 1)$ qui est manifestement meilleure pour les deux ?

Dans les cas b et c, le constat est le même car comme au chapitre précédent, on suppose qu'un agent n'a rien contre le fait de se coordonner s'il n'y perd pas. Dans les cas d et e, on peut déduire que le point de coordination sera $(1, 1)$ en utilisant les deux arguments précédents. Reste le cas f. Dans ce cas, rien ne nous permet de dégager un point particulier sans faire une hypothèse supplémentaire sur le protocole de coordination.

Pour ce qui est de classer les configurations de coordination, l'agent doit donc comparer les projections des points résultats possibles d'une coordination. Cela revient à comparer les projections sur son axe de ces points dans les configurations précédentes. En d'autres termes, il doit comparer $1/2$, 1 et $1/2, 1$. Analysons les caractéristiques du cadre dans lequel on compare ces ensembles. Ce sont des résultats dont seulement un va se réaliser et cela dépendra d'une procédure de coordination sur laquelle on ne suppose rien. On est donc bien dans un cadre d'ignorance (incertitude stricte). Les mêmes principes que précédemment doivent donc guider le classement. Et on arrive donc au suivant

$$1 \succ 1, 1/2 \succ 1/2$$

Nous pouvons voir facilement que si nous étendons encore le cadre au cas de n niveaux, il faudra définir de manière plus formelle et générale l'ensemble des résultats potentiels d'une coordination.

Nous avons donc maintenant une idée de la manière de comparer deux ensembles d'opportunités de coordination. En particulier, nous savons comment évaluer l'effet d'un tel ensemble sur notre satisfaction future. Ce deuxième critère ne peut être raisonnablement pris en compte sans le premier (qui était d'évaluer la projection du SP dans sa totalité) car, en plus du fait que la coordination peut échouer, les agents peuvent simplement ne pas être intéressés par une coordination. La simple existence d'une communication entre les agents peut pousser un agent à envisager le premier cas. Et entre deux SPs présentant des opportunités de coordination similaires, un agent préfèrera celui qui lui sera préférable en cas d'échec de la coordination. Le deuxième cas de figure peut se présenter quand la coordination n'apporte rien de plus à un agent, par exemple quand il est sûr d'avoir la même satisfaction quel que soit le mélange réalisé. Il s'agit donc de faire la balance entre les influences des deux critères (la projection du SP dans sa globalité et celle des opportunités de coordination qu'il offre). Comme on est dans un cadre d'ignorance, il semble incohérent de penser pouvoir donner un poids à chacun des critères. Cela reviendrait en effet à supposer que l'agent possède une estimation probabiliste des chances de réussite d'une coordination. De la même manière que dans le cas binaire, une procédure lexicographique nous semble la plus indiquée. Elle permet de prendre en compte les deux critères. Le deuxième est, en effet, un raffinement du premier plus qu'un critère différent. Les deux estimations partagent comme objet la satisfaction que l'agent est susceptible de recevoir.

6.4 Conclusion

Dans ce chapitre nous avons proposé un modèle pour la planification multi-agent. Ce modèle est destiné à des applications réalistes dans lesquelles un agent manque d'information à propos du monde qui l'entoure, monde qu'il partage avec d'autres agents doués eux aussi de raisonnement et de buts propres. Le modèle pallie l'inadéquation du modèle éprouvé de la planification classique à ce cas en relâchant certaines de ses hypothèses. En particulier, l'agent n'est plus supposé certain du résultat de l'exécution de son plan. Le fait que l'incertitude vienne avant tout de la présence des autres agents incite à utiliser les concepts et raisonnements de la théorie des jeux. Même si les actions sont déterministes et les plans linéaires, les manières dont s'intercalent les actions de deux plans exécutés en même temps peuvent donner des résultats finaux différents. A chaque choix d'un ensemble de plans (un pour chaque agent) correspond un tel ensemble de résultats dont on se propose de donner une représentation graphique. Cette représentation permet d'expliciter les critères servant à évaluer les ensembles de résultats car c'est sur la base de cette évaluation que l'agent fondera sa décision. Le modèle s'appliquant au cas d'agents égoïstes qui recherchent chacun avant tout leur intérêt propre, les agents se préoccupent d'abord de la projection sur leur satisfaction de l'ensemble de résultats. Dans ce but, on peut se baser sur les préceptes de la théorie de la décision sous ignorance. Chaque agent est face à une décision (choix de plan étant donnés ceux des autres agents) qui induit un ensemble de résultat sans qu'il ait à sa disposition par exemple une loi de probabilité régissant la réalisation de ces résultats. Le cadre de la planification multi-agent offre une possibilité supplémentaire aux agents. C'est celle de se coordonner sur un résultat particulier de l'ensemble de résultats. Cette possibilité doit être prise en compte dans le choix d'un plan. Nous proposons le faire de manière lexicographique en examinant d'abord la globalité des résultats possibles, puis, pour des ensembles indiscernables par cette première comparaison, de comparer les ensembles formés par les résultats possibles d'une coordination.

Chapitre 7

Décision multi-agent sous incertitude stricte : jeux à résultats multiples

7.1 Introduction

Le premier chapitre de la contribution a été consacré à la justification axiomatique des critères de décision dans le cas de l'ignorance complète. Il s'agit de guider un agent qui doit choisir entre plusieurs actions complètement incertaines. Plus précisément, on considère le cas où l'agent ne possède aucune information (probabiliste ou même qualitative) sur la conséquence d'une action qui se réalisera finalement. Le deuxième chapitre a ensuite porté sur un cadre particulier de décision. Il s'agit d'un agent qui doit produire un plan (suite d'actions) dans le cas où il partage son environnement avec d'autres agents planificateurs. La présence des autres agents introduit une incertitude sur le résultat d'exécution d'un plan. Nous avons vu que le cadre de la décision sous incertitude offre des outils appropriés à l'agent dans une telle situation. Toutefois, la nature stratégique de la décision à prendre (provenant du fait que les autres agents sont capables de raisonner et forment aussi des plans pour atteindre leurs buts) appelle à l'utilisation des concepts de la théorie des jeux dans le choix du "meilleur plan". Bien qu'il existe des travaux traitant de jeux dont les évaluations des stratégies prennent en compte une incertitude non probabiliste sur l'issue des profils de stratégies, les types de jeux utilisées ne correspondent pas parfaitement au cas considéré.

Afin de pallier ce manque, le but de ce chapitre est d'introduire et d'étudier des jeux qualitatifs à résultats multiples. Plus généralement, dans ce type de jeux, chaque profil de stratégies est associé à un ensemble de conséquences possibles. Cette multiplicité des conséquences fait qu'une donnée des jeux classiques, à savoir la valeur pour le joueur d'un profil de stratégies, essentielle pour la résolution du jeu, devient ici non évidente à évaluer. Si maintenant le joueur pouvait comparer les ensembles de conséquences, au moyen d'une relation d'évaluation qui soit cohérente avec ses préférences sur les conséquences, il pourrait attribuer une valeur aux profils de stratégies associés et sur cette base comparer les stratégies dont il dispose et finalement en choisir une. Il pourrait pour ce faire utiliser les notions classiques de solutions. C'est la voie que nous suivons ici.

Pour chaque profil de stratégies, les joueurs ont la possibilité de se coordonner, c'est-à-dire de se mettre d'accord sur un élément particulier de l'ensemble de conséquences associé. La coordination peut se faire seulement si tous les joueurs sont d'accord. Les motivations pour introduire la coordination dans ce type de jeux sont similaires à celles qui motivent l'introduction de la notion d'équilibre corrélé. L'idée est que les joueurs peuvent discuter à propos de l'ensemble de conséquences associées à un profil de stratégies avant de choisir leurs stratégies, afin de se coordonner pour réaliser une conséquence intéressante pour tous. Quand les joueurs ne veulent pas (ne peuvent pas) s'accorder sur une conséquence, la conséquence effective est un unique élément de l'ensemble de conséquences associé mais on ne peut prédire lequel. On peut s'imaginer pour le concevoir qu'un agent extérieur (e.g. la Nature) choisit une conséquence dans l'ensemble associé. Dans ce cas, les joueurs ne connaissent pas la distribution de probabilités utilisée par cet agent dans son choix. De tels jeux peuvent aussi être utiles pour modéliser des scenarii dans lesquels il est inapproprié d'utiliser des distributions de probabilités sur les conséquences. Halpern et Tuttle soulignent le besoin de tels modèles sans probabilités dans [Halpern & Tuttle 1993]. Ils avancent en particulier que "certains choix dans les systèmes distribués doivent être vus comme intrinsèquement non déterministes (ou, peut être mieux, non probabilistes), et il est inapproprié, à la fois philosophiquement et pragmatiquement, de modéliser de manière probabiliste ce qui est intrinsèquement non déterministe". Les agents se trouvent dans une situation d'ignorance complète à propos de la conséquence effective du choix d'un profil de stratégies. La seule information dont ils disposent est l'ensemble de conséquences associé à chaque profil de stratégies. Dans la suite, nous faisons l'hypothèse supplémentaire que les utilités des joueurs sont qualitatives. Cela signifie que les utilités sont définies sur des échelles ordinales, et pas nécessairement sur un sous-ensemble de réels. En particulier, aucune notion d'intensité de préférence ne peut être dérivée.

Soit l'exemple suivant qui illustre l'intérêt des jeux à résultats multiples sur une version modifiée du dilemme des prisonniers.

Exemple 7.1 Deux suspects sont arrêtés par la police. la police n'a que des preuves indirectes insuffisantes pour les condamner, et offre alors aux deux prisonniers (séparément) un marché identique : si l'un témoigne (stratégie D) contre l'autre et que celui-ci reste silencieux (stratégie C), le traître aura seulement à purger une peine d'une année d'emprisonnement (une satisfaction de 5) et le complice silencieux recevra une peine de 6 années (une satisfaction de 0).

Si les deux se taisent, l'inculpation sera seulement basée sur les preuves et le résultat du procès dépendra alors de la crédibilité du témoignage des suspects durant celui-ci. S'ils ne sont pas suffisamment convaincants, il auront quand même une peine de trois années de prison (une satisfaction de 3), mais s'ils donnent des versions cohérentes (au moyen d'une coordination quelconque), ils sortiront libres (une satisfaction de 6).

Si chacun trahit l'autre, chacun recevra une peine de 4 ou 5 années (une satisfaction de respectivement 2 ou 1), en fonction de leurs témoignages durant le procès. Ainsi, encore une fois, ils peuvent s'accorder sur une histoire qui, même si elle ne les innocente plus, minimise leurs responsabilités et leur assure la peine la plus légère.

Dans ce scénario, les joueurs sont confrontés à des situations menant chacune à différents résultats possibles. Les joueurs peuvent se coordonner sur un résultat particulier. S'il ne le font pas, n'importe quel résultat pourra se réaliser.

Sous forme stratégique, le jeu associé est représenté à la Table 7.1.

	C	D
C	(6,6) (3,3)	(0,5)
D	(5,0)	(1,1) (1,2) (2,1) (2,2)

TABLE 7.1 – Dilemme des prisonniers avec incertitude et coordination possible.

Dans ce jeu, il y a une incertitude sur le résultat obtenu avec le profil de stratégies (C,C), et sur celui obtenu avec le profil de stratégies (D,D). Cette incertitude ne peut être probabilisée et peut être dissipée si les deux agents se coordonnent. Notons que pour presque toutes les distributions de probabilités envisageables sur les résultats, le jeu est un dilemme des prisonniers.

La possibilité de coordination change les choses puisque, grâce à elle, les profils (C,C), et le résultat (6,6) sont réalisables.

Ces jeux qualitatifs à résultats multiples peuvent modéliser nombre de situations décrivant des problèmes de planification multi-agent dans lesquelles il n'y a pas de planificateur central et les agents sont autonomes, ce qui signifie que chacun a ses propres buts et est prêt à se coordonner si et seulement si il y voit son intérêt (sur ce point ces problèmes diffèrent de ceux de partage des tâches).

Dans ce chapitre, nous prenons le point de vue d'un agent qui est sur le point d'agir dans un environnement tel que décrit plus haut ; il y a de l'incertitude liée à la présence d'autres agents qui eux aussi agissent de manière autonome pour atteindre leurs buts. L'agent n'a que des informations minimales sur les autres (leurs actions possibles et leurs conséquences) et il doit prendre ses décisions a priori. Notre modèle répond à la question suivante : quelles décisions stratégiques doit-t-il prendre dans un tel cadre, i.e., quelles actions choisir et si oui ou non il a besoin de solliciter (ou d'accepter) une éventuelle coordination ?

7.2 Jeux qualitatifs à résultats multiples

Nous allons maintenant donner une définition des jeux qualitatifs à résultats multiples que nous proposons. Nous partons d'une définition usuelle d'un jeu sous forme stratégique.

Définition 7.1 (Jeu sous forme stratégique) *Un* jeu sous forme stratégique *est un triplet* $\Gamma = (N, (S_i)_{i \in [1,n]}, \mu)$*. N est un ensemble d'agents, S_i est un ensemble non vide de stratégies pour l'agent i, $S = \times_{i \in N} S_i$ est l'ensemble des profils de stratégies et $\mu : S \longrightarrow I\!R^n$ est la fonction d'utilité du jeu qui associe $s \in S$ à un résultat $\mu(s) = (\mu_1(s), \cdots, \mu_n(s))$. Ainsi, $\mu_i(s)$ est l'utilité de l'agent i pour le profil $s \in S$.*

Dans les jeux à résultats multiples, on associe chaque profil à un ensemble non vide de résultats. En d'autres termes, une fois que chacun des joueurs a choisi sa stratégie, aucun d'eux ne connaît exactement le résultat obtenu (pour lui et pour les autres). Tout ce qu'ils savent est que ce résultat appartient à un ensemble de résultats associé au profil de stratégies.

Définition 7.2 (Jeu à résultats multiples) Un jeu à résultats multiples sous forme stratégique *est un triplet* $\Gamma = (N, (S_i)_{i \in [1,n]}, U)$*. N est un ensemble d'agents, S_i est un ensemble non vide de stratégies pour l'agent i, $S = \times_{i \in N} S_i$ est l'ensemble des profils de stratégies et $U : S \longrightarrow 2^{I\!R^n} \setminus \{\emptyset\}$ est la fonction d'utilité du jeu qui associe $s \in S$ à un ensemble fini non vide de résultats.*

Il est doublement intéressant de noter que la définition précédente peut être étendue quasiment sans modification au cas où les préférences des agents sont seulement qualitatives (i.e., les utilités ne sont pas forcément définies sur $I\!R$). D'abord cela montre que notre cadre peut s'appliquer à un plus grand nombre de situations de décision stratégique. Ensuite, si l'on ne possède pas d'information probabiliste, on ne peut espérer obtenir une échelle numérique de préférences sur les profils de stratégies. Des préférences sur les conséquences effectives ne sont donc pas un handicap et sont même plus cohérentes avec le cadre. En effet, si on est face à des agents possédant une utilité numérique sur les conséquences, c'est seulement l'information ordinale qu'elle porte (une conséquence qui donne une utilité plus grande est plus préférée) qui nous est vraiment utile. Donnons alors la définition suivante, cette fois qualitative, d'un jeu à résultats multiples.

