



HAL
open science

Apprentissage de préférences sous incertitude sévère

Loïc Adam

► **To cite this version:**

Loïc Adam. Apprentissage de préférences sous incertitude sévère. Informatique [cs]. Université de Technologie de Compiègne, 2023. Français. NNT : 2023COMP2768 . tel-04529292v2

HAL Id: tel-04529292

<https://theses.hal.science/tel-04529292v2>

Submitted on 22 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Par Loïc ADAM

*Apprentissage de préférences
sous incertitude sévère*

Thèse présentée
pour l'obtention du grade
de Docteur de l'UTC



Soutenue le 23 octobre 2023

Spécialité : Informatique : Unité de recherche Heudyasic
(UMR-7253)

D2768



Thèse de doctorat

Apprentissage de préférences sous incertitude sévère

Spécialité : Informatique

Présentée et soutenue publiquement le 23 octobre 2023 par

Loïc ADAM

DIRECTION DE THÈSE ET COMPOSITION DU JURY :

- Directeur de thèse [Sébastien DESTERCKE](#)
Directeur de recherche CNRS, Université de technologie de Compiègne
- Rapporteurs [Hélène FARGIER](#)
Directrice de recherche CNRS, Université Toulouse III
[Wassila OUERDANE](#)
Maîtresse de conférences, CentraleSupélec
- Examineurs [Nadjet BOURDACHE](#)
Maîtresse de conférences, Université de Caen Normandie
[Frédéric KORICHE](#)
Professeur des universités, Université d'Artois
[Sylvain LAGRUE](#)
Professeur des universités, Université de technologie de Compiègne
[Olivier SPANJAARD](#)
Maître de conférences, Sorbonne Université

*À mon grand-père Bernard, parti trop tôt.
À mon grand-père Noël, que j'aurais aimé connaître.*

Remerciements

Je tiens tout d'abord à remercier mon directeur de thèse, Sébastien Destercke, qui m'a encadré durant ces trois années. Sébastien était toujours bienveillant et disponible pour répondre à mes très nombreuses questions et à mes doutes, souvent avec humour, et ce, malgré mes difficultés à m'exprimer clairement...

Je remercie les deux rapporteuses de cette thèse, Hélène Fargier et Wassila Ouerdane, qui ont lu attentivement ce manuscrit et m'ont fait des retours très pertinents. Je remercie aussi l'intégralité de mon jury, qui m'a suggéré avec bienveillance de nombreuses améliorations et pistes de recherche durant ma défense, que j'espère pouvoir explorer durant la suite de mon parcours.

Merci à toutes les personnes avec lesquelles j'ai interagi au laboratoire. Plus particulièrement Dritan qui m'a encouragé à poursuivre en master et en thèse ; ainsi qu'Antoine, Arthur, Benjamin, David, Ghislaine, Jean-Benoist, Jean-Paul, Khaled, Marius, Nathalie, Sabine, Stéphane, Sylvain L., Sylvain R. et Yves.

Je souhaite aussi remercier les amis qui m'ont accompagné durant ces trois années. Je ne pourrai pas citer tout le monde, mais je tiens à particulièrement remercier au laboratoire les imposteurs Lahcene, Maxime, Nicolas, Sana et Soudouss ; ainsi que Henoïk, Jérôme et Michael. À l'extérieur, un grand merci à Ano, Magali, Mario, Miki, Nekky, Nicolas, Nora, Senuty, Stenu, Steveh et Sylvain.

Enfin, je tiens bien évidemment à remercier très chaleureusement mes parents, mon frère Florian et mes deux mamies pour leur soutien sans faille, y compris dans les moments les plus difficiles.

Résumé

Les travaux présentés dans ce manuscrit s'inscrivent dans le cadre de l'aide à la décision multicritère, et plus particulièrement dans la problématique du choix, où un analyste aide à un décideur à prendre la meilleure décision entre plusieurs alternatives multicritères explicites et en nombre fini. L'élicitation de préférences, c'est-à-dire la formalisation des préférences du décideur par l'analyste, permet de déterminer quelle alternative correspond le mieux aux préférences du décideur. Nous nous intéressons dans cette thèse à l'élicitation incrémentale basée sur l'optimisation d'un critère de décision, qui permet de garantir une recommandation finale de qualité, tout en minimisant la charge cognitive pour le décideur. Néanmoins, dans ce type d'élicitation, les approches robustes basées sur des ensembles sont très sensibles aux différentes sources d'incertitude, et les approches probabilistes n'offrent que des garanties en espérance sur la recommandation finale.

Notre objectif ici est de proposer une nouvelle approche qui permet de gérer toutes les sources d'incertitude, conservant la plupart des avantages des approches ensemblistes. Pour cela, nous modélisons des préférences incertaines et nous étendons l'approche ensembliste avec la théorie des possibilités, un cadre des probabilités imprécises qui permet de modéliser de l'incertitude. Nous montrons dans cette thèse que notre approche permet de détecter de l'incohérence, quelle qu'en soit la source, tout en offrant des recommandations de meilleure qualité que les approches ensemblistes en cas d'incohérence. De plus, nous avons étudié dans cette thèse plusieurs moyens d'inférer malgré l'incohérence, par exemple sans modifier les informations préférentielles du décideur et sans restaurer la cohérence, ou en restaurant la cohérence via la fusion d'information. Nous étayons nos propositions par des expériences synthétiques qui confirment leur intérêt.

Table des matières

Liste des Notations	vii
0 Introduction	1
1 Modélisation et élicitation de préférences incertaines	5
1.1 Introduction	5
1.2 Décision multicritère	7
1.2.1 Dominance de Pareto	8
1.2.2 Agrégation de préférences	11
1.2.3 Décision sous incertitude ensembliste	23
1.3 Élicitation de préférences	29
1.3.1 Objectif d'élicitation	31
1.3.2 Élicitation non incrémentale	33
1.3.3 Élicitation incrémentale	35
1.4 Gestion de l'incertitude et fusion d'information	46
1.4.1 Probabilités imprécises	47
1.4.2 Théorie des possibilités	50
1.4.3 Principe et opérateurs de fusion	54
1.4.4 Fusion en théorie des possibilités : T-normes et T-conormes .	61
1.5 Conclusion	67
2 Élicitation possibiliste de préférences	69
2.1 Introduction	69
2.2 Élicitation avec de l'incertitude	73

2.2.1	Erreurs dans les approches ensemblistes	73
2.2.2	Compléments en théorie des possibilités	75
2.2.3	Information préférentielle possibiliste	76
2.3	Regret possibiliste et adaptation du CSS	80
2.3.1	Extension du regret au cas possibiliste	80
2.3.2	Extension avec l'heuristique CSS	85
2.3.3	Éléments de complexité	88
2.4	Expériences	92
2.4.1	Protocole expérimental	93
2.4.2	Convergence et lien avec la T-norme et le niveau de confiance	93
2.4.3	Détection d'erreurs de la part du décideur	96
2.4.4	Détection d'un modèle de préférence erroné	99
2.5	Conclusion et perspectives	100
3	Choix du critère de décision	103
3.1	Introduction	103
3.2	Approche Optimiste	105
3.2.1	Élicitation robuste avec un gain Maximax	105
3.2.2	Optimalité potentielle	108
3.2.3	Convergence	109
3.2.4	Complexité computationnelle	110
3.3	Expériences	110
3.3.1	Protocole expérimental	111
3.3.2	Résultats	112
3.3.3	Discussion des résultats	116
3.4	Conclusion et perspectives	119
4	Manipuler l'incohérence	121
4.1	Introduction	121
4.2	Compléments en théorie des possibilités	123
4.3	Résoudre l'incohérence par la fusion d'information	125
4.3.1	Inférer malgré les incohérences	127
4.3.2	Résoudre les incohérences par la fusion d'information	129

4.4	Expériences	136
4.4.1	Protocole expérimental	136
4.4.2	Nombre de réponses incorrectes détectées (PCSS)	139
4.4.3	Comparaison des méthodes pour gérer l'incertitude (PCSS)	142
4.4.4	Influence de la règle de décision et élicitation par batch	147
4.4.5	Discussion des résultats	149
4.5	Conclusion et perspectives	150
5	Conclusion	161
A	MCSs multidimensionnels	165
A.1	Introduction	165
A.2	Compléments sur les MCSs de polytopes	166
A.3	Énumération des MCSs d'hyperrectangles alignés sur les axes	167
A.4	Approximation de polytopes avec des hyperrectangles	170
A.4.1	Hyperrectangle inscrit aligné sur les axes d'hypervolume maximal (MVAIR)	171
A.4.2	Hyperrectangle englobant aligné sur les axes d'hypervolume minimal (MVABH)	172
A.4.3	Encadrement des MCSs	173
A.5	Illustration sur une estimation linéaire	174
A.5.1	Estimer les modèles linéaires possibles	174
A.5.2	Application	176
A.6	Conclusion et perspectives	179
B	Bibliographie	181
	Liste des Figures	199
	Liste des Tables	203
	Liste des Algorithmes	205

Liste des Notations

Modélisation de préférences et agrégation de préférences

x	Alternative
\mathbb{X}	Sous-ensemble des alternatives disponibles
$u_j(x)$	Performance d'une alternative x par rapport au critère j
\mathcal{Q}	Ensemble des critères
\succsim	Préférence
\succsim_P	Pareto-dominance faible
\succ_P	Pareto-dominance forte
\gg_P	Pareto-dominance stricte
\mathbf{ND}_P	Alternatives non dominées (Pareto-dominance forte)
f_ω	Fonction d'agrégation de critères de paramètre ω
Ω	Ensemble des valeurs possibles de ω
\mathbf{WS}_ω	Somme pondérée
\mathbf{OWA}_ω	Somme pondérée ordonnée
C_v	Intégrale de Choquet avec une capacité v

Élicitation de préférences

E_k	Sous-ensemble des valeurs possibles de ω après k étapes
$\mathbf{R}_\omega(x, y)$	Regret
$\mathbf{PMR}(x, y, \Omega)$	Regret maximal par paire
$\mathbf{MR}(x, \Omega)$	Regret maximal
$\mathbf{mMR}(\Omega)$	Regret maximal minimal
PO_Ω	Ensemble des solutions Ω -optimales

(E_i, α_i)	Information préférentielle possibiliste
α	Niveau de confiance
$\mathbf{PEMR}(x, y, \pi_k)$	Extension du regret maximal par paire
$\mathbf{EMR}(x, \pi_k)$	Extension du regret maximal
$\mathbf{mEMR}(\pi_k)$	Extension du regret maximal minimal
$\{\Omega_1, \dots, \Omega_P\}$	Partition de Ω en polytopes constants par rapport à ω

Théorie des possibilités

π	Distribution de possibilité
Π	Mesure de possibilité
N	Mesure de nécessité
π^α	Coupe alpha
$\mathbf{Inc}(\pi)$	Degré d'incohérence
\mathbb{E}	Espérance
$\underline{\mathbb{E}}$	Espérance inférieure
$\overline{\mathbb{E}}$	Espérance supérieure

Fusion d'information

S	Information
\mathcal{S}	Ensemble d'informations
\mathbf{T}	T-norme
\perp	T-conorme
c	Sous-ensemble cohérent maximal
\mathcal{C}	Ensemble des sous-ensembles cohérent maximal

Annexe

\mathcal{P}	Ensemble de polytopes P_i
\mathcal{H}	Ensemble d'hyperrectangles H_i
I_H^d	Projection d'un hyperrectangle H sur la dimension d
H^I	Projection interne d'un hyperrectangle H
H^O	Projection externe d'un hyperrectangle H

Chapitre 0

Introduction

L'aide à la décision multicritère (MCDA) est une branche de la recherche opérationnelle qui s'intéresse à l'évaluation de plusieurs alternatives ayant chacune plusieurs critères, avec une analyste qui a pour but d'aider un décideur à prendre une décision parmi ces alternatives. Un critère est une caractéristique d'une alternative sur laquelle le décideur se base pour prendre une décision, par exemple le prix d'une voiture ou sa vitesse. La difficulté de ces problèmes provient des critères qui sont en général conflictuels : augmenter le gain d'un critère fait souvent diminuer le gain d'autres critères. Par exemple, une voiture moins chère est souvent moins performante. Le décideur doit donc le plus souvent faire des compromis sur les différents critères, et l'analyste doit l'aider à déterminer quels compromis sont les plus adaptés. L'aide à la décision multicritère s'appuie alors sur de nombreuses disciplines comme les mathématiques, l'informatique, la philosophie, l'économie ou encore la sociologie pour comprendre comment un décideur prend une décision, et ainsi déterminer la meilleure stratégie pour l'assister. Cette branche de recherche est assez récente avec les premiers travaux datant des années 1960-1970 (Benayoun *et al.*, 1971; Keeney et Raiffa, 1976; Zionts et Wallenius, 1976).

Afin d'aider le décideur à prendre une décision, il est impératif de déterminer ses préférences, pour comprendre quels critères et quels compromis il privilégie. Demander directement au décideur comment il juge exactement une alternative paraît compliqué et donc illusoire. En gestion des connaissances, il est courant d'élucider les préférences du décideur, c'est-à-dire de formaliser les préférences pour pouvoir les exploiter. Pour cela, l'analyste décide d'une famille de modèle de

préférence pour le décideur, et à travers plusieurs questions, l'analyste détermine le modèle de préférence qui correspond au mieux aux préférences du décideur, ce qui permet ensuite de déterminer quelle décision prendre à partir des alternatives disponibles. Il existe de nombreux modèles pour modéliser l'attrait pour une alternative multicritère, par exemple la moyenne pondérée, la somme pondérée ordonnée (OWA, Yager (1988) ou l'intégrale de Choquet (Choquet, 1953), ce dernier modèle permettant de modéliser des interactions complexes entre les critères.

L'intérêt de l'aide à la décision multicritère est son applicabilité à de nombreuses situations. Ainsi, il est possible de concevoir des outils de recommandation et d'aide à la décision dans différents contextes, par exemple pour juger différents candidats pour un poste, gérer des budgets, recommander des produits... Néanmoins, dans des contextes plus critiques, comme dans le médical ou le militaire, il paraît primordial de garantir une aide à la décision robuste à de petits changements, d'autant plus que les informations sur les préférences du décideur peuvent être incertaines (manque de concentration, alternatives trop semblables ou distantes...). La gestion de l'incertitude est une problématique de plus en plus étudiée, avec notamment l'émergence des probabilités imprécises (Augustin *et al.*, 2014) qui offrent différents cadres pour représenter et traiter l'incertitude. En manipulant correctement l'incertitude, il est possible de bien l'évaluer et donc de bien l'analyser et de la corriger, ce qui permet d'améliorer le processus d'élicitation, notamment si le problème initial est mal modélisé ou que le décideur est peu fiable.

Positionnement de la thèse Dans cette thèse, nous avons choisi de travailler avec des modèles numériques pour représenter les préférences d'un individu, comme nous le détaillons dans la Section 1.2. Nous étudions aussi un type particulier d'élicitation : l'élicitation incrémentale robuste (Wang et Boutilier, 2003). Nous présentons néanmoins dans la Section 1.3 d'autres approches d'élicitation, tout en justifiant notre intérêt pour l'approche incrémentale robuste. Cette approche en question, minimisant l'effort cognitif demandé et offrant de fortes garanties de performance, n'est pas robuste à certaines formes d'incertitude, justifiant des travaux qui permettent de modéliser et de manipuler cette incertitude. Dans cette thèse, nous modélisons l'incertitude dans les informations préférentielles, et non dans les alternatives ou dans le modèle de préférence. Nous nous focalisons alors sur

la théorie des possibilités (Dubois et Prade, 2001), un des cadres des probabilités imprécises présenté dans la Section 1.4, pour modéliser et éliciter nos informations préférentielles incertaines.

Le premier objectif de cette thèse est de proposer une approche efficace d'élicitation incrémentale robuste qui permet de manipuler des préférences incertaines et de détecter de l'incohérence, tout en ayant les garanties fortes au niveau de la convergence et de qualité de la recommandation finale (au moins quand il n'y a aucune incohérence entre les informations préférentielles), les approches de l'état de l'art offrant seulement l'un des deux. Un deuxième objectif est d'ensuite proposer des stratégies pour inférer malgré l'incohérence, mais aussi des outils pour rétablir la cohérence, afin de retrouver des recommandations de qualité malgré les erreurs encourues initialement durant le processus d'élicitation.

Organisation du document Ce document se décompose en six chapitres, dont un d'introduction et un de conclusion que nous ne présenterons pas ici :

- Le Chapitre 1 introduit de nombreuses notions qui seront utilisées dans les chapitres suivants. Plus particulièrement, nous voyons dans ce chapitre comment modéliser des préférences dans un problème de décision multicritère, nous faisons ensuite un rapide état de l'art des différentes méthodes d'élicitation de préférences, et nous voyons enfin comment gérer de l'incertitude.
- Le Chapitre 2 présente notre méthode d'élicitation basée sur la théorie des possibilités, qui permet de manipuler des informations préférentielles incertaines et de détecter de l'incohérence, quelle qu'en soit la source.
- Le Chapitre 3 compare deux règles de décision pour recommander une alternative à un décideur durant un processus d'élicitation : une approche pessimiste standard dans la littérature et une approche optimiste.
- Le Chapitre 4 propose de restaurer la cohérence une fois que notre méthode d'élicitation possibiliste détecte de l'incohérence, principalement quand le décideur est la source de l'incohérence. Nous proposons deux types d'approche : une qui ne cherche pas à modifier les informations préférentielles, et une qui modifie la manière dont les informations préférentielles sont combinées en une information synthétique.

Les chapitres 2, 3 et 4 contiennent les différentes contributions de cette thèse. De plus, à la fin de chacun de ces chapitres, nous mettons en avant les différentes perspectives de recherche, ainsi que les contributions dans des conférences et des revues scientifiques. Le Chapitre 5 de conclusion ne revient que rapidement sur les perspectives de recherches.

Modélisation et élicitation de préférences incertaines

1.1 Introduction

Prendre une décision est un acte omniprésent dans la vie de tous les jours. Quand nous allons aux courses, nous hésitons souvent entre plusieurs produits, mais sans conséquences majeures si nous choisissons le ou les mauvais. Par contre, lors d'un achat important (voiture, ordinateur, etc) ou d'un choix de carrière, un mauvais choix peut se montrer désastreux, rendant la prise de décision à la fois capitale et difficile. Dès lors, il est donc important de proposer à l'individu, appelé décideur dans la littérature, une aide qui va l'aider à faire le meilleur choix.

Dans ce travail, nous nous intéressons à l'aide à la décision multicritère, c'est-à-dire que l'attrait pour une alternative dépend de plusieurs facteurs (par exemple le prix, la qualité...). L'aide à la décision multicritère est un objet de recherche depuis la fin des années 1960-début des années 1970 (voir Cochrane et Zeleny (1975); Thiriez et Zionts (1975); Keeney et Raiffa (1976); Bell *et al.* (1977)). La difficulté provient du possible conflit entre les différents critères, qui doivent être pris en compte simultanément. Ainsi, un produit est rarement à la fois peu cher et très qualitatif, mais un compromis. Il est souvent impossible de déterminer si une alternative est meilleure qu'une autre sans avoir connaissance des préférences du décideur : une alternative chère et de qualité n'est pas intrinsèquement meilleure qu'une alternative pas chère et de mauvaise qualité.

Le plus souvent, l'aide à la décision est faite par une analyste spécialisée dans la théorie de la décision. Le travail de l'analyste peut se diviser en deux phases :

- Une *phase de structuration*, où l'analyste doit déterminer les alternatives possibles, leurs conséquences, établir les critères d'évaluation, collecter des informations sur les préférences subjectives du décideur, etc.
- Une *phase de résolution*, où l'analyste propose une recommandation pertinente à partir des informations préférentielles récoltées.

La phase de structuration est très importante, car elle résume toutes les informations à prendre en compte pour déterminer la recommandation à donner et est à l'origine de différentes recherches (par exemple von Winterfeldt (1980); Buede (1986); Belton et Stewart (2010)). En plus des différents critères conflictuels à prendre en compte, d'autres difficultés existent dans la phase de structuration, par exemple des opinions divergentes (choix social) ou des conséquences non prévisibles avec certitude (décision dans l'incertain). Remarquons qu'en plus, il existe d'autres difficultés : les alternatives peuvent être très nombreuses, et les préférences du décideur très complexes. Il est donc important de correctement modéliser ses préférences. Or, les préférences du décideur peuvent être incertaines pour de nombreuses raisons (incertitude, fatigue, alternatives difficiles à comparer...), rendant le processus encore plus complexe, nécessitant de correctement capturer et gérer cette incertitude.

En phase de résolution, trois grandes problématiques sont usuellement distinguées (Roy, 1996) :

- La problématique de *choix* : aider à sélectionner une ou plusieurs alternatives que le décideur préfère.
- La problématique de *rangement* : aider à classer les différentes alternatives par ordre de préférence du décideur.
- La problématique de *tri* : aider à affecter les différentes alternatives dans différentes catégories prédéfinies, selon les préférences du décideur.

Dans ce chapitre, nous commençons par étudier la problématique de la modélisation de préférences en décision multicritère (Section 1.2). Par la suite, nous présentons différentes approches proposées dans la littérature pour éliciter des

préférences, en nous focalisant sur l'élicitation incrémentale (Section 1.3). Enfin, nous étudions différents outils pour représenter l'incertitude, plus particulièrement la théorie des possibilités, et comment fusionner des informations incohérentes afin de manipuler des préférences incertaines et potentiellement inconsistantes dans les chapitres suivants (Section 1.4).

1.2 Décision multicritère

Un problème de décision multicritère est caractérisé par la prise en compte de plusieurs critères qui représentent différents objectifs du décideur, par exemple minimiser le prix, maximiser la qualité ou l'esthétique... Nous supposons que nous avons plusieurs critères $j \in \{1, \dots, q\} = \mathcal{Q}$. Chaque critère a pour domaine X_j . L'ensemble $X = X_1 \times \dots \times X_q$ est l'espace des critères qui décrit l'ensemble des décisions possibles. Nous supposons que chaque alternative x peut être décrite par le vecteur $(u_1(x), \dots, u_q(x)) \in X$, où $u_i(x)$ est la performance (l'utilité) de l'alternative par rapport au j^{e} critère, prenant valeur dans \mathbb{R}^+ . Par la suite, nous notons $\mathbb{X} \subseteq X$ le sous-ensemble des alternatives disponibles parmi toutes les décisions possibles. Nous posons aussi $x_j = u_j(x)$ pour simplifier les notations. Enfin, nous supposons que les décisions sont d'autant meilleures que leurs performances x_j sont élevées.

Chaque critère $j \in \mathcal{Q}$ induit un préordre total sur \mathbb{X} noté \geq_j , représentant les préférences du décideur par rapport au j^{e} critère. Cette relation est définie par :

$$\forall x, y \in \mathbb{X}, x \geq_j y \Leftrightarrow x_j \geq y_j,$$

où $x \geq_j y$ indique que x est au moins aussi bon que y sur le critère j . \geq_j est un préordre total sur x_j , c'est-à-dire une relation binaire avec les propriétés suivantes :

- (Réflexivité) : $\forall x \in \mathbb{X}, x \geq_j x$,
- (Transitivité) : $\forall x, y, z \in \mathbb{X}, (x \geq_j y \wedge y \geq_j z) \Rightarrow x \geq_j z$,
- (Totalité) : $\forall x, y \in \mathbb{X}, x \geq_j y \vee y \geq_j x$.

Notons que \geq_j n'est pas un ordre : nous n'avons pas la propriété d'antisymétrie, car $\exists x, y \in \mathbb{X}$ tel que $x_j = y_j \not\Rightarrow x = y$ (les autres critères peuvent être différents). Les relations \geq_j permettent de comparer les alternatives par rapport au j^{e} critère. Néanmoins, elles ne permettent pas de résoudre le problème de décision : les critères

n'évoluent pas nécessairement de la même manière. Par exemple, augmenter un critère peut faire diminuer les autres. De plus, il est très rare d'avoir une alternative qui soit meilleure que toutes les autres sur tous les critères.

Exemple 1. *Présentation d'un problème de décision multicritère*

Nous supposons qu'une cliente souhaite acheter un fromage parmi une sélection limitée de fromage. Chaque fromage est une alternative multicritère parmi un ensemble d'alternatives disponibles \mathbb{X} représenté dans la Table 1.1. Le préordre sur le goût \geq_1 nous indique que le brie à la truffe est le choix le plus gouteux, tandis que le cheddar américain est le pire. A contrario, le préordre sur le prix \geq_2 nous indique que le cheddar américain est le moins cher, tandis que le prix à la truffe est le plus cher, donc le pire choix. Par conséquent, il n'est pas possible de prendre une décision, car le goût évolue inversement au prix.

Nom	Goût	1/Prix
Cheddar américain	0	9
Emmental	4	6
Mozzarella	5	5
Gorgonzola	7	3
Brie à la truffe	8	1

TABLE 1.1 – Ensemble d'alternatives \mathbb{X}

Dans la suite de cette section, nous allons voir comment prendre une bonne décision. Tout d'abord dans la Sous-Section 1.2.1 nous présentons la dominance de Pareto qui est une règle de dominance très naturelle pour sélectionner des alternatives multicritères. Ensuite, dans la Sous-Section 1.2.2 nous nous intéresserons à l'agrégation de critères pour sélectionner une alternative en particulier. Enfin, dans la Sous-Section 1.2.3 nous verrons comment décider dans un contexte incertain.

1.2.1 Dominance de Pareto

La notion d'optimalité en décision multicritère est subjective et dépend des préférences du décideur. Néanmoins, il existe des règles de dominance qui permettent

de présélectionner des alternatives dans \mathbb{X} qui sont intéressantes quel que soit le décideur. Nous étudions ici la dominance de Pareto, assez naturelle pour un décideur supposé rationnel.

L'idée est de rejeter une alternative x s'il existe une alternative y qui est « meilleure » sur tous les critères simultanément. En effet, x n'est pas intéressante (pour un décideur rationnel), car il est possible d'avoir une alternative y qui « améliore » des critères sans en sacrifier. En fonction de ce que nous entendons par « meilleure », la Pareto-dominance est traduite par les trois relations suivantes :

Définition 1. *Pareto-dominance faible*

Une alternative $x \in \mathbb{X}$ est dominée faiblement au sens de Pareto par une alternative $y \in \mathbb{X}$, noté $y \succsim_P x$, si et seulement si $\forall j \in \mathcal{Q}, y_j \geq x_j$.

Définition 2. *Pareto-dominance forte*

Une alternative $x \in \mathbb{X}$ est dominée fortement au sens de Pareto par une alternative $y \in \mathbb{X}$, noté $y \succ_P x$, si et seulement si $\forall j \in \mathcal{Q}, y_j \geq x_j$ et $\exists j \in \mathcal{Q}$ tel que $y_j > x_j$.

Définition 3. *Pareto-dominance stricte*

Une alternative $x \in \mathbb{X}$ est dominée strictement au sens de Pareto par une alternative $y \in \mathbb{X}$, noté $y \gg_P x$, si et seulement si $\forall j \in \mathcal{Q}, y_j > x_j$.

La Pareto-dominance forte nécessite une amélioration stricte sur au moins un des critères contrairement à la Pareto-dominance faible. La Pareto-dominance stricte nécessite une amélioration stricte sur tous les critères. Notons que ces trois relations de dominance sont liées entre elles : Pareto-dominance stricte \Rightarrow Pareto-dominance \Rightarrow Pareto-dominance faible (la réciproque étant fausse).

Dans la suite de ce manuscrit, nous allons plus particulièrement nous intéresser à la Pareto-dominance forte. Nous sélectionnons alors les alternatives qui ne sont pas fortement Pareto-dominées :

Définition 4. *Pareto-optimalité*

L'ensemble $ND_P(\mathbb{X})$ des alternatives non dominées au sens de la dominance forte de Pareto est défini par $ND_P(\mathbb{X}) = \{x \in \mathbb{X} \mid \nexists y \in \mathbb{X} \text{ t.q. } y \succ_P x\}$.

Exemple 2. *Front de Pareto et dominance*

Nous reprenons les alternatives présentées dans l'Exemple 1 en rajoutant trois nouvelles alternatives : le fromage blanc avec $FB = (2, 6)$, le roquefort avec $RO = (7, 2)$ et l'édam avec $ED = (4, 4)$. Sur la Figure 1.1 est représentée le front de Pareto. Les 5 alternatives présentes dans la Table 1.1 sont Pareto-optimales et sont donc sur le front de Pareto. Les alternatives FB et RO sont fortement Pareto-dominées : $u_{Goût}(FB) < u_{Goût}(EM)$ et $u_{1/Prix}(RO) < u_{1/Prix}(GO)$. Néanmoins, elles ne sont pas strictement Pareto-dominées, au contraire d' ED : $u_j(ED) < u_j(MO) \forall j \in \{1, 2\}$. Nous avons alors $ND_P(\mathbb{X}) = \{CA, EM, MO, GO, BT\}$.

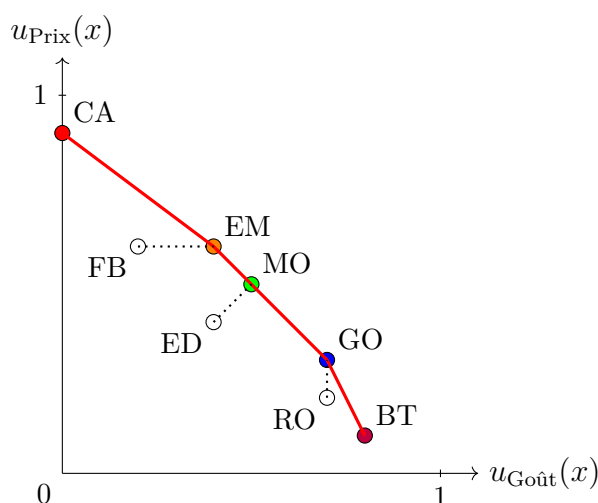


FIGURE 1.1 – Exemple d'un front de Pareto et d'une alternative dominée

Comme nous avons pu le voir dans l'Exemple 2, la dominance de Pareto permet d'écartier des alternatives forcément inintéressantes quel que soit le décideur, mais ne permet que rarement de choisir une alternative, car aucun ordre total n'est fourni sur \mathbb{X} . Ainsi, $ND_P(\mathbb{X})$ contient sûrement beaucoup d'alternatives, dont la plupart ne seront pas intéressantes pour le décideur, dépendant de ses préférences subjectives. Nous avons donc besoin d'établir un ordre entre les alternatives, et pour cela une technique classique est l'agrégation multicritère, qui permet d'évaluer chaque alternative.

1.2.2 Agrégation de préférences

L'agrégation multicritère est une opération de synthèse entre des informations de même type. Elle consiste principalement à réunir les différentes performances des critères avec une fonction d'agrégation $f : \mathbb{R}^q \rightarrow \mathbb{R}$ pour obtenir une nouvelle relation de préférence \succeq entre les différentes alternatives. $x \succeq y$ signifie que le décideur préfère l'alternative x à l'alternative y . Dans la littérature, il existe deux approches majeures (Grabisch et Perny, 2003) :

« **Agréger puis comparer** » (AC) : une évaluation globale est attribuée à chaque alternative avec une fonction d'agrégation $f : \mathbb{R}^q \rightarrow \mathbb{R}$, puis les évaluations sont comparées. Le plus souvent nous avons $x \succeq y$ si et seulement si $f(x) \geq f(y)$.

C'est une approche classique dans la littérature, à la fois simple d'utilisation et de compréhension, basée sur la Théorie de la Valeur Multi-Attributs (MAVT) (Keeney et Raiffa, 1976). Elle est intéressante pour notre problématique de choix, car elle force un ordre total, par la transitivité et la comparabilité entre toutes les alternatives. Néanmoins, cette approche nécessite une information riche, avec plusieurs hypothèses :

- Les critères doivent être commensurables. Néanmoins, il est possible de faire en sorte que les critères soient commensurables (bien que cela peut être long et cognitivement difficile pour le décideur). Nous supposons ici avoir accès à des critères commensurables.
- Les interactions entre critères doivent être connues, indiquant que la famille de fonction d'agrégation doit être bien établie. Les critères ne sont pas toujours additifs (parfois super-additifs ou sous-additifs).
- Les alternatives doivent être comparables en réalité.

« **Comparer puis agréger** » (CA) : tout d'abord x et y sont comparées critère par critère à l'aide des fonctions $\Phi_j : \mathbb{R}^2 \rightarrow \mathbb{R}$, puis les indices de performance $\Phi_j(x_j, y_j)$ sont agrégés avec une fonction d'agrégation $f : \mathbb{R}^q \rightarrow \mathbb{R}$ pour déduire si $x \succeq y$.

Si nous avons peu d'information sur les préférences du décideur, l'approche « Comparer puis agréger » est alors intéressante, car elle ne nécessite pas la com-

mesurabilité des critères. Dans cette approche, nous retrouvons les méthodes de surclassements, dont notamment les méthodes ELECTRE, PROMETEE ou ORESTE (Roy, 1985; Roy et Bouyssou, 1993; Maystre *et al.*, 1994). La relation de surclassement, que nous ne présenterons pas ici, à néanmoins plusieurs inconvénients :

- Elle n'est pas nécessairement complète.
- Elle n'est pas nécessairement transitive.
- En général, elle ne permet pas un rangement total des alternatives.

Avoir un ordre total entre les alternatives est indispensable pour faire un rangement ou même un choix. La méthode « Agréger puis comparer », malgré ses hypothèses assez fortes, est donc notre choix ici. Nous supposons avoir accès à une information suffisamment riche, notamment sur les interactions entre critères.

Propriétés des fonctions d'agrégation en approche « Agréger puis comparer »

Pour un inventaire détaillé sur l'agrégation multicritère, nous renvoyons à Grabisch *et al.* (2009). Nous évoquons ici quelques propriétés désirables pour ces fonctions d'agrégation de critères :

- *Idempotence* : une généralisation de l'unanimité. f est idempotente si et seulement si $f(a, \dots, a) = a$.
- *Monotonie* : $f : \mathbb{R}^q \rightarrow \mathbb{R}$ est monotone si et seulement si elle est croissante par rapport à chaque critère :

$$x_j \geq x'_j \Rightarrow f(x_1, \dots, x_j, \dots, x_q) \geq f(x_1, \dots, x'_j, \dots, x_q).$$

Notons que la fonction peut aussi être strictement monotone. Dans l'approche AC, la monotonie revient à assurer la Pareto optimalité.

- *Compromis* : $f : \mathbb{R}^q \rightarrow \mathbb{R}$ est une fonction de compromis si et seulement si l'évaluation globale des critères est comprise entre la conjonction et la disjonction des critères. En supposant que min soit l'opérateur de conjonction et que max soit l'opérateur de disjonction, nous avons :

$$\min_{j \in \mathcal{Q}} x_j \leq f(x_1, \dots, x_q) \leq \max_{j \in \mathcal{Q}} x_j.$$

D'autres propriétés sont plus techniques, par exemple la stabilité pour le changement d'échelle linéaire, assurant que changer d'échelle n'influencera pas l'ordre final des préférences, la continuité ou encore l'associativité. Notons que d'autres propriétés sont beaucoup moins désirables : la symétrie supposerait que nous accorderions la même importance aux différents critères, ce qui est rarement le cas (on aurait par exemple $f(x_1, x_2) = f(x_2, x_1)$). Notons enfin que certaines propriétés sont dépendantes entre elles (Marichal, 1998; Grabisch *et al.*, 2009) :

- Le compromis implique l'idempotence.
- L'idempotence et la monotonie impliquent le compromis.
- La stabilité pour le changement d'échelle linéaire implique l'idempotence.

Nous renvoyons à (Marichal, 1998; Grabisch *et al.*, 2009) pour les preuves.

Nous allons maintenant voir quelques opérateurs d'agrégation qui sont très utilisés dans la littérature. Notons que nous ne parlons pas des opérateurs de conjonction et de disjonction pour le moment, car ils ne sont pas très utilisés dans le cadre de la décision multicritère. Néanmoins, nous les présentons dans la Sous-Section 1.4.3, dans le cadre de la fusion d'information.

Somme pondérée (WS)

La somme pondérée, aussi connue sous le nom de moyenne arithmétique pondérée, est l'un des opérateurs d'agrégation les plus utilisés de fait de sa simplicité d'utilisation et de sa compréhension intuitive.

Définition 5. Somme pondérée

Soit un vecteur de pondération $\omega \in [0, 1]^q$ avec $\sum_{i=1}^q \omega_j = 1$. La somme pondérée d'une alternative x est définie par :

$$WS_\omega(x) = \sum_{j \in \mathcal{Q}} \omega_j x_j. \quad (1.1)$$

Remarquons que si les poids sont égaux, nous avons une moyenne arithmétique classique. Une somme pondérée définit la relation de préférence suivante :

$$x \succeq y \Leftrightarrow WS_\omega(x) \geq WS_\omega(y). \quad (1.2)$$

Exemple 3. Somme pondérée

Nous reprenons le sous-ensemble d'alternatives disponibles \mathbb{X} de l'Exemple 1 et nous supposons que nous agrégeons les critères des différentes alternatives selon une somme pondérée ayant pour poids $\omega = (0.7, 0.3)$, préférant donc les alternatives ayant du goût. Nous représentons sur la Figure 1.2 l'évolution du score de chaque alternative dans Ω , et nous notons sur la Table 1.2 les scores de chaque alternative pour $\omega = (0.7, 0.3)$. Nous pouvons alors établir le rangement $BT \succ GO \succ MO \succ EM \succ CA$ et donc suggérer BT comme étant le meilleur choix d'après cette fonction d'agrégation.

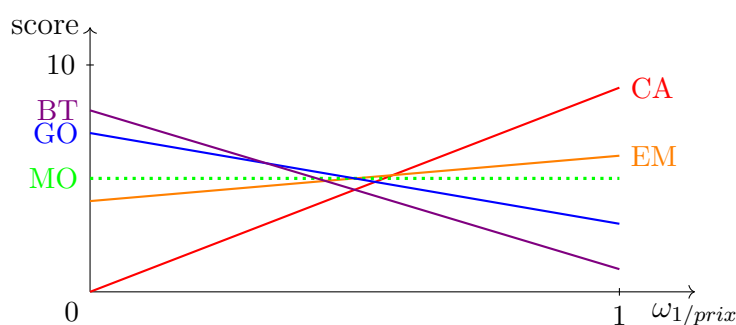


FIGURE 1.2 – Score de \mathbb{X} pour une somme pondérée en fonction de $\omega \in \Omega$

Nom	Goût	1/Prix	Score
CA	0	9	2.7
EM	4	6	4.6
MO	5	5	5
GO	7	3	5.8
BT	8	1	5.9

TABLE 1.2 – Score de \mathbb{X} pour une somme pondérée (0.7, 0.3)

D'autres types de moyennes existent, ainsi que leur version pondérée, que nous pouvons retrouver à partir de la définition de la moyenne quasi-arithmétique :

$$M_f(x_1, \dots, x_q) = f^{-1} \left(\sum_{j \in \mathcal{Q}} \omega_j f(x_j) \right), \quad (1.3)$$

où f est une fonction continue et injective. Nous retrouvons la moyenne géométrique ($f(x) = \ln(x)$) ou encore la moyenne harmonique ($f(x) = \frac{1}{x}$). Néanmoins, si toutes les moyennes sont monotones (stricte), idempotentes et de compromis, seule la moyenne arithmétique est stable au changement d'échelle linéaire.

Remarquons que la moyenne pondérée a tendance à écarter certaines solutions Pareto-optimales de compromis : en regardant l'Exemple 3, l'alternative MO n'est jamais strictement la meilleure alternative, au mieux égale avec d'autres. Pourtant, cette alternative est un compromis entre goût et prix qui peut intéresser un décideur. Cette limitation justifie la recherche vers d'autres opérateurs d'agrégation, notamment la somme pondérée ordonnée, qui attribue des poids sur les rangs.

Somme pondérée ordonnée (OWA)

La somme pondérée ordonnée, aussi connu sous le nom d'OWA (Ordered Weighted Average) (Yager, 1988), est un opérateur qui attribue un poids au rang de chaque performance au lieu de chaque critère comme le ferait la moyenne pondérée.

Définition 6. Soit un vecteur de pondération $\omega \in [0, 1]^q$ avec $\sum_{i=1}^q \omega_j = 1$. Notons aussi (j) la j^e performance la plus faible¹. La somme pondérée ordonnée d'une alternative x est définie par :

$$OWA_{\omega}(x) = \sum_{j \in \mathcal{Q}} \omega_j x_{(j)}. \quad (1.4)$$

Autrement dit, nous avons $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(q)}$ et nous attribuons les poids ω_i en fonction de ce rangement. Une OWA définit la relation de préférence suivante :

$$x \succeq y \Leftrightarrow OWA_{\omega}(x) \geq OWA_{\omega}(y). \quad (1.5)$$

Remarquons que l'OWA généralise plusieurs opérateurs traditionnels :

- $\min(x)$ si $\omega_1 = 1$ et $\omega_j = 0$, $\forall j \in \{2, \dots, q\}$.
- $\max(x)$ si $\omega_q = 1$ et $\omega_j = 0$, $\forall j \in \{1, \dots, q-1\}$.
- Plus généralement, le quantile ϵ si $\omega_{\epsilon} = 1$ et $\omega_j = 0$, $\forall j \in \mathcal{Q} \setminus \{\epsilon\}$.
- Moyenne arithmétique si $\omega_j = 1/q$, $\forall j \in \mathcal{Q}$.

1. Dans certaines conventions, c'est la j^e plus forte.

Si l'OWA généralise la moyenne arithmétique, elle ne généralise pas la moyenne arithmétique pondérée : il n'est pas toujours possible de trouver une pondération de rangs qui donne le même résultat qu'une pondération de critères.

Exemple 4. Somme pondérée ordonnée

Toujours avec les mêmes alternatives, nous agrégeons les critères selon une somme pondérée ordonnée ayant pour poids $\omega = (0.55, 0.45)$. Cela indique que la plus petite performance à un poids de 0.55, et la plus grande un poids de 0.45². Nous représentons sur la Figure 1.3 l'évolution du score de chaque alternative dans Ω , et nous notons sur la Table 1.3 les scores de chaque alternative pour $\omega = (0.55, 0.45)$. Nous pouvons établir alors le rangement $MO \succ EM \succ GO \succ BT \succ CA$ et donc suggérer MO comme étant le meilleur choix d'après cette fonction d'agrégation. Nous voyons sur la Figure 1.3 que MO peut être strictement le meilleur choix (avec une moyenne pondérée ce ne serait jamais le cas).

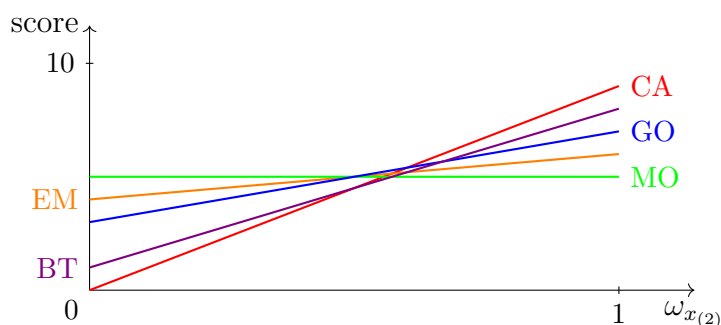


FIGURE 1.3 – Score de \mathbb{X} pour une somme pondérée ordonnée en fonction de $\omega \in \Omega$

Les opérateurs OWA sont idempotents, monotones, de compromis et stables pour le changement d'échelle linéaire. Par contre, ils ont aussi l'inconvénient d'être symétriques : par exemple, $OWA_{\omega}(10, 8) = OWA_{\omega}(8, 10)$, car les poids sont attribués selon l'ordre décroissant des performances. Or, cela modéliserait que les critères ont la même importance pour le décideur, ce qui n'est généralement pas le cas. Dans la littérature, il existe des extensions de l'opérateur OWA, par exemple la moyenne ordonnée doublement pondérée (WOWA) (Torra, 1997), paramétrée

2. $\omega = (1, 0)$ favorisera les alternatives équilibrées, $\omega = (0, 1)$ les alternatives maximisant un critère, et donc potentiellement déséquilibrées.

Nom	Goût	1/Prix	$0.55x_{(1)}$	$0.45x_{(2)}$	Score
CA	0	9	0	4.05	4.05
EM	4	6	2.2	2.7	4.9
MO	5	5	2.75	2.25	5
GO	7	3	1.65	3.15	4.8
BT	8	1	0.55	3.6	4.15

TABLE 1.3 – Score de \mathbb{X} pour une somme pondérée ordonnée $(0.55, 0.45)$

par deux vecteurs : les poids des rangs et le poids des critères. Elle permet de généraliser à la fois l'OWA et la somme pondérée. Nous allons présenter un autre opérateur généralisant la somme pondérée, l'OWA et la WOWA, permettant de mieux prendre en compte les interactions entre critères : l'intégrale de Choquet.

Intégrale de Choquet

L'intégrale de Choquet (Choquet, 1953; Schmeidler, 1986; Grabisch, 1996) est un opérateur qui associe un poids à chaque sous-ensemble de critères, plutôt qu'à chaque critère ou chaque rang. Elle est paramétrée par une fonction d'ensemble v appelée capacité qui associe à chaque sous-ensemble de critères un poids.

Définition 7. Capacité (normalisée)

La fonction $v : 2^{\mathcal{Q}} \rightarrow [0, 1]$ est une capacité normalisée si et seulement si :

- $v(\emptyset) = 0$ et $v(\mathcal{Q}) = 1$ (normalisation),
- $v(A) \leq v(B) \forall A \subseteq B \subseteq \mathcal{Q}$ (monotonie).

Définition 8. Intégrale de Choquet

Soit une capacité normalisée v . Notons aussi (j) la j^e performance la plus faible et $X_{(j)} = \{k \mid u_k(x) \geq u_{(j)}(x), \forall k \in \mathcal{Q}\}$, l'ensemble des critères mieux évalués que le j^e plus faible. L'intégrale de Choquet d'une alternative x est définie par :

$$C_v(x) = \sum_{j \in \mathcal{Q}} [x_{(j)} - x_{(j-1)}] v(X_{(j)}), \quad (1.6)$$

$$= \sum_{j \in \mathcal{Q}} [v(X_{(j)}) - v(X_{(j+1)})] x_{(j)}. \quad (1.7)$$

avec $x_{(0)} = 0$ et $X_{(q+1)} = \emptyset$.

L'intégrale de Choquet définit la relation de préférence suivante :

$$x \succeq y \Leftrightarrow \sum_{j \in \mathcal{Q}} [x_{(j)} - x_{(j-1)}] v(X_{(j)}) \geq \sum_{j \in \mathcal{Q}} [y_{(j)} - y_{(j-1)}] v(Y_{(j)}). \quad (1.8)$$

L'intégrale de Choquet permet de généraliser plusieurs opérateurs vu dans les précédents paragraphes :

- Somme pondérée : capacité v additive par rapport aux poids ω_j de la somme, c'est-à-dire $\forall A \subseteq \mathcal{Q}, v(A) = \sum_{j \in A} \omega_j$.
- OWA : v symétrique, c'est-à-dire $v(A) = v(B), \forall A, B \subseteq \mathcal{Q}$ tels que $|A| = |B|$. $u_{(i)}(x)$ dépend de la taille de l'ensemble $X_{(i)}$ et donc de l'indice i .
- WOWA (Torra, 1997) : v est une capacité telle que $\forall A \subseteq \mathcal{Q}, v(A) = \phi\left(\sum_{j \in A} \omega_j\right)$, ϕ une fonction croissante avec $\phi(0) = 0$ et $\phi(1) = 1$.

De plus, l'intégrale de Choquet, au travers du choix de la capacité v , permet de modéliser d'autres comportements du décideur. Lorsque la capacité est convexe :

$$\forall A, B \subseteq \mathcal{Q}, v(A \cup B) \geq v(A) + v(B) - v(A \cap B), \quad (1.9)$$

alors l'intégrale de Choquet privilégie les solutions ayant des performances équilibrées (Chateaufort et Tallon, 2002; Galand *et al.*, 2010). La capacité permet aussi de modéliser des interactions entre critères (Grabisch et Labreuche, 2010) :

- $v(\{i, j\}) > v(\{i\}) + v(\{j\})$: interaction positive entre i et j (coopération entre les critères).
- $v(\{i, j\}) < v(\{i\}) + v(\{j\})$: interaction négative entre i et j (conflit entre les critères).
- $v(\{i, j\}) = v(\{i\}) + v(\{j\})$: aucune interaction entre i et j .

Exemple 5. *Intégrale de Choquet*

Nous reprenons le sous-ensemble des alternatives disponibles \mathbb{X} de l'Exemple 1, à la différence que les alternatives ont cette fois-ci un troisième critère : la facilité de cuisiner avec. Les alternatives mises-à-jour sont présentées dans la Table 1.4. Pour déterminer le score de chaque alternative, nous utilisons cette

fois-ci une intégrale de Choquet dont la capacité v est indiquée dans la Table 1.5. En étudiant les interactions entre les critères, nous pouvons voir une interaction positive entre le goût et le prix ($v(\{1, 2\}) > v(\{1\}) + v(\{2\})$), semblant indiquer que le décideur apprécie particulièrement un fromage peu cher et goûteux. Par contre, il n'y a pas d'interactions entre les autres critères ($v(\{1, 3\}) = v(\{1\}) + v(\{3\})$ et $v(\{2, 3\}) = v(\{2\}) + v(\{3\})$).

Nom	Goût	1/Prix	Cuisine	Score
CA	0	9	8	3.4
EM	4	6	6	4.9
MO	5	5	7	5.2
GO	7	3	2	3.8
BT	8	1	0	2.7

TABLE 1.4 – Score de \mathbb{X} pour une intégrale de Choquet de capacité v

X	\emptyset	$\{1\}$	$\{2\}$	$\{3\}$	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$	\mathcal{Q}
$v(X)$	0	0.3	0.2	0.1	0.6	0.4	0.3	1

TABLE 1.5 – Capacité v

Le calcul des scores présentés dans la Table 5 est le suivant :

- $C_v(CA) = 0 + [8 - 0]v(\{2, 3\}) + [9 - 8] \times v(\{2\}) = 8 \times 0.3 + 0.2 = 2.6,$
- $C_v(EM) = 4 + [6 - 4]v(\{2, 3\}) + [6 - 6] \times v(\{2, 3\}) = 4 + 2 \times 0.3 = 4.6,$
- $C_v(MO) = 5 + [5 - 5]v(\{1, 2, 3\}) + [7 - 5] \times v(\{3\}) = 5 + 2 \times 0.1 = 5.2,$
- $C_v(GO) = 2 + [3 - 2]v(\{1, 2\}) + [7 - 3] \times v(\{1\}) = 2 + 0.6 + 4 \times 0.3 = 3.8,$
- $C_v(BT) = 0 + [1 - 0]v(\{1, 2\}) + [8 - 1] \times v(\{1\}) = 1 \times 0.6 + 7 \times 0.3 = 2.7.$

Nous avons ainsi $MO \succ EM \succ GO \succ BT \succ CA$. La meilleure alternative selon C_v est donc MO .

Une autre formulation de cet opérateur est basée sur l'inverse de Möbius de la capacité (voir par exemple Rota (1964)) :

Définition 9. *Inverse de Möbius*

L'inverse de Möbius $m : 2^{\mathcal{Q}} \rightarrow \mathbb{R}$ associée à une fonction d'ensemble $v : 2^{\mathcal{Q}} \rightarrow \mathbb{R}$ est définie par :

$$\forall A \subseteq \mathcal{Q}, m(A) = \sum_{A' \subseteq A} (-1)^{|A \setminus A'|} v(A'), \quad (1.10)$$

où les valeurs $m(A)$, $\forall A \subseteq \mathcal{Q}$ sont les masses de Möbius.

Il est aussi possible de retrouver la capacité à partir des masses de Möbius grâce à la transformation Zêta :

$$\forall A \subseteq \mathcal{Q}, v(A) = \sum_{A' \subseteq A} m(A'). \quad (1.11)$$

Il existe une équivalence entre une capacité normalisée et son inverse de Möbius :

Proposition 1. *Lien entre capacité et inverse de Möbius*

v est une capacité normalisée si et seulement si son inverse de Möbius satisfait :

- $m(\emptyset) = 0$,
- $\sum_{A \subseteq \mathcal{Q}} m(A) = 1$,
- $\sum_{A' \subseteq A} m(A') \geq 0 \forall A \in \mathcal{Q}$ (monotonie).

La preuve de cette proposition est donnée par Chateauneuf et Jaffray (1989). Il est possible de trouver plus d'informations sur le lien entre inverse de Möbius et capacité, notamment quand les capacités sont convexes, dans les travaux de Chateauneuf et Jaffray (1995).

Il est aussi possible d'écrire l'intégrale de Choquet d'une alternative x en fonction des masses de Möbius (Chateauneuf et Jaffray, 1989) :

$$C_v(x) = \sum_{A \subseteq \mathcal{Q}} m(A) \min_{j \in A} x_j. \quad (1.12)$$

Nous pouvons alors remarquer que l'intégrale de Choquet est une somme pondérée dans un espace à 2^q dimensions dont les poids sont les masses de Möbius. Notons la différence avec l'opérateur somme pondérée : les poids peuvent être négatifs, même si $\sum_{A \subseteq \mathcal{Q}} m(A) = 1$.

Exemple 6. *Inverse de Möbius*

Nous reprenons la capacité présentée dans la Table 1.5 et nous cherchons à déterminer son inverse de Möbius. Voici quelques calculs des masses de Möbius présentées dans la Table 1.6 :

- $m_v(\{1\}) = (-1)^{|\{\emptyset\}|} v(\{1\}) = 0.3,$
- $m_v(\{1, 2\}) = (-1)^{|\{2\}|} v(\{1\}) + (-1)^{|\{1\}|} v(\{2\}) + (-1)^{|\{\emptyset\}|} v(\{1, 2\})$
 $= -0.3 - 0.2 + 0.6 = 0.1,$
- $m_v(\{\mathcal{Q}\}) = (-1)^{|\{2,3\}|} v(\{1\}) + (-1)^{|\{1,3\}|} v(\{2\}) + (-1)^{|\{1,2\}|} v(\{3\}) +$
 $(-1)^{|\{3\}|} v(\{1, 2\}) + (-1)^{|\{2\}|} v(\{1, 3\}) + (-1)^{|\{1\}|} v(\{2, 3\}) + (-1)^{|\{\emptyset\}|} v(\{\mathcal{Q}\})$
 $= 0.3 + 0.2 + 0.1 - 0.6 - 0.4 - 0.3 + 1 = 0.3.$

A	\emptyset	$\{1\}$	$\{2\}$	$\{3\}$	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$	\mathcal{Q}
$v(A)$	0	0.3	0.2	0.1	0.6	0.4	0.3	1
$m_v(A)$	0	0.3	0.2	0.1	0.1	0	0	0.3

TABLE 1.6 – Inverse de Möbius de v

Il est possible de réduire la dimension de la somme en considérant des capacités k -additives (Grabisch, 1996) :

Définition 10. *Capacité k -additive*

Une capacité v est dite k -additive si son inverse de Möbius m respecte les propriétés suivantes :

- $m_v(A) = 0 \forall A \subseteq \mathcal{Q}$ avec $|A| > k,$
- $m_v(A) \neq 0$ pour au moins un $A \subseteq \mathcal{Q}$ avec $|A| = k.$

Remarquons que si $k = 1$ nous avons trivialement une capacité additive, et plus particulièrement une somme pondérée classique. Avec une capacité k -additive, nous réduisons le nombre de paramètres à $\binom{k}{n}$, ce qui est intéressant quand k est petit. Nous perdons néanmoins une partie de la capacité de l'intégrale de Choquet à modéliser des interactions entre des groupes de critères.

Exemple 7. *Capacité 2-additive*

Partons d'une inverse de Möbius pour déterminer une capacité 2-additive. Soit une inverse de Möbius présentée dans la Table 1.7. Il est possible de vérifier que cette inverse de Möbius correspond bien à une capacité normalisée ($m(\emptyset) = 0, \sum_{A \subseteq \mathcal{Q}} m(A) = 1$ et la monotonie est bien respectée), ainsi qu'à une capacité 2-additive ($m(\mathcal{Q}) = 0$ et $\exists A$ tel que $m(A) \neq 0$ et $|A| = 2$). Nous pouvons alors déterminer la capacité 2-additive correspondante à partir de la transformation Zêta dont voici quelques calculs :

- $v(\{1\}) = m_v(\emptyset) + m_v(\{1\}) = 0.2,$
- $v(\{1, 3\}) = m_v(\emptyset) + m_v(\{1\}) + m_v(\{3\}) + m_v(\{1, 3\}) = 0.5,$
- $v(\{\mathcal{Q}\}) = \sum_{A \subseteq \mathcal{Q}} = 1.$

A	\emptyset	$\{1\}$	$\{2\}$	$\{3\}$	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$	\mathcal{Q}
$m_v(A)$	0	0.2	0.1	0.5	0.4	-0.2	0	0
$v(A)$	0	0.2	0.1	0.5	0.7	0.5	0.6	1

TABLE 1.7 – Capacité 2-additive v et son inverse de Möbius

L'intégrale de Choquet est un opérateur idempotent, monotone, de compromis et stable pour le changement d'échelle linéaire. C'est un opérateur à la fois puissant pour modéliser des comportements complexes, mais aussi intéressant mathématiquement. Nous vous invitons à vous référer aux travaux de Schmeidler (1986); Marichal (2000); Köbberling et Wakker (2003); Labreuche (2012) pour approfondir cet outil. Néanmoins, c'est une fonction d'agrégation avec un grand nombre de paramètres (2^q pour une capacité complète), ce qui limite fortement son utilisation. Notons que plusieurs travaux récents se portent sur l'apprentissage de capacités pour les intégrales de Choquet (Bresson *et al.*, 2021; Herin *et al.*, 2022). Bien que déjà très coûteuse, il existe des modèles plus puissants que l'intégrale de Choquet, par exemple l'intégrale de Choquet à bi-capacité (Grabisch et Labreuche, 2005).

Dans la Sous-Section suivante, nous allons voir comment prendre une décision sous incertitude totale, quand par exemple le score associé à une alternative dépend d'évènements dont les vraisemblances ne sont pas connues.

