



HAL
open science

Une approche constructive de la décision multicritère

Michel Grabisch

► **To cite this version:**

Michel Grabisch. Une approche constructive de la décision multicritère. *Traitement du Signal*, 2005, 22 (4), pp.321-337. halshs-00268356

HAL Id: halshs-00268356

<https://shs.hal.science/halshs-00268356>

Submitted on 31 Mar 2008

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.

Une approche constructive de la décision multicritère

Michel GRABISCH

Université Paris I - Panthéon-Sorbonne; LIP6

8, rue du Capitaine Scott, 75015 Paris, France

email `Michel.Grabisch@lip6.fr`

1 Introduction

La fusion de données, au sens où on l'entend habituellement, consiste à combiner plusieurs représentations du monde, souvent partielles et entachées d'erreur, fournies par des sources ou systèmes indépendants, afin d'obtenir une nouvelle vision du monde, plus précise et raffinée¹. Dans ce cas, il s'agit grosso modo de combiner les distributions d'incertitude sur les différents attributs décrivant le monde en question. On peut cependant avoir un point de vue plus large et considérer la fusion de représentations logiques, qui décrivent ou restreignent les états possibles du monde, ou la fusion des préférences, c'est-à-dire se préoccuper non pas du monde tel qu'il est mais tel qu'on voudrait qu'il soit [1]. C'est sur ce dernier aspect que nous allons consacrer cet article, en considérant que les préférences sur un même ensemble d'objets proviennent de différents critères ou points de vue. Les préférences, une fois combinées, devant déboucher sur un choix ou un classement des objets en question, on parle de *décision en présence de critères multiples* ou plus simplement de *décision multicritère*.

Nous allons tenter de donner une vision générale de la décision multicritère, sans prétendre à l'exhaustivité, mais que nous allons essayer de construire pas à pas, en tentant de garder un cheminement logique, où les concepts ne seront introduits que par nécessité organique. La taille d'un article ne nous permettra pas de détailler beaucoup les concepts

¹Pour une présentation générale de la fusion de données, voir par exemple les deux numéros spéciaux de *Int. J. of Intelligent Systems* Vol.16 (10,11), en particulier le chapitre introductif [1]

et les approches, ce que nous compenserons par des références bibliographiques. Comme références générales sur la décision multicritère, citons les ouvrages de Roubens et Vincke [43], de Roy et Bouyssou [47], et de Pomerol et Barba-Romero [40]. On pourra également consulter le chapitre de Grabisch et Perny [25] et Grabisch et Labreuche [22].

2 Qu'est ce que la décision multicritère ?

Une définition simple pourrait être la suivante :

Décision multicritère : action de synthétiser des informations relatives à différents *points de vue* ou *aspects* concernant un ensemble d'objets, et choisir un ou plusieurs objets parmi cet ensemble.

On peut distinguer deux contextes d'applications très différents :

- l'*aide à la décision*, où une personne (le *décideur*) tente de prendre une décision en dialoguant avec un système, le dialogue permettant de construire peu à peu un modèle de décision censé refléter les préférences et la stratégie de décision du décideur. Le temps mis pour construire ce modèle est très variable (de l'ordre de l'heure à plusieurs jours). En dernier ressort, c'est le décideur qui prend la décision finale ;
- la *décision automatique*, où cette fois le système prend lui-même la décision, en un temps très court. Il s'agit de situations où un décideur humain ne saurait faire face à la complexité du problème (en particulier le nombre d'objets) et aux contraintes de temps.

Citons quelques exemples qui illustrent ces deux situations :

- (E1) choix d'emplacement d'une centrale, d'une autoroute, d'une ligne de TGV (aide à la décision) ;
- (E2) achat d'un nouveau véhicule, d'un appartement (aide à la décision) ;
- (E3) allocation de ressources, planification, contrôle aérien (décision automatique) ;
- (E4) sélection de candidats à un poste, de projets, évaluation d'élèves.

Détaillons maintenant les différentes composantes d'un problème de décision multicritère. L'*ensemble des objets* désigne selon les cas l'ensemble des solutions potentielles ou l'ensemble des objets (au sens le plus abstrait du terme) d'intérêt pour le décideur

(E1 : emplacements, tracés, E2 : catalogue de voitures, E3 : schémas d'allocation, E4 : projets, etc.), que l'on appelle encore *alternatives* ou *actions*. Cet ensemble d'objet peut être décrit de façon explicite sous forme de liste, et sera alors fini, souvent de petite taille, ou de façon implicite (ensemble de solutions potentielles limité par des contraintes), et il sera en général de grande taille, éventuellement infini, continu.

Chaque objet est décrit par un ensemble d'*attributs* ou *points de vue*, qui sont des caractéristiques mesurables (par un capteur physique ou humain : la caractéristique peut être d'essence qualitative) de l'objet en question (E1 : coût des travaux, impact sur l'environnement, nuisance aux populations, E2 : prix, performances, esthétique, confort, consommation, etc.). On appelle *critère* l'association d'un point de vue et d'une préférence : le décideur exprime une préférence sur les différentes valeurs possibles de l'attribut (E1 : coût faible, peu d'impact sur l'environnement, peu de nuisance aux populations, E2 : bonnes performances, confortable, esthétique sportive, etc.).

On peut distinguer plusieurs types de problèmes de décision :

- *choix* : choisir le ou les meilleurs objets ;
- *classification, tri* : mettre dans des catégories pré-définies (bon, acceptable, à rejeter, etc.) ;
- *rangement* : produire un ordre partiel ou total sur les objets ;
- *scoring* : attribuer un score à chaque objet.

On remarquera que ces types de problèmes sont rangés dans un ordre croissant de difficulté, et que solutionner l'un deux revient à pouvoir solutionner ceux qui se trouvent placés au-dessus.

La méthodologie générale pour résoudre un problème de décision multicritère combine une opération de *comparaison* et d'*agrégation* (ou combinaison, fusion). L'ordre dans lequel sont faites ces opérations va déterminer deux grands types d'approche possible pour la décision multicritère [25], que nous allons détailler dans la section suivante.

3 Les deux approches principales

Introduisons quelques notations. $X = \{a, b, \dots\}$ désigne l'ensemble des objets ou alternatives, et \succeq la *relation de préférence* sur X , une relation binaire le plus souvent,

supposée réflexive. On utilisera soit une notation préfixe $\succeq(a, b)$, soit infixe $a \succeq b$ pour signifier que a est préféré au sens large à b . La notation préfixe est utilisée pour signifier que \succeq est considéré comme une fonction sur X^2 . Si la relation est binaire, alors \succeq prend les valeurs 1 ou 0, et si la relation est *floue* ou *valuée* [13], elle prend ses valeurs dans $[0, 1]$. On notera \sim la partie symétrique de la relation \succeq , c'est-à-dire $a \sim b$ ssi $a \succeq b$ et $b \succeq a$.

On note $N := \{1, \dots, n\}$ l'ensemble d'indices des critères. On appelle *fonction de scorage* ou *fonction d'utilité* sur le critère j une fonction $g_j : X \mapsto \mathbb{R}$, pour $j \in N$. La quantité $g_j(a)$ est le score de a selon le critère j , noté a_j .

L'approche *agréger puis comparer* consiste à calculer

$$\succeq(a, b) = \phi(\psi(a_1, \dots, a_n), \psi(b_1, \dots, b_n)) \quad (1)$$

tandis que l'approche *comparer puis agréger* calcule

$$\succeq(a, b) = \psi(\phi_1(a_1, b_1), \dots, \phi_n(a_n, b_n)) \quad (2)$$

$\phi(x, y)$ est une fonction de comparaison, croissante en x et décroissante en y , et $\phi(x, x) = 1$.

Par exemple, pour une relation binaire :

$$\phi(x, y) = \begin{cases} 1 & \text{si } y - x \leq p \\ 0 & \text{sinon.} \end{cases} \quad (3)$$

Si $p > 0$, \sim n'est pas transitive et donne lieu à un *quasi-ordre* [39]. Détaillons maintenant ces deux approches.

3.1 L'approche "agréger puis comparer"

La principale représentante de cette approche est la théorie de l'*utilité multi-attributs* (MAUT) [28]. La représentation des préférences s'écrit sous la forme :

$$a \succeq b \Leftrightarrow \psi(u_1(a_1), \dots, u_n(a_n)) \geq \psi(u_1(b_1), \dots, u_n(b_n)). \quad (4)$$

Cette équation est un cas particulier de (1), où ϕ est la fonction définie en (3) avec $p = 0$. MAUT fait l'hypothèse que \succeq est une relation binaire complète et transitive, hypothèse sans laquelle l'équation ci-dessus ne pourrait être écrite. La conséquence de

cette supposition est que toutes les alternatives sont comparables, une propriété qu'il n'est pas toujours souhaitable d'avoir car l'on sait que pour un décideur humain, il existe des situations comportant des alternatives incomparables.

Il est d'usage d'écrire l'ensemble des alternatives X sous la forme d'un produit cartésien $X := X_1 \times \cdots \times X_n$, où X_i représente l'ensemble des valeurs possibles de l'attribut i . Les fonctions $u_i : X_i \rightarrow \mathbb{R}$ s'appellent les *fonctions d'utilité*, et $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ est la *fonction d'agrégation*.

Une condition importante pour que le modèle (4) ait un sens est que les fonctions d'utilité soient *commensurables*, c'est-à-dire que $u_i(a_i) = u_j(b_j)$ implique que a_i et b_j induisent pour le décideur le même degré de satisfaction. Nous reviendrons par la suite sur cette notion afin de la préciser davantage.