Définition 7.3 (Jeu qualitatif à résultats multiples) *Soient $\chi_1, \cdots, \chi_n$ des ensembles totalement ordonnés par, respectivement $\leq_1, \cdots, \leq_n$. Un* jeu qualitatif à résultats multiples sous forme stratégique *est un triplet $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$. N est un ensemble d'agents, S_i est un ensemble non vide de stratégies pour l'agent i, $S = \times_{i\in N} S_i$ est l'ensemble des profils de stratégies et $U : S \longrightarrow 2^{\chi_1\times\cdots\times\chi_n} \setminus \{\emptyset\}$ est la fonction d'utilité du jeu qui associe $s \in S$ à un ensemble fini non vide de résultats.*

Notons que l'échelle d'utilité $\leq_i$ de l'agent i peut être différente de celles utilisées par les autres agents. On peut alors imaginer un grand nombre de scenarii dans lesquelles une partie des agents possèdes des utilités numériques alors que le reste ne peut exprimer ses préférences que qualitativement. Par exemple des humains en prises avec des machines, ou des joueurs plus ou moins intelligents et disposant d'une information plus ou moins précise. Tant que chaque joueur peut exprimer au moins une préférence ordinale, notre modèle peut s'appliquer. Afin de faciliter la lecture, nous appellerons *nombres* les éléments des ensembles χ_i.

7.3 Résolution

Rappelons que, dans un jeu à résultats multiples, un profil de stratégies ne donne généralement pas un unique résultat, mais un ensemble de résultats. Ainsi l'utilité du joueur pour un profil donné n'est pas un unique nombre, mais un ensemble de tels nombres, un pour chaque résultat possible. Nous appelons cet ensemble la projection du profil pour le joueur.

Dans les jeux bayésiens, les espérances du joueur concernant une stratégie sont calculées comme la valeur moyenne de ses meilleures réponses pour chaque type d'adversaire pondérées par les probabilités de ces types. Dans les jeux prébayésiens, la solution développée dans [Ashlagi *et al.* 2006],[Ashlagi *et al.* 2007] se base aussi sur le concept de meilleure réponse. Une stratégie est une meilleure réponse à une stratégie de l'adversaire si ses espérances pour les types possibles sont les meilleures quand il est confronté à cette stratégie. Les espérances sont calculées pour une stratégie comme le résultat minimum pour le profil de stratégies quand on fait varier le type de l'adversaire. Dans les jeux à résultats multiples, les types ne sont pas pertinents. Tout ce dont on dispose c'est de l'ensemble de résultats associé à chaque profil de stratégies séparément. En particulier, on ne possède pas d'information permettant d'affirmer "tel type d'agent aura telle utilité pour une de mes stratégies et telle utilité pour une autre de mes stratégies". Les attentes des joueurs doivent donc être calculées en revenant à la définition : ce que l'agent peut attendre en jouant une stratégie. Un joueur sait que pour un profil de stratégies donné, il peut s'attendre à avoir l'un des résultats de l'ensemble associé. Ses attentes dépendent donc de ces résultats. Les concepts de solutions classiques en théorie des jeux sont basés sur les notions de domination et de meilleure réponse. Afin d'utiliser ces notions dans notre modèle, nous devons être capables d'estimer les attentes d'un joueur quand aucune probabilité n'est donnée (ou quand le cadre est qualitatif) ce qui signifie alors être capable de comparer des ensembles de résultats. Il n'existe pas de manière unique, canonique, de le faire. Mais plusieurs définitions sont possibles. Toutefois, on peut d'ores et déjà dégager deux points sur lesquels se basera cette comparaison. D'abord, l'agent "voit" un ensemble de résultat à travers le prisme de ses intérêts. Cela signifie que ce qui importe pour lui dans chacun des résultats possibles, c'est la satisfaction qu'il lui rapportera. Celle-ci consiste en la composante correspondant à l'agent dans le vecteur de satisfactions formant le résultat. On peut donc dire que l'agent voit la projection de l'ensemble de résultats sur sa composante. C'est un ensemble de nombres. Il doit donc d'abord savoir comparer des ensembles de nombres. Ensuite, l'opportunité offerte de se coordonner avec les autres agents doit manifestement être prise en compte car elle peut changer l'appréciation que l'agent forme à propos d'un ensemble de résultats en en réduisant les éléments à ceux pouvant être des points plausibles de coordination. La suite de cette section traite de la prise en compte de ces deux points.

7.3.1 Évaluation pour un joueur

Définissons d'abord formellement la notion de projection.

Définition 7.4 (Projection) *Considérons $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soit $s \in S$ un profil de stratégies. La* projection *par rapport à $i \in [1, n]$ de U donne les utilités possibles de l'agent i pour chaque profil $s \in S$. C'est une application $U_i : S \longrightarrow 2^{\mathcal{X}_i} \setminus \{\emptyset\}$.*

Quand aucune distribution de probabilités n'est donnée, les espérances des joueurs sont communément réduites comme dans les jeux prébayésiens au plus mauvais résultat possible, i.e., celui garanti pour la stratégie. Comme exposé plus haut, ce critère de comparaison des stratégies manque de pouvoir prescriptif. On peut alors se tourner vers la théorie de la décision sous incertitude et plus précisément la théorie de la décision sous ignorance complète dont les conditions d'applications correspondent parfaitement au cas présent. La théorie de la décision sous ignorance complète s'applique au cas où l'agent est face à un choix entre des actions à l'issue incertaines dont il connaît pour chacune uniquement l'ensemble de ses conséquences possibles (une seule conséquence par action se réalisera finalement). Nous avons vu au chapitres 2 et 5 que les critères de choix raisonnables sont ceux qui prennent en compte les conséquences extrémales d'une action, i.e., les conséquences qui sont associées aux valeurs minimales et maximales pour l'agent. Redéfinissons dans le cadre présent quelques critères qui remplissent ces conditions. Remarquons que nous utilisons ces critères initialement prévus pour un seul agent dans un cadre multi-agent. C'est légitime puisque nous supposons que le joueur compare des résultats de profils de stratégies (c'est-à-dire le résultat de son action, l'action de son adversaire étant fixée).

Une première possibilité est une comparaison au pire des cas, elle correspond au critère de même nom utilisé en théorie de la décision sous incertitude stricte.

Définition 7.5 (Critère Min) *Soit $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soient s_1 et s_2 deux profils de stratégies. Pour $i \in [1, n]$, $s_1 \leq_{\min_i} s_2$ ssi*

$$\min(U_i(s_1), \leq_i) \leq_i \min(U_i(s_2), \leq_i).$$

Ce critère prudent compare deux profils comme si la pire conséquence pour chacun se réalisait possède un pendant optimiste comparant, lui, les profils selon les meilleures conséquences pour l'agent.

Définition 7.6 (Critère Max) *Soit $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soient s_1 et s_2 deux profils de stratégies. Pour $i \in [1, n]$, $s_1 \leq_{\max_i} s_2$ ssi*

$$\max(U_i(s_1), \leq_i) \leq_i \max(U_i(s_2), \leq_i).$$

Ces critère peuvent être agrégés lexicographiquement en un critère minmax en utilisant le meilleur résultat possible.

Définition 7.7 (Critère minmax) *Soit $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soient s_1 et s_2 deux profils de stratégies. Pour $i \in [1, n]$,*

$$s_1 \leq_{minmax_i} s_2 \text{ ssi } s_1 \leq_{lex(\leq_{\min_i}, \leq_{\max_i})} s_2.$$

Les deux précédents critères définissent des préordres totaux sur les profils de stratégie. Ils permettent alors d'ordonner totalement tous les profils. Le critère suivant est basé sur les relations d'intervalles [Fishburn 1985]. Son principal inconvénient est que le préordre correspondant est en général seulement partiel.

Définition 7.8 (Critère produit) *Soit* $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ *un jeu qualitatif à résultats multiples. Soient* s_1 *et* s_2 *deux profils de stratégies. Pour* $i \in [1, n]$*,*

$$s_1 \leq_{produit_i} s_2 \;\; ssi \;\; s_1 \leq_{\min_i} s_2 \text{ et } s_1 \leq_{\max_i} s_2.$$

En utilisant un de ces critères dont nous avons montré les bonnes propriétés pour le présent cadre au chapitre 5, l'agent est capable de comparer deux ensembles de projections correspondant à deux ensembles de résultats de deux profils de stratégies. Sur cette base, nous sommes prêts à définir l'évaluation finale d'un agent pour les profils de stratégies :

Définition 7.9 (Evaluation d'un agent) *Soit* $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ *un jeu qualitatif à résultats multiples. Soient* $i \in N$ *et* $\leq_c$ *un préordre sur les ensembles de nombres qui est une extension des préférences du joueur* i *sur les projections des résultats. Soit* $\preccurlyeq_i^c$ *l'évaluation pour le joueur des profils. Alors le joueur* i *évalue deux profils* s *et* s' *de* S *de la manière suivante* $s \preccurlyeq_i^c s'$ *ssi* $s \leq_{\gamma_c^i} s'$*.*

Avec cette évaluation on peut maintenant utiliser toutes les notions de la théorie classique des jeux, comme l'équilibre de Nash, pour résoudre ces jeux.

Quand une relation d'évaluation $\preccurlyeq_i^c$ est totale, on peut la représenter en utilisant des réels et alors résoudre le jeu classique sous forme stratégique correspondant. Bien sûr les stratégies mixtes ne peuvent pas être utilisées dans notre cadre puisque l'intensité des préférences ne peut être ici inférée. La préférence entre deux profils de stratégies est seulement ordinale.

Notons que, même si la relation d'évaluation $\leq_{c_i}$ est partielle, on peut encore utiliser les notions de théorie des jeux, notamment l'équilibre de Nash (en stratégies pures), pour résoudre ces jeux. Cela peut être exprimé comme suit :

Définition 7.10 (Equilibre de Nash) *Soit* $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ *un jeu qualitatif à résultats multiples. Un profil de stratégies* s *est un équilibre de Nash si aucun agent ne peut avoir une meilleur évaluation en choisissant unilatéralement une autre stratégie, i.e.,* s *est un équilibre de Nash si et seulement si* $\nexists i \in N$ $\nexists s' \in S_i$ *t.q* $(s_{-i}, s_i) \prec_i (s_{-i}, s')$*.*

Illustrons les notions jusqu'ici introduites à travers la résolution de l'exemple du pont précédemment présenté.

Application à l'exemple du pont Représentons le jeu à résultats multiples associé à ce problème de planification multi-agent. Nous restreignons la longueur des plans considérés à deux actions (observons que des plans de plus grande longueur sont inutiles). Le jeu est donné à la Table 7.2.

Ce jeu peut maintenant être transformé en un jeu qualitatif classique auquel on pourra appliquer des notions de solution de la théorie des jeux telles que l'équilibre de Nash. Cette transformation s'opère comme suit. Pour chaque profil de plans, on identifie d'abord le minimum et le maximum des projections des résultats pour le joueur. En se basant sur ces valeurs, on peut maintenant ordonner les profils. Pour un profil donné, l'utilité qui lui est associée dans le jeu qualitatif est son rang d'après cet ordre. Ainsi, en fonction du critère choisi pour comparer les profils, le jeu qualitatif obtenu peut être différent et conduire à des solutions différentes.

Illustrons ce processus sur deux exemples de profils de stratégie en prenant le point de vue de l'agent 1. Soient les deux profils $S^1 = (B^1C^1, B^2)$ et $S^2 = (B^1J^1, B^2C^2)$ dont les ensembles de résultats sont donnés à la figure 7.1.

Afin de comparer ces deux profils de stratégies (puis les deux ensembles de résultats associés) on considère les projections des ensembles de résultats pour le joueur 1 qui sont donnés à la figure 7.2.

	ϵ	B^2	C^2	B^2C^2	C^2B^2
ϵ	(0,0)	(0,0)	(0,0)	(0,1)	(0,0)
B^1	(0,0)	(0,0)	(0,0)	(0,0) (0,1)	(0,0)
C^1	(0,0)	(0,0)	(0,0)	(0,1)	(0,0)
J^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1)	(0.5,0)
B^1C^1	(1,0)	(1,0) (0,0)	(1,0)	(1,1) (0,1) (1,0)	(1,0) (0,0)
B^1J^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1) (0.5,0)	(1,0)
C^1B^1	(0,0)	(0,0)	(0,0)	(0,1) (0.0)	(0,0)
C^1J^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1)	(0.5,0)
J^1B^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,0) (0.5,1)	(0.5,0)
J^1C^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1)	(0.5,0)

TABLE 7.2 – Jeux à résultats multiples

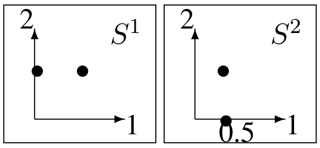


FIGURE 7.1 – Deux profils

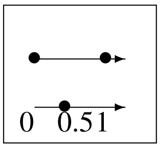


FIGURE 7.2 – Projections des ensembles de résultats

Maintenant pour comparer ces projections on utilise l'un des critères donnés plus haut, e.g., le critère Min, qui donne la relation suivante : $\{0.5\} \geq_{Min} \{0, 1\}$. On peut alors conclure que l'agent préfèrera le profil S^2 à S^1 ce qui est exprimé en assignant une utilité de 1 à S^2 et 0 à S^1 (ou n'importe quels autres nombres compatibles avec cette préférences sur les profils).

Critère Min Nous obtenons le jeu qualitatif donné à la Table 7.3 :

Ce jeu peut être simplifié sans changer la solution en enlevant les plans nuls (un plan est dit nul s'il ne conduit jamais l'agent à un état but quel que soit ce que fait l'autre agent). Toutefois, nous gardons le plan ϵ dans la version simplifiée dans la Table 7.4 même s'il est nul.

Dans ce jeu les profils de plans $(B^1.C^1, B^2.C^2)$,$(J^1, B^2.C^2)$,$(J^1.C^1, B^2.C^2)$,$(C^1.J^1, B^2.C^2)$, $(J^1.B^1, B^2.C^2)$ et $(B^1.J^1, B^2.C^2)$ sont tous des équilibres de Nash. Il n'y a pourtant pas d'équilibre Pareto optimal. L'analyse au pire cas peut aider l'agent 1 à décider d'utiliser le plan J^1 mais n'est d'aucune aide pour l'agent 2. Prédire le résultat final du jeu n'est pas évident.