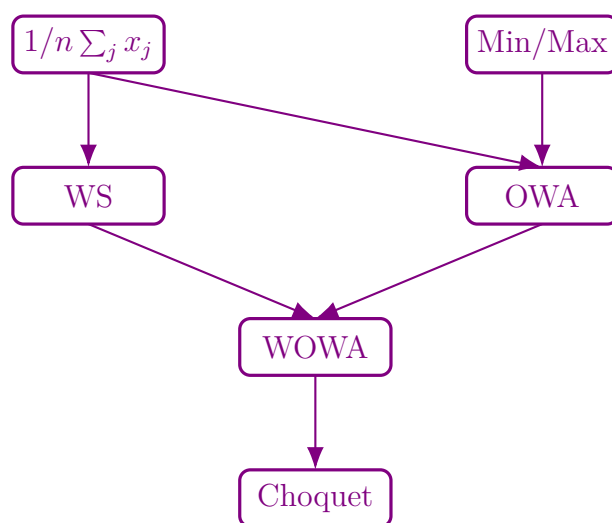


FIGURE 1.4 – Lien entre les opérateurs d'agrégation

1.2.3 Décision sous incertitude ensembliste

Admettons que nous voulions toujours faire un choix entre plusieurs alternatives $x \in \mathbb{X}$ et que chaque alternative a un score associé $f_\omega(x)$, où f est une fonction d'agrégation de critères et ω son paramètre comme vu dans la Sous-Section 1.2.2. Dans cette Sous-section, nous supposons connaître la famille de la fonction d'agrégation f_ω (WS, OWA, Choquet...), mais pas exactement la valeur du paramètre ω : nous n'avons qu'un ensemble Ω de valeurs que peut prendre ω . Nous souhaitons alors prendre une décision dans l'incertain. Deux des principales branches de la décision dans l'incertain sont la décision sous risque (Wu *et al.*, 2004), où l'incertitude est représentée par une distribution de probabilité sur les événements, et la décision sous incertitude totale lorsque nous n'avons aucune information sur la vraisemblance des événements. Nous vous invitons à vous reporter par exemple au livre de Gilboa (2009) pour une introduction à la décision dans l'incertain.

Dans la suite de cette sous-section, nous allons nous intéresser à la décision sous incertitude totale, et plus particulièrement aux approches permettant d'obtenir un rangement complet des alternatives (et donc de faire un choix optimal pour le décideur) en supposant que le décideur prenne une décision selon un comportement particulier (optimiste, pessimiste, entre deux, etc). Néanmoins, il existe aussi des approches permettant de ne pas faire une supposition sur le comportement du

décideur, ou de juste connaître les alternatives potentiellement optimales (Troffaes, 2007)). Ces dernières approches, que nous n'évoquerons pas, ne permettent pas toujours d'avoir un ordre total et donc de faire un choix (contraire à notre objectif).

Maximax : comportement optimiste

L'approche Maximax privilégie l'alternative avec le score maximal dans le plus favorable des scénarios, soit une approche optimiste qui ne se soucie pas des scénarios les plus défavorables.

Définition 11. *Maximax*

$\bar{f}(x)$ est le score d'une alternative x selon le critère Maximax tel que :

$$\bar{f}(x) = \max_{\omega \in \Omega} f_{\omega}(x). \tag{1.13}$$

Nous avons $x \succ_{MM} y$ si $\bar{f}(x) > \bar{f}(y)$.

Exemple 8. *Maximax*

Nous reprenons les alternatives de l'Exemple 1. Nous supposons que f_{ω} est une moyenne pondérée dont nous ne connaissons pas le paramètre ω . Nous avons l'ensemble des valeurs possibles de ω noté $\Omega = \{\omega | 0 \leq \omega_i \leq 1 \text{ et } \sum \omega_i = 1, \forall i \in \{1, 2\}\}$. Nous cherchons à déterminer la meilleure alternative selon le critère Maximax. Sur la Figure 1.5 est représenté le score de chaque alternative en fonction de $\omega \in \Omega$, avec une mise en avant du score maximal de chaque alternative. Par exemple, l'alternative BT a un score maximal de 8 atteint pour $\omega = (1, 0)$.

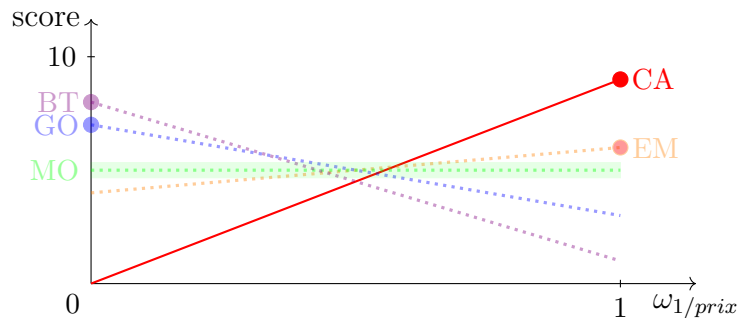


FIGURE 1.5 – Illustration de l'approche Maximax

Nous avons $CA \succ_{MM} BT \succ_{MM} GO \succ_{MM} EM \succ_{MM} MO$, l'alternative CA est donc recommandée selon le critère *Maximax*. Remarquons que l'alternative CA est la meilleure dès que le prix est important pour le décideur, mais dès que le goût devient important, elle devient la pire. L'approche *Maximax* (optimiste) peut donc donner une très mauvaise recommandation dans le pire cas.

Maximin : comportement pessimiste

L'approche Maximin privilégie l'alternative avec le score maximal dans le plus défavorable des scénarios, soit une approche pessimiste qui ne s'intéresse pas aux scénarios plus favorables.

Définition 12. Maximin

$\underline{f}(x)$ est le score d'une alternative x selon le critère Maximin tel que :

$$\underline{f}(x) = \min_{\omega \in \Omega} f_{\omega}(x). \quad (1.14)$$

Nous avons $x \succ_{Mm} y$ si $\underline{f}(x) > \underline{f}(y)$.

Exemple 9. Maximin

En partant du même contexte que l'Exemple 8, nous cherchons à déterminer la meilleure alternative selon le critère Maximin. Sur la Figure 1.6 est représenté le score de chaque alternative en fonction de $\omega \in \Omega$, avec une mise en avant du score minimal de chaque alternative. Par exemple, l'alternative GO a un score minimal de 3 atteint pour $\omega = (0, 1)$.

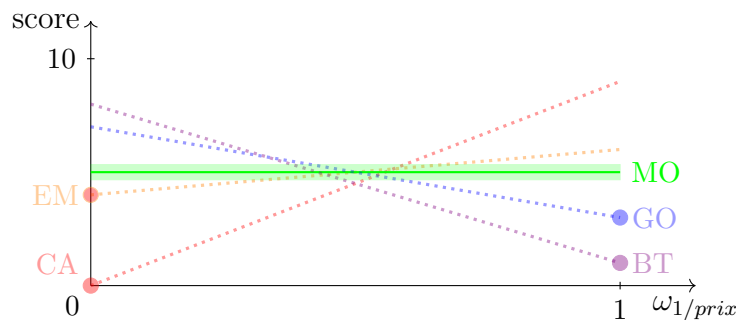


FIGURE 1.6 – Illustration de l'approche Maximin

Nous avons $MO \succ_{Mm} EM \succ_{Mm} GO \succ_{Mm} BT \succ_{Mm} CA$, l'alternative MO est donc recommandée selon le critère Maximin. Remarquons que l'alternative MO est celle qui maximise le score dans le pire des cas, mais pour toute valeur de ω , il existe une alternative qui a au moins un aussi bon score que MO (et souvent strictement mieux).

Hurwicz : comportement intermédiaire

L'approche Hurwicz (Arrow et Hurwicz, 1972) est un intermédiaire entre l'approche Maximax et l'approche Maximin.

Définition 13. *Hurwicz*

$f_{H(\alpha)}(x)$ est le score d'une alternative x selon le critère d'Hurwicz tel que :

$$f_{H(\alpha)}(x) = \alpha \bar{f}(x) + (1 - \alpha) \underline{f}(x), \tag{1.15}$$

où $\alpha \in [0, 1]$ est un indice d'optimisme.

Quand $\alpha = 1$ nous avons l'approche Maximax, et quand $\alpha = 0$ nous avons l'approche Maximin. Nous avons $x \succ_{H(\alpha)} y$ si $f_{H(\alpha)}(x) > f_{H(\alpha)}(y)$.

Exemple 10. *Hurwicz*

En partant du même contexte que l'Exemple 8, nous cherchons à déterminer la meilleure alternative selon le critère Hurwicz, avec $\alpha = 0.55$. Sur la Table 1.8 est représenté le score minimal et maximal de chaque alternative, ainsi que le score selon le critère Hurwicz.

	$\bar{f}(\cdot)$	$\underline{f}(\cdot)$	$f_{H(0.55)}(\cdot)$
CA	9	0	4.95
EM	6	4	5.1
MO	5	5	5
GO	7	3	5.2
BT	8	1	4.85

TABLE 1.8 – Illustration de l'approche Hurwicz

Nous avons $GO \succ_{H(0.55)} EM \succ_{H(0.55)} MO \succ_{H(0.55)} CA \succ_{H(0.55)} BT$, l'alternative MO est donc recommandée selon le critère Hurwicz. Ici, nous avons choisi un α particulier, mais il n'est pas forcément évident de trouver le bon compromis entre pessimiste et optimiste.

Regret Maximax : comportement prudent

L'approche du regret Maximax de Savage (1951) est différente des autres approches, car elle s'intéresse à la notion de regret :

Définition 14. Regret

Le regret de choisir une alternative $x \in \mathbb{X}$ par rapport à une alternative $y \in \mathbb{X}$ est défini par :

$$R_\omega(x, y) = f_\omega(y) - f_\omega(x). \quad (1.16)$$

Le regret exprime combien en performance nous aurions gagné (ou perdu) si nous avions pris y au lieu de x . Si le regret est positif, cela indique que y aurait été un meilleur choix. Si le regret est négatif, x est effectivement le meilleur choix.

Définition 15. Regret maximal par paire

Soit l'ensemble des valeurs possibles de ω noté Ω . Le regret maximal par paire entre $x, y \in \mathbb{X}^2$ sachant Ω est défini par :

$$PMR(x, y, \Omega) = \max_{\omega \in \Omega} R_\omega(x, y). \quad (1.17)$$

Le regret maximal par paire exprime le regret maximal de choisir x à la place de y sachant toutes les valeurs possibles de paramètres $\omega \in \Omega$. Ainsi, un regret maximal élevé indique que dans un scénario au moins (le pire scénario), l'alternative x est très défavorable par rapport à y . Néanmoins, cet indicateur ne nous permet pas de savoir l'intérêt de x dans les autres scénarios.

Définition 16. Regret maximal

Le regret maximal de choisir $x \in \mathbb{X}$ sachant Ω est défini par :

$$MR(x, \Omega) = \max_{y \in \mathbb{X}} PMR(x, y, \Omega). \quad (1.18)$$

Ce regret maximal (par alternative) nous permet d'identifier les alternatives qui minimisent le regret perçu dans le pire des scénarios. Une nouvelle fois, il ne permet pas de savoir comment se comporte l'alternative dans des scénarios plus favorables (s'ils existent).

Définition 17. *Regret maximal minimal*

Le regret maximal minimal du sous-ensemble d'alternatives disponibles \mathbb{X} sachant Ω est défini par :

$$mMR(\Omega) = \min_{x \in \mathbb{X}} MR(x, \Omega). \tag{1.19}$$

Nous avons alors $x^* = \arg mMR(\Omega)$ l'alternative avec le regret minimal dans le pire scénario, correspond au choix optimal (en termes de regret).

Exemple 11. *Regret Minimax*

Nous partons du même contexte que l'Exemple 8, mais pour simplifier les calculs et rendre les illustrations plus explicites, nous allons nous focaliser que sur trois alternatives : CA, MO et BT. Nous commençons par le calcul du regret maximal par paire (PMR). Le PMR est le plus grand regret possible entre deux alternatives. Sur la Figure 1.7 est représenté le PMR entre MO et CA, ainsi qu'entre MO et BT. $\max_{\omega \in \Omega} R_{\omega}(MO, CA) = 4$, atteint pour $\omega = (0, 1)$, tandis que $\max_{\omega \in \Omega} R_{\omega}(MO, BT) = 3$, atteint pour $\omega = (1, 0)$. Remarquons que la fonction déterminant le regret n'est pas commutative, et qu'en général $PMR(MO, CA, \Omega) \neq PMR(CA, MO, \Omega)$ (ici égale à 5).

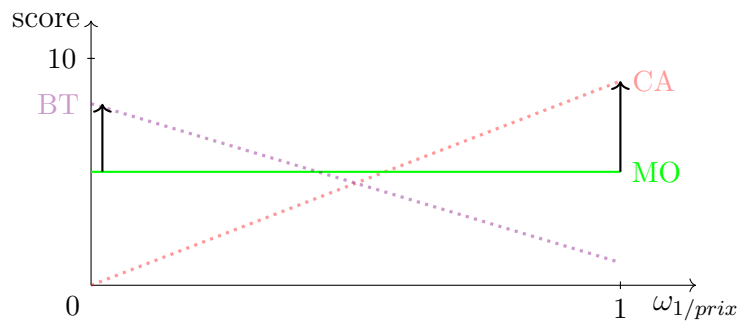


FIGURE 1.7 – Illustration de l'approche du regret Minimax : visualisation du PMR

Nous regardons ensuite le regret maximal pour une alternative donnée (MR). Le regret maximal de x est le plus grand regret maximal par paire entre x et n'importe

quelle alternative $y \in \mathbb{X}, y \neq x$. Déterminer le regret maximal de MO est facile ici, car nous avons déjà calculé le PMR entre MO et CA , et entre MO et BT . Nous avons alors $MR(MO) = PMR(MO, CA, \Omega) = 4$. Le calcul des autres MR donne $MR(CA) = 8$ et $MR(BT) = 8$. Sur la Figure 1.8 est représenté le MR de chaque alternative. Une alternative étant préférée à une autre selon le critère du regret Minimax si son regret maximal est inférieur, nous avons $MO \succ_{MR} CA =_{MR} BT$, l'alternative MO est donc recommandée selon le critère du regret Minimax.

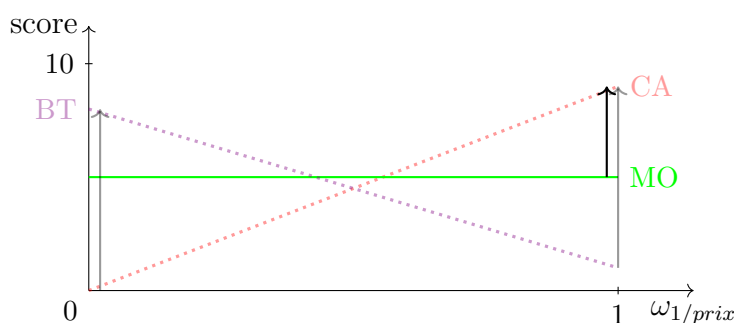


FIGURE 1.8 – Illustration de l'approche du regret Minimax : visualisation du MR

1.3 Élicitation de préférences

L'élicitation des préférences a pour objectif de construire un système d'aide à la décision pouvant proposer des recommandations pertinentes à un décideur, afin de l'aider à prendre une décision. L'idée est de récolter les informations préférentielles les plus utiles possibles, d'explicitier les préférences du décideur afin de les formaliser (pour pouvoir ensuite les exploiter). Pour cela, il est nécessaire d'interagir avec le décideur. Un système d'élicitation et de recommandation de qualité doit respecter plusieurs éléments :

- Il doit minimiser le nombre de questions posées au décideur.
- Chaque question doit aider à récolter un maximum d'information.
- L'effort cognitif pour le décideur doit être minimisé.
- La complexité et le temps de calcul du système doivent être minimisés.
- Éviter les erreurs.

Remarquons que ces objectifs sont pour la plupart liés : minimiser le nombre de questions permet de minimiser l'effort cognitif, le temps de calcul, ainsi que le risque de commettre des erreurs. Néanmoins, d'autres objectifs sont parfois difficiles à satisfaire ensemble : des questions trop simples ne parviennent pas toujours à obtenir une information préférentielle suffisante. Il est donc important, mais aussi difficile, de trouver des questions adéquates permettant de remplir le plus d'objectifs possibles. Les questions pour le décideur peuvent être de différents types, dont en voici des exemples :

- Demander à classer $k > 2$ alternatives selon ses préférences.
- Déterminer si le score associé à une réponse est supérieur à $t \in \mathbb{R}$.
- Si une solution x ne lui plaît pas, le décideur indique quels critères est-il prêt à sacrifier et à quel point.
- Comparer une paire d'alternatives x et y , puis indiquer laquelle est préférée ($x \succeq y$ ou $y \succeq x$).

Remarquons que les trois premières questions sont en général cognitivement difficiles pour le décideur, augmentant le risque de fatigue et d'erreurs, les erreurs ayant un impact potentiellement très négatif sur la qualité des recommandations finales. Le quatrième type de question est assez simple pour le décideur et souvent utilisé en élicitation de préférences. Ainsi, les comparaisons de paire sont simples, mais permettent pourtant de bien modéliser des comportements complexes (au prix d'un nombre de questions plus grand). En supposant que les préférences du décideur soient modélisées par un modèle particulier, l'information portée par la comparaison de paires est enrichie.

Exemple 12. *Comparaison de paires*

En prenant le sous-ensemble d'alternatives disponibles \mathbb{X} présentée dans la Table 1.1 et en supposant que nous ayons l'information $EM \succeq GO$, nous savons juste que GO n'est pas le choix optimal par rapport aux préférences du décideur et aux autres alternatives disponibles. Néanmoins, en supposant que les préférences du décideur soient modélisées par une somme pondérée WS_ω , nous pouvons déduire l'information préférentielle suivante : $4\omega_1 + 6\omega_2 \geq 7\omega_1 + 3\omega_2 \Rightarrow 3\omega_2 \geq 3\omega_1 \Rightarrow \omega_2 \geq \omega_1$. Nous pouvons en déduire les informations suivantes :

- Le score maximal de *BR* dans ce contexte est atteint pour $\omega = (0.5, 0.5)$ et il est de 4.5, or le score minimal d'*EM* est atteint pour $\omega = (0.5, 0.5)$ et il est de 5. *BR* n'est donc pas optimal.
- Le score de *CA* pouvant aller de 4.5 à 9 dans ce contexte et celui d'*EM* de 5 à 6, ces deux alternatives ne dominent pas les autres. *MO* a toujours un score de 5 et donc *EM* fait toujours aussi bien ou mieux, donc on peut la considérer comme dominée (mais pas strictement).

Nous ne ferons pas un état de l'art exhaustif des différentes méthodes d'élicitation de préférences, celles-ci étant trop nombreuses et dans des contextes très différents. À la place, nous présenterons tout d'abord l'objectif d'élicitation que nous poursuivons, puis nous évoquerons quelques méthodes classiques d'élicitation non incrémentale, avant d'enfin justifier notre choix pour l'élicitation incrémentale. Nous vous proposons plusieurs ressources intéressantes en gestion des préférences et en élicitation de préférences : Jacquet-Lagrèze et Siskos (2001); Doumpos et Zopounidis (2011); Pigozzi *et al.* (2016). Notons aussi les états de l'art sur l'élicitation de Benabbou (2017) et Bourdache (2020) qui sont très complets et recourent en partie celui de ce manuscrit.

1.3.1 Objectif d'élicitation

Nous considérons que les préférences du décideur sont modélisées par une fonction d'agrégation f_ω paramétrée par ω (voir la Sous-Section 1.2.2 pour plus d'informations). L'élicitation peut alors porter sur deux objets : les valeurs des performances ou la valeur des paramètres de la fonction f_ω . Les deux problématiques sont donc les suivantes :

Élicitation des valeurs des performances : les performances $u_j(x)$, $j \in \mathcal{Q}$ associées par le décideur à chaque alternative x ne sont pas connues précisément. L'élicitation a alors pour but de déterminer ces performances. Les méthodes d'élicitation font en général l'hypothèse que la fonction f_ω et ses paramètres sont précisément connus.

Nous n'allons pas nous intéresser à cette problématique, car nous supposons que les valeurs des performances sont connues avec précision au moment de l'échange

avec le décideur, et qu'elles n'évoluent pas avec le temps. Citons néanmoins quelques méthodes classiques en élicitation par valeur de performance : la méthode UTA (performance additive) (Jacquet-Lagrange et Siskos, 1982; Siskos et Yannacopoulos, 1985) qui permet d'obtenir un ensemble de valeurs de performances compatibles avec les informations disponibles en triant les alternatives par ordre de préférence, la méthode MACBETH (Bana e Costa et Vansnick, 1995) qui élicite les préférences du décideur par comparaison de paires de solutions et ensuite par programmation linéaire déterminer des performances compatibles.

Élicitation des valeurs des paramètres : les arbitrages faits par le décideur entre les différentes alternatives ne sont pas connus avec précision. Nous supposons que les performances $u_j(x), j \in \mathcal{Q}$ sont connues avec précision et que nous connaissons le type d'arbitrage du décideur (pondération, équilibre entre les critères, interaction complexe entre critères...) et que nous supposons une famille d'opérateur d'agrégation (WS, OWA, WOWA, Choquet...). L'élicitation se porte alors sur les paramètres ω de la fonction f_ω modélisant les préférences.

Dans le cadre de cette thèse, nous nous intéresserons à cette problématique, dont nous présenterons des méthodes dans la suite. L'objectif est donc de trouver les valeurs réelles du paramètre de notre fonction d'agrégation. Nous noterons ω^* la vraie valeur du paramètre que nous cherchons à estimer. Plutôt que de donner une estimation ponctuelle $\hat{\omega}^*$ de la vraie valeur du paramètre, pouvant être arbitraire, les méthodes dites robustes (Greco *et al.*, 2010; Angilella *et al.*, 2010) considèrent l'ensemble des valeurs de ω compatibles avec les informations préférentielles disponibles. Ces méthodes ne nous permettent pas d'avoir une estimation précise de la valeur du paramètre ω , mais donnent de meilleures garanties sur la qualité de la recommandation finale proposée.

Exemple 13. *Élicitation des valeurs des paramètres*

Nous reprenons le sous-ensemble des alternatives disponibles \mathbb{X} de l'Exemple 1. Supposons que les préférences du décideur soient modélisées par une somme pondérée WS_ω . Nous cherchons à déterminer la valeur du paramètre ω qui corresponde le mieux aux préférences du décideur. Cette valeur, notée ω^ , est supposée ici égale à $(0.6, 0.4)$. En l'absence d'information, l'ensemble initial des paramètres possibles*

est Ω , contraint par $0 \leq \omega_i \leq 1$ et $\sum_i \omega_i = 1$, $i \in \{1, 2\}$.

Nous supposons que nous avons recueilli préalablement les deux informations préférentielles suivantes :

- Information 1 : $EM \succeq CA$ (soit $4\omega_1 \geq 3\omega_2 \Rightarrow \omega_2 \leq 4/7$),
- Information 2 : $GO \succeq BT$ (soit $\omega_1 \leq 2\omega_2 \Rightarrow \omega_2 \geq 1/3$).

En supposant que ces informations soient exactes, nous pouvons alors déduire un encadrement de la valeur de ω^* , comme représenté sur la Figure 1.9. Ainsi, nous pouvons estimer que $\omega^* \in [1/3, 4/7]$.

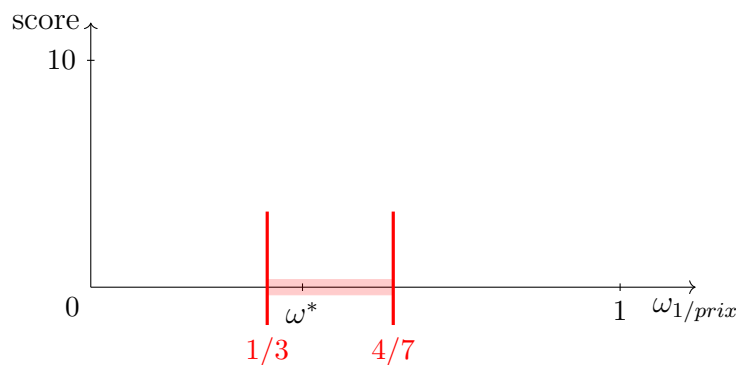


FIGURE 1.9 – Encadrement de la valeur ω^*

Notons qu'ici, nous avons une estimation robuste de la valeur ω^* , et non une estimation ponctuelle $\hat{\omega}^*$. Le choix de $\hat{\omega}^*$ influence la recommandation faite au décideur : si $\hat{\omega}^* = (2/3, 1/3)$, alors la recommandation est GO ; si $\hat{\omega}^* = (3/7, 4/7)$, alors la recommandation est EM .

1.3.2 Élicitation non incrémentale

L'élicitation non incrémentale se distingue de l'élicitation incrémentale par la nature des exemples. Un exemple en apprentissage est une donnée contenant une information sur les préférences du décideur. En élicitation non incrémentale, chaque exemple ne dépend pas des autres, tandis qu'en élicitation incrémentale, les nouveaux exemples dépendent des précédents.

En élicitation non incrémentale, l'algorithme doit, à partir des données qui lui sont fournies, déterminer les valeurs des paramètres précisément tout en étant

représentatif des données. L'objectif est donc de trouver au mieux les valeurs des paramètres, trouver la meilleure recommandation n'étant qu'une conséquence. Ici, un exemple est une information binaire sur les préférences du décideur, en supposant des comparaisons par paire : à la question $x \succeq y$, la réponse est *oui* ou *non*, soit les deux classes possibles. Nous voyons ici un lien avec les problèmes d'apprentissage automatique, plus particulièrement ceux de classification : l'idée est que l'algorithme choisi doit respecter tous les exemples (nos données classifiées), et donc chercher les paramètres pouvant respecter des contraintes. Notons qu'il peut y avoir plusieurs paramètres respectant les contraintes, et donc il est possible d'utiliser ensuite des algorithmes d'optimisation pour trouver une solution optimale. Une fois les paramètres de la fonction d'agrégation f_ω estimés, le modèle décisionnel peut être réutilisé quand les critères et les préférences sont les mêmes, quelles que soient les nouvelles alternatives présentées.

Sans être exhaustif, les approches d'élicitation indirectes de la valeur du paramètre d'une fonction d'agrégation existent au moins depuis les années 1970 (Srinivasan et Shocker, 1973; Pekelman et Sen, 1974). Pour traiter l'imprécision ou l'incertitude, il existe notamment la méthodologie SMAA (Analyse d'Acceptabilité Multicritère Stochastique) (Lahdelma *et al.*, 1998; Tervonen et Figueira, 2008) et des approches précurseurs (Bana e Costa, 1986). Pour les intégrales de Choquet, il est aussi possible de déterminer les poids des capacités de manière à respecter des informations préférentielles (Marichal et Roubens, 2000), y compris avec l'approche SMAA (Angilella *et al.*, 2015).

Limitations : les algorithmes d'élicitation non incrémentale sont très sensibles à la quantité et à la qualité des exemples disponibles : si les exemples ne sont pas assez nombreux ou représentatifs, le modèle final pourrait être assez éloigné des préférences du décideur, amenant à une recommandation potentiellement sous-optimale. Malheureusement, plus le nombre d'exemples est grand, plus il y a un risque d'apparition d'incohérences, ce qui pourrait soit arrêter l'algorithme sans obtenir de modèle décisionnel, soit le faire converger vers des paramètres erronés, amenant à une recommandation finale de mauvaise qualité. Par conséquent, nous n'avons pas de garantie *à priori* sur la qualité de la recommandation. De plus, il n'est pas possible pour le décideur de corriger une incohérence, ni pour nous de

choisir le nombre d'exemples ou leur nature, ce qui pourrait limiter le nombre de questions à poser d'une part, et d'autre part ajuster le modèle s'il est jugé mauvais.

Exemple 14. *Impact d'un mauvais exemple*

Dans l'Exemple 13 nous avons vu une illustration simple d'élicitation des valeurs des paramètres. Remarquons qu'à partir de deux exemples, l'intervalle donnant l'estimation du paramètre ω^* est assez large, car le nombre d'exemples est très faible. Pour améliorer la recommandation, il est donc nécessaire d'augmenter le nombre d'exemples. Supposons que nous ajoutons un nouvel exemple erroné : $EM \succeq MO$ (soit $\omega_1 \leq \omega_2 \Rightarrow \omega_2 \geq 1/2$). Sur la Figure 1.10 nous voyons l'impact de ce mauvais exemple : à partir des trois exemples, nous pouvons conclure à tort que $\omega^* \in [1/2, 4/7]$.

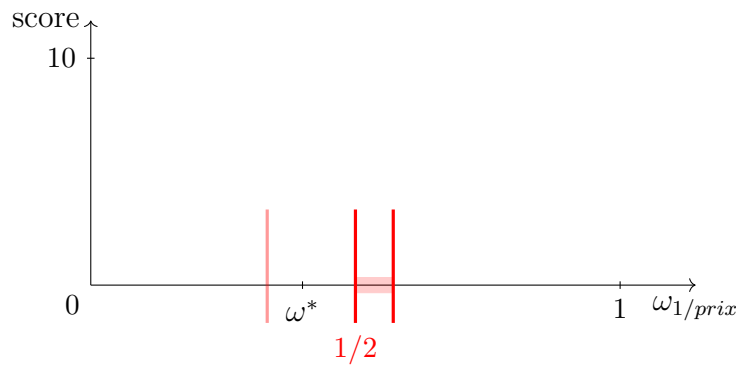


FIGURE 1.10 – Encadrement erroné de la valeur ω^*

1.3.3 Élicitation incrémentale

L'élicitation incrémentale (ou active) a pour objectif de proposer une recommandation optimale en alternant questions préférentielles judicieusement choisies et exploration de l'espace des paramètres possibles. Les questions sont ainsi posées progressivement au décideur, choisies en fonction des réponses déjà obtenues précédemment, tout en réduisant le plus que possible l'espace des paramètres possibles. Remarquons que l'objectif n'est pas de déterminer précisément la valeur des paramètres (nous avons un sous-ensemble de paramètres possibles réduit jusqu'à

avoir une recommandation de qualité), mais bel et bien de proposer la meilleure recommandation.

Ce type d'approche est désormais assez standard dans la littérature (par exemple : Lu et Boutilier (2011), Benabbou *et al.* (2017), Bourdache et Perny (2019)). Nous allons présenter quelques méthodes importantes en élicitation incrémentale : les approches interactives, la méthode ISMAUT et l'élicitation incrémentale fondée sur l'optimisation d'un critère (usuellement le regret Minimax).

Méthodes interactives

L'élicitation interactive (Benayoun *et al.*, 1971; Zionts et Wallenius, 1976; Korhonen et Wallenius, 1988; Vanderpooten et Vincke, 1989; Miettinen, 1999) a pour but d'explorer (sans toutes les énumérer) l'ensemble des solutions Pareto-optimales en fonction des différentes interactions avec le décideur. L'idée générale est d'évaluer des solutions avec une instance des paramètres de la fonction d'agrégation f_ω , puis de présenter une ou plusieurs solutions (choisies judicieusement) au décideur. S'il est satisfait, l'algorithme s'arrête. Dans le cas contraire, l'algorithme continue.

Une des premières méthodes interactives est l'approche STEM (STEp Method) (Benayoun *et al.*, 1971). Une solution Pareto-optimale est déterminée à l'aide d'une norme de Tchebychev pondérée augmentée, puis est présentée au décideur. S'il n'est pas satisfait de la recommandation, il doit indiquer sur quel critère il accepte de dégrader la performance, et ce, jusqu'à quel point. Ce type de question est cognitivement difficile et plusieurs travaux se sont portés sur la simplification des questions, arrivant notamment à une variante de l'approche STEM se basant sur des comparaisons de paires d'alternatives (Lokman *et al.*, 2018).

Une limitation des approches interactives est l'absence de garantie sur l'optimalité sur la recommandation finale : le décideur s'arrête dès qu'il le souhaite, en théorie dès qu'une alternative lui convient, même si l'estimation de ω^* reste très partielle. En pratique, nous pouvons très bien imaginer que le décideur est lassé de répondre à des questions et qu'il souhaite s'arrêter, alors même que la solution ne lui convient pas parfaitement. De plus, le décideur n'a pas nécessairement une bonne vision de l'espace des solutions, et il peut s'arrêter de bonne foi avec une solution qu'il juge convenable, alors qu'il en existe de meilleures.

Méthode ISMAUT

La méthode ISMAUT (Imprecisely Specified Multiattribute Utility Theory) (White *et al.*, 1984) est l'une des premières méthodes d'élicitation incrémentale développée pour l'élicitation des paramètres. Pour cette méthode, les préférences du décideur sont supposées être additives (donc une somme pondérée entre les différents critères).

Au départ, la méthode définit le sous-ensemble des paramètres possibles $E \subseteq \Omega$, ainsi que les solutions dominées ND_E au sens de l'utilité additive pondérée :

$$ND_E(\mathbb{X}) = \{x \in \mathbb{X} \mid \nexists y \in \mathbb{X} \text{ tel que } y \succeq_E x \text{ et } \neg(x \succeq_E y)\} \quad (1.20)$$

avec \succeq_E une relation binaire définie par $x \succeq_E y \Leftrightarrow \forall \omega \in E, f_\omega(x) \leq f_\omega(y)$. L'ensemble ND_E est alors présenté au décideur. Dans le cas où il peut choisir l'alternative qu'il préfère, l'algorithme s'arrête. Dans le cas où il n'arrive pas à choisir l'alternative qu'il préfère, des nouvelles informations sont récoltées pour réduire E . La récolte d'information peut se faire selon différents types de questions (par exemple comparaison par paire). Le but est, à chaque étape, de réduire l'ensemble ND jusqu'à que le décideur soit capable de prendre une décision. La méthode ISMAUT se focalisant sur la représentation de l'incertitude concernant les valeurs de ω , les problématiques d'exploration de l'espace des solutions et du choix des questions, problématiques importantes pour limiter l'effort cognitif du décideur, ne sont abordées qu'ultérieurement (par exemple Holloway et White III (2003)).

Demander à un décideur de choisir son alternative préférée parmi plusieurs n'est pas une très bonne idée, le décideur pouvant avoir des difficultés à choisir une unique alternative parmi un ensemble d'alternatives trop proches, trop éloignées ou/et trop nombreuses. La méthode présentée ensuite cherche à réellement minimiser l'effort cognitif pour le décideur, notamment en faisant en sorte que le décideur n'ait qu'à choisir parmi des paires d'alternatives, choisies judicieusement.

Élicitation incrémentale fondée sur le critère du regret Minimax

Le critère de décision du regret Minimax (Savage, 1951) est notamment utilisé dans les problèmes de décision dans l'incertain, comme nous l'avons vu dans la Sous-Section 1.2.3. Employé initialement pour déterminer les valeurs de performance de

chaque critère, il a été adapté ensuite pour déterminer les valeurs des paramètres ω d'une fonction d'agrégation f_ω (Salo et Hamalainen, 2001; Boutilier *et al.*, 2006, 2009) et peut être utilisé y compris sur des fonctions complexes comme l'intégrale de Choquet (Benabbou *et al.*, 2017). Ce nouveau type d'approche cherche à minimiser l'effort cognitif pour le décideur, en posant des questions simples (le plus souvent des comparaisons de paires) tout en cherchant à maximiser le gain d'information.

Pour minimiser l'effort cognitif, le critère de décision est utilisé pour choisir judicieusement la paire d'alternative à présenter au décideur. Le regret Minimax est intéressant, car il donne des garanties fortes sur la qualité de la recommandation, tout en étant moins conservatif que le critère Maximin (Wang et Boutilier, 2003). Néanmoins, d'autres critères sont aussi intéressants sur d'autres points, notamment le critère Maximax comme nous le verrons dans le Chapitre 3. Pour le calcul du regret Minimax, nous nous basons sur la Sous-Section 1.2.3. Ainsi, nous retrouvons les différentes notations qui sont le regret, le PMR, le MR et le mMR.

Algorithme 1 : Élicitation incrémentale fondée sur le critère du regret Minimax

Données : Sous-ensemble d'alternatives disponibles \mathbb{X} , Nombre maximal de questions q_{\max} , seuil de regret τ .

Résultat : La recommandation optimale $x^* = \text{mMR}(E_k)$.

$k = 0, E_0 = \Omega$;

tant que $k < q_{\max}$ *et* $\text{mMR}(E_k) > \tau$ **faire**

Choisir deux alternatives x et y ;
Demander au décideur si $x \succeq y$ ou $x \preceq y$;
si le décideur répond $x \succeq y$ alors
$E_{k+1} = E_k \cap \{\omega x \preceq y\}$
sinon $E_{k+1} = E_k \cap \{\omega x \preceq y\}$;
$k = k + 1$;

fin

L'Algorithme 1 détaille le fonctionnement de l'élicitation incrémentale fondée sur le critère du regret Minimax. À chaque étape k , deux alternatives sont choisies et présentées au décideur. Le décideur doit indiquer quelle alternative il préfère.

Dès lors, le sous-ensemble des valeurs possibles de ω à l'étape k , noté E_k (avec $E_0 = \Omega$) est mis à jour avec la frontière de décision linéaire obtenue à partir de la réponse du décideur. L'algorithme s'arrête dès que le budget de questions est dépassé, ou quand $mMR(E_k)$ est plus petit qu'un seuil τ . L'Exemple 15 illustre le calcul du mMR initial (pour Ω). Ultérieurement, l'Exemple 16 illustre la première étape $k = 1$ de l'algorithme, avec une heuristique pour le choix des alternatives à présenter au décideur.

Sous réserve que deux hypothèses soient respectées :

- Le décideur indique toujours correctement quelle alternative il préfère lors de comparaisons de paires d'alternatives (ne commettant ainsi jamais d'erreurs).
- La famille de fonction d'agrégation de critères (WS, OWA, Choquet...) modélise correctement les préférences du décideur, et donc les interactions entre les critères.

Alors l'algorithme converge vers une recommandation optimale :

Proposition 2. *Soit E_k le sous-ensemble des paramètres possibles obtenu après la k^e question. Sachant que nous avons E_k et E_{k+1} , les sous-ensembles obtenus après les questions k et $k + 1$ avec $\emptyset \neq E_{k+1} \subseteq E_k$, alors si les hypothèses précédentes sont respectées, nous avons :*

$$PMR(x, y, E_k) \geq PMR(x, y, E_{k+1}), \quad (1.21)$$

$$MR(x, E_k) \geq MR(x, E_{k+1}), \quad (1.22)$$

$$mMR(E_k) \geq mMR(E_{k+1}). \quad (1.23)$$

Démonstration. PMR. Supposons que nous avons une fonction f et deux ensembles E, E_i tels que $E_i \subseteq E$. Nous avons $\max_{x \in E} f(x) \geq \max_{x \in E_i} f(x)$, le maximum de E étant soit dans E_i ou dans $E \setminus E_i$. Nous remplaçons f par le PMR, E par E_k et E_i par E_{k+1} , car $E_{k+1} \subseteq E_k$. (1.21) est alors démontré.

Les preuves pour le MR et le mMR s'ensuivent, comme ils sont le maximum et le minimum pris sur les valeurs du PMR qui sont décroissantes pour n'importe quelle paire x, y quand nous considérons de plus petits ensembles, comme nous venons de le démontrer.

□

Remarquons qu'en pratique, le regret décroît strictement (voir par exemple Braziunas (2012)). En plus de déterminer une recommandation à chaque itération, le regret Minimax nous donne une distance entre la recommandation et la vraie solution optimale (le regret).

Exemple 15. *Choix initial*

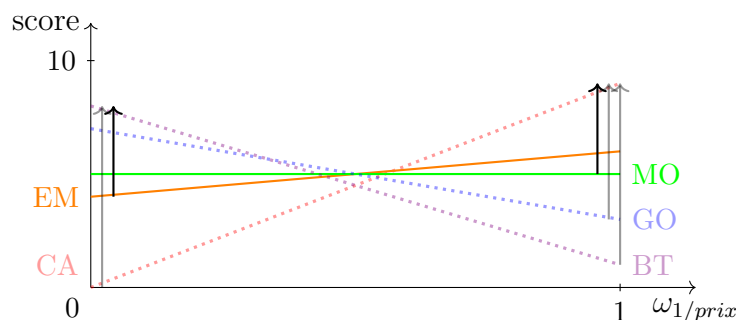
Nous reprenons le sous-ensemble des alternatives disponibles de l'Exemple 1. Nous supposons que les préférences du décideur sont représentées par une somme pondérée WS_ω de poids $\omega = (0.6, 0.4)$. En l'absence d'information, l'ensemble initial des paramètres possibles est Ω , contraint par $0 \leq \omega_i \leq 1$ et $\sum_i \omega_i = 1$, $i \in \{1, 2\}$

La première étape est de calculer le regret maximal entre chaque paire d'alternatives : $PMR(x, y, \Omega) = \max_{\omega \in \Omega} R_\omega(x, y) \forall x, y \in \mathbb{X}^2$. Ce regret maximum par paire est déterminable en optimisant $\max_{\omega \in \Omega} f_\omega(y - x) \forall x, y \in \mathbb{X}^2$. Comme nous optimisons une fonction linéaire sur un polytope convexe Ω , ce problème d'optimisation est résolu facilement et exactement avec un algorithme d'optimisation linéaire, par exemple le simplexe. Par exemple, le regret maximal entre le brie truffé et l'emmental est $PMR(BT, EM, \Omega) = 5$: en effet, pour $\omega = (0, 1)$ nous avons $WS_\omega(BT) = 1$ et $WS_\omega(EM) = 6$, donnant un regret de 5. Les autres regrets maximaux par paire sont présentés dans la Table 1.9.

x/y	CA	EM	MO	GO	BT	MR
CA	0	4	5	7	8	8
EM	3	0	1	3	4	4
MO	4	1	0	2	3	4
GO	6	3	2	0	1	6
BT	8	5	4	2	0	8
mMR						4

TABLE 1.9 – Détermination du mMR initial de l'Exemple 15

L'étape suivante est de calculer le regret maximal pour chaque alternative, soit $MR(x) = \max_{y \in \mathbb{X}} PMR(x, y, \omega) \forall x \in \mathbb{X}$. Les regrets maximaux ainsi calculés sont aussi présentés dans la Table 1.9, mais aussi sur la Figure 1.11. Sur cette représentation, le MR de chaque alternative est représentée par une flèche.

FIGURE 1.11 – Score des alternatives et MR pour $\omega \in \Omega$

Nous pouvons enfin choisir une alternative à recommander au décideur en prenant l'alternative ayant le regret maximal minimal : $mMR(\Omega) = \min_{x \in \mathbb{X}} MR(x, \Omega)$. $mMR(\Omega) = 4$ indiquant que le meilleur choix en l'absence d'information est EM ou MO, tous deux ayant le même regret maximal minimal.

Remarquons que les solutions initialement proposées dans l'Exemple 15 sont non optimales par rapport aux préférences réelles, mais inconnues, du décideur. En effet, en l'absence d'information sur celles-ci, le regret Minimax va privilégier des solutions qui sont le moins regrettées dans le pire cas, qui sont généralement équilibrées. Bien entendu, le principe de l'élicitation est de recueillir des informations préférentielles pour ensuite proposer une solution correspondant au mieux aux préférences du décideur. Le choix des questions est très important, d'autant plus que le but de cette méthode est de minimiser l'effort cognitif du décideur, et donc d'utiliser le moins de questions possibles pour arriver à une recommandation de qualité. Choisir aléatoirement les paires d'alternatives à comparer n'est pas une très bonne idée, certaines comparaisons ne pouvant apporter que peu d'information. Déterminer les questions les plus informatives à chaque itération est compliqué, du fait du nombre potentiellement élevé de questions possibles (par exemple, si nous avons n alternatives, cela fait $\binom{n}{2}$ comparaisons de paires possibles), d'autant plus qu'en théorie, il faudrait prendre en compte toutes les futures questions et réponses possibles à chaque itération (Boutillier, 2002), bien qu'en pratique l'évaluation se fasse sans prendre en compte les questions futures (par exemple Viappiani et Boutillier (2009)).

Stratégie de la solution courante (CSS) : La stratégie de la solution courante (CSS) est une heuristique couramment employée pour déterminer de bonnes questions, qui est très efficace en pratique pour réduire les valeurs du mMR (Boutilier *et al.*, 2006). La paire présentée au décideur à l'itération k est composée de la meilleure recommandation courante $x^* = \arg \text{mMR}(E)$, qui est optimale par rapport au critère du regret Maximax; et de l'alternative $y^* = \arg \max_{y \in \mathbb{X}} \text{PMR}(x^*, y, E_k)$, qui est l'alternative réalisant le plus grand regret possible avec x^* . L'Algorithme 2 est une variante de l'Algorithme 1 avec prise en compte de l'heuristique CSS pour choisir les alternatives à présenter au décideur.

Algorithme 2 : Élicitation incrémentale fondée sur l'heuristique CSS

Données : Sous-ensemble d'alternatives disponibles \mathbb{X} , nombre maximal de questions q_{\max} , seuil de regret τ .

Résultat : La recommandation optimale x^* .

$k = 0, E_0 = \Omega;$

tant que $k < q_{\max}$ *et* $\text{mMR}(E_k) > \tau$ **faire**

$x^* = \text{mMR}(E_k) ;$
$y^* = \arg \max_{y \in \mathbb{X}} \text{PMR}(x^*, y, E_k) ;$
Demander au décideur si $x^* \succeq y^*$ ou $x^* \preceq y^*$;
si le décideur répond $x^* \succeq y^*$ alors
$E_{k+1} = E_k \cap \{\omega x^* \succeq y^*\}$
sinon $E_{k+1} = E_k \cap \{\omega x^* \preceq y^*\};$
$k = k + 1;$

fin

Exemple 16. CSS

Nous continuons l'Exemple 15 en utilisant la stratégie CSS pour déterminer quelle paire d'alternatives le décideur doit comparer, plus particulièrement la meilleure recommandation courante avec l'alternative qui a le pire regret maximal avec celle-ci. Nous avons vu que la meilleure recommandation courante est EM ou MO. Nous prenons au hasard EM. Nous avons $\max_{y \in \mathbb{X}} \text{PMR}(EM, y, \Omega) = 4$ et nous choisissons donc comme opposant BT. En supposant que le décideur soit rationnel, il indiquera $BT \succeq EM$, ce qui nous permet de rajouter une nouvelle contrainte

linéaire sur l'ensemble des paramètres possibles $\Omega : BT \succeq EM \Rightarrow 8\omega_1 + \omega_2 \geq 4\omega_1 + 6\omega_2 \Rightarrow 4\omega_1 \geq 5\omega_2$. Notons ce sous-ensemble de paramètres possibles E_1 . Le score des alternatives en fonction de ce nouvel sous-ensemble est représenté dans la Figure 1.12. Nous pouvons alors calculer de nouveau le PMR et le MR comme présenté dans la Table 1.10.

x/y	CA	EM	MO	GO	BT	MR
CA	0	4	5	7	⑧	8
EM	-8/9	0	1	3	④	4
MO	-1	-1/9	0	2	③	3
GO	-11/9	-3/9	-2/9	0	①	1
BT	-8/9	0	1/9	③/9	0	③/9
mMR						3/9

TABLE 1.10 – PMR après la première question de l'Exemple 16

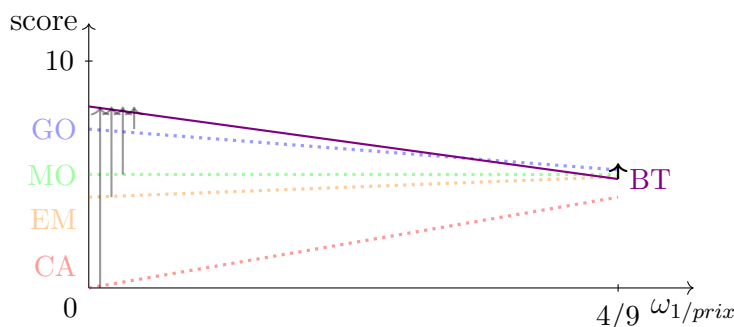


FIGURE 1.12 – Score des alternatives et MR pour $\omega \in E_1$

Nous pouvons ainsi déterminer que $mMR(E_1) = 3/9$ pour l'alternative BT, qui est alors la recommandation courante. Il serait possible de diminuer encore le mMR jusqu'à qu'il s'approche ou atteigne 0 (ou que le décideur soit lassé de répondre à des questions) pour avoir de meilleures garanties sur la recommandation finale.

Limitations : Nous avons vu dans la proposition 2 que l'algorithme converge quand des hypothèses sont respectées (décideur qui répond toujours correctement

comme un oracle, et modèle suffisamment riche pour modéliser les préférences du décideur). Si les hypothèses ne sont pas respectées, de l'incohérence peut apparaître, ayant l'une des conséquences suivantes :

- L'algorithme est incapable de trouver un paramètre compatible avec toutes les informations préférentielles données. Cela arrive quand au moins deux réponses sont contradictoires. Nous avons alors $E_k = \emptyset$. Avec l'heuristique CSS, il est impossible d'avoir ce cas (Benabbou *et al.*, 2015).
- Le paramètre ω^* correspondant aux réelles préférences du décideur est exclu. Ainsi, $\omega^* \notin E_k$, et comme nous avons vu dans la Proposition 2 que $E_{k+1} \subseteq E_k$, alors ω^* sera exclu à chaque étape de l'élicitation. Pire encore, L'algorithme converge alors vers une mauvaise estimation de ω , donnant un modèle différent des préférences du décideur. Cela arrive quand les réponses ne sont pas contradictoires, mais au moins une est contradictoire avec les préférences du décideur. Avec l'heuristique CSS nous aurons toujours ce cas si une erreur est commise.

Le premier cas est plus facile à gérer : nous sommes au courant qu'il y a de l'incohérence, aucune recommandation n'est faite et il est donc possible d'essayer de la corriger. Dans le second, nous ne savons pas qu'il y a de l'incohérence : nous aurons une recommandation finale de mauvaise qualité sans le savoir. La stratégie CSS a pour avantage de garantir qu'aucune réponse du décideur ne va contredire des informations préférentielles précédentes, même s'il commet des erreurs ou que la fonction d'agrégation ne peut pas représenter les préférences du décideur. Néanmoins, cela veut aussi dire que la stratégie CSS peut nous faire converger vers le mauvais modèle de préférence si les hypothèses initiales ne sont pas respectées, sans avoir le moindre signal sur la présence d'incohérence.

Il est important de noter qu'il existe d'autres approches pour gérer l'inconsistance, autre que celle que nous allons présenter au long de ce manuscrit. Celles-ci cherchent un modèle qui minimise une erreur moyenne. Cela peut se faire par exemple via la minimisation des moindres carrés (Grabisch *et al.*, 2008) ; la maximisation d'une marge (Teso *et al.*, 2017) ; ou même avec des probabilités bayésiennes, qui consistent à spécifier une loi de probabilité appliquée sur les différents modèles, mise à jour en fonction des informations préférentielles, utilisant l'espérance pour

faire des inférences (Eric *et al.*, 2007; Viappiani et Boutilier, 2010; Bourdache *et al.*, 2019). Remarquons que ces approches sont efficaces pour gérer l'inconsistance, mais même si toutes les réponses sont cohérentes, elles n'offrent pas le même niveau de garantie sur la qualité de la recommandation finale que le regret Minimax. En effet, elles n'offrent au mieux que des garanties en moyenne, sous réserve que le modèle d'incertitude utilisé est correct.

Objectif	Choix questions	
	Non incrémentale	Incrémentale
Performance	Jacquet-Lagreze et Siskos (1982)	
Paramètre (+ Incertitude)	Srinivasan et Shocker (1973)	Boutilier <i>et al.</i> (2006)
	Lahdelma <i>et al.</i> (1998)	Probabilités : Bourdache <i>et al.</i> (2019) Ensembles : nos travaux

TABLE 1.11 – Récapitulatif des approches d'élicitation et exemples de références

Dans la Table 1.11 est proposé un récapitulatif des différentes méthodes d'élicitation présentées dans cette section, avec un exemple de travail représentatif si c'est pertinent. Il existe de nombreuses méthodes d'élicitation de valeur de paramètre, incrémentale ou non, avec gestion de l'incertitude ou non. Néanmoins, la gestion de l'incertitude se fait souvent par des approches se basant sur des probabilités. Nous avons indiqué juste avant que les approches à base de probabilités posent parfois un problème. Nos travaux proposent une approche ensembliste pour gérer l'incertitude en élicitation incrémentale de valeur de paramètres. Nous proposons dans le Chapitre 2 notre solution pour détecter de l'incohérence, tout en revenant sur les problèmes des approches minimisant une erreur moyenne, tandis que dans le Chapitre 4 nous proposons diverses manières de gérer l'incohérence.

1.4 Gestion de l'incertitude et fusion d'information

Dans les sections précédentes, nous supposons avoir accès à une information certaine sur les préférences du décideur³, et que celui-ci est capable de donner ses préférences de manière fiable. Nous savons par exemple, lors d'une comparaison de paire, si x est préférée ou non à une alternative y . Il ne peut pas avoir la moindre expression d'une incertitude de la part du décideur.

Manipuler des données certaines rend le travail d'élicitation plus facile, mais pouvons-nous réellement demander en permanence des informations certaines au décideur ? En effet, le décideur peut être incertain de son choix, par exemple s'il est distrait, fatigué après plusieurs questions, ou que les alternatives qu'il doit comparer ne sont pas suffisamment distinguables. Par exemple, indiquer son fromage préféré entre la Mozzarella et le Maroilles est facile si on est intolérant à l'odeur du fromage, mais choisir entre de la Mozzarella et du fromage blanc est assez compliqué si seul l'odeur du fromage nous importe. Faire un choix ferme est alors compliqué, et peut mener au mauvais choix. Comme nous l'avons vu dans la section précédente, la plupart des algorithmes d'élicitation mènent à une mauvaise recommandation en présence de réponses erronées. Il est donc utile d'avoir un modèle permettant de représenter l'incertitude du décideur dans le choix de ses réponses.

De plus, le décideur nous donne plusieurs informations préférentielles durant l'élicitation. Il est souhaitable d'obtenir une seule information préférentielle synthétique à partir des différentes informations partielles. Or, les informations préférentielles peuvent être incomplètes, et même dorénavant inexactes, rendant l'obtention d'une information synthétique difficile.

Dans cette section, nous avons pour objectif de présenter les modèles mathématiques utilisés dans nos travaux pour gérer ces problématiques, tout en évoquant très rapidement d'autres solutions, plus particulièrement pour la gestion de l'incertitude. L'utilisation de ces différents modèles pour modéliser des préférences incertaines et faire de l'élicitation est détaillée dans le Chapitre 2.

3. Remarquons qu'une information certaine peut aussi être fausse (affirmer que $2+2=5$) ou/et imprécise (affirmer que $2+2$ est compris entre 3 et 5)

1.4.1 Probabilités imprécises

Dans de nombreux problèmes, nous supposons qu'il est possible de répondre de manière sûre à des questions. Précédemment, nous avons supposé qu'un décideur est capable d'indiquer quelle alternative il préfère parmi deux alternatives. Pourtant, il n'est pas toujours possible de répondre de manière sûre : c'est l'incertitude. L'incertitude arrive notamment quand on ne sait pas si une proposition est vraie ou fausse, ou quand on ne sait pas si un évènement va se produire. Ainsi, il est difficile d'indiquer avec certitude, y compris pour les météorologues professionnels, s'il va pleuvoir demain. Par contre, il est tout à fait possible d'exprimer s'il est probable/peu probable qu'il pleuve demain en fonction des informations disponibles.

Probabilités et critiques

Un outil classique pour mesurer la vraisemblance d'un évènement sont les probabilités, permettant de mesurer l'incertitude d'un évènement A . La théorie des probabilités est fondée sur trois axiomes de probabilité (Kolmogorov, 1950) :

1. Pour tout évènement A : $0 \leq P(A) \leq 1$.
2. Soit Ω l'univers (l'ensemble des issues possibles). Nous avons alors $P(\Omega) = 1$.
3. Toute famille dénombrable d'évènements deux à deux disjoints satisfait $P(\cup A_i) = \sum P(A_i)$.

Plusieurs interprétations des probabilités existent (Hájek, 2019), dont notamment :

- L'approche *fréquentiste*, initialement exposée par Venn (1888), où les probabilités mesurent une fréquence d'occurrence, par exemple le nombre de jours qu'il a plu par rapport au nombre de jours par an pour avoir la probabilité qu'il pleuve demain. L'approche est alors empirique et se base sur des données.
- L'approche *subjective* (ou Bayésienne), se basant notamment sur les travaux de De Finetti (1937, 1972), qui exprime le degré de croyance par rapport à un évènement. Dans l'approche initiale de De Finetti, $P(A)$ est la somme qu'un agent est prêt à parier s'il reçoit 1 quand A arrive. L'approche est donc subjective et se base initialement sur les connaissances et l'intuition, bien qu'il soit aussi possible de faire du Bayésien avec des données.

Malgré ces différentes interprétations, et les très nombreux travaux sur les probabilités, plusieurs critiques envers les probabilités subsistent. Celles-ci sont anciennes (par exemple la critique de Boole (1854), qui exprime un doute sur la capacité des probabilités de mesurer une émotion de l'esprit), et assez nombreuses (Bradley, 2019). Parmi certaines critiques, nous trouvons :

- En théorie des probabilités, il est supposé que toute incertitude est modélisable par une probabilité. Pourtant, les informations sont souvent partielles (population non représentative/limitée, ou connaissances partielles).
- En plus de pouvoir tout modéliser par une probabilité, il est supposé que toute relation est complète (préférences par exemple), or certains auteurs (Kaplan, 1998; Joyce, 1999) suggèrent qu'il faudrait faire une distinction entre incomplétude (pas assez d'information pour décider) et incomparabilité (quelles que soient les informations obtenues, il est impossible de comparer deux évènements).
- Les données peuvent être imprécises, non fiables ou en faible nombre. Pour l'approche fréquentiste, c'est un problème majeur. Pour l'approche bayésienne, des solutions existent, mais nécessitent souvent des lois de probabilité très complexes.
- Modéliser l'ignorance avec les probabilités se fait en général avec une loi uniforme : si nous ne savons pas du tout s'il va pleuvoir ou faire beau demain, nous indiquons une probabilité de $1/2$ pour chaque situation. Mais comment distinguer le fait qu'il y a une chance sur deux qu'il pleuve demain, ou tout simplement que nous ne savons pas le temps de demain ?

Notons aussi le paradoxe de Ellsberg (1961) qui montre que les décisions prises par des individus ne sont pas toujours consistantes avec l'utilité espérée subjective (théorie des probabilités), et que les individus préfèrent par exemple prendre un risque quantifié, qu'un risque inconnu/non calculable.

Probabilités imprécises

Toute une littérature s'est développée pour répondre aux critiques des probabilités. Ici, nous nous intéressons aux probabilités imprécises (Augustin *et al.*, 2014),

qui sont un cadre formel englobant et étendant les probabilités pour permettre l'incomplétude, l'imprécision et l'indécision, et donc de manipuler des informations incomplètes⁴. Si le développement des probabilités imprécises est assez récent, notamment sous l'impulsion du livre de Walley (1991), les idées sont beaucoup plus anciennes. Nous pouvons citer notamment la critique de Boole (1854), les travaux de Keynes (1921) sur les probabilités incomplètes, les expériences de Ellsberg (1961) ou encore les travaux de Smith (1961); Dempster (1967); Shafer (1976); Walley et Fine (1982) que nous allons évoquer. Nous pouvons voir que les probabilités imprécises sont utilisées dans des domaines autres que les statistiques ou la théorie de la décision (notamment en économie et en philosophie).