Un des ingrédients de base du modèle est la fonction d'agrégation ψ . En général, on impose que $\min \leq \psi \leq \max$ et que ψ soit croissante selon chaque argument. Ces deux conditions impliquent que $\psi(a, a, \dots, a) = a, \forall a$. Donnons-en quelques exemples.

- *les moyennes pondérées* (arithmétique, géométrique, harmonique, etc.) La forme générale est la suivante :

$$M_f(a_1, \dots, a_n) = f^{-1} \left(\sum_{i=1}^n w_i f(a_i) \right),$$

avec f strictement croissante et continue, $w_i \in [0, 1]$ et $\sum_{i=1}^n w_i = 1$.

- *minimum et maximum pondérés* [12]

$$\begin{aligned} \text{wmin}_w(a_1, \dots, a_n) &= \bigwedge_{i=1}^n [(1 - w_i) \vee a_i] \\ \text{wmax}_w(a_1, \dots, a_n) &= \bigvee_{i=1}^n [w_i \wedge a_i] \end{aligned}$$

où \wedge, \vee désignent min, max respectivement. On a $w_i \in [0, 1]$ et $\bigvee_{i=1}^n w_i = 1$. On retrouve le minimum et le maximum en prenant $w_i = 1, \forall i \in N$.

- *les moyennes pondérées ordonnées (OWA)* [55]

$$\text{OWA}_w(a_1, \dots, a_n) = \sum_{i=1}^n w_i a_{\sigma(i)}$$

avec $w_i \in [0, 1]$ tels que $\sum_{i=1}^n w_i = 1$, et σ est une permutation sur N telle que $a_{\sigma(1)} \leq \cdots \leq a_{\sigma(n)}$. De la sorte, on retrouve comme cas particuliers le minimum, le maximum, la médiane, les statistiques d'ordre, etc.

Pour plus de détails, nous renvoyons le lecteur à [24, 25].

3.2 L’approche “comparer puis agréger”

Les méthodes dites de *surclassement* sont les représentantes principales de cette approche, et plus particulièrement les différentes versions de la méthode ELECTRE, fondée par Bernard Roy [44, 45, 46]. Les méthodes de surclassement sont basées sur le *principe de concordance-discordance* :

a surclasse b si la majorité des critères sont en faveur de a, et s’il n’y a aucun critère fortement en désaccord avec cette affirmation.

Ce principe peut se traduire de la façon générale suivante :

$$\succeq(a, b) = h(\mu(C(a \succeq b)), 1 - \nu(C(b \gg a))) \quad (5)$$

avec $C(a \succeq b)$ l’ensemble des critères en faveur de “a surclasse b”, $C(b \gg a)$ l’ensemble des critères contre l’affirmation “a surclasse b”, μ, ν des mesures d’importance de coalitions de critères, enfin h , une fonction d’agrégation conjonctive (minimum, produit, ou toute norme triangulaire).

Ces méthodes agrègent des préférences et non pas des scores comme dans les méthodes “agréger puis comparer”. Ainsi, tout problème de commensurabilité disparaît car les préférences sont par construction homogènes entre elles. Ces méthodes sont donc particulièrement adaptées aux cas où il est difficile de s’assurer que les scores soient commensurables, au cas où certains critères sont quantitatifs et d’autres qualitatifs, etc.

Une autre caractéristique intéressante est que l’on n’obtient pas nécessairement un ordre total sur les alternatives, au contraire des méthodes du type MAUT. Cela offre une plus grande souplesse de modélisation des préférences.

L’inconvénient de ces méthodes réside dans la phase finale de décision, quand il s’agit, après agrégation des préférences, d’en déduire un ordre partiel ou total sur les alternatives. Ces difficultés s’expliquent par un résultat fondamental que nous allons présenter plus loin. Auparavant, nous présentons brièvement quelques méthodes du type ELECTRE.

La méthode ELECTRE I [44] considère en entrée les informations suivantes :

- un ensemble de n fonctions de scorage, jouant un rôle similaire aux fonctions d’utilité, appelées par Roy *critères*, g_1, \dots, g_n ;

- un ensemble de poids d'importance w_1, \dots, w_n , pour les critères g_1, \dots, g_n ;
- un seuil de veto $v_j(g_j) > 0$ pour chaque critère ;
- un niveau de concordance s .

Pour chaque paire (a, b) , on calcule le degré de concordance de “ a surclasse b ”

$$c(a \succ b) = \frac{1}{W} \sum_{g_j(a) \geq g_j(b)} w_j, \quad W = \sum_j w_j.$$

On dit que $a \succ b$ (a surclasse b) si $c(a \succ b) \geq s$ (concordance) et si $\forall j$ tel que $g_j(a) < g_j(b)$, on a $g_j(b) - g_j(a) < v_j(g_j)$ (pas de discordance). La décision se fait de la façon suivante : on cherche un ensemble $K \subseteq X$ tel que $\forall b \in X \setminus K$, il existe $a \in K, a \succ b$, et $\forall a, b \in K, a \not\succeq b$. K est appelé le *noyau* du graphe (X, \succ) . Il existe et est unique si le graphe ne contient pas de circuit, une condition qu'il est cependant impossible de garantir.

ELECTRE III [46] est une version floue d'ELECTRE I, au sens où sur chaque critère, le degré de concordance $c_j(a \succ b)$ est valué sur $[0, 1]$. De la même façon, un seuil flou de veto $D_j(a \succ b)$ est défini, comme par exemple sur la Figure 1. La procédure de décision

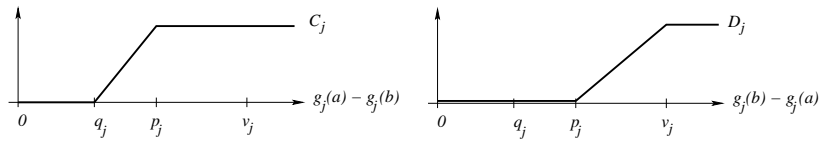


FIG. 1 – Exemple de seuil flou pour la méthode ELECTRE III.

est compliquée, et nous ne la détaillons pas ici.

ELECTRE TRI (Roy, Bouyssou 1993) est une variante d'ELECTRE dont le but est de faire du tri dans k catégories pré-définies et ordonnées C_1, \dots, C_k . La catégorie j est définie par des seuils g_1^j, \dots, g_n^j sur chaque critère, et $g_i^j > g_i^{j-1}$ pour tout i, j . On assigne l'alternative a à l'une des catégories en calculant une relation de surclassement sur l'ensemble d'alternatives $\{a, C_1, \dots, C_k\}$.

3.3 Agrégation des préférences : le théorème d'Arrow

Agréger des préférences selon plusieurs critères est équivalent à un problème de vote, où les alternatives jouent le rôle de candidats et les critères celui de votants. Chaque

votant exprime sa préférence par un préordre sur l'ensemble des candidats. Le problème est alors le suivant : *Comment choisir le meilleur candidat en prenant en compte toutes les préférences exprimées par les votants, tout en respectant un certain nombre de règles rationnelles (par exemple, pas de dictateur) ?*

Répondre à cette question est l'objet de la théorie du choix social. Beaucoup de méthodes ont été proposées dès l'époque de la révolution française, par exemple par Borda et le Marquis de Condorcet, pour ne citer que les plus connues. Cependant aucune de ces méthodes ne s'est avérée satisfaisante, car il était toujours possible de trouver un exemple où certains critères rationnels n'étaient pas satisfaits (on pourra lire avec intérêt une synthèse des principaux paradoxes de vote dans [27]). En 1951, l'économiste Kenneth Arrow a apporté une réponse pour le moins surprenante à ces questions. En supposant que chaque votant exprime ses préférences sous forme d'un préordre total (réflexif, complet, transitif) sur l'ensemble des candidats, il a montré que si l'on imposait comme conditions de rationalité les quatre suivantes :

1. Le résultat de l'agrégation des préférences doit être de nouveau un préordre total ;
2. Il ne doit pas y avoir de dictateur, c'est-à-dire un votant qui imposerait systématiquement son choix ;
3. Si $a \succ b$ pour tous les votants, alors $a \succ b$ doit être dans la préférence agrégée (unanimité) ;
4. La préférence de a sur b ne doit pas dépendre des préférences sur c, d, \dots (indépendance),

alors, à partir de 3 votants, *il n'existe aucune procédure de vote satisfaisant toutes ces conditions* [2]. Autrement dit, toute procédure de vote viole nécessairement l'une de ces quatre conditions.

Ce résultat explique les difficultés rencontrées lors de la phase de décision des méthodes "comparer puis agréger", puisque précisément on commence par construire pour chaque critère une comparaison entre alternatives, que l'on tente ensuite d'agréger pour obtenir un préordre global.

Dans la suite, nous ne développons pas plus avant l'approche "comparer puis agréger", et tentons de bâtir une méthodologie complète sur la base de MAUT.

4 Une méthodologie multicritère basée sur MAUT

4.1 Échelles et fonctions d'utilité

La théorie du mesurage [29, 42] permet de définir rigoureusement la notion d'échelle de mesure, et plus généralement l'acte de mesurer une grandeur. Soit (A, \succeq) un ensemble d'objets munis d'une relation binaire. Une *échelle de mesure (numérique)* est un homomorphisme u de (A, \succeq) dans (\mathbb{R}, \geq) . On appelle la fonction $u : A \rightarrow \mathbb{R}$ une *fonction d'utilité*.