	ϵ	B^2	C^2	B^2C^2	C^2B^2
ϵ	$0,0$	$0,0$	$0,0$	$0,1$	$0,0$
B^1	$0,0$	$0,0$	$0,0$	$0,0$	$0,0$
C^1	$0,0$	$0,0$	$0,0$	$0,1$	$0,0$
J^1	$1,0$	$1,0$	$1,0$	$1,1$	$1,0$
B^1C^1	$2,0$	$0,0$	$2,0$	$0,0$	$0,0$
B^1J^1	$1,0$	$1,0$	$1,0$	$1,0$	$1,0$
C^1B^1	$0,0$	$0,0$	$0,0$	$0,0$	$0,0$
C^1J^1	$1,0$	$1,0$	$1,0$	$1,1$	$1,0$
J^1B^1	$1,0$	$1,0$	$1,0$	$1,0$	$1,0$
J^1C^1	$1,0$	$1,0$	$1,0$	$1,0$	$1,0$

TABLE 7.3 – Jeu qualitatif

	ϵ	B^2C^2
ϵ	$0,0$	$0,1$
J^1	$1,0$	$1,1$
B^1C^1	$2,0$	$0,0$
B^1J^1	$1,0$	$1,0$
C^1J^1	$1,0$	$1,1$
J^1B^1	$1,0$	$1,0$
J^1C^1	$1,0$	$1,1$

TABLE 7.4 – Jeu simplifié

Critère Minmax Les résultats du chapitre 5 indiquent qu'on peut obtenir une classification plus fine que celle du critère prudent Min et ainsi espérer une solution plus claire au jeu en utilisant par exemple le critère Minmax. On obtient alors le jeu qualitatif à la Table 7.5.

	ϵ	B^2	C^2	B^2C^2	C^2B^2
ϵ	$0,0$	$0,0$	$0,0$	$0,2$	$0,0$
B^1	$0,0$	$0,0$	$0,0$	$0,1$	$0,0$
C^1	$0,0$	$0,0$	$0,0$	$0,2$	$0,0$
J^1	$2,0$	$2,0$	$2,0$	$2,2$	$2,0$
B^1C^1	$3,0$	$1,0$	$3,0$	$1,1$	$1,0$
B^1J^1	$2,0$	$2,0$	$2,0$	$2,1$	$2,0$
C^1B^1	$0,0$	$0,0$	$0,0$	$0,1$	$0,0$
C^1J^1	$2,0$	$2,0$	$2,0$	$2,2$	$2,0$
J^1B^1	$2,0$	$2,0$	$2,0$	$2,1$	$2,0$
J^1C^1	$2,0$	$2,0$	$2,0$	$2,2$	$2,0$

TABLE 7.5 – Jeu qualitatif

Encore une fois, le jeu peut être simplifié en enlevant les plans nuls. Le résultat est le jeu représenté à la Table 7.6.

	ϵ	B^2C^2
ϵ	$0,0$	$0,3$
J^1	$2,0$	$2,2$
B^1C^1	$3,0$	$1,1$
B^1J^1	$2,0$	$2,1$
C^1J^1	$2,0$	$2,2$
J^1B^1	$2,0$	$2,1$
J^1C^1	$2,0$	$2,2$

TABLE 7.6 – Jeu simplifié

On peut observer que, les équilibres de Nash de ce jeu sont $(J^1, B^2.C^2)$,$(J^1.B^1, B^2.C^2)$, $(B^1.J^1, B^2.C^2)$, $(C^1.J^1, B^2.C^2)$ et $(J^1.C^1, B^2.C^2)$. On peut conclure que l'agent 2 utilisera le plan $B^2.C^2$. Pour l'agent 1, il peut utiliser n'importe quel plan contenant l'action J^1.

L'utilisation du critère Minmax, prenant en compte à la fois le pire et le meilleur résultat pour chaque profil, permet ainsi aux joueurs de trouver une solution naturelle au problème.

7.3.2 Coordination sous ignorance

Nous supposons que les joueurs peuvent dissiper l'incertitude associée à un profil, si tous sont d'accord pour le faire (en se coordonnant). Nous avons dit plus haut que l'incertitude sur le résultat d'un profil de stratégies peut être exprimé comme suit : quand le profil de stratégies est choisi, son "vrai" résultat est un unique résultat dans l'ensemble de résultats associés. Dissiper l'incertitude signifie que les agents se mettent d'accord pour contourner cette incertitude et décider ensemble de quel sera le "vrai" résultat. Comme les joueurs sont supposés rationnels, la coordination s'effectue seulement si elle bénéficie à tous. Plus exactement elle n'est pas irrationnelle quand elle bénéficie strictement à quelques joueurs, les autres joueurs lui étant indifférents. Il est aisé de voir le bénéfice qu'un agent peut retirer d'un processus de coordination : elle peut lui permettre d'éviter le pire résultat associé au profil de stratégie à propos duquel se fait la coordination. Ainsi, dans le cas d'une coordination, il semble inutile de considérer la totalité de la projection de l'ensemble des résultats pour le joueur. Seuls les résultats possibles d'une coordination (ce que nous appelons points de consensus) importent. Comme la décision sur la stratégie jouée prend place ex ante, c'est- à-dire avant que le jeu commence et qu'aucune coordination ne soit possible, ce que le joueur doit prendre en compte est la possibilité de coordination associée à un profil de stratégies. Une hypothèse doit être faite à propos de l'engagement des joueurs envers une possibilité de coordination. Est-ce qu'un joueur peut raisonnablement croire que quand un adversaire est intéressé par une coordination, cet adversaire l'acceptera une fois les stratégies choisies ? Nous supposons que la réponse est oui car la représentation du jeu est supposée porter en elle toute l'information pertinente pour la prise de décision des agents. En d'autres termes, si l'on voit en examinant les résultats d'un profil de stratégies qu'il est dans l'intérêt d'un joueur de se coordonner, alors aucune information cachée ne viendra changer cela. Mais comment déterminer l'intérêt d'un agent pour un point de coordination ? Il existe plusieurs critères à prendre en compte dont nous avons présenté l'intuition au chapitre 6. Nous allons dans la suite en expliciter les motivations.

• *L'incertitude de l'autre joueur* : supposons que, dans un jeu à deux joueurs, nous ayons les trois profils suivants

-a- (0,0) (1,1)

-b- (0,1) (1,1)

-c- (0,1) (1,0)

qui sont représentés graphiquement à la figure 7.3.

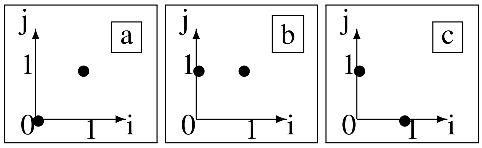


FIGURE 7.3 – Ensembles possibles de résultats

Pour ces trois profils la projection de l'agent 1 est la même $(\{0,1\})$, mais la projection de l'agent 2 varie. L'objectif est de déterminer si l'agent 1 doit évaluer différemment ces trois situations. Clairement le profil c est le moins intéressant pour l'agent 1 puisqu'il n'y a pas de possibilité de se coordonner pour ce profil. En effet, pour le profil c les deux agents sont clairement antagonistes (un agent obtient 0 quand l'autre obtient 1), et il n'y a donc pas d'espoir de se coordonner. Maintenant, il reste à déterminer s'il y a une différence entre les profils a et b pour l'agent 1. Pour les deux profils, les agents peuvent se coordonner pour atteindre le résultat $(1,1)$. Toutefois, quand dans le profil a les deux agents sont intéressés par cette coordination, pour le profil b l'agent 2 n'a pas de raison pour rejeter une coordination mais pas non plus de raison forte pour en accepter une. Alors l'agent 1 préférera le profil a au profil b puisque 2 aura pour a plus de motivation à se coordonner. Ainsi, afin de discriminer ces deux profils, l'agent 1 doit prendre en compte l'incertitude qu'a l'agent 2 sur ses résultats.

– *points de consensus*
  Nous avons besoin d'identifier les possibilités de coordination. C'est le but des points de consensus.

Définition 7.11 (Points de consensus) *Soit $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soit $s \in S$ un profil de stratégies. Un* point de consensus *pour s est un résultat $w \in U(s)$ t.q. $\exists J \subset N$, $\forall j \in J$, $\forall w' \in U(s)$, $w'_j \sim w_j$ et $\forall i \in N \setminus J$, $\exists w" \in U(s)$ t.q. $w"_i < w_i$. L'ensemble de tous les points de consensus associés à s est noté* $\mathsf{cp}(s)$.

Les points de consensus sont simplement des résultats tels que pour chaque agent, il y a au moins un résultat pire, ou alors tous les résultats du profil sont équivalents. Ainsi, quand un point de consensus existe, tous les agents ont un intérêt à se coordonner sur un tel point (il garantit d'éviter le pire résultat), ou au moins n'ont pas de raison de le refuser (tous les résultats sont équivalents). Les points de consensus sont fortement reliés aux points individuellement strictement rationnels considérés dans la théorie du marchandage (cf. par exemple [Thomson 1994]). Dans cette théorie, le critère de rationalité individuelle stricte **SIR** requiert qu'un agent ne considère que les points de l'ensemble de marchandage qu'il préfère strictement à ce qu'il peut attendre s'il n'y a pas d'arrangement. C'est une hypothèse minimale qu'on peut faire si l'on veut définir une solution rationnelle. **SIR** peut être traduite dans notre cadre par la demande suivante. Un agent est d'accord pour se coordonner sur un résultat w appartenant à un ensemble de résultats $U(s)$ associé à un profil de stratégies s seulement si sa récompense attendue w_i est strictement meilleure que le minimum de la projection de son ensemble de résultats, qui est la pire récompense qu'il peut obtenir sans coordination. La seconde partie de la définition précédente spécifie qu'un agent i peut se coordonner sur un résultat quand ce dernier domine strictement un autre élément de l'ensemble de résultats par rapport à sa relation de préférence $\leq_i$. On peut facilement voir que cela répond à la demande de **SIR**.

La première partie de la définition de points de consensus traite le cas où tous les résultats dans l'ensemble de résultats ont la même valeur pour un agent. Dans ce cas, l'agent n'a pas de raison de refuser une coordination puisqu'il n'y a pas d'autre résultat qu'il préfère (de manière usuelle, nous supposons que toute l'information sur les préférences de l'agent est codée par la fonction d'utilité, une approche analogue est utilisée pour la définition de la solution "dictatoriale" (**DS**) dans [Thomson 1994]).

La proposition suivante rend explicite la relation entre les points de consensus d'un profil de stratégies et les points Pareto optimaux :

Proposition 7.1 *Soit* $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ *un jeu qualitatif à résultats multiples. Soit* $s \in S$ *un profil de stratégies.*

$$\mathsf{pf}(\mathsf{cp}(s)) \subseteq \mathsf{pf}(s).$$

Preuve
Supposons que $\exists\omega \in \mathsf{pf}(\mathsf{cp}(s))$ et $\omega \notin \mathsf{pf}(s)$.
Alors $\exists\omega' \in \mathsf{pf}(s)$ qui pareto domine ω puisque $\omega \notin \mathsf{pf}(s)$ et alors par définition de $\mathsf{cp}(s)$, $\omega' \in \mathsf{cp}(s)$.
Mais $\omega \in \mathsf{pf}(\mathsf{cp}(s))$ induit $\nexists\omega'' \in \mathsf{cp}(s)$ qui pareto domine ω. Contradiction.

□

Nous avons aussi que :

Proposition 7.2 *Soit* $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ *un jeu qualitatif à résultats multiples. Soit* $s \in S$ *un profil de stratégies. Si* $\mathsf{cp}(s) \neq \emptyset$ *alors*

$$\mathsf{cp}(s) \cap \mathsf{pf}(s) \neq \emptyset.$$

Preuve
Supposons que $\mathsf{cp}(s) \neq \emptyset$. Alors $\mathsf{pf}(\mathsf{cp}(s)) \neq \emptyset$. Puisque nous avons $\mathsf{pf}(\mathsf{cp}(s)) \subseteq \mathsf{pf}(s)$, nous déduisons $\mathsf{pf}(s) \neq \emptyset$ et de là la proposition. □

Ainsi parmi les points de consensus il y a certains points Pareto optimaux, mais pas tout ceux-ci en général ; il n'est pas non plus vrai en général que tout point de consensus est Pareto optimal.

Comme expliqué plus haut, quand une coordination est possible, toute la projection d'un profil n'est pas intéressante. L'ensemble $\gamma(s)$ des résultats pertinents d'un profil de stratégies s est défini comme l'ensemble des points de consensus Pareto optimaux, quand cet ensemble n'est pas vide :

Définition 7.12 (Résultats pertinents) *Soit* $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ *un jeu qualitatif à résultats multiples. Soit* $s \in S$ *un profil de stratégies.*

$$\gamma(s) = \begin{cases} U(s) & \textit{si } \mathsf{cp}(s) = \emptyset \\ \mathsf{pf}(\mathsf{cp}(s)) & \textit{sinon} \end{cases}$$

La raison pour choisir seulement les points de consensus Pareto optimaux est claire : si les agents se mettent d'accord sur une coordination, alors ils vont choisir un résultat parmi les plus intéressants. On peut remarquer que nous disons que les agents ne vont pas simplement se mettre d'accord sur n'importe quel point Pareto optimal dans l'ensemble de résultats. La raison est qu'un résultat peut être Pareto optimal sans être un résultat possible d'une coordination. Sur la figure 7.4 les points Pareto optimaux

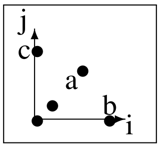


FIGURE 7.4 – POINTS PERTINENTS

sont a, b, c. Il est clair que le agent i n'acceptera jamais de se coordonner sur le point c puisque sa projection pour i est minimale. De même pour le agent j et le point b. Ainsi le seul point pertinent dans ce cas est le point a.

Supposons maintenant que nous avons un préordre $\leq_c$ pour comparer des ensembles de nombres (nous donnerons des exemples de tels préordres dans la section suivante). Nous pouvons l'utiliser pour évaluer chaque profil. Plus exactement, comme l'évaluation est qualitative, nous pouvons comparer chaque paire de profils, en construisant une relation $\preccurlyeq_i$, qui indique comment comparer des profils pour l'agent i.