L'objectif n'est pas de présenter en détail les probabilités imprécises. Tout juste, nous évoquerons quelques-uns des modèles mathématiques les plus utilisés dans le cadre des probabilités imprécises, avant de nous focaliser sur l'un d'entre eux dans la Sous-Section 1.4.2. Parmi les modèles les plus utilisés, nous trouvons :

- Les *probabilités inférieures et supérieures* (Smith, 1961; Dempster, 1967), qui donnent une borne $P(A) \in [\underline{P}(A), \overline{P}(A)]$ sur la probabilité d'un évènement A (plutôt qu'un seul nombre).
- Les *prévisions inférieures et supérieures* (Walley, 1991), qui sont une généralisation des probabilités inférieures et supérieures. Nous avons un ensemble de loi de probabilité (et pas une seule), supposé fermé et convexe, que nous appelons ensemble crédal, dont nous cherchons les bornes.
- La théorie des *fonctions de croyance*, aussi connue sur le nom de *théorie de l'évidence* ou *théorie de Dempster–Shafer* (Dempster, 1967; Shafer, 1976; Denoeux, 2019). Un degré de croyance est attribué à des ensembles d'éléments via des masses m , avec deux mesures induites : la croyance et la plausibilité.
- Les *mesures de possibilité et de nécessité*, utilisés en *théorie des possibilités* (Zadeh, 1978; Dubois, 1986; Dubois et Prade, 2001) que nous évoquerons en détail dans la Sous-Section 1.4.2. Un degré de possibilité π est attribué à chaque élément, avec deux mesures induites : la possibilité et la nécessité.

4. Remarquons que le nom de probabilité imprécise est peu justifié et critiqué par la communauté elle-même, car les outils en eux-mêmes ne sont pas des probabilités rendues imprécises, mais bel et bien des outils pour gérer l'imprécision (entre autres).

Ces modèles sont liés les uns les autres. Nous ne détaillerons pas ces liens⁵. La Figure 1.13 montre ces différents liens, avec une flèche indiquant que le modèle pointé est une généralisation du modèle d'où part la flèche.

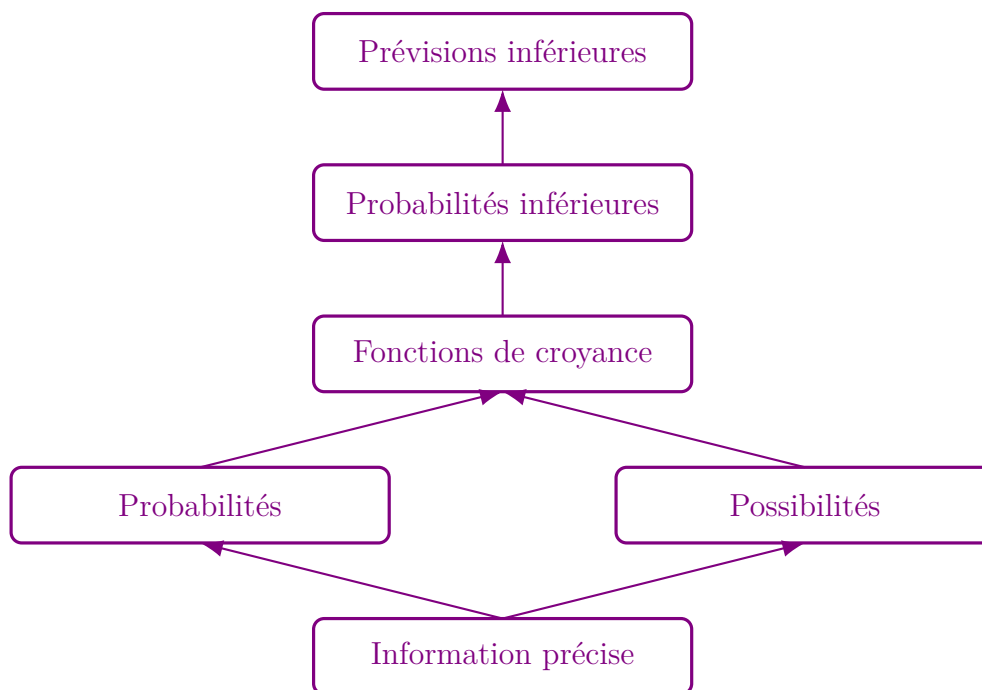


FIGURE 1.13 – Liens entre différents modèles des probabilités imprécises

1.4.2 Théorie des possibilités

La théorie des possibilités est une théorie mathématique alternative aux probabilités pour gérer certains types d'incertitude. Introduite formellement par Zadeh (1978) comme une extension de ses travaux sur la logique floue et les ensembles flous, elle a néanmoins une histoire un peu plus ancienne, avec des approches évoquant les idées de possibilité et de nécessité (par exemple l'économiste Shackle (1961) pour une théorie de la décision avec des degrés de possibilité). Cette théorie a été

5. Ces liens sont explicités dans la revue de Augustin *et al.* (2014). Pour le lien entre la théorie des possibilités et la théorie des fonctions de croyance : la théorie des possibilités est un cas particulier de la théorie des fonctions de croyance, où ce qui s'appelle les ensembles focaux sont imbriqués les uns dans les autres.

particulièrement développée par Dubois et Prade dans plusieurs de leurs travaux (Dubois, 1986; Dubois et Prade, 2001).

Dans la logique floue, qui est une logique non-classique polyvalente, les valeurs de vérité des variables prennent valeur dans $[0, 1]$, alors qu'en logique mathématique les valeurs de vérité sont faux (0) et vrai (1). Dans les ensembles flous, le degré d'appartenance d'un élément dans un ensemble prend valeur dans $[0, 1]$, alors que dans les théories des ensembles plus traditionnelles, un élément appartient ou non à un ensemble (Zadeh, 1965). Ces éléments sont en lien avec la définition d'une distribution de possibilité :

Définition 18. *Distribution de possibilité*

Une distribution de possibilité π est une application $\pi(\Omega) \rightarrow [0, 1]$ partant d'un ensemble de valeurs Ω vers une échelle totalement ordonnée dans l'intervalle unitaire $[0, 1]$, avec les axiomes suivants :

- $\pi(\emptyset) = 0$,
- $\pi(\Omega) = 1$,
- $\pi(U \cup V) = \max(\pi(U), \pi(V))$ si $U \cap V = \emptyset$.

$\pi(\omega)$ mesure ainsi la plausibilité d'une valeur de paramètre $\omega \in \Omega$. La convention suivante est associée à une distribution de possibilité :

- $\pi(\omega) = 0$ indique qu'une valeur est impossible.
- $\pi(\omega) = 1$ indique qu'une valeur est totalement possible⁶.

Une distribution π est dite consistante (ou normalisée) si $\max_{\omega \in \Omega} \pi(\omega) = 1$. Dans le cas contraire, le deuxième axiome n'est plus respecté.

Définition 19. *Spécificité*

Une distribution π_1 est dite au moins aussi spécifique qu'une distribution π_2 si et seulement si $\pi_2(\omega) \leq \pi_1(\omega) \forall \omega \in \Omega$.

π_1 est alors plus restrictive et plus informative. Remarquons deux distributions particulières :

6. Mais pas forcément certaine.

- La distribution de possibilité la moins spécifique avec $\pi(\omega) = 1 \forall \omega \in \Omega$, correspondant à une ignorance totale (tous les états sont possibles).
- La distribution de possibilité la plus spécifique avec $\pi(\omega') = 1$ et $\pi(\omega) = 0 \forall \omega \neq \omega'$, correspondant à une connaissance précise.

Pour obtenir des informations sur un évènement $U \subseteq \Omega$, nous introduisons deux mesures duales : la mesure de possibilité $\Pi : 2^\Omega \rightarrow [0, 1]$ et la mesure de nécessité $N : 2^\Omega \rightarrow [0, 1]$:

Définition 20. *Mesures de possibilité et de nécessité*

Soit une distribution de possibilité π , un évènement U et son complémentaire U^c . La mesure de possibilité Π et la mesure de nécessité N de U sont définies par :

$$\Pi(U) = \sup_{\omega \in U} \pi(\omega), \quad (1.24)$$

$$N(U) = \inf_{\omega \notin U} 1 - \pi(\omega) = 1 - \Pi(U^c). \quad (1.25)$$

La mesure de possibilité évalue à quel point un évènement U est consistant avec nos connaissances. La mesure de nécessité évalue à quel point U est certain à partir de nos connaissances. Nous avons les liens suivants :

- $\Pi(U) = 1 \Rightarrow N(U) = 1$,
- $\Pi(U) < 1 \Rightarrow N(U) = 0$.

Remarquons que ces mesures ne sont pas des mesures de probabilités. Néanmoins, $\Pi(U)$ peut être vu comme une probabilité supérieure de $P(U)$ (Walley, 1996). Si π est consistant, alors $N(U)$ est une probabilité inférieure de $P(U)$. Une distribution de possibilité définit alors un unique ensemble crédal (ensemble de lois de probabilité) $K = \{P | N(U) \leq P(U) \leq \Pi(U), \forall U \subseteq \Omega\}$.

Exemple 17. *Possibilité et fromage*

Supposons que nous voulons représenter nos connaissances sur les préférences du décideur en matière de fromage. Supposons que ses préférences puissent être représentées par les états suivants : $S = \{\text{aucun goût, plutôt le prix, équilibré, plutôt le goût, grand amateur}\}$, et nous voulons avoir des mesures sur l'évènement U « il s'intéresse plutôt au prix qu'au goût, ou autant ». Nous avons les informations suivantes grâce à ses amis :

- Tous indiquent que jamais il ne prendrait un fromage que selon son prix.
- Tous, sauf un, indiquent qu'il considère plutôt le prix que le goût, tout en s'intéressant au goût.
- Le dernier ami indique qu'il ne sait rien au sujet de son ami, hormis la première affirmation.

À partir de ces informations :

- La possibilité qu'il n'ait aucun goût en matière de fromage est nulle, l'état étant impossible (il considère au moins un peu le goût du fromage).
- La possibilité qu'il considère plutôt le prix, ou autant le prix que le goût, est de 1 (rien ne dit que ces états ne peuvent pas arriver).
- La possibilité qu'il favorise le goût au prix, ou seulement le goût, est faible : seul un ami a indiqué que ces états sont possibles (du fait de son ignorance). En supposant que le décideur à 10 amis, nous pouvons attribuer par exemple une possibilité de 0.1.

Nous pouvons alors établir une distribution de possibilité qui est représentée dans la Table 1.12, à laquelle nous rajoutons les probabilités de chaque état.

s	aucun goût	plutôt le prix	équilibré	plutôt le goût	grand amateur
$\pi(s)$	0	1	1	0.1	0.1

TABLE 1.12 – Distribution de possibilité associée aux préférences d'un décideur

Calculons maintenant le degré de possibilité et de nécessité de l'évènement U . En supposant que $U = \{ \text{aucun goût, plutôt le prix, équilibré} \}$, nous avons alors :

- $\Pi(U) = \sup_{s \in U} \pi(s) = \max(\pi(\text{aucun goût}), \pi(\text{plutôt le prix}), \pi(\text{équilibré})) = 1$,
- $N(U) = 1 - \Pi(U^c) = 1 - \max(\pi(\text{plutôt le goût}), \pi(\text{grand amateur})) = 1 - 0.1 = 0.9$.

Le lien avec les probabilités n'est pas évident, mais nous pouvons avoir des bornes sur $P(U)$. En effet, $\Pi(U)$ est une probabilité supérieure, et comme nous avons une distribution de possibilité consistante, $N(U)$ est une probabilité inférieure. Par conséquent, pour notre évènement U spécifique, nous avons $0.9 \leq P(U) \leq 1$.

1.4.3 Principe et opérateurs de fusion

Cette sous-section présente les principes de base de la fusion d'information partielle, ainsi que quelques opérateurs de fusion essentiels (conjonction prudente, disjonction prudente et adaptation). Elle se base particulièrement sur le travail de Dubois *et al.* (2016), détaillant les principes et les opérateurs, et ce, dans différents cadres (général, booléen, ensembles, théorie des possibilités, etc). Dans la Sous-Section 1.4.4, nous présenterons une généralisation de la conjonction et la disjonction pour la théorie des possibilités.

Définition

L'objectif de la fusion d'information est d'obtenir une information synthétique la plus précise possible à partir d'informations incomplètes. Soit $\mathcal{S} = \{S_1, \dots, S_n\}$ un ensemble d'informations chacune cohérente. Chaque élément d'information est formalisé par une théorie mathématique (ensemble, ensemble flou, distribution de probabilité, etc). Un opérateur de fusion n -aire f_n est une application de \mathcal{S}^n à \mathcal{S} avec $f_n(S_1, \dots, S_n) = S$. Nous notons f à la place de f_n quand il n'y a pas d'ambiguïté. De plus, la fusion d'information cherche en général à respecter trois principes :

- Les différentes sources jouent le même rôle : aucune source ne doit être préférée en absence d'information sur leur fiabilité.
- Autant de sources que possible sont considérées comme fiables, pour obtenir le résultat le plus précis que possible. Le résultat doit être fidèle à ce qu'apportent les différentes informations en termes de contenu.
- La fusion d'information doit résoudre les conflits entre les différentes sources, sans en rejeter ou en favoriser sans aucune raison.

Propriétés

Plusieurs propriétés sont intéressantes en fusion d'information :

1. *Unanimité* : si toutes les sources considèrent une situation comme possible, alors le résultat de la fusion aussi. De même, si toutes les sources considèrent une situation comme impossible, alors le résultat de la fusion aussi.

2. *Monotonicit  de l'information* : si un ensemble d'informations \mathcal{S}_1 est moins informatif qu'un ensemble coh rent d'informations \mathcal{S}_2 , alors $f(\mathcal{S}_1)$ ne peut pas  tre plus informatif que $f(\mathcal{S}_2)$.
3. *Maintien de la coh rence* : fusionner des informations coh rentes entre elles doit donner un r sultat coh rent.
4. *Optimisme* : en l'absence d'information sur la fiabilit  des sources, on doit assumer qu'autant de sources que possible doivent  tre fiables, tout en respectant leur coh rence mutuelle ( viter des sources incoh rentes).
5. * quit * : le r sultat de la fusion doit garder quelque chose de chaque information.
6. *Insensibilit    l'ignorance totale* : une source ne donnant aucune information, c'est- -dire l'ignorance totale, ne doit pas influencer le r sultat de la fusion.
7. *Commutativit * : l'ordre des sources ne doit pas changer le r sultat de la fusion.
8. *Engagement minimal* : l'op rateur de fusion doit  tre le moins informatif possible tout en respectant les autres propri t s.

La propri t  d'optimisme stipule qu'on doit exploiter les informations disponibles au maximum, tandis que l'engagement minimal emp che le r sultat de la fusion d' tre plus informatif que les informations disponibles. Les autres propri t s sont assez naturelles, mais comme nous allons le voir dans la suite, certains op rateurs couramment utilis s (notamment la conjonction prudente et la disjonction optimiste) ne respectent pas n cessairement toutes ces propri t s dans toutes les situations.

Conjonction prudente et disjonction optimiste

Ces deux op rateurs de fusion sont essentiels, et couramment utilis s :

D finition 21. Conjonction prudente

Un op rateur conjonctif est une fonction commutative $c_n : \mathcal{S}^n \rightarrow \mathcal{S}$ telle que $c_n(S_1, \dots, S_n) \subseteq S_i, \forall i \in \{1, \dots, n\}$. Cet op rateur conjonctif est dit prudent s'il respecte la propri t  d'engagement minimal.

Cet opérateur est intéressant, car il respecte toutes les propriétés quand les informations sont mutuellement consistantes. Dans le cas contraire, les propriétés de maintien de la cohérence, d'équité et d'engagement minimal ne sont plus respectées.

En logique booléenne, l'opérateur \wedge (et) est une conjonction prudente. En théorie des ensembles, l'intersection (\cap) est aussi une conjonction prudente.

Définition 22. *Disjonction optimiste*

Un opérateur disjonctif est une fonction commutative $\delta_n : \mathcal{S}^n \rightarrow \mathcal{S}$ telle que $S_i \subseteq \delta_n(S_1, \dots, S_n)$, $\forall i \in \{1, \dots, n\}$. Cet opérateur disjonctif est dit optimiste s'il respecte la propriété d'engagement maximal (il doit être le plus informatif possible).

Cet opérateur ne respecte pas la propriété d'optimisme (malgré son nom), d'insensibilité à l'ignorance totale et d'engagement minimal quand les informations sont mutuellement consistantes.

En logique booléenne, l'opérateur \vee (ou) est une disjonction optimiste. En théorie des ensembles, l'union (\cup) est aussi une disjonction optimiste.

Exemple 18. *Illustration de la conjonction prudente et de la disjonction optimiste (en théorie des ensembles)*

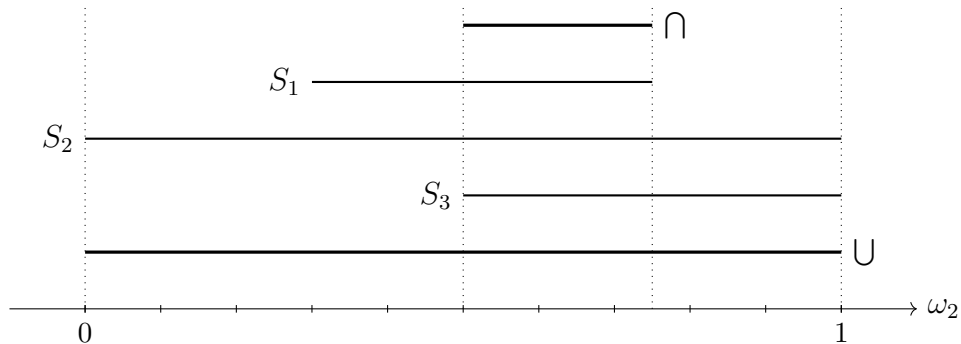


FIGURE 1.14 – Illustration d’une intersection et d’une union d’informations quand les informations sont cohérentes

Sur les Figures 1.14 et 1.15, nous pouvons voir une illustration de la conjonction prudente et de la disjonction optimiste. Dans les deux cas, les trois sources représentées donnent une information sur les préférences du décideur. Le but est d’obtenir l’information la plus précise possible sur la valeur du paramètre ω . Dans

ce contexte particulier, la conjonction prudente correspond à l'intersection et la disjonction optimiste à l'union des différentes informations.

Sur la Figure 1.14, les trois sources sont cohérentes. Par conséquent, l'intersection respecte toutes les propriétés d'un opérateur de fusion et nous permet d'obtenir une bonne estimation de la valeur de ω , avec $\omega_2 \in [0.5, 0.75]$. En se basant sur l'union des sources, on obtient que $\omega_2 \in [0, 1]$. Nous voyons bien l'influence d'une information creuse S_2 sur le résultat de la fusion.

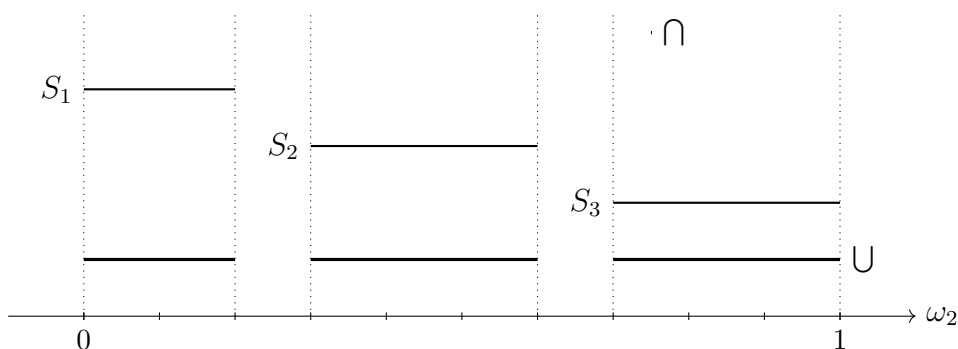


FIGURE 1.15 – Illustration d'une intersection et d'une union d'informations quand les informations sont incohérentes

Sur la Figure 1.15, les trois sources ne sont pas cohérentes. Comme nous pouvons le voir, l'intersection est sensible à l'incohérence et ne permet plus d'avoir un résultat cohérent. A contrario, l'union n'est pas sensible à l'incohérence, et nous obtenons une information synthétique cohérente, avec $\omega_2 \in [0, 2] \cup [3, 6] \cup [7, 10]$.

La conjonction prudente est intéressante quand les informations sont cohérentes, et la disjonction optimiste est intéressante quand les informations sont totalement incohérentes. Nous allons maintenant présenter un nouvel opérateur qui permet de s'adapter à la nature des informations.

Adaptation, Sous-ensembles cohérents maximaux et k -quota

Nous présentons ici plusieurs approches qui permettent de gérer le conflit, c'est-à-dire d'obtenir une information synthétique intéressante malgré des sources contradictoires. Ces approches sont introduites et rapidement illustrées dans ce paragraphe, mais elles ne seront utilisées que dans le Chapitre 4.

Opérateur adaptatif : nous commençons par présenter l'opérateur adaptatif dans le cas où nous n'avons que deux informations.

Définition 23. *Adaptatif (deux informations)*

Un opérateur adaptatif est défini à partir d'un opérateur conjonctif prudent c_2 et d'un opérateur disjonctif optimiste δ_2 :

$$a(S_1, S_2) = \begin{cases} c_2(S_1, S_2) & \text{si les informations sont consistantes,} \\ \delta_2(S_1, S_2) & \text{sinon.} \end{cases}$$

Les opérateurs adaptatifs respectent toutes les propriétés des opérateurs de fusion. Si les informations sont totalement cohérentes, nous retrouvons un opérateur conjonctif prudent. Si elles sont totalement incohérentes, nous retrouvons un opérateur disjonctif optimiste.

Remarquons que dans le travail de Dubois *et al.* (2016), deux opérateurs adaptatifs sont présentés (prudent et imprudent) en fonction de si on ne garde que les valeurs totalement possibles, ou aussi les valeurs partiellement possibles. Cela n'a pas d'importance quand les valeurs sont dans $\{0, 1\}$ (logique booléenne, théorie des ensembles classique), mais en logique multivaluée et dans les ensembles flous, cette différence est importante.

Sous-ensembles cohérents maximaux : en se plaçant dans un contexte ensembliste, la généralisation des opérateurs adaptatifs pour n éléments se fait à partir des sous-ensembles cohérents maximaux (MCS) (Rescher et Manor, 1970) :

Définition 24. *MCS*

Un sous-ensemble cohérent maximum $c \subseteq \{1, \dots, n\}$ est un ensemble d'indices respectant les deux contraintes suivantes :

- Cohérence : $\bigcap_{i \in c} S_i \neq \emptyset$,
- Maximalité : $\forall j \notin c, (\bigcap_{i \in c} S_i) \cap S_j = \emptyset$.

Remarquons que nous avons plusieurs MCSs dès que les sources sont incohérentes entre elles.

Définition 25. *Opérateur de fusion basé sur les MCSs*

Notons \mathcal{C} l'ensemble des sous-ensembles cohérents maximums d'un ensemble $\mathcal{S} = \{S_1, \dots, S_n\}$. L'opérateur de fusion basé sur les MCSs est une généralisation d'un opérateur adaptatif à n éléments défini par :

$$f_n^{MCS}(\mathcal{S}_1, \dots, \mathcal{S}_n) = \bigcup_{c \in \mathcal{C}} \bigcap_{i \in c} S_i. \quad (1.26)$$

Nous retrouvons dès lors le comportement adaptatif : quand plusieurs sources sont cohérentes, nous appliquons un opérateur conjonctif prudent (intersection). Un opérateur disjonctif optimiste (union) est alors appliqué sur les différents résultats partiels. Toutes les propriétés des opérateurs de fusion sont respectées. Nous reviendrons sur l'utilisation et le calcul des MCSs dans le Chapitre 4.

Exemple 19. *MCS*

Supposons que nous avons quatre sources qui nous donnent des informations sur les préférences du décideur $\mathcal{S} = \{S_1, S_2, S_3, S_4\}$. Sur la Figure 1.16, nous pouvons voir que les informations ne sont pas cohérentes entre elles, notamment S_1 et S_2 , ainsi que S_1 et S_3 .

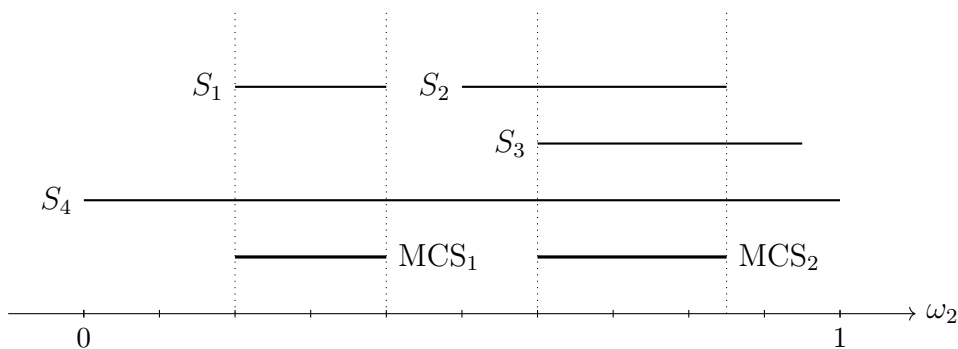


FIGURE 1.16 – MCS de sources incohérentes

Les différents MCSs sont aussi représentés sur la Figure 1.16. Remarquons que, par exemple, $\{2, 3\}$ n'est pas un MCS, mais seulement un sous-ensemble cohérent, car il est possible d'ajouter l'information S_4 tout en gardant la cohérence. Par opposition, $\{1, 2, 3, 4\}$ n'est pas un MCS, car S_1 et S_2 ne sont pas cohérentes entre elles, et par conséquent forment une conjonction vide. Il n'y a que deux MCSs : $\{1, 4\}$ et $\{2, 3, 4\}$. Finalement, en appliquant l'opérateur de fusion basé sur les MCSs, nous obtenons $\omega_2 \in [0.2, 0.4] \cup [0.6, 0.85]$.

k -quota : une autre solution permettant de combler les lacunes de l'opérateur conjonctif (contradictions quand beaucoup de sources) et de l'opérateur disjonctif optimiste (peu informatif quand beaucoup de sources) est d'utiliser un opérateur de fusion dit de k -quota (Dubois *et al.*, 1999; Everaere *et al.*, 2010) :

Définition 26. *k -quota*

Soit n informations. Soit k le nombre maximal d'informations mutuellement consistantes. Un opérateur de k -quota consiste en les opérations suivantes :

- *Appliquer un opérateur conjonctif sur les ensembles de k sources cohérentes.*
- *Appliquer un opérateur disjonctif sur les résultats partiels.*

Nous pouvons voir un lien avec les MCSs, à la seule différence que les MCSs peuvent être de taille variable, alors qu'ici les ensembles contiennent k sources cohérentes, k étant le plus grand possible pour respecter le principe d'optimisme. Remarquons que cet opérateur ne respecte pas la propriété d'équité, certaines sources pouvant être oubliées : si une ou plusieurs sources ne sont pas dans des ensembles cohérents de taille k , alors ces informations seront oubliées, brisant la propriété d'équité. Une illustration se trouve dans l'Exemple 20

Exemple 20. *k -quota*

En reprenant directement l'Exemple 19, nous remarquons que le nombre maximum d'informations mutuellement cohérentes est de 3 (S_2, S_3 et S_4). Par conséquent, l'opérateur de fusion à partir des k -quotas, avec $k = 3$, ne va considérer que les ensembles cohérents de trois sources. Le seul ensemble cohérent de trois sources est $\{2, 3, 4\}$. Nous voyons bien que la première source est oubliée par l'opérateur de fusion, brisant la propriété d'équité. Finalement, l'opérateur de fusion à partir des k -quotas, avec $k = 3$, va nous indiquer que $\omega_2 \in [0.6, 0.85]$ (omettant l'information portée par la première source $\omega_2 \in [0.2, 0.4]$).

Tous les opérateurs que nous avons présentés dans cette sous-section sont généralisables à la théorie des possibilités. Dans la sous-section qui suit, nous allons généraliser la conjonction et la disjonction pour la théorie des possibilités⁷.

7. L'article de Dubois *et al.* (2016) propose plusieurs opérateurs de fusion en théorie des possibilités, notamment pour généraliser les MCSs. Dans nos travaux, nous n'avons pas besoin de ces généralisations des MCSs.

1.4.4 Fusion en théorie des possibilités : T-normes et T-conormes

Dans cette sous-section, nous nous intéressons à deux opérations binaires qui nous permettent de généraliser la conjonction et la disjonction en logique multivaluée, et par extension à la théorie des possibilités. Ces opérateurs sont les normes triangulaires (T-normes) pour la conjonction, et les conormes triangulaires (T-conormes) pour la disjonction (voir Klement *et al.* (2000) pour une introduction).

Normes triangulaires

Définition 27. T-norme

Une T-norme est une opération binaire $T : [0, 1]^2 \rightarrow [0, 1]$ qui $\forall a, b, c \in [0, 1]^3$ respecte les axiomes suivants :

- Commutativité : $T(a, b) = T(b, a)$,
- Monotonie : $T(a, b) \leq T(a, c) \forall b \leq c$,
- Associativité : $T(a, T(b, c)) = T(T(a, b), c)$,
- Condition de limite : $T(a, 1) = a$.

La T-norme est une opération algébrique binaire sur l'intervalle $[0, 1]$. Dans la littérature, cette opération est notée soit avec l'expression préfixée T, soit avec l'opérateur binaire infixé $*$: $T(a, b) = a * b$.

Initialement, les T-normes ont été introduites comme une généralisation de l'inégalité triangulaire à des espaces métriques probabilistes (Menger, 1942) avec des axiomes différents. Les axiomes présentés précédemment sont liés à des travaux plus récents (Schweizer et Sklar, 1958, 1960; Schweizer et Sklar, 1961).

Remarquons qu'il existe un nombre infini de T-normes, et plusieurs familles. Ici, nous présentons que 4 T-normes classiques dans la littérature :

- Minimum : $T_M(a, b) = \min(a, b)$,
- Produit : $T_P(a, b) = a.b$,
- Łukasiewicz : $T_L(a, b) = \max(a + b - 1, 0)$,
- Produit drastique : $T_D(a, b) = \begin{cases} 0 & \text{si } (a, b) \in [0, 1]^2, \\ \min(x, y) & \text{sinon.} \end{cases}$

Les contours en 3D de ces T-normes sont représentées sur la Figure 1.17⁸. Ces T-normes sont intéressantes pour plusieurs raisons :

- T_M est la plus grande (par point) des T-normes et sert en logique floue pour les conjonctions faibles. Elle est continue.
- T_P sert en logique floue pour les conjonctions fortes. Elle est stricte (continue et monotone) et Archimédienne⁹.
- T_L sert aussi en logique floue pour les conjonctions fortes, mais dans le cadre de la logique de Łukasiewicz. $T_L(a, b) \leq T_P(a, b)$ et elle est continue, nilpotente¹⁰ et Archimédienne.
- T_D est la plus petite (par point) des T-normes. Elle n'est pas continue (seulement continue à droite) et elle est Archimédienne.

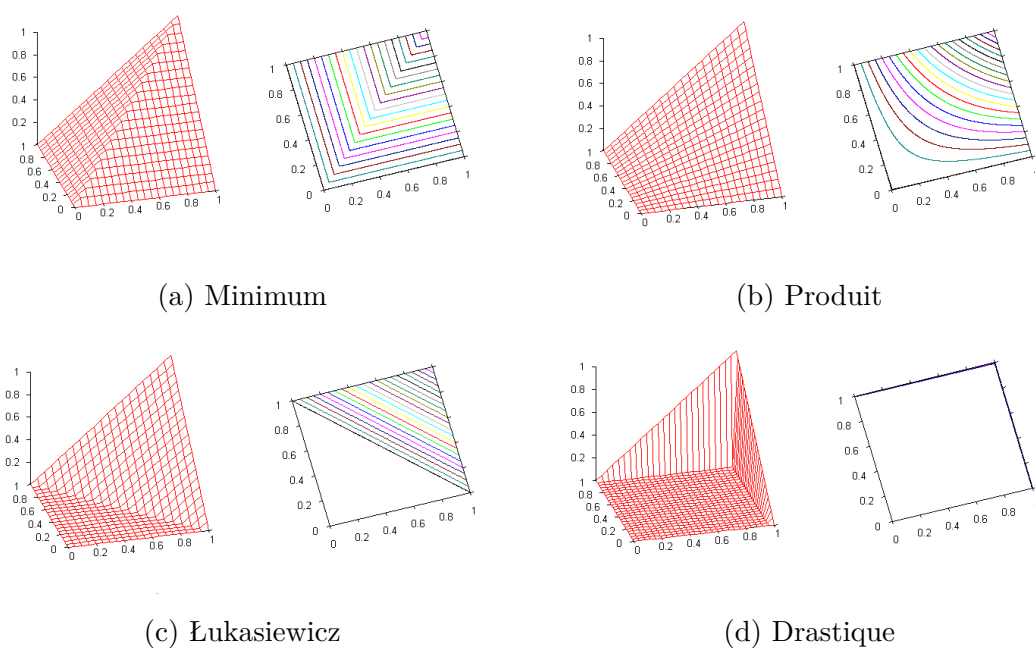


FIGURE 1.17 – T-normes usuelles

8. Source des images : Wikipédia.

9. $\forall (a, b) \in]0, 1[^2$ avec $a > b$, $\exists n \in \mathbb{N}$ tel que $a * a * \dots * a$ (n fois) $< b$.

10. $\forall a \in]0, 1[$, $\exists n \in \mathbb{N}$ tel que $a * a * \dots * a$ (n fois) $= 0$.

Pour toute T-norme T_i , nous avons $T_D(a, b) \leq T_i(a, b) \leq T_M(a, b)$. T_M , T_P et T_L sont aussi des copules¹¹ et pour toute copule C nous avons, $T_L \leq C \leq T_M$ (limites de Fréchet-Hoeffding)

Exemple 21. *Calcul de T-normes*

Prenons deux informations S_1 et S_2 telles que :

$$S_1 = \begin{cases} 1 & \text{si } \omega_2 \in [0.2, 0.5], \\ 0.7 & \text{si } \omega_2 \in]0.5, 1], \\ 0 & \text{sinon.} \end{cases} \quad S_2 = \begin{cases} 1 & \text{si } \omega_2 \in [0.3, 0.6], \\ 0.5 & \text{si } \omega_2 \in [0, 0.3[\cup]0.6, 1], \\ 0 & \text{sinon.} \end{cases}$$

Pour faire une conjonction de ces deux informations, nous utilisons les 4 T-normes usuelles, comme représentées sur la Figure 1.20. Nous pouvons voir que nous avons $\forall s_1 \in S_1, \forall s_2 \in S_2, T_M(s_1, s_2) \geq T_P(s_1, s_2) \geq T_L(s_1, s_2) \geq T_D(s_1, s_2)$. Par ailleurs, les différentes T-normes sont différentes pour des valeurs de s_1 et s_2 différentes de 0 et de 1. Notons que si $s_1 + s_2 < 1$ pour $\omega_2 \in]0.6, 0.1]^2$, nous aurions eu $T_L(s_1, s_2) = T_D(s_1, s_2)$

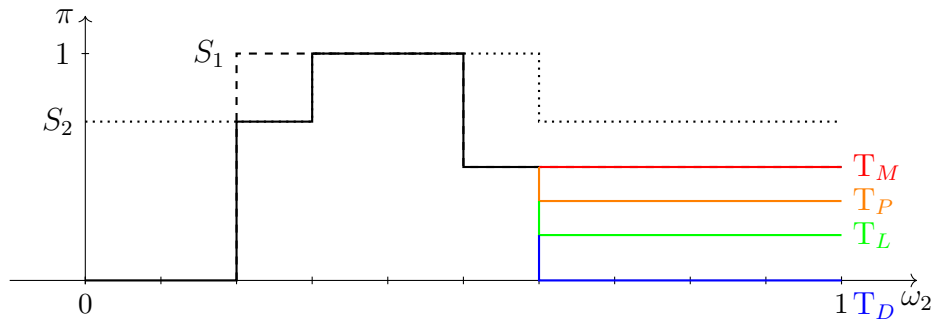


FIGURE 1.18 – Comparaison des T-normes usuelles

11. Les copules sont des fonctions de répartition définies sur $[0, 1]^d$ et dont les marges sont uniformes sur $[0, 1]$. Remarquons que toutes les copules ne sont pas des T-normes et toutes les T-normes ne sont pas de copules, mais un lien fort existe entre les deux. Pour plus d'information, nous vous renvoyons au chapitre 9 de Klement *et al.* (2000) et à Nelsen (1999).

Conormes triangulaires

Les T-conormes ont été introduites par Schweizer et Sklar (1961) comme l'opération duale des T-normes.

Définition 28. T-conorme

Une T-conorme est une opération binaire $\perp : [0, 1]^2 \rightarrow [0, 1]$ qui $\forall a, b, c \in [0, 1]^3$ respecte les axiomes suivants :

- Commutativité : $\perp(a, b) = \perp(b, a)$,
- Monotonie : $\perp(a, b) \leq \perp(a, c) \forall b \leq c$,
- Associativité : $\perp(a, \perp(b, c)) = \perp(\perp(a, b), c)$,
- Condition de limite : $\perp(a, 0) = a$.

Remarquons ainsi que les T-normes et les T-conormes ne diffèrent axiomatiquement que par la condition de limite (pour les T-normes : 1, pour les T-conormes : 0). Nous avons ensuite le lien suivant :

Proposition 3. Dualité entre T-normes et T-conormes

Une fonction $\perp : [0, 1]^2 \rightarrow [0, 1]$ est une T-conorme si et seulement s'il existe une T-norme T telle que $\forall a, b \in [0, 1]^2$:

$$\perp(a, b) = 1 - T(1 - a, 1 - b). \quad (1.27)$$

De même, une fonction $T : [0, 1]^2 \rightarrow [0, 1]$ est une T-norme si et seulement s'il existe une T-conorme \perp telle que $\forall a, b \in [0, 1]^2$:

$$T(a, b) = 1 - \perp(1 - a, 1 - b). \quad (1.28)$$

Démonstration. Trivialement, nous voyons que si les axiomes de commutativité, de monotonie et d'associativité sont respectés pour la T-norme (réciproquement pour la T-conorme), alors ils le sont aussi pour la T-conorme (réciproquement pour la T-norme).

Pour vérifier le dernier axiome pour les T-conormes, si $b = 0$ nous avons $\perp(a, 0) = 1 - T(1 - a, 1 - 0) = 1 - T(1 - a, 1) = 1 - (1 - a) = a$, confirmant que l'axiome sur la condition de limite est respecté. Le même raisonnement peut être appliqué pour $b = 1$ avec les T-normes, concluant ainsi la preuve. \square

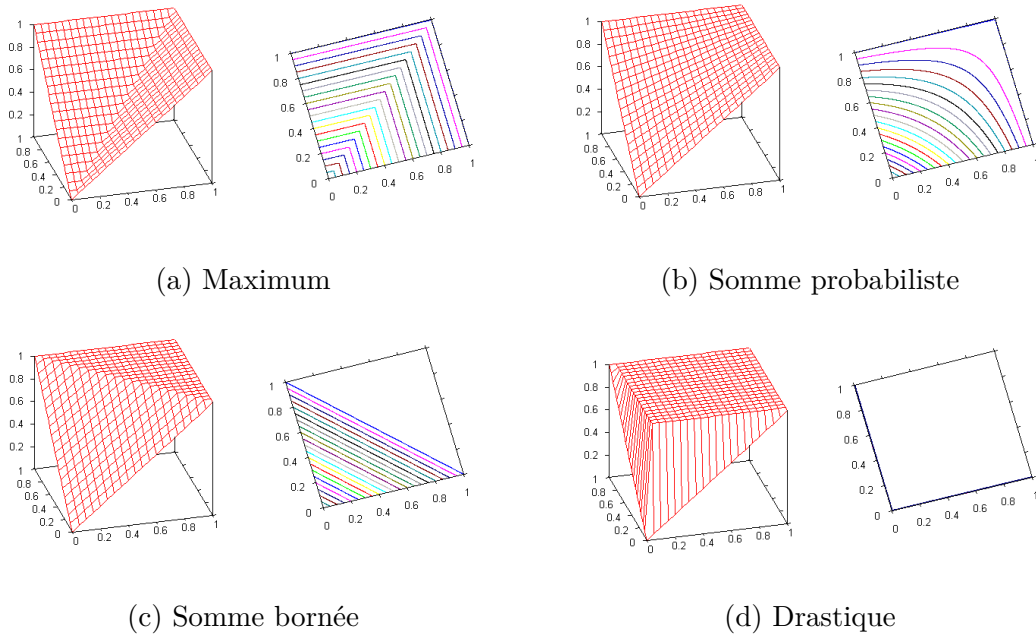


FIGURE 1.19 – T-conormes usuelles

Remarquons qu'il existe aussi un nombre infini de T-conormes. De nouveau, nous présentons que 4 T-conormes classiques dans la littérature, qui sont en fait les duales des T-normes usuelles :

- Maximum : $\perp_M(a, b) = \max(a, b)$,
- Somme probabiliste : $\perp_P(a, b) = a + b - a.b = 1 - (1 - a).(1 - b)$,
- Somme bornée/Lukasiewicz : $\perp_L(a, b) = \min(a + b, 1)$,
- Somme drastique : $\perp_D(a, b) = \begin{cases} 1 & \text{si } (a, b) \in]0, 1]^2, \\ \max(x, y) & \text{sinon.} \end{cases}$

Les contours en 3D de ces T-conormes sont représentées sur la Figure 1.19¹². Ces T-conormes ont des propriétés analogues aux T-normes classiques, que nous ne rappellerons pas ici. Remarquons néanmoins :

- \perp_M sert en logique floue pour les disjonctions faibles. \perp_P et \perp_L servent en logique floue pour les disjonctions fortes.

12. Source des images : Wikipédia.

- \perp_M est la T-conorme la plus petite (par point), tandis que \perp_D est la plus grande (par point). Nous avons alors pour toute T-conorme T_i , $\perp_D(a, b) \geq T_i(a, b) \leq \perp_M(a, b)$. De plus, $\perp_P(a, b) \leq \perp_L(a, b)$.

Exemple 22. *Calcul de T-conormes*

Prenons deux informations S_3 et S_4 telles que :

$$S_3 = \begin{cases} 1 & \text{si } \omega_2 \in [0.2, 0.5], \\ 0.3 & \text{si } \omega_2 \in]0.5, 1], \\ 0 & \text{sinon.} \end{cases} \quad S_4 = \begin{cases} 1 & \text{si } \omega_2 \in [0.3, 0.6], \\ 0.5 & \text{si } \omega_2 \in [0, 0.3[\cup]0.6, 1], \\ 0 & \text{sinon.} \end{cases}$$

Pour faire une disjonction de ces deux informations, nous utilisons les 4 T-conormes usuelles, comme représentées sur la Figure 1.20. Nous pouvons voir que nous avons $\forall s_1 \in S_1, \forall s_2 \in S_2, \perp_D(s_1, s_2) \geq \perp_L(s_1, s_2) \geq \perp_P(s_1, s_2) \geq \perp_M(s_1, s_2)$. Par ailleurs, les différentes T-conormes sont différentes pour des valeurs de s_1 et s_2 différentes de 0 et de 1. Notons que les informations sont différentes de l'Exemple 21, sinon nous aurions eu pour $\omega_2 \in]0.6, 0.1]^2$, $\perp_L(s_1, s_2) = \perp_D(s_1, s_2)$ (car $s_1 + s_2 > 1$).

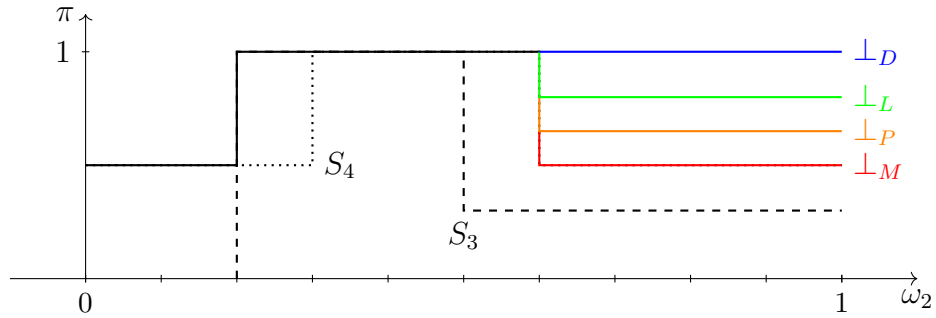


FIGURE 1.20 – Comparaison des T-conormes usuelles

1.5 Conclusion

Dans ce chapitre, nous avons tout d'abord vu comment modéliser des problèmes multicritères, notamment quand un décideur doit faire un choix entre plusieurs alternatives, chacune ayant plusieurs facteurs (prix, qualité, etc). Dans le cas général, il n'est pas possible de toutes les comparer, comme les critères sont souvent contradictoires (augmenter un critère fait diminuer les autres). Pour régler ce problème, nous avons présenté plusieurs fonctions d'agrégation de critères qui attribuent un score pour chaque alternative, ce qui permet de les comparer. Par ailleurs, nous avons aussi montré qu'il est possible de prendre une décision sous incertitude totale, par exemple quand le score des alternatives dépend d'évènements dont nous ne connaissons pas la probabilité.

Ensuite, nous avons recensé différents algorithmes d'élicitation pour formaliser les préférences du décideur, et ainsi l'aider à faire un choix entre plusieurs alternatives. Nous avons porté notre attention sur l'élicitation incrémentale, celle-ci permettant de minimiser la charge cognitive du décideur tout en interagissant directement avec lui. Plus particulièrement, nous nous intéressons à un algorithme se basant sur l'optimisation d'un regret Minimax, apportant de fortes garanties sur la recommandation finale, et sur l'approche CSS, qui permet de sélectionner des questions (comparaison de deux alternatives) plutôt efficaces. Néanmoins, cette approche ne nous permet pas de détecter une incohérence, et si le décideur se trompe dans ses réponses, ou que l'on a mal modélisé les préférences du décideur, l'algorithme converge potentiellement vers une mauvaise recommandation.

Enfin, nous avons présenté les probabilités imprécises, une alternative aux probabilités pour modéliser l'incertitude, qui permettent de répondre à plusieurs critiques des probabilités. Nous avons notamment étudié la théorie des possibilités, qui est une solution relativement simple et efficace pour modéliser l'incertitude. Nous avons par la suite analysé différents agrégateurs qui permettent de fusionner différentes informations partielles, et possiblement incertaines, en une seule information synthétique. Nous avons aussi présenté des généralisations de ces outils à des informations multivaluées, permettant ainsi de fusionner des informations résultantes de la théorie des possibilités.

Points importants vus dans ce chapitre :

- Prendre une décision entre plusieurs alternatives multicritères, y compris sous incertitude totale.
- Présenter des algorithmes d'élicitation pour formaliser les préférences du décideur. Le choix se porte sur un minimisant la charge cognitive.
- Modéliser l'incertitude avec la théorie des possibilités et fusionner différentes informations possiblement incertaines.

Suite : l'objectif est de détecter des incohérences dans l'élicitation, tout en gardant les avantages de l'algorithme incrémental basé sur l'approche CSS (des garanties fortes sur la qualité de la recommandation, une convergence vers cette recommandation, une faible charge cognitive, etc). Pour cela, nous allons appliquer les différents éléments que nous avons vus dans ce chapitre, notamment la théorie des possibilités qui permet de modéliser une incertitude dans les réponses du décideur.

Élicitation possibiliste de préférences

2.1 Introduction

La modélisation des préférences d'individus ou de populations est un sujet de longue date dans des domaines tels que l'économie, les statistiques, l'intelligence artificielle et la recherche opérationnelle. Cependant, les problématiques liées au traitement de l'incertitude, de l'incohérence et d'informations incomplètes sont toujours activement discutées, en particulier dans l'intelligence artificielle (Fürnkranz et Hüllermeier, 2011; Viappiani et Boutilier, 2020).

Dans ce chapitre, nous nous intéressons à la gestion de l'incertitude et de l'incohérence lors de l'élicitation des préférences d'un seul décideur. Ce type de problème est souvent rencontré dans l'aide à la décision multicritère ou l'intelligence artificielle, tandis que d'autres domaines tels que les statistiques, l'apprentissage automatique ou l'économie tendent à se concentrer sur l'apprentissage à partir de populations. Il existe dans la littérature spécifique à notre contexte deux types principaux d'approche pour traiter l'incertitude et l'incohérence (par rapport au modèle de préférence choisi) :

Approches ensemblistes

Le premier type d'approche consiste à utiliser des méthodes robustes, basées sur des ensembles, dans lesquelles nous considérons des ensembles de modèles

potentiels et faisons des déductions à partir de ces ensembles. Chaque information sur les préférences réduit l'ensemble des modèles possibles, cet ensemble mis à jour représentant notre incertitude quant au modèle optimal. Ces approches sont similaires au concept d'espace de version dans l'apprentissage automatique, comme l'ont noté certains auteurs (Prade *et al.*, 2009a).

Lorsque les informations disponibles aboutissent à un ensemble de modèles possibles, différentes stratégies d'inférence peuvent être utilisées. L'une d'entre elles consiste à faire des inférences qui soient vraies pour tous les modèles possibles, ce qui donne lieu à des méthodes dites robustes (Corrente *et al.*, 2013) qui peuvent parfois fournir des résultats partiels, en recommandant par exemple plusieurs alternatives. Une autre méthode consiste à faire des déductions précises en présence d'incertitude, généralement en adoptant une attitude prudente à partir des critères du Maximin ou du regret Minimax (Boutilier, 2013; Benabbou *et al.*, 2017; Benabbou et Perny, 2018). Ces approches prennent en compte les pires scénarios pour recueillir de nouvelles informations sur les préférences et fournissent des outils d'inférence (par exemple, pour choisir une alternative à recommander) avec de fortes garanties. Nous nous concentrerons sur cette dernière approche dans nos expériences, car il est plus facile de comparer différentes approches lorsqu'il s'agit de faire des inférences et des recommandations précises.

Il convient de noter que les approches d'élicitation utilisant le regret Minimax évitent les incohérences en recueillant des informations préférentielles qui sont nécessairement cohérentes avec les réponses précédentes et la famille de modèles choisie (voir par exemple (Benabbou *et al.*, 2017)). Toutefois, cette approche repose sur des hypothèses très fortes, et notamment sur le fait que le décideur et l'analyste sont supposés ne pas commettre d'erreur respectivement dans leurs réponses ou dans la sélection du modèle. Nous avons évoqué ces approches ensemblistes dans la Section 1.3.

Approches probabilistes

La deuxième approche pour traiter les incohérences consiste à trouver un modèle qui minimise une certaine forme d'erreur moyenne. Cela peut être réalisé par des techniques d'optimisation telles que la minimisation des moindres carrés

(Grabisch *et al.*, 2008), la maximisation de la marge (Teso *et al.*, 2017), ou en utilisant les probabilités bayésiennes pour maintenir une distribution de probabilité sur les modèles possibles et la mettre à jour lorsque de nouvelles informations sont disponibles, tout en faisant des inférences à l'aide de l'opérateur d'espérance classique (Eric *et al.*, 2007; Viappiani et Boutilier, 2010; Bourdache *et al.*, 2019).

Bien que ces approches soient efficaces pour traiter les incohérences, elles ne peuvent généralement pas fournir le même niveau de garanties fortes que les approches basées sur les ensembles. Elles n'offrent que des garanties statistiques ou en espérance, à condition que le modèle d'incertitude utilisé soit correct. Ceci est particulièrement contraignant, car les modèles à base de probabilité nécessitent souvent une grande quantité de données ou un fort biais inductif pour la spécification, ainsi que des outils statistiques pour la validation, ce qui peut être difficile à concilier avec l'apprentissage des préférences individuelles. Par exemple, une hypothèse courante est que la probabilité d'erreur est proportionnelle à la distance entre deux alternatives (Bourdache *et al.*, 2019), mais cela n'est pas toujours évident. Il peut être plus facile de comparer des solutions similaires (par exemple, celles qui ne diffèrent que sur deux critères (Hammond *et al.*, 1998)) que des solutions très différentes. En outre, les approches probabilistes ne font pas de distinction claire entre les situations avec et sans incohérences, car les masses de probabilité doivent être partagées entre plusieurs hypothèses concurrentes et exclusives.

Exemple 23. *Approches probabilistes et conflit*

Nous reprenons l'ensemble des alternatives disponibles \mathbb{X} de l'Exemple 1 en supposant que les préférences du décideur soient modélisées par une moyenne pondérée de paramètre $\omega^ = (0.75, 0.25)$. Imaginons que le décideur réponde dans un premier temps qu'il préfère MO à EM avec une probabilité de 0.6, et qu'il réponde dans un deuxième temps qu'il préfère BT à EM avec une probabilité de 0.7. Nous avons alors les frontières de décision $\omega_1 \geq \omega_2$ et $4\omega_1 \geq 5\omega_2$. Les deux réponses sont cohérentes entre elles et en adéquation avec les préférences du décideur. La Table 2.1 indique les différents éléments pour déterminer la distribution de probabilité $p(\omega)$. La Figure 2.1 représente $p(\omega)$.*

\cap	BT \succ EM	BT \preceq EM
MO \succ EM	$[\frac{5}{9}, 1]$	$[0.5, \frac{5}{9}]$
MO \preceq EM	\emptyset	$[0, 0.5]$

$p(\omega)$	BT \succ EM	BT \preceq EM
MO \succ EM	$0.6 \times 0.7 = 0.42$	$0.6 \times 0.3 = 0.18$
MO \preceq EM	$0.4 \times 0.7 = 0.28$	$0.4 \times 0.3 = 0.12$

TABLE 2.1 – Détermination de la distribution de probabilité $p(\omega)$ de l'Exemple 23

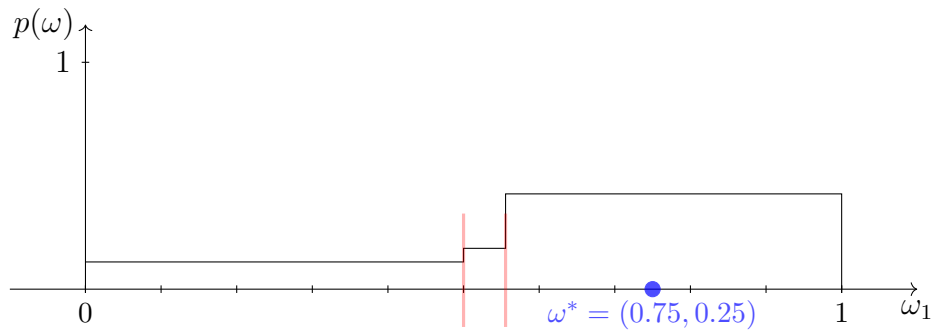


FIGURE 2.1 – Distribution de probabilité $p(\omega)$ de l'Exemple 23

Nous remarquons que $p(\emptyset) = 0.4 \times 0.7 = 0.28$, ce qui indique que nous avons un conflit non nul, alors que les réponses sont cohérentes entre elles et en adéquation avec le modèle de préférence initial. De plus, la distribution de probabilité n'est pas normalisée : $\int_0^1 p(\omega) d\omega \approx 0.26 < 1$, alors que pour chaque réponse prise individuellement, cette intégrale est égale à 1. Cela montre que la théorie des probabilités ne permet pas de faire de distinction claire entre des situations avec et sans incohérences, comme du conflit sera observé quoi qu'il arrive dès que nous avons plusieurs informations préférentielles.

Solution basée sur la théorie des possibilités

Dans ce chapitre, mais aussi tout le long de ce manuscrit, nous étudions une troisième approche consistant à utiliser les distributions de possibilités pour

modéliser des informations préférentielles incertaines, et la théorie des possibilités (Dubois et Prade, 1998) pour raisonner avec les informations fournies. L'intérêt d'une telle approche est qu'elle reste cohérente avec une approche ensembliste, puisque les ensembles sont une instance spéciale des distributions de possibilités, tout en produisant une quantification non binaire de l'incohérence. Elle met également à notre disposition divers outils pour traiter une telle incohérence, ces outils étendant les opérations de la théorie des ensembles et de la logique telles que la conjonction et la disjonction plutôt que les opérateurs basés sur l'espérance, comme nous l'avons vu dans la Section 1.4. Ces approches possibilistes constituent donc une extension intéressante de l'approche ensembliste, tout en offrant un point de vue sur le traitement de l'incertitude qui diffère de celui des probabilités (car la théorie des possibilités et la théorie des probabilités ne subsument ni l'une ni l'autre). Dans ce chapitre, nous nous intéressons seulement à la méthode et à la détection d'incohérence. La gestion de l'incohérence détectée est traitée dans le Chapitre 4.

Remarquons qu'une approche basée sur les probabilités imprécises (la théorie des possibilités étant un cadre spécifique, comme nous l'avons vu dans la Sous-section 1.4.1) n'est pas totalement nouvelle, Guillot et Destercke (2019) ayant déjà proposé une première approche basée sur les fonctions de croyance. Néanmoins, nous verrons dans la Section 2.3 les avantages d'une approche basée sur la théorie des possibilités par rapport à une approche basée sur les fonctions de croyance.

Nous introduisons plusieurs éléments nécessaires pour présenter notre méthode et son intérêt dans la Section 2.2. Nous discutons ensuite de notre méthode possibiliste et de son intérêt dans la Section 2.3. Enfin, la Section 2.4 présente quelques expériences simulées qui confirment l'intérêt potentiel de notre approche.

2.2 Élicitation avec de l'incertitude

2.2.1 Erreurs dans les approches ensemblistes

Comme nous l'avons évoqué dans l'introduction, les approches ensemblistes sont très utiles et peuvent fournir des garanties fortes concernant les inférences, à condition que les informations fournies par le décideur soient correctes. Toutefois, ces hypothèses ne sont pas réalistes dans de nombreuses applications, et il est

souhaitable de tenir compte des erreurs possibles (dans les réponses du décideur ou dans le choix du modèle de préférence par l'analyste) grâce à une modélisation raffinée de l'incertitude. Dans l'Exemple 14 nous avons déjà montré qu'une erreur peut négativement influencer un algorithme d'élicitation. Prenons un exemple similaire, mais en se plaçant en élicitation incrémentale avec l'heuristique CSS, comme nous l'avons présenté dans la Sous-Section 1.3.3 :

Exemple 24 (Une simple erreur ruine l'élicitation incrémentale). *À partir de l'ensemble des alternatives disponibles \mathbb{X} de l'Exemple 1, supposons que les préférences du décideur soient modélisées par une moyenne pondérée de paramètre $\omega^* = (0.75, 0.25)$. Supposons que le décideur doit comparer les alternatives BT et EM et en choisir une. Le décideur, par manque de concentration ou par incertitude, déclare par erreur que $BT \preceq EM$. Nous avons alors la frontière de décision suivante :*

$$WS_\omega(BT) \leq WS_\omega(EM) \Rightarrow 8\omega_1 + \omega_2 \leq 4\omega_1 + 6\omega_2 \Rightarrow 4\omega_1 \leq 5\omega_2.$$

Nous pouvons voir sur la Figure 2.2 qu'après cette affirmation du décideur, le paramètre ω^* , correspondant aux vraies préférences du décideur, n'est pas dans le sous-ensemble des valeurs possibles de ω , noté $E_1 \subseteq \Omega$. Or, comme nous nous plaçons dans le cadre de l'élicitation incrémentale à partir de l'heuristique CSS, nous aurons toujours après chaque nouvelle question un nouveau sous-ensemble de paramètres possibles inclus dans E_1 ($E_{k+1} \subseteq E_k \forall k$, cf. la Sous-Section 1.3.3), et donc ω^* sera toujours exclu de E_k . Ainsi, l'estimateur $\hat{\omega}^*$ converge irrémédiablement vers une mauvaise valeur de ω^* , quelles que soient les réponses futures du décideur.