En psychologie, il a été montré (voir par exemple les travaux de Slovic [51]) que notre façon de juger, évaluer et prendre des décisions est guidée par ce qu'on appelle l'*affect*. Ce mot désigne la qualité spécifique de «bon» ou «mauvais», telle qu'elle est ressentie consciemment ou non, et délimitant des zones de stimuli de qualités positive et négative.

De cela, on peut retenir deux conséquences essentielles :

- dans une évaluation ou choix multi-attributs, les attributs dont la perception est peu précise ou sans point de repère ont peu d'importance dans la décision finale. Dans le même ordre d'idée, exprimer un attribut sous forme de proportion ou de pourcentage a plus d'impact que si on l'exprime de manière absolue. La raison est encore une fois l'absence de repère dans le cas de l'expression sous forme absolue. C'est le phénomène de la *dominance de la proportion* ;
- le caractère *bipolaire* de l'affect, c'est-à-dire construit sur deux pôles opposés (bon/mauvais, positif/négatif), est central dans l'acte de décision, et il importe de pouvoir le représenter correctement.

La représentation utilisant une échelle bipolaire, c'est-à-dire un seul axe pour coder l'affect allant du négatif au positif, a été le courant dominant en recherche (voir les travaux d'Osgood *et col.* en 1957 [37]). Récemment, Cacioppo *et col.* ont proposé l'usage de deux échelles unipolaires séparées, l'une pour la partie positive de l'affect, l'autre pour la partie négative [7]. La motivation pour une telle approche est que l'on peut très bien ressentir pour le même objet à la fois un sentiment positif *et* un sentiment négatif, sans qu'il soit possible de les fondre en un seul sentiment *résultant* (par exemple, manger du chocolat procure un plaisir gustatif, mais on peut se sentir en même temps coupable de gourmandise). Une étude récente de Peters et Slovic [38] ayant pour but de comparer les

deux paradigmes, n'a cependant pas permis de trancher de façon claire.

Nous basons notre approche sur l'utilisation d'une échelle bipolaire. Donnons-en une définition plus formelle, ce qui nous permettra d'introduire des niveaux remarquables dans les échelles, qui formeront la base de notre construction. Considérons A comme étant un ensemble de «niveaux» ou valeurs possibles prises par un attribut. Il peut exister dans A une valeur particulière e , dite *niveau neutre*, qui a la propriété que si $a \succ e$, alors a est jugé «bon», tandis que si $e \succ a$, alors a est jugé «mauvais». On pose par commodité $u(e) = 0$; ce faisant, les nombres positifs correspondent aux bonnes valeurs de l'attribut, et les nombres négatifs aux mauvaises valeurs.

Un niveau neutre existe chaque fois que la relation \succsim correspond à deux notions appariées et opposées liées à un affect. Ainsi en est-il des relations «plus attractif que», «meilleur que», «aimer plus que», dont les paires opposées correspondantes sont respectivement attirance/répulsion, bon/mauvais, aimer/détester. Au contraire, des relations comme «plus prioritaire que», «plus permis que», «appartient plus à la catégorie C que» ne tombent pas dans cette catégorie, et donc n'ont pas de niveau neutre.

Une *échelle bipolaire* a un niveau neutre, tandis qu'une *échelle unipolaire* n'en a pas. Pour les échelles bipolaires, le domaine image de u est \mathbb{R} (*bipolaire non borné*) ou un intervalle symétrique de \mathbb{R} (*bipolaire borné*).

En général, une échelle unipolaire a un plus petit élément a , c'est-à-dire $b \succeq a$ pour tout $b \in A$. On pose $u(a) := 0$, et on adopte une notation commune $\mathbf{0}$ pour le niveau neutre des échelles bipolaires et le plus petit élément des échelles unipolaires.

Une échelle a un plus grand élément s'il existe $a_1 \in A$ tel que $a_1 \succeq a$ pour tout $a \in A$. Une échelle unipolaire ayant un plus grand élément est dite *bornée*. Une échelle bipolaire est bornée si elle a un plus grand et un plus petit élément. Si nous reprenons les exemples d'échelles unipolaires ci-dessus, les relations «plus permis que» et «appartient plus à la catégorie C que» ont un plus petit élément, que l'on peut exprimer respectivement par «interdit» (au sens strict), et «hors de la catégorie C ». Par contre, il n'existe pas de plus petit élément pour «plus prioritaire que» car on peut toujours trouver quelque chose de moins prioritaire. Nous notons $\mathbf{1}$ le plus grand élément, et $-\mathbf{1}$ le plus petit élément d'une échelle bipolaire, quand ils existent.

4.2 Commensurabilité des échelles ; la méthode MACBETH

Ainsi qu'on l'a expliqué dans la section 3.1, le problème crucial de l'approche MAUT est de pouvoir assurer la commensurabilité des échelles u_1, \dots, u_n . Les niveaux remarquables $\mathbf{0}$, $\mathbf{1}$ introduits ci-dessus vont nous aider à résoudre ce problème, par l'intermédiaire de la méthode MACBETH. Celle-ci a été proposée par Vansnick et Bana e Costa en 1994 [3, 4, 5], dans le but de construire des échelles numériques à partir d'informations ordinales, et pour un problème multicritère, de construire les poids d'importance sur les critères afin de les utiliser dans une somme pondérée.

Supposons qu'il soit possible de construire une *échelle d'intervalle* u_i sur chaque X_i , c'est-à-dire une échelle définie à une transformation affine positive près :

$$u_i \longrightarrow \alpha_i u_i + \beta_i, \quad \alpha_i > 0, \quad \beta_i \in \mathbb{R}.$$

Il suffit alors de déterminer sur chaque échelle deux points ayant un sens absolu pour le décideur pour fixer toutes les échelles les unes par rapport aux autres, et ainsi résoudre le problème de la commensurabilité. Nous choisissons pour cela sur chaque X_i l'élément $\mathbf{0}_i$, qui est le niveau neutre (cas bipolaire) ou le plus petit élément (cas unipolaire) de X_i , et l'élément $\mathbf{1}_i$ si l'échelle est bornée. De la sorte, $u_i(\mathbf{0}_i) = 0$ et $u_i(\mathbf{1}_i) = 1$, $i = 1, \dots, n$.

Que faire du cas où les échelles ne sont pas bornées ? Il faut recourir alors au *niveau satisfaisant*, que l'on notera également $\mathbf{1}$. MACBETH, qui utilise ce niveau, le définit comme suit : *le niveau satisfaisant est considéré comme bon et tout à fait acceptable si le décideur pouvait l'obtenir, même si des éléments plus attractifs peuvent exister*. L'existence d'un tel niveau, sur laquelle repose la méthode MACBETH, et par là-même notre construction, est attestée par au moins deux théories :

- la théorie de la *rationalité bornée (satisficing bounded rationality)* de l'économiste Herbert Simon [49, 50, 41] : dans une situation réelle, donc complexe par essence (par ex. : jeu d'échecs) et souvent en information incomplète (par ex. : recherche de nourriture par un animal), le décideur ou l'agent ne cherche pas à *optimiser*, mais à *satisfaire* : celui-ci choisira toute solution qui lui procure un niveau de satisfaction jugé suffisant.
- la nécessité pour le décideur d'avoir des niveaux de référence (Slovic), ainsi qu'on l'a expliqué ci-dessus. Dans le même ordre d'idée, le phénomène de dominance

de la proportion va dans le même sens : exprimer une valeur en pourcentage ou en proportion est bien faire référence à un plus grand élément (ou en tout état de cause un niveau remarquable considéré comme standard), auquel on compare toutes les valeurs.

Ce niveau satisfaisant $\mathbf{1}_i$ étant déterminé sur chaque X_i , on pose $u_i(\mathbf{1}_i) = 1$.

Ces niveaux neutre et satisfaisant ayant été déterminés sur chaque X_i , la méthode MACBETH peut s'appliquer. La *procédure d'interrogation de base* est la suivante. Soit $A = \{a, b, c, \dots\}$ un ensemble fini d'objets ou alternatives. Pour chaque paire $(a, b) \in A^2$, on pose au décideur les questions suivantes :

- ❶ *Est-ce que a est plus attractif que b ? (oui/non)*
- ❷ *Si oui, est-ce que la différence d'attractivité est très faible, faible, modérée, forte, très forte, extrême ? (choisir seulement une catégorie)*

Sur la base des réponses, MACBETH détecte les éventuelles incohérences, et après correction, construit une échelle d'intervalle (non unique).

Cette procédure générale est utilisée pour construire les échelles u_i de façon à ce qu'elles soient commensurables. Pour chaque X_i , on construit l'ensemble d'alternatives fictives $A_i := \{(\mathbf{0}_1, \dots, \mathbf{0}_{i-1}, x_i^j, \mathbf{0}_{i+1}, \dots, \mathbf{0}_n) \mid x_i^j \in X_i\}$, et on applique la procédure d'interrogation sur cet ensemble. La commensurabilité est assurée en posant $u_i(\mathbf{1}_i) = 1$, $u_i(\mathbf{0}_i) = 0$ pour tout i , ce qui fixe les échelles.