$\leq_{c_i}$ permet de comparer les utilités de l'agent i pour les deux profils considérés. Afin de prendre en compte la possibilité de coordination et l'incertitude des autres agents sur les résultats, $\leq_{c_i}$ doit être raffinée de deux manières :

• La première revient à comparer différentes possibilités de coordination :

Définition 7.13 (Critère des points pertinents) *Soit $\leq_{c_i}$ une relation sur les ensembles de nombres. Alors la relation associée $\leq_{\gamma_c^i}$ est définie par : pour tous profils de stratégies s_1 et s_2,*

$$s_1 \leq_{\gamma_c^i} s_2 \;\; ssi \;\; \gamma(s_1) \leq_{c_i} \gamma(s_2).$$

• La seconde évalue le nombre d'autres agents qui n'ont pas d'incertitude sur le résultat pour un profil de stratégies donné :

Définition 7.14 (Critère de dépendence) *Soit $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soit i un agent, s un profil de stratégies et $E(s, i) = \{j \in N \setminus \{i\}$ t.q. $\forall w, w^{'} \in U(s),\ w \sim_j w^{'}$ et $\exists w, w^{'} \in U(s),\ w <_i w^{'}\}$. La relation $\leq_{d(i)}$ sur les profils de stratégies est définie par : soient s_1 et s_2 deux profils de stratégies,*

$$s_1 \leq_{d(i)} s_2 \; ssi \begin{cases} |E(s_1, i)| \geq |E(s_2, i)| \\ ou \\ \mathsf{cp}(s_1) = \emptyset \end{cases}$$

Ce raffinement est utile seulement quand il y a des points de consensus (cela explique la deuxième partie de la définition) ; il discrimine les cas où il y a un intérêt mutuel pour la coordination entre tous les agents, des cas où certains agents sont indifférents. Clairement, moins il y a d'agents indifférents, mieux c'est.

Sur cette base, nous sommes prêts à définir l'évaluation finale d'un agent pour les profils de stratégies :

Définition 7.15 (Evaluation d'un agent) *Soit $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soient $i \in N$ et $\leq_c$ un préordre sur les ensembles de nombres. L'evaluation par l'agent i des profils de stratégies de S est donnée par la relation $\preccurlyeq_i^c = lex(\leq_{\gamma_c^i}, \leq_{c_i}, \leq_{d(i)})$.*

L'utilisation de $lex(\leq_{\gamma_c^i},\leq_{c_i},\leq_{d(i)})$ nous permet de prendre en compte les trois dimensions de l'évaluation de manière adéquate. Nous commençons par comparer l'ensemble de nombres associé à chaque projection, mais en se restreignant aux points où une coordination est possible dans le profil. Pour les profils équivalents pour ce premier critère, nous examinons la totalité de la projection. Et quand deux profils sont encore équivalents, nous prenons en compte les cas où d'autres agents sont indifférents dans un profil, qui sont moins intéressants que les cas où tous les agents ont besoin de se coordonner.

Avec cette évaluation on peut maintenant, comme dans le cas sans coordination, utiliser des concepts de solution de la théorie classique des jeux, comme l'équilibre de Nash, pour résoudre ces jeux. Et encore une fois comme dans le cas sans coordination, quand une relation d'évaluation $\preccurlyeq_i^c$ est totale, on peut la représenter en utilisant des réels et alors juste résoudre le jeu classique sous forme stratégique correspondant. Bien sûr les stratégies mixtes ne peuvent pas être utilisées dans notre cadre puisque l'intensité des préférences ne peut être ici inférée.

Proposition 7.3
• *Si $\leq_{c_i}$ est un préordre sur les ensembles de nombres, alors $\preccurlyeq_i^c$ est un préordre sur les profils de stratégies.*
• *Si $\leq_{c_i}$ est total alors $\preccurlyeq_i^c$ est total.*

Preuve
Prouvons d'abord que si $\leq_c$ est un préordre total alors $\leq_{\gamma_c^i}$ est aussi un préordre total.
Soit $s \in S$,$s \leq_{\gamma_c^i} s$ ssi $\gamma(s) \leq_{c_i} \gamma(s)$ ce qui est vrai.
Soit $s_1, s_2, s_3 \in S$, $s_1 \leq_{\gamma_c^i} s_2$ et $s_2 \leq_{\gamma_c^i} s_3$ ssi $\gamma(s_1) \leq_{c_i} \gamma(s_2)$ et $\gamma(s_2) \leq_{c_i} \gamma(s_3)$. Cela induit $\gamma(s_1) \leq_{c_i} \gamma(s_3)$ et alors $s_1 \leq_{\gamma_c^i} s_3$.
Donc $\leq_{\gamma_c^i}$ est un préordre.
Finalement $\leq_{\gamma_c^i}$ est total. En effet $\forall s_1, s_2 \in S$, $\gamma(s_1) \leq_{c_i} \gamma(s_2)$ ou $\gamma(s_2) \leq_{c_i} \gamma(s_1)$ ce qui est équivalent à $s_1 \leq_{\gamma_c^i} s_2$ ou $s_2 \leq_{\gamma_c^i} s_1$.

Par ailleurs, $\leq_{d(i)}$ est un préordre total :
D'abord, pour $s \in S$, $s{\leq_{d(i)}}s$. En effet, $|E(s,i)| = |E(s,i)|$.
Ensuite, soit $s_1, s_2, s_3 \in S$. $s_1{\leq_{d(i)}}s_2$ et $s_2{\leq_{d(i)}}s_3$ induit $|E(s_1,i)| \leq |E(s_2,i)| \leq |E(s_3,i)|$ ou $\mathsf{cp}(s_1) = \mathsf{cp}(s_2) = \mathsf{cp}(s_3) = \emptyset$ ce qui signifie $s_1{\leq_{d(i)}}s_3$.
Finalement $\leq_{d(i)}$ est total.

Nous pouvons maintenant conclure en notant que $\preccurlyeq_i$ est la relation lexicographique associée à trois préordres totaux et est alors un préordre total. □

Résolution de l'exemple du dilemme des prisonniers Expliquons comment résoudre notre exemple 7.1 avec cette approche. Utilisons le critère min. Pour l'agent 1 nous avons les projections suivantes des profils de stratégies :
$U_1((C,C)) = \{3,6\}, U_1((C,D)) = \{0\}$,
$U_1((D,C)) = \{5\}$, $U_1((D,D)) = \{1,2\}$.
Nous avons les ensembles suivants de points pertinents pour les profils de stratégies :
$\gamma((C,C)) = \{(6,6)\}, \gamma((C,D)) = \{(0,5)\}$,
$\gamma((D,C)) = \{(5,0)\}, \gamma((D,D)) = \{(2,2)\}$.
La première comparaison des profils de stratégies (utilisant $\leq_{\gamma_c^1}$) donne ici directement un ordre strict sur les profils. Donc la relation lexicographique $\preccurlyeq_1$ est telle que

$$(C,D) \prec_1 (D,D) \prec_1 (D,C) \prec_1 (C,C).$$

Les deux agents ayant des positions symétriques, nous déduisons l'ordre suivant sur les profils de stratégies pour l'agent 2 :

$$(D,C) \prec_2 (D,D) \prec_2 (C,D) \prec_2 (C,C).$$

Il est facile de vérifier que le profil (C,C) est le seul équilibre de Nash Pareto optimal dans ce jeu. Alors la solution pour le jeu qualitatif à résultats multiples est que les deux agents choisissent la stratégie C et se coordonnent (sur une histoire cohérente), ce qui leur permettra de sortir libres.

Comme le critère Min donne ici un ordre strict sur les profils de stratégies, utiliser un autre critère qui le raffine (comme le Minmax et le produit) donnera la même solution.

Quand les relations $\preccurlyeq_i$ sont totales, il est possible de traduire le jeu qualitatif à résultats multiples en un jeu classique sous forme stratégique, simplement en assignant des nombres aux classes d'équivalence. Ainsi nous obtenons le jeu sous forme stratégique de la table 7.7 qui n'est plus maintenant un dilemme des prisonniers. Notre modèle nous a permis d'envisager la situation comme un jeu de coordination. On peut facilement vérifier que (C,C) est un équilibre de Nash.

	C	D
C	(4,4)	(1,3)
D	(3,1)	(2,2)

TABLE 7.7 – LE JEU SOUS FORME STRATÉGIQUE CORRESPONDANT

Résolution de l'exemple du pont Reprenons l'exemple développé à la section 6.2.4 et résolu dans le cas de préférences binaires. Nous montrons comment le cadre des jeux à résultats multiples permet de le résoudre dans le cas de préférences plus riches.

Représentons, dans la table 7.8 le jeu à résulats multiples associé à ce problème de planification multi-agent. Nous restreignons la longueur des plans examinés à deux actions (observons que les plans de longueur supérieure à 2 sont inutiles).

Ce jeu peut être transformé en un jeu qualitatif sous forme stratégique pour lequel on pourra appliquer les notions de théorie des jeux telles que les équilibres de Nash. Cette transformation se passe de la manière suivante. Pour chaque profil de plans on identifie d'abord le minimum et le maximum des projections des résultats pour le joueur ainsi que le minimum et le maximum des projections des points de consensus (quand ils existent). On parcourt ensuite l'ensemble des profils de plans tout en appliquant un algorithme de tri pour comparer les profils avec la relation d'évaluation $\preccurlyeq_i^c$ définie plus haut. Dans le jeu obtenu, le résultat associé à un profil de plan est simplement l'indice qui lui sera assigné dans le tableau trié obtenu.

Illustrons cette démarche sur deux exemples de profiles de stratégies en prenant le point de vue de l'agent 1. Soient les deux profils $S^1 = (B^1C^1, B^2)$et $S^2 = (B^1C^1, B^2C^2)$ dont les ensembles de résultats sont donnés à la figure 7.5.

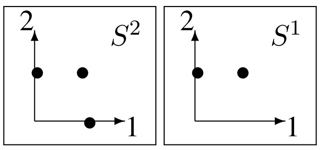


FIGURE 7.5 – Deux profils

La première étape est de considérer pour chaque profil la projection des points pertinents pour le joueur 1 (figure 7.6). L'axe horizontal représente les préférences ascendantes de l'agent 1. On observe

	ϵ	B^2	C^2	B^2C^2	C^2B^2
ϵ	(0,0)	(0,0)	(0,0)	(0,1)	(0,0)
B^1	(0,0)	(0,0)	(0,0)	(0,0) (0,1)	(0,0)
C^1	(0,0)	(0,0)	(0,0)	(0,1)	(0,0)
J^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1)	(0.5,0)
B^1C^1	(1,0)	(1,0) (0,0)	(1,0)	(1,1) (0,1) (1,0)	(1,0) (0,0)
B^1J^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1) (0.5,0)	(1,0)
C^1B^1	(0,0)	(0,0)	(0,0)	(0,1) (0.0)	(0,0)
C^1J^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1)	(0.5,0)
J^1B^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,0) (0.5,1)	(0.5,0)
J^1C^1	(0.5,0)	(0.5,0)	(0.5,0)	(0.5,1)	(0.5,0)

TABLE 7.8 – Jeu à résultats multiples

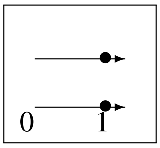


FIGURE 7.6 – POINTS PERTINENTS

que les deux profils sont équivalents pour le premier critère. La deuxième étape est alors de considérer la projection de l'ensemble des résultats pour le joueur 1 (figure 7.7).

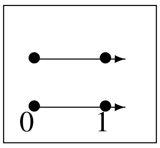


FIGURE 7.7 – RÉSULTATS

Les deux profils sont encore équivalents pour ce deuxième critère. Enfin vient la comparaison des profils selon le critère de dépendance. Ce critère montre que l'agent préfère le profil S^2 car pour le profil S^1 la projection de l'agent 2 est constante et il n'aura donc pas d'intérêt à se coordonner avec l'agent 1.

On peut observer que les agents ne peuvent obtenir une meilleure évaluation que celle obtenue pour $(B^1.C^1, B^2.C^2)$ en changeant unilatéralement de stratégie, ce qui signifie que ce profil est un équilibre de Nash. Une vérification systématique permet de constater que c'est le seul. Ainsi, il est plausible que l'agent 1 va exécuter le plan $B^1.C^1$ et l'agent 2 va exécuter le plan $B^2.C^2$. Les deux agents se coordonneront pour que la séquence d'action exécutée leur assure une satisfaction de $(1, 1)$, c'est-à-dire une des exécutions $B^1.C^1.B^2.C^2$ ou $B^2.C^2.B^1.C^1$ sera réalisée.

	ϵ	B^2	C^2	B^2C^2	C^2B^2
ϵ	0,0	0,0	0,0	0,3	0,0
B^1	0,0	0,0	0,0	0,1	0,0
C^1	0,0	0,0	0,0	0,3	0,0
J^1	1,0	1,0	1,0	1,3	1,0
B^1C^1	4,0	2,0	4,0	3,2	2,0
B^1J^1	1,0	1,0	1,0	1,1	1,0
C^1B^1	0,0	0,0	0,0	0,1	0,0
C^1J^1	1,0	1,0	1,0	1,1	1,0
J^1B^1	1,0	1,0	1,0	1,1	1,0
J^1C^1	1,0	1,0	1,0	1,3	1,0

TABLE 7.9 – jeu sous forme stratégique

7.4 Finesse de l'échelle d'évaluation

Dans cette section, nous nous concentrons sur le nombre de classes d'équivalence induites par $\preccurlyeq_i^c$ et montrons que cette relation est assez fine. Cette analyse est utile car même si un concept de solution est rationnel, si le préordre résultant divise l'ensemble de profils de stratégies en un très petit nombre de classes, le concept n'est d'aucune aide pour un agent qui souhaite l'utiliser pour prendre sa décision (il n'est pas assez discriminant).

Commençons par considérer la relation d'évaluation des profils possibles de stratégies quand les utilités des agents sont binaires, i.e., 0 (insatisfait), et 1 (satisfait). Pour de tels jeux, avec deux agents, il y a seulement 4 résultats possibles (0,0), (0,1), (1,0), (1,1), et 15 ensembles de résultats. Rappelons-en la représentation à la figure 7.8.

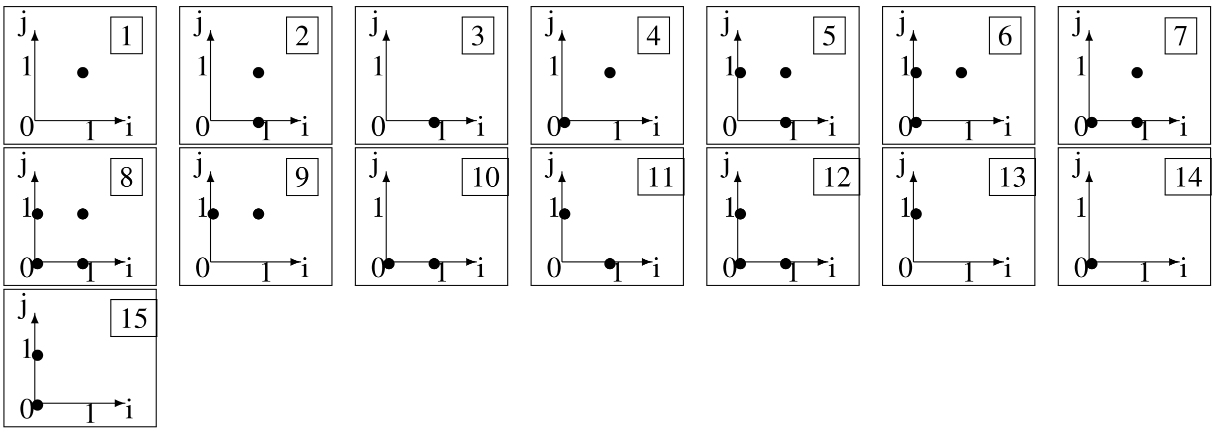


FIGURE 7.8 – ENSEMBLES DE RÉSULTATS POSSIBLES DANS LE CAS BINAIRE

Avec le critère Min, l'évaluation d'un agent ordonne les 15 ensembles de résultats dans 4 classes d'équivalence. Avec le critère Minmax, l'évaluation d'un agent ordonne les 15 ensembles de résultats dans 5 classes d'équivalence. La proposition suivante stipule de manière évidente que les critères Minmax et produit sont plus discriminants que le critère Min.