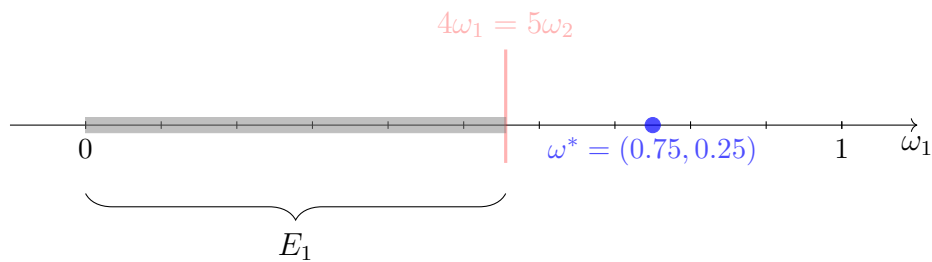


FIGURE 2.2 – Une mauvaise réponse amène à une mauvaise estimation de ω^* dans l'Exemple 24

Les approches probabilistes sont une solution à ce problème, mais elles ne fournissent pas les mêmes garanties que celles de la Proposition 2, comme nous l'avons indiqué dans l'introduction de ce chapitre, en ne fournissant que des garanties en espérance ou statistiques. En outre, ces approches ne remettent généralement pas en question la famille de modèles, car toutes les probabilités sont renormalisées après chaque réponse du décideur : la renormalisation empêche d'identifier la présence d'une incohérence, quelle qu'en soit la source. L'une de nos principales motivations dans ce travail est d'abandonner les hypothèses fortes de l'approche robuste, tout en préservant ses garanties solides. Comme nous allons le voir par la suite, une solution se base sur une modélisation possibiliste des préférences du décideur, en lui demandant (ou à l'analyste) de fournir un niveau de confiance sur ses préférences.

2.2.2 Compléments en théorie des possibilités

Dans la Sous-section 1.4.2 nous avons déjà introduit la théorie des possibilités. Ici, nous complétons cette présentation avec une notion importante pour la suite de ce chapitre, celle de la coupe-alpha :

Définition 29. *Coupe-alpha*

Soit une distribution de possibilité π . Sa coupe-alpha π^α est le sous-ensemble défini par :

$$\pi^\alpha = \{\omega \in \Omega : \pi(\omega) \geq \alpha\}. \quad (2.1)$$

Cette coupe-alpha contient tous les éléments ω qui ont un degré de possibilité supérieur à α .

Exemple 25. *Coupe-alpha*

Supposons que nous avons de nouveau des informations sur les préférences en matière de fromage du décideur, et que nous obtenons la distribution de possibilité présentée dans la Figure 2.3. Supposons maintenant que nous voulons une coupe-alpha de cette distribution de possibilité avec $\alpha = 0.5$. Nous avons alors $\pi^{0.5} = [0, 0.6]$. En effet, ces éléments ont un degré de possibilité supérieur à 0.5 : $\pi(\omega) = 0.6$ quand $\omega \in [0, 0.2[$ et $\pi(\omega) = 1$ quand $\omega \in [0.2, 0.6]$. Par contre, quand $\omega \in]0.6, 1]$,

alors le degré de possibilité est inférieur à 0.5, expliquant la non-inclusion de ces valeurs dans $\pi^{0.5}$.

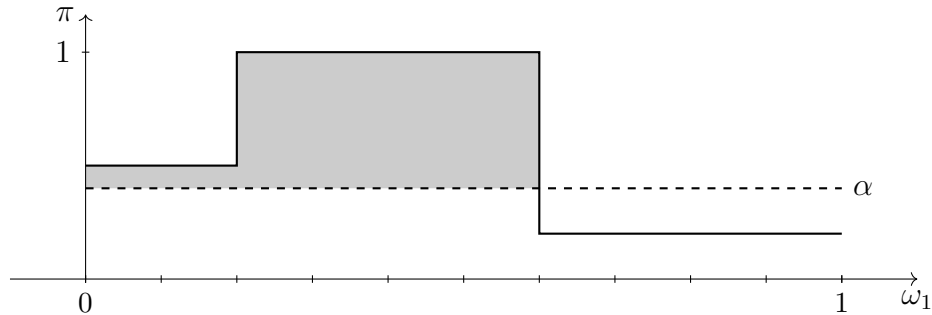


FIGURE 2.3 – Exemple d'une coupe-alpha avec $\alpha = 0.5$

Dans la suite de ce Chapitre, ainsi que dans les suivants, nous considérons des distributions non normalisées π telles que $\max_{\omega \in \Omega} \pi_{\omega} < 1$. Nous pouvons obtenir un indicateur sur l'incohérence :

Définition 30. *Degré d'incohérence*

Soit une distribution de possibilité non normalisée π . L'incohérence de l'information disponible est quantifiée par l'indicateur $Inc(\pi)$, défini par :

$$Inc(\pi) = 1 - \max_{\omega \in \Omega} \pi_{\omega}. \quad (2.2)$$

Cette notion est illustrée dans l'Exemple 27.

2.2.3 Information préférentielle possibiliste

Pour contrer les limitations de l'approche ensembliste, nous considérons des informations préférentielles possibilistes prenant la forme (E, α) , où $E \subseteq \Omega$ est un sous-ensemble de modèles possibles et α est le niveau de confiance que l'assertion E soit vraie. Ce degré peut être interprété comme une mesure de nécessité sur l'ensemble E : $N(E) = \alpha$. Il est dès lors possible d'associer une distribution de possibilité $\pi_{(E, \alpha)}$ à cette information, cette distribution étant celle qui est la moins informative, tout en satisfaisant $N(E) = \alpha$. $\pi_{(E, \alpha)}$ est ainsi définie par :

$$\pi_{(E, \alpha)}(\omega) = \begin{cases} 1 & \text{si } \omega \in E, \\ 1 - \alpha & \text{sinon.} \end{cases} \quad (2.3)$$

En particulier, $\alpha = 1$ indique une certitude totale sur l'appartenance de ω à E (et nous revenons à une information ayant valeur dans un ensemble), tandis que $\alpha = 0$ équivaut à une ignorance totale sur l'appartenance ou non de ω à E . L'équation (2.3) peut être interprétée comme modélisant une information en logique possibiliste (Dubois et Prade, 2004), et la plupart des outils et des raisonnements utilisés dans ce chapitre et les suivants pourraient être interprétés à travers le prisme d'une telle logique¹. Dans ce manuscrit, nous considérons qu'un ensemble E_i est le résultat d'une comparaison d'une paire d'alternatives $(x, y) \in \mathbb{X}^2$, où le décideur peut indiquer $x \succeq y$ ou $y \succeq x$. Nous notons par $E_{x \succeq y}$ et $E_{y \succeq x}$ les sous-ensembles de Ω résultants de chaque réponse possible du décideur. Pour compléter ces notations, nous pouvons noter $x \succeq_\alpha y$ et $y \succeq_\alpha x$ la préférence du décideur pour une alternative par rapport à une autre, sachant un niveau de confiance α .

Exemple 26 (Information possibiliste). *Nous reprenons l'Exemple 24, mais cette fois-ci le décideur ajoute un niveau de confiance $\alpha_1 = 0.7$ avec sa réponse, indiquant qu'il est plutôt confiant de celle-ci, mais pas totalement certain. Nous pouvons ainsi calculer la distribution $\pi_{(E_1, \alpha_1)}$, aussi représentée sur la Figure 2.4 :*

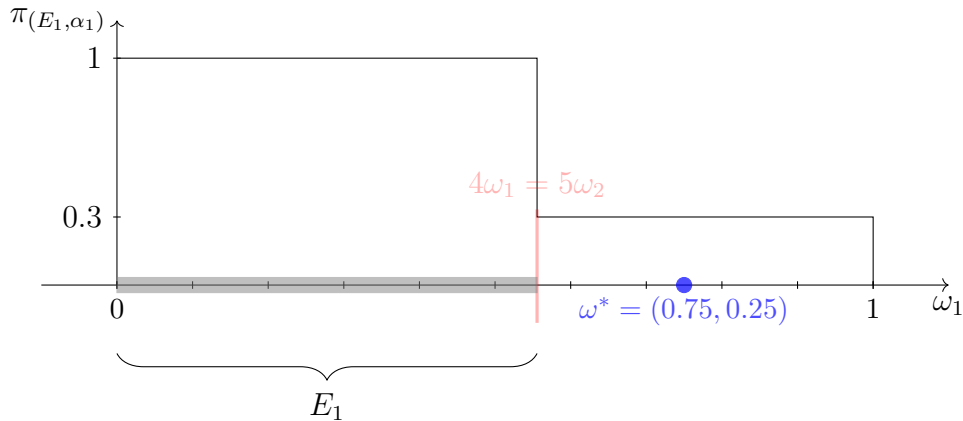


FIGURE 2.4 – Information préférentielle $\pi_{(E_1, \alpha_1)}(\omega)$ de l'Exemple 26 permettant d'atténuer l'impact de l'erreur de l'Exemple 24

1. Il ne faut pas confondre cette interprétation avec l'idée d'utiliser la logique possibiliste pour représenter les préférences (Benferhat *et al.*, 2002), dans laquelle les niveaux α_i représentent les intensités des préférences et non la quantification de l'incertitude.

$$\pi_{(E_1, \alpha_1)}(\omega) = \begin{cases} 1 & \text{si } \omega \in E = \{\omega \mid 4\omega_1 \leq 5\omega_2\}, \\ 0.3 & \text{sinon.} \end{cases}$$

L'intérêt de l'approche possibiliste est ici de ne pas écarter l'ensemble $\Omega \setminus E_1$, qui contient ω^* dans cet exemple particulier, la possibilité de ce sous-ensemble n'étant pas nulle. Ainsi, en supposant que le décideur fasse ensuite de son mieux pour ne donner que des bonnes réponses, l'algorithme d'élicitation va pouvoir converger vers une estimation correcte de ω^* .

En pratique, nous recueillons plusieurs informations (E_i, α_i) , $i \in \{1, \dots, n\}$ au cours du processus d'élicitation, chaque information correspondant à une distribution $\pi_{(E_i, \alpha_i)}$. Les distributions $\pi_{(E_i, \alpha_i)}$ peuvent alors être combinées ou fusionnées en une seule distribution en étendant les opérateurs d'ensemble classiques tels que la conjonction (ET logique) et la disjonction (OU logique). L'utilisation de ces opérateurs permet également une interprétation plus facile des opérations effectuées (Destercke *et al.*, 2008; Pichon *et al.*, 2014; Dubois *et al.*, 2016). En particulier, si nous n'avons aucune raison de penser que les éléments d'information $\pi_{(E_i, \alpha_i)}$ ne sont pas fiables, la manière la plus sensée de les combiner est la conjonction, comme nous l'avons vu dans la Sous-Section 1.4.3. En théorie des possibilités, cette conjonction se fait généralement par l'utilisation d'un opérateur T-norme, comme nous l'avons vu dans la Sous-Section 1.4.4. π_{\cap} est définie comme la conjonction des distributions $\pi_{(E_i, \alpha_i)}$ avec une T-norme (notée par l'opérateur binaire infixé $*$) :

$$\pi_{\cap}(\omega) = \pi_{(E_1, \alpha_1)}(\omega) * \dots * \pi_{(E_n, \alpha_n)}(\omega). \quad (2.4)$$

Exemple 27. *fusion d'information préférentielle possibiliste et détection d'une incohérence*

Nous continuons l'Exemple 26 en rajoutant une nouvelle réponse du décideur : nous supposons que nous lui demandons de comparer les alternatives BT et GO, et qu'il réponde cette fois-ci correctement qu'il préfère BT à GO avec un niveau de confiance $\alpha_2 = 0.9$, soit une très forte confiance. Nous avons alors la frontière de décision suivante :

$$WS_{\omega}(BT) \geq WS_{\omega}(GO) \Rightarrow 8\omega_1 + \omega_2 \geq 7\omega_1 + 3\omega_2 \Rightarrow \omega_1 \geq 2\omega_2.$$

Nous déduisons alors la distribution $\pi_{(E_2, \alpha_2)}$, représentée aussi sur la Figure 2.5 :

$$\pi_{(E_2, \alpha_2)}(\omega) = \begin{cases} 1 & \text{si } \omega \in E = \{\omega | \omega_1 \geq 2\omega_2\}, \\ 0.1 & \text{sinon.} \end{cases}$$

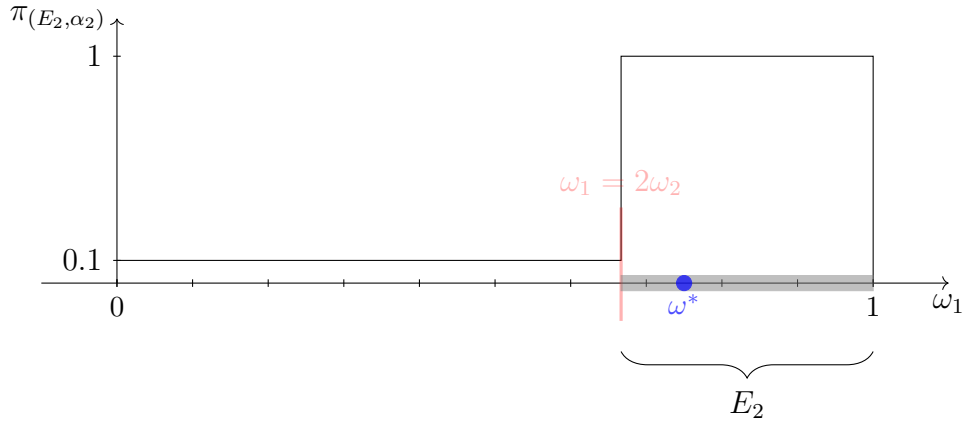


FIGURE 2.5 – Information préférentielle $\pi_{(E_2, \alpha_2)}(\omega)$ de l'Exemple 27

Cherchons désormais à fusionner les informations (E_1, α_1) et (E_2, α_2) pour obtenir une seule distribution de possibilité π_\cap , représentée sur la Figure 2.6. En supposant que nous utilisons la T -norme produit, nous avons :

$$\pi_\cap(\omega) = \begin{cases} T_P(1, 1 - \alpha_2) = 0.1 & \text{si } \omega_1 < 5/9, \\ T_P(1 - \alpha_1, 1 - \alpha_2) = 0.03 & \text{si } 5/9 \leq \omega_1 \leq 2/3, \\ T_P(1 - \alpha_1, 1) = 0.3 & \text{si } \omega_1 > 2/3. \end{cases}$$

Nous pouvons voir que la distribution π_\cap est non normalisée : son maximum est de 0.3. Cela indique qu'une incohérence a été détectée : les deux informations sont en contradiction, l'une d'entre elle indiquant que le décideur préfère fortement le prix que le goût, l'autre indiquant que ce même décideur préfère fortement le goût au prix. Il est possible de mesurer le degré d'incohérence qui est $\text{Inc}(\pi_\cap) = 1 - \max_{\omega \in \Omega} \pi_\cap = 0.7$.

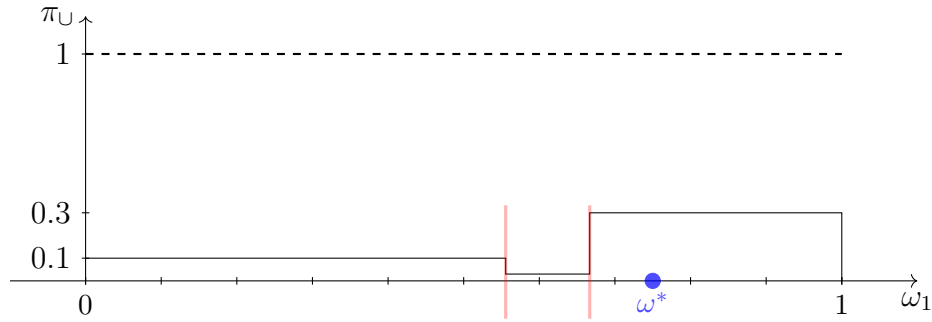


FIGURE 2.6 – Exemple d’une distribution de possibilité non normalisée, avec $\text{Inc}(\pi) = 0.7$

Nous venons de voir dans cette section que la théorie des possibilités est capable de modéliser correctement des informations préférentielles incertaines, de fusionner ces informations préférentielles en une information préférentielle synthétique, potentiellement d’empêcher que l’algorithme d’élicitation converge vers une mauvaise recommandation à la suite d’une erreur, et même de potentiellement détecter de l’incohérence. Les deux derniers points, bien que vérifiés sur de simples illustrations, méritent d’être confirmés au travers d’expérience. Néanmoins, avant de faire des expériences synthétiques, il reste à savoir comment recommander des alternatives au décideur, et surtout comment choisir des questions à lui présenter. Pour ce faire, nous proposons une extension possibiliste du regret Minimax, ainsi que du critère CSS, que nous détaillons en Section 2.3.

2.3 Regret possibiliste et adaptation du CSS

2.3.1 Extension du regret au cas possibiliste

À partir des informations préférentielles (E_i, α_i) , et plus particulièrement des E_i , nous avons une partition finie $\{\Omega_1, \dots, \Omega_P\}$ de Ω , l’ensemble initial des valeurs de ω possibles. Comme nous avons choisi des fonctions d’agrégation linéaires en leur paramètre, chaque information préférentielle (E_i, α_i) nous permet d’établir une frontière de décision linéaire sur Ω . Par conséquent, chaque élément Ω_i est un

polytope convexe² où $\forall \omega, \omega' \in \Omega_i$ nous avons $\pi(\omega) = \pi(\omega')$, signifiant que chaque polytope Ω_i est constant vis-à-vis de la distribution de possibilité. Nous verrons par la suite que cette partition de Ω nous est très utile pour déterminer une alternative à recommander au décideur dans notre cadre possibiliste.

Exemple 28 (Partition de l'ensemble Ω). *Reprenons l'Exemple 27, mais cette fois-ci la première information préférentielle donnée par le décideur est corrigée pour retrouver de la cohérence (au lieu de $BT \preceq_{\alpha_1} EM$, nous avons bien $BT \succeq_{\alpha_1} EM$). Ainsi, nous avons une nouvelle distribution de possibilité π_{\cap} représentée en Figure 2.7 :*

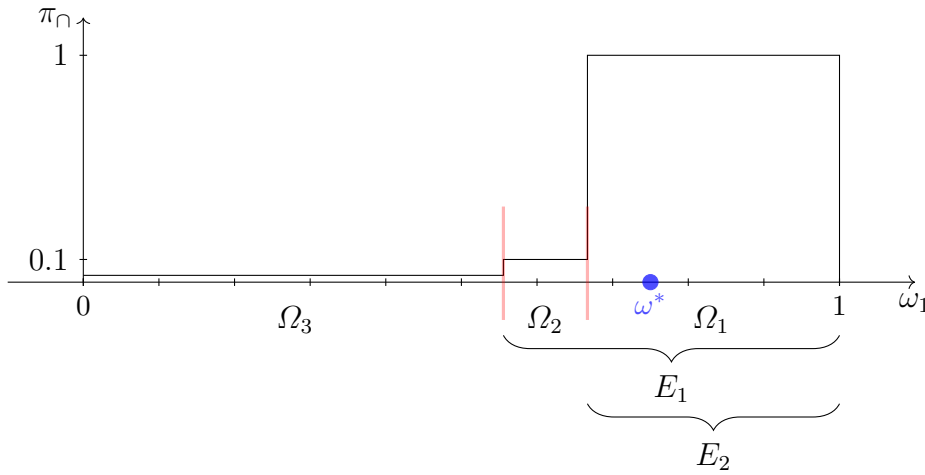


FIGURE 2.7 – Illustration de la partition d'un ensemble Ω en éléments Ω_i

Nous retrouvons bien l'idée que pour chaque polytope Ω_i , ses éléments sont constants par rapport à la distribution de possibilité :

$$\begin{cases} \pi(\omega) = T_P(0.3, 0.1) = 0.03 & \text{si } \omega \in \Omega_1, \\ \pi(\omega) = T_P(1, 0.1) = 0.1 & \text{si } \omega \in \Omega_2, \\ \pi(\omega) = T_P(1, 1) = 1 & \text{si } \omega \in \Omega_3. \end{cases}$$

2. Ω est un polytope convexe, et chaque information préférentielle (E_i, α_i) induit une contrainte linéaire sur Ω , formant donc des polytopes convexes.

Rappel des équations du regret Minimax

Nous avons évoqué en détail dans la Sous-Section 1.2.3 comment déterminer le regret Minimax, et dans la Sous-Section 1.3.3 comment l'appliquer à l'élicitation incrémentale, ainsi que l'heuristique CSS. Dans les paragraphes qui suivent, nous allons étendre ces notions à une approche possibiliste. Pour faciliter la lecture, nous rappelons ici les équations qui permettent de déterminer le regret Minimax, sans les détails qui se trouvent déjà dans les sous-sections mentionnées.

Le regret de choisir une alternative $x \in \mathbb{X}$ par rapport à une alternative $y \in \mathbb{X}$ est défini par :

$$R_\omega(x, y) = f_\omega(y) - f_\omega(x). \quad (2.5)$$

Soit l'ensemble des valeurs possibles de ω noté Ω . Le regret maximal par paire entre $x, y \in \mathbb{X}^2$ sachant Ω est défini par :

$$\text{PMR}(x, y, \Omega) = \max_{\omega \in \Omega} R_\omega(x, y). \quad (2.6)$$

Le regret maximal de choisir $x \in \mathbb{X}$ sachant Ω est défini par :

$$\text{MR}(x, \Omega) = \max_{y \in \mathbb{X}} \text{PMR}(x, y, \Omega). \quad (2.7)$$

Le regret maximal minimal du sous-ensemble d'alternatives disponibles \mathbb{X} sachant Ω est défini par :

$$\text{mMR}(\Omega) = \min_{x \in \mathbb{X}} \text{MR}(x, \Omega), \quad (2.8)$$

avec $x^* = \arg \text{mMR}(\Omega)$ l'alternative avec le regret minimal dans le pire scénario, correspond au choix optimal (en termes de regret).

Nouvelles équations du regret Minimax au cas possibiliste

Le principe général de cette extension du regret Minimax est de considérer les différentes coupes-alpha d'une distribution de possibilité, et de calculer les différentes notions de regret (dans le cas ensembliste) pour chaque coupe, puis de moyenner en fonction des coupes-alpha. Cette approche est standard dans les approches possibilistes (voir par exemple Hüllermeier (2014)).

Définition 31 (EPMR). *Soit une distribution de possibilité π_k . Le regret maximal par paire entre $x, y \in \mathbb{X}^2$ sachant π_k est défini comme étant le PMR moyenné selon les différentes coupes-alpha de π_k :*

$$EPMR(x, y, \pi_k) = \sum_{i=1}^n (\alpha_i - \alpha_{i+1}) PMR(x, y, \pi_k^{\alpha_i}), \quad (2.9)$$

où $1 = \alpha_1 > \dots > \alpha_n > \alpha_{n+1} = 0$ sont les valeurs distinctes de π_k .

Remarquons que, dans le cas où π_k prend valeur dans un ensemble, c'est-à-dire que les valeurs distinctes α_i ne sont que 0 et 1, nous retrouvons le PMR standard de l'Équation (2.6). De plus, si $\pi_k^1 = \emptyset$, indiquant que π_k est non normalisée et donc que de l'incohérence a été détectée, nous devons définir $PMR(x, y, \emptyset)$. Sachant que dans ce chapitre, nous ne nous intéressons qu'à la détection d'incohérence et non à sa gestion, ce problème n'est pas encore important à traiter. Nous détaillerons cette problématique dans le Chapitre 4, et nous supposerons que $PMR(x, y, \emptyset) = 0$, permettant de conserver la convergence vu dans la Propriété 7 (Guillot et Destercke, 2019).

Définition 32 (EMR). *Le regret maximal de choisir $x \in \mathbb{X}$ sachant π_k est défini comme étant le MR moyenné selon les différentes coupes-alpha de π_k :*

$$EMR(x, \pi_k) = \sum_{i=1}^n (\alpha_i - \alpha_{i+1}) MR(x, \pi_k^{\alpha_i}). \quad (2.10)$$

De nouveau, si π_k prend valeur dans un ensemble, nous retrouvons l'Équation (2.7). Remarquons qu'un autre choix aurait pu se porter sur le MER, où nous aurions $MER(x, \pi_k) = \max_{y \in \mathbb{X}} EPMR(x, y, \pi_k)$. Les deux approches se rapportent à l'approche ensembliste quand π_k prend valeur dans un ensemble, mais ne sont pas équivalentes. Notre choix s'est porté sur le EMR, car sous l'hypothèse que ω^* appartienne à une coupe-alpha $\pi_k^{\alpha_i}$, il nous paraît plus logique de déterminer au sein de cette coupe-alpha quelle alternative y maximise le regret avec x , sachant que pour différentes coupes-alpha cette alternative y pourrait être différente.

Définition 33 (mEMR). *Le regret maximal minimal du sous-ensemble d'alternatives disponibles \mathbb{X} sachant π_k est défini par :*

$$mEMR(\pi_k) = \min_{x \in \mathbb{X}} EMR(x, \pi_k). \quad (2.11)$$

Nous avons vu que $\text{EMR}(x, \pi_k)$ se réduit à l'Équation (2.7) quand π_k prend valeur dans un ensemble, donc $\text{mEMR}(\pi_k)$ se réduit aussi à l'Équation (2.8). Un autre choix aurait pu se porter sur $\text{EmMR} = \sum_{i=1}^n (\alpha_i - \alpha_{i+1}) \text{mMR}(\pi_k^{\alpha_i})$, qui n'est pas équivalent au mEMR , mais tous deux se réduisent à l'Équation (2.8) dans le cas où π_k prend valeur dans un ensemble. Néanmoins, le EmMR pourrait nous donner plusieurs meilleures solutions, ce qui est problématique quand nous voulons faire le choix d'une seule alternative à recommander (mais est intéressant pour étudier la robustesse des alternatives recommandées par exemple). Le mEMR ne donne qu'une seule alternative, ce qui correspond mieux à notre problématique.

Lorsque π_k est normalisée et qu'il n'y a aucune incohérence, nous avons vu dans la Sous-section 1.4.2 que les mesures de nécessité (N) et de possibilité (Π) peuvent être vues respectivement comme une probabilité supérieure et une probabilité inférieure. Cela signifie que l'EPMR, l'EMR et le mEMR peuvent être interprétés comme des espérances supérieures sur un ensemble de probabilités. En l'absence d'incohérences, le mEMR reste une limite supérieure du regret réel. Ce lien est perdu en cas d'erreurs.

Exemple 29 (Formulation de l'EPMR et du EMR). *Reprenons l'information préférentielle possibiliste de l'Exemple 26. Nous avons 2 valeurs distinctes pour la distribution de possibilité $\pi_{(E_1, \alpha_1)}$ qui sont 1 et 0.3. Nous avons alors :*

$$\text{EPMR}(x, y, \pi_{(E_1, \alpha_1)}) = (1 - 0.3)\text{PMR}(x, y, [5/9, 1]) + 0.3\text{PMR}(x, y, [0, 1]).$$

x/y	CA	EM	MO	GO	BT
CA	0	4	5	7	8
EM	0.278	0	1	3	4
MO	0.5	.222	0	2	3
GO	0.944	0.667	0.444	0	1
BT	1.78	1.5	1.28	0.833	0

TABLE 2.2 – EPMR de l'information préférentielle $\pi_{(E_1, \alpha_1)}$ de l'Exemple 26

Nous avons déjà calculé dans le chapitre précédent $\text{PMR}(x, y, [5/9, 1])$ (Table 1.10) et $\text{PMR}(x, y, [0, 1])$ (Table 1.9). Nous obtenons alors l'EPMR de la Table 2.2.

La formule de l'EMR est alors :

$$MR(x, \pi_{(E_1, \alpha_1)}) = 0.7 MR(x, [5/9, 1]) + 0.3 MR(x, [0, 1]).$$

Nous obtenons alors l'EMR de la Table 2.3.

x/y	CA	EM	MO	GO	BT
MR	8	4	3.3	2.5	2.63

TABLE 2.3 – EMR de l'information préférentielle $\pi_{(E_1, \alpha_1)}$ de l'Exemple 26

Dès lors, nous pouvons déterminer que $mEMR = 2.5$ pour l'alternative GO, qui est alors la meilleure recommandation courante. Nous remarquons d'une part que l'alternative recommandée est différente de celle de l'algorithme d'élicitation ensembliste présenté dans l'Exemple 16 et que d'autre part l'EMPR ainsi que l'EMR sont tous deux plus grands que leurs équivalents non possibilistes (PMR et MR). Nous formaliserons cette observation avec la Propriété 5.

2.3.2 Extension avec l'heuristique CSS

Nous proposons désormais une application de l'heuristique CSS à notre approche possibiliste (PCSS) pour sélectionner les questions, qui étend donc la stratégie CSS à des informations préférentielles possibilistes. La stratégie est présentée dans l'Algorithme 3. Nous supposons que le décideur fournit un choix unique lorsqu'on lui présente une paire, ce qui se traduit par une préférence non stricte. De plus, que ce soit au niveau des polytopes ou de l'optimisation linéaire, l'inclusion ou non de frontières n'aura aucune conséquence sur les calculs pratiques et les résultats numériques.

Étant donné que l'EMPR, l'EMR et le mEMR se réduisent tous à leurs homologues ensemblistes quand $\alpha = 1$ pour toutes les réponses, notre approche PCSS étend l'approche ensembliste CSS standard.

Algorithme 3 : Élicitation incrémentale possibiliste avec l'heuristique CSS (PCSS)

Données : Sous-ensemble d'alternatives disponibles \mathbb{X} , nombre maximal de questions q_{\max} , seuil de regret τ , incohérence maximale I_{\max} , T-norme T

Résultat : La recommandation optimale $x^* = \arg \text{mEMR}(\pi_k)$

$k = 0, \pi_0 = \mathbb{I}_\Omega;$

tant que $k < q_{\max}$, $\text{mEMR}(\pi_k) > \tau$ **et** $\text{Inc}(\pi_k) < I_{\max}$ **faire**

$x^* = \text{mEMR}(\pi_k);$

$y^* = \arg \max_{y \in \mathbb{X}} \text{EPMR}(x^*, y, \pi_k);$

Demander au décideur si $x^* \succeq y^*$ ou $x^* \preceq y^*$, avec un niveau de confiance $\alpha_k;$

si la décideur répond $x^* \succeq_{\alpha_k} y^*$ **alors**

$\pi_{k+1} = T(\pi_k, \pi_{x^* \succeq_{\alpha} y^*})$

sinon $\pi_{k+1} = T(\pi_k, \pi_{x^* \preceq_{\alpha} y^*});$

$k = k + 1;$

fin

Propriétés de l'approche PCSS Notre objectif est de conserver les propriétés intéressantes des approches ensemblistes (et notamment avec l'heuristique CSS) lorsque leurs hypothèses sont vérifiées (bon choix de modèle et réponses correctes), c'est-à-dire de converger vers la bonne recommandation avec des garanties élevées, tout en autorisant les erreurs. Nous montrons que la première condition est satisfaite, tandis que d'autres expériences dans la Section 2.4 démontrent que les incohérences, quel que soit leur type, peuvent être détectées par la stratégie PCSS.

L'approche ensembliste avec l'heuristique CSS possède trois propriétés importantes :

1. Monotonicité : le mMR diminue à chaque itération.
2. Garanties fortes : le regret ressenti de la recommandation est au moins aussi mauvais que le mMR calculé.
3. Non-conflit : une question, quelle que soit sa réponse, n'entre jamais en conflit avec les réponses précédentes.

Pour le PCSS, nous voulons conserver les deux premières propriétés lorsque nous n'avons pas d'informations contradictoires. Ces propriétés garantissent que le PCSS converge et fournit des recommandations robustes dans le cas idéal. La troisième propriété n'est pas conservée, car nous voulons détecter les contradictions.

Tout d'abord, nous démontrons que nous retrouvons la convergence lorsque qu'aucune incohérence n'est observée.

Proposition 4 (Monotonicité). *Soit π_k et π_{k-1} deux distributions de possibilité successives issues de notre approche PCSS. En supposant qu'elles soient normalisées (aucune incohérence), nous avons :*

1. $EPMR(x, y, \pi_k) \leq EPMR(x, y, \pi_{k-1})$,
2. $EMR(x, \pi_k) \leq EMR(x, \pi_{k-1})$,
3. $mEMR(\pi_k) \leq mEMR(\pi_{k-1})$.

Démonstration. Notons par $\mathcal{A}^k = \{\alpha_1, \dots, \alpha_m\}$ les valeurs distinctes et ordonnées de π_k , et par $\mathcal{A}^{k-1} = \{\alpha_1, \dots, \alpha_n\}$ celles de π_{k-1} . Considérons l'ensemble ordonné $\mathcal{B} = \mathcal{A}^k \cup \mathcal{A}^{k-1} = \{\beta_1, \dots, \beta_\ell\}$, nous avons alors :

$$EPMR(x, y, \pi_k) = \sum_{i=1}^{\ell} (\beta_i - \beta_{i+1}) PMR(x, y, \pi_{\pi_k}^{\beta_i}),$$

$$EPMR(x, y, \pi_{k-1}) = \sum_{i=1}^{\ell} (\beta_i - \beta_{i+1}) PMR(x, y, \pi_{\pi_{k-1}}^{\beta_i}),$$

$\pi_k \subseteq \pi_{k-1}$ implique que $\pi_{\pi_k}^{\beta_i} \subseteq \pi_{\pi_{k-1}}^{\beta_i}$ et que $PMR(x, y, \pi_{\pi_k}^{\beta_i}) \leq PMR(x, y, \pi_{\pi_{k-1}}^{\beta_i})$, complétant la preuve pour l'EPMR.

Des arguments similaires peuvent être utilisés pour prouver les inégalités avec l'EMR et le mEMR.

□

Nous montrons maintenant que notre approche possibiliste préserve les garanties fortes de l'approche ensembliste, et qu'elle est même plus prudente en cas de réponses incertaines, car les regrets correspondants seront plus élevés.

Proposition 5 (Garanties fortes). *Soit π_k une distribution de possibilité et E_k le sous-ensemble des valeurs possibles de ω obtenus à partir de la même séquence de questions et de réponses, avec $\alpha_i < 1$ dans le premier cas et $\alpha_i = 1$ dans le deuxième. Nous avons alors :*

1. $PMR(x, y, E_k) \leq EPMR(x, y, \pi_k)$,
2. $MR(x, E_k) \leq EMR(x, \pi_k)$,
3. $mMR(E_k) \leq mEMR(\pi_k)$.

Démonstration. Nous observons que $\pi_{x \succeq_{\alpha} y} \subseteq \pi_{x \succeq_{\beta} y}$ dès que $\alpha \geq \beta$. Cela indique que toutes les réponses avec $\alpha = 1$ sont incluses dans les réponses avec $\alpha < 1$.

Comme nous avons $\pi_{\Omega^k} = \mathbb{I}_{\Omega^k} \subseteq \pi_k$, alors nous pouvons reproduire la preuve de la Proposition 4.

□

L'approche PCSS, en plus de respecter la monotonie, est encore plus prudente que l'approche ensembliste avec une heuristique CSS. Le choix de la norme T influence également la prudence : notre approche converge plus lentement si nous prenons T_M plutôt que T_P ou T_L , comme nous le verrons dans la Sous-section 2.4.2. Avec une grande prudence et l'assouplissement de la propriété 3 du CSS, nous pouvons donc détecter les incohérences, comme nous le verrons dans les Sous-sections 2.4.3 et 2.4.4.

2.3.3 Éléments de complexité

Notre approche est plus coûteuse sur le plan computationnel qu'une approche ensembliste. Ainsi, l'ensemble des informations préférentielles définissent une partition finie $\{\Omega_1, \dots, \Omega_P\}$ de Ω . Dans une approche ensembliste classique, seul un polytope Ω_i est considéré (le seul avec une possibilité de 1, où on est certain que ω^* se trouve³, les autres ayant une possibilité de 0) comme le sous-ensemble des valeurs possibles de ω . Par opposition, dans l'approche possibiliste que nous proposons, chaque polytope Ω_i doit être considéré, avec $\pi(\Omega_i) \in [0, 1]$, car nous n'avons aucune certitude dans quel polytope se trouve ω^* , y compris si aucune

3. Sous réserve, bien sûr, qu'aucune erreur n'ait été commise.

erreur n'a été commise. De plus, dans chacun de ces polytopes, nous faisons des opérations potentiellement coûteuses (déterminer le PMR par exemple, qui revient à $n(n-1)$ optimisations linéaires, n étant le nombre d'alternatives). Plusieurs facteurs permettent néanmoins de réduire la complexité de notre approche.

Réutilisation des calculs de regret et imbrication des coupes-alpha. Le calcul du PMR pour un sous-ensemble (et donc pour une coupe-alpha π^α) peut être subdivisé en utilisant la propriété :

$$\text{PMR}(x, y, \pi^\alpha) = \max_{\Omega_i \in \pi^\alpha} \text{PMR}(x, y, \Omega_i), \quad (2.12)$$

ce qui est intéressant, car les calculs effectués durant les précédentes étapes de l'élicitation peuvent être réutilisés pour les polytopes Ω_i n'ayant pas évolués.

De plus, un attrait de la théorie des possibilités est l'imbrication des coupes-alpha. Ainsi, nous avons :

$$\pi^{\alpha_i} \supseteq \pi^{\alpha_j}, \quad \forall \alpha_i \leq \alpha_j. \quad (2.13)$$

Dès lors, les calculs effectués pour une coupe peuvent être partiellement réutilisés pour d'autres coupes, ce qui est très utile pour le calcul de l'EPMR et du EMR.

Exemple 30 (Réutilisation des PMR). *Dans l'Exemple 28, supposons que nous avons calculé $\text{PMR}(x, y, \Omega_1)$, $\text{PMR}(x, y, \Omega_2)$ et $\text{PMR}(x, y, \Omega_3)$. Nous pouvons alors rapidement calculer le PMR de chaque coupe-alpha :*

$$\begin{aligned} \text{PMR}(x, y, \pi_\cap^1) &= \text{PMR}(x, y, \Omega_1), \\ \text{PMR}(x, y, \pi_\cap^{0.1}) &= \max[\text{PMR}(x, y, \Omega_1), \text{PMR}(x, y, \Omega_2)], \\ \text{PMR}(x, y, \pi_\cap^{0.03}) &= \max[\text{PMR}(x, y, \Omega_1), \text{PMR}(x, y, \Omega_2), \text{PMR}(x, y, \Omega_3)]. \end{aligned}$$

Par ailleurs, nous voyons que chaque coupe-alpha ici correspond à l'ajout d'un polytope. Il serait donc possible de déterminer par exemple le PMR de la coupe-alpha $\pi_\cap^{0.1}$ à partir du PMR de la coupe-alpha π_\cap^1 et du PMR du polytope Ω_2 .

Limitation du nombre de polytopes. La partition $\{\Omega_1, \dots, \Omega_P\}$ de Ω est déduite à partir de n informations préférentielles possibilistes (E_i, α_i) . En théorie, nous pourrions croire que le nombre de polytopes convexes Ω_i augmente exponentiellement pour atteindre 2^n , chaque information préférentielle induisant un

hyperplan qui délimite l'espace des valeurs possibles de ω en deux. Néanmoins, il est possible de montrer que le nombre de polytopes augmente polynomialement avec chaque information préférentielle pour un nombre de paramètres p fixé, et que le nombre total de polytopes est limité par un nombre inférieur à 2^n , d'autant plus que p est faible (Schläfli et Wild, 2013, P 39) :

Théorème 1. *Soit E un sous-ensemble convexe borné d'un espace euclidien F de dimension q . Soit H un ensemble de n hyperplans dans F tel que chaque hyperplan sépare à chaque fois F en deux sous-ensembles. Le nombre de sous-ensembles obtenus par l'intersection des n hyperplans de H dans F est de 2^n .*

Soit P le nombre de sous-ensembles non-vides parmi ces 2^n sous-ensembles. Nous avons alors une borne Λ_q^n sur le nombre P :

$$P \leq \Lambda_q^n = \sum_{i=0}^q \binom{n}{i} \leq 2^n. \quad (2.14)$$

La preuve de ce théorème est donnée par Schläfli et Wild (2013). Nous pouvons voir dans notre contexte que $E = \Omega$, H est l'ensemble des hyperplans déduits des informations préférentielles possibilistes (E_i, α_i) , et $q = p$ le nombre de paramètres à estimer. De plus, pour plusieurs fonctions d'agrégation de préférences comme la moyenne pondérée ou l'OWA, nous avons certes p paramètres, mais seulement $p - 1$ degrés de liberté ($\sum_{i=1}^p \omega_i = 1$), indiquant qu'il est possible de travailler dans une dimension $p - 1$. Nous utiliserons donc Λ_{p-1}^n dans ces cas. Notons d'ailleurs que Λ_{p-1}^n est une borne qui n'est pas forcément atteinte, et que potentiellement, nous avons un grand nombre de polytopes vides.

Exemple 31 (Nombre de polytopes). *Dans l'Exemple 28 nous avons $p = 2$ et $n = 2$. Sur la Figure 2.7 de l'Exemple 28, nous n'avons pas quatre polytopes comme nous pourrions l'imaginer ($2^2 = 4$), mais seulement trois : il est impossible de trouver une contrainte linéaire qui permet de couper à la fois E_1 et E_1^c et qui soit différente de celle séparant E_1 et E_1^c . Nous voyons d'ailleurs que pour chaque nouvelle information préférentielle, la droite définie par la contrainte linéaire ne peut couper qu'un seul polytope existant. Pour confirmer cette observation, et en se rappelant que dans cet exemple les préférences du décideur sont modélisées par une*

moyenne pondérée, nous calculons Λ_{p-1}^n , ce qui donne :

$$\Lambda_{p-1}^n = \sum_{i=0}^{p-1} \binom{n}{i} = 1 + n.$$

Dans la Table 2.4 nous donnons la valeur de Λ_p^n en fonction de plusieurs valeurs de p et de n , ainsi que le nombre 2^n pour comparaison. Nous voyons bien que pour des valeurs de p faibles, le nombre de polytopes est différent de 2^n .

$p \backslash n$	5	10	15
3	16	56	121
4	26	176	576
5	31	386	1941
2^n	32	1 024	32 768

TABLE 2.4 – Calcul de Λ_p^n pour différentes valeurs de n et p , en comparaison avec le nombre 2^n

Importance du nombre de critères. Nous avons vu que le nombre de polytopes Ω_i reste faible, d'autant plus que le nombre de paramètres p est faible (et donc le nombre le critère q). Si pour une moyenne pondérée ou une OWA nous avons $p = q - 1$ à estimer, pour l'intégrale de Choquet nous en avons $p = 2^q - 2$ à estimer⁴. Il serait donc préférable que le nombre de critères reste donc faible, d'autant plus que le nombre de paramètres, ainsi que le nombre de contraintes qui est lié au nombre de questions n , influence le nombre de points extrêmes pour chaque polytope, et donc la complexité d'un problème d'optimisation linéaire (Motzkin, 1957; McMullen, 1970). Pour rappel, déterminer le PMR d'un couple d'alternatives est un problème d'optimisation linéaire pour les fonctions d'agrégation que nous avons présentées.

Fort heureusement, en aide à la décision multicritère, le nombre de critères est en général faible (moins d'une dizaine en général, plutôt aux alentours de 5), ainsi que le nombre de questions (15-20 au maximum) ce qui indique que

4. Nous travaillons avec des capacités normalisées, donc nous savons que $v(\emptyset) = 0$ et $v(\mathbb{Q}) = 1$.

notre approche est applicable en aide à la décision multicritère. Les calculs exacts sont donc possibles, mais si le nombre de questions augmente trop, ou que nous travaillons avec beaucoup de paramètres (intégrale de Choquet), il est possible d'utiliser des calculs approximatifs (par exemple en ignorant les coupes-alpha ayant des poids très faibles).

Comparaison avec l'approche de Guillot et Destercke (2019)

Notre approche basée sur la théorie des possibilités, en plus de modéliser simplement et naturellement des préférences incertaines, est plus efficace d'un point de vue computationnel qu'une approche basée sur des fonctions de croyance :

- Comme indiqué précédemment, les coupes-alpha sont imbriquées les unes dans les autres, réduisant le coût computationnel. Avec les fonctions de croyance, on utilise ce qui s'appelle des ensembles focaux, et ceux-ci ne sont pas nécessairement imbriqués, ce qui empêche une réutilisation de certains calculs.
- Les règles de fusion que nous utilisons dans notre approche, basée sur les T-normes, créent moins de coupes-alpha que d'ensembles focaux que dans l'approche des fonctions de croyance. C'est particulièrement le cas avec les T-normes se basant sur le minimum T_M , où le nombre de coupes-alpha est égale au nombre de niveaux de confiance α distincts.

2.4 Expériences

Dans cette section, nous réalisons des simulations numériques⁵ afin d'étudier le comportement de notre méthode et de démontrer ses avantages par rapport à une approche ensembliste standard. La confirmation de cet intérêt nécessiterait des expériences basées sur des données réelles, ce qui dépasse le cadre de cette thèse.

5. https://github.com/LoicAdam/Possibilist_Elicitation

2.4.1 Protocole expérimental

Nous voulons recommander à un décideur l'alternative qu'il devrait préférer parmi 35 alternatives multicritères qui sont Pareto non dominées (aucune alternative n'est pire que les autres, quelles que soient les préférences du décideur). Chaque alternative a quatre critères avec $\mathcal{X}_i = [0, 1]$. Chaque alternative x est générée aléatoirement par rapport à une loi uniforme, c'est-à-dire $x_i \sim \mathcal{U}(0, 1) \forall i \in \{1, \dots, 4\}$. Nous ajoutons également une autre contrainte sur les critères pour nous assurer que les alternatives ne sont pas Pareto-dominées : pour chaque alternative x , la somme des critères est approximativement égale à 2, c'est-à-dire $\sum_{i=1}^4 x_i \approx 2$.

Dans ces expériences, la fonction d'agrégation de critères est soit une somme pondérée, soit une OWA, paramétrée toutes les deux par un ensemble de poids tels que $0 \leq \omega_i \leq 1$ et $\sum_{i=1}^4 \omega_i = 1$. Les poids de chaque fonction d'agrégation sont générés aléatoirement selon une loi de Dirichlet avec, sauf indiqué explicitement que non, l'hyperparamètre $\beta = (1, 1, 1, 1)$. La loi de Dirichlet nous donne des poids dont la somme est égale à un, et cet hyperparamètre spécifique nous garantit que les poids sont générés uniformément sur Ω ⁶.

Chaque processus d'élicitation est ensuite répété 200 fois, afin d'obtenir un échantillon de taille raisonnable.

2.4.2 Convergence et lien avec la T-norme et le niveau de confiance

La première expérience illustre l'influence de la T-norme et du niveau de confiance α sur la convergence de notre méthode possibiliste. Au cours de cette expérience, nous avons regardé l'évolution du mEMR en fonction d'un niveau de confiance α pour une T-norme donnée, le décideur répondant toujours correctement avec un niveau de confiance α . Nous avons ensuite comparé les trois T-normes

6. Lorsqu'une loi de Dirichlet a pour hyperparamètre $\beta = \mathbf{1}$, elle est équivalente à une loi uniforme sur le $(K-1)$ -simplexe ouvert standard. Générer des poids selon une loi uniforme, puis les diviser par leur somme, n'est pas équivalent : les poids ainsi divisés seront plus courants autour du centre du simplexe.

usuelles⁷ : $T_M(a, b) = \min(a, b)$, $T_P(a, b) = a.b$ et $T_L(a, b) = \max(a + b - 1, 0)$. Pour chaque simulation, la fonction d'agrégation est une somme pondérée. À chaque étape, avons calculé le mEMR divisé par le mEMR initial. Nous avons ensuite calculé la moyenne des ratios pour chaque niveau de confiance et pour chaque T-norme.

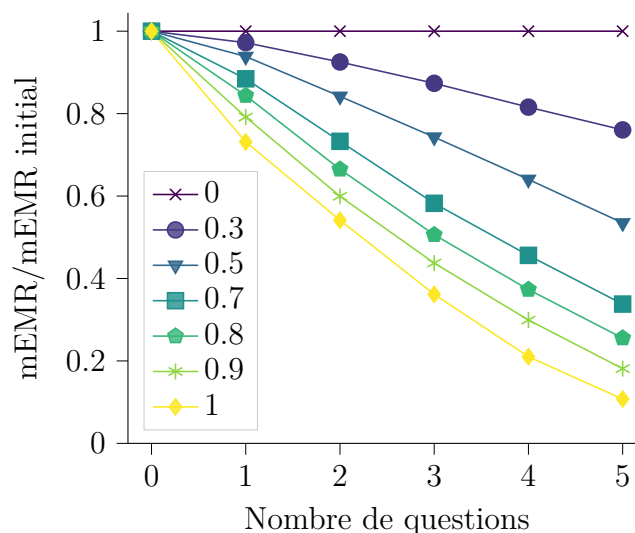


FIGURE 2.8 – Diminution du mEMR en fonction de α avec T_P

Sur la Figure 2.8 nous voyons l'évolution du mEMR en fonction du niveau de confiance pour T_P . Nous voyons que plus la confiance est élevée, plus la convergence est rapide. Rappelons que $\alpha = 1$ équivaut à une approche ensembliste standard, et que $\alpha = 0$ indique une ignorance totale. Le comportement de notre approche possibiliste est donc en accord avec nos attentes : moins la confiance du décideur est élevée, plus l'élicitation met du temps pour converger vers une alternative à recommander, permettant de questionner le modèle et donc de détecter une incohérence si elle existe. Remarquons que quand $\alpha = 0$, la méthode ne convergera jamais : nous avons à chaque étape une ignorance totale sur les préférences.

7. Nous n'étudierons pas la quatrième T-norme usuelle, la T-norme drastique, que nous avons vu dans la Sous-section 1.4.4. D'une part, elle est semblable à T_M quand une des informations préférentielle a une possibilité égale à 1. D'autre part, la possibilité de plusieurs polytopes va très vite atteindre 0, ce qui n'est pas un comportement souhaité si nous souhaitons détecter de l'incohérence.

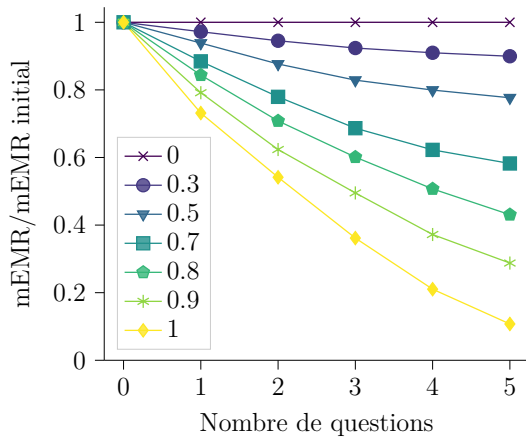


FIGURE 2.9 – Diminution du mEMR en fonction de α avec T_M

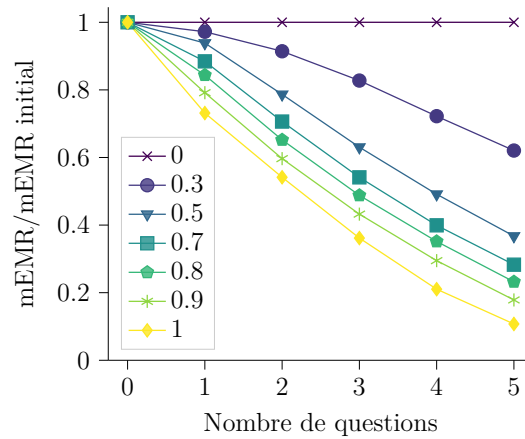


FIGURE 2.10 – Diminution du mEMR en fonction de α avec T_L

Les Figures 2.9 et 2.10 montrent l'évolution pour les autres T-normes. Malgré un comportement similaire, l'évolution de la convergence est différente. La Figure 2.11 permet de comparer la convergence pour un niveau de confiance fixé. Ainsi, plus la T-norme est petite ($T_L \leq T_P \leq T_M$), plus $\pi(\omega)$ diminue pour $\omega \notin E_k$ (voir la Figure 1.18) et plus la méthode converge rapidement.

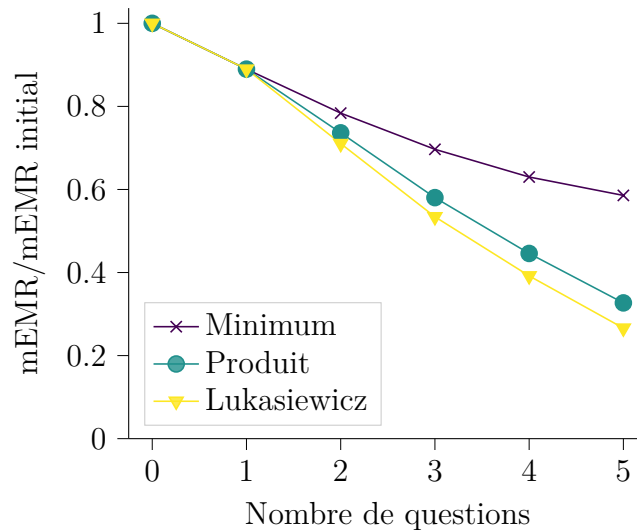


FIGURE 2.11 – Diminution du mEMR en fonction de la T-norme avec $\alpha = 0.7$

À partir des résultats précédents, notre choix de T-norme dans la suite de ce manuscrit va se porter sur la T-norme produit. En effet, en plus d'être un choix classique en logique floue pour les conjonctions fortes, nous avons ces observations :

- T_{\min} pourrait être trop prudente, mettant beaucoup plus de temps à converger. Toutefois, plus la convergence vers des petites valeurs de distribution est rapide, moins il y a de chances d'interroger les modèles $\omega \notin E_k$ et d'enquêter sur d'éventuelles incohérences. Cette T-norme pourrait être justifiée dans le cas où l'élicitation converge rapidement vers une estimation de ω^* , par exemple quand il y a très peu d'alternative ou de critères. Néanmoins, avec 4 à 5 critères et plusieurs dizaines d'alternatives, l'algorithme PCSS ne converge pas instantanément, même avec les autres T-normes.
- L'écart entre T_{prod} et T_{Luk} est faible, ce qui est à la fois surprenant, mais aussi intéressant, car cela suggère que T_{prod} soit suffisante pour fusionner des informations préférentielles.
- T_{Luk} est nilpotente, ce qui nous permet de réduire le coût computationnel de notre approche (plusieurs polytopes avec une possibilité de 0). Néanmoins, cette même propriété est problématique pour manipuler de l'incohérence. En effet, nous avons $T_{\text{Luk}}(a, b) = \max(a+b-1, 0)$: dès que $a+b \leq 1$, la possibilité sera de 0. Or cela peut être problématique si, à cause d'une multiplication du nombre de réponses fausses, la possibilité du polytope contenant ω^* devient 0, car nous aurions le même problème que les approches ensemblistes. De plus, si les réponses sont fortement contradictoires, il n'est pas impossible que la distribution de possibilité soit nulle pour tout ω .

2.4.3 Détection d'erreurs de la part du décideur

La deuxième expérience met en évidence que notre méthode est capable de détecter de l'incohérence quand celle-ci vient du décideur. Dans cette expérience, lorsque de l'incohérence est détectée, l'élicitation est interrompue. L'idée est de prouver que notre méthode est capable de détecter les incohérences et donc de s'arrêter avant que l'élicitation fournisse une estimation $\hat{\omega}^*$ qui s'écarte trop de ω . Nous ne cherchons pas ici à inférer malgré l'incohérence, ou même de corriger l'incohérence, ce travail étant présenté dans le Chapitre 4.

Pour chaque simulation, la fonction d'agrégation de critères est une moyenne pondérée, nous avons choisi la T-norme produit, et le niveau de confiance donné par le décideur ou l'analyste est fixe, prenant valeur dans $\alpha \in \{0.5, 0.7, 0.9, 1\}$. Nous modélisons l'incohérence du décideur en supposant qu'il réponde correctement 70% du temps, et qu'il donne une réponse aléatoire 30% du temps, correspondant à un niveau de confiance de 70%. Chaque simulation se termine dès que le degré d'incohérence est supérieur à 0, lorsque 15 questions ont été répondues ou lorsque que $\text{mEMR} = 0$. Nous déterminons $x = \arg \text{mEMR}$ et nous calculons le regret de choisir l'alternative x par rapport à l'alternative x^* telle que $x^* = \arg \max_{x \in \mathbb{X}} \text{WS}_{\omega^*}(x)$. Nous moyennons enfin le degré d'incohérence et le regret réel pour chaque niveau de confiance, et nous calculons un intervalle de confiance de 95% avec $\text{IC} = [\bar{x} \pm t_{199, 1 - \frac{0.05}{2}} \frac{S^*}{\sqrt{200}}]$ on \bar{x} , S^* étant l'écart type corrigé de l'ensemble de regrets réels et \bar{x} la moyenne empirique des regrets réels.

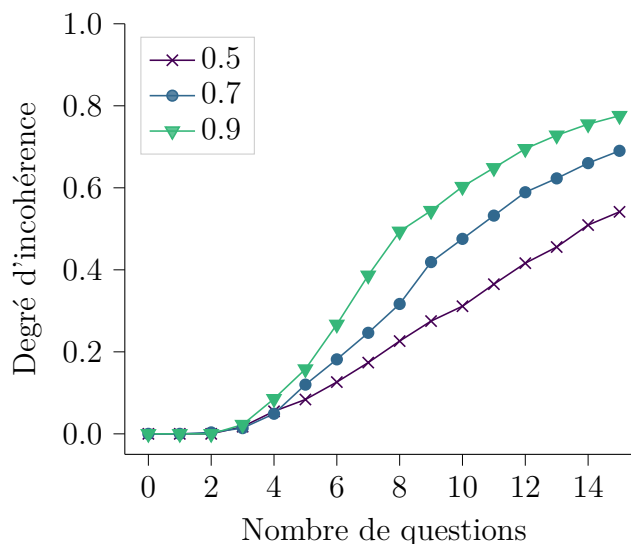


FIGURE 2.12 – Degré d'incohérence en cas d'erreur du décideur pour différentes valeurs de α et un taux d'erreur de 15% environ

La Figure 2.12 montre que les incohérences entre les réponses du décideur sont rapidement détectées (parfois dès la troisième question), ce qui confirme que notre cadre est pertinent en pratique pour étudier les erreurs potentielles du décideur, sans qu'il soit nécessaire de modifier la stratégie d'élicitation pour rechercher

activement les erreurs. Sans surprise, des valeurs élevées de α tendent à créer des incohérences plus importantes lorsqu'elles apparaissent. Remarquons que nous n'avons pas indiqué le degré d'incohérence pour $\alpha = 1$, car logiquement il est tout le temps égal à 0, car il n'est pas possible de détecter de l'incohérence avec une stratégie ensembliste basée sur une heuristique CSS.