MACBETH considère que ψ est la somme pondérée $\sum w_i a_i$, et il reste donc à déterminer les poids w_i . Il est facile de voir que le poids w_i est lié au score de l'alternative fictive $(\mathbf{1}_i, \mathbf{0}_{ic}) := (\mathbf{0}_1, \dots, \mathbf{0}_{i-1}, \mathbf{1}_i, \mathbf{0}_{i+1}, \dots, \mathbf{0}_n)$. En effet

$$\psi(\mathbf{1}_i, \mathbf{0}_{ic}) = 0.w_1 + \dots + 0.w_{i-1} + 1.w_i + 0.w_{i+1} + \dots = w_i.$$

En utilisant la procédure d'interrogation de MACBETH, il suffit alors de construire une échelle d'intervalle pour «scorer» les alternatives $(\mathbf{1}_i, \mathbf{0}_{ic})$, $i = 1, \dots, n$, et de la fixer en imposant $\sum_{i=1}^n w_i = 1$, $w_i \geq 0$.

4.3 Les limitations de la somme pondérée

Bien qu'utilisée dans la majorité des méthodes, la somme pondérée est en fait une fonction d'agrégation aux possibilités très limitées, qui ne peut convenir à modéliser les

préférences d'un décideur d'une façon un tant soit peu générale. Prenons ici un exemple simple avec 2 critères afin d'illustrer cela.

Soit $X = X_1 \times X_2$, et supposons les fonctions d'utilité u_1, u_2 connues. On considère 3 alternatives a, b, c , dont les utilités sont :

$$\begin{aligned} u_1(a_1) = 0.4 & \quad u_1(b_1) = 0 & \quad u_1(c_1) = 1 \\ u_2(a_2) = 0.4 & \quad u_2(b_2) = 1 & \quad u_2(c_2) = 0. \end{aligned}$$

Supposons que le décideur ne veule pas d'alternatives comportant des critères non satisfaits, si bien que ses préférences sont $a \succ b \sim c$. Ces préférences induisent les relations suivantes sur les poids d'importance w_1, w_2 :

$$\begin{aligned} b \sim c & \Leftrightarrow w_1 = w_2 \\ a \succ b & \Leftrightarrow 0.4(w_1 + w_2) > w_2, \end{aligned}$$

équivalent à $0.8w_2 > w_2$, ce qui est impossible. La somme pondérée ne permet pas de représenter ces préférences, pourtant naturelles. Nous allons proposer dans ce qui suit des modèles plus généraux, mais qui s'inscrivent dans la continuité de la construction générale que nous venons de faire.

5 Vers des modèles plus généraux

5.1 Notion de capacité

Notons comme précédemment $N := \{1, \dots, n\}$, l'ensemble d'indices des critères. L'idée de base est de généraliser la procédure d'interrogation du décideur qui détermine les poids des critères, en demandant au décideur ses préférences sur l'ensemble des alternatives dites *binaires*, c'est-à-dire du type $(\mathbf{1}_A, \mathbf{0}_{A^c})$, $A \subseteq N$. La notation $(\mathbf{1}_A, \mathbf{0}_{A^c})$ signifie que l'alternative vaut $\mathbf{1}_i$ pour les critères $i \in A$, et $\mathbf{0}_i$ sinon. De la sorte, on construit une échelle d'intervalle $\mu : \mathcal{P}(N) \longrightarrow \mathbb{R}$, attribuant un score à chaque alternative binaire. On fixe cette échelle d'intervalle en posant

$$\mu(\emptyset) = 0, \quad \mu(N) = 1. \tag{6}$$

D'autre part, si $A \subseteq B$, on dit que l'alternative $(\mathbf{1}_B, \mathbf{0}_{B^c})$ *domine* l'alternative $(\mathbf{1}_A, \mathbf{0}_{A^c})$, au sens où sur chaque critère, la première est au moins aussi bonne que la deuxième. Il

est donc naturel d'imposer que le score obtenu par la première soit au moins aussi bon que celui de la deuxième, ce qui donne la relation :

$$A \subseteq B \Rightarrow \mu(A) \leq \mu(B). \quad (7)$$

On appelle *capacité* [8] ou *mesure floue* [52] toute fonction $\mu : \mathcal{P}(N) \longrightarrow [0, 1]$ vérifiant les deux conditions (6) et (7), où $\mathcal{P}(N) = 2^N$ désigne l'ensemble des parties de N .

D'un point de vue des ensembles ordonnés, une capacité est simplement une fonction du treillis $(\mathcal{P}(N), \subseteq)$ des parties de N dans le treillis linéaire $([0, 1], \leq)$, *isotone*, c'est-à-dire préservant la relation d'ordre (Eq. (7)), et les plus grand et plus petit éléments (Eq. (6)). Il est par conséquent commode de représenter une capacité sous la forme du treillis des parties de N (Fig. 2).

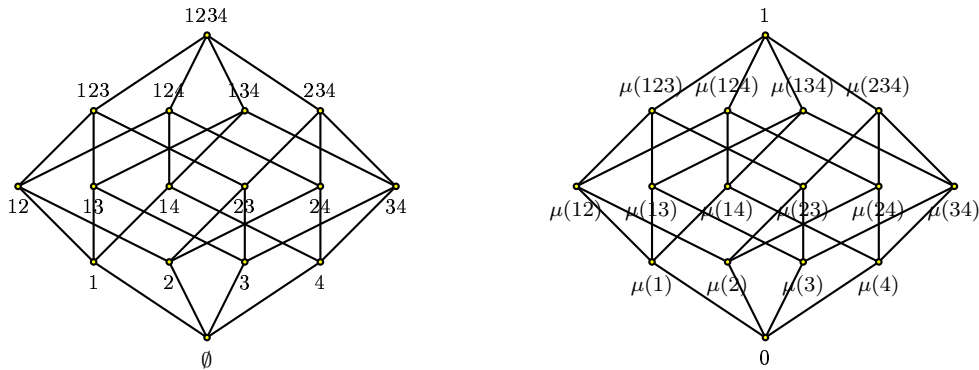


FIG. 2 – Treillis $(\mathcal{P}(N), \subseteq)$ (gauche), et treillis des valeurs de μ (droite).

5.2 Détermination de la fonction ψ

Étant donné que $u_i(\mathbf{1}_i) = 1$, $u_i(\mathbf{0}_i) = 0$, $\forall i \in N$, on en déduit :

$$\psi(\mathbf{1}_A, \mathbf{0}_{A^c}) = \mu(A), \quad \forall A \subseteq N. \quad (8)$$

Ainsi μ détermine ψ sur les sommets de l'hypercube $[0, 1]^n$. On peut donc voir le problème de la détermination de ψ comme un problème d'interpolation entre les sommets, où la fonction est connue. Ce nombre de sommets étant exponentiellement élevé, il convient de chercher une formule d'interpolation la plus simple possible.

Cherchons une interpolation linéaire qui utilise le moins de sommets possible pour chaque point. Cela revient à partitionner $[0, 1]^n$ en polyèdres $\mathcal{A}_1, \dots, \mathcal{A}_q$, dont les sommets

sont pris parmi les sommets de l'hypercube $[0, 1]^n$. Pour tout point $x \in [0, 1]^n$, on note $\mathcal{A}(x)$ le polyèdre qui le contient. On cherche une formule d'interpolation linéaire en x :

$$\psi(x) = \sum_{A \in \mathcal{A}(x)} \left[\sum_{i=1}^n \alpha_i(A) x_i \right] \psi(\mathbf{1}_A, \mathbf{0}_{A^c}), \quad (9)$$

avec $\alpha_i(A) \in \mathbb{R}$, $i = 1, \dots, n$, $\forall A$ sommet de $\mathcal{A}(x)$.

On peut montrer qu'il y a une unique solution à ce problème, avec des polyèdres contenant le moins de sommets possible, décrite comme suit [19] :

- $q = n!$
- chaque polyèdre est défini par une permutation σ sur N :

$$\mathcal{A}_\sigma = \{x \in [0, 1]^n \mid x_{\sigma(1)} \leq x_{\sigma(2)} \leq \dots \leq x_{\sigma(n)}\}$$

- chaque \mathcal{A}_σ contient n sommets + l'origine $(0, 0, \dots, 0)$;
- pour tout $x \in \mathcal{A}_\sigma$, la formule d'interpolation est :

$$\psi(x) = \sum_{i=1}^n [x_{\sigma(i)} - x_{\sigma(i-1)}] \mu(\{\sigma(i), \dots, \sigma(n)\}). \quad (10)$$

La formule (10) est en fait l'*intégrale de Choquet* de x (considéré comme une fonction sur N), par rapport à μ , que l'on notera $\mathcal{C}_\mu(x)$ [8, 35]. La figure 3 illustre l'interpolation dans le cas $n = 2$, où la formule d'interpolation devient, pour $x_1 \leq x_2$:

$$\begin{aligned} \psi(x) &= x_1 \psi(\mathbf{1}, \mathbf{1}) + (x_2 - x_1) \psi(\mathbf{0}, \mathbf{1}) \\ &= x_1 + (x_2 - x_1) \mu(\{x_2\}). \end{aligned}$$

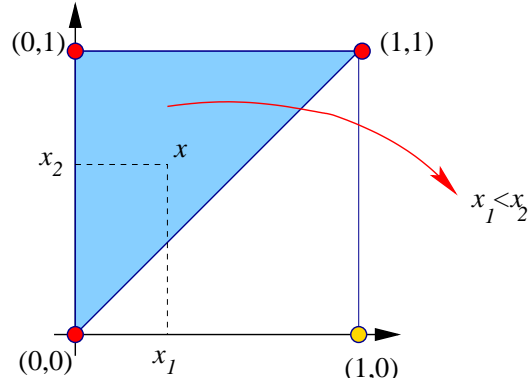
En résumé, nous avons trouvé le modèle le plus simple possible, prenant en compte les préférences du décideur sur toutes les alternatives binaires :

- pour tout $A \subseteq N$, $\mu(A)$ est le score de l'alternative binaire $(\mathbf{1}_A, \mathbf{0}_{A^c})$;
- le score d'une alternative $x \in X$ est calculé par l'intégrale de Choquet :

$$u(x) = \mathcal{C}_\mu(u_1(x_1), \dots, u_n(x_n)),$$

avec, pour tout $a \in [0, 1]^n$ (en fait cette formule s'étend à tout $a \in \mathbb{R}^n$) et σ tel que $a_{\sigma(1)} \leq \dots \leq a_{\sigma(n)}$:

$$\mathcal{C}_\mu(a) := \sum_{i=1}^n [a_{\sigma(i)} - a_{\sigma(i-1)}] \mu(\{\sigma(i), \dots, \sigma(n)\}). \quad (11)$$

FIG. 3 – Interpolation linéaire : cas $n = 2$.