Proposition 7.4 *Soit $\Gamma = (N, (S_i)_{i\in[1,n]}, U)$ un jeu qualitatif à résultats multiples. Soient $i \in N$ et $s, s^{'} \in S$. Alors $s \preccurlyeq_i^{MinMax} s^{'} \Rightarrow s \preccurlyeq_i^{Min} s^{'}$ et $s \preccurlyeq_i^{product} s^{'} \Rightarrow s \preccurlyeq_i^{Min} s^{'}$.*

Preuve
Soit $i \in N$ et $s, s^{'} \in S$. On a $s \preccurlyeq_i^{Minmax} s^{'} \Rightarrow \min(U_i(s_1), \leq_i) \leq_i \min(U_i(s_2), \leq_i) \Rightarrow s \preccurlyeq_i^{Min} s^{'}$.
et $s \preccurlyeq_i^{product} s^{'} \Rightarrow \min(U_i(s_1), \leq_i) \leq_i \min(U_i(s_2), \leq_i) \Rightarrow s \preccurlyeq_i^{Min} s^{'}$. □

On peut facilement calculer le nombre de classes d'équivalence induite par l'évaluation d'un agent quand le critère Min est utilisé et que les échelles sont finies :

Proposition 7.5 *Si les échelles d'évaluation des agents ont pour cardinal m, alors en utilisant le critère Min, la relation d'équivalence induite par $\preccurlyeq_i^{Min}$ possède m^2 classes d'équivalence dans le cas de 2 agents. Et plus généralement, elle possède $\frac{n}{2} * m^2 + (1 - \frac{n}{2}) * m$ classes d'équivalence dans le cas de n agents.*

Preuve
Le premier critère différencie les profils $s, s^{'} \in S$ tels que $\min(U_i(\gamma(s_1)), \leq_i) \nsim_i \min(U_i(\gamma(s_2)), \leq_i)$. Cela génère m groupes (m possibilités pour $\min(U_i(\gamma(s)), \leq_i)$. Dans chaque groupe, le second critère différencie les profils $s, s^{'} \in S$ tels que $\min(U_i(s_1), \leq_i) \nsim_i \min(U_i(s_2), \leq_i)$. Soit G_l un tel groupe où $\min(U_i(\gamma(s_1)), \leq_i) \sim_i l$, alors puisque $\min(U_i(s_1), \leq_i) \leq_i \min(U_i(\gamma(s_1)), \leq_i)$ on obtient l sous-groupes. Dans ces sous-groupes, celui dans lequel $\min(U_i(\gamma(s_1)), \leq_i) \sim_i \min(U_i(s_1), \leq_i) \sim_i l$ n'offre de toute évidence pas d'opportunité de coordination. Pour le reste des sous-groupes, le dernier critère les sépare en n classes, une pour chaque nombre $\mid E(s, i) \mid$. Finalement, nous obtenons :

$$nbclasses(n, Min, m) = \sum_{l=1}^{m} (1 + n * (l - 1)).$$

En particulier, pour n=2,

$$nbclasses(n, Min, 2) = \sum_{l=1}^{m} (1 + 2 * (l - 1)) = m^2.$$

□

La proposition 7.4 montre qu'avec les critères Minmax et produit le nombre de classes d'équivalence ne peut être plus petit.

7.5 Liens avec la théorie des jeux classique

Une hypothèse essentielle à la résolution d'un jeu bayésien / prébayésien est qu'un joueur connaît son type au moment d'entrer dans le jeu, i.e., il est en possession de son information privée. Cette hypothèse peut sembler naturelle, pourtant elle restreint implicitement l'incertitude considérée dans le jeu à une incertitude sur le comportement de l'autre agent (qui est décrit par le type de celui-ci). Le cadre que nous proposons ne fait pas une telle hypothèse. Il permet encore de représenter l'incertitude sur les comportements des autres joueurs, mais en plus il permet de modéliser d'autres sources d'incertitude provenant de la présence des autres agents. Nous sommes particulièrement intéressés par la modélisation de l'incertitude due à l' intercalation incontrôlable des actions des différents agents. Nous pensons que cette hypothèse est cohérente avec la volonté de modéliser des agents qui sont aussi autonomes et, pour reprendre une terminologie utilisée dans le cadre des systèmes distribués, "aussi peu connectés que possible". L'hypothèse d'un joueur qui ne connaît pas son type n'est pas un concept complètement exotique en théorie des jeux. En effet, les solutions obtenues peuvent être considérées comme des solutions ex ante classiques. Les jeux prébayésiens situent leur raisonnement pendant la phase intermédiaire durant laquelle le joueur connaît seulement son propre type. La phase ex post vient quand les types de tous les joueurs sont révélés. Dans ce sens-là, notre cadre suppose moins d'information qu'un jeu prébayésien car il se situe temporellement dans une phase plus en amont de l'acquisition de l'information, et ce, même si dans les deux cas, les probabilités ne sont pas disponibles. Les jeux bayésiens définissent un profil d'action/stratégie comme une application qui associe un état du monde à un résultat. Mais une telle application fournit déjà de l'information. Pour qu'un joueur puisse raisonner stratégiquement en utilisant cet outil, il doit connaître le résultat de chaque action dans chaque état du monde. Ceci n'est pas toujours possible. Une structure plus naturelle de l'information est de considérer qu'un agent associe simplement un profil de stratégies à un ensemble de résultats possibles. Comme nous l'avons déjà énoncé, un résultat associé à un profil de stratégies donné peut être considéré comme une description partielle du monde dans lequel le profil s'est réalisé. La description de ce monde devient complète en spécifiant un résultat pour chaque profil de stratégies, c'est-à-dire une conséquence dans chaque ensemble de conséquences. Ainsi, un jeu à résultats multiples peut être considéré comme un jeu prébayésien où chaque profil de types est associé à un (ou plusieurs) élément(s) du produit cartésien des ensembles de conséquences des profils de stratégies. L'inverse n'est pas toujours vrai. Certaines actions qui peuvent facilement être décrites dans le cadre bayésien ne font pas sens dans le nôtre. Il pourrait sembler à première vue que notre modèle nous empêche de représenter des actions. Nous pensons qu'alternativement, ce point révèle la correspondance état-conséquence dans le cadre bayésien comme déjà une information. Supposons en effet qu'une action a ait comme ensemble de conséquences c, c' et qu'une action a' conduise aux conséquences $c_1, c_1^{'}$ et que cette situation soit représentée par le jeu prébayésien à la table 7.10. On peut clairement voir que ce jeu porte l'information supplémentaire que le cas où a donne c et a' donne $c_1^{'}$ est impossible, ce qui nous éloigne quelque peu de l'ignorance complète désirée.

	a'	b'		a'	b'
a	(c, c_1)	(0,5)	a	$(c', c_1^{'})$	(0,5)
b	(5,0)	(1,1)	b	(5,0)	(1,1)

TABLE 7.10 – un jeu prébayésien contient déjà de l'information

Et voici quasiment le même argument exprimé de manière probabiliste. Supposons que dans le même état du monde l'action a donne la conséquence c et l'action a' donne la conséquence c', alors le joueur peut supposer que $p(c/a) = p(c'/a')$. Ce n'est généralement pas le cas pour nous. Afin d'illustrer la différence dans les situations qu'il convient de modéliser avec les jeux prébayésiens, revenons au dilemme de l'impala. Les deux matrices symbolisaient deux types de lions : rapide ou lent. La justification en était que le succès de la fuite de l'impala dépend de la caractéristique du lion que l'impala ignore : sa vitesse. Si nous tentons de mieux prendre en compte la réalité, nous notons qu'il y a beaucoup d'autres facteurs qui peuvent influencer le résultat de la course : le vent, le terrain, la densité végétale... Au lieu de diviser les mondes possibles en deux (de manière aléatoire) comme le ferait un impala (pré) bayésien, notre impala devrait simplement associer le profil (fuir, attaquer) à ses résultats possibles sans plus de conjectures. Le jeux à résultat multiples qui s'en suit est décrit à la table 7.11.

Dans la suite nous donnons quelques résultats qui illustrent la relation entre jeux qualitatifs à résultats multiples et jeux usuels sous forme stratégique. Afin de pouvoir effectuer la comparaison, nous supposons ici que les utilités sont des nombres réels. La première proposition montre que les jeux quali-

	attaquer	ne pas attaquer
fuir	(1,0) (0,1)	(1,0)
se cacher	(0,1)	(1,0)

TABLE 7.11 – Dilemme de l'impala à résultats multiples.

tatifs à résultats multiples sont une généralisation (qualitative) des jeux sous forme stratégique.

Proposition 7.6 *Quand dans un jeu qualitatif à résultats multiples Γ, il y a seulement un résultat associé à chaque profil de stratégies, les équilibres de Nash (en stratégies pures) de Γ vu comme un jeu qualitatif à résultats multiples sont exactement les équilibres de Nash (en stratégies pures) de Γ vu comme un jeu classique.*

Preuve
Soit $\Gamma = (N, (S_i)_{i \in [1,n]}, U)$ un jeu qualitatif à résultats multiples. Quand il existe seulement un résultat par profil de stratégies, la fonction d'utilité U se réduit à une fonction classique $\mu : S \longrightarrow \chi_1 \times \cdots \times \chi_n$.
La projection des résultats d'un profil de stratégie $s \in S$ se réduit à un point ;
$\forall i \in N,\ min(U_i(s), \leq_i) = max(U_i(s), \leq_i) = \mu_i(s)$. Alors le critère utilisé $c \in$ Min, Minmax, produit se trivialise vers

$$s1 \preccurlyeq_i^c s2 \ ssi \ \mu_i(s1) \leq_i \mu_i(s2).$$

De plus, le critère de dépendance ne peut jamais être satisfait. En effet, sa définition requiert l'existence d'au moins deux résultats associés à chaque profil de stratégies. □

Dans les jeux qualitatifs à résultats multiples, l'incertitude sur le résultat exact peut être modélisée en ajoutant un joueur supplémentaire, par exemple la Nature, qui choisit un résultat pour chaque profil avec une certaine probabilité, inconnue des autres joueurs, quand ils ne se coordonnent pas. Appelons cette représentation du problème le jeu étendu associé et explorons quelques résultats à propos de cette représentation.

Proposition 7.7 *Soit Γ un jeu qualitatif à résultats multiples. Si un profil de stratégies est un équilibre de Nash du jeu étendu associé pour n'importe quelle distribution de probabilités sur les résultats alors ce profil de stratégies est un équilibre de Nash de Γ.*

Preuve
Soit $\Gamma = (N, (S_i)_{i \in [1,n]}, U)$ un jeu qualitatif à résultats multiples.
Soit $NE(\Gamma)$ un ensemble d'équilibres de Nash Γ.
Soit n_{tot} le nombre total de profils de stratégies de Γ.
Soit CRIT = $\{Min, Minmax, produit\}$.
Soit $i \in N$.
Pour $k \in [1, n_{tot}]$, soit $p_k \in \Delta(U(s_k))$ une distribution de probabilités sur l'ensemble des résultats.
Soit p un vecteur de telles distributions de probabilités.
Alors $p \in \Delta = \Delta(U(s_1) \times \cdots \times U(s_{n_{tot}}))$

Soit $\Gamma(p)$ le jeu étendu associé à Γ obtenu à partir de p.
Soit s_{eq} un profil de stratégie tel que pour chaque $p \in \Delta$, $s_{eq} \in NE(\Gamma(p))$.
Alors pour $\omega \in U(s_{eq})$, pour chaque $s^{'} \in S_i$, pour un certain $k \in [1, ntot]$ tel que $U(s_{eq_{-i}}, s^{'}) = U(s_k)$ et, $\omega^{'} \in U(s_k)$, nous avons $s \in NE(\Gamma(p))$ où p est tel que $p_{eq}(\omega) = 1$ et $p_k(\omega^{'}) = 1$.
Alors nous avons $\omega_i^{'} \leq_i \omega_i$.
Cela étant vrai pour chaque tel couple $\omega, \omega^{'}$ nous pouvons conclure :

$\forall s^{'} \in S_i, max(U_i((s_{-i}, s^{'})), \leq_i) \leq_i min(U_i(s), \leq_i)$ ce qui induit que $\forall c \in CRIT, (s_{-i}, s^{'}) \preccurlyeq_i^c s$ et donc $s \in NE(\Gamma)$. □

Ce résultat signifie que si un profil de stratégies est objectivement (i.e., quel que soit le résultat réalisé) un équilibre de Nash, alors il est reconnu comme tel par le modèle des jeux qualitatifs à résultats multiples.

Il existe aussi un résultat dual, qui met en lumière une certaine rationalité des équilibres de Nash des jeux qualitatifs à résultats multiples.

Proposition 7.8 *Si un profil de stratégies est un équilibre de Nash d'un jeu qualitatif à résultats multiples, alors il existe une distribution de probabilités sur les résultats telle que ce profil est un équilibre de Nash du jeu étendu associé correspondant.*

Preuve

Soit $\Gamma = (N, (S_i)_{i \in [1,n]}, U)$ un jeu qualitatif à résultats multiples. Soit $NE(\Gamma)$ un ensemble d'équilibres de Nash de Γ.
Soit n_{tot} le nombre total des profils de stratégies de Γ.
Soit CRIT = $\{Min, Minmax, produit\}$.

Pour $k \in [1, n_{tot}]$, soit $p_k \in \Delta(U(s_k))$ une distribution de probabilités sur l'ensemble de résultats. Soit p un vecteur de telles distributions de probabilités.
Alors $p \in \Delta = \Delta(U(s_1) \times \cdots \times U(s_{n_{tot}}))$
Soit $\Gamma(p)$ le jeu étendu associé à Γ obtenu à partir de p.
Soit s_{eq} un profil de stratégies tel que pour chaque $p \in \Delta$, $s_{eq} \in NE(\Gamma(p))$.