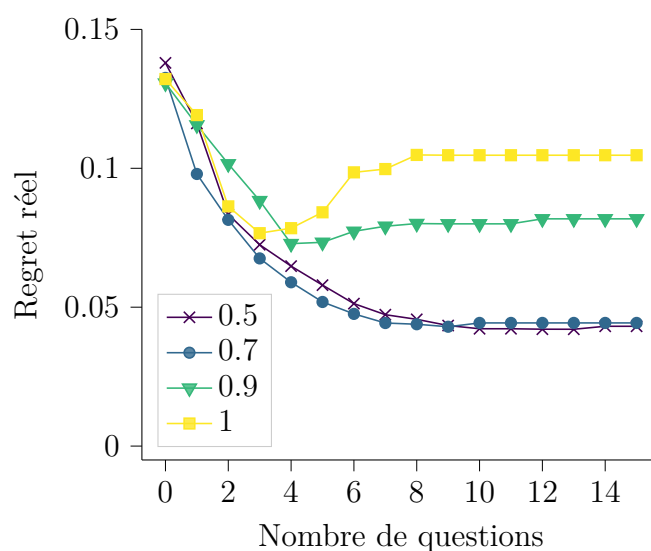


FIGURE 2.13 – Regret réel en cas d'erreur du décideur pour différentes valeurs de α et un taux d'erreur de 15% environ

Nous observons sur la Figure 2.13 et sur la Table 2.5 que, en ce qui concerne le regret réel, les méthodes ensemblistes ($\alpha = 1$) sont celles qui donnent les pires recommandations. Ce comportement était prévisible, car ces méthodes ne permettent pas de détecter d'incohérence, et donc celles-ci ne s'arrêtent que lorsque le budget de question est dépassé, ou que le mEMR est faible. À chaque étape, le mEMR calculé diminue, mais le regret réel peut augmenter, car nous nous éloignons de la bonne valeur ω^* . Nous observons également que plus le niveau de confiance est faible, plus le regret réel est faible, notre méthode PCSS étant de plus en plus prudente. Néanmoins, nous remarquons qu'au bout d'un moment diminuer le niveau de confiance n'a plus vraiment d'influence. Quoi qu'il en soit, la Figure 2.13 montre que lorsque les réponses sont incertaines, notre approche est plus performante qu'une approche ne prenant pas en compte l'incertitude.

	Inférieure	Moyenne	Supérieure
$\alpha = 0.5$.032	.043	.055
$\alpha = 0.7$.030	.044	.058
$\alpha = 0.9$.063	.081	.100
$\alpha = 1$.082	.105	.127

TABLE 2.5 – Intervalle de confiance sur le regret réel en cas d’erreur du décideur pour différentes valeurs de α et un taux d’erreur de 15% environ

2.4.4 Détection d’un modèle de préférence erroné

La troisième expérience souligne que nous pouvons détecter un modèle de préférence erroné, surtout s’il est très éloigné du vrai modèle. Pour chaque simulation, le modèle de préférence est supposé être une moyenne pondérée avec quatre critères. Nous avons également utilisé la T-norme produit et un niveau de confiance fixe $\alpha = 0.7$. Le décideur répond toujours correctement selon son vrai modèle de préférence (qui n’est donc pas une moyenne pondérée). Chaque simulation se termine lorsque 15 questions ont été répondues, ou lorsque $mEMR = 0$. Le but est à chaque étape de voir dans combien de répétitions de l’incohérence a été relevée, indiquant que nous avons détecté un mauvais modèle de préférence. Nous avons testé trois modèles de préférence pour modéliser le vrai modèle de préférence du décideur, qui est supposé par erreur être une moyenne pondérée. :

- Modèle aléatoire : le décideur répond purement aléatoirement.
- OWA 1 : les poids sont générés aléatoirement selon une loi de Dirichlet avec l’hyperparamètre $\beta = (0.85, 0.05, 0.05, 0.05)$. Ces modèles devraient être en général éloignés d’une moyenne pondérée.
- OWA 2 : les poids sont générés aléatoirement selon une loi de Dirichlet avec l’hyperparamètre $\beta = (0.25, 0.25, 0.25, 0.25)$. Ces modèles devraient être en général proches d’une moyenne pondérée.

La Figure 2.14 montre que lorsque le modèle réel est très éloigné du modèle postulé (modèle aléatoire), de l’incohérence est rapidement détectée, et nous pouvons rapidement conclure que le modèle supposé est erroné. Lorsque les modèles se

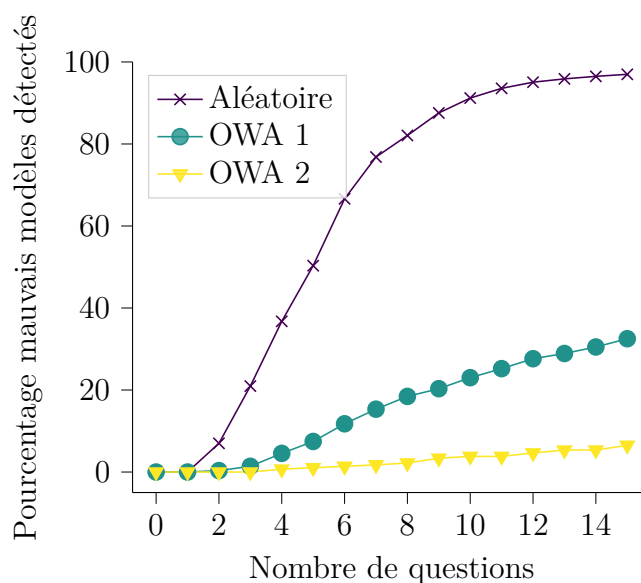


FIGURE 2.14 – Détection d'une fonction d'agrégation de critères erronée

comportent de manière similaire (OWA 2), cette détection peut prendre plus de temps. Dans le Chapitre 4 nous étudierons un peu plus l'incohérence provenant d'un mauvais choix de modèle.

2.5 Conclusion et perspectives

Dans ce chapitre, nous avons mis en avant la problématique de détection de l'incohérence pour les méthodes ensemblistes d'élicitation basées sur l'heuristique CSS, et les limitations des approches probabilistes. Nous avons proposé notre propre solution, basée sur la théorie des possibilités, qui étend formellement les ensembles et est donc cohérente avec les approches robustes et leurs propriétés associées. Elle permet une évaluation progressive de l'incohérence présente dans les informations préférentielles et peut bénéficier d'une littérature très riche sur la façon de traiter de telles situations. De plus, comparé à d'autres théories comme les fonctions de croyance ou les probabilités imprécises, le coût de calcul supplémentaire est limité.

Les expériences sur des données simulées ont montré que notre approche permet de détecter de l'incohérence, que ce soit en cas d'erreur de la part du décideur, ou

par le mauvais choix d'une fonction d'agrégation de critères. En plus de détecter de l'incohérence, notre approche permet de limiter les impacts des erreurs, en limitant la convergence vers une mauvaise estimation du modèle de préférence du décideur.

Points importants vus dans ce chapitre :

- Limitations des approches ensemblistes et probabilistes pour détecter des erreurs de la part du décideur (mauvaises réponses) ou de l'analyste (modèle).
- Proposition d'une approche basée sur la théorie des possibilités qui étend l'approche ensembliste.
- Possibilité de détecter de l'incohérence et de limiter son impact.

Perspectives :

- Dans ce chapitre, nous avons seulement cherché à détecter une incohérence, mais nous n'avons pas cherché à la manipuler, ni même à la réparer (par exemple, supprimer ou corriger les informations préférentielles erronées, ou changer un mauvais modèle de préférence). Nous proposons dans le Chapitre 4 différentes approches pour manipuler et réparer l'incohérence.
- Améliorer le coût computationnel de notre approche : malgré différents éléments pour réduire la complexité (voir la Sous-section 2.3.3), le nombre de polytopes à considérer reste assez grand, et dans chacun de ces polytopes, nous faisons des opérations coûteuses (comme calculer le PMR, soit environ $n^2 - n$ optimisations linéaires). Nous proposons dans le Chapitre 3 d'étudier une autre règle de décision, qui permet notamment de réduire le coût dans chaque polytope à environ $2n - 1$ optimisations linéaires. Néanmoins, une perspective de recherche intéressante serait de voir si notre approche pourrait se faire sans une recherche exhaustive à travers tous les polytopes, en se focalisant que sur les plus intéressants.
- Étendre notre approche à d'autres modèles qui ont des structures très différentes, comme les modèles lexicographiques (Booth *et al.*, 2010) qui ne sont pas basés sur une évaluation numérique, nécessitant d'adapter la notion de regret ou d'utiliser une autre notion pour choisir des questions et pour recommander une alternative. Nous n'avons pas étudié cette perspective.

- Nous avons déjà mentionné que notre cadre est très proche de la logique possibiliste, et pourrait en fait être lu comme une instanciation de celle-ci si nous nous limitons à la Section 2.2. Il serait donc très intéressant de voir comment le traitement de l'incohérence dans de telles logiques (Dubois et Prade, 2015) peut aider dans notre cadre actuel.

Contributions :

- Adam, Loïc et Destercke, Sébastien. Possibilistic preference elicitation by minimax regret. Dans : *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence, UAI 2021, 27-29 juillet 2021, en ligne*. PMLR, 2021. p. 718-727.
- Adam, Loïc et Destercke, Sébastien. Élicitation possibiliste de préférences avec un regret Minimax. Dans : *30e Rencontres Francophones sur la Logique Floue et ses Applications, LFA 2021, 21-22 octobre 2021, Paris, France*. Cépadués Editions, 2021. p. 27-34.

Choix du critère de décision

3.1 Introduction

Nous avons vu dans le Chapitre 1 que l'élicitation incrémentale est une approche très intéressante, puisque chaque nouvelle question prend en compte les informations préférentielles fournies précédemment. Dans la littérature, l'une des principales approches de l'élicitation incrémentale est l'approche robuste, basée sur une optimisation du regret Minimax (Boutilier, 2013; Bourdache et Perny, 2017). Si les hypothèses sous-jacentes sont satisfaites, l'intérêt de cette approche est que, du fait de sa position pessimiste (minimiser le regret dans la pire situation), elle offre de fortes garanties quant à l'alternative recommandée. Elles convergent également en un nombre raisonnable d'étapes vers une bonne recommandation, étant donné que l'espace des modèles possibles est garanti de se réduire après chaque question. Dans ce chapitre, nous travaillons avec ces approches robustes, mais nous pourrions facilement adapter nos résultats à l'approche possibiliste présentée dans le Chapitre 2, ce que nous faisons d'ailleurs dans le Chapitre 4.

L'approche du regret Minimax (Savage, 1951) est un choix populaire pour prendre des décisions sous incertitude, car elle minimise le regret dans le pire des cas. Une telle règle de décision fournit des recommandations plutôt sûres, correspondant à une situation dans laquelle le décideur est plutôt pessimiste, craignant un scénario catastrophique, aussi rare que ce scénario puisse être. Cependant, le regret Minimax n'est qu'une règle de décision parmi tant d'autres en cas d'incertitude (voir Troffaes (2007) pour un compte rendu de ces règles), et comme toutes les autres règles,

elle présente des inconvénients qui peuvent être problématiques. Ainsi, le regret Minimax est sensible à l'ajout d'alternatives non pertinentes et ne garantit pas des recommandations potentiellement optimales, c'est-à-dire des recommandations qui sont les meilleures pour au moins un modèle particulier dans l'ensemble des modèles possibles.

Dans ce chapitre, nous envisageons un point de vue quelque peu opposé, en proposant une approche optimiste, celle du Maximax, pour formuler des recommandations. Il existe plusieurs raisons d'étudier cette règle : tout d'abord, les approches optimistes sont déjà utilisées dans de nombreux domaines en présence d'incertitude, par exemple pour traiter des données manquantes (Cabannes *et al.*, 2020) ou pour identifier des modèles optimaux (Maron et Moore, 1997). Ainsi, l'adoption d'un point de vue optimiste dans l'élicitation des préférences, partageant de nombreuses similitudes avec le cadre d'apprentissage susmentionné, semble pertinente. Ensuite, les recommandations se basant sur la règle de décision du Maximax ne souffrent pas des mêmes inconvénients que nous avons mentionnés pour le regret Minimax¹. Ces recommandations peuvent donc être acceptables dans des situations dans lesquelles celles basées sur le regret Minimax ne le sont pas. Enfin, une telle approche peut être plus efficace sur le plan computationnel que celles basées sur un regret, puisque calculer un regret implique généralement la comparaison de paires d'alternatives, alors que des règles de décision comme le Maximax impliquent généralement des calculs que sur l'ensemble des alternatives.

Nous présentons dans la Section 3.2 une approche optimiste que nous appelons gain Maximax, en adaptant l'heuristique de la stratégie de la solution courante (CSS). Nous discutons également de plusieurs propriétés intéressantes de cette approche. Nous complétons cette présentation dans la Section 3.3 avec quelques expériences sur des données simulées, dont l'objectif est d'étudier des situations dans lesquelles une approche optimiste augmente également la qualité des recommandations.

1. La règle du Maximax n'offre pas, cependant, les mêmes garanties de robustesse sur les recommandations.

3.2 Approche Optimiste

Le regret Maximax est une règle de décision pessimiste : le décideur souhaite l'alternative qui minimise la perte maximale, c'est-à-dire dans le pire des cas. Le décideur a donc une aversion pour le risque et ne voit pas d'inconvénient à ce que le gain soit inférieur en moyenne. Toutefois, il n'est pas évident, dans un cadre préférentiel, que le décideur soit toujours averse au risque plutôt qu'opportuniste. C'est pourquoi nous étudions l'approche du gain Maximax et son adaptation CSS directe, où le décideur veut maximiser son gain dans le meilleur des cas. Le choix de l'approche du gain Maximax peut être justifié de différentes manières : nous montrons plus loin dans cette section qu'une alternative suggérée par une approche du gain Maximax est la meilleure possible pour au moins une situation (un modèle ω). Ce n'est pas nécessairement le cas avec un regret Minimax, ce qui signifie qu'un décideur réticent au risque peut choisir une option qui sera nécessairement sous-optimale quelle que soit la valeur de ω . Cela ne signifie pas qu'une stratégie est meilleure que l'autre, mais simplement qu'elles ont des propriétés différentes, car le choix peut dépendre de l'appétence du décideur de prendre des risques pour maximiser son gain possible. Dans cette section, nous expliquons pourquoi une telle approche est possiblement une alternative intéressante à l'approche du regret Minimax.

Notons qu'il serait possible de voir des règles de décision encore plus pessimistes que le regret Minimax, comme le Maximin, qui est basée sur un gain comme le Maximax, mais dans le pire scénario (tandis que le Maximax s'intéresse au gain dans la meilleure des scénarios). Si cette règle a des propriétés intéressantes que n'a pas le regret Minimax (l'insensibilité aux alternatives non pertinentes, plus conservateur que le regret Minimax), il a été montré que le regret Minimax est plus intéressant dans le cadre de l'élicitation incrémentale (Wang et Bouilher, 2003). Par conséquent, nous ne nous intéresserons ici qu'à une approche optimiste.

3.2.1 Élicitation robuste avec un gain Maximax

Dans cette sous-section, nous définissons notre approche du gain Maximax en appliquant la règle de décision du Maximax.

Définition 34. *Gain*

Le gain associé à une alternative $x \in \mathbb{X}$ pour une fonction d'agrégation de critères f_ω et son paramètre ω est défini par :

$$G_\omega(x) = f_\omega(x). \quad (3.1)$$

Le gain équivaut tout simplement au score de l'alternative selon une fonction d'agrégation pour une valeur de paramètre donnée. Nous parlons de gain en supposant que la performance d'un critère ne puisse pas être négative. Contrairement au regret, le gain ne dépend pas d'une autre alternative.

Nous avons déjà défini le Maximax (MM) dans la Sous-section 1.2.3 comme étant le score maximal que pouvait avoir une alternative $x \in \mathbb{X}$ sur l'ensemble des valeurs possibles de ω noté Ω . Cette valeur ne dépend encore une fois que de l'alternative x et non d'autres alternatives, contrairement au MR du regret Minimax.

Définition 35. *Gain Maximax maximal*

Le gain Maximax maximal du sous-ensemble d'alternatives disponibles \mathbb{X} sachant Ω est défini par :

$$MMM(\Omega) = \max_{x \in \mathbb{X}} \underbrace{\max_{\omega \in \Omega} G_\omega(x)}_{\text{Maximax de } x}. \quad (3.2)$$

Nous avons alors $x^* = \arg \text{MMM}(\Omega)$, l'alternative avec le gain maximal dans le meilleur des scénarios, correspond au choix optimal (en termes de gain). Pour poser des questions au décideur qui soient pertinentes, nous continuons à considérer l'heuristique CSS et nous prenons $y^* = \arg \max_{y \in \mathbb{X}} \text{PMR}(x^*, y, \Omega')$ comme deuxième alternative à présenter au décideur (la première étant x^*).

Exemple 32. *Choix initial avec une approche optimiste*

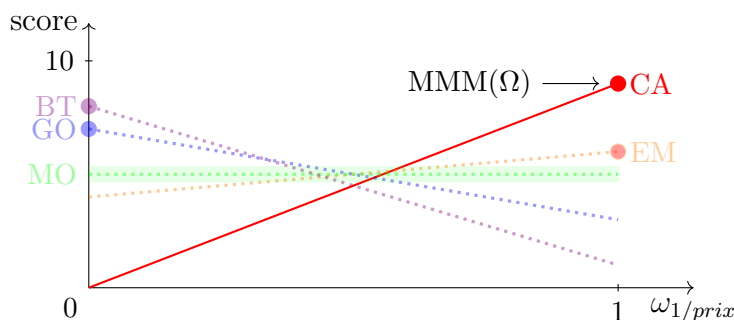
Nous reprenons le sous-ensemble des alternatives disponibles de l'Exemple 1. Nous supposons que les préférences du décideur sont représentées par une somme pondérée WS_ω de poids $\omega = (0.6, 0.4)$. En l'absence d'information, l'ensemble initial des paramètres possibles est Ω , contraint par $0 \leq \omega_i \leq 1$ et $\sum_i \omega_i = 1$, $i \in \{1, 2\}$.

Nous commençons par calculer le Maximax pour chaque alternative $x \in \mathbb{X}$ avec $\max_{\omega \in \Omega} f_\omega(x)$. Comme nous optimisons une fonction linéaire sur un polytope

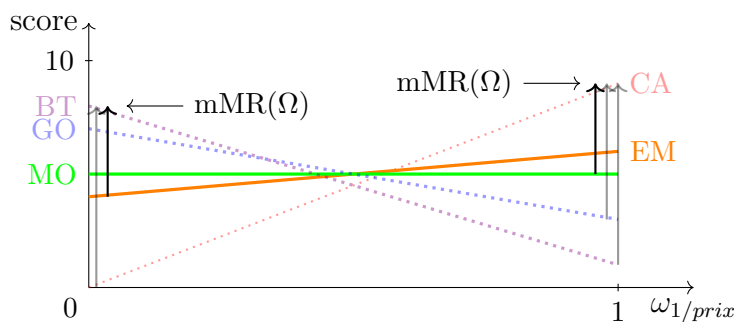
x	CA	EM	MO	GO	BT
Maximax	9	6	5	7	8

TABLE 3.1 – Détermination du Maximax maximal initial de l'Exemple 15

convexe Ω , ce problème d'optimisation est résolu facilement et exactement. Les gains Maximax pour chaque alternative sont présentés dans la Table 3.1 ainsi que sur la Figure 3.1. Par exemple, le gain maximal de CA est de 9 pour $\omega = (0, 1)$.

FIGURE 3.1 – Maximax pour chaque alternative $x \in \mathbb{X}$

$MMM(\Omega) = \max_{x \in \mathbb{X}} \max_{\omega \in \Omega} f_{\omega}(x) = 9$, indiquant que l'alternative à recommander au décideur en l'absence d'information est CA. Pour comparaison, nous avons représenté sur la Figure 3.2 quelle aurait été l'alternative recommandée avec un regret Minimax. Nous pouvons voir que la recommandation dans cet exemple est différente en fonction du critère de décision.

FIGURE 3.2 – MR pour chaque alternative $x \in \mathbb{X}$

3.2.2 Optimalité potentielle

Nous introduisons tout d'abord la notion des solutions potentiellement Ω -optimales (Bourdache et Perny, 2017) :

Définition 36. *Solutions Ω -optimales*

Soit l'ensemble des alternatives disponibles \mathbb{X} et l'ensemble des valeurs possibles de ω noté Ω . L'ensemble PO_Ω est défini par :

$$PO_\Omega = \{x : \exists \omega \in \Omega, x \in \arg \max_{x \in \mathbb{X}} f_\omega(x)\}. \quad (3.3)$$

En d'autres termes, une alternative $x \in \mathbb{X}$ est potentiellement Ω -optimale si x est la meilleure alternative pour au moins une valeur de $\omega \in \Omega$.

Une approche d'élicitation optimiste basée sur un gain Maximax est intéressante du fait de la propriété suivante, qui montre que l'alternative recommandée peut être la meilleure (ce qui n'est pas garanti par l'approche du regret Minimax) :

Proposition 6. *Garantie d'optimalité potentielle*

L'alternative $x^* = \arg \text{MMM}(\Omega)$ recommandée par une approche basée sur un gain Maximax est potentiellement Ω -optimale : $x^* \in PO_\Omega$.

Démonstration. Considérons une valeur du paramètre ω pour laquelle est obtenue $x^* = \max_{x \in \mathbb{X}} [\max_{\omega \in \Omega} f_\omega(x)]$. Il est clair que pour cette valeur ω , qui est dans Ω , x^* est la meilleure alternative, donc x^* est une alternative potentiellement Ω -optimale. \square

Exemple 33. *Le regret Minimax ne garantit pas des solutions potentiellement Ω -optimales*

Sur la Figure 3.3 sont représentées trois alternatives : CA, BT, ainsi que MO', une version alternative de MO qui est beaucoup moins intéressante au niveau du score. MO' est l'alternative recommandée selon la règle de décision du regret Minimax. Comme nous pouvons le voir, une approche basée sur le regret Minimax ne permet pas toujours d'offrir des recommandations Ω -optimales, montrant qu'elle ne respecte pas la Proposition 6. Néanmoins, nous pouvons voir sur cette même figure que la recommandation selon le gain Maximax peut être un mauvais choix pour certaines valeurs de ω ($\omega \in [0, 0.5[$), signifiant qu'il peut être important de vérifier si nous nous trouvons dans de telles situations.

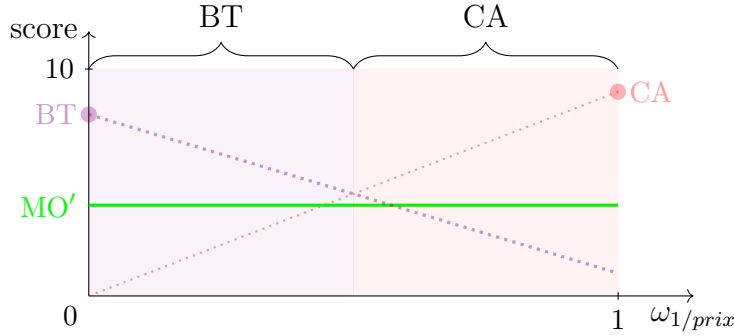


FIGURE 3.3 – Visualisation de l'optimalité de trois alternatives

3.2.3 Convergence

Avec le regret Minimax et l'heuristique CSS, l'élicitation converge de manière prouvée vers le modèle optimal ω^* tant qu'aucune erreur n'est commise. Après la k^e réponse, il est garanti que $\emptyset \neq E_k \subseteq E_{k-1}$, et que l'inclusion sera stricte sous certaines hypothèses. Dans le cas du gain CSS, nous avons la même propriété :

Proposition 7. *Convergence de l'approche du gain CSS*

Soit la paire d'alternative $x^* = \arg \text{MMM}(E_k)$, $y^* = \arg \max_{y \in \mathbb{X}} \text{PMR}(x^*, y, E_k)$ choisies à la k^e étape de l'élicitation selon l'heuristique CSS. Soit $E_k \subseteq \Omega$ le sous-ensemble des valeurs possibles de ω à l'étape k . Nous avons alors :

$$E_k \cap E_{x^* \preceq y^*} \neq \emptyset, \quad (3.4)$$

$$E_k \cap E_{\omega | x^* \succeq y^*} \neq \emptyset. \quad (3.5)$$

Démonstration. Nous prouvons que quelle que soit la réponse, l'intersection avec E_k est non-vide :

- $E_k \cap E_{x^* \preceq y^*} \neq \emptyset$: immédiat comme $x^* \in PO_{E_k}$. En effet, d'après la Proposition 6, il existe un modèle $\omega \in E_k$ tel que $f_\omega(x^*) \succeq f_\omega(y^*)$.
- $E_k \cap E_{\omega | x^* \succeq y^*} \neq \emptyset$: comme y^* est choisi selon le CSS, alors $y^* = \arg \max_{y \in \mathbb{X}} \text{PMR}(x^*, y, E_k)$. Comme $\text{PMR}(x^*, y^*, E_k) = \max_{\omega \in E_k} [f_\omega(y^*) - f_\omega(x^*)] \geq 0$, alors nous avons $\omega \in E_k$ (par exemple, l'alternative pour laquelle le PMR est atteint) avec $f_\omega(y^*) \geq f_\omega(x^*)$, concluant la preuve.

□

Cette proposition nous indique que notre sous-ensemble de valeurs possibles de ω se réduit après chaque réponse de manière non dégénérée, ce qui nous garantit de converger vers la valeur ω^* sous réserve qu'aucune erreur ne soit commise. Notons que nous avons une inclusion stricte $E_k \subset E_{k-1}$ si $\arg \text{MMM}(E_{k-1})$ est unique et si le PMR correspondant est strictement positif.

3.2.4 Complexité computationnelle

Un autre avantage d'une approche optimiste basée sur un gain Maximax est sa plus faible complexité computationnelle. Avec une approche pessimiste basée sur un regret Minimax, le calcul du PMR pour toutes les paires possibles (x, y) telles que $x, y \in \mathbb{X}^2$ et $x \neq y$ est équivalent à la résolution de $n^2 - n$ problèmes d'optimisation linéaire, où $n = |\mathbb{X}|$.

Avec une approche optimiste, le calcul du MMM pour toutes les alternatives $x \in \mathbb{X}$ est équivalent à la résolution de seulement n problèmes d'optimisation linéaire. En utilisant l'heuristique CSS, il faut rajouter $n - 1$ problèmes d'optimisation linéaire pour déterminer le PMR entre x^* et les autres alternatives. Au total, avec l'heuristique CSS, cela fait $2n - 1$ problèmes d'optimisation linéaire. Les problèmes d'optimisation linéaire des deux approches sont similaires (hormis la fonction de coût). Ce coût computationnel inférieur est très intéressant pour les problèmes de décision comportant un grand nombre d'alternatives.

3.3 Expériences

Dans la Section 3.2 nous avons présenté une approche d'élicitation incrémentale basée sur l'optimisation d'un gain (Maximax) plutôt que l'optimisation d'un regret (Minimax). Si cette approche possède des propriétés intéressantes comme un meilleur coût computationnel, une insensibilité aux alternatives non pertinentes, ou encore une garantie que chaque recommandation soit la meilleure dans au moins un scénario ; nous devons encore vérifier si cette approche peut fournir des performances raisonnables, voire meilleures qu'une approche standard basée sur l'optimisation d'un regret. Cette section apporte quelques éléments de réponse, en comparant les stratégies CSS de regret et de gain dans différentes situations.

3.3.1 Protocole expérimental

Nous avons effectué des simulations numériques² pour comparer les performances des deux approches dans différents contextes, dans lesquels nous changeons le type des alternatives considérées, ainsi que le type de modèle. Le décideur est supposé toujours répondre correctement à 10 questions selon ses préférences.

Le premier élément de comparaison est le choix des alternatives :

- Dans un premier cadre, nous avons généré aléatoirement 100 alternatives Pareto-optimales avec huit critères à partir d'une loi uniforme : $x \in [0, 1]^8$. Nous ajoutons également une autre contrainte sur les critères pour nous assurer que les alternatives ne sont pas Pareto-dominées : pour chaque alternative x , la somme des critères est approximativement égale à 4, c'est-à-dire $\sum_{i=1}^8 x_i \approx 4$. Nous avons alors des alternatives dont l'utilité moyenne est approximativement la même, mais dont les valeurs de performance de chaque critère sont en général assez élevées. Dès lors, le décideur n'a pas besoin de faire de gros compromis (si le décideur souhaite se focaliser sur un critère, les autres peuvent néanmoins avoir des valeurs relativement correctes). Nous considérons alors ces alternatives comme de *bonnes alternatives*.
- Dans un second cadre, nous avons généré aléatoirement 100 alternatives avec huit critères à partir d'une loi de Dirichlet : $x \in [0, 1]^8$ et $\sum_{i=1}^8 x_i = 1$. Nous avons une nouvelle fois des alternatives dont l'utilité moyenne est la même, mais dont les valeurs de performance de chaque critère sont en général faibles. Dès lors, le décideur doit faire des compromis importants (s'il souhaite se focaliser sur un critère, tous les autres auront des valeurs très faibles). Nous considérons alors ces alternatives comme de *mauvaises alternatives*.

Le deuxième élément de comparaison est le choix d'une fonction d'agrégation de critères f_ω pour estimer l'utilité d'une alternative. Nous avons comparé les deux approches avec quatre fonctions différentes, toutes générées aléatoirement à partir de différentes lois de Dirichlet de paramètre β :

- WSB : une somme pondérée « équilibrée », où toutes les valeurs des critères sont proches. $\beta = 1000 \cdot (1/8, 1/8, \dots, 1/8)$.

2. https://github.com/LoicAdam/Elicitation_Optimist

- WSU : une somme pondérée « déséquilibrée », où certains critères peuvent avoir des valeurs significativement plus élevées que les autres.
 $\beta = (1/8, 1/8, \dots, 1/8)$.
- OWAU : une OWA « injuste », qui favorise les critères ayant les valeurs les plus élevées. $\beta = 50.(1/36, 2/36, \dots, 8/36)$.
- OWAR : une OWA « redistributive », qui favorise les critères ayant les valeurs les plus faibles. $\beta = 50.(8/36, 7/36, \dots, 1/36)$.

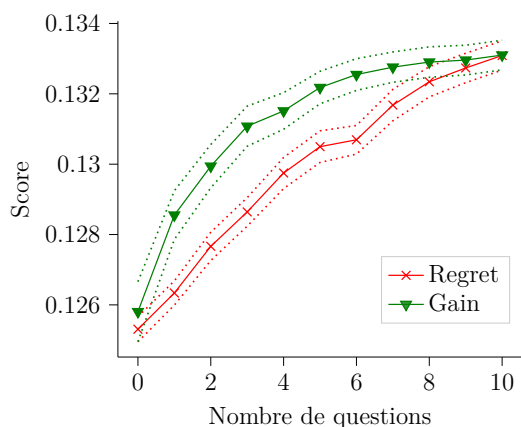
Nous proposons deux mesures pour évaluer la qualité de prédiction de chaque approche. La première mesure est le score réel de la recommandation actuelle, calculé à partir de la fonction d'agrégation f_{ω^*} qui représente les préférences réelles du décideur. La seconde mesure est la position de la recommandation actuelle par rapport aux autres alternatives, compte tenu du score réel de chaque alternative. 1 signifie que nous avons la meilleure alternative et 100 la plus mauvaise.

Nous réduisons également la variabilité des deux mesures avec 200 simulations :

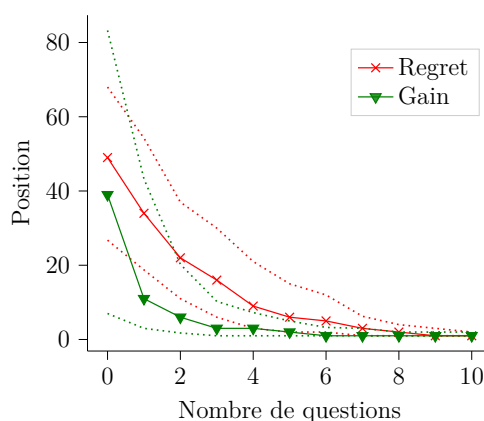
- Pour le score réel, nous supposons la normalité et nous calculons à la fois la moyenne du score réel $\overline{\text{score}}$ sur les 200 simulations et un intervalle de confiance de 95% avec $IC = [\overline{\text{score}} \pm t_{199, 1 - \frac{0.05}{2}} \frac{S^*}{\sqrt{200}}]$ on \bar{x} , S^* étant l'écart type corrigé du score réel en prenant toutes les simulations.
- Pour la position réelle, nous ne ferons pas l'hypothèse de normalité, d'autant plus que la position réelle prend valeur dans \mathbb{N}^+ . Par conséquent, nous prenons la médiane de la position réelle à partir des 200 simulations, ainsi que le premier quartile et le deuxième quartile pour voir 75% des données.

3.3.2 Résultats

Sur la Figure 3.4 nous pouvons voir qu'une approche basée sur un gain fournit de bien meilleures recommandations en présence de mauvaises alternatives pour un modèle de préférence WSB qu'une approche basée sur un regret, bien qu'au bout de 10 réponses, les deux approches sont semblables. Par contre, nous pouvons voir sur la Figure 3.5 que ce n'est plus le cas dès que les alternatives sont bonnes, et que l'approche basée sur un regret est même légèrement meilleure.

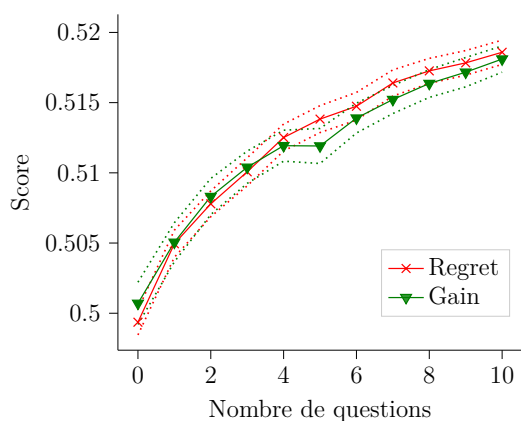


3.4.a : Score

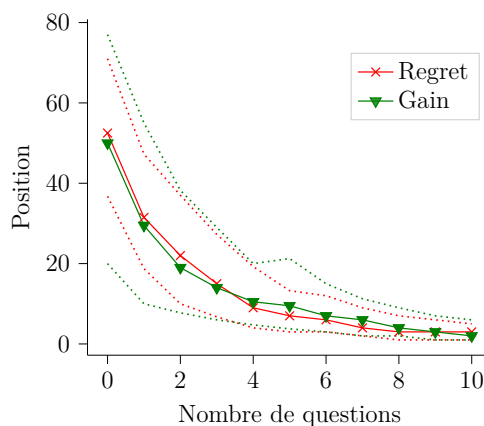


3.4.b : Position

FIGURE 3.4 – Score et position avec de mauvaises alternatives et un modèle WSB



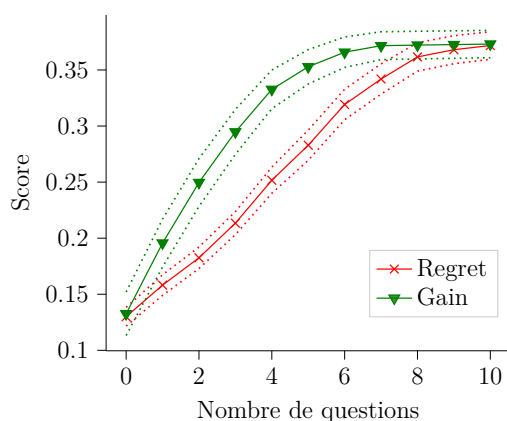
3.5.a : Score



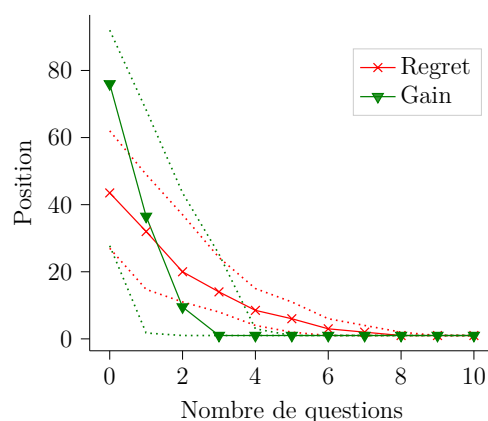
3.5.b : Position

FIGURE 3.5 – Score et position avec de bonnes alternatives et un modèle WSB

Sur la Figure 3.6, nous pouvons voir qu'en présence de mauvaises alternatives, la différence entre les deux approches avec un modèle WSU est similaire à la différence entre les deux approches avec un modèle WSB. En présence de bonnes alternatives, comme nous pouvons le voir sur la Figure 3.7, l'approche basée sur un gain est cette fois-ci légèrement meilleure que l'approche basée sur un regret. Les différences sont néanmoins de plus en plus minimales quand le nombre de réponses augmente.

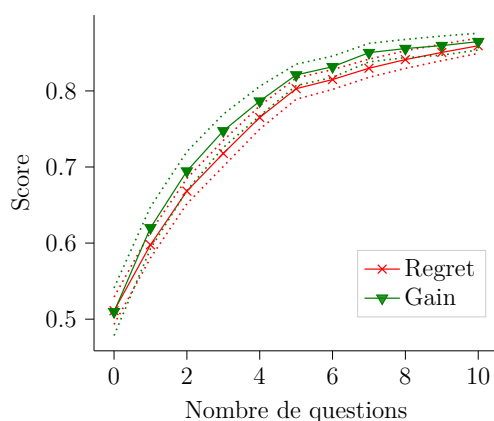


3.6.a : Score

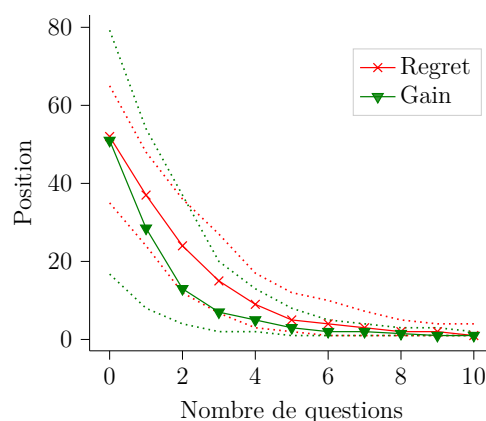


3.6.b : Position

FIGURE 3.6 – Score et position avec de mauvaises alternatives et un modèle WSU



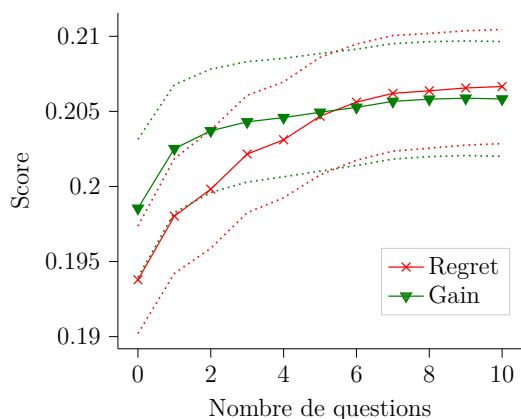
3.7.a : Score



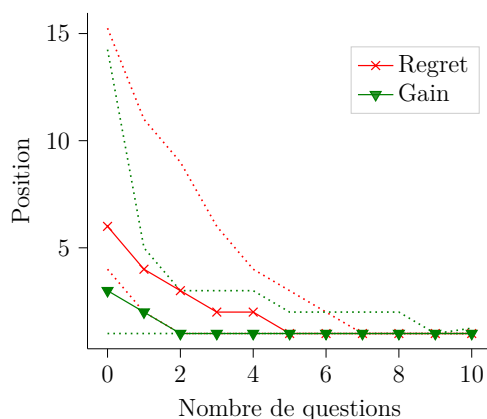
3.7.b : Position

FIGURE 3.7 – Score et position avec de bonnes alternatives et un modèle WSU

Sur les Figures 3.8 et 3.9, nous pouvons voir qu’au départ l’approche basée sur un gain donne de meilleures recommandations qu’une approche basée sur un regret quand nous avons un modèle OWAU. Néanmoins, à partir de la cinquième question, les deux approches sont semblables et l’approche basée sur un regret devient même très légèrement plus intéressante.

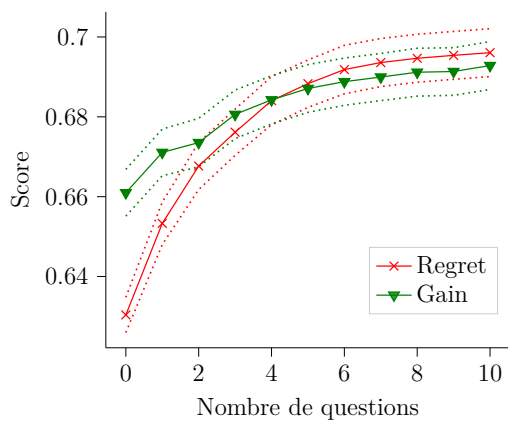


3.8.a : Score

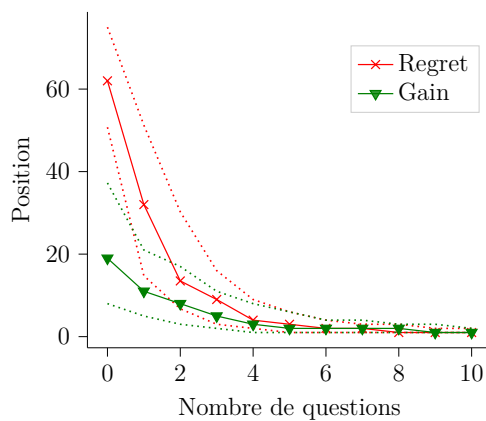


3.8.b : Position

FIGURE 3.8 – Score et position avec de mauvaises alternatives et un modèle OWAU



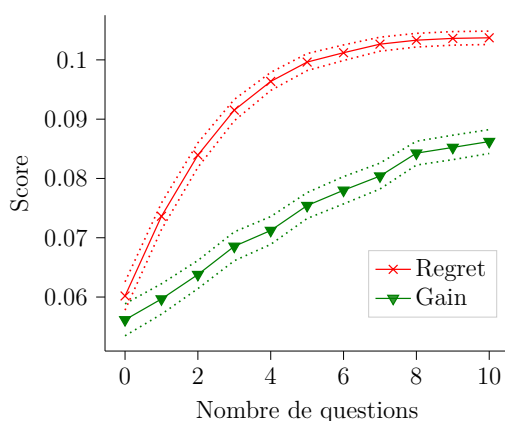
3.9.a : Score



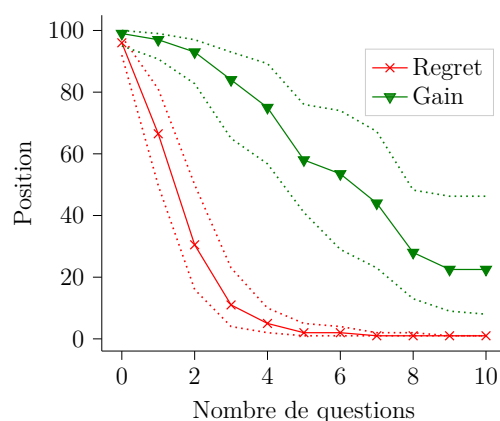
3.9.b : Position

FIGURE 3.9 – Score et position avec de bonnes alternatives et un modèle OWAU

Enfin, sur les Figures 3.10 et 3.11 nous pouvons voir qu'avec un modèle redistributif comme l'OWAR, l'approche basée sur un gain donne des recommandations bien plus mauvaises que l'approche basée sur un regret, quel que soit le type d'alternative. Remarquons tout de même que l'approche basée sur un gain s'améliore au fur et à mesure que le nombre de questions augmente, et que nos expériences s'arrêtent seulement après 10 questions.

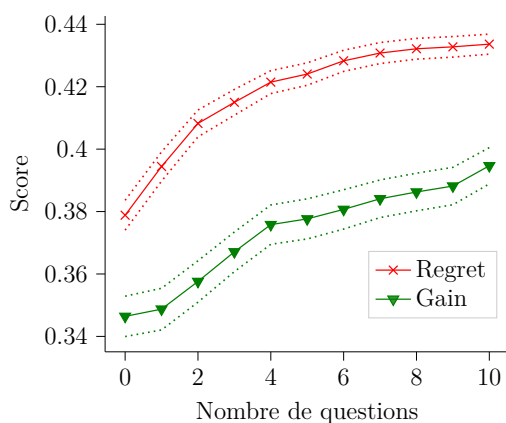


3.10.a : Score

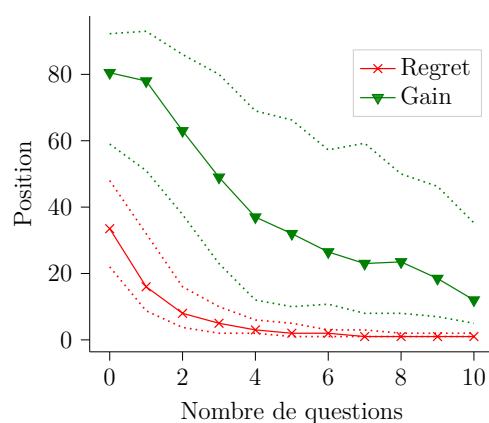


3.10.b : Position

FIGURE 3.10 – Score et position avec de mauvaises alternatives et un modèle OWAR



3.11.a : Score



3.11.b : Position

FIGURE 3.11 – Score et position avec de bonnes alternatives et un modèle OWAR

3.3.3 Discussion des résultats

Les Tables 3.2 et 3.3 résument les différents graphiques, en donnant le score réel moyen et la position réelle médiane en fonction du nombre de questions (5 et 10) et du contexte (modèle et alternatives). En prenant de mauvaises alternatives, l'approche CSS basée sur un gain semble donner des performances nettement meilleures ou similaires dans trois des quatre scénarios. Toutefois, les deux approches

Méthode	Mauvaises alternatives				Bonnes alternatives			
	WSB	WSU	OWAU	OWAR	WSB	WSU	OWAU	OWAR
Gain ₅	.132	.353	.205	.0754	.512	.821	.687	.378
Regret ₅	.130	.283	.205	.0996	.514	.803	.688	.424
Gain ₁₀	.133	.373	.206	.0862	.518	.865	.693	.395
Regret ₁₀	.133	.372	.207	.104	.519	.860	.696	.434

TABLE 3.2 – Score réel moyen de l’alternative recommandée après 5 et 10 questions selon différents contextes

Méthode	Mauvaises alternatives				Bonnes alternatives			
	WSB	WSU	OWAU	OWAR	WSB	WSU	OWAU	OWAR
Gain ₅	2	1	1	58	9.5	3	2	32
Regret ₅	6	6	1	2	7	5	3	2
Gain ₁₀	1	1	1	22.5	2	1	1	12
Regret ₁₀	1	1	1	1	3	1	1	1

TABLE 3.3 – Position réelle médiane de l’alternative recommandée après 5 et 10 questions selon différents contextes

tendent à converger rapidement vers des recommandations de qualité similaire après une dizaine de questions. Si nous entrons dans les détails :

- Pour les modèles équilibrés et déséquilibrés (WSB et WSU), l’approche basée sur un gain est plus intéressante que l’approche basée sur un regret.
- Pour les modèles injustes (OWAU), l’approche basée sur un gain trouve la meilleure alternative après seulement une ou deux questions, ce qui est très intéressant. Cependant, l’approche basée sur un regret la trouve après environ cinq questions, et la différence de score n’est pas réellement significative quel que soit le nombre de questions.
- Pour les modèles redistributifs (OWAR), l’approche basée sur un gain est généralement moins efficace que l’approche basée sur un regret.

En prenant de bonnes alternatives, l'approche basée sur un gain semble moins efficace ou semblable que l'approche basée sur un regret pour trois des scénarios. Néanmoins, les deux approches convergent de nouveau vers la meilleure solution après une dizaine de questions. En rentrant de nouveau dans les détails :

- Pour les modèles équilibrés, l'approche basée sur un gain est légèrement moins bonne que l'approche basée sur un regret.
- Pour les modèles déséquilibrés, l'approche basée sur un gain est meilleure que l'approche basée sur un regret, mais elles convergent rapidement vers la même recommandation.
- Pour les modèles injustes, l'approche basée sur un gain trouve au départ de meilleures alternatives que l'approche basée sur un regret, mais après seulement 3-4 questions, les deux approches deviennent relativement similaires.
- Pour les modèles redistributifs, l'approche basée sur un gain est généralement moins efficace que l'approche basée sur un regret.

Dans la Table 3.4 nous résumons ces différences en termes de performances entre l'approche basée sur un gain et l'approche basée sur un regret. La principale conclusion à tirer de cette table est que la nature du véritable modèle sous-jacent et des alternatives peuvent avoir un impact considérable sur les résultats.

	Mauvaises alternatives	Bonnes alternatives
Équilibrée	++	–
Déséquilibrée	++	+
Injuste	~	~
Redistributive	--	--

TABLE 3.4 – Intérêt de l'approche basée sur un gain (Maximax) par rapport à une approche basée sur un regret (Minimax) en termes de qualité des recommandations (+ : plus intéressante, – : moins intéressante)

À notre avis, cette observation a deux conséquences importantes : la première est que la manière dont les simulations sont effectuées dans le cadre de la validation des méthodes d'élicitation de préférences peut avoir un impact considérable sur

les résultats de ces simulations, ce qui nécessite à la fois des études théoriques plus approfondies sur les situations dans lesquelles une heuristique donnée a des chances de mieux fonctionner, et des simulations prenant en compte un large éventail de situations. La deuxième est, qu'en l'absence d'un fort biais inductif ou d'une connaissance affinée des alternatives, le choix d'une technique d'élicitation peut difficilement être basée sur des exigences de performance, et devrait donc se concentrer sur les axiomes que nous voulons satisfaire dans un problème donné.

3.4 Conclusion et perspectives

Nous avons étudié dans ce chapitre l'utilisation d'une approche optimiste utilisant la règle de décision du Maximax pour recommander une alternative dans une élicitation incrémentale robuste de préférence, au lieu d'une approche pessimiste utilisant un regret Minimax. Nous avons démontré qu'une telle approche optimiste possède les mêmes propriétés de convergence que l'approche standard basée sur les regrets, et qu'elle présente des propriétés intéressantes en termes d'optimalité et de calcul. Des expériences sur des données simulées ont montré qu'une approche optimiste peut être plus efficace dans certains contextes, notamment en fonction du type d'alternative et de modèle de préférence. Plus généralement, notre travail montre que le choix d'une règle de décision n'est pas anodin, et qu'il est intéressant de s'intéresser au type de données, ainsi qu'aux propriétés que l'on souhaite sur nos alternatives.

Points importants vus dans ce chapitre :

- Méthode d'élicitation robuste incrémentale basée sur le gain Maximax.
- Des propriétés différentes, une même convergence et un coût computationnel plus intéressant.
- Performances différentes par rapport à l'approche basée sur un regret Minimax en fonction du modèle de préférence et du type d'alternatives.

Perspectives :

- Nos travaux ont montré que le choix des alternatives a un impact sur les performances des deux approches. Nous pensons qu'il pourrait être intéressant d'étudier plus précisément l'influence des alternatives sur le calcul du regret. Cela pourrait être utile pour déterminer la meilleure stratégie de choix des alternatives dans de futurs travaux. La Section 3.2 donne quelques avantages et inconvénients en termes de propriétés et d'axiomes qui sont similaires aux avantages et inconvénients des approches optimistes et pessimistes trouvables dans d'autres contextes (Hüllermeier *et al.*, 2019).
- La sélection de la règle de décision selon une exigence de performance nécessite clairement des études théoriques plus approfondies concernant les situations dans lesquelles les différentes heuristiques sont plus performantes. Une manière plus empirique de résoudre cette question pourrait consister à caractériser les problèmes d'élicitation par diverses mesures de qualité (voir, par exemple, Ciomek *et al.* (2017)) et de voir si nous pouvons prédire la stratégie optimale/gagnante à partir de cela, en nous inspirant des méthodes d'apprentissage automatique (Tornede *et al.*, 2020).
- Notons que si nous avons examiné le problème avec une approche gourmande, dont l'intérêt est son efficacité et son agnosticisme par rapport au nombre de questions restantes, il peut également être intéressant (mais aussi beaucoup plus difficile) de considérer la version séquentielle de notre problème de décision (voir, par exemple, Holloway et White III (2003)).

Contributions :

- Adam, Loïc et Destercke, Sébastien. Incremental elicitation of preferences : optimist or pessimist? Dans : *Algorithmic Decision Theory : 7th International Conference, ADT 2021, Toulouse, France, 3-5 novembre 2021*. Springer International Publishing, 2021. p. 71-85.

Manipuler l'incohérence

4.1 Introduction

Dans ce chapitre, nous nous intéressons à la manipulation de l'incohérence apparaissant lors d'un processus d'élicitation. En se plaçant dans le cadre de l'élicitation incrémentale ensembliste (Boutillier, 2013; Benabbou, 2017; Benabbou et Perny, 2018), de l'incohérence apparaît lorsque, pour une fonction d'agrégation de critères f_ω donnée, aucune des valeurs de ω ne peut correspondre à toutes les préférences observées du décideur. Cette incohérence peut avoir deux sources : le décideur qui commet des erreurs (incertitude, manque de concentration, changement de point de vue...) ou l'analyste qui fait un mauvais choix de modèle de préférence f_ω , ne permettant pas de représenter correctement les préférences du décideur. L'incohérence dans ce cadre particulier est un signal binaire tout ou rien, et les stratégies courantes pour traiter l'incohérence incluent la suppression ou la modification de certaines informations sur les préférences, soit en minimisant une fonction objectif telle que le nombre de préférences non respectées (Mousseau *et al.*, 2003; Schmitt *et al.*, 2006), soit en interagissant directement avec le décideur pour déterminer les informations préférentielles erronées (Mayag *et al.*, 2011).

Avec l'heuristique de la stratégie de la solution courante (CSS) (Boutillier *et al.*, 2006), couramment utilisée dans ces approches pour choisir des questions optimales, il n'est plus possible de détecter l'incohérence : le processus d'élicitation converge vers une mauvaise estimation des préférences du décideur, quoique fasse le décideur ou l'analyste. Dans le Chapitre 2, nous avons étudié une autre approche utilisant

des distributions de possibilité pour modéliser des informations préférentielles incertaines, et la théorie des possibilités (Dubois et Prade, 1998) pour raisonner avec les informations fournies. Cette approche permet de détecter l'incohérence et de la quantifier graduellement, quelle que soit la source de l'incohérence. Si arrêter l'élicitation dès que l'incohérence détectée dépasse un certain seuil permet d'éviter la dégradation de la recommandation finale, il serait intéressant de restaurer la cohérence en ne manipulant ou en ne gardant que les réponses correctes du décideur, ou en choisissant un modèle de préférence puissant correctement modéliser les préférences du décideur. Ainsi, la qualité de la recommandation finale serait augmentée, et il serait même possible de continuer le processus d'élicitation.

La théorie des possibilités met à notre disposition plusieurs outils pour traiter une telle incohérence, ces outils étendant les opérateurs de la théorie des ensembles et de la logique tels que la conjonction et la disjonction. En particulier, nous examinons la résolution des incohérences provenant du décideur du point de vue de la fusion d'information (Sous-section 1.4.3 et Dubois *et al.* (2016)), car le traitement des incohérences est un problème typique de ce type d'approche. Bien que l'idée d'appliquer la théorie des possibilités à la modélisation des préférences ne soit pas nouvelle (Prade *et al.*, 2009b), ce chapitre enrichit cette approche en considérant des méthodes issues de la fusion d'information pour traiter les incohérences, et en vérifiant par des expériences synthétiques que ces propositions méthodologiques ont un certain intérêt empirique.

Après avoir introduit dans la Section 4.2 des compléments sur la théorie des possibilités, la Section 4.3 discute ensuite des incohérences possibles dans les informations préférentielles observées et de la manière de les traiter, que ce soit avec ou sans les opérateurs de fusion d'information. Enfin, la Section 4.4 présente quelques expériences démontrant l'intérêt potentiel de notre approche pour traiter les incohérences. Bien que nous nous concentrons sur des modèles numériques associant une valeur numérique à chaque alternative, la plupart des notions de ce chapitre peuvent être facilement appliquées à des modèles non numériques, en particulier celles de la Section 4.3. Notons que dans la Section de conclusion 4.5 nous mettons en avant des perspectives intéressantes pour réparer une erreur de modèle ou pour distinguer les deux types d'erreur.

4.2 Compléments en théorie des possibilités

Nous avons déjà présenté dans la Sous-section 1.4.2 la théorie des possibilités. Ici, nous introduisons deux nouvelles notions : l'espérance inférieure et l'espérance supérieure. Ces deux notions ont déjà été utilisées implicitement dans la Section 2.3 pour déterminer le regret Minimax dans une élicitation possibiliste.

Définition 37. *Espérances inférieure et supérieure*

Soit une fonction $f_\omega : \mathbb{R}^q \rightarrow \mathbb{R}$ dont le paramètre $\omega \in \Omega$ n'est pas connu avec précision. $\underline{\mathbb{E}}_\pi[f_\omega]$ et $\overline{\mathbb{E}}_\pi[f_\omega]$ sont respectivement l'espérance inférieure et l'espérance supérieure de f_ω , définis par :

$$\underline{\mathbb{E}}_\pi[f_\omega] = \int_0^1 \min_{\omega \in \pi^\alpha} f_\omega \, d\alpha, \quad (4.1)$$

$$\overline{\mathbb{E}}_\pi[f_\omega] = \int_0^1 \max_{\omega \in \pi^\alpha} f_\omega \, d\alpha, \quad (4.2)$$

où π est une distribution de possibilité associée à la fonction f_ω , et π^α est la coupe-alpha de cette distribution de possibilité.