Si nous revenons à notre exemple précédent (Section 4.3), il est maintenant facile de le résoudre. Les contraintes imposées par les préférences deviennent :

$$b \sim c \Leftrightarrow \mu(1) = \mu(2)$$

$$a \succ b \Leftrightarrow 0.4\mu(1, 2) > \mu(1),$$

soit $\mu(1) < 0.4$. Il suffit de prendre par exemple $\mu(1) = \mu(2) = 0.3$.

On trouvera dans [23, 30] une axiomatisation de ce modèle.

5.3 L'intégrale de Choquet vs. les fonctions d'agrégation classiques

L'intégrale de Choquet contient comme cas particuliers bon nombre des exemples de fonctions d'agrégation classiques que nous avons donné dans la section 3.1. La somme pondérée $\sum_i w_i a_i$ coïncide avec l'intégrale de Choquet quand la capacité est *additive*, c'est-à-dire vérifiant $\mu(A) = \sum_{i \in A} \mu(\{i\})$ pour tout $A \subseteq N$. La correspondance est :

$$w_i = \mu(\{i\}), \quad i = 1, \dots, n.$$

Les sommes pondérées ordonnées (OWA) correspondent à l'intégrale de Choquet quand la capacité est *symétrique*, c'est-à-dire vérifiant $\mu(A) = \mu(B)$ si $|A| = |B|$. La correspondance s'écrit, pour $i = 1, \dots, n$:

$$w_i = \mu(\{1, \dots, n - i + 1\}) - \mu(\{1, \dots, n - i\}); \quad \mu(\{1, \dots, i\}) = \sum_{j=0}^{i-1} w_{n-j}.$$

Par conséquent, le minimum, le maximum, la médiane, et les statistiques d'ordre sont couverts par l'intégrale de Choquet. Par exemple, la capacité correspondant au minimum est définie par $\mu_{\min}(A) = 0$ pour tout $A \neq N$.

Nous renvoyons le lecteur à [24, 25, 31, 36] pour plus de détails sur l'intégrale de Choquet vue comme une fonction d'agrégation.

5.4 Interprétation du modèle : importance et interaction entre critères

Le modèle basé sur l'intégrale de Choquet et les capacités que nous venons d'introduire semble très riche par rapport à celui de la somme pondérée. Le prix à payer cependant est élevé, car le modèle comporte $2^n - 2$ paramètres à déterminer, qui sont les valeurs de la capacité sur toutes les parties de N (en fait les scores de toutes les alternatives binaires), sauf \emptyset et N lui-même. Par exemple, avec 5 critères, on arrive déjà à 30 paramètres, contre 4 pour une somme pondérée.

Cette complexité exponentielle peut être prohibitive, et deux questions essentielles viennent à l'esprit :

1. Existe-t-il un moyen d'interpréter la capacité μ afin de faciliter son élicitation ?
2. Existe-t-il un moyen de simplifier le modèle sans trop porter atteinte à sa généralité, en se situant quelque part entre l'intégrale de Choquet et la somme pondérée, ou la somme pondérée ordonnée ?

Nous allons voir que la réponse à ces deux questions est positive. Dans cette section, nous abordons la première, en introduisant les concepts d'importance de critère et d'interaction entre critères. La deuxième question sera abordée dans la section suivante.

Afin d'interpréter notre modèle, une question naturelle est la suivante :

Quelle est l'importance du critère i dans un modèle donné ?

La réponse est en fait moins simple à donner qu'il n'y paraît, car en fait elle dépend de la fonction d'agrégation choisie. Pour la somme pondérée, il semble naturel de prendre w_i comme poids d'importance du critère i . Par contre, w_i n'est manifestement pas le poids d'importance du critère i dans une somme pondérée ordonnée, puisque w_i ne s'applique non pas sur le critère i mais sur le i ème critère moins bien satisfait.

En considérant les critères comme des joueurs dans un jeu coopératif, nous empruntons au domaine de la théorie des jeux la notion d'*indice de pouvoir* ou de *valeur*. Cette idée semble avoir été donnée pour la première fois par Murofushi [33]. On considère donc un ensemble de joueurs N , l'ensemble des *coalitions* de joueurs $\mathcal{P}(N)$, et une fonction $v : \mathcal{P}(N) \rightarrow \mathbb{R}$ représentant la somme d'argent ou le bénéfice (au sens abstrait) que recevra chaque coalition si le jeu est joué. L'exemple suivant, emprunté à I. Dragan [11], illustre bien cette situation.

EXEMPLE : LE PROBLÈME DES 3 «LADIES»

Trois dignes ladies Agatha, Betty et Cathy (que nous nommerons A, B, C) sont des habituées du théâtre où elles aiment à se rendre régulièrement. Le directeur du théâtre, qui les connaît bien, désire leur faire un petit cadeau, et leur écrit à chacune une lettre. Dans la lettre pour lady A, il offre une remise de £10 sur le prochain spectacle. Cependant, si lady A vient avec son amie lady B, elles auront une remise de £40, si elle vient avec son amie lady C, la remise sera de £50, et de £90 si elles viennent toutes les trois.

Il écrit une lettre semblable à lady B, lui offrant les remises suivantes

S	remise $v(S)$
B	£20
A,B	£40
B,C	£60
A,B,C	£90

et une lettre à lady C, avec les remises suivantes

S	remise $v(S)$
C	£30
A,C	£50
B,C	£60
A,B,C	£90

La semaine suivante, lors de leur réunion autour du thé de 5 heures, elles sont très excitées et ne peuvent s'empêcher de parler des lettres qu'elles ont reçues. Après concertation, elles décident d'aller toutes les trois au prochain spectacle, et ainsi de gagner une remise de £90. Passé ce moment d'euphorie,

elles se regardent et s'exclament : mais comment va-t-on partager les £90 entre nous? Facile, dit lady A, chacun reçoit £30 et le tour est joué. Pas du tout, rétorque lady C, car moi en allant toute seule c'est ce que j'aurais eu, alors que vous, Agatha, vous n'auriez eu que £10! La discussion qui s'ensuivit fut peu digne de sujets de sa majesté et faillit mettre en péril leur amitié.

Comment donc aider ces pauvres ladies à résoudre leur problème?

On a ici clairement une situation modélisée par un jeu coopératif v sur $N = \{A, B, C\}$. Le problème est de partager $v(N)$ sur les joueurs de façon équitable, selon leur mérite, c'est-à-dire de définir une fonction $\phi_v : N \rightarrow \mathbb{R}$ telle que $\sum_i \phi_v(i) = v(N)$. Cette fonction est dite *valeur* de v , ou *indice de pouvoir* dans le cas où les joueurs sont des partis politiques (jeux de vote). Une solution au problème ci-dessus est d'imposer des conditions de rationalité (axiomes) et de voir quelles sont les solutions possibles. Shapley a proposé les axiomes suivants [48] :

1. **Partage** : $\sum_i \phi_v(i) = v(N)$;
2. **Équité** : ϕ ne doit pas dépendre de la façon dont sont nommés les joueurs ;
3. **Joueur nul** : si un joueur i est tel que sa contribution à toute coalition est nulle ($v(A \cup i) = v(A)$ pour toute coalition A), alors $\phi_v(i) = 0$;
4. **Linéarité** : ϕ_v est une fonction linéaire de v , i.e. $\phi_{\alpha v + \beta v'} = \alpha \phi_v + \beta \phi_{v'}$.

Shapley a montré qu'il existe une unique fonction ϕ_v vérifiant ces quatre axiomes, dite *valeur de Shapley* :

$$\phi_v(i) = \sum_{K \subseteq N \setminus i} \frac{(n - |K| - 1)! |K|!}{n!} [v(K \cup \{i\}) - v(K)]. \quad (12)$$

Cette formule est directement utilisable pour exprimer l'importance *globale*² du critère i pour une capacité μ .

Si l'indice d'importance (globale) d'un critère est un élément fondamental pour l'interprétation du modèle, elle ne suffit pas pour décrire les différentes attitudes de décision induites par une capacité. La notion d'*interaction* entre les critères i, j exprime le fait que les importances individuelles des critères i, j ne s'accumulent pas pour former l'importance de la coalition $\{i, j\}$. Examinons le cas où $n = 2$, et considérons les quatre

²Par opposition à la quantité $\mu(\{i\})$, qui représente l'importance du critère i seul.

alternatives suivantes (voir Fig. 4), en faisant l'hypothèse que les 2 critères ont même indice d'importance :

- $x = (0_1, 0_2)$;
- $y = (1_1, 0_2)$;
- $z = (0_1, 1_2)$;
- $t = (1_1, 1_2)$.