$s_{eq} \in NE(\Gamma) \Leftrightarrow \forall s^{'} \in S_i, (s_{-i}, s^{'}) \leq_{\gamma_c^i} s$.
Cela induit
$\min(U_i(\gamma(s_{-i}, s^{'})), \leq_i) \leq_i \min(U_i(\gamma(s)), \leq_i)$.
Soit $k \in [1, ntot]$ tel que $U(s_{eq_{-i}}, s^{'}) = U(s_k)$ et $\omega^k \in U(s_k)$ tel que $\omega_i^k \sim_i \min(U_i(\gamma(k)))$.
Soit $p \in \Delta$ une distribution de probabilités telle que $p_k(\omega^k) = 1$. Supposons de plus que $p_{eq}(\omega) = 0$ pour $\omega \in U(s_{eq}) \setminus U(\gamma(s_{eq}))$. Alors nous avons $\omega_i^k \leq_i \sum_{\omega \in s_{eq}} (p_{eq}(w) * w_i)$.
En effet $\forall w \in U(s_{eq}), \omega_i^k \sim_i \min(U_i(\gamma(s_{-i}, s^{'})) \leq_i \min(U_i(\gamma(s)), \leq_i) \leq_i \omega_i$.
Ceci étant vrai pour chaque k et pour chaque $s^{'} \in S_i$, nous avons par définition $s \in NE(\Gamma(p))$. □

Donc si un profil de stratégies n'est un équilibre de Nash pour aucun jeu étendu associé alors ce n'est pas un équilibre de Nash du jeu qualitatif à résultats multiples.

7.6 Conclusion

Dans ce chapitre, nous avons introduit et étudié les jeux qualitatifs à résultats multiples. Dans ce type de jeux, chaque profil de stratégies des joueurs est associé non pas à un résultat unique (un vecteur de satisfactions) comme dans les jeux classiques, mais à un ensemble de tels résultats. Les joueurs ne possèdent comme information que cet ensemble de résultats. En particulier, ils ne possèdent pas de distribution de probabilités sur les résultats formant l'ensemble. Ces jeux sont particulièrement adaptés à des situations extrêmement pauvres en information dont nous avons montré qu'elles ne pouvaient être capturées de manière satisfaisante par les modèles de théorie des jeux sous incertitude que sont par exemple les jeux bayésiens ou les jeux prébayésiens. Nous avons montré que ces jeux pouvaient être résolus au moyen des concepts classiques de théorie des jeux, notamment l'équilibre de Nash. Utiliser ces concepts requiert de se baser sur la théorie de la décision sous ignorance pour expliciter les préférences des agents sur les ensembles de résultats à partir de leurs préférences sur les résultats.

Conclusion

Quand un agent n'est pas sûr du résultat de ses actions, il doit trouver ailleurs que dans la comparaison des conséquences de ses actions avec ses buts l'évaluation de ses actions. Il doit définir ses chances d'atteindre son but. Une fois ses chances évaluées, il choisira l'action qui lui donne la meilleure évaluation. Son choix repose donc sur la manière d'évaluer sa situation. Or c'est exactement le rôle du modèle de permettre à l'agent de se représenter en termes utiles l'état du monde qui l'entoure et les perspectives qui s'offrent à lui, d'où son importance.

Dans cette thèse, nous nous sommes intéressés au cadre le plus pauvre en information dans lequel l'agent ne possède comme information que celle, minimale, consistant en l'ensemble des conséquences d'un choix. Nous avons exhibé un ensemble d'axiomes qui utilisent un minimum de conditions pour déterminer les caractéristiques d'un comportement rationnel. Cette caractérisation est réalisée au moyen de théorèmes qui stipulent que la comparaison entre deux actions se base sur une relation de comparaison entre les couples formés par les éléments extrémaux (au regard de la préférence de l'agent) de leurs ensembles respectifs de conséquences. La famille de critères délimitée comporte notamment le critère prudent Min qui compare seulement pour deux actions leurs conséquences les moins préférées pour l'agent. Elle ne se limite toutefois pas à ce critère et comporte notamment le critère Max qui compare pour deux actions leurs conséquences préférées pour l'agent et ouvre alors la possibilité d'exprimer une disposition plus optimiste de l'agent. Les autres éléments de la famille de critères peuvent être vus comme des manières différentes de combiner ces deux critères.

La planification est un cadre intéressant d'application de la décision dans l'incertitude. Le modèle y acquiert une importance plus criante encore car les contraintes de temps et de calcul deviennent très fortes dans des conditions d'application réalistes. De plus, la planification présente un défi supplémentaire pour l'agent par rapport à un cadre de décision ponctuelle. L'agent doit, en effet, y choisir une suite d'actions qui introduisent autant de points d'incertitude. Une incertitude particulièrement intéressante est celle générée par la présence d'autres agents planificateurs. Ce cas est très courant dans des applications réelles. Il est intéressant car il permet de nouvelles possibilités sur le raisonnement à propos des actions. L'agent peut en effet utiliser la rationalité supposée des autres agents présents pour prévoir leurs actions futures et même profiter des possibilités de coordination possibles. On a d'abord étudié le cas de préférences dichotomiques dans lequel on peut circonscrire toutes les configurations possibles. Nous avons introduit un modèle de planification qui étend celui de la planification classique au cadre multi-agent. Une notion primordiale dans ce nouveau modèle est celle de mélanges de plans. Un mélange de plans est une manière possible d'intercaler les actions des deux plans. Étant donné que nous supposons, comme en planification classique, que les actions des agents sont déterministes, un mélange induit de manière sûre une conséquence unique. L'incertitude des agents, une fois fixés leurs plans respectifs, porte alors sur le mélange qui se réalisera finalement. On propose pour traiter cette incertitude en utilisant les profils de satisfaction. Un profil de satisfaction est une représentation graphique, pour un ensemble de plans des agents, des satisfactions que tirent les agents des différents mélanges possibles. Cette représentation permet alors d'utiliser les résultats montrés dans la partie de la thèse concernant la décision sous incertitude pour évaluer, pour chaque agent, la situation dans laquelle il se trouve quand l'ensemble de plans spécifié est exécuté. Un agent peut en particulier évaluer la possibilité de coordination offerte sur un mélange et visible dans le profil de satisfaction. Un mélange peut faire l'objet d'une coordination entre les agents s'il permet par exemple à l'agent d'éviter une issue moins préférée. Cela se traduit alors par le fait que la coordination ne peut pas se faire autour de mélanges strictement dominés au sens de Pareto. Cela s'exprime de manière aisée sur un profil de satisfaction. On a ensuite esquissé l'extension du modèle au cas de préférences plus riches par le rajout d'un échelon dans l'échelle de préférences. Une structure de préférences plus riche induit des configurations plus nombreuses et appelle une évaluation plus fine des possibilités de coordination. On montre alors que la procédure introduite dans le cas de préférences dichotomiques peut être étendue à cette nouvelle situation.

Afin de baser l'étude sur une démarche axiomatique, nous nous sommes intéressé au cadre de la théorie des jeux qui offre le meilleur terrain d'expression des interactions entre agents rationnels. Nous avons alors introduit un nouveau type de jeux, les jeux qualitatifs à résultats multiples. Ce sont des jeux dans lesquels chacun des joueurs associe à un profil de stratégies donné non pas un résultat unique mais un ensemble de tels résultats. Nous avons montré que la situation que ces jeux décrivent ne peut être capturée de manière satisfaisante par les modèles existants en théorie des jeux. Nous avons aussi montré comment résoudre de tels jeux en nous appuyant les résultats de décision sous incertitude pour comparer les ensembles des résultats des profils de stratégies puis en utilisant les concepts éprouvés de la théorie des jeux classique tels que l'équilibre de Nash. Notre modèle de jeux permet de prendre en compte les possibilités d'interaction dans un cadre très pauvre en information dans lequel l'évaluation des ensembles de résultats ne peut être ramenée au cadre classique en utilisant par exemple le concept d'utilité espérée. L'une des caractéristiques les plus intéressantes des résultats et modèles que nous avons exhibés est qu'ils s'appliquent aussi bien à des préférences numériques qu'ordinales, ce dernier cas est particulièrement cohérent avec une information rare et essentiellement qualitative.

Dans la construction du modèle de décision multi-agent, nous nous sommes appuyés sur des concepts de théorie des jeux qui ne prennent pas en compte, quand le nombre d'agents dépasse deux, les possibilités de coalitions, temporaires ou définitives, entre des sous-groupes d'agents. Il sera sans doute intéressant de caractériser les critères de formation de coalitions dans le cadre de l'ignorance complète. Du point de vue de l'application du modèle à des situations réelles, le prolongement naturel de notre travail est celui du choix d'un langage de représentation adapté pour exploiter les concepts du modèle. Bien sûr, le langage dépendra avant tout du domaine dont on veut représenter une situation en utilisant le modèle. Nous pensons que l'élaboration du modèle est une étape importante. Les algorithmes qui seront par la suite élaborés dans le langage choisi tireront du modèle leur pertinence et leur efficacité.

Bibliographie

[Aghassi & Bertsimas 2006] M. AGHASSI et D. BERTSIMAS. « Robust game theory ». *Mathematical Programming*, 107(1) :231273, 2006.

[Alexander 1975] E. R. ALEXANDER. « The limits of uncertainty : A note ». *Theory and Decision*, 6 :363–370, 1975.

[Allais 1953] M. ALLAIS. « Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école Américaine ». *Econometrica*, 21 :503–546, 1953.

[Anant *et al.* 1990] T. C. A. ANANT, B. MUKHERJI, et K. BASU. « BargainingWithout Convexity, Generalizing the Kalai-Smorodinsky Solution ». *EconomicsLetters*, 33 :115–119, 1990.

[Arrow & Hurwicz 1977] K. J. ARROW et L. HURWICZ. « An optimality criterion for decision-making under ignorance ». Dans *Studies in Resource Allocation Processes*, pages 463–471. Cambridge UniversityPress, 1977.

[Arrow 1959] K. J. ARROW. « Rational choice functions and orderings ». *Economica N.S*, 26 :121– 127, 1959.

[Ashlagi *et al.* 2006] I. ASHLAGI, D. MONDERER, et M. TENNENHOLTZ. « Resource selection games with unknown number of players ». Dans *AAMAS*, pages 819–825, 2006.

[Ashlagi *et al.* 2007] I. ASHLAGI, D. MONDERER, et M. TENNENHOLTZ. « Routing games with an unknown set of active players ». Dans *AAMAS*, page 195, 2007.

[Aumann 1962] R. J. AUMANN. « Utility Theory without the Completeness Axiom ». *Econometrica, Vol. 30, No. 3*, pages 445–462, 1962.

[Aumann 1974] R. J. AUMANN. « Subjectivity and correlation in randomized strategies ». *Journal of Mathematical Economics*, 1(1) :67–96, 1974.

[Aumann 1987] R. J. AUMANN. « Correlated Equilibrium as an Expression of Bayesian Rationality ». *Econometrica*, 55(1) :1–18, 1987.

[Bacchus & Grove 1996] F. BACCHUS et A. GROVE. « Utility independence in a qualitative decision theory ». Dans *In Proceedings of KR'96*, pages 542–552. Morgan Kaufmann, 1996.

[Bai & Zhang 2008] Q. BAI et M. ZHANG. « A flexible and reasonable mechanism for self-interested agent team formingSource ». *Multiagent and Grid Systems*, 4 :85–101, 2008.

[Barbera & Pattanaik 1984] S. BARBERA et P.K. PATTANAIK. « Extending an order on a set to the powerset : some remarks on Kannai and Peleg's approach ». *Journal of EconomicTheory*, 32 :185–191, 1984.

[Barbera *et al.* 1984] S. BARBERA, C. R. BARRETT, et P. K. PATTANAIK. « On some axioms for ranking sets of alternatives ». *Journal of Economic Theory*, 33(2) :301–308, 1984.

[Barbera *et al.* 2001] S. BARBERA, W. BOSSERT, et P. K. PATTANIK. « Ranking Sets of Objects ». Rapport Technique, Universite de Montreal, Departement de sciences economiques, 2001. available at http ://ideas.repec.org/p/mtl/montde/2001-02.html.

[Barbera *et al.* 2004] S. BARBERA, W. BOSSERT, et P.K. PATTANAIK. « Ranking Sets of Objects ». Rapport Technique, 2004.

[Bernoulli 1954] D. BERNOULLI. « Specimen theorie novae de mensura sortis (1738) ». *Econometrica*, 22 :23–36, 1954.

[Bhattacharyya 2008] A. BHATTACHARYYA. « Median-based Rules for Decision-making under CompleteIgnorance ». *Soumis à Economic Theory*, 2008.

[Binmore 1999] K. BINMORE. *Jeux et théorie des jeux*. DeBoeck Université, 1999.

[Blythe 1999] J. BLYTHE. « Decision-theoretic planning ». *AI Magazine*, 1999.

[Borodin & El-Yaniv 1998] A. BORODIN et R. EL-YANIV. *Online Computation andCompetitive Analysis*. Cambridge University Press, Cambridge, 1998.

[Bossert 1989] W. BOSSERT. « On the extension of preferences over a set to the power set : An axiomatic characterization of a quasi-ordering ». *Journal of Economic Theory*, 49(1) :84–92, 1989.

[Bossert 1997] W. BOSSERT. « Uncertainty aversion in nonprobabilistic decision models ». *Mathematical Social Sciences*, 34(3) :191–203, 1997.

[Bossert 2001] W. BOSSERT. « Choices, consequences, and rationality ». *Synthese*, 129 no3 :343–369, 2001.

[Bossert *et al.* 1994] W. BOSSERT, P. K. PATTANAIK, et Y. XU. « Ranking Opportunity Sets : An Axiomatic Approach ». *Journal of Economic Theory*, 63(2) :326–345, 1994.

[Bossert *et al.* 2000] W. BOSSERT, P. K. PATTANAIK, et Y. XU. « Choice under complete uncertainty : axiomatic characterizations of some decision rules ». *Journal of Economic Theory*, pages 295–312, 2000.

[Boutilier 1994] C. BOUTILIER. « Toward a Logic for Qualitative Decision Theory ». Dans *Proceedings of the International Conference on the Principles of Knowledge Representation and Reasoning (KR'94)*, pages 75–86, 1994.

[Boutilier *et al.* 1999] C. BOUTILIER, T. DEAN, et S. HANKS. « Decision-Theoretic Planning : Structural Assumptions and Computational Leverage ». *JAIR*, 11 :1–94, 1999.

[Bouyssou *et al.* 2006] D. BOUYSSOU, D. DUBOIS, M. PIRLOT, et H. PRADE. *Concepts et méthodes pour l'aide à la décision - risque et incertain*, volume 2 de *Traité IC2*. Lavoisier, http ://www.editions-hermes.fr/, 2006.