Notons que si π prend un nombre fini de valeurs distinctes $1 = \alpha_1 > \dots > \alpha_n > \alpha_{n+1} = 0$ (comme c'est le cas dans ce manuscrit), les Équations (4.1) et (4.2) deviennent :

$$\underline{\mathbb{E}}_\pi[f_\omega] = \sum_{i=1}^n (\alpha_i - \alpha_{i+1}) \min_{\omega \in \pi^{\alpha_i}} f_\omega, \quad (4.3)$$

$$\overline{\mathbb{E}}_\pi[f_\omega] = \sum_{i=1}^n (\alpha_i - \alpha_{i+1}) \max_{\omega \in \pi^{\alpha_i}} f_\omega. \quad (4.4)$$

Exemple 34 (Bornes de l'espérance). *Nous illustrons la notion d'espérance inférieure et supérieure à partir de l'Exemple 28, dont nous rappelons la distribution de possibilité π_\cap sur la Figure 4.1. Nous nous intéressons ici à la fonction de score de GO, supposée être une moyenne pondérée et notée $WS_\omega(GO) = 7\omega_1 + 3\omega_2$, représentée ici sur la même figure. Nous avons trois valeurs distinctes pour π_\cap : 1, 0.1 et 0.03. Nous avons donc trois coupes-alpha : $\pi_\cap^1 = [2/3, 1]$, $\pi_\cap^{0.1} = [5/9, 1]$ et $\pi_\cap^{0.03} = [0, 1]$.*

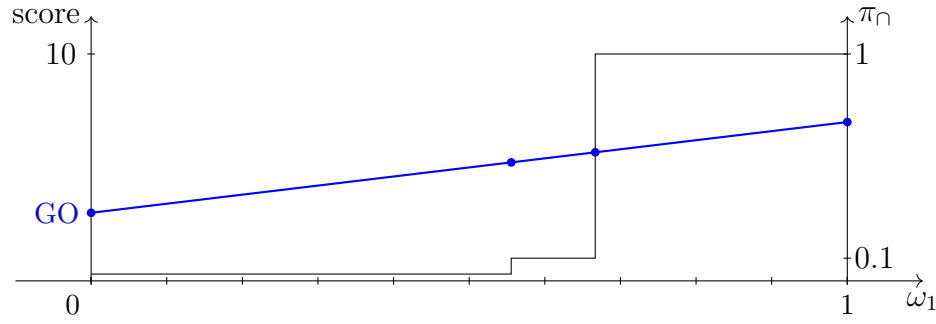


FIGURE 4.1 – Rappel de la distribution de possibilité π_\cap de l'Exemple 28 et superposition du score de l'alternative GO

Le calcul de l'espérance inférieure est le suivant :

$$\begin{aligned} \underline{\mathbb{E}}_{\pi_\cap}[f_\omega(GO)] &= (1 - 0.1) \min_{\omega \in [2/3, 1]} f_\omega(GO) + (0.1 - 0.03) \min_{\omega \in [5/9, 1]} f_\omega(GO) \\ &\quad + (0.03 - 0) \min_{\omega \in [0, 1]} f_\omega(GO) \\ &= 0.9(7 \times 2/3 + 3 \times 1/3) + 0.07(7 \times 5/9 + 3 \times 4/9) + 0.03(3 \times 1) \\ &= 0.9 \times 17/3 + 0.07 \times 47/9 + 0.03 \times 3 \\ &\approx 5.55. \end{aligned}$$

Le calcul de l'espérance supérieure est ici trivial : $\max_{\omega \in \pi_\cap^\alpha} f_\omega(GO) = 7$, quelle que soit la coupe-alpha. Par conséquent, $\overline{\mathbb{E}}_{\pi_\cap}[f_\omega(GO)] = 7$.

À partir de ces notations, nous pouvons redéfinir les règles de décision que nous avons présentées dans la Sous-section 1.2.3, mais cette fois-ci en présence d'une distribution de possibilité (et donc d'informations incertaines) :

- Le Maximax (MM), qui consiste à prendre l'alternative qui maximise le score f_ω dans le meilleur des scénarios. C'est une approche optimiste pour choisir une alternative. Nous avons ici $x^* = \arg \max_{x \in \mathbb{X}} \overline{\mathbb{E}}(f_\omega(x))$.
- Le Maximin (Mm), qui consiste à prendre l'alternative qui maximise le score f_ω dans le pire des scénarios. C'est une approche pessimiste pour choisir une alternative. Nous avons ici $x^* = \arg \max_{x \in \mathbb{X}} \underline{\mathbb{E}}(f_\omega(x))$.
- Le regret Minimax (mMR), qui consiste à prendre l'alternative qui minimise son regret maximal avec les autres alternatives. C'est une approche pessimiste,

mais moins conservatrice que le Maximin, pour choisir une alternative. Nous avons déjà défini dans la Sous-section 2.3.1 une extension du regret Minimax au cas possibiliste, utilisant la notion d'espérance supérieure dans le calcul du EPMR et du EMR (Équations (2.9) - (2.10)).

4.3 Résoudre l'incohérence par la fusion d'information

Comme nous l'avons déjà indiqué dans le Chapitre 2, l'un des intérêts de l'approche possibiliste est de pouvoir modéliser de l'incertitude dans les réponses du décideur et ainsi d'éviter de s'écarter de la bonne estimation de ω^* en cas d'erreur. Si certaines réponses sont incohérentes entre elles, une distribution possibiliste permet de quantifier l'incohérence de manière progressive, plutôt que d'avoir une information tout ou rien comme font les approches ensemblistes.

Les incohérences observées dans les informations préférentielles et dans les inférences qui s'ensuivent s'expliquent principalement de deux manières :

- **Erreur de modèle**, provenant d'un mauvais choix de fonction d'agrégation de critères f_ω par l'analyste. Ainsi, la fonction n'est pas en mesure de rendre compte des informations préférentielles données, même si celles-ci sont toutes correctes. Par exemple, une intégrale de Choquet est capable de modéliser des préférences complexes, y compris des interactions positives ou négatives entre critères, tandis qu'une moyenne pondérée ne permet pas de modéliser ces interactions entre critères.
- **Erreur du décideur**, provenant du décideur qui a commis des erreurs lors de la formulation de ses préférences, rendant ainsi les réponses conjointement incompatibles avec tout modèle $\omega \in \Omega$. Il convient de noter que ces erreurs peuvent avoir différentes origines : le décideur pouvait ne pas être certain de ses réponses, ou a peut-être changé de point de vue au cours du processus d'élicitation.

Il est possible de tenter de réparer les erreurs, et donc de rendre les informations préférentielles cohérentes, de plusieurs manières :

- Modifier ou élargir l'espace de modélisation, c'est-à-dire en passant vers des modèles plus expressifs, par exemple d'une moyenne pondérée vers une intégrale de Choquet k -monotone (Herin *et al.*, 2022). Nous abordons cette stratégie de réparation dans la Section 4.5.
- Changer la manière dont nous fusionnons les informations préférentielles (Destercke *et al.*, 2008; Pichon *et al.*, 2014). Il existe notamment plusieurs opérateurs de fusion d'information qui permettent de gérer des informations incohérentes entre elles (Dubois *et al.*, 2016). Nous allons voir dans la Sous-section 4.3.2 deux de ces opérateurs permettant de rétablir la cohérence.
- Supprimer directement les réponses erronées du décideur de manière minimale, rendant ainsi les réponses restantes cohérentes entre elles (Mousseau *et al.*, 2003). Un des opérateurs de la Sous-section 4.3.2 permet de déterminer des sous-ensembles cohérents de réponses. Sachant que les bonnes réponses sont cohérentes entre elles, nous pouvons supprimer des mauvaises réponses si le bon sous-ensemble est choisi (mais pas forcément toutes les mauvaises réponses, certaines pouvant être cohérentes avec les bonnes)¹.

Nous pouvons remarquer que certaines stratégies semblent plus adaptées à un certain type d'erreur : modifier l'espace de modélisation semble mieux convenir aux erreurs de modèle, tandis que supprimer les réponses erronées ou changer la règle de fusion semble mieux convenir aux erreurs du décideur. Néanmoins, différencier les deux types d'erreurs sans avoir d'interactions avec le décideur ou sans informations supplémentaires est en soi un problème difficile, que nous évoquons dans la Section 4.5. Dans ce chapitre, nous nous intéressons principalement aux erreurs qui proviennent du décideur et nous examinons des stratégies pour traiter ces erreurs dans le cadre de la théorie des possibilités, avec des outils de fusion d'information associés. Concrètement, nous examinons le cas où π_{\cap} est sous-normalisée, c'est-à-dire $\text{Inc}(\pi_{\cap}) > 0$.

1. Remarquons que dans notre contexte, il serait aussi possible de corriger directement les réponses erronées, permettant de transformer de l'information erronée en information utile. En effet, comme le décideur doit choisir à chaque fois son alternative préférée entre x et y , si nous savons que $x \succeq y$ est faux, alors $x \preceq y$ est nécessairement vrai. Néanmoins, si supprimer une bonne information ne va pas créer de nouveaux conflits, « corriger » une bonne information peut créer de nouveaux conflits.

Nous détaillerons dans cette section deux stratégies : ne pas remettre en question la fusion conjonctive des informations et adapter notre inférence en se basant sur les Équations (4.1) et (4.2) ; ou changer l'opérateur de fusion des informations, modifiant ainsi nos hypothèses sur les sources d'information. La première solution est facile à mettre en œuvre, mais ne fournit que peu d'éléments sur la source des incohérences, tandis que la seconde stratégie est plus complexe, mais fournit une analyse supplémentaire de l'incohérence. Il convient également de noter que les deux stratégies ne sont pas incompatibles : nous utilisons par exemple dans la Section 4.4 la première stratégie pour obtenir des alternatives à présenter au décideur malgré l'incohérence observée, et la seconde pour obtenir une recommandation finale.

4.3.1 Inférer malgré les incohérences

L'existence d'une incohérence $\text{Inc}(\pi_\cap) > 0$ implique que $\pi_\cap^1 = \emptyset$. Cela signifie que si nous voulons faire des inférences sans corriger les incohérences comme vu dans les Équations (4.3)-(4.4), il faut définir des minima et des maxima sur l'ensemble vide. Bien qu'il soit possible de définir une infinité de stratégies pour tenir compte de cette situation, les deux stratégies suivantes sont des solutions classiques :

- Considérer que $\min_{\omega \in \emptyset} f_\omega = \max_{\omega \in \emptyset} f_\omega = 0$. Cela revient simplement à ignorer les informations incohérentes. Ceci est, dans une certaine mesure, similaire aux procédures d'inférence en logique possibiliste en présence d'incohérence (Dubois et Prade, 2004). Un avantage est que, si $\pi' \subseteq \pi$, alors $[\underline{\mathbb{E}}_{\pi'}[f_\omega], \overline{\mathbb{E}}_{\pi'}[f_\omega]] \subseteq [\underline{\mathbb{E}}_\pi[f_\omega], \overline{\mathbb{E}}_\pi[f_\omega]]$, ce qui maintient une certaine monotonie par rapport au gain d'information, étant donné que des distributions de possibilité plus précises conduiront à des déductions plus précises.
- Considérer que $\min_{\omega \in \emptyset} f_\omega = \min_{\omega \in \Omega} f_\omega$ et $\max_{\omega \in \emptyset} f_\omega = \max_{\omega \in \Omega} f_\omega$. Cela revient à transformer le conflit en ignorance et à avoir un point de vue très conservateur à ce sujet. Nous pouvons également considérer qu'il s'agit d'une normalisation de la distribution de possibilités en prenant $\pi' = \pi + \text{Inc}(\pi_\cap)$.

Cette manière de résoudre les incohérences ne modifie ni notre hypothèse par rapport aux réponses et aux informations données précédemment (elles ne sont pas modifiées), ni la façon dont nous les combinons (conjonctivement). Il convient

de noter que cette approche évite en quelque sorte de rechercher les sources d'incohérence, et qu'elle ne fait que les ignorer ou les transformer en ignorance (un concept différent). Par conséquent, ces stratégies ne semblent légitimes que lorsque l'incohérence et ses effets sont susceptibles d'être limités et qu'il n'est pas nécessaire d'analyser en détails la situation conflictuelle.

Exemple 35 (Inférer avec de l'incohérence sans modifier les informations préférentielles). *Nous reprenons l'Exemple 34, avec comme différence que la distribution de possibilité n'est plus normalisée, comme nous pouvons le voir sur la Figure 4.2. Comme nous avons $Inc(\pi_\cap) = 0.7$, nous devons déterminer les extrema sur l'ensemble vide pour inférer.*

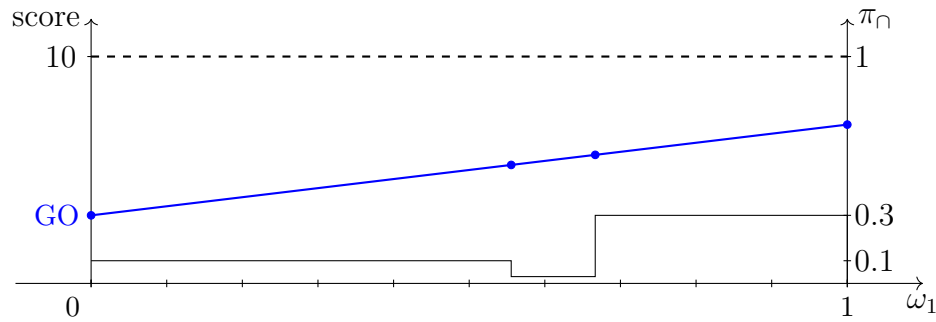


FIGURE 4.2 – Distribution non normalisée π_\cap de l'Exemple 35 et superposition du score de l'alternative GO

Nous avons trois valeurs distinctes pour π_\cap : 0.3, 0.1 et 0.03. Nous avons donc quatre coupes-alpha : $\pi_\cap^1 = \emptyset$, $\pi_\cap^{0.3} = [2/3, 1]$, $\pi_\cap^{0.1} = [0, 5/9] \cup [2/3, 1]$ et $\pi_\cap^{0.03} = [0, 1]$. Nous définissons alors le minimum et le maximum sur l'ensemble vide de deux manières différentes :

- $\min_{\omega \in \emptyset} f_\omega(GO) = \max_{\omega \in \emptyset} f_\omega(GO) = 0$.
- $\min_{\omega \in \emptyset} f_\omega(GO) = \min_{\omega \in \Omega} f_\omega(GO) = 3$ (et pour le max 7).

Nous pouvons alors déterminer l'espérance inférieure :

$$\begin{aligned}
\mathbb{E}_{\pi_{\cap}}[f_{\omega}(GO)] &= 0.7 \min_{\omega \in \emptyset} f_{\omega}(GO) + 0.2 \min_{\omega \in [2/3,1]} f_{\omega}(GO) \\
&+ 0.07 \min_{\omega \in [0,5/9] \cup [2/3,1]} f_{\omega}(GO) + 0.03 \min_{\omega \in [0,1]} f_{\omega}(GO) \\
&= 0.7 \min_{\omega \in \emptyset} f_{\omega}(GO) + 0.2 \times 17/3 + 0.07 \times 3 + 0.03 \times 3 \\
&\approx 0.7 \min_{\omega \in \emptyset} f_{\omega}(GO) + 1.43.
\end{aligned}$$

L'espérance inférieure est égale à environ 1.43 si on prend $\min_{\omega \in \emptyset} = 0$ (le conflit est ignoré), ou à environ 3.53 si on prend $\min_{\omega \in \emptyset} f_{\omega}(GO) = \min_{\omega \in \Omega} f_{\omega}(GO) = 3$ (le conflit équivaut à de l'ignorance).

Pour l'espérance supérieure :

$$\begin{aligned}
\overline{\mathbb{E}}_{\pi_{\cap}}[f_{\omega}(GO)] &= 0.7 \max_{\omega \in \emptyset} f_{\omega}(GO) + 0.2 \max_{\omega \in [2/3,1]} f_{\omega}(GO) \\
&+ 0.07 \max_{\omega \in [0,5/9] \cup [2/3,1]} f_{\omega}(GO) + 0.03 \max_{\omega \in [0,1]} f_{\omega}(GO) \\
&= 0.7 \max_{\omega \in \emptyset} f_{\omega}(GO) + 0.2 \times 7 + 0.07 \times 7 + 0.03 \times 7 \\
&= 0.7 \max_{\omega \in \emptyset} f_{\omega}(GO) + 2.1.
\end{aligned}$$

L'espérance supérieure est donc égale à environ 2.1 si on prend $\max_{\omega \in \emptyset} = 0$, ou à 7 si on prend $\max_{\omega \in \emptyset} f_{\omega}(GO) = \max_{\omega \in \Omega} f_{\omega}(GO) = 7$.

Une autre façon d'inférer malgré les incohérences consisterait à normaliser la distribution π_{\cap} , pour revenir à une situation cohérente. Cette normalisation peut se faire de différentes manières (Oussalah, 2002), mais peuvent être plus difficiles à interpréter que nos deux propositions. C'est pourquoi nous ne les étudions pas ici, bien que la deuxième façon de traiter l'incohérence puisse être considérée comme une normalisation spécifique, comme nous l'avons déjà mentionné.

4.3.2 Résoudre les incohérences par la fusion d'information

Une deuxième stratégie pour résoudre les incohérences consiste à modifier la façon dont nous combinons les différentes sources d'information pour que l'incohérence disparaisse. Une telle approche ne modifie pas les informations préférentielles

que nous recevons, mais constitue un outil pratique pour formuler ou tester différentes hypothèses à leur sujet. Par exemple, une règle conjonctive aboutissant à π_\cap suppose que toutes les sources sont fiables et fournissent des informations dignes de confiance. Il est clair que si $\text{Inc}(\pi_\cap) > 0$, cette hypothèse ne peut pas être vraie, et d'autres peuvent être étudiées. Vouloir que la fusion soit cohérente avec chaque information initiale est discutable en modélisation des préférences, car si le modèle est fiable, l'incohérence résulte nécessairement d'une erreur dans les réponses du décideur. Toutefois, il est également raisonnable de supposer que la plupart des réponses des décideurs sont correctes. Nous examinons ci-dessous quelques opérateurs de fusion qui s'inscrivent dans ce cadre.

ℓ -parmi- k

Il est tentant d'utiliser un opérateur puissant résoudre les incohérences dans une certaine mesure, mais dont le résultat tend à être cohérent avec la plupart (mais pas nécessairement toutes) des informations préférentielles initiales. Comme il est également difficile de savoir quelle réponse du décideur est erronée, il est naturel d'envisager des opérateurs qui traitent les sources de manière anonyme (c'est-à-dire dont le résultat reste inchangé si les indices des informations préférentielles π_i sont permutés). L'hypothèse selon laquelle ℓ sources parmi les k considérées sont correctes correspond à un tel opérateur. Si $\mathcal{S} = \{\pi_1, \dots, \pi_k\}$ sont les sources d'information considérées, alors la distribution résultant de la supposition ℓ/k est :

$$\pi_{\ell/k}(\omega) = \bigcup_{\mathcal{L} \subseteq \mathcal{S}, |\mathcal{L}|=\ell} \left(\bigcap_{\pi_i \in \mathcal{L}} \pi_i(\Omega) \right), \quad (4.5)$$

où \cap et \cup sont remplacés par une T-norme et une T-conorme duale correspondante (dans notre cas, nous prenons la T-norme produit et la T-conorme somme probabiliste, présentées dans la Sous-section 1.4.4). Cet opérateur de fusion est un exemple d'opérateur des ℓ -quota appliqué à la théorie des possibilités (Sous-section 1.4.3).

Idéalement, une réparation minimale devrait consister à trouver une valeur ℓ la plus proche possible de k . Deux étapes sont alors nécessaires pour déterminer ℓ^* :

1. Trouver la plus grande valeur ℓ^* tel que $\pi_{\ell^*/k}$ soit normalisée.
2. Déterminer la distribution de possibilité $\pi_{\ell^*/k}$.

Bien que ces opérations puissent sembler complexes, nous proposons une méthode efficace en pratique. Nous supposons que, pour chaque élément Ω_i de la partition $\{\Omega_1, \dots, \Omega_P\}$ de Ω (déterminée à partir des différentes informations préférentielles), nous avons un vecteur associé $\vec{\pi}_i = (\pi_1(\Omega_i), \dots, \pi_k(\Omega_i))$. Ce vecteur permet, pour un Ω_i donné, de savoir quelles informations préférentielles indiquent que ω^* appartient à Ω_i . L'algorithme 4 permet de trouver facilement ℓ^* , basé sur l'idée simple que $\pi_{\ell/k}$ sera normalisé s'il existe au moins un élément Ω_i tel qu'au moins ℓ niveaux de confiance ont une valeur d'un ². L'Algorithme 4 consiste alors à trouver la plus grande valeur satisfaisant cette contrainte. Il est d'une complexité linéaire en fonction du nombre P d'éléments, et donc assez rapide une fois que P est fixé.

Algorithme 4 : Algorithme pour trouver ℓ^*

Données : Sources $\mathcal{S} = \{\pi_1, \dots, \pi_k\}$

Résultat : ℓ^* maximal pour restaurer la cohérence

$\ell^* = k$;

pour $j \in \{1, \dots, P\}$ **faire**

si $|\{\pi_i(\Omega_j) : \pi_i(\Omega_j) = 1, i = 1, \dots, k\}| < \ell^*$ **alors**
 | $\ell^* = |\{\pi_i(\Omega_j) : \pi_i(\Omega_j) = 1, i = 1, \dots, k\}|$

fin

fin

La stratégie de réparation ℓ -parmi- k consiste alors à calculer simplement $\pi_{\ell^*/k}$ et à inférer en utilisant cette distribution. Comme $\pi_{\ell^*/k}$ peut être évalué par élément, son calcul reste abordable tant que k n'est pas trop élevé, ce qui est généralement le cas lorsque que nous demandons des informations au décideur.

Exemple 36 (Réparation ℓ -parmi- k). *Nous supposons maintenant que le décideur donne quatre informations préférentielles, avec des niveaux de confiance α_i qui sont respectivement 0.9, 0.5, 0.7, et 0.3. Nous supposons que les préférences du décideur sont modélisées par une moyenne pondérée de paramètre $\omega^* = (0.6, 0.4)$. Les réponses du décideur sont présentées sur la Figure 4.3, sous forme d'ensembles et non de distributions de possibilité pour faciliter la lecture initiale.*

2. Pour n'importe quelle T-norme, nous avons $a_1 * \dots * a_\ell = 1 \iff a_1 = \dots = a_\ell = 1$. Dès lors, nous ne pouvons pas avoir un résultat unitaire si les ℓ niveaux de confiance ne sont pas un.

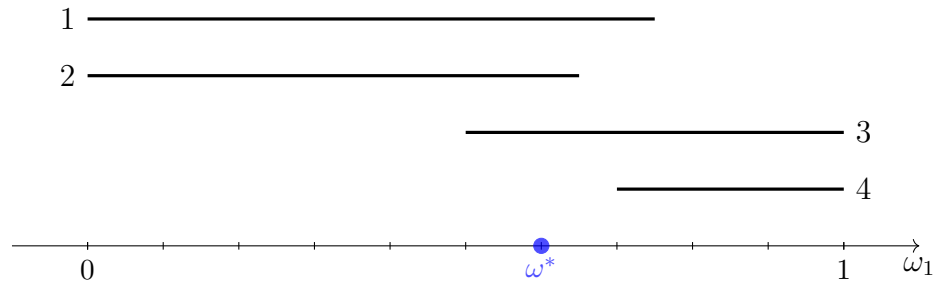


FIGURE 4.3 – Réponses données par le décideur dans l'Exemple 36, la réponse 4 étant fausse

Nous pouvons voir que la quatrième réponse est erronée, le décideur étant incertain ou non concentré, ce qui a conduit à la détection d'une certaine incohérence, comme le montre la Figure 4.4, avec $Inc(\pi_\cap) = 0.3$. Notre objectif est de traiter les incohérences par la fusion d'information.

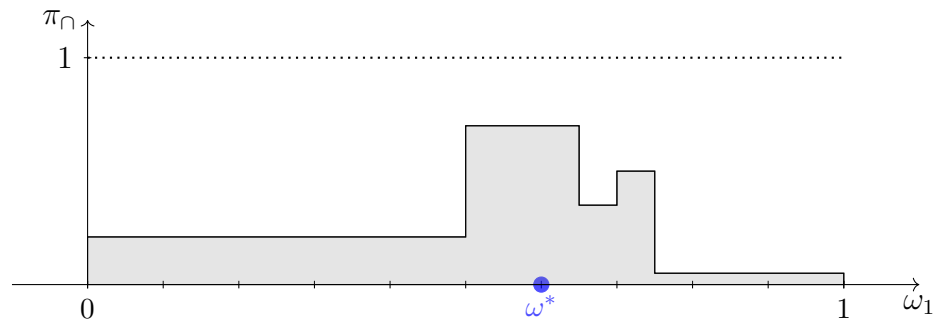


FIGURE 4.4 – Fusion d'information préférentielle avec incohérence dans l'Exemple 36

Nous utiliserons ici l'algorithme de réparation ℓ -parmi- k . Dans ce cas, nous pouvons facilement déterminer le ℓ pour rétablir la cohérence : $\ell = 3$, donc la cohérence est rétablie en supprimant seulement une seule réponse, plus particulièrement la réponse 2 ou la réponse 4. Nous calculons alors $\pi_{3/4}$ selon l'équation (4.5). Étant donné $\mathcal{S} = \{\pi_1, \dots, \pi_4\}$, la première étape est de déterminer tous les sous-ensembles $\mathcal{L} \subseteq \mathcal{S}$ tels que $|\mathcal{L}| = 3$. Nous obtenons alors $\mathcal{L}_1 = \{\pi_1, \pi_2, \pi_3\}$, $\mathcal{L}_2 = \{\pi_1, \pi_2, \pi_4\}$, $\mathcal{L}_3 = \{\pi_1, \pi_3, \pi_4\}$ et $\mathcal{L}_4 = \{\pi_2, \pi_3, \pi_4\}$. Nous déterminons alors les quatre distributions de possibilité associées $\pi_{\mathcal{L}_i}$ avec une T -norme produit. Par

exemple, $\pi_{\mathcal{L}_1}(\omega) = \prod_{i=1}^3 \pi_i(\omega)$. Après cela, nous déterminons $\pi_{3/4}$ avec une T -conorme somme probabiliste. Comme une T -conorme est commutative et associative, cela peut être fait itérativement en l'appliquant sur des paires de distributions, en se rappelant que combiner deux distributions de possibilité π_1 et π_2 avec cette T -conorme se fait ainsi :

$$\pi_{\cup}(\omega) = \pi_1(\omega) + \pi_2(\omega) - \pi_1(\omega) \cdot \pi_2(\omega).$$

Le résultat final est illustré sur la Figure 4.5. Comme prévu, la cohérence est rétablie et la distribution de possibilité qui en résulte est maximale sur deux sous-ensembles distincts de Ω , dans lesquels au moins trois réponses sont cohérentes, ce qui est le cas pour les réponses 1, 2 et 3 lorsque $\omega_1 \in [0.5, 0.65]$ et pour les réponses 1, 3 et 4 lorsque $\omega_1 \in [0.7, 0.75]$. Comme l'indique cette remarque, cette approche ne garantit pas que l'ensemble des valeurs de ω les plus plausibles est convexe, même si chaque réponse pointe individuellement vers un sous-ensemble convexe de valeurs de ω les plus plausibles. Cependant, ce type de cas ne se produit qu'en cas de désaccord, et il serait possible de l'analyser en détail avec le décideur.

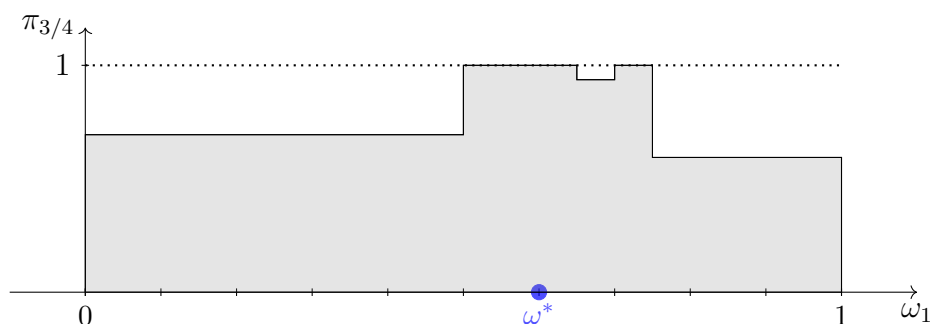


FIGURE 4.5 – Fusion d'information préférentielle dans l'Exemple 36 corrigée avec l'algorithme de réparation ℓ -parmi- k

Sous-ensembles cohérents maximaux

Plutôt que d'utiliser un opérateur de conjonction qui équivaut à un ET logique, nous pouvons utiliser d'autres opérateurs logiques qui tentent de traiter les situations conflictuelles et incohérentes, principalement en trouvant des compromis entre les

comportements conjonctifs et disjonctifs. Une approche courante consiste par exemple à utiliser la notion de sous-ensembles cohérents maximaux (MCS), que nous avons déjà présenté dans la Sous-section 1.4.3. Dans notre contexte, étant donné un ensemble $\mathcal{S} = \{\pi_1, \dots, \pi_k\}$ d'informations préférentielles, nous définissons un sous-ensemble $\mathcal{L} \subseteq \mathcal{S}$ comme étant un MCS si le résultat

$$\pi_{\mathcal{L}} = \bigcap_{\pi_i \in \mathcal{L}} \pi_i$$

de leur combinaison est tel que³ $\text{Inc}(\pi_{\mathcal{L}}) = 0$ et $\text{Inc}(\pi_{\mathcal{K}}) > 0$ pour tout $\mathcal{K} \supset \mathcal{L}$.

Une manière classique de rétablir la cohérence par la fusion d'information, héritée d'idées en logique (Rescher et Manor, 1970), consiste simplement à prendre toutes les MCSs et à effectuer une disjonction de toutes les conjonctions des MCSs, comme vu dans l'Équation (1.26). Cependant, une telle approche donnera généralement des résultats assez imprécis en présence de valeurs aberrantes ou d'erreurs, principalement parce qu'elle aboutit à une distribution dont l'intersection avec n'importe laquelle des informations préférentielles initiales n'est pas vide. Nous voulons donc considérer un seul MCS, idéalement contenant au moins toutes les réponses correctes du décideur.

Lister tous les MCSs d'un ensemble d'informations est très coûteux : contrairement à l'Algorithme 4, nous devons considérer tous les sous-ensembles possibles, donc au pire 2^K sous-ensembles. En supposant que le nombre d'informations reste raisonnable, il est possible de dresser la liste de tous les MCSs. Une stratégie consisterait à ne considérer que les MCSs de taille ℓ donnée par l'Algorithme 4, en supposant que la plupart des informations soient correctes. Cependant, comme nous le verrons dans la Section 4.4, bien que cette heuristique puisse être intéressante lorsqu'elle est associée au degré de confiance moyen, elle ne donne généralement pas le MCS le plus intéressant.

Exemple 37 (Réparation avec des MCSs). *En conservant le même cadre que dans l'Exemple 36, nous voulons cette fois-ci restaurer la cohérence au moyen d'un MCS, et plus précisément d'un MCS de taille maximale.*

3. Une notion plus faible serait de considérer que $\text{Inc}(\pi_{\mathcal{K}}) < 1$ et $\text{Inc}(\pi_{\mathcal{L}}) = 1$, mais cela ne rétablirait pas la cohérence, et ne serait pas utile ici, car les distributions ont pour support Ω .

Comme dans l'exemple précédent, nous avons quatre réponses, l'une d'entre elle étant incorrecte, et nous savons que $\ell = 3$. Dès lors, nous devons tout d'abord déterminer les MCSs \mathcal{L} tels que $|\mathcal{L}| = 3$. Comme il n'est pas possible d'avoir un MCS \mathcal{L} avec $|\mathcal{L}| > 3$ (sinon ℓ ne serait pas 3), il est suffisant de ne déterminer que les sous-ensembles cohérents, c'est-à-dire les sous-ensembles \mathcal{L} tels que $\text{Inc}(\pi_{\mathcal{L}}) = 0$. $\mathcal{L}_1 = \{\pi_1, \pi_2, \pi_3\}$ et $\mathcal{L}_3 = \{\pi_1, \pi_3, \pi_4\}$ sont les seuls sous-ensembles cohérents de taille ℓ . Nous prenons ensuite le MCS qui maximise la moyenne des niveaux de confiance associés. Nous avons $\alpha_{\mathcal{L}_1} = \frac{\alpha_1 + \alpha_2 + \alpha_3}{3} = 0.7$, et $\alpha_{\mathcal{L}_3} = \frac{\alpha_1 + \alpha_3 + \alpha_4}{3} \approx 0.63$, indiquant que $\hat{\mathcal{L}} = \mathcal{L}_1$.

La distribution de possibilité $\pi_{\hat{\mathcal{L}}}$ est illustrée sur la Figure 4.6. Comme nous pouvons le voir, la cohérence est rétablie et, contrairement à l'algorithme de réparation ℓ -parmi- k , nous avons la garantie que l'ensemble des valeurs de ω les plus plausibles forme un ensemble convexe, si chaque information préférentielle pointe vers un ensemble convexe aussi, grâce à la seule utilisation d'opérateurs conjonctifs.

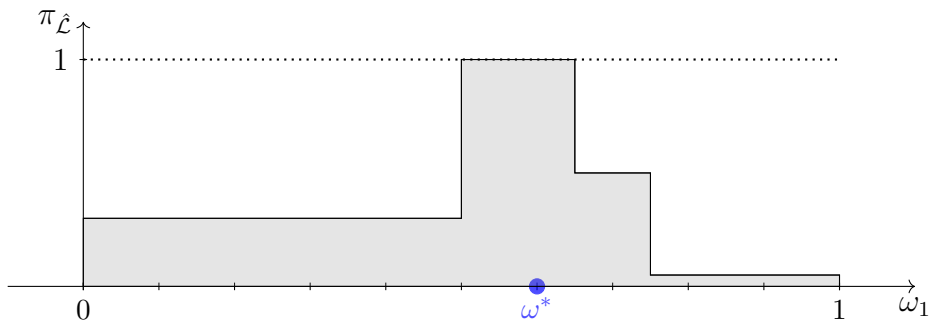


FIGURE 4.6 – Fusion d'information préférentielle dans l'Exemple 37 corrigée avec un MCS

Par rapport aux approches de la Sous-section 4.3.1, la modification de la façon dont nous combinons les éléments d'information est généralement plus intensive sur le plan du calcul, mais présente l'avantage de fournir potentiellement des informations intéressantes au décideur ou à l'analyste. Par exemple, le nombre ℓ résultant de l'Algorithme 4 nous donne une limite inférieure du nombre d'erreurs commises, tandis que l'ensemble des réponses constituant un MCS fournit un sous-ensemble intéressant de réponses que l'on pourrait soumettre au décideur.

4.4 Expériences

Dans cette section, nous réalisons plusieurs expériences synthétiques⁴ pour voir comment nos différentes approches se comportent lorsque des incohérences apparaissent. Ces expériences mettent en avant l'intérêt des approchées présentées dans ce chapitre lors du traitement d'informations préférentielles incertaines.

La confirmation de cet intérêt dans des situations appliquées nécessiterait des expériences dans le monde réel, ce qui dépasse le cadre de cette thèse. Dans ce chapitre, nos expériences se concentrent sur l'évaluation de nos méthodes de traitement des incohérences, les expériences du Chapitre 2 ayant déjà montré que l'approche possibiliste peut être plus performante qu'une approche ensembliste en présence d'erreurs. Afin d'évaluer nos approches, nous allons considérer un problème de recommandation dans lequel une alternative doit être recommandée sur la base des informations préférentielles disponibles.

Si dans la plupart des expériences, nous allons considérer l'algorithme d'élicitation PCSS que nous avons détaillé précédemment (Sous-section 2.3.2), dans la Sous-section 4.4.4 nous allons nous intéresser à plusieurs règles de décision et à de l'élicitation par batch, un cas d'élicitation non incrémentale. Ainsi, toutes les informations préférentielles sont données en une seule fois, contrairement à l'élicitation incrémentale où à chaque étape l'information préférentielle demandée au décideur dépend des questions et des réponses obtenues précédemment. Plus simplement, en élicitation par batch, les réponses du décideur ne vont pas influencer le choix des futures questions.

4.4.1 Protocole expérimental

Nous voulons recommander à un décideur l'alternative qu'il devrait préférer parmi 50 alternatives multicritères qui sont Pareto non dominées (aucune alternative n'est pire que les autres, quelles que soient les préférences du décideur). Chaque alternative a quatre critères avec $\mathcal{X}_i = [0, 1]$. Chaque alternative x est générée aléatoirement par rapport à une loi uniforme, c'est-à-dire $x_i \sim \mathcal{U}(0, 1) \forall i \in \{1, \dots, 4\}$.

4. https://github.com/LoicAdam/Possibilist_Elicitation_Fusion et https://github.com/LoicAdam/Possibilist_Elicitation_Fusion_Random

Nous ajoutons également une autre contrainte sur les critères pour nous assurer que les alternatives ne sont pas Pareto-dominées : pour chaque alternative x , la somme des critères est approximativement égale à 2, c'est-à-dire $\sum_{i=1}^4 x_i \approx 2$. Nous avons également réalisé des expériences avec moins et plus de critères (c'est-à-dire 3 et 5 critères), mais les conclusions obtenues au cours de ces expériences ne diffèrent pas de celles présentées ici.

Dans ces expériences, la fonction d'agrégation de critères est une somme pondérée paramétrée par un ensemble de poids tels que $0 \leq \omega_i \leq 1$ et $\sum_{i=1}^4 \omega_i = 1$. Les poids de chaque fonction d'agrégation sont générés aléatoirement selon une loi de Dirichlet avec l'hyperparamètre $\beta = (1, 1, 1, 1)$. La loi de Dirichlet nous donne des poids dont la somme est égale à un, et cet hyperparamètre spécifique nous garantit que les poids sont générés uniformément sur Ω .

Pour toutes les expériences, sauf une, nous appliquons l'algorithme PCSS présenté dans l'Algorithme 3 pour trouver la recommandation optimale. Pour l'expérience présentée dans la Sous-section 4.4.4, nous utilisons un algorithme d'élicitation non incrémental. Le décideur doit répondre à 15 questions. Les niveaux de confiance α_i fournis avec chaque réponse sont générés aléatoirement en fonction d'un scénario, en utilisant soit une loi bêta $\mathcal{B}(a, b)$ soit une loi uniforme $\mathcal{U}(0, 1)$:

- $\alpha_i \sim \mathcal{B}(7, 2)$ dans un scénario optimiste, où le décideur est certain de ses choix.
- $\alpha_i \sim \mathcal{U}(0, 1)$ dans un scénario intermédiaire.

Notons qu'un scénario pessimiste, dans lequel le décideur est très peu confiant dans ses choix et commet beaucoup d'erreurs, n'est pas réaliste, car le décideur est censé faire de son mieux pour nous aider et ne pas se montrer particulièrement contradictoire.

Nous modélisons l'incertitude du décideur de la manière suivante : étant donné un niveau de confiance α_i , la probabilité que le décideur réponde nécessairement correctement est α_i , et aléatoirement $1 - \alpha_i$. Lorsque le décideur répond au hasard, nous considérons que la probabilité d'obtenir une réponse incorrecte est de 50%. Globalement, le décideur a une probabilité $\alpha_i + \frac{(1-\alpha_i)}{2}$ de répondre correctement, et une probabilité $\frac{(1-\alpha_i)}{2}$ de répondre incorrectement. Nous nous assurons également qu'au moins une réponse soit incorrecte pour conserver l'expérience (sinon, toutes

nos approches de traitement des incohérences aboutissent au même résultat). Un tel échantillonnage est assez commun en théorie des possibilités et avec les fonctions de croyance, où cette probabilité correspond à la probabilité pignistique (Smets *et al.*, 1989) qui est équivalente à la valeur de Shapley en théorie des jeux. Notons que n'importe quelle probabilité dans l'ensemble \mathcal{P} induit par la distribution de possibilité serait aussi valide, par exemple en échantillonnant une probabilité $\beta_i \in [\alpha_i, 1]$ que la réponse du décideur soit correcte pour chaque question. Dans les expériences que nous allons présenter, changer cette probabilité ne modifie que marginalement les résultats et ne change pas les conclusions. Par conséquent, nous présentons les résultats que pour le choix classique de la probabilité pignistique.

Avec l'élicitation possibiliste (PCSS) sans correction, nous devons inférer malgré les incohérences, comme présenté dans la Sous-section 4.3.1, en utilisant l'une des deux stratégies suivantes : soit ignorer le conflit avec $\min_{\omega \in \emptyset} f_\omega = \max_{\omega \in \emptyset} f_\omega = 0$, ou considérer le conflit et le transformer en ignorance avec $\min_{\omega \in \emptyset} f_\omega = \min_{\omega \in \Omega} f_\omega$ et $\max_{\omega \in \emptyset} f_\omega = \max_{\omega \in \Omega} f_\omega$. En élicitation possibiliste avec le regret Minimax, nous avons $\text{PMR}(x, y, \emptyset) = 0$ ou $\text{PMR}(x, y, \emptyset) = \text{PMR}(x, y, \Omega)$. Nous comparons ensuite plusieurs méthodes d'élicitation, ainsi que des méthodes de fusion d'information :

- Pour toutes les expériences, sauf celle présentée dans la Sous-section 4.4.4, l'élicitation ensembliste classique CSS, dans laquelle le décideur est censé ne pas commettre d'erreurs. Cette méthode devrait être moins performante lorsque que le décideur donne de mauvaises réponses. Nous proposons également une stratégie similaire à la stratégie PCSS, mais en utilisant les probabilités pour gérer l'incertitude, ce qui signifie que la réponse (E_i, α_i) est traduite en $P(E_i) = \alpha_i$ et $P(E_i^c) = 1 - \alpha_i$. Le regret minimal et les autres éléments de la stratégie CSS deviennent alors de simples espérances sur la partition obtenue, qui est la même que pour le PCSS (ce qui signifie que les approches possibiliste et probabiliste ont la même complexité). Cela nous permet de comparer notre approche possibiliste avec une approche probabiliste.
- L'élicitation PCSS (avec l'une des stratégies permettant d'inférer malgré les incohérences) sans correction supplémentaire, telle que présentée dans l'Algorithme 3.
- Notre algorithme de réparation ℓ -parmi- k , présenté dans la Sous-section 4.3.2,

appliqué au résultat de l'élicitation, tel que présenté dans l'Algorithme 4.

- Différentes stratégies basées sur les MCSs, présentées dans la Sous-section 4.3.2, à nouveau appliquées au résultat de l'élicitation. Pour trois stratégies, un MCS est sélectionné par une heuristique : soit un MCS aléatoire de la plus grande taille ; soit un MCS parmi les MCSs de plus grande taille, dont les réponses maximisent la confiance moyenne $\bar{\alpha} = \frac{1}{|\text{MCS}|} \sum_{i \in \text{MCS}} \alpha_i$; ou le MCS dont les réponses maximisent la confiance moyenne. Avec la dernière stratégie MCS, nous considérons le MCS qui corrige le plus l'incohérence, qui nous sert de point de référence pour évaluer la pertinence des heuristiques.
- Un algorithme simple et naïf pour restaurer la cohérence en relaxant toutes les contraintes linéaires $Ax < b$ associées à chaque réponse du décideur. Pour ce faire, pour chaque réponse i , nous ajoutons une variable δ_i telle que $A_i x - \delta_i < b_i$ et nous essayons de trouver les valeurs de δ_i telles que la cohérence soit restaurée et que $\sum_{i=1}^n \delta_i$ soit minimale.

Chaque scénario est ensuite répété 300 fois, afin d'obtenir un échantillon de taille raisonnable.

4.4.2 Nombre de réponses incorrectes détectées (PCSS)

Nous voulons tout d'abord voir si notre méthode de fusion ℓ -parmi- k est capable de détecter le nombre de réponses incorrectes données par le décideur, étant une information utile à connaître à la fois pour l'analyste et pour le décideur. Il convient de rappeler que nous ne savons pas si une réponse donnée par le décideur est erronée ou non (et nous n'avons donc aucune information *a priori* sur le nombre de réponses erronées). En l'absence d'autres informations, l'analyste ne peut se fier qu'aux réponses précédemment données et à leur niveau de confiance : un niveau de confiance élevé signifie que nous sommes convaincus que la réponse est correcte, tandis qu'un niveau de confiance faible signifie simplement que nous ne savons pas si l'information est correcte ou non. Il convient de souligner qu'un faible niveau de confiance n'est pas interprété ici comme un signe que l'information donnée est probablement fausse, puisqu'il s'agit d'une mesure de nécessité, auquel cas un faible niveau indique que nous ne savons pas si l'information donnée est vraie ou

fausse. Cela contraste avec une interprétation probabiliste, où une probabilité de 0 implique avec certitude que l'information est fausse.

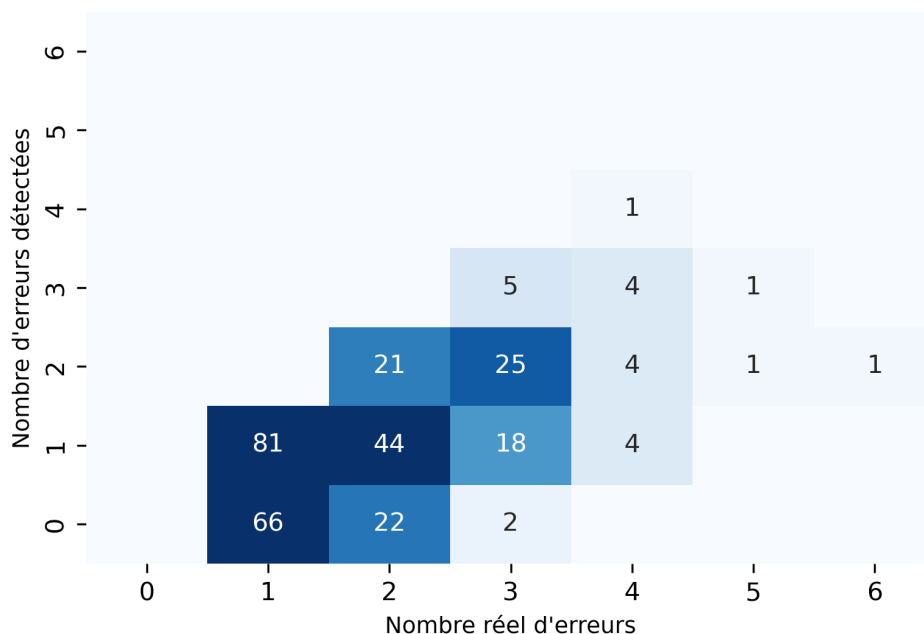


FIGURE 4.7 – Nombre d'erreurs détectées par rapport au nombre réel d'erreurs (scénario optimiste)

La Figure 4.7 indique le nombre de répétitions, compte tenu d'un nombre réel d'erreurs et en supposant que le décideur soit très confiant dans ses choix ($\alpha_i \sim \mathcal{B}(7, 2)$), pour lesquelles l'Algorithme 4 a détecté ℓ erreurs. Notons que dans ce cadre spécifique, nous ignorons le conflit, mais les résultats sont très similaires si nous transformons le conflit en ignorance. Dans ce contexte, un décideur commet généralement entre 1 et 4 erreurs, rarement plus, sur 15 réponses. L'intérêt de cette figure est de voir si notre approche permet d'estimer à chaque fois avec précision le nombre réel d'erreurs, et ce quel qu'il soit : si notre approche est un bon estimateur, alors nous verrons sur la Figure 4.7 une diagonale, tandis que si elle sous-estime de temps en temps le nombre d'erreurs, alors nous verrons plutôt un triangle inférieur. Remarquons que notre approche ne peut pas surestimer le nombre d'erreurs. Comme nous pouvons le voir, la plupart du temps, notre méthode de fusion est capable

d'estimer avec précision le nombre de fausses réponses lorsqu'il est faible. La différence entre le nombre retourné par l'Algorithme 4 et le nombre réel d'erreurs s'explique facilement : une mauvaise réponse ne contredit pas nécessairement toutes les bonnes réponses, ce qui signifie qu'une mauvaise réponse ne crée pas automatiquement des incohérences (pensons par exemple au cas où la première réponse est erronée).

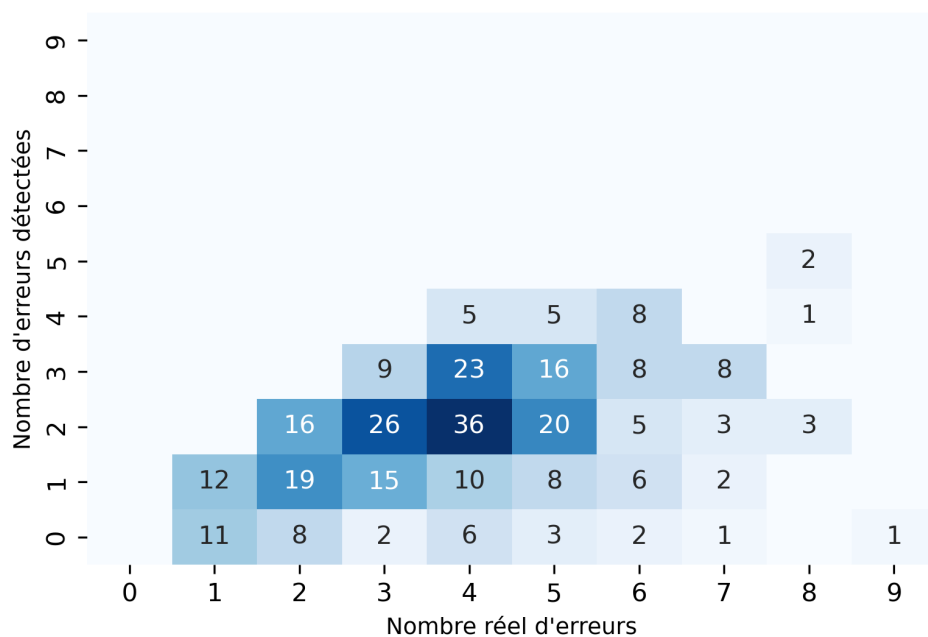


FIGURE 4.8 – Nombre d'erreurs détectées par rapport au nombre réel d'erreurs (scénario uniforme)

La Figure 4.8 montre le nombre d'erreurs détectées comme sur la Figure 4.7, mais dans le cas où le décideur n'est pas nécessairement sûr de ses choix ($\alpha_i \sim \mathcal{U}(0, 1)$). Cette fois, le nombre de réponses incorrectes est beaucoup plus élevé, avec 1 à 7 choix incorrects parmi 15 réponses la plupart du temps. Notre méthode de fusion est moins efficace pour détecter le nombre réel de réponses incorrectes, mais cela est conforme aux résultats précédents : étant donné que le nombre de choix incorrects est beaucoup plus élevé, notre méthode est moins capable de détecter toutes les réponses incorrectes. Plus le nombre d'erreurs est élevé, plus la probabilité que

plusieurs erreurs soient cohérentes entre elles augmente, ce qui rend leur détection encore plus difficile.

4.4.3 Comparaison des méthodes pour gérer l'incertitude (PCSS)

Dans ce chapitre, nous nous sommes concentrés sur le problème de recommandation d'une alternative à l'aide d'un modèle numérique. L'utilisation d'un modèle numérique est avantageux, car nous disposons d'une mesure numérique directe des performances des différentes approches. Pour ce faire, nous calculons sur les différentes répétitions la moyenne du regret réel $R_{\omega^*}(x^*, x_i^{\text{opt}})$ entre l'alternative x^* recommandée par une méthode et x_i^{opt} la meilleure alternative étant donné le vrai modèle ω^* d'une expérience. Nous désignons cette moyenne par :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n R_{\omega^*}(x^*, x_i^{\text{opt}}). \quad (4.6)$$

Notons que nous n'avons conservé que les répétitions pour lesquelles une incohérence a été détectée (sinon toutes les méthodes de traitement des incohérences coïncident), soit 210 répétitions pour le scénario optimiste et 266 pour le scénario intermédiaire. Nous avons également déterminé un intervalle de confiance $\text{IC} = [\bar{x} \pm t_{n-1, 1-\frac{\alpha}{2}} \frac{S^*}{\sqrt{n}}]$ sur \bar{x} , S^* étant l'écart-type corrigé des regrets réels, n étant le nombre de répétitions conservées, et $\alpha = 0.05$. Notons que nous avons vérifié la normalité à chaque fois avec des tests de Shapiro-Wilk.

Afin de comparer les différentes méthodes et de déterminer si les différences de regret réel sont significatives, nous avons effectué des tests statistiques de différence sur des échantillons appariés avec un niveau de signification de 5%. Comme nous ne voulions pas supposer que les différences soient normalement distribuées (ce qui est confirmé dans la plupart des exemples par des tests de Shapiro-Wilk), nous utilisons des tests non paramétriques, ceux des rangs signés de Wilcoxon, pour déterminer si les différences sont négligeables. Nous fournissons la p -value p associée aux tests, et si $p < 0.05$, nous pouvons supposer que les différences sont statistiquement significatives, compte tenu de la taille de notre échantillon.

Les Figures 4.9 et 4.10 présentent des intervalles de confiance sur le regret réel moyen. Chaque figure est divisée en trois parties :

- La partie supérieure contient des stratégies de correction des incohérences qui n'utilisent pas les informations possibilistes (les niveaux de confiance) : l'algorithme naïf qui relâche les contraintes, un MCS aléatoire ayant une taille maximale et un MCS théorique qui minimise le regret réel. Les deux premières approches devraient être surpassées par les approches utilisant les informations possibilistes, tandis que la dernière approche nous aide à estimer la qualité de nos heuristiques pour trouver des MCSs ;
- La partie centrale contient nos différents algorithmes de fusion d'information (ℓ -parmi- k et les MCSs obtenus par heuristique qui utilisent les degrés de confiance) qui devraient être comparés à l'algorithme PCSS pour déterminer si les stratégies de fusion donnent de meilleures recommandations que simplement spécifier la valeur de $\min_{\omega \in \emptyset} f(\omega)$.
- La partie inférieure contient les algorithmes CSS, probabiliste et PCSS, afin de déterminer si le simple traitement de l'incohérence par notre algorithme PCSS est utile ou non, et de le comparer aux approches standard.

Une première question est de savoir si la manière dont nous traitons l'incohérence dans le choix incrémental des questions (c'est-à-dire, en ignorant le conflit ou en considérant le conflit comme de l'ignorance pendant l'élicitation possibiliste) dans la Sous-section 4.3.1 a un impact sur le résultat final et la recommandation. Compte tenu du regret réel moyen avec l'élicitation classique ($\bar{x} = 0.0504$ dans le scénario fort, 0.125 dans le scénario uniforme), nous pouvons voir qu'ignorer l'incohérence est une assez bonne stratégie lors du processus d'élicitation (réduction du regret réel dans le scénario fort : 0.0423 , $p = 1.59e - 16$. Scénario uniforme : 0.0880 , $p = 5.79e - 24$), car elle présente un gain systématique par rapport à l'approche basée sur les ensembles, qui est significatif dans nos deux scénarios. *A contrario*, nous constatons que la transformation de l'incohérence en ignorance lors du processus d'élicitation dégrade généralement les résultats en termes de recommandation (réduction du regret réel dans le scénario fort : -0.146 , $p = 4.647e - 32$. Scénario uniforme : -0.0978 , $p = 4.449e - 23$). Ainsi, bien que cette option soit assez courante dans la littérature, nos résultats indiquent clairement que confondre l'incohérence et l'ignorance est ici un choix très risqué et préjudiciable, du moins lorsqu'il s'agit de choisir les questions à présenter au décideur.

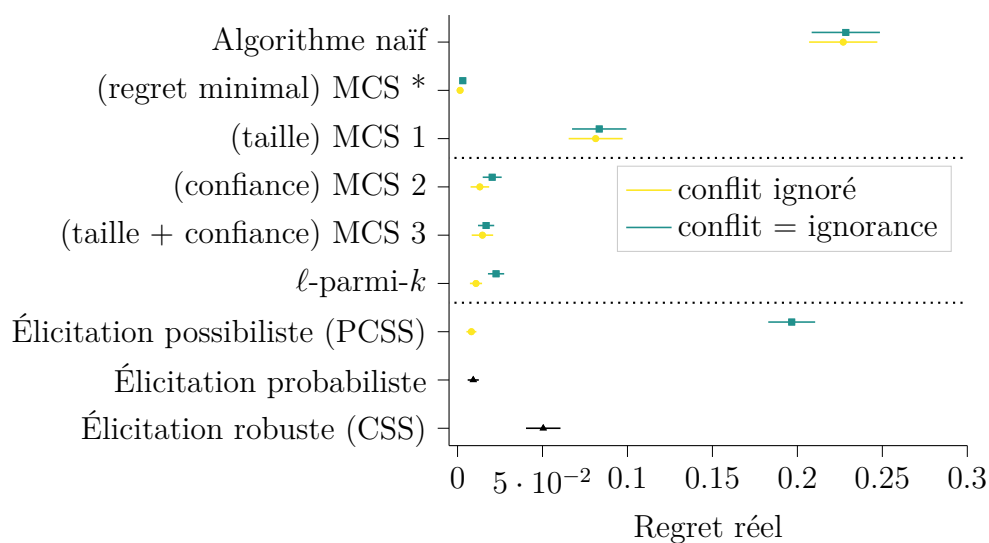


FIGURE 4.9 – Intervalles de confiance sur le regret réel dans le scénario optimiste. Plus le regret réel est faible, meilleure est la recommandation finale.

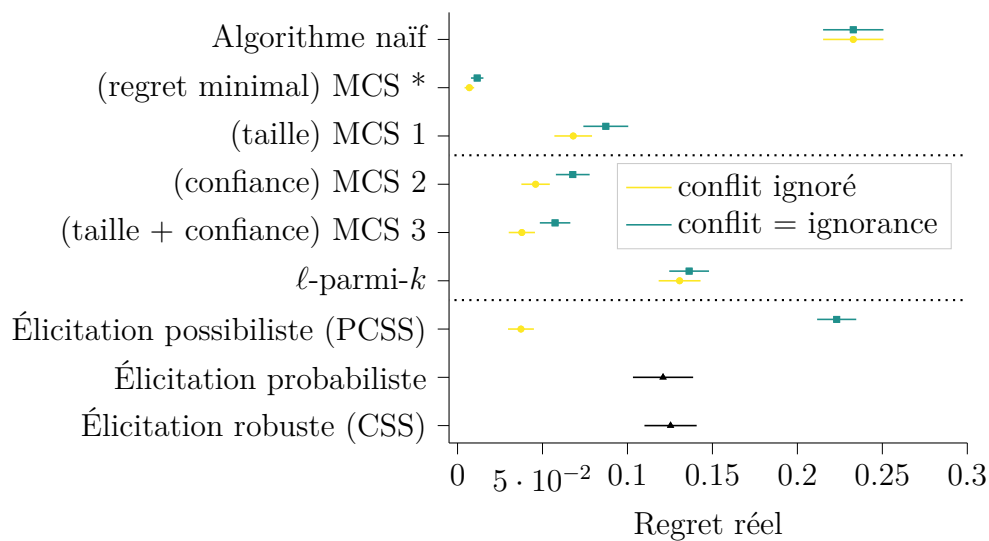


FIGURE 4.10 – Intervalles de confiance sur le regret réel dans le scénario intermédiaire. Plus le regret réel est faible, plus la recommandation finale est bonne.