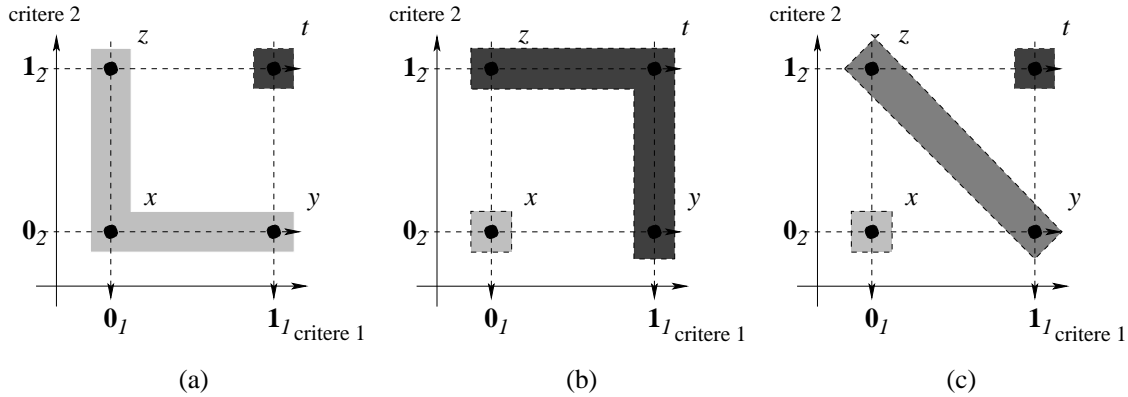


FIG. 4 – Différents cas d'interaction.

Il est clair que t est meilleur que x , mais les préférences sur les autres paires dépendent du sujet. On peut se situer entre les deux extrêmes suivants :

situation extrême 1 : on pose $\mu(\{1\}) = \mu(\{2\}) = 0$, ce qui revient aux préférences $x \sim y \sim z$ (figure 4 (a)). Cela signifie que pour le sujet, les deux critères doivent être satisfaits pour avoir un objet satisfaisant, et la satisfaction de seulement l'un des deux ne suffit pas. On dit que les critères sont *complémentaires*.

situation extrême 2 : on pose $\mu(\{1\}) = \mu(\{2\}) = 1$, ce qui équivaut aux préférences $y \sim z \sim t$ (figure 4 (b)). Dans ce cas, le sujet estime que la satisfaction de l'un des deux critères suffit pour avoir un objet satisfaisant, et satisfaire les deux critères n'est pas nécessaire. On dit que les critères sont *substitutifs*.

Dans ces deux situations, les critères ne sont pas indépendants dans le sens où la satisfaction de l'un influence sur l'utilité de l'autre pour avoir un objet globalement satisfaisant (nécessaire dans le premier cas, inutile dans le second). On dit alors qu'il y a *interaction* entre les critères.

Une situation sans interaction est telle que la satisfaction de chaque critère apporte sa propre contribution à la satisfaction globale, ce qui se traduit par :

$$\mu(\{1, 2\}) = \mu(\{1\}) + \mu(\{2\}) \quad (13)$$

(additivité) (voir Fig. 4 (c)). Dans la situation 1, on a $\mu(\{1, 2\}) > \mu(\{1\}) + \mu(\{2\})$, avec l'inégalité inversée pour la situation 2. Cela suggère que l'interaction I_{12} entre les critères 1 et 2 devrait être définie par :

$$I_{12} := \mu(\{1, 2\}) - \mu(\{1\}) - \mu(\{2\}) + \mu(\emptyset). \quad (14)$$

L'interaction est positive quand les critères sont complémentaires, et négative quand ils sont substitutifs.

Dans le cas de plus de 2 critères, la définition de l'interaction est similaire à celle de l'indice de Shapley, c'est-à-dire que toutes les coalitions de N doivent être prises en compte. La définition suivante a été proposée par Murofushi et Soneda [34], pour une paire de critères i, j :

$$I_{ij} := \sum_{K \subseteq N \setminus \{i, j\}} \frac{(n - k - 2)!k!}{(n - 1)!} [\mu(K \cup \{i, j\}) - \mu(K \cup \{i\}) - \mu(K \cup \{j\}) + \mu(K)]. \quad (15)$$

On a $I_{ij} > 0$ (resp. $< 0, = 0$) pour des critères complémentaires (resp. substitutifs, indépendants). La définition de cet indice a été ensuite étendue par l'auteur à toute coalition de critères $A \subseteq N$ [15] :

$$I_{\mu}(A) := \sum_{K \subseteq N \setminus A} \frac{(n - k - |A|)!k!}{(n - |A| + 1)!} \sum_{L \subseteq A} (-1)^{|A| - |L|} \mu(K \cup L). \quad (16)$$

(nous omettrons l'indice μ s'il n'y a pas d'ambiguïté) On remarque que l'on a $I_{ij} = I(\{i, j\})$, et aussi $I(\{i\}) = \phi(i)$, l'indice d'importance de Shapley. Pour cette raison, on appelle la fonction I l'*indice d'interaction de Shapley*. Il est facile de voir que, quand la capacité est additive, on a $I(A) = 0$ pour tout A tel que $|A| > 1$.

La fonction I peut être vue comme une transformée inversible et linéaire sur l'ensemble des capacités (ou plus généralement des jeux coopératifs v). Ceci est illustré par la figure 5. Les nombres $I(S)$ ne forment pas un treillis, par contre, ils vérifient deux relations représentées par des surfaces grisées dans la figure 5 : $\sum_{i \in N} I(i) = 1$, $\sum_{S \subseteq N} B_s I(S) = 0$, où les B_s sont les nombres de Bernoulli (voir [10, 15, 17, 18] pour tous les détails sur ces transformations).

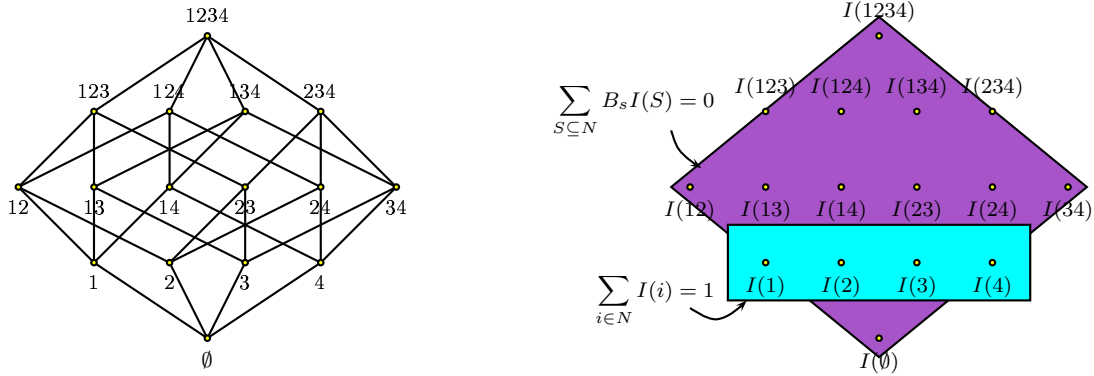


FIG. 5 – Treillis $(\mathcal{P}(N), \subseteq)$ et valeurs de I .

5.5 Capacités k -additives et capacités p -symétriques

Les notions ci-dessus sont à la base de la définition des capacités k -additives, qui vont nous permettre de couvrir les intermédiaires entre la somme pondérée et l'intégrale de Choquet dans toute sa généralité.

Une capacité est dite k -additive si son indice d'interaction I est tel que $I(A) = 0$ quand $|A| > k$ [15]. Une capacité 1-additive ne peut représenter de phénomène d'interaction puisque seuls les indices $I(\{i\})$ (i.e. les indices d'importance) sont non nuls. Il est facile de voir que c'est en fait une capacité additive au sens usuel.

Une capacité 2-additive considère seulement des interactions pour des paires de critères; elle nécessite seulement $\frac{n(n+1)}{2} - 1$ paramètres pour être entièrement déterminée. Plus généralement, une capacité k -additive nécessite $1 + n + \binom{n}{2} + \dots + \binom{n}{k} - 2$ paramètres. La figure 6 montre une représentation d'une capacité 2-additive.

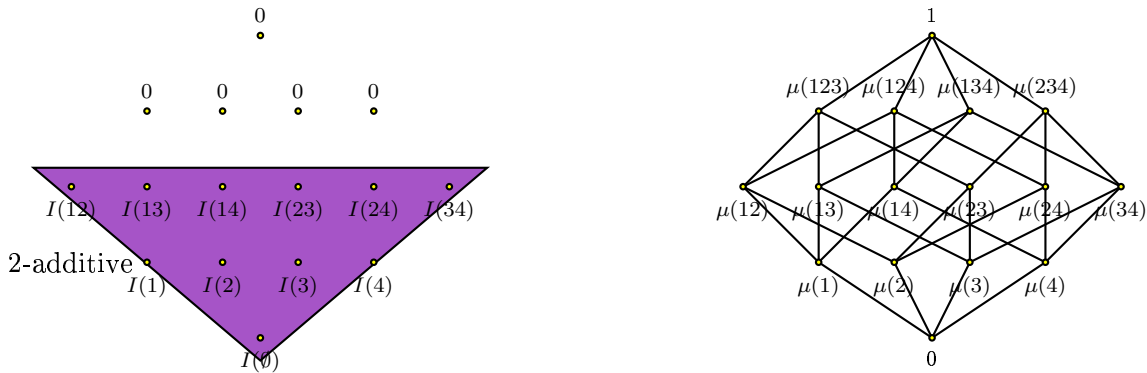


FIG. 6 – Valeurs de $I(S)$, $S \subseteq N$, pour une capacité 2-additive, et treillis des valeurs de μ obtenu par transformation inverse I^{-1} .