[Bowling & Veloso 2003] M. BOWLING et M. VELOSO. « Simultaneous Adversarial Multi-Robot Learning ». Dans *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 699–704, 2003.

[Brafman & Tennenholtz 1996] R. I. BRAFMAN et M. TENNENHOLTZ. « On the Foundations of Qualitative Decision Theory ». pages 1291–1296. MIT Press, 1996.

[Chernoff 1954] H. CHERNOFF. « Rational selection of decision functions ». *Econometrica*, 22 :422443, 1954.

[Cimatti & Roveri 1999] A. CIMATTI et M. ROVERI. « Conformant Planning via Model Checking ». Dans *ECP*, pages 21–34, 1999.

[Cimatti *et al.* 2004] A. CIMATTI, M. ROVERI, et P. BERTOLI. « Conformant planning via symbolic model checking and heuristic search ». *Artificial Intelligence*, 159(1-2) :127–206, 2004.

[Claus & Boutilier 1997] C. CLAUS et C. BOUTILIER. « The dynamics of reinforcement learning in cooperative multiagentsystems ». Dans *Workshop on Multi-Agent Learning*, page 602608, 1997.

[Clement & Barrett 2003] B. J. CLEMENT et A. C. BARRETT. « Continual coordination through shared activities ». Dans *AAMAS*, pages 57–64, 2003.

[Cohen & Levesque 1991] P. R. COHEN et H. J. LEVESQUE. « Confirmations and Joint Action ». Dans *IJCAI*, pages 951–959, 1991.

[Conley & Wilkie 1991] J. P.. CONLEY et S. WILKIE. « The bargaining problem without convexity : Extending the egalitarian and Kalai-Smorodinsky solutions ». *Economics Letters*, 36(4) :365–369, 1991.

[de Finetti 1937] B. de FINETTI. La Prévision : Ses Lois Logiques, ses Sources Subjectives. Dans *Annales de l'Institut Henri Poincaré 7*, pages 1– 68. 1937.

[de Weerdt *et al.* 2003] M. de WEERDT, A. BOS, H., et C. WITTEVEEN. « A Resource Logic for Multi-Agent Plan Merging ». *Ann. Math. Artif. Intell.*, 37(1-2) :93–130, 2003.

[Dean & Kanazawa 1989] T. DEAN et K. KANAZAWA. « Persistence and probabilistic inference ». *IEEE Transactions on Systems, Man, and Cybernetics*, 19(3) :574–585, 1989.

[Decker & Lesser 1992] K. S. DECKER et V. R. LESSER. « Generalizing the partial global planning algorithm ». *International Journal of Intelligent and Cooperative Information Systems*, 1 :319–346, 1992.

[Decker *et al.* 2000] K. DECKER, J. LI, et Y. DEMAZEAU. « Coordinating mutually exclusive resources using gpgp ». *Autonomous Agents and Multi-Agent Systems*, 3 :200, 2000.

[Doyle & Thomason 1999] J. DOYLE et R. H. THOMASON. « Background to qualitative decision theory ». *AI magazine*, 20 :55–68, 1999.

[Drummond 1985] Mark DRUMMOND. « Refining and Extending the Procedural Net ». Dans *IJCAI*, pages 1010–1012, 1985.

[Dubois & Prade 1988] D. DUBOIS et H. PRADE. *Possibility theory*. Plenum Press, 1988.

[Dubois *et al.* 2000] D. DUBOIS, L. GODO, H. PRADE, et A. ZAPICO. « Advances in Qualitative Decision Theory : Refined Rankings ». Dans *IBERAMIA-SBIA*, pages 427–436, 2000.

[Dubois *et al.* 2002] D. DUBOIS, H. FARGIER, H. PRADE, et P. PERNY. « Qualitative Decision Theory : From Savage's Axioms to Nonmonotonic Reasoning ». *Journal of the ACM*, 49(4) :455–495, 2002.

[Dubois *et al.* 2003] D. DUBOIS, H. FARGIER, H. PRADE, et P. PERNY. « Qualitative decision theory with preference relations and comparative uncertainty : An axiomatic approach ». *Artificial Intelligence*, 148 :219–260, 2003.

[Dunne & van der Hoek 2004] P.E. DUNNE et W. van der HOEK. « Representation and complexity in boolean games ». Dans *JELIA*, page 347359, 2004.

[Durfee & Lesser 1991] E. H. DURFEE et V. R. LESSER. « Partial global planning : A coordination framework for distributed hypothesis formation ». *IEEE Transactions on Systems, Man, and Cybernetics*, 21 :1167–1183, 1991.

[Eliaz 2006] K. ELIAZ. « Indifference or Indecisiveness ? The Choice-Theoretic Foundations of Incomplete Preferences ». *Games and Economic Behavior*, 56(1) :61–86, 2006.

[Ellsberg 1961] D. ELLSBERG. « Risk, Ambiguity, and the Savage Axioms ». *The Quarterly Journal of Economics*, 75, No. 4 :643–669, 1961.

[Ephrati & Rosenschein 1993] E. EPHRATI et J. S. ROSENSCHEIN. « Distributed Consensus Mechanisms for Self-Interested Heterogeneous Agents ». Dans *CoopIS*, pages 71–79, 1993.

[Ephrati *et al.* 1995] E. EPHRATI, M. E. POLLACK, et J. S. ROSENSCHEIN. « A Tractable Heuristic that Maximizes Global Utility through Local Plan Combination ». Dans *ICMAS*, pages 94–101, 1995.

[Fargier & Sabbadin 2003a] H. FARGIER et R. SABBADIN. « Qualitative decision under uncertainty : Back to expected utility ». *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI'2003)*, pages 303–308, 2003.

[Fargier & Sabbadin 2003b] H. FARGIER et R. SABBADIN. « Qualitative Decision under Uncertainty : Back to Expected Utility ». Dans *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'03)*, pages 303–308, 2003.

[Fargier & Sabbadin 2005] H. FARGIER et R. SABBADIN. « Qualitative decision under uncertainty : Back to expected utility ». *Artificial Intelligence*, 164 :245–280, 2005.

[Feldman & Sproull 1977] J.A. FELDMAN et R.F. SPROULL. « Decision Theory and Artificial Intelligence II : The Hungry Monkey ». *Cognitive Science*, 2 :158–192, 1977.

[Fikes & Nilsson 1971] R. FIKES et N. J. NILSSON. « STRIPS : A New Approach to the Application of Theorem Proving to Problem Solving ». *Artificial Intelligence*, 2(3/4) :189–208, 1971.

[Fishburn 1985] P.C. FISHBURN. *Interval Orders and Interval Graphs*. John Wiley, New York, 1985.

[Florea 1999] A. M. FLOREA. « Cost-based cooperation of self-interested agents ». Dans *In Proc I. Dumitrache, editor, Proc. of the 12 th International Conference on Control Systems and Computer Science (CSCS12*, pages 62–67, 1999.

[Foulser *et al.* 1992] D. E. FOULSER, M. LI, et Q. YANG. « Theory and Algorithms for Plan Merging ». *Artificial Intelligence*, 57(2–3) :143–181, 1992.

[Geert *et al.* 2007] J. GEERT, D. FRANK, et M. JOHN-JULES. « Achieving cooperation among selfish agents in the air traffic management domain using signed money ». Dans *AAMAS '07 : Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 1–3, 2007.

[Genesereth & Nilsson 1987] M. GENESERETH et N. J. NILSSON. *The Logical Foundations of Artificial Intelligence*. Morgan Kaufmann, 1987.

[Georgeff 1983] M. P. GEORGEFF. « Communication and Interaction in Multi-Agent Planning ». Dans *AAAI*, pages 125–129, 1983.

[Good 1983] I. J. GOOD. *Good Thinking : The Foundations of Probability and Its Applications*. Minneapolis : University of Minnesota Press, 1983.

[Grant *et al.* 1997] S. GRANT, A. KAJII, et B. POLAK. « Weakening the Sure-Thing Principle : Decomposable Choice under Uncertainty ». STICERD - Theoretical Economics Paper Series 339, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE, 1997.

[Gravel *et al.* 2008] N. GRAVEL, T. MARCHANT, et A. SEN. « Ranking completely uncertain decisions by the uniform expected utility criterion ». *Third World Congress of the Game Theory Society (GAME'08)*, 2008.

[Green 1969] C. GREEN. « Application of Theorem Proving to Problem Solving ». Dans *IJCAI*, pages 219–240, 1969.

[Grosz *et al.* 1999] B. J. GROSZ, L. HUNSBERGER, et S. KRAUS. « Planning and Acting Together ». *AI Magazine*, 20(4) :23–34, 1999.

[Haddawy & Hanks 1998] P. HADDAWY et S. HANKS. « Interactive and Mixed-Initiative Decision-Theoretic Systems ». *AI Magazine*, 19(3) :133, 1998.

[Halpern & Tuttle 1993] J. HALPERN et M. TUTTLE. « knowledge, probability, and adversaries ». *Journal of the ACM 40 :4*, pages 917–962, 1993.

[Harrenstein *et al.* 2001] P. HARRENSTEIN, W. van der HOEK, J. MEYER, et C. WITTEVEEN. « Boolean games ». Dans *TARK*, page 287298, 2001.

[Harsanyi 1962] J. C. HARSANYI. « Bargaining in ignorance of the opponent's utility function ». *Journal ofConflict Resolution*, 6 :29–38, 1962.

[Harsanyi 1968] J. C. HARSANYI. « Games with incomplete information played by Bayesian players ». *Management Science*, 14 :159–182, 320–334, 486–502, 1967-1968.

[Herrero 1989] M. J. HERRERO. « The nash program : Non-convex bargaining problems ». *Journal of Economic Theory*, 49(2) :266–277, 1989.

[Hertzberg 1990] J. HERTZBERG. « Planen ». *KI*, 4(1) :22, 1990.

[Holmstrom & Myerson 1983] B. HOLMSTROM et R. B. MYERSON. « Efficient and durale decission rules with incomplete information ». *Econometrica : Journal of the Econometric Society*, 51 :1799 1820, 1983.

[Holzman & Monderer 2004] R. HOLZMAN et D. MONDERER. « Characterization of Ex Post Equilibrium in the VCGCombinatorial Auctions ». *Games and Economic Behavior*, 47 :87103, 2004.

[Holzman *et al.* 2004] R. HOLZMAN, N. KFIR-DAHAV, D. MONDERER, et M. TENNENHOLTZ. « Bundling equilibriumin combinatorial auctions ». *Games and Economic Behavior*, 47 :104123, 2004.

[Houthakker 1950] H. S. HOUTHAKKER. « Revealed preference and the utility function ». *Economica*, 17(66) :59174, 1950.

[Howard 1960] R. A. HOWARD. *Dynamic Programming and Markov Processes*. The M.I.T. Press, 1960.

[Huberman & Clearwater 1995] B. A. HUBERMAN et S. H. CLEARWATER. « A Multi-Agent System for Controlling Building Environments ». Dans *ICMAS*, pages 171–176, 1995.

[Hurwicz 1951] L. HURWICZ. « Some specification problems and application to econometric models (abstract) ». *Econometrica*, 19 :343–344, 1951.

[Hyafil & Bacchus 2003] N. HYAFIL et F. BACCHUS. « Conformant Probabilistic Planning via CSPs ». Dans *ICAPS*, pages 205–214, 2003.

[Hyafil & Boutilier 2004] N. HYAFIL et C. BOUTILIER. « Regret minimizing equilibria and mechanisms for games withstrict type uncertainty ». Dans *In Proceedings of the 20th Annual Conference on Uncertainty inArtificial Intelligence (UAI-04)*, page 268277, 2004.

[Jaynes 2003] E. T. JAYNES. « Probability Theory : The Logic of Science ». Dans *Studies in Resource Allocation Processes*. Cambridge UniversityPress, 2003.

[Jeffreys 1939] H. JEFFREYS. *Theory of Probability*. Clarendon Press, Oxford, 1939.

[Jennings 1996] N. R. JENNINGS. « Coordination techniques for distributed artificial intelligence ». Dans *Foundations of Distributed Artificial Intelligence*, pages 187–210. John Wiley and Sons, 1996.

[J.Nash 1951] J.NASH. « Non-cooperative games ». *Annals of Mathematics*, 54 :286–295, 1951.

[Kahneman & Tversky 1979] D. KAHNEMAN et A. TVERSKY. « Prospect Theory : An Analysis of Decision under Risk ». *Econometrica*, 47 :263–291, 1979.

[Kalai & Rosenthal 1996] E. KALAI et R.W. ROSENTHAL. « Arbitration of Two-Party Disputes under Ignorance ». *International Journal of Game Theory*, 7 :6572, 1996.

[Kalai & Smorodinsky 1975] E. KALAI et M. SMORODINSKY. « Other Solutions to Nash's Bargaining Problem ». *Econometrica*, 43(3) :513–518, 1975.

[Kalai 1977] E. KALAI. « Proportional Solutions to Bargaining Situations : Interpersonal Utility Comparisons ». Discussion Papers 179, Northwestern University, Center for Mathematical Studies in Economics and Management Science, Mar 1977.

[Kaneko 1980] M. KANEKO. « An Extension of Nash Bargaining Problem and theNash Social Welfare Function ». *Theory and Decision*, 12 :135–148, 1980.

[Kannai & Peleg 1984] Y. KANNAI et B. PELEG. « A note on the extension of an order on a set to the power set ». *Journal of Economic Theory*, 32(1) :172–175, 1984.

[Keynes 1921] J. M. KEYNES. *Treatise on Probability*. Macmillan And Co, 1921.

[Klemisch-Ahlert 1993] M. KLEMISCH-AHLERT. « Freedom of choice : a comparison of different rankings of opportunity sets ». *Social Choice and Welfare*, 10 :189–207, 1993.

[Knight 1921] F. H. KNIGHT. *Risk, Uncertainty and Profit*. Boston, MA : Hart, Schaffner Marx ;Houghton Mifflin Company, 1921.

[Koehler 1998] J. KOEHLER. « Planning under Resource Constraints ». Dans *ECAI*, pages 489–493, 1998.

[Kowalski 1979] R. KOWALSKI. *Logic for Problem Solving*. 1979.

[Kreps 1988] D. M. KREPS. *Notes on the theory of choice*. Boulder, Colorado : Westview Press, 1988.

[Krithivasan *et al.* 1999] K. KRITHIVASAN, A. SCHIRRA, et P.I. VIJAYKUMAR. « Velocity planning for a robot moving along the shortest straightline path among moving obstacles ». Dans *The Canadian Conference on Computational Geometry*, 1999.

[Kuhn 1953] H. W. KUHN. *Contributions to the Theory of Games, II (AM-28)*. Princeton University Press, 1953.