En complément de cette première question, nous voulons savoir si l'approche possibiliste est meilleure que l'approche probabiliste pour traiter l'incohérence. Rappelons que nous avons déjà montré dans l'Exemple 23 qu'une approche basée

sur les probabilités ne peut pas détecter un conflit de manière fiable, puisqu'un conflit peut apparaître même sans incohérence. Par conséquent, même si une approche basée sur la théorie des probabilités donnait de meilleures recommandations, notre algorithme PCSS aurait toujours un grand intérêt. Dans le scénario dans lequel le décideur commet peu d'erreurs, et comme le montre la Figure 4.9, l'élicitation probabiliste et le PCSS donnent tous deux des résultats similaires, sans différence statistiquement significative ($\bar{x}_{\text{prob}} = 0,00924$, $\bar{x}_{\text{PCSS}} = 0,00815$, $p = 0,785$). Néanmoins, dans un scénario dans lequel le décideur est moins confiant et commet plus d'erreurs, comme présenté sur la Figure 4.10, une approche probabiliste améliore à peine la qualité de la recommandation en moyenne, et nous obtenons des résultats médiocres par rapport à l'approche possibiliste lorsqu'elle ignore les conflits ($\bar{x}_{\text{prob}} = 0,121$, $\bar{x}_{\text{PCSS}} = 0,0373$, $p = 6,94e - 17$). Cela montre qu'en plus de fournir des informations utiles sur l'incohérence, une approche possibiliste peut être compétitive, en particulier lorsque le nombre d'erreurs augmente.

Une deuxième question est de savoir s'il existe une différence, en termes de qualité de recommandation, entre le simple traitement de l'incohérence par la redéfinition de $\text{PMR}(x, y, \emptyset)$, et l'utilisation de stratégies de fusion plus élaborées. Avant d'entrer dans les détails, nous pouvons voir quelques premiers éléments :

- La restauration naïve de la cohérence donne toujours les pires recommandations, et toutes les méthodes sont meilleures quelle que soit la stratégie de redéfinition.
- Toutes les méthodes de fusion donnent de meilleures recommandations que l'élicitation classique, sauf en choisissant un MCS aléatoire de la plus grande taille dans le scénario optimiste, et la méthode de fusion ℓ -parmi- k dans le scénario intermédiaire.
- Toutes les méthodes de fusion donnent de meilleures recommandations quand le conflit n'est pas considéré comme de l'ignorance. Compte tenu de ce résultat, nous nous concentrerons sur la simple ignorance des conflits dans le reste de cette sous-section.

En plus des graphiques précédents, la Table 4.1 compare les différentes stratégies de fusion en considérant comme méthode de référence $\text{PMR}(x, y, \emptyset) = 0$. Pour les MCSs, nous ne donnons ici les résultats que pour celui de taille maximale dont

	Optimiste	Uniforme
ℓ -parmi- k	-0.00270	-0.0934
p -value	0.0430	4.21e-31
MCS heuristique	-0.00499	-0.008312
p -value	0.122	0.00483
MCS minimisant \bar{x}	0.00666	0.0304
p -value	1.78e-08	4.76e-23

TABLE 4.1 – Différence de regret réel entre l'ignorance du conflit et les méthodes de fusion d'information, une valeur positive signifiant que les méthodes de fusion ont un regret réel plus faible en moyenne

les réponses minimisent la confiance moyenne, et pour celui qui minimise le regret réel. Étant donné une stratégie de fusion et un scénario, nous fournissons deux nombres. Le premier nombre est la différence en moyenne de regret réel entre la stratégie de fusion et la simple inférence $\text{PMR}(x, y, \emptyset) = 0$, la différence étant $\frac{1}{n} \sum_{i=1}^n R_{\omega_i}(x_i^{\text{fusion}}, x_i^{\text{infer}})$. Un nombre négatif signifie qu'en moyenne la stratégie de fusion donne de moins bonnes recommandations, tandis qu'un nombre positif signifie qu'en moyenne la stratégie de fusion donne de meilleures recommandations. Le deuxième nombre est la p -value, qui permet de déterminer si la différence en qualité des recommandations est statistiquement significative ou non.

Nous remarquons que les résultats sont assez mitigés : la méthode de fusion ℓ -parmi- k et le MCS déterminé par une heuristique ne réduisent pas en moyenne le regret réel, et pire encore, ils augmentent légèrement en moyenne le regret réel. Ce comportement reste néanmoins moyen et, dans certains cas, ces méthodes réduisent le regret réel. Notons cependant que la stratégie de fusion ℓ -parmi- k est peu performante dans le scénario uniforme. Nous voyons également que le MCS minimisant le regret réel réduit le regret réel en moyenne, quel que soit le scénario, ce qui est intéressant.

4.4.4 Influence de la règle de décision et élicitation par batch

Afin d'étayer les résultats obtenus avec l'algorithme PCSS, nous avons effectué un test simple dans lequel les questions n'ont pas été choisies selon l'algorithme PCSS, mais de manière totalement aléatoire. De plus, nous voulons connaître l'influence de la règle de décision sur les résultats. Pour cela, nous considérons trois règles de décision usuelles déjà présentées dans la Sous-Section 4.2 : le Maximax (MM), le Maximin (Mm) et le regret Minimax (mMR).

Les Figures 4.11 et 4.12 montrent les performances obtenues par les différentes approches pour les scénarios optimiste et uniforme, respectivement. Elles sont divisées en trois parties, comme les figures 4.9 et 4.10, la seule différence se situant dans la partie inférieure : elle ne contient que l'algorithme PCSS avec les deux stratégies de traitement de l'incohérence (ignorer le conflit ou transformer le conflit en ignorance).

Nous pouvons observer les résultats suivants :

- Si nous nous concentrons sur la règle de décision du regret Minimax (mMR), nous pouvons constater que même avec des questions aléatoires, les résultats observés dans la sous-section 4.4.3 se maintiennent.
- La règle de décision du Maximax (MM) fournit en général des recommandations très similaires au regret Minimax, bien qu'avec un regret réel légèrement plus élevé en moyenne.
- La règle de décision du Maximin (Mm) fournit les pires recommandations quelle que soit la méthode (sauf pour le MCS aléatoire de plus grande taille). En outre, si nous utilisons uniquement une redéfinition, la règle de décision du Maximin donne de très mauvaises recommandations, quelle que soit la redéfinition. Cependant, en utilisant des méthodes de fusion pour réparer l'incohérence, les recommandations deviennent similaires à celles des autres règles de décision.

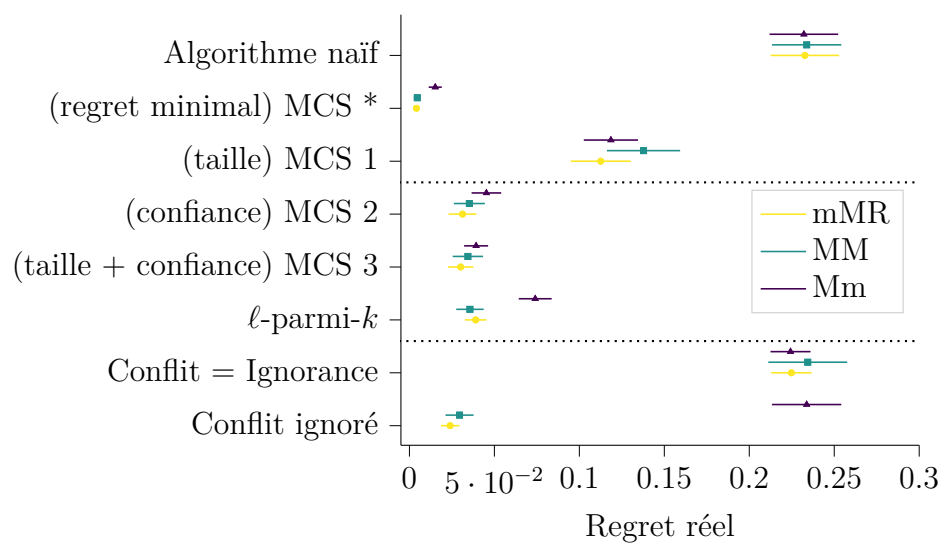


FIGURE 4.11 – Intervalles de confiance sur le regret réel pour chaque méthode dans le scénario optimiste avec une élicitation par batch et trois règles de décision

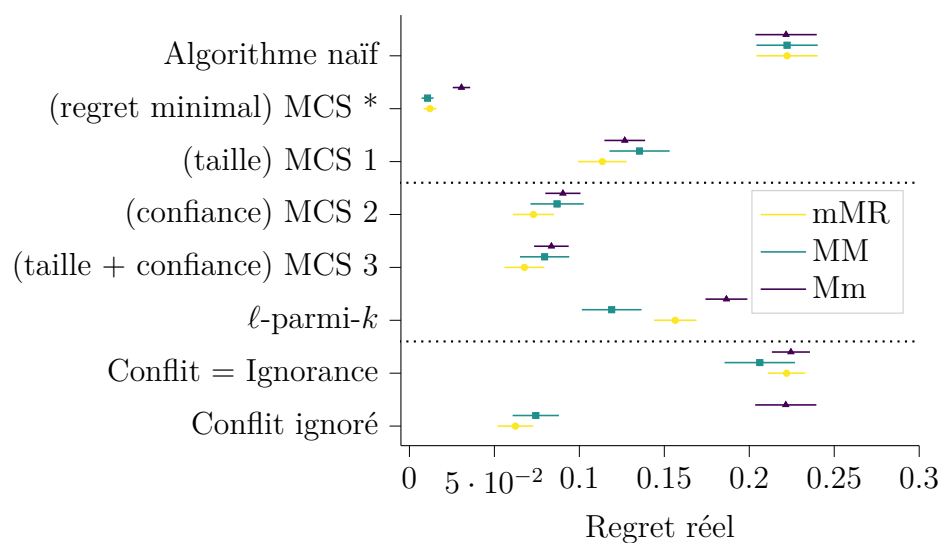


FIGURE 4.12 – Intervalles de confiance sur le regret réel pour chaque méthode dans le scénario intermédiaire avec une élicitation par batch et trois règles de décision

4.4.5 Discussion des résultats

Sur la base des résultats précédents, nous pouvons tenter de tirer quelques premières conclusions générales :

- Il semble que les méthodes de fusion puissent améliorer les recommandations par rapport aux stratégies d'inférence tolérantes aux incohérences lorsque ces dernières fournissent des recommandations plutôt mauvaises au départ. Dans les autres cas, l'utilisation de stratégies de fusion d'information n'apporte au mieux aucune amélioration supplémentaire (à une exception près), ce qui semble normal si nous partons d'une situation déjà bonne.
- Les avantages numériques bruts de l'utilisation de stratégies de fusion d'information semblent être limités pour la plupart des méthodes de fusion, étant donné que l'augmentation de la puissance de calcul requise ne semble pas toujours être compensée par une amélioration significative de la qualité des recommandations. Il convient toutefois de rappeler que ces règles de fusion présentent également un intérêt analytique, en ce sens qu'elles peuvent nous fournir des informations utiles sur le nombre potentiel d'erreurs, comme le montre la Figure 4.7. Toutefois, l'intérêt réel de ces propriétés analytiques est difficile à évaluer dans le cadre d'une expérience purement synthétique, et nous aimerions poursuivre nos recherches en pratique.
- Le MCS qui minimise le regret réel ne peut pas toujours être trouvé par une simple heuristique. Cependant, en triant les MCSs en fonction de leur taille et de leur confiance moyenne, un MCS minimisant le regret réel est généralement trouvé parmi les premiers (rang moyen dans le scénario optimiste : 1,5, le MCS spécifique classé entre le 1er et le 2e rang pour 90 % des répétitions ; rang moyen dans le scénario intermédiaire : 3,02, le MCS spécifique classé entre le 1er et le 4e rang pour 80 % des répétitions). Il est donc utile de présenter ces ensembles au décideur, afin qu'il puisse les examiner de plus près et éventuellement déterminer les bonnes réponses.
- L'utilisation du Maximax au lieu de regret Minimax n'a qu'un léger impact sur la qualité des recommandations. Ces résultats semblent en adéquation avec les observations du Chapitre 3, où après un certain nombre de questions, le regret

Minimax et le Maximax convergent tous les deux vers une recommandation conforme aux préférences du décideur. La règle de décision du Maximin a un impact négatif sur la qualité des recommandations, en particulier lorsque aucune correction n'est appliquée. Cependant, nos méthodes de fusion peuvent restaurer la qualité des recommandations de la règle de décision du Maximin à des niveaux acceptables.

Il convient de noter que la généralisation des remarques ci-dessus, si elles semblent intuitivement raisonnables, doit être vérifiée par d'autres expériences synthétiques ou réelles. En effet, nous avons considéré une méthode d'élicitation incrémentale spécifique qui est connue pour fournir de bonnes performances en général, et nous avons considéré des questions très simples dont les réponses fournissent des informations très limitées sur le modèle : cela signifie d'une part que les actions correctives auront un impact limité (puisque le point de départ est susceptible de montrer déjà de bonnes performances), et d'autre part que de nombreuses informations préférentielles peuvent être cohérentes les unes avec les autres, même en cas de mauvaises réponses.

4.5 Conclusion et perspectives

Nous avons vu dans ce chapitre comment manipuler l'incohérence provenant du décideur une fois qu'elle a été détectée par une approche d'élicitation possibiliste. Nous avons vu deux manières de procéder : soit en inférant malgré les incohérences, sans modifier les informations préférentielles, soit en changeant la manière dont les informations sont fusionnées, prenant en compte que certaines informations préférentielles sont erronées. Avec les outils de fusion, il est également possible d'obtenir une analyse intéressante de la situation. En particulier, le fait que la plupart des règles de fusion puissent être directement associées à des opérateurs logiques ou à des hypothèses claires sur les sources apporte une lisibilité supplémentaire qui devrait être appréciée par un décideur

Pour illustrer ces aspects, nous nous sommes concentrés sur l'apprentissage incrémental des préférences par le biais du regret Minimax, en choisissant une moyenne pondérée et quelques règles de fusion sélectionnées dans nos expériences

empiriques. Ces expériences confirment l'intérêt potentiel de nos approches et nous ont permis d'identifier certaines de leurs limites, c'est-à-dire en identifiant certaines situations où elles ont montré de faibles performances. Nous avons également obtenu des résultats similaires avec différentes règles de décision et des questions aléatoires.

Points importants vus dans ce chapitre :

- Possibilité d'inférer malgré l'incohérence, ou de réparer l'incohérence par des outils de fusion d'information.
- Les outils de fusion d'information apportent des informations supplémentaires intéressantes, notamment sur les informations possiblement erronées.
- Ignorer l'incohérence ou réparer l'incohérence permet dans plusieurs cas d'améliorer la qualité des recommandations par rapport à une approche ensembliste CSS.

Perspectives :

Application à d'autres modèles : comme nous l'avons déjà mentionné, plusieurs aspects de ce chapitre, notamment ceux présentés dans la Section 4.3, ainsi que le cadre général de notre approche possibiliste présentée dans la Section 2.2, peuvent être étendus à des situations différentes de celle de nos expériences synthétiques. Nous avons déjà évoqué une extension aux modèles lexicographiques dans la Section 2.5, mais nous pouvons ajouter :

- Appliquer l'approche d'élicitation basée sur le regret et les techniques de traitement des incohérences à d'autres modèles ou situations. Par exemple, l'approche basée sur le regret Minimax a également été appliquée à des problèmes combinatoires multiobjectifs (Benabbou et Perny, 2018), à des modèles numériques plus complexes tels que les intégrales de Choquet (Benabbou *et al.*, 2017; Martin et Perny, 2021) ou même à des modèles plus qualitatifs tels que les intégrales de Sugeno (Rico et Viappiani, 2020).
- Extension du cadre proposé pour prendre en compte d'autres types de tâches ou d'inférences que la recommandation d'une seule alternative. Les tâches

classiques consistent à classer toutes les alternatives, éventuellement partiellement, et à les classer dans des catégories ordonnées (voir par exemple (Bouyssou *et al.*, 2006, Ch. 7)). Un point de départ simple consisterait à envisager des stratégies d'élicitation selon le regret Minimax pour de telles tâches (Benabbou *et al.*, 2017; Özpeynirci *et al.*, 2022) et d'y rattacher simplement notre extension possibiliste. Il faudrait toutefois préciser comment l'inférence correspondante (classement ou attribution d'une catégorie) peut être effectuée compte tenu de la distribution finale de possibilité sur les modèles.

Changement de modèle : nous avons évoqué dans la Section 4.3 la réparation d'erreurs de modèle par un changement de modèle. Il serait intéressant d'évaluer à quel point un changement de modèle permet de réparer l'incohérence provenant d'une erreur de modèle. Nous présentons ici une première série d'expériences préliminaires⁵ qui répond à certaines questions tout en ayant des points à approfondir, en se basant sur le protocole expérimental de l'expérience de la Sous-section 2.4.4. Ainsi, nous voulons recommander à un décideur l'alternative qu'il devrait préférer parmi 35 alternatives avec 4 critères qui sont Pareto non dominées. Au niveau de la fonction d'agrégation des critères, nous avons deux scénarios :

- Les préférences du décideur sont modélisées par une intégrale de Choquet dont la capacité est générée aléatoirement, mais à tort l'analyste a supposé que le modèle de préférence est une somme pondérée (WS) ou une OWA.
- Les préférences du décideur sont modélisées par une OWA, mais à tort l'analyste a supposé que le modèle de préférence est une somme pondérée.

Au niveau du procédé d'élicitation, nous avons de nouveau deux scénarios :

- Élicitation PCSS (incrémentale).
- Élicitation possibiliste par batch (non incrémentale), questions aléatoires.

Nous arrêtons l'élicitation dès qu'une incohérence est détectée, ou que le décideur a répondu à 15 questions. Le décideur est supposé toujours répondre correctement selon son vrai modèle de préférence et être confiant de ses choix ($\alpha_i \sim \mathcal{B}(7, 2)$). Enfin, nous répétons 200 fois le processus d'élicitation.

5. https://github.com/LoicAdam/Model_Elicitation

Sur la Figure 4.13, en partant sur une élicitation PCSS et un vrai modèle qui est une intégrale de Choquet, nous pouvons voir que seulement une partie des mauvais modèles a été détectée (environ 34 à 38 % des mauvais modèles au bout de 15 questions sur 200 répétitions). Une première interrogation est de savoir pourquoi ces mauvais modèles ne sont pas détectés. Nous avons effectué le même type d'expérience, mais avec une élicitation par batch, pour voir si une stratégie de questionnement aléatoire est meilleure que l'élicitation PCSS pour remettre en cause le modèle supposé. Les résultats pour l'élicitation par batch sont présentés sur la Figure 4.14. Nous pouvons voir que le taux de détection est beaucoup plus élevé et dès les premières questions. Cette observation semble justifier la recherche d'un algorithme d'élicitation possibiliste particulier, qui cherche à remettre en question le modèle avec des questions spécifiques, pour détecter les erreurs au plus tôt⁶.

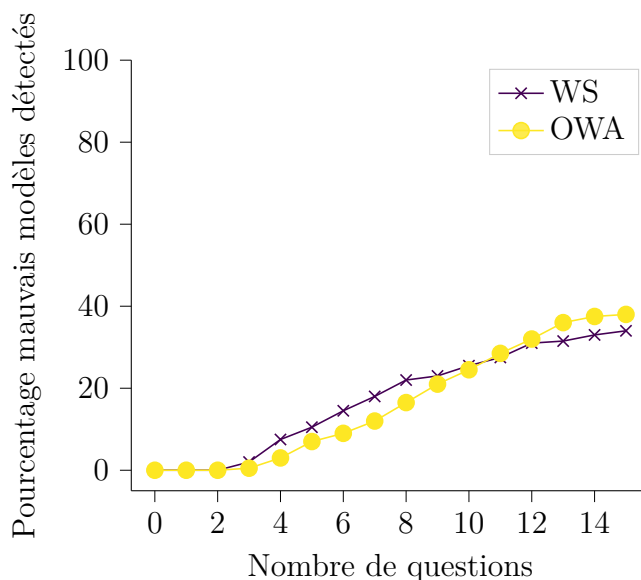


FIGURE 4.13 – Détection d'une erreur de modèle avec une élicitation incrémentale, le vrai modèle étant une intégrale de Choquet

6. Remarquons néanmoins que l'élicitation par batch donne des recommandations qui ont un regret réel plus fort dans nos expériences, indiquant que l'élicitation PCSS reste meilleure pour recommander de bonnes alternatives.

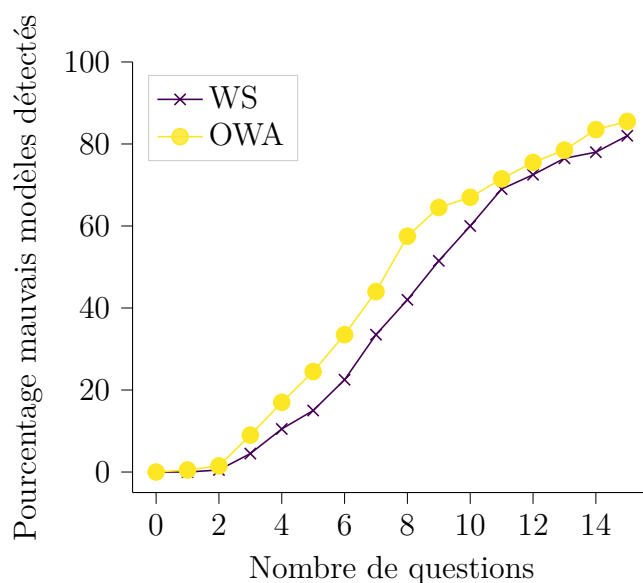


FIGURE 4.14 – Détection d’une erreur de modèle avec une élicitation par batch, le vrai modèle étant une intégrale de Choquet

Pour mesurer l’intérêt d’un changement de modèle, nous avons remplacé le modèle erroné (WS ou OWA) par une intégrale de Choquet pour les répétitions dans lesquelles de l’incohérence a été détectée. Nous avons ensuite mesuré le regret réel de la recommandation finale des mauvais modèles détectés, des mauvais modèles détectés et corrigés par un changement de modèle (Choquet) et des mauvais modèles non détectés. Le calcul du regret réel est présenté au début de la Section 4.4. En supposant le modèle erroné être une WS, la moyenne du regret réel et son écart type corrigé sont présentés sur la Table 4.2 (les statistiques pour une OWA étant très similaires). Nous pouvons tout d’abord voir que les mauvais modèles non détectés ont en général un regret réel très faible, indiquant que le mauvais modèle permet dans ces cas précis de capturer correctement les préférences du décideur⁷. Une deuxième observation est que le regret réel diminue que faiblement en moyenne en changeant le mauvais modèle par le bon modèle. Pire encore, pour certaines répétitions, le regret réel augmente. Une question est donc de savoir pourquoi le

7. Dès lors, il serait intéressant de voir dans de futurs travaux si les intégrales de Choquet correspondantes sont semblables aux mauvais modèles (l’intégrale de Choquet étant une généralisation de la WS et de l’OWA), ou si ce regret faible est expliqué autrement.

regret réel ne diminue que faiblement en moyenne, alors qu'on s'attendrait à ce qu'il diminue très fortement, ayant le vrai modèle de préférence du décideur.

Méthode	\bar{x}	σ_x
Mauvais modèle détecté	.0948	.0817
Mauvais modèle détecté et corrigé	.0678	.0741
Mauvais modèle non détecté	.00866	.0217

TABLE 4.2 – Statistiques sur le regret réel pour des erreurs de modèle (WS supposé, Choquet réel) lors d'une élicitation PCSS

Pour répondre en partie à cette dernière question, nous nous intéressons au mEMR dans le cas où un mauvais modèle a été détecté, étant un indicateur de la qualité de la recommandation courante, sous réserve qu'il n'y ait aucune erreur. Nous calculons le mEMR selon le modèle non corrigé (ici WS) et le modèle corrigé (Choquet). Les résultats sont présentés dans la Table 4.3. Nous pouvons voir une différence flagrante au niveau du mEMR : quand un mauvais modèle a été détecté, le mEMR est relativement faible (sans être zéro), indiquant que nous nous approchons d'une recommandation optimale selon le modèle de préférence supposé et les informations préférentielles disponibles, alors que le mEMR est élevé quand le mauvais modèle a été corrigé.

Méthode	\bar{x}	σ_x
Mauvais modèle détecté	.0815	.0778
Mauvais modèle détecté et corrigé	.310	.0770

TABLE 4.3 – Statistiques le mEMR pour des erreurs de modèle (WS supposé, Choquet réel) lors d'une élicitation PCSS

Pour essayer de comprendre pourquoi le mEMR est élevé, nous avons fait le même type d'expérience, mais avec un vrai modèle qui est une OWA et le modèle supposé une WS. En effet, l'intégrale de Choquet n'a pas le même nombre de paramètres ($2^4 = 16$) qu'une OWA ou qu'une WS (4). Nous voulons ainsi voir si la

diminution du mEMR n'est pas principalement dû à la dimensionnalité plus grande du problème. Nous présentons ici les résultats pour une élicitation incrémentale, les résultats pour une élicitation par batch ne changeant pas les conclusions (seulement un regret réel un peu plus élevé en moyenne). Sur les Tables 4.4 et 4.5 sont indiqués des statistiques sur le regret réel et le mEMR avec ces nouveaux modèles. Nous voyons que cette fois-ci, corriger un mauvais modèle va fortement diminuer le regret réel, et même diminuer légèrement le mEMR.

Méthode	\bar{x}	σ_x
Mauvais modèle détecté	.103	.0841
Mauvais modèle détecté et corrigé	.00649	.0140
Mauvais modèle non détecté	.0238	.0331

TABLE 4.4 – Statistiques sur le regret réel pour des erreurs de modèle (WS supposé, OWA réel) lors d'une élicitation PCSS

Méthode	\bar{x}	σ_x
Mauvais modèle détecté	.0601	.0648
Mauvais modèle détecté et corrigé	.0306	.0272

TABLE 4.5 – Statistiques le mEMR pour des erreurs de modèle (WS supposé, OWA réel) lors d'une élicitation PCSS

Nous pouvons donc supposer que l'écart observé avec l'intégrale de Choquet est principalement dû à l'augmentation du nombre de paramètres à estimer : le nombre et la nature des informations préférentielles ne permettant pas d'estimer au mieux ces paramètres, une stratégie de questionnement optimale pour une WS n'étant pas nécessairement optimale pour une intégrale de Choquet. Ce dernier point sur la stratégie de questionnement optimale pour un modèle, mais pas un autre, serait intéressant de développer d'un point de vue théorique. De plus, nous n'avons pas de mesure sur la distance entre l'OWA et la WS : il serait intéressant de déterminer si la distance entre les modèles influe ou non la qualité des recommandations finales. Ce même raisonnement pourrait être appliqué avec les intégrales de Choquet.

Distinction du type d'erreur : une dernière perspective intéressante serait de distinguer les erreurs de modèle de celles du décideur, comme nous l'avions évoqué dans la Section 4.3, notamment pour traiter au mieux l'erreur. Il serait intéressant d'avoir un outil automatique qui permet de détecter le type d'erreur, avec le moins d'information extérieure possible. Ainsi, nous voudrions éviter par exemple de demander au décideur s'il a commis des erreurs. Nous voudrions aussi autant que possible éviter d'utiliser les niveaux de confiance, car s'il pourrait sembler logique que le décideur soit confiant de ses réponses en cas d'erreur de modèle et moins confiant en cas d'erreur de sa part, cela devrait être confirmé par des expériences dans le monde réel, sans oublier que les niveaux de confiance pourraient être donnés par l'analyste pour permettre de détecter de l'incohérence.

Pour cela, nous avons exploré une piste liée à un de nos travaux sur les MCSs présenté dans l'Annexe A. Ainsi, ce travail semble indiquer que les MCSs peuvent distinguer les deux types d'erreur :

- En cas d'erreur de modèle, le nombre de MCSs serait faible, avec peu, voire aucun élément en commun entre les MCSs, comme la Figure 4.15 l'illustre. Ainsi, nous voyons que les informations préférentielles ne sont pas toutes cohérentes entre elles, mais nous pouvons voir deux MCSs avec aucun élément en commun. Dans cette illustration particulière, nous avons un modèle linéaire par morceau, dépendant ici de l'ordonnancement des critères. Le modèle de base est donc vrai que localement, et non globalement. L'OWA est par exemple une moyenne pondérée par morceau, ou nous avons un ordonnancement différent des poids ω_i en fonction de l'ordre des performances des critères.

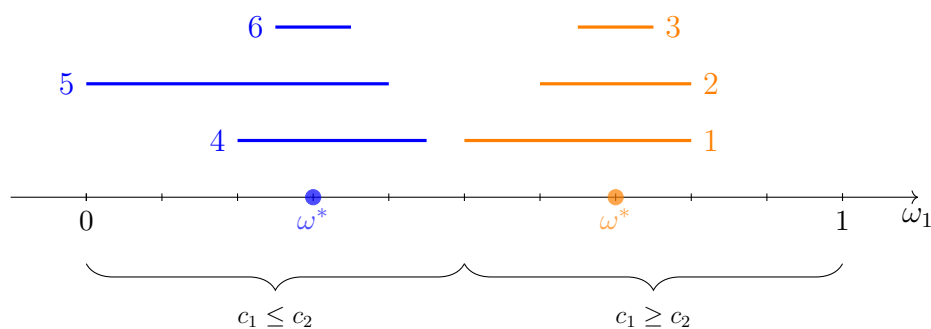


FIGURE 4.15 – Illustration d'un cas idéal d'erreur de modèle

- En cas d'erreurs de la part du décideur, le nombre de MCSs serait élevé, avec possiblement plusieurs éléments en commun entre les MCSs, les bonnes réponses étant cohérentes entre elles, mais pas les mauvaises, comme la Figure 4.16 l'illustre. Ici, nous voyons trois bonnes réponses (1, 2 et 3) et trois mauvaises réponses (4, 5 et 6), ainsi que quatre MCSs : $\mathcal{C} = \{\{2, 3, 4\}, \{1, 2, 3, 5\}, \{1, 2, 6\}, \{1, 3, 6\}\}$. Ainsi, les réponses 1, 2 et 3 sont très souvent présentes, tandis que 4, 5 et 6 apparaissent plus sporadiquement et jamais ensemble.

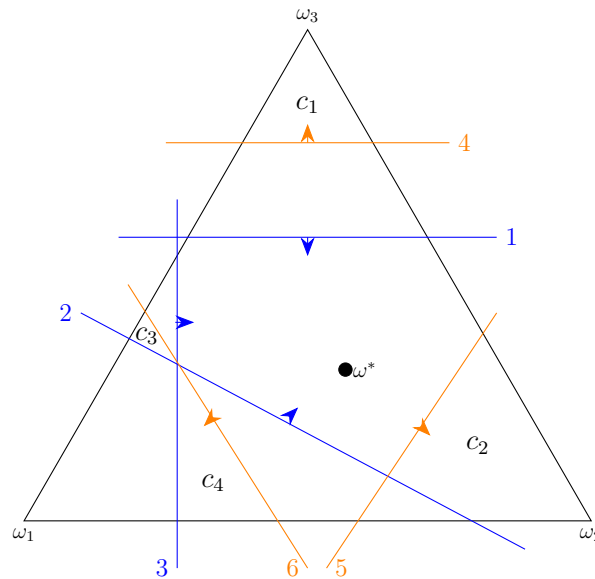


FIGURE 4.16 – Illustration d'un cas idéal d'erreurs du décideur

Nous avons tenté d'utiliser ces observations dans notre contexte d'élicitation de préférences, en étudiant les différents MCSs et leurs propriétés (nombre de MCSs, nombre d'éléments, occurrence des éléments, confiance moyenne, co-occurrence...) pour voir s'il était possible de distinguer les deux types d'erreur avec un algorithme de machine learning. Néanmoins, sur de premiers résultats très préliminaires, nous faisons à peine mieux que de décider aléatoirement (environ 55 % de prédictions correctes). Une explication possible serait que les informations préférentielles soient trop simples (contrainte linéaire sur l'ensemble des valeurs possibles de ω) pour être exploitables. Sans meilleures statistiques sur les MCSs (notamment sur la connexité des informations préférentielles, pour voir si certaines apparaissent plus souvent ensemble, ou si les MCSs sont relativement disjoints) ou à exploiter des

MCSs différents (par exemple, regarder les MCSs des informations préférentielles corrigées, les mauvaises réponses initiales seraient alors cohérentes entre elles, et les bonnes réponses initiales pas nécessairement cohérentes entre elles), une perspective serait de distinguer le type d'erreur avec d'autres stratégies de questionnement qui donneraient des informations beaucoup plus riches, en demandant par exemple au décideur de comparer plusieurs alternatives à chaque fois ($p - 1$ contraintes linéaires pour p alternatives comparées), voir de les classer ($\binom{p}{2}$ contraintes linéaires). Cela risque néanmoins d'augmenter fortement la charge cognitive du décideur, et nous rappelons qu'un des intérêts de l'approche incrémentale à base d'un regret Minimax est de ne pas trop surcharger cognitivement le décideur.

Contributions :

- Adam, Loïc et Destercke, Sébastien. Handling inconsistency in (numerical) preferences using possibility theory. Dans : *Information Fusion*. Elsevier, 2023.
- Adam, Loïc et Destercke, Sébastien. Handling inconsistency in (numerical) preferences using possibility theory. Dans : *17e European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU 2023, 19-22 septembre 2023, Arras, France*. Springer Cham, 2023.
- Adam, Loïc et Destercke, Sébastien. Traiter les Incohérences dans les Préférences (numériques) à l'aide de la Théorie des Possibilités. Dans : *32e Rencontres Francophones sur la Logique Floue et ses Applications, LFA 2023, 9-10 novembre 2023, Bourges, France*. Cépadués Editions, 2023.

NB : le numéro des pages de ces travaux sont inconnus lors de la publication de ce manuscrit.

Chapitre 5

Conclusion

Un premier objectif de cette thèse était de proposer une approche d'élicitation efficace pour manipuler des préférences incertaines tout en ayant des garanties fortes au niveau de la convergence et de la qualité de la recommandation finale. Un deuxième objectif était de proposer des stratégies pour inférer malgré l'incohérence, mais aussi des outils pour rétablir la cohérence. Nous revenons ensemble dans ce chapitre de conclusion sur les apports de cette thèse dans les deux problématiques initiales. Nous évoquons aussi quelques perspectives qui sont toutes détaillées dans les chapitres correspondants.

Approche possibiliste pour détecter l'incohérence

Dans le Chapitre 1, nous avons montré comment modéliser des préférences, et justifié notre choix pour l'élicitation incrémentale, basée sur l'optimisation d'un regret Minimax et sur l'heuristique CSS. Cette approche est intéressante, car elle offre de fortes garanties sur la convergence et sur la qualité de la recommandation finale, tout en limitant le coût cognitif pour le décideur. Néanmoins, cette approche n'est pas robuste à l'incertitude, que ce soit de la part du décideur ou du choix de modèle. Dès lors, dans le Chapitre 2, nous avons proposé une approche étendant l'approche précédente, qui permet de manipuler des préférences incertaines et de détecter de l'incohérence. Notre approche se base sur la théorie des possibilités, présentée dans le Chapitre 1, pour modéliser des préférences incertaines et pour les manipuler durant le processus d'élicitation. Nous avons montré formellement que notre approche possibiliste est une extension des approches ensemblistes initiales, et

qu'elle garde les mêmes garanties fortes. Nous avons vérifié au travers d'expériences synthétiques que notre approche permet bien de détecter de l'incohérence quelle que soit sa source, et de fournir des recommandations de meilleure qualité que les approches initiales, notamment en arrêtant l'élicitation une fois que l'incohérence est détectée afin que l'algorithme ne converge pas vers une mauvaise solution.

Notre approche possibiliste permet de détecter des incohérences, mais elle est assez coûteuse sur le plan computationnel malgré plusieurs propriétés pour limiter ce coût. Une piste parmi d'autres proposée dans le Chapitre 3 est d'utiliser un autre critère, celui du Maximax, permettant d'obtenir des recommandations en général semblables au regret Minimax, tout en ayant un coût computationnel moins lourd. Une autre perspective intéressante serait de regarder les liens de notre approche possibiliste avec la logique possibiliste, pour s'en inspirer afin de traiter l'incohérence, et de voir si notre approche peut être appliquée à d'autres modèles et d'autres situations (modèle lexicographique, problèmes combinatoires, tâches de rangement...). Enfin, il serait intéressant d'appliquer notre méthode sur des données réelles, afin de confirmer ou non son intérêt théorique.

Inférence et réparation de l'incohérence

Nous avons présenté dans le Chapitre 4 deux stratégies pour inférer malgré l'incohérence, qui permettent d'obtenir une recommandation finale sans modifier les informations préférentielles provenant du décideur ou la manière dont elles sont combinées, ni de restaurer la cohérence. Après avoir présenté dans le Chapitre 1 comment fusionner des informations potentiellement incertaines et incohérentes, nous avons proposé dans le Chapitre 4 deux opérateurs de fusion d'information, dans le but de restaurer la cohérence en cas d'incertitude de la part du décideur. Un premier opérateur permet de restaurer la cohérence en considérant que ℓ informations préférentielles parmi k , sans supposer quelles sont les mauvaises réponses, tandis qu'un deuxième opérateur permet de sélectionner un sous-ensemble d'informations préférentielles cohérentes, supposé contenir les bonnes réponses du décideur et potentiellement quelques mauvaises. Au travers d'expériences synthétiques, nous avons montré qu'une des stratégies d'inférence permet d'obtenir des recommandations bien meilleures que les approches ne prenant pas en compte l'incertitude.

De plus, nous avons montré que si les opérateurs de fusion ne permettent pas nécessairement d'améliorer la qualité des recommandations par rapport à une simple inférence, nous obtenons des informations supplémentaires très intéressantes pour estimer quelles sont les informations préférentielles erronées.

Dans cette thèse, nous avons étudié en détail la restauration de la cohérence en cas d'erreurs de la part du décideur. Néanmoins, l'incohérence peut aussi provenir d'un mauvais choix de modèle de préférence par l'analyste. Pour réparer ce type d'erreur, il nous semble plus pertinent de changer le modèle de préférence. Si nous avons montré à la fin du Chapitre 4 qu'il était possible de changer de modèle pour restaurer la cohérence, nous avons aussi mis en avant que changer le modèle ne permet pas toujours d'améliorer la qualité des recommandations finales. Nous supposons alors qu'une stratégie de questionnement pour un modèle donné n'est pas optimale pour un autre modèle, mais cela mériterait des recherches plus approfondies, ainsi que des preuves théoriques. Enfin, nous avons rapidement évoqué la problématique de différencier erreur du décideur (aléatoire) et erreur de modèle (systémique) qui nous semble très importante pour déterminer la source de l'incohérence et la stratégie à utiliser pour restaurer la cohérence. En dehors de quelques pistes portées par l'Annexe A, qui ne semblent pas être vérifiées en élicitation de préférences, cette problématique reste ouverte. Des pistes seraient d'obtenir des informations préférentielles plus riches (en changeant le type de question, quitte à augmenter la charge cognitive) ou à proposer un algorithme d'élicitation qui cherche à remettre en cause le modèle de préférence sélectionné.

MCSs multidimensionnels

A.1 Introduction

Ce chapitre annexe traite directement du problème de fusion d'information (Dubois *et al.*, 2016), sans lien direct avec l'élicitation de préférences, bien que nous verrons à la fin une piste intéressante. Dans ce problème, une tâche difficile est la gestion de conflits entre des sources d'information contradictoires. Le plus souvent, cette tâche est obligatoire, notamment dans les situations dans lesquelles nous souhaitons obtenir une estimation synthétique à partir de toutes les sources. En outre, l'analyse des causes de la contradiction et la tentative d'expliquer son apparition sont importantes, car elles peuvent donner des indications importantes.

Dubois *et al.* (1999) ont passé en revue différentes méthodes d'agrégation des conflits, certaines nécessitant des données supplémentaires telles que la fiabilité des différentes sources. Or, il est parfois difficile, voire impossible, de disposer de données supplémentaires. Nous avons déjà abordé dans la Sous-section 1.4.3 plusieurs outils pour fusionner des informations, dont une méthode assez attrayante pour traiter les contradictions, qui ne nécessite aucune information supplémentaire : les sous-ensembles cohérents maximaux (MCS). Les MCSs sont des ensembles de sources cohérentes aussi grands que possible. Les MCSs ont été utilisés dans le passé à la fois en logique (Manor et Rescher, 1970) et dans des contextes numériques (Destercke *et al.*, 2008). Nous illustrons ici leur application aux problèmes d'estimation.

La détection et l'énumération des MCSs sont des problèmes NP-difficiles, les intervalles étant une exception bien connue (Dubois *et al.*, 2000). Dans cet article,

nous montrons que nous pouvons étendre cette exception à certains hyperrectangles, qui peuvent à leur tour être utilisés comme approximations de polytopes.

Dans la Section A.2 nous introduisons très rapidement quelques éléments sur les polytopes. Nous montrons ensuite dans la Section A.3 qu'il est facile d'énumérer les MCS d'un ensemble d'hyperrectangles alignés sur les axes. Par la suite, dans la Section A.4 nous montrons qu'il est facile d'approximer des polytopes avec des hyperrectangles alignés sur les axes. Enfin, la Section A.5 illustre notre approche sur des problèmes de régression linéaire.

A.2 Compléments sur les MCSs de polytopes

Représentations d'un polytope Dans ce chapitre, nous considérons que les sources d'information sont dans l'espace euclidien de dimension \mathbb{R}^D . Dès lors, les polytopes sont des outils très polyvalents pour modéliser des informations qui prennent valeur dans un ensemble. Nous considérons ici que les polytopes sont nécessairement convexes et bornés. Dès lors, nous pouvons définir les polytopes de deux manières équivalentes :

Définition 38. *H-représentation d'un polytope*

Soit un ensemble de demi-espaces fermés dans \mathbb{R}^D , un demi-espace fermé étant défini par une inégalité linéaire $a_1x_1 + \dots + a_dx_d \leq b$. La H-représentation d'un polytope convexe borné est alors définie comme l'intersection d'un nombre fini de demi-espaces fermés, représentable par le système linéaire d'inégalités $Ax \leq b$.

Définition 39. *V-représentation d'un polytope*

Soit une enveloppe convexe définie comme étant le plus petit ensemble convexe qui contient tous les points d'un ensemble dans \mathbb{R}^D . La V-représentation d'un polytope convexe borné est alors l'enveloppe convexe d'un ensemble fini de points, où l'ensemble fini doit contenir l'ensemble des points extrêmes du polytope.

Le passage d'une représentation à une autre est difficile. Passer d'une H-représentation vers une V-représentation est le problème d'énumération des sommets (Motzkin *et al.*, 1953; Fukuda et Prodon, 1995). Il s'agit d'un problème NP-hard dans le cas général (Khachiyan *et al.*, 2009), sauf pour les polytopes bornés avec

167A.3. Énumération des MCSs d'hyperrectangles alignés sur les axes

un système d'inégalités non dégénéré (Avis et Fukuda, 1992). Passer d'une V-représentation vers une H-représentation est le problème d'énumération de facetes (généralisation des faces pour des polytopes en d dimensions), lié au calcul de l'enveloppe convexe qui est aussi un problème difficile dès que le nombre de dimensions dépasse 3 (Avis et Bremner, 1995).

MCSs de polytopes Rappelons rapidement la définition d'un MCS que nous avons déjà présenté dans la Sous-section 1.4.3. Nous supposons que nous avons un ensemble d'information $\mathcal{S} = \{S_1, \dots, S_N\}$ pour lequel l'intersection \cap est bien définie. Un MCS $c \subseteq \{1, \dots, N\}$ est un ensemble d'indices tel que :

- $\bigcap_{i \in c} S_i \neq \emptyset$,
- $\forall j \notin c, \bigcap_{i \in c} S_i \cap S_j = \emptyset$.

Déterminer si deux polytopes définis par leurs V-représentations forment une intersection non-vide est un problème NP-hard (Tiwary, 2008). Le même problème peut être résolu facilement en H-représentation avec un algorithme de programmation linéaire. Trouver un MCS unique en H-représentation est donc facile : nous ajoutons des polytopes un par un, et nous avons un MCS dès qu'il n'est plus possible d'en ajouter un autre sans avoir une intersection vide. Vérifier si un ensemble est un MCS est également facile : nous vérifions que l'intersection des H-polytopes n'est pas vide et maximale. Cependant, l'énumération de tous les MCSs d'un ensemble de polytopes $\mathcal{P} = \{P_1, \dots, P_N\}$ nécessite dans le pire des cas de considérer tous les sous-ensembles de \mathcal{P} , donc au plus 2^N ensembles. Lorsque le nombre de polytopes est important, il devient impossible d'énumérer tous les MCS. Dans la Section A.3, nous proposons un algorithme efficace pour énumérer tous les MCSs par le biais d'une approximation des polytopes avec des hyperrectangles alignés sur les axes.

A.3 Énumération des MCSs d'hyperrectangles alignés sur les axes

L'énumération de tous les MCSs d'un ensemble de polytopes est un problème difficile dans le cas général. Cependant, des algorithmes polynomiaux existent

pour déterminer les MCSs d'intervalles (Dubois *et al.*, 2000). Nous montrerons dans cette section que ces résultats peuvent également être utilisés dans le cas où nous considérons un ensemble \mathcal{H} d'hyperrectangles alignés sur les axes, afin de déterminer efficacement l'ensemble des MCSs $\mathcal{C}_{\mathcal{H}}$.

Nous notons par $I_{H_i}^d \in \mathbb{R}$ la projection de H_i sur la d^e dimension de l'espace \mathbb{R}^D . Nous avons une équivalence importante entre l'intersection d'hyperrectangles et l'intersection de leurs projections :

Proposition 8. *Équivalence intersection d'hyperrectangles et de projections*

Soit un ensemble $\mathcal{H} = \{H_1, \dots, H_N\}$ d'hyperrectangles alignés sur les axes dans l'espace \mathbb{R}^D , et soient leurs projections $I_{H_i}^d \in \mathbb{R}$ sur les différentes dimensions $d \in \{1, \dots, D\}$. Nous avons alors :

$$\bigcap_{H_i \in \mathcal{H}} H_i \neq \emptyset \iff \bigcap_{\mathcal{H}} I_{H_i}^d \neq \emptyset \quad \forall d \in \{1, \dots, D\}. \quad (\text{A.1})$$

Démonstration. Pour voir l'équivalence, il suffit d'observer que $\times_{d=1}^D \cap_{i=1}^N I_{H_i}^d = \bigcap_{H_i \in \mathcal{H}} H_i$. Cela signifie que pour n'importe quel point $u \in \mathbb{R}^D$ tel que $u_d \in \cap_{i=1}^N I_{H_i}^d \quad \forall d \in \{1, \dots, D\}$, alors ce point sera aussi dans $\bigcap_{H_i \in \mathcal{H}} H_i$. Notons que cela n'est vrai que pour les hyperrectangles alignés sur les axes. \square

Le corollaire suivant, qui est simplement une négation de la Proposition 8, est utile dans une démonstration ultérieure :

Corollaire 1.

$$\bigcap_{\mathcal{H}} H_i = \emptyset \iff \exists d \in \{1, \dots, D\} \text{ tel que } \bigcap_{\mathcal{H}} I_{H_i}^d = \emptyset. \quad (\text{A.2})$$

Dans la preuve suivante, nous montrons que les MCSs d'hyperrectangles alignés sur les axes peuvent être trouvés exactement en combinant les MCSs de leurs projections, qui peuvent être trouvées en temps polynomial, comme les projections sont des intervalles et qu'il a été montré qu'il est possible de trouver les MCSs d'intervalles rapidement (Dubois *et al.*, 2000).

Proposition 9. *Soit un ensemble $\mathcal{H} = \{H_1, \dots, H_N\}$ d'hyperrectangles alignés sur les axes dans l'espace \mathbb{R}^D , son ensemble de MCSs $\mathcal{C}_{\mathcal{H}}$, ainsi que l'ensemble de MCSs $\mathcal{C}_{I_{\mathcal{H}}^d}$ de ses projections dans la d^e dimension. Nous avons alors :*

$$\mathcal{C}_{\mathcal{H}} = \left\{ \bigcap_{d=1}^D c_d \mid c_d \in \mathcal{C}_{I_{\mathcal{H}}^d} \quad \forall d \in \{1, \dots, D\}, \bigcap_{d=1}^D c_d \neq \emptyset \right\}. \quad (\text{A.3})$$

169A.3. Énumération des MCSs d'hyperrectangles alignés sur les axes

Démonstration. Nous procédons en montrant une double inclusion pour un MCS donné $c_h \in \mathcal{C}_{\mathcal{H}}$:

- Nous montrons tout d'abord que $c_i \in \left\{ \bigcap_{d=1}^D c_d \mid c_d \in \mathcal{C}_{I_{\mathcal{H}}^d} \right\}$ tel que $c_h \subseteq c_i$, c'est-à-dire que c_i est une approximation externe de c_h . Comme $\bigcap_{i \in c_h} H_i \neq \emptyset$ (étant un MCS), la Proposition 8 nous indique que pour n'importe quelle dimension d , $\bigcap_{i \in c_h} I_{H_i}^d \neq \emptyset$, indiquant qu'il y aura un MCS $c_d \in \mathcal{C}_{I_{\mathcal{H}}^d}$ tel que $c_h \subseteq c_d$. Comme cela est vrai pour tout $d \in \{1, \dots, D\}$, nous avons alors $c_h \subseteq \bigcap_{d=1}^D c_d$ pour un ensemble de c_d , montrant l'inclusion.
- Pour démontrer l'autre inclusion, nous considérons un ensemble $c_d, d \in \{1, \dots, D\}$ de MCSs sur la dimension d qui sont des approximations externes $c_h \subseteq c_d$ dont nous connaissons l'existence grâce à la première partie de la preuve. Nous démontrons alors que $j \notin c_h$ implique $j \notin \bigcap_{d=1}^D c_d$, et donc que $\bigcap_{d=1}^D c_d \subseteq c_h$. Pour voir cela, considérons un ensemble d'hyperrectangles $H_k, k \in c_h \cup \{j\}$. Dès lors, par le Corollaire 1, il existera une dimension d telle que $\bigcap_{k \in c_h \cup \{j\}} I_{H_k}^d = \emptyset$, pourtant $\bigcap_{k \in c_h} I_{H_k}^d \neq \emptyset$ (c_h étant un MCS). Cela montre que $j \notin c_h$ implique $j \notin \bigcap_{d=1}^D c_d$.

□

La Proposition 9 nous fournit une approche intéressante pour obtenir les MCSs : nous commençons par projeter les hyperrectangles sur les différentes dimensions, afin d'obtenir des intervalles. Ensuite, nous énumérons les MCSs sur chaque dimension, ce qui est faisable en un temps polynomial. Enfin, nous déterminons $\mathcal{C}_{\mathcal{H}}$ comme l'ensemble des sources communes à toutes les énumérations, c'est-à-dire la combinaison conjonctive (intersection) des différentes sources pour chaque élément c_h de $\mathcal{C}_{\mathcal{H}}$. D'après la Proposition 9 et l'Équation (A.3), nous obtenons exactement l'ensemble des MCSs.

Exemple 38. *Illustration de l'approche pour déterminer les MCSs*

Sur la Figure A.1 est représenté un ensemble $\mathcal{H} = \{H_1, \dots, H_5\}$ d'hyperrectangles alignés sur les axes dans l'espace \mathbb{R}^2 . Nous pouvons voir trois intersections non-vides et maximales entre H_2 et H_3 , entre H_4 et H_5 , ainsi qu'avec le singleton H_1 qui ne forme aucune intersection non-vide avec les autres hyperrectangles. Le but est de retrouver ces sous-ensembles avec l'algorithme que nous avons présenté.

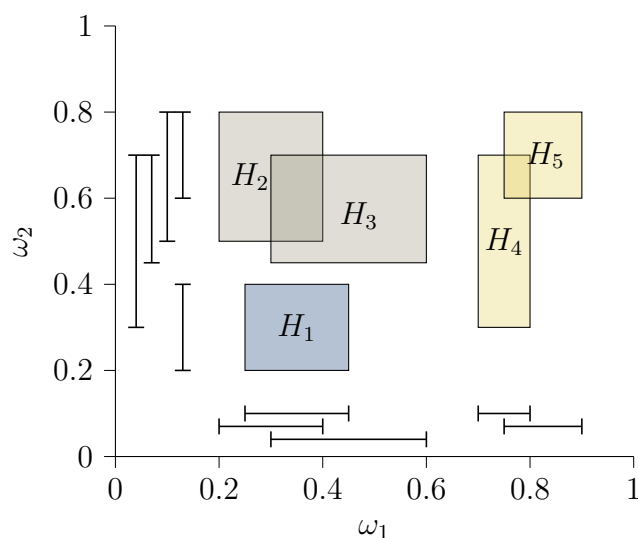


FIGURE A.1 – Illustration de notre algorithme pour déterminer les MCSs d'un ensemble d'hyperrectangles alignés sur les axes.

Nous avons représenté les projections de ces hyperrectangles sur chacune des dimensions. Nous pouvons alors déterminer que :

- Sur la première dimension : $\mathcal{C}_{I_{\mathcal{H}}^1} = \{\{1, 2, 3\}, \{4, 5\}\}$.
- Sur la deuxième dimension : $\mathcal{C}_{I_{\mathcal{H}}^2} = \{\{2, 3, 4, 5\}, \{1, 4\}\}$.

En combinant ces ensembles de MCSs, nous trouvons quatre sous-ensembles cohérents : $\{\{2, 3\}, \{1\}, \{4, 5\}, \{4\}\}$. En enlevant le dernier sous-ensemble qui n'est pas maximal, nous retrouvons les MCSs réels des hyperrectangles de la Figure A.1.

A.4 Approximation de polytopes avec des hyperrectangles

Nous avons vu dans la section précédente que nous pouvons déterminer les MCSs d'un ensemble d'hyperrectangles alignés sur les axes. Nous cherchons désormais à approximer des polytopes généraux avec des hyperrectangles alignés sur les axes spécifiques.

A.4.1 Hyperrectangle inscrit aligné sur les axes d'hypervolume maximal (MVAIR)

L'hyperrectangle inscrit d'hypervolume maximal, noté MVIR, est l'hyperrectangle d'hypervolume maximal qui est entièrement contenu dans un polytope. Étant donné un polytope $P = \{x \mid Ax \leq b\}$ et un hyperrectangle H paramétré par les longueurs de ses côtés $l = (l_1, \dots, l_D)$, la recherche du MVIR est formulée ainsi :

$$\max_l \left\{ \prod_l : H \subseteq \mathcal{P} \right\}. \quad (\text{A.4})$$

Dans le cas général, trouver le MVIR nécessite de résoudre un problème d'optimisation non convexe avec un nombre exponentiel de contraintes (Behroozi, 2019). Cependant, si nous considérons H comme un hyperrectangle aligné sur les axes, noté MVAIR, le problème d'optimisation devient convexe, résoluble assez efficacement.

Un MVAIR est une approximation interne d'un polytope, ce qui signifie que l'approximation est entièrement contenue dans le polytope :

Proposition 10. *Soit un ensemble de polytopes \mathcal{P} et un ensemble \mathcal{P}^I d'approximations internes. Nous avons alors :*

$$\bigcap_{\mathcal{P}^I} P_i^I \neq \emptyset \Rightarrow \bigcap_{\mathcal{P}} P_i \neq \emptyset, \quad (\text{A.5})$$

$$\bigcap_{\mathcal{P}} P_i = \emptyset \Rightarrow \bigcap_{\mathcal{P}^I} P_i^I = \emptyset. \quad (\text{A.6})$$

Les contraposées sont en général fausses : si les polytopes ont une intersection non-vide, celle des approximations internes peut être vide. Si les approximations internes ont une intersection vide, celle des polytopes peut être non-vide.

Démonstration. Par définition, nous avons $u \in P_i^I \Rightarrow u \in P_i$, comme $P_i^I \subseteq P_i$.

- $\bigcap_{\mathcal{P}^I} P_i^I \neq \emptyset \Rightarrow \exists u$ tel que $u \in P_i^I, \forall P_i^I \in I$, et donc $u \in P_i, \forall P_i \in \mathcal{P}$, comme un point de P_i^I est forcément inclus dans P_i . Nous avons donc une intersection non-vide pour \mathcal{P} qui contient au moins u .
- $\bigcap_{\mathcal{P}} P_i = \emptyset \Rightarrow \forall u, \exists P_i \in \mathcal{P}$ tel que $u \notin P_i$, dès lors $u \notin P_i \Rightarrow u \notin P_i^I$, comme tous les points de P_i^I sont inclus dans P_i . Nous obtenons alors une intersection vide pour \mathcal{P}^I .

□

A.4.2 Hyperrectangle englobant aligné sur les axes d'hy-pervolume minimal (MVABH)

L'hyperrectangle englobant d'hy-pervolume minimal, noté MVBH, est l'hyperrectangle d'hy-pervolume minimal qui englobe complètement un polytope. Étant donné un polytope $P = \{x \mid Ax \leq b\}$ et un hyperrectangle H paramétré par les longueurs $l = (l_1, \dots, l_D)$ de ses côtés, la recherche du MVBH est formulée ainsi :

$$\min_l \left\{ \prod_l : P \subseteq H \right\}. \quad (\text{A.7})$$

Résoudre ce problème peut s'avérer difficile en fonction de la représentation du polytope, car nous avons besoin de son enveloppe convexe. Puisque ici un polytope est défini par sa H-représentation, le problème est NP-hard dans le cas général comme nous l'avons vu dans la Section A.2, avec néanmoins des exceptions que nous pourrions exploiter. Néanmoins, en considérant que l'hyperrectangle doit être aligné sur les axes, et le nouveau problème noté MVABH, nous avons ajouté des contraintes sur les bords, ce qui rend le problème trivial à résoudre, puisqu'il nous suffit de trouver les extrema de notre polytope sur chaque dimension.

Un MVABH est une approximation externe de notre polytope, ce qui signifie que l'approximation contient entièrement le polytope :

Proposition 11. *Soit un ensemble de polytopes \mathcal{P} et l'ensemble \mathcal{P}^O des approximations externes correspondantes. Nous avons alors :*

$$\bigcap_{\mathcal{P}} P_i \neq \emptyset \Rightarrow \bigcap_{\mathcal{P}^O} P_i^O \neq \emptyset, \quad (\text{A.8})$$

$$\bigcap_{\mathcal{P}^O} P_i^O = \emptyset \Rightarrow \bigcap_{\mathcal{P}} P_i = \emptyset. \quad (\text{A.9})$$

Les contraposées sont en général fausses : si les approximations externes ont une intersection non-vide, celle des polytopes peut être vide. Si les polytopes ont une intersection vide, celle des approximations extérieures peut être non-vide.