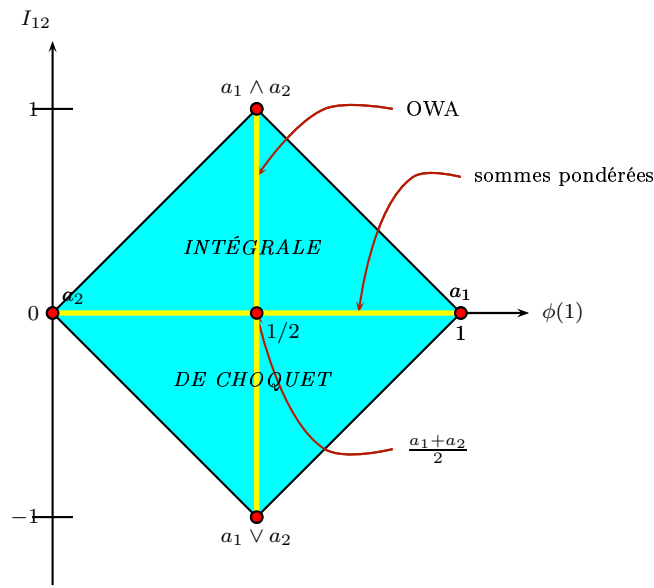


FIG. 7 – Ensemble des intégrales de Choquet pour $n = 2$.

Examinons avec plus d'attention le modèle 2-additif. Celui-ci est particulièrement intéressant puisqu'il est économique et permet néanmoins de représenter des interactions entre 2 critères. On montre que dans ce cas, l'intégrale de Choquet se met sous la forme suivante, pour $a \in \mathbb{R}^n$:

$$\mathcal{C}_\mu(a) = \sum_{I_{ij} > 0} (a_i \wedge a_j) I_{ij} + \sum_{I_{ij} < 0} (a_i \vee a_j) |I_{ij}| + \sum_{i=1}^n a_i (\phi(i) - \frac{1}{2} \sum_{j \neq i} |I_{ij}|). \quad (17)$$

On remarque que si $I_{ij} > 0$, alors l'agrégation est du type *conjonctif* (critères reliés par un minimum), et que si $I_{ij} < 0$, alors l'agrégation est de type *disjonctif* (critères reliés par un maximum). Enfin, il reste un terme linéaire faisant intervenir la valeur de Shapley. On peut montrer que $\mathcal{C}_\mu(a)$ est une combinaison convexe des a_i (dictateurs), des termes $a_i \wedge a_j$ (conjonctions) et $a_i \vee a_j$ (disjonctions) [14]. La figure 7 montre l'ensemble des intégrales de Choquet pour $n = 2$ (elles sont donc du type 2-additif), où l'on peut voir clairement cette propriété de convexité [16].

Une autre possibilité de définir une hiérarchie de modèles est de considérer les capacités p -symétriques, qui généralisent la notion de capacité symétrique [32]. On dit que $A \subseteq N$ est un *sous-ensemble d'indifférence* pour μ si pour tout $B_1, B_2 \subseteq A$ tels que $|B_1| = |B_2|$, $\mu(C \cup B_1) = \mu(C \cup B_2)$, $\forall C \subseteq N \setminus A$. Remarquons que si A est sous-ensemble d'indifférence, alors tout $A' \subseteq A$ l'est aussi. D'autre part, tout singleton est un ensemble

d'indifférence.

On dit que μ est *p-symétrique* si la partition de N la plus grossière en sous-ensembles d'indifférence a exactement p sous-ensembles A_1, \dots, A_p (on appelle une telle partition une *base* de μ). Une capacité 1-symétrique est une capacité symétrique au sens ordinaire, i.e. $\mu(A)$ dépend seulement de $|A|$.

Avec une base (A_1, \dots, A_p) , tout sous-ensemble $B \subseteq N$ est représenté par un vecteur (b_1, \dots, b_p) , où $b_i = |B \cap A_i|$. Il est facile de voir qu'une capacité p -symétrique avec une base (A_1, \dots, A_p) nécessite $\prod_{i=1}^p (|A_i| + 1)$ coefficients pour être définie. Il est intéressant de considérer la représentation sous forme de treillis de telles capacités (voir figures 8 et 9).

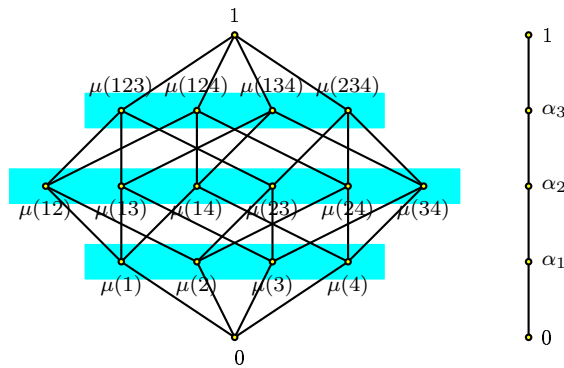


FIG. 8 – Capacité 1-symétrique. Sur le treillis $\mathcal{P}(N)$, les surfaces grisées réunissent les valeurs de μ identiques. Le treillis des valeurs de μ se réduit ainsi à celui figuré à droite.

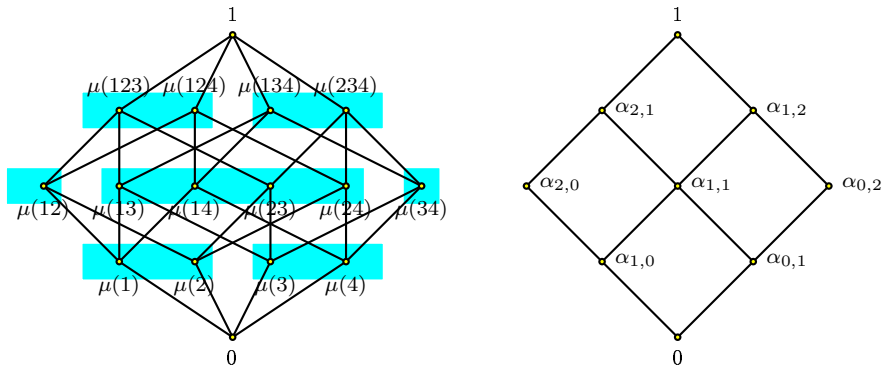


FIG. 9 – Capacité 2-symétrique avec la base $(\{1, 2\}, \{3, 4\})$. Mêmes conventions de représentation que pour la figure 8. Il y a $3 \times 3 - 2$ paramètres, formant le treillis 3^2 .

6 Les modèles bipolaires

6.1 Un exemple embarrassant

Les modèles à base de capacités présentés ci-dessus permettent une représentation très souple des préférences, si bien qu'il est possible de les adapter à des attitudes de décision très variées. Considérons l'exemple suivant, où le décideur exprime sa stratégie de décision sous forme de règles [22].

EXEMPLE : ÉVALUATION DES ÉTUDIANTS

Un directeur de faculté veut évaluer ses étudiants à partir de leurs notes en mathématiques (M), statistiques (S) et langues (L). Plutôt que de fixer des coefficients d'importance pour chaque matière et d'utiliser une somme pondérée (il a été convaincu à la lecture de notre section 4.3), il pense que l'importance des matières dépend en fait du profil de l'étudiant en question, plus particulièrement de ses aptitudes en mathématiques. Il exprime les deux règles suivantes :

(R1) *Pour un étudiant bon en mathématiques, les langues sont plus importantes que les statistiques.*

(R2) *Pour un étudiant mauvais en mathématiques, les statistiques sont plus importantes que les langues.*

La raison qui sous-tend ces deux règles est que la faculté étant de type scientifique, on ne saurait admettre un étudiant faible dans les deux matières scientifiques, et on désire si possible des étudiants également bons en langues.

Il considère les deux étudiants A et B.

	mathématiques (M)	statistiques (S)	langues (L)
étudiant A	14	16	7
étudiant B	14	15	8

D'après la règle (R1), on a sans ambiguïté $A \prec B$. Si l'on désire représenter cette préférence par une intégrale de Choquet, ceci implique $\mu(\{M, S\}) + \mu(\{S\}) < 1$, ainsi qu'on peut facilement le vérifier. Le directeur considère maintenant les étudiants C et D.

	mathématiques (M)	statistiques (S)	langues (L)
étudiant C	9	16	7
étudiant D	9	15	8

Cette fois, d'après la règle (R2), on a $C \succ D$, ce qui implique que $\mu(\{M, S\}) + \mu(\{S\}) > 1$. Mais ceci mène à une contradiction !

Il s'avère donc impossible de représenter ces préférences par une intégrale de Choquet.