[Kushmerick *et al.* 1994] N. KUSHMERICK, S. HANKS, et D. S. WELD. « An Algorithm for Probabilistic Least-Commitment Planning ». Dans *AAAI*, pages 1073–1078, 1994.

[Kyburg & Smokler 1964] H. KYBURG et H. SMOKLER. *Studies in Subjective Probability*. Wiley, 1964.

[Lang 1996] Jérôme LANG. « Conditional Desires and Utilities : an Alternative Logical Approach to Qualitative Decision Theory ». Dans *ECAI*, pages 318–322, 1996.

[Lang *et al.* 2000] J. LANG, F. GARCIA, et AI MOUADDIB. *Modèles Décisionnels de Markov*. Temps, Espace et évolution, éditions Cépadus-Editions, 2000.

[Lehmann 1959] E. L. LEHMANN. *Testing Statistical Hypotheses*. John Wiley, 1959.

[Lehmann 1996] D. LEHMANN. « Generalized qualitative probability : Savage revisited ». Dans *In : Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence, UAI96*, pages 381–388. Morgan Kaufmann, 1996.

[Lesser 1998] V. R. LESSER. « Reflections on the Nature of Multi-Agent Coordination and Its Implications for an Agent Architecture ». Dans *Autonomous Agents and Multi-Agent Systems*, pages 89–111, 1998.

[Lifschitz 2000] V. LIFSCHITZ. « Review : M. Shanahan, Solving the Frame Problem ». *Artif. Intell.*, 123(1-2) :265–268, 2000.

[Littman *et al.* 1995] M. L. LITTMAN, T. DEAN, et L. P. KAELBLING. « On the Complexity of Solving Markov Decision Problems ». Dans *UAI*, pages 394–402, 1995.

[Lovejoy 1991] W.S. LOVEJOY. « A survey of algorithmic methods for partially observed Markov decision processes ». *Ann. Oper. Res.*, 28(1-4) :47–66, 1991.

[Luce & Raiffa 1957] R. LUCE et H. RAIFFA. *Games and Decisions*. J. Wiley, New York, 1957.

[Malone & Crowston 1994] T. W. MALONE et K. CROWSTON. « The interdisciplinary study of coordination ». *ACM Computing Surveys*, 26 :87–119, 1994.

[Marthy & Hayes 1969] J. MARTHY et P.J. HAYES. « Some Philosophical Problems fromthe Standpoint of Artificial Intelligence ». *Machine Intelligence 4*, 1969.

[Maskin 1979] E. MASKIN. « Decision Making Under Ignorance with Implications for Social Choice ». *Theory and Decision*, pages 319–337, 1979.

[Melliti *et al.* 2005] T. MELLITI, S. HADDAD, A. SUNA, et Amal El FALLAH-SEGHROUCHNI. « Web-MASI : Multi-Agent Systems Interoperability Using a Web Services Based Approach ». Dans *IAT*, pages 739–742, 2005.

[Milnor 1954] J. MILNOR. *Games against nature*. Wiley, New York, 1954.

[Monderer & Tennenholtz 2006] D. MONDERER et M. TENNENHOLTZ. « Strong mediated equilibrium ». Dans *AAAI*, 2006.

[Mongin 1997] P. MONGIN. « Expected Utility Theory ». *Handbook of Economic Methodology*, pages 342–350, 1997.

[Munson 1971] J. H. MUNSON. « Robot Planning, Execution, and Monitoring in an Uncertain Environment ». Dans *IJCAI*, pages 338–349, 1971.

[Nagel 1939] E. NAGEL. *Principles of the Theory of Probability*. University of Chicago Press, 1939.

[Nash 1950] J. NASH. « The Bargaining Problem ». *Econometrica*, 18 (2) :155162, 1950.

[Nehring & Puppe 1996] K. NEHRING et C. PUPPE. « Continuous Extensions of an Order on a Set to the Power Set ». *Journal of Economic Theory*, 68(2) :456–479, 1996.

[Nilsson 1982] N.J. NILSSON. *Principles of Artificial Intelligence*. Springer, 1982.

[Nitzan & Pattanaik 1984] S. NITZAN et P.K. PATTANAIK. « Median-based extensions of an ordering over a set to the power set :an axiomatic characterization ». *Journal of Economic Theory*, 34 :252–261, 1984.

[Nwana *et al.* 1996] H. S. NWANA, L. C. LEE, et N. R. JENNINGS. « Co-ordination in Software Agent Systems ». *The British Telecom Technical Journal*, 14(4) :79–88, 1996.

[Palacios & Geffner 2006] H. PALACIOS et H. GEFFNER. « Compiling Uncertainty Away : Solving Conformant Planning Problems using a Classical Planner (Sometimes) ». Dans *AAAI*, 2006.

[Palacios & Geffner 2007] H. PALACIOS et H. GEFFNER. « From Conformant into Classical Planning : Efficient Translations that May Be Complete Too ». Dans *ICAPS*, pages 264–271, 2007.

[Pattanaik & Peleg 1984] P. K. PATTANAIK et B. PELEG. « An axiomatic characterization of the lexicographic maximin extension of an ordering over a set to the power set ». *Social Choice and Welfare*, 1 :113–122, 1984.

[Pearl 1996] Judea PEARL. « Decision Making Under Uncertainty ». *ACM Comput. Surv.*, 28(1) :89–92, 1996.

[Puppe 1995] C. PUPPE. « Freedom of choice and rational decisions ». *Social Choice and Welfare*, 12 :137–153, 1995.

[Ramsey 1931] F.P. RAMSEY. « Truth and probability ». Dans *Foundations of Mathematics and other Logical Essays*, pages 156–198. Routledge & P. Kegan, London, 1931.

[Rathod & des Jardins 2004] P. RATHOD et M. des JARDINS. « Stable Team Formation Among Self-Interested Agents ». Dans *AAAI Workshop on Forming and Maintaing Coalitions in Adaptive Multiagent Systems*, 2004.

[Richter 1971] M. K. RICHTER. *Rational choice*. J. S. Chipman, L. Hurwicz, M. K. Richter, and H. F.Sonnenschein, eds., Preferences, Utility, and Demand : A Minnesota Symposium. New York : Harcourt, Brace, Jovanovich, 1971.

[Rosenschein 1995] J. S. ROSENSCHEIN. « Multiagent Planning as a Social Process : Voting, Privacy, and Manipulation ». Dans *ICMAS*, pages 431–431, 1995.

[Roth 1977] A. E. ROTH. « Individual Rationality and Nash's Solution to the Bargaining Problem ». *Mathematics of operations research*, 2 :64–65, 1977.

[Rozenfeld & Tennenholtz 2007] O. ROZENFELD et M. TENNENHOLTZ. « Routing mediators ». Dans *International Joint Conferences on Artificial Intelligence*, page 14881493, 2007.

[Rubinstein *et al.* 1992] Ariel RUBINSTEIN, Zvi SAFRA, et William THOMSON. « On the Interpretation of the " Bargaining Solution and Its Extension to Non-expected Utility Preferences ». *Econometrica*, 60(5) :1171–86, 1992.

[Sabbadin 2000] Régis SABBADIN. « Empirical Comparison of Probabilistic and Possibilistic Markov Decision Processes Algorithms ». Dans *ECAI*, pages 586–590, 2000.

[Sacerdoti 1975] E. D. SACERDOTI. « The nonlinear nature of plans ». Dans *Proceedings of the 4th International Joint Conference on Artificial Intelligence (IJCAI-75)*, pages 206–214, 1975.

[Sacerdoti 1977] E.D. SACERDOTI. *A Structure for Plans Behavior*. Elsevier, 1977.

[Samuelson 1938] P. A. SAMUELSON. « A note on the pure theory of consumer's behavior ». *Economica N.S*, page 353354, 1938.

[Samuelson 1950] P. A. SAMUELSON. « The problem of integrability in utility theory ». *Economica*, 17(68) :355385, 1950.

[Sandholm & Lesser 1997] T. SANDHOLM et V. R. LESSER. « Coalitions Among Computationally Bounded Agents ». *Artificial Intelligence*, 94(1-2) :99–137, 1997.

[Savage 1954] L.J. SAVAGE. *The Foundations of Statistics*. J. Wiley, 1954. second revised edition, 1972.

[Schank & Abelson 1975] R.C. SCHANK et R.P. ABELSON. « Scripts, Plans, and Knowledge ». Dans *In of the 4th International Joint Conference on Artificial Intelligence (IJCAI-75)*, volume 1, pages 151–157, 1975.

[Schelling 1960] T. C. SCHELLING. *The strategy of conflict*. Cambridge, MA : Harvard University Press, 1960.

[Sen 1969] A. K. SEN. « Quasi-transitivity, rational choice and collective decisions ». *Review ofEconomic Studies*, page 381393, 1969.

[Sen 1970] A.K. SEN. *Collective Choice and Social Welfare*. North Holland, Amsterdam, 1970.

[Shackle 1955] G. L. S. SHACKLE. « A Non-Additive Measure of Uncertainty ». *Reprinted in Uncertainty in Economics*, 1955.

[Shehory & Kraus 1998] O. SHEHORY et S. KRAUS. « Methods for Task Allocation via Agent Coalition Formation ». *Artificial Intelligence*, 101(1-2) :165–200, 1998.

[Shoham & Tennenholtz 1995] Y. SHOHAM et M. TENNENHOLTZ. « On Social Laws for Artificial Agent Societies : Off-Line Design ». *Artificial Intelligence*, 73(1-2) :231–252, 1995.

[Smith & Weld 1999] D. SMITH et D. WELD. « Temporal Planning with Mutual Exclusion Reasoning ». Dans *IJCAI*, 1999.

[Smith 1980] R. G. SMITH. « The Contract Net Protocol : High-LevelCommunication and Control in a DistributedProblem Solver ». *IEEE Transactions on computers*, C-29 :1104–1113, 1980.

[Stiglitz 2005] J. E. STIGLITZ. « The Ethical Economist ». *Foreign Affairs*, 2005.

[Tate 1977] A. TATE. « Generating Project Networks ». Dans *IJCAI*, pages 888–893, 1977.

[ter Mors *et al.* 2004] A. ter MORS, J. VALK, et C. WITTEVEEN. « Coordinating Autonomous Planners ». Dans *IC-AI*, 2004.

[Thomson 1994] W. THOMSON. « Cooperative models of bargaining ». *Handbook of Game Theory with Economic Applications*, pages 1237–1284, 1994.

[Thorndyke *et al.* 1981] P. W. THORNDYKE, D. MCARTBUR, et S. J. CAMMARATA. « AUTOPILOT : A Distributed Planner for Air Fleet Control ». Dans *IJCAI*, pages 171–177, 1981.

[Tsamardinos *et al.* 2000] I. TSAMARDINOS, M. E. POLLACK, et J. F. HORTY. « Merging Plans with Quantitative Temporal Constraints, Temporally Extended Actions, and Conditional Branches ». Dans *AIPS*, pages 264–272, 2000.

[Uzawa 1956] H. UZAWA. « Note on preferences and axioms of choice ». *Annals of the Institute ofStatistical Mathematics*, page 3540, 1956.

[Vickrey 1961] W. VICKREY. « Counterspeculation, Auctions, and Competitive Sealed Tenders ». *Journal of Finance*, 1961.

[von Martial 1992] Frank von MARTIAL. « Coordinating Plans of Autonomous Agents ». *Springer*, 1992.

[von Neumann & Morgenstern 1947] J. von NEUMANN et O. MORGENSTERN. *Theory of games and economic behaviour*. Princeton University Press, Princeton, 1947. 2nd edition.

[von Neumann 1928] J. von NEUMANN. « Zur Theorie der Gesellschaftsspiele ». *Mathematische Annalen*, 100 :295300(1959 p. 13), 1928.

[Wald 1950] A. WALD. *Statistical Decision Functions*. John Wiley, 1950.

[Walsh & Wellman 1998] W. E. WALSH et M. P. WELLMAN. « A Market Protocol for Decentralized Task Allocation ». Dans *ICMAS*, pages 325–332, 1998.

[Walsh & Wellman 1999] W. E. WALSH et M. P. WELLMAN. « Efficiency and Equilibrium in Task Allocation Economies with Hierarchical Dependencies ». Dans *IJCAI*, pages 520–526, 1999.

[Wellman & Doyle 2000] M. P. WELLMAN et J. DOYLE. « Modular utility representation for decision-theoretic planning ». *First International Conference on AI Planning Systems*, 2000.

[Wellman 1990] M. P. WELLMAN. « The STRIPS Assumption for Planning Under Uncertainty ». Dans *AAAI*, pages 198–203, 1990.

[Wilkins 1988] D. WILKINS. *Practical Planning*. Morgan Kaufmann, 1988.

[Wittgenstein 1986] L. WITTGENSTEIN. *Tractatus logico-philosophicus*. Halsted Press, New York, 1986.

[Wood 1983] S. WOOD. « Dynamic World Simulation for Planning With Multiple Agents ». Dans *IJCAI*, pages 69–71, 1983.

Publications

[Ben Larbi *et al.* 2007a] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « Extending Classical Planning to the Multi-agent Case : A Game-Theoretic Approach ». Dans *9th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty(ECSQARU'07)*, pages 731–742. Springer, 2007.

[Ben Larbi *et al.* 2007b] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « Planification multi-agent et diagnostique stratégique ». Dans *4èmes journées francophones sur les Modèles Formels de l'Interaction (MFI'07)*, pages 25–36, 2007.

[Ben Larbi *et al.* 2008a] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « A Characterization of an Optimality Criterion for Decision Making under Complete Ignorance ». Dans *Workshop on Decision, Games and Logic(DGL'08)*, Amsterdam, 2008.

[Ben Larbi *et al.* 2008b] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « A Characterization of an Optimality Criterion for Decision Making under Complete Ignorance ». Dans *12th International Workshop on Non-Monotonic Reasoning(NMR'08)*, pages 268–274, Sydney, 2008.

[Ben Larbi *et al.* 2008c] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « Jeux Qualitatifs à Résultats Multiples ». Dans *Journées Francophones sur les Systèmes Multi-Agents(JFSMA'08)*, 2008.

[Ben Larbi *et al.* 2008d] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « A model for multiple outcomes games ». Dans *20th IEEE International Conference on Tools with Artificial Intelligence(ICTAI'08)*, pages 27–34, vol1, 2008.

[Ben Larbi *et al.* 2008e] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « Qualitative multiple outcomes games with consensus ». Dans *Third World Congress of the Game Theory Society(GAME 2008)*, 2008.

[Ben Larbi *et al.* 2010] Ramzi BEN LARBI, Sébastien KONIECZNY, et Pierre MARQUIS. « A Characterization of Optimality Criteria for Decision Making under Complete Ignorance ». Dans *Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning (KR'10)*, page (à paraître), 2010.