Démonstration. Par définition, nous avons $u \in P_i \Rightarrow u \in P_i^O$, comme $P_i \subseteq P_i^O$. Nous pouvons procéder comme dans la preuve de la Proposition 10, en remplaçant \mathcal{P}^I par \mathcal{P} et \mathcal{P} par \mathcal{P}^O . \square

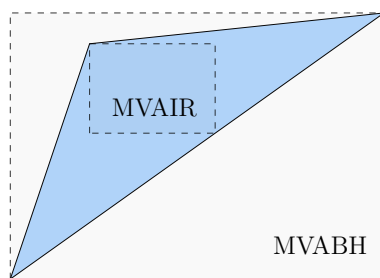


FIGURE A.2 – MVAIR et MVABR d'un polytope

A.4.3 Encadrement des MCSs

Notre objectif était de déterminer tous les MCSs d'un ensemble de polytopes. Avec des approximations internes et externes de nos polytopes, il est possible de donner des bornes sur l'ensemble des MCSs :

Proposition 12. *Étant donné $\mathcal{C}_{\mathcal{P}}$ les MCSs de l'ensemble des polytopes \mathcal{P} , $\mathcal{C}_{\mathcal{P}^I}$ et $\mathcal{C}_{\mathcal{P}^O}$ respectivement les MCSs de l'ensemble des approximations internes \mathcal{P}^I et de l'ensemble des approximations externes \mathcal{P}^O , nous avons la relation suivante :*

$$\forall c_{\mathcal{P}} \in \mathcal{C}_{\mathcal{P}}, \exists c_{\mathcal{P}^I} \in \mathcal{C}_{\mathcal{P}^I}, \exists c_{\mathcal{P}^O} \in \mathcal{C}_{\mathcal{P}^O}, \text{ tel que } c_{\mathcal{P}^I} \subseteq c_{\mathcal{P}} \subseteq c_{\mathcal{P}^O}. \quad (\text{A.10})$$

Démonstration. Nous prouvons tout d'abord $\exists c_{\mathcal{P}^I}$, puis $\exists c_{\mathcal{P}^O}$:

- Si $c_{\mathcal{P}} \in \mathcal{C}_{\mathcal{P}}$ est un singleton (aucune intersection avec un autre polytope), nous avons trivialement que $c_{\mathcal{P}^I} = c_{\mathcal{P}}$. Considérons maintenant $\bigcap_{P_i \in \mathcal{P}} P_i \neq \emptyset$. Étant donné que $P_i^I \subseteq P_i$, nous pouvons avoir $P_i^I \notin \bigcap_{P_i \in \mathcal{P}} P_i \wedge P_i \in \bigcap_{P_i \in \mathcal{P}} P_i$, mais nous ne pouvons pas avoir $P_i^I \in \bigcap_{P_i \in \mathcal{P}} P_i \wedge P_i \notin \bigcap_{P_i \in \mathcal{P}} P_i, \forall P_i \in \mathcal{P}$. Dès lors, $\bigcap_{P_i \in \mathcal{P}} P_i$ contient autant ou moins d'approximations internes que de polytopes, et aucune approximation interne dont le polytope ne serait pas dans $\bigcap_{P_i \in \mathcal{P}} P_i$. Nous avons donc au moins un $c_{\mathcal{P}^I}$ tel que $c_{\mathcal{P}^I} \subseteq c_{\mathcal{P}}$.
- Si $c_{\mathcal{P}} \in \mathcal{C}_{\mathcal{P}}$ contient tous les polytopes de l'espace, nous avons trivialement que $c_{\mathcal{P}^O} = c_{\mathcal{P}}$. Considérons maintenant $\bigcap_{P_i \in \mathcal{P}} P_i \neq \emptyset$. Étant donné que $P_i \subseteq P_i^O$, nous pouvons avoir $P_i \notin \bigcap_{P_i \in \mathcal{P}} P_i \wedge P_i^O \in \bigcap_{P_i \in \mathcal{P}} P_i$, mais nous ne pouvons pas avoir $P_i \in \bigcap_{P_i \in \mathcal{P}} P_i \wedge P_i^O \notin \bigcap_{P_i \in \mathcal{P}} P_i, \forall P_i \in \mathcal{P}$. Dès lors, $\bigcap_{P_i \in \mathcal{P}} P_i$ contient autant ou plus d'approximations externes que de polytopes. Nous avons donc au moins un $c_{\mathcal{P}^O}$ tel que $c_{\mathcal{P}} \subseteq c_{\mathcal{P}^O}$.

□

A.5 Illustration sur une estimation linéaire

Dans cette section, nous illustrons notre méthode sur de petits problèmes de régression linéaire imprécise à seulement deux dimensions¹. L'objectif de cette section est purement illustratif, afin de montrer l'utilité potentielle de notre approche lors d'estimation d'un point de vue logique et théorique.

Étant donné un ensemble de données $\{y_i, x_i\}_{i=1}^N$ avec une seule entrée et une seule sortie, un modèle linéaire suppose que la relation entre la variable de réponse y_i et la variable d'entrée x_i est linéaire :

$$y_i = \beta_0 + \beta_1 x_i. \quad (\text{A.11})$$

Habituellement, les y_i sont considérés comme observés avec un bruit (normal) ϵ , et une régression statistique est effectuée. Nous adopterons ici un point de vue logique : celui de l'espace des versions (Mitchell, 1982). Nous supposons que les données sont à valeurs fixes, et nous examinons tous les modèles linéaires compatibles.

A.5.1 Estimer les modèles linéaires possibles

Nous supposons que nous observons des données imprécises $R_i = ([y_i, \bar{y}_i], [x_i, \bar{x}_i])$. Étant donné l'équation cartésienne d'une droite $L = \{(x, y) \mid ax + by = c\}$ et deux points imprécis R_i et R_j , trouver toutes les droites qui coupent les deux rectangles est formalisé comme suit :

$$\mathcal{L}_{ij} = \{(a, b) \mid (L \cap R_i \neq \emptyset) \wedge (L \cap R_j \neq \emptyset)\}. \quad (\text{A.12})$$

La Figure A.3 est une illustration de ce problème, où nous cherchons toutes les droites coupant nos points imprécis.

n points peuvent être coupés par une seule ligne si et seulement si tous les \mathcal{L}_{ij} correspondants ($\binom{n}{2}$ au total) ont une intersection non vide, c'est-à-dire que leurs indices appartiennent à un seul sous-ensemble cohérent. Elle est maximale si aucun autre point ne peut être intersecté par la même ligne. Il est difficile d'énumérer toutes les MCSs, mais si nous approximons les différents \mathcal{L}_{ij} avec des rectangles alignés sur les axes, alors nous pouvons utiliser l'algorithme de la Section A.3.

1. https://github.com/LoicAdam/MCS_Multi

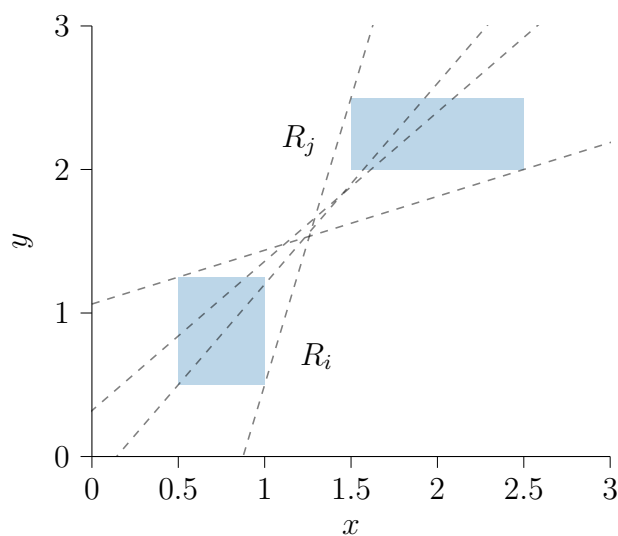


FIGURE A.3 – Illustration de plusieurs droites respectant l'Équation (A.12)

Déterminer le rectangle $H_{ij} = [a_i, \bar{a}_i] \times [b_i, \bar{b}_i]$ qui est d'aire minimale, aligné sur les axes, et englobant \mathcal{L}_{ij} , donc le MVABH, équivaut à trouver les valeurs minimale et maximale des paramètres a et b . Cela peut se faire assez facilement, car maximiser (respectivement minimiser) a est équivalent à minimiser (respectivement maximiser) b . L'approximation de \mathcal{L}_{ij} , notée $\hat{\mathcal{L}}_{ij}$, est représentée sur la Figure A.4.

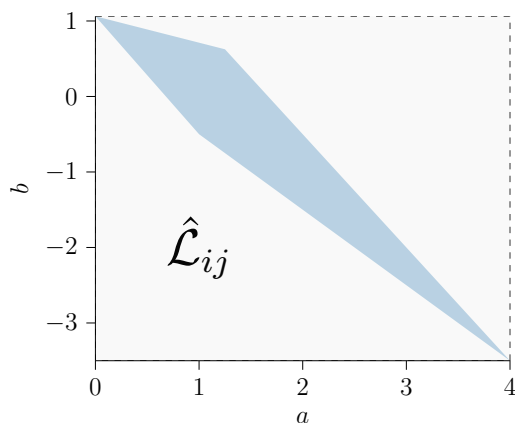


FIGURE A.4 – Approximation de l'ensemble des valeurs possibles de L par un MVABH

A.5.2 Application

Nous commençons avec un modèle effectivement linéaire, mais avec une donnée aberrante (ici le deuxième point), comme illustré sur la Figure A.5. Comme nous pouvons le voir, le modèle de régression statistique, en raison de la valeur aberrante, ne capture pas le véritable modèle.

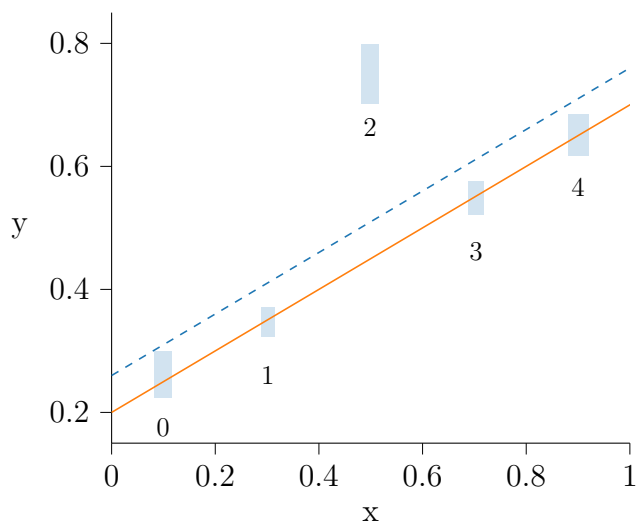


FIGURE A.5 – Exemple d'un problème de régression linéaire imprécise. La ligne en pointillés correspond à la régression linéaire incluant la valeur aberrante. La ligne continue correspond à la vérité terrain

La Figure A.6 montre les cinq MCSs (et leur intersection) obtenus en utilisant des approximations de \mathcal{L}_{ij} . Nous pouvons voir que le plus grand MCS est c_3 , et qu'il inclut les vrais paramètres, tandis que les autres MCSs sont plus petites et mettent en évidence les éventuelles aberrations du modèle : nous voyons que l'observation 2 se trouve dans tous les MCSs restantes, mais pas dans le plus grand. Il convient de noter qu'une telle approche est très différente des régressions imprécises standards, inspirées des moindres carrés (Ferraro *et al.*, 2010).

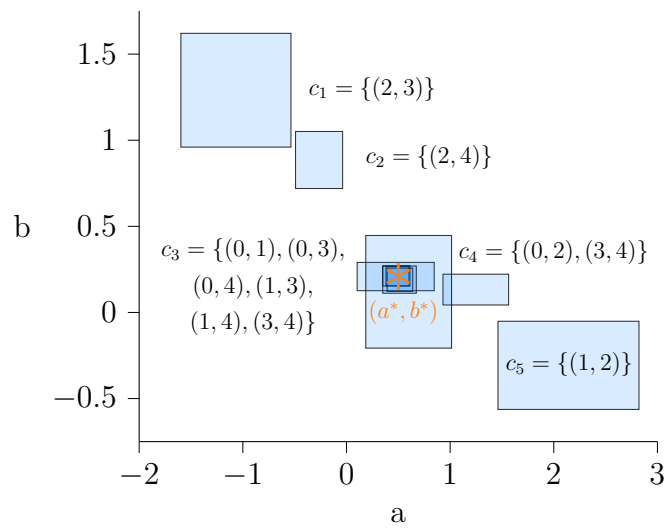


FIGURE A.6 – Approximation par des rectangles des différents \mathcal{L}_{ij} de l'exemple montré sur la Figure A.5

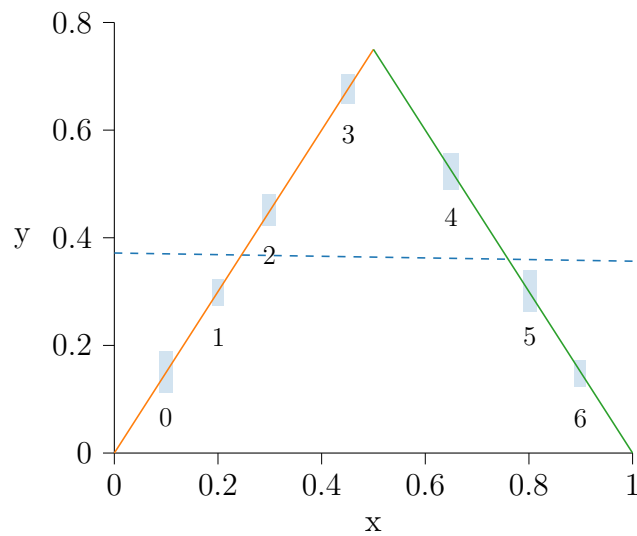


FIGURE A.7 – Exemple d'un problème de régression linéaire par morceaux imprécise. La ligne en pointillé correspond à la régression linéaire complète. Les lignes continues correspondent aux vérités terrain par morceaux

La deuxième illustration considère des observations non erronées (le modèle passe par toutes les observations), mais où l'hypothèse du modèle est erronée, car nous avons une régression linéaire par morceaux comme le montre la Figure A.7, et non une régression linéaire. Nous avons deux partitions sur x : $[0, 0.5]$ et $[0.5, 1]$. Les deux vérités de base sont très différentes, et une régression linéaire statistique unique (en pointillé) sur l'ensemble du domaine x s'ajuste mal aux données.

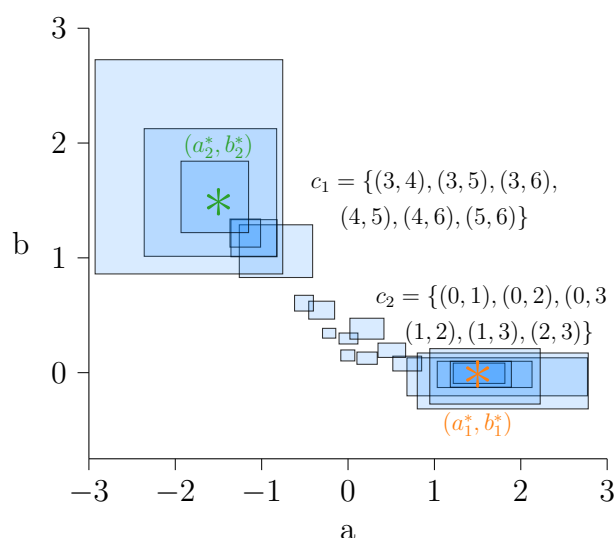


FIGURE A.8 – Approximation par des rectangles des différents \mathcal{L}_{ij} de l'exemple montré sur la Figure A.7

La Figure A.8 montre ce qui se passe lorsque nous envisageons une approche MCS dans ce cas. Au lieu d'avoir un grand MCS et d'autres plus petits, nous avons maintenant deux MCSs assez grands c_1 et c_2 , ainsi que d'autres isolés. Ces deux MCSs distincts indiquent que l'erreur se situe probablement dans le modèle, qui peut n'être valide que localement. Cela montre également qu'il n'est pas évident de savoir si l'observation 3 appartient à une régression locale ou à l'autre, puisqu'elle se trouve dans les deux MCSs. Cette ambiguïté fait également que c_1 n'inclut pas le vrai paramètre (a_2^*, b_2^*) .

A.6 Conclusion et perspectives

Les MCSs sont des outils théoriquement très intéressants pour traiter des observations ou des éléments d'information contradictoires, mais difficile à dénombrer en pratique. Nous avons montré qu'une telle énumération était plus facile avec les hyperrectangles, car nous pouvons utiliser des algorithmes polynomiaux connus sur leur projection en forme d'intervalle. Nous avons montré qu'il était possible d'approximer des polytopes avec des hyperrectangles alignés sur les axes, pour avoir des approximations internes et externes de nos polytopes. Nous avons illustré leur utilisation possible dans des problèmes d'estimation, en montrant que leurs différents comportements pouvaient fournir des informations utiles (observations erronées contre modèles erronés).

Points importants vus dans ce chapitre :

- Énumérer les MCSs de polytopes est difficile, mais faisable pour des hyperrectangles alignés sur les axes.
- Possibilité d'approximer des polytopes avec des hyperrectangles alignés sur les axes.
- Applicable sur des problèmes d'estimation.

Perspectives :

- Il serait intéressant de faire des comparaisons entre ces approches d'estimation et les approches bayésiennes en cas de mauvaise spécification du modèle, à l'instar de travaux antérieurs visant à résoudre des problèmes inverses (Shinde *et al.*, 2021).
- Les résultats de l'illustration sont très intéressants, car ils sembleraient indiquer que les MCSs soient utiles pour distinguer des observations erronées de modèles erronés. Or, cela voudrait indiquer qu'en utilisant des MCSs en élicitation de préférences, et plus particulièrement sur les informations préférentielles, il pourrait être possible de distinguer des erreurs provenant du décideur, de celles provenant d'une mauvaise spécification du modèle de

préférence. Dans la fin du Chapitre 4, plus particulièrement dans les perspectives, nous évoquons ce problème, en regardant si nous pouvons effectivement clairement distinguer les deux types d'erreur avec les MCSs, ce qui nous permettrait d'appliquer des corrections adaptées. Néanmoins, pour le moment, nous n'arrivons pas à distinguer aussi clairement les deux types d'erreurs.

Contributions :

- Adam, Loïc et Destercke, Sébastien. Multi-dimensional maximal coherent subsets made easy : illustration on an estimation problem. Dans : *Building Bridges between Soft and Statistical Methodologies for Data Science, SMPS 2022, 14-16 septembre 2022, Valladolid, Espagne*. Cham : Springer International Publishing, 2022. p. 1-8.
- Adam, Loïc et Destercke, Sébastien. Sous-ensembles maximaux cohérents multidimensionnels rendus faciles : illustration sur un problème d'estimation. Dans : *31e Rencontres Francophones sur la Logique Floue et ses Applications, LFA 2022, 20-21 octobre 2022, Toulouse, France*. Cepadués Editions, 2022. p. 79-86.

Annexe B

Bibliographie

- ANGILELLA, S., CORRENTE, S. et GRECO, S. (2015). Stochastic multiobjective acceptability analysis for the choquet integral preference model and the scale construction problem. *European Journal of Operational Research*, 240(1):172–182. (Cité à la page 34.)
- ANGILELLA, S., GRECO, S. et MATARAZZO, B. (2010). Non-additive robust ordinal regression : A multiple criteria decision model based on the choquet integral. *European Journal of Operational Research*, 201(1):277–288. (Cité à la page 32.)
- ARROW, K. J. et HURWICZ, L. (1972). An optimality criterion for decision-making under ignorance. *Uncertainty and expectations in economics*, 1. (Cité à la page 26.)
- AUGUSTIN, T., COOLEN, F. P., DE COOMAN, G. et TROFFAES, M. C., éditeurs (2014). *Introduction to imprecise probabilities*. John Wiley & Sons. (Cité aux pages 2, 48, and 50.)
- AVIS, D. et BREMNER, D. (1995). How good are convex hull algorithms? *In Proceedings of the eleventh annual symposium on Computational geometry*, pages 20–28. (Cité à la page 167.)
- AVIS, D. et FUKUDA, K. (1992). A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete & Computational Geometry*, 8(3):295–313. (Cité à la page 167.)

- BANA E COSTA, C. A. (1986). A multicriteria decision aid methodology to deal with conflicting situations on the weights. *European Journal of Operational Research*, 26(1):22–34. (Cité à la page 34.)
- BANA E COSTA, C. A. et VANSNICK, J.-C. (1995). General overview of the macbeth approach. *Advances in multicriteria analysis*, pages 93–100. (Cité à la page 32.)
- BEHROOZI, M. (2019). Largest inscribed rectangles in geometric convex sets. *arXiv preprint arXiv :1905.13246*. (Cité à la page 171.)
- BELL, D. E., KEENEY, R. L. et RAIFFA, H., éditeurs (1977). *Conflicting objectives in decisions*. John Wiley & Sons. (Cité à la page 5.)
- BELTON, V. et STEWART, T. (2010). Problem structuring and multiple criteria decision analysis. *Trends in multiple criteria decision analysis*, pages 209–239. (Cité à la page 6.)
- BENABBOU, N. (2017). *Procédures de décision par élicitation incrémentale de préférences en optimisation multicritère, multi-agents et dans l'incertain*. Thèse de doctorat, Université Pierre et Marie Curie-Paris VI. (Cité aux pages 31 and 121.)
- BENABBOU, N., GONZALES, C., PERNY, P. et VIAPPANI, P. (2015). Minimax regret approaches for preference elicitation with rank-dependent aggregators. *EURO journal on Decision processes*, 3:29–64. (Cité à la page 44.)
- BENABBOU, N. et PERNY, P. (2018). Interactive resolution of multiobjective combinatorial optimization problems by incremental elicitation of criteria weights. *EURO journal on decision processes*, 6(3):283–319. (Cité aux pages 70, 121, and 151.)
- BENABBOU, N., PERNY, P. et VIAPPANI, P. (2017). Incremental elicitation of choquet capacities for multicriteria choice, ranking and sorting problems. *Artificial Intelligence*, 246:152–180. (Cité aux pages 36, 38, 70, 151, and 152.)
- BENAYOUN, R., DE MONTGOLFIER, J., TERGNY, J. et LARITCHEV, O. (1971). Linear programming with multiple objective functions : Step method (stem). *Mathematical programming*, 1(1):366–375. (Cité aux pages 1 and 36.)

- BENFERHAT, S., DUBOIS, D. et PRADE, H. (2002). Towards a possibilistic logic handling of preferences. *Aiding Decisions with Multiple Criteria : Essays in Honor of Bernard Roy*, pages 315–337. (Cité à la page 77.)
- BOOLE, G. (1854). *An investigation of the laws of thought : on which are founded the mathematical theories of logic and probabilities*, volume 2. Walton and Maberly. (Cité aux pages 48 and 49.)
- BOOTH, R., CHEVALEYRE, Y., LANG, J., MENGIN, J. et SOMBATTHEERA, C. (2010). Learning conditionally lexicographic preference relations. *In ECAI*, volume 10, pages 269–274. (Cité à la page 101.)
- BOURDACHE, N. (2020). *Élicitation incrémentale des préférences pour l'optimisation multi-objectifs : modèles non-linéaires, domaines combinatoires et approches tolérantes aux erreurs*. Thèse de doctorat, Sorbonne université. (Cité à la page 31.)
- BOURDACHE, N. et PERNY, P. (2017). Anytime algorithms for adaptive robust optimization with owa and wowa. *In International Conference on Algorithmic Decision Theory*, pages 93–107. Springer. (Cité aux pages 103 and 108.)
- BOURDACHE, N. et PERNY, P. (2019). Active preference learning based on generalized gini functions : Application to the multiagent knapsack problem. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:7741–7748. (Cité à la page 36.)
- BOURDACHE, N., PERNY, P. et SPANJAARD, O. (2019). Incremental elicitation of rank-dependent aggregation functions based on bayesian linear regression. *In IJCAI-19-Twenty-Eighth International Joint Conference on Artificial Intelligence*, pages 2023–2029. International Joint Conferences on Artificial Intelligence Organization. (Cité aux pages 45 and 71.)
- BOUTILIER, C. (2002). A POMDP formulation of preference elicitation problems. *In AAAI/IAAI*, pages 239–246. Edmonton, AB. (Cité à la page 41.)
- BOUTILIER, C. (2013). Computational decision support : Regret-based models for optimization and preference elicitation. *Comparative Decision Making : Analysis*

- and Support Across Disciplines and Applications*, pages 423–453. (Cité aux pages 70, 103, and 121.)
- BOUTILIER, C., PATRASCU, R., POUPART, P. et SCHUURMANS, D. (2006). Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence*, 170(8-9):686–713. (Cité aux pages 38, 42, 45, and 121.)
- BOUTILIER, C., REGAN, K. et VIAPPANI, P. (2009). Online feature elicitation in interactive optimization. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 73–80. (Cité à la page 38.)
- BOUYSSOU, D., MARCHANT, T., PIRLOT, M., TSOUKIAS, A. et VINCKE, P. (2006). *Evaluation and decision models with multiple criteria : Stepping stones for the analyst*, volume 86. Springer Science & Business Media. (Cité à la page 152.)
- BRADLEY, S. (2019). Imprecise probabilities. *Computer Simulation Validation : Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, pages 525–540. (Cité à la page 48.)
- BRAZIUNAS, D. (2012). *Decision-theoretic elicitation of generalized additive utilities*. University of Toronto (Canada). (Cité à la page 40.)
- BRESSON, R., COHEN, J., HÜLLERMEIER, E., LABREUCHE, C. et SEBAG, M. (2021). Neural representation and learning of hierarchical 2-additive choquet integrals. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, pages 1984–1991. (Cité à la page 22.)
- BUEDE, D. M. (1986). Structuring value attributes. *Interfaces*, 16(2):52–62. (Cité à la page 6.)
- CABANNES, V., RUDI, A. et BACH, F. (2020). Structured prediction with partial labelling through the infimum loss. In *International Conference on Machine Learning*, pages 1230–1239. PMLR. (Cité à la page 104.)
- CHATEAUNEUF, A. et JAFFRAY, J.-Y. (1989). Some characterizations of lower probabilities and other monotone capacities through the use of möbius inversion. *Mathematical social sciences*, 17(3):263–283. (Cité à la page 20.)

- CHATEAUNEUF, A. et JAFFRAY, J.-Y. (1995). Local möbius transforms of monotone capacities. In *Symbolic and Quantitative Approaches to Reasoning and Uncertainty : European Conference, ECSQARU'95 Fribourg, Switzerland, July 3–5, 1995 Proceedings 3*, pages 115–124. Springer. (Cité à la page 20.)
- CHATEAUNEUF, A. et TALLON, J.-M. (2002). Diversification, convex preferences and non-empty core in the choquet expected utility model. *Economic Theory*, 19:509–523. (Cité à la page 18.)
- CHOQUET, G. (1953). Théorie des capacités. In *Annales de l'Institut Fourier*, volume 5, pages 31–295. (Cité aux pages 2 and 17.)
- CIOMEK, K., KADZIŃSKI, M. et TERVONEN, T. (2017). Heuristics for prioritizing pair-wise elicitation questions with additive multi-attribute value models. *Omega*, 71:27–45. (Cité à la page 120.)
- COCHRANE, J. L. et ZELENY, M., éditeurs (1975). *Multiple Criteria Decision Making*. Springer. (Cité à la page 5.)
- CORRENTE, S., GRECO, S., KADZIŃSKI, M. et SŁOWIŃSKI, R. (2013). Robust ordinal regression in preference learning and ranking. *Machine Learning*, 93(2): 381–422. (Cité à la page 70.)
- DE FINETTI, B. (1937). La prévision : ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68. (Cité à la page 47.)
- DE FINETTI, B. (1972). *Probability, Induction and Statistics : The Art of Guessing*. J. Wiley. (Cité à la page 47.)
- DEMPSTER, A. P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics*, 38(2):325 – 339. (Cité à la page 49.)
- DENOEUX, T. (2019). Decision-making with belief functions : a review. *International Journal of Approximate Reasoning*, 109:87–110. (Cité à la page 49.)

- DESTERCKE, S., DUBOIS, D. et CHOJNACKI, E. (2008). Possibilistic information fusion using maximal coherent subsets. *IEEE Transactions on Fuzzy Systems*, 17(1):79–92. (Cité aux pages 78, 126, and 165.)
- DOUMPOS, M. et ZOPOUNIDIS, C. (2011). Preference disaggregation and statistical learning for multicriteria decision support : A review. *European Journal of Operational Research*, 209(3):203–214. (Cité à la page 31.)
- DUBOIS, D. (1986). Belief structures, possibility theory and decomposable confidence measures on finite sets. *Computers and Artificial Intelligence*, 5(5):403–416. (Cité aux pages 49 and 51.)
- DUBOIS, D., FARGIER, H. et PRADE, H. (2000). Multiple-sources informations fusion—a practical inconsistency-tolerant approach. In *8th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2000)*. Consejo Superior de Investigaciones Científicas. (Cité aux pages 165 and 168.)
- DUBOIS, D., LIU, W., MA, J. et PRADE, H. (2016). The basic principles of uncertain information fusion. an organised review of merging rules in different representation frameworks. *Information Fusion*, 32:12–39. (Cité aux pages 54, 58, 60, 78, 122, 126, and 165.)
- DUBOIS, D. et PRADE, H. (1998). Possibility theory : qualitative and quantitative aspects. In *Quantified representation of uncertainty and imprecision*, pages 169–226. Springer. (Cité aux pages 73 and 122.)
- DUBOIS, D. et PRADE, H. (2001). Possibility theory, probability theory and multiple-valued logics : A clarification. *Annals of mathematics and Artificial Intelligence*, 32:35–66. (Cité aux pages 3, 49, and 51.)
- DUBOIS, D. et PRADE, H. (2004). Possibilistic logic : a retrospective and prospective view. *Fuzzy sets and Systems*, 144(1):3–23. (Cité aux pages 77 and 127.)
- DUBOIS, D. et PRADE, H. (2015). Inconsistency management from the standpoint of possibilistic logic. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 23(Suppl. 1):15–30. (Cité à la page 102.)

- DUBOIS, D., PRADE, H. et YAGER, R. (1999). Merging fuzzy information. *Fuzzy sets in approximate reasoning and information systems*, pages 335–401. (Cité aux pages 60 and 165.)
- ELLSBERG, D. (1961). Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, 75(4):643–669. (Cité aux pages 48 and 49.)
- ERIC, B., FREITAS, N. et GHOSH, A. (2007). Active preference learning with discrete choice data. *Advances in neural information processing systems*, 20. (Cité aux pages 45 and 71.)
- EVERAERE, P., KONIECZNY, S. et MARQUIS, P. (2010). Disjunctive merging : Quota and gmin merging operators. *Artificial Intelligence*, 174(12-13):824–849. (Cité à la page 60.)
- FERRARO, M. B., COPPI, R., RODRÍGUEZ, G. G. et COLUBI, A. (2010). A linear regression model for imprecise response. *International Journal of Approximate Reasoning*, 51(7):759–770. (Cité à la page 176.)
- FUKUDA, K. et PRODON, A. (1995). Double description method revisited. In *Franco-Japanese and Franco-Chinese Conference on Combinatorics and Computer Science*, pages 91–111. Springer. (Cité à la page 166.)
- FÜRNKRANZ, J. et HÜLLERMEIER, E. (2011). *Preference learning*. Springer. (Cité à la page 69.)
- GALAND, L., PERNY, P. et SPANJAARD, O. (2010). A branch and bound algorithm for choquet optimization in multicriteria problems. In *Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems : Proceedings of the 19th International Conference on Multiple Criteria Decision Making, Auckland, New Zealand, 7th-12th January 2008*, pages 355–365. Springer. (Cité à la page 18.)
- GILBOA, I. (2009). *Theory of decision under uncertainty*, volume 45. Cambridge university press. (Cité à la page 23.)
- GRABISCH, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European journal of operational research*, 89(3):445–456. (Cité aux pages 17 and 21.)

- GRABISCH, M., KOJADINOVIC, I. et MEYER, P. (2008). A review of methods for capacity identification in choquet integral based multi-attribute utility theory : Applications of the kappalab r package. *European journal of operational research*, 186(2):766–785. (Cité aux pages 44 and 71.)
- GRABISCH, M. et LABREUCHE, C. (2005). Bi-capacities—ii : the choquet integral. *Fuzzy sets and systems*, 151(2):237–259. (Cité à la page 22.)
- GRABISCH, M. et LABREUCHE, C. (2010). A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, 175:247–286. (Cité à la page 18.)
- GRABISCH, M., MARICHAL, J.-L., MESIAR, R. et PAP, E. (2009). *Aggregation functions*, volume 127. Cambridge University Press. (Cité aux pages 12 and 13.)
- GRABISCH, M. et PERNY, P. (2003). Agrégation multicritère. In *Logique floue, principes, aide à la décision*, Traité IC2, série Informatique et Systèmes d’Information, pages 81–120. Hermès-Lavoisier. (Cité à la page 11.)
- GRECO, S., SŁOWIŃSKI, R., FIGUEIRA, J. R. et MOUSSEAU, V. (2010). *Robust Ordinal Regression*, pages 241–283. Springer US, Boston, MA. (Cité à la page 32.)
- GUILLOT, P.-L. et DESTERCKE, S. (2019). Preference elicitation with uncertainty : Extending regret based methods with belief functions. In *International Conference on Scalable Uncertainty Management*, pages 289–309. Springer. (Cité aux pages 73, 83, and 92.)
- HAMMOND, J. S., KEENEY, R. L. et RAIFFA, H. (1998). Even swaps : A rational method for making trade-offs. *Harvard business review*, 76:137–150. (Cité à la page 71.)
- HERIN, M., PERNY, P. et SOKOLOVSKA, N. (2022). Learning utilities and sparse representations of capacities for multicriteria decision making with the bipolar choquet integral. In *Multidisciplinary Workshop on Advances in Preference Handling, IJCAI*. (Cité aux pages 22 and 126.)

- HOLLOWAY, H. A. et WHITE III, C. C. (2003). Question selection for multi-attribute decision-aiding. *European Journal of Operational Research*, 148(3):525–533. (Cité aux pages 37 and 120.)
- HÜLLERMEIER, E. (2014). Learning from imprecise and fuzzy observations : Data disambiguation through generalized loss minimization. *International Journal of Approximate Reasoning*, 55(7):1519–1534. (Cité à la page 82.)
- HÜLLERMEIER, E., DESTERCCKE, S. et COUSO, I. (2019). Learning from imprecise data : adjustments of optimistic and pessimistic variants. *In International Conference on Scalable Uncertainty Management*, pages 266–279. Springer. (Cité à la page 120.)
- HÁJEK, A. (2019). Interpretations of Probability. *In ZALTA, E. N., éditeur : The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2019 édition. (Cité à la page 47.)
- JACQUET-LAGREZE, E. et SISKOS, J. (1982). Assessing a set of additive utility functions for multicriteria decision-making, the uta method. *European journal of operational research*, 10(2):151–164. (Cité aux pages 32 and 45.)
- JACQUET-LAGRÈZE, E. et SISKOS, Y. (2001). Preference disaggregation : 20 years of mcda experience. *European Journal of Operational Research*, 130(2):233–245. (Cité à la page 31.)
- JOYCE, J. M. (1999). *The foundations of causal decision theory*. Cambridge University Press. (Cité à la page 48.)
- KAPLAN, M. (1998). *Decision theory as philosophy*. Cambridge University Press. (Cité à la page 48.)
- KEENEY, R. L. et RAIFFA, H. (1976). *Decisions with multiple objectives : preferences and value trade-offs*. John Wiley & Sons. (Cité aux pages 1, 5, and 11.)
- KEYNES, J. M. (1921). *A treatise on probability*. Macmillan. (Cité à la page 49.)

- KHACHIYAN, L., BOROS, E., BORYS, K., GURVICH, V. et ELBASSIONI, K. (2009). Generating all vertices of a polyhedron is hard. *In Twentieth Anniversary Volume* :, pages 1–17. Springer. (Cité à la page 166.)
- KLEMENT, E., MESIAR, R. et PAP, E. (2000). *Triangular norms*, volume 8. Kluwer Academic Publishers. (Cité aux pages 61 and 63.)
- KÖBBERLING, V. et WAKKER, P. P. (2003). Preference foundations for nonexpected utility : A generalized and simplified technique. *Mathematics of Operations Research*, 28(3):395–423. (Cité à la page 22.)
- KOLMOGOROV, A. (1950). *Foundations of the theory of probability*. Chelsea Publishing Company. (Cité à la page 47.)
- KORHONEN, P. et WALLENIUS, J. (1988). A pareto race. *Naval Research Logistics (NRL)*, 35(6):615–623. (Cité à la page 36.)
- LABREUCHE, C. (2012). An axiomatization of the choquet integral and its utility functions without any commensurability assumption. *In Advances in Computational Intelligence : 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2012, Catania, Italy, July 9-13, 2012, Proceedings, Part IV 14*, pages 258–267. Springer. (Cité à la page 22.)
- LAHDELMA, R., HOKKANEN, J. et SALMINEN, P. (1998). Smaa-stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, 106(1):137–143. (Cité aux pages 34 and 45.)
- LOKMAN, B., KÖKSALAN, M., KORHONEN, P. J. et WALLENIUS, J. (2018). An interactive approximation algorithm for multi-objective integer programs. *Computers & Operations Research*, 96:80–90. (Cité à la page 36.)
- LU, T. et BOUTILIER, C. (2011). Robust approximation and incremental elicitation in voting protocols. *In IJCAI*, volume 1, pages 287–293. (Cité à la page 36.)
- MANOR, R. et RESCHER, N. (1970). On inference from inconsistent premises. *Theory and Decision*, 1:179–219. (Cité à la page 165.)

- MARICHAL, J.-L. (1998). *Aggregation operators for multicriteria decision aid*. Thèse de doctorat, University of Liège, Liège, Belgique. (Cité à la page 13.)
- MARICHAL, J.-L. (2000). An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria. *IEEE transactions on fuzzy systems*, 8(6):800–807. (Cité à la page 22.)
- MARICHAL, J.-L. et ROUBENS, M. (2000). Determination of weights of interacting criteria from a reference set. *European journal of operational Research*, 124(3):641–650. (Cité à la page 34.)
- MARON, O. et MOORE, A. W. (1997). The racing algorithm : Model selection for lazy learners. *Artificial Intelligence Review*, 11(1):193–225. (Cité à la page 104.)
- MARTIN, H. et PERNY, P. (2021). Incremental preference elicitation with bipolar choquet integrals. In *International Conference on Algorithmic Decision Theory*, pages 101–116. Springer. (Cité à la page 151.)
- MAYAG, B., GRABISCH, M. et LABREUCHE, C. (2011). Dealing with inconsistencies in the representation of ordinal information by a 2-additive choquet integral. In *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology*, pages 119–126. Atlantis Press. (Cité à la page 121.)
- MAYSTRE, L. Y., PICTET, J. et SIMOS, J. (1994). *Méthodes multicritères ELECTRE : description, conseils pratiques et cas d'application à la gestion environnementale*, volume 8. PPUR presses polytechniques. (Cité à la page 12.)
- MCMULLEN, P. (1970). The maximum numbers of faces of a convex polytope. *Mathematika*, 17(2):179–184. (Cité à la page 91.)
- MENGER, K. (1942). Statistical metrics. *Proceedings of the National Academy of Sciences of the United States of America*, 28(12):535–537. (Cité à la page 61.)
- MIETTINEN, K. (1999). *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media. (Cité à la page 36.)
- MITCHELL, T. M. (1982). Generalization as search. *Artificial intelligence*, 18(2):203–226. (Cité à la page 174.)

- MOTZKIN, T. S. (1957). Comonotone curves and polyhedra. *Bull. Amer. Math. Soc.*, 63. (Cité à la page 91.)
- MOTZKIN, T. S., RAIFFA, H., THOMPSON, G. L. et THRALL, R. M. (1953). The double description method. *Contributions to the Theory of Games*, 2(28):51–73. (Cité à la page 166.)
- MOUSSEAU, V., FIGUEIRA, J., DIAS, L., da SILVA, C. G. et CLIMACO, J. (2003). Resolving inconsistencies among constraints on the parameters of an mcda model. *European Journal of Operational Research*, 147(1):72–93. (Cité aux pages 121 and 126.)
- NELSEN, R. B. (1999). *An introduction to copulas*. Springer. (Cité à la page 63.)
- OUSSALAH, M. (2002). On the normalization of subnormal possibility distributions : New investigations. *International Journal of General Systems*, 31(3):277–301. (Cité à la page 129.)
- ÖZPEYNIRCI, Ö., ÖZPEYNIRCI, S. et MOUSSEAU, V. (2022). A decomposition based minimax regret approach for inverse multiple criteria sorting problem. *4OR*, pages 1–25. (Cité à la page 152.)
- PEKELMAN, D. et SEN, S. K. (1974). Mathematical programming models for the determination of attribute weights. *Management Science*, 20(8):1217–1229. (Cité à la page 34.)
- PICHON, F., DESTERCKE, S. et BURGER, T. (2014). A consistency-specificity trade-off to select source behavior in information fusion. *IEEE transactions on cybernetics*, 45(4):598–609. (Cité aux pages 78 and 126.)
- PIGOZZI, G., TSOUKIAS, A. et VIAPPANI, P. (2016). Preferences in artificial intelligence. *Annals of Mathematics and Artificial Intelligence*, 77:361–401. (Cité à la page 31.)
- PRADE, H., RICO, A. et SERRURIER, M. (2009a). Elicitation of sugeno integrals : A version space learning perspective. In *International Symposium on Methodologies for Intelligent Systems*, pages 392–401. Springer. (Cité à la page 70.)

- PRADE, H., RICO, A., SERRURIER, M. et RAUFASTE, E. (2009b). Eliciting sugeno integrals : Methodology and a case study. *In European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 712–723. Springer. (Cité à la page 122.)
- RESCHER, N. et MANOR, R. (1970). On inference from inconsistent premisses. *Theory and decision*, 1(2):179–217. (Cité aux pages 58 and 134.)
- RICO, A. et VIAPPIANI, P. (2020). Incremental elicitation of capacities for the sugeno integral with a maximin approach. *In International Conference on Scalable Uncertainty Management*, pages 156–171. Springer. (Cité à la page 151.)
- ROTA, G. C. (1964). On the foundations of combinatorial theory i. theory of möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2(4):340–368. (Cité à la page 19.)
- ROY, B. (1985). *Méthodologie multicritère d'aide à la décision*. Economica Paris. (Cité à la page 12.)
- ROY, B. (1996). *Multicriteria methodology for decision aiding*, volume 12. Springer Science & Business Media. (Cité à la page 6.)
- ROY, B. et BOUYSSOU, D. (1993). *Aide multicritère à la décision : méthodes et cas*. Economica Paris. (Cité à la page 12.)
- SALO, A. A. et HAMALAINEN, R. P. (2001). Preference ratios in multiattribute evaluation (prime)-elicitation and decision procedures under incomplete information. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 31(6):533–545. (Cité à la page 38.)
- SAVAGE, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical association*, 46(253):55–67. (Cité aux pages 27, 37, and 103.)
- SCHLÄFLI, L. et WILD, H. (2013). *Theorie der vielfachen Kontinuität*, volume 38. Springer-Verlag. (Cité à la page 90.)
- SCHMEIDLER, D. (1986). Integral representation without additivity. *Proceedings of the American mathematical society*, 97(2):255–261. (Cité aux pages 17 and 22.)

- SCHMITT, M., MARTIGNON, L. et RON, D. (2006). On the complexity of learning lexicographic strategies. *Journal of Machine Learning Research*, 7(1). (Cité à la page 121.)
- SCHWEISER, B. et SKLAR, A. (1961). Associative functions and statistical triangle inequalities. *Publicationes Mathematicae Debrecen*, 8:169–186. (Cité aux pages 61 and 64.)
- SCHWEIZER, B. et SKLAR, A. (1958). Espaces métriques aléatoires. *Comptes Rendus Académie Science Paris*, 247(23):2092–2094. (Cité à la page 61.)
- SCHWEIZER, B. et SKLAR, A. (1960). Statistical metric spaces. *Pacific J. Math*, 10(1):313–334. (Cité à la page 61.)
- SHACKLE, G. L. S. (1961). *Decision, order, and time in human affairs*. CUP Archive. (Cité à la page 50.)
- SHAFER, G. (1976). *A mathematical theory of evidence*, volume 42. Princeton university press. (Cité à la page 49.)
- SHINDE, K., FEISSEL, P. et DESTERCKE, S. (2021). Dealing with inconsistent measurements in inverse problems : set-based approach. *International Journal for Uncertainty Quantification*, 11(3). (Cité à la page 179.)
- SISKOS, Y. et YANNACOPOULOS, D. (1985). Utastar : An ordinal regression method for building additive value functions. *Investigação Operacional*, 5(1):39–53. (Cité à la page 32.)
- SMETS, P. *et al.* (1989). Constructing the pignistic probability function in a context of uncertainty. *In UAI*, volume 89, pages 29–40. (Cité à la page 138.)
- SMITH, C. A. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society : Series B (Methodological)*, 23(1):1–25. (Cité à la page 49.)
- SRINIVASAN, V. et SHOCKER, A. D. (1973). Estimating the weights for multiple attributes in a composite criterion using pairwise judgments. *Psychometrika*, 38(4):473–493. (Cité aux pages 34 and 45.)

- TERVONEN, T. et FIGUEIRA, J. R. (2008). A survey on stochastic multicriteria acceptability analysis methods. *Journal of Multi-Criteria Decision Analysis*, 15(1-2):1–14. (Cité à la page 34.)
- TESO, S., PASSERINI, A. et VIAPPIANI, P. (2017). Constructive preference elicitation for multiple users with setwise max-margin. *In Algorithmic Decision Theory : 5th International Conference, ADT 2017, Luxembourg, Luxembourg, October 25–27, 2017, Proceedings 5*, pages 3–17. Springer. (Cité aux pages 44 and 71.)
- THIRIEZ, H. et ZIONTS, S., éditeurs (1975). *Multiple Criteria Decision Making*. Springer. (Cité à la page 5.)
- TIWARY, H. R. (2008). On the hardness of computing intersection, union and minkowski sum of polytopes. *Discrete & Computational Geometry*, 40(3):469–479. (Cité à la page 167.)
- TORNEDE, A., WEVER, M. et HÜLLERMEIER, E. (2020). Extreme algorithm selection with dyadic feature representation. *In International Conference on Discovery Science*, pages 309–324. Springer. (Cité à la page 120.)
- TORRA, V. (1997). The weighted owa operator. *International Journal of Intelligent Systems*, 12(2):153–166. (Cité aux pages 16 and 18.)
- TROFFAES, M. C. (2007). Decision making under uncertainty using imprecise probabilities. *International journal of approximate reasoning*, 45(1):17–29. (Cité aux pages 24 and 103.)
- VANDERPOOTEN, D. et VINCKE, P. (1989). Description and analysis of some representative interactive multicriteria procedures. *In Models and Methods in Multiple Criteria Decision Making*, pages 1221–1238. Elsevier. (Cité à la page 36.)
- VENN, J. (1888). *The logic of chance*. Macmillan, 3. édition. (Cité à la page 47.)
- VIAPPIANI, P. et BOUTILIER, C. (2009). Regret-based optimal recommendation sets in conversational recommender systems. *In Proceedings of the third ACM conference on Recommender systems*, pages 101–108. (Cité à la page 41.)

- VIAPPANI, P. et BOUTILIER, C. (2010). Optimal bayesian recommendation sets and myopically optimal choice query sets. *Advances in neural information processing systems*, 23. (Cité aux pages 45 and 71.)
- VIAPPANI, P. et BOUTILIER, C. (2020). On the equivalence of optimal recommendation sets and myopically optimal query sets. *Artificial Intelligence*, 286:103328. (Cité à la page 69.)
- von WINTERFELDT, D. (1980). Structuring decision problems for decision analysis. *Acta Psychologica*, 45(1-3):71–93. (Cité à la page 6.)
- WALLEY, P. (1991). *Statistical reasoning with imprecise probabilities*, volume 42. Springer. (Cité à la page 49.)
- WALLEY, P. (1996). Measures of uncertainty in expert systems. *Artificial intelligence*, 83(1):1–58. (Cité à la page 52.)
- WALLEY, P. et FINE, T. L. (1982). Towards a frequentist theory of upper and lower probability. *The Annals of Statistics*, 10(3):741–761. (Cité à la page 49.)
- WANG, T. et BOUTILIER, C. (2003). Incremental utility elicitation with the minimax regret decision criterion. In *Ijcai*, volume 3, pages 309–316. (Cité aux pages 2, 38, and 105.)
- WHITE, C. C., SAGE, A. P. et DOZONO, S. (1984). A model of multiattribute decisionmaking and trade-off weight determination under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-14(2):223–229. (Cité à la page 37.)
- WU, G., ZHANG, J. et GONZALEZ, R. (2004). Decision under risk. *Blackwell handbook of judgment and decision making*, pages 399–423. (Cité à la page 23.)
- YAGER, R. R. (1988). On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on systems, Man, and Cybernetics*, 18(1):183–190. (Cité aux pages 2 and 15.)
- ZADEH, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353. (Cité à la page 51.)

ZADEH, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 1(1):3–28. (Cité aux pages 49 and 50.)

ZIONTS, S. et WALLENIS, J. (1976). An interactive programming method for solving the multiple criteria problem. *Management science*, 22(6):652–663. (Cité aux pages 1 and 36.)

Liste des Figures

1.1	Exemple d'un front de Pareto et d'une alternative dominée	10
1.2	Score de \mathbb{X} pour une somme pondérée en fonction de $\omega \in \Omega$	14
1.3	Score de \mathbb{X} pour une somme pondérée ordonnée en fonction de $\omega \in \Omega$	16
1.4	Lien entre les opérateurs d'agrégation	23
1.5	Illustration de l'approche Maximax	24
1.6	Illustration de l'approche Maximin	25
1.7	Illustration de l'approche du regret Minimax : visualisation du PMR	28
1.8	Illustration de l'approche du regret Minimax : visualisation du MR	29
1.9	Encadrement de la valeur ω^*	33
1.10	Encadrement erroné de la valeur ω^*	35
1.11	Score des alternatives et MR pour $\omega \in \Omega$	41
1.12	Score des alternatives et MR pour $\omega \in E_1$	43
1.13	Liens entre différents modèles des probabilités imprécises	50
1.14	Illustration d'une intersection et d'une union d'informations quand les informations sont cohérentes	56
1.15	Illustration d'une intersection et d'une union d'informations quand les informations sont incohérentes	57
1.16	MCS de sources incohérentes	59
1.17	T-normes usuelles	62
1.18	Comparaison des T-normes usuelles	63
1.19	T-conormes usuelles	65
1.20	Comparaison des T-conormes usuelles	66
2.1	Distribution de probabilité $p(\omega)$ de l'Exemple 23	72

2.2	Une mauvaise réponse amène à une mauvaise estimation de ω^* dans l'Exemple 24	74
2.3	Exemple d'une coupe-alpha avec $\alpha = 0.5$	76
2.4	Information préférentielle $\pi_{(E_1, \alpha_1)}(\omega)$ de l'Exemple 26 permettant d'atténuer l'impact de l'erreur de l'Exemple 24	77
2.5	Information préférentielle $\pi_{(E_2, \alpha_2)}(\omega)$ de l'Exemple 27	79
2.6	Exemple d'une distribution de possibilité non normalisée, avec $\text{Inc}(\pi) = 0.7$	80
2.7	Illustration de la partition d'un ensemble Ω en éléments Ω_i	81
2.8	Diminution du mEMR en fonction de α avec T_P	94
2.9	Diminution du mEMR en fonction de α avec T_M	95
2.10	Diminution du mEMR en fonction de α avec T_L	95
2.11	Diminution du mEMR en fonction de la T-norme avec $\alpha = 0.7$	95
2.12	Degré d'incohérence en cas d'erreur du décideur pour différentes valeurs de α et un taux d'erreur de 15% environ	97
2.13	Regret réel en cas d'erreur du décideur pour différentes valeurs de α et un taux d'erreur de 15% environ	98
2.14	Détection d'une fonction d'agrégation de critères erronée	100
3.1	Maximax pour chaque alternative $x \in \mathbb{X}$	107
3.2	MR pour chaque alternative $x \in \mathbb{X}$	107
3.3	Visualisation de l'optimalité de trois alternatives	109
3.4	Score et position avec de mauvaises alternatives et un modèle WSB	113
3.5	Score et position avec de bonnes alternatives et un modèle WSB	113
3.6	Score et position avec de mauvaises alternatives et un modèle WSU	114
3.7	Score et position avec de bonnes alternatives et un modèle WSU	114
3.8	Score et position avec de mauvaises alternatives et un modèle OWAU	115
3.9	Score et position avec de bonnes alternatives et un modèle OWAU	115
3.10	Score et position avec de mauvaises alternatives et un modèle OWAR	116
3.11	Score et position avec de bonnes alternatives et un modèle OWAR	116
4.1	Rappel de la distribution de possibilité π_\cap de l'Exemple 28 et superposition du score de l'alternative GO	124

4.2	Distribution non normalisée π_\cap de l'Exemple 35 et superposition du score de l'alternative GO	128
4.3	Réponses données par le décideur dans l'Exemple 36, la réponse 4 étant fausse	132
4.4	Fusion d'information préférentielle avec incohérence dans l'Exemple 36	132
4.5	Fusion d'information préférentielle dans l'Exemple 36 corrigée avec l'algorithme de réparation ℓ -parmi- k	133
4.6	Fusion d'information préférentielle dans l'Exemple 37 corrigée avec un MCS	135
4.7	Nombre d'erreurs détectées par rapport au nombre réel d'erreurs (scénario optimiste)	140
4.8	Nombre d'erreurs détectées par rapport au nombre réel d'erreurs (scénario uniforme)	141
4.9	Intervalles de confiance sur le regret réel dans le scénario optimiste. Plus le regret réel est faible, meilleure est la recommandation finale.	144
4.10	Intervalles de confiance sur le regret réel dans le scénario intermédiaire. Plus le regret réel est faible, plus la recommandation finale est bonne.	144
4.11	Intervalles de confiance sur le regret réel pour chaque méthode dans le scénario optimiste avec une élicitation par batch et trois règles de décision	148
4.12	Intervalles de confiance sur le regret réel pour chaque méthode dans le scénario intermédiaire avec une élicitation par batch et trois règles de décision	148
4.13	Détection d'une erreur de modèle avec une élicitation incrémentale, le vrai modèle étant une intégrale de Choquet	153
4.14	Détection d'une erreur de modèle avec une élicitation par batch, le vrai modèle étant une intégrale de Choquet	154
4.15	Illustration d'un cas idéal d'erreur de modèle	157
4.16	Illustration d'un cas idéal d'erreurs du décideur	158
A.1	Illustration de notre algorithme pour déterminer les MCSs d'un ensemble d'hyperrectangles alignés sur les axes.	170

A.2	MVAIR et MVABR d'un polytope	173
A.3	Illustration de plusieurs droites respectant l'Équation (A.12)	175
A.4	Approximation de l'ensemble des valeurs possibles de L par un MVABH175	
A.5	Exemple d'un problème de régression linéaire imprécise. La ligne en pointillés correspond à la régression linéaire incluant la valeur aberrante. La ligne continue correspond à la vérité terrain	176
A.6	Approximation par des rectangles des différents \mathcal{L}_{ij} de l'exemple montré sur la Figure A.5	177
A.7	Exemple d'un problème de régression linéaire par morceaux imprécise. La ligne en pointillé correspond à la régression linéaire complète. Les lignes continues correspondent aux vérités terrain par morceaux	177
A.8	Approximation par des rectangles des différents \mathcal{L}_{ij} de l'exemple montré sur la Figure A.7	178

Liste des Tables

1.1	Ensemble d'alternatives \mathbb{X}	8
1.2	Score de \mathbb{X} pour une somme pondérée (0.7, 0.3)	14
1.3	Score de \mathbb{X} pour une somme pondérée ordonnée (0.55, 0.45)	17
1.4	Score de \mathbb{X} pour une intégrale de Choquet de capacité v	19
1.5	Capacité v	19
1.6	Inverse de Möbius de v	21
1.7	Capacité 2-additive v et son inverse de Möbius	22
1.8	Illustration de l'approche Hurwicz	26
1.9	Détermination du mMR initial de l'Exemple 15	40
1.10	PMR après la première question de l'Exemple 16	43
1.11	Récapitulatif des approches d'élicitation et exemples de références	45
1.12	Distribution de possibilité associée aux préférences d'un décideur	53
2.1	Détermination de la distribution de probabilité $p(\omega)$ de l'Exemple 23	72
2.2	EPMR de l'information préférentielle $\pi_{(E_1, \alpha_1)}$ de l'Exemple 26	84
2.3	EMR de l'information préférentielle $\pi_{(E_1, \alpha_1)}$ de l'Exemple 26	85
2.4	Calcul de Λ_p^n pour différentes valeurs de n et p , en comparaison avec le nombre 2^n	91
2.5	Intervalle de confiance sur le regret réel en cas d'erreur du décideur pour différentes valeurs de α et un taux d'erreur de 15% environ	99
3.1	Détermination du Maximax maximal initial de l'Exemple 15	107
3.2	Score réel moyen de l'alternative recommandée après 5 et 10 questions selon différents contextes	117

3.3	Position réelle médiane de l'alternative recommandée après 5 et 10 questions selon différents contextes	117
3.4	Intérêt de l'approche basée sur un gain (Maximax) par rapport à une approche basée sur un regret (Minimax) en termes de qualité des recommandations (+ : plus intéressante, - : moins intéressante)	118
4.1	Différence de regret réel entre l'ignorance du conflit et les méthodes de fusion d'information, une valeur positive signifiant que les méthodes de fusion ont un regret réel plus faible en moyenne	146
4.2	Statistiques sur le regret réel pour des erreurs de modèle (WS supposé, Choquet réel) lors d'une élicitation PCSS	155
4.3	Statistiques le mEMR pour des erreurs de modèle (WS supposé, Choquet réel) lors d'une élicitation PCSS	155
4.4	Statistiques sur le regret réel pour des erreurs de modèle (WS supposé, OWA réel) lors d'une élicitation PCSS	156
4.5	Statistiques le mEMR pour des erreurs de modèle (WS supposé, OWA réel) lors d'une élicitation PCSS	156

Liste des Algorithmes

1	Élicitation incrémentale fondée sur le critère du regret Minimax . . .	38
2	Élicitation incrémentale fondée sur l'heuristique CSS	42
3	Élicitation incrémentale possibiliste avec l'heuristique CSS (PCSS) .	86
4	Algorithme pour trouver ℓ^*	131