Implicitement, le directeur considère que 10 est la limite entre les bonnes et les mauvaises notes. Il y a donc existence d'un niveau frontière, ou comme nous l'avons introduit dans la section 4.1, d'un niveau neutre, ce qui fait que l'échelle de notation $[0, 20]$ est bipolaire. Or l'intégrale de Choquet telle que nous l'avons introduite ne peut "voir" une échelle bipolaire. En effet, on a la propriété suivante : $\mathcal{C}_\mu(\alpha x + \beta) = \alpha \mathcal{C}_\mu(x) + \beta \mu(N)$, autrement dit l'intégrale de Choquet est insensible au changement d'échelle de différence, et aucun point de l'échelle ne peut jouer un rôle particulier. Il convient donc de trouver une version bipolaire de l'intégrale de Choquet.

6.2 Généralisation de l'intégrale de Choquet pour les échelles bipolaires

Jusqu'à présent, nous avons construit notre modèle en considérant sur chaque attribut les niveaux remarquables $\mathbf{0}_i$ (neutre ou plus petit élément) et $\mathbf{1}_i$ (satisfaisant). Dans une échelle bipolaire, il y a un troisième niveau remarquable, le niveau $-\mathbf{1}_i$ qui est symétrique du niveau satisfaisant. Il semble donc opportun d'utiliser conjointement ces trois niveaux pour construire notre nouveau modèle.

Un premier modèle très simple consisterait à supposer une symétrie entre les parties positive (au-dessus du neutre) et négative (au-dessous du neutre). Il suffirait donc de considérer les alternatives binaires comme précédemment et de poser :

$$\psi(-1_A, 0_{A^c}) = -\psi(1_A, 0_{A^c}). \quad (18)$$

Pour un vecteur de scores $x \in \mathbb{R}^n$ quelconque, on écrit que :

$$\psi(x) := \psi(x^+) - \psi(x^-) \quad (19)$$

avec $x_i^+ := x_i \vee 0$ et $x_i^- = (-x_i)^+$, $i = 1, \dots, n$. Quand ψ est l'intégrale de Choquet, nous obtenons l'intégrale de Šipoš [54] ou *intégrale de Choquet symétrique* [9] :

$$\tilde{C}_\mu(x) := C_\mu(x^+) - C_\mu(x^-). \quad (20)$$

Appliquons l'intégrale de Choquet symétrique à l'exemple ci-dessus. Il est facile de vérifier que les préférences du directeur sur A, B, C, D impliquent que $\mu(\{S\}) < \mu(\{L\})$ et $\mu(\{S\}) > \mu(\{L\})$, ce qui est toujours insoluble.

Un second modèle serait de ne pas supposer une parfaite symétrie entre les parties négatives et positives, mais simplement une séparation. En plus des alternatives binaires, il faut considérer les *alternatives binaires négatives* $(-\mathbf{1}_A, \mathbf{0}_{A^c})$, $A \subseteq N$. Les alternatives binaires donnent lieu à une capacité μ^+ , tandis que les alternatives binaires négatives donnent lieu à une capacité μ^- . En procédant comme dans (20), on obtient :

$$C_{\mu^+, \mu^-}(x) := C_{\mu^+}(x^+) - C_{\mu^-}(x^-), \quad \forall x \in \mathbb{R}^n. \quad (21)$$

Ce modèle est celui proposé par Tversky and Kahnemann sous le nom de "Cumulative Prospect Theory Model" (modèles des perspectives cumulées), que nous abrègerons par CPT [53]. En fait, on s'aperçoit que ce modèle non plus ne peut représenter les préférences du directeur sur A, B, C, D, car on trouve les inégalités suivantes à satisfaire : $\mu^+(\{S\}) < \mu^-(\{L\})$ et $\mu^+(\{S\}) > \mu^-(\{L\})$, ce qui mène toujours à une impossibilité.

Un troisième modèle envisageable serait de ne plus supposer la séparabilité des parties positives et négatives. Il faut dans ce cas considérer toutes les combinaisons possibles des trois niveaux $-\mathbf{1}, \mathbf{0}, \mathbf{1}$ sur les attributs, ce qu'on appelle une *alternative ternaire* $(\mathbf{1}_A, -\mathbf{1}_B, \mathbf{0}_{(A \cup B)^c})$, $\forall A, B \subseteq N$, $A \cap B = \emptyset$. En notant

$$\mathcal{Q}(N) := \{(A, B) \in \mathcal{P}(N) \times \mathcal{P}(N) \mid A \cap B = \emptyset\}, \quad (22)$$

construire une échelle de différence à l'aide de MACBETH en donnant un score global à toutes les alternatives ternaires conduit à une fonction $v : \mathcal{Q}(N) \rightarrow \mathbb{R}$, que nous appelons une *bi-capacité*. Les trois modèles que nous venons d'introduire sont représentés sur la figure 10. Étudions d'un peu plus près les bi-capacités. Nous allons montrer qu'elles permettent de résoudre notre problème, et que bien sûr, elles contiennent les deux premiers modèles.

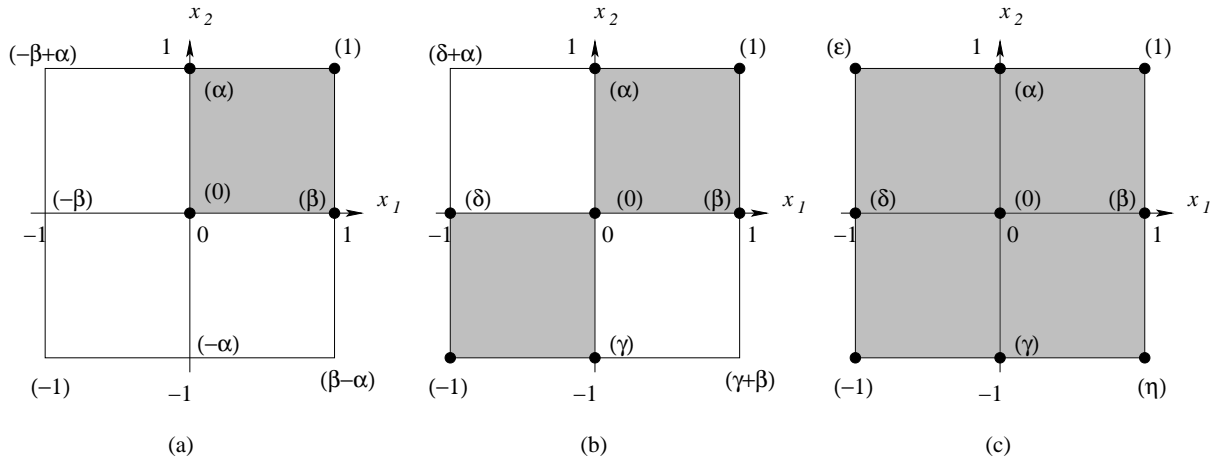


FIG. 10 – Cas d’une échelle bipolaire : modèles symétrique (a), CPT (b), bipolaire général (c). Les ronds noirs indiquent les alternatives fictives déterminant le modèle, les valeurs entre parenthèses des évaluations globales ; les scores dans les parties non grisées sont calculés à partir des scores dans les parties grisées.

6.3 Bi-capacités et intégrale de Choquet

Une fonction $v : \mathcal{Q}(N) \longrightarrow \mathbb{R}$ est une *bi-capacité* [21, 20] si elle vérifie :

(i) $v(\emptyset, \emptyset) = 0$

(ii) $A \subseteq B$ implique $v(A, \cdot) \leq v(B, \cdot)$ et $v(\cdot, A) \geq v(\cdot, B)$.

De plus, v est *normalisée* si $v(N, \emptyset) = 1 = -v(\emptyset, N)$. Une bi-capacité est du *type CPT* si elle peut s’écrire sous la forme :

$$v(A, B) = \nu^+(A) - \nu^-(B)$$

où ν^+, ν^- sont des capacités ordinaires. Une bi-capacité est *symétrique* si elle est du type CPT avec $\nu^+ = \nu^-$. Il est facile de voir que l’on retrouve ainsi les deux premiers modèles pour les alternatives binaires et binaires négatives. Il reste à prouver que ceci reste vrai pour tous les vecteurs de score $x \in \mathbb{R}^n$, mais pour cela il faut définir ce qu’est l’intégrale de Choquet pour une bi-capacité.

Essayons de trouver une définition adéquate pour l’intégrale de Choquet, et pour cela procédons comme dans le cas classique, c’est-à-dire par interpolation. Prenons le cas $n = 2$ (voir figure 11). L’idée est de se ramener par une symétrie adéquate au premier quadrant $\{x_1 \geq 0, x_2 \geq 0\}$, où l’on sait que l’interpolateur linéaire le plus simple est l’intégrale de

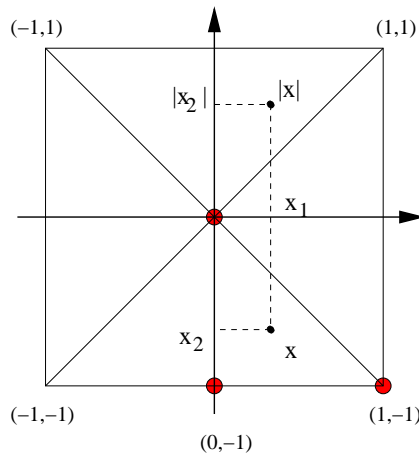


FIG. 11 – Interpolation sur une échelle bipolaire : cas $n = 2$.

Choquet. Par exemple, supposons que $x_1 \geq 0$, $x_2 \leq 0$, $|x_1| \leq |x_2|$. Alors nous pouvons écrire directement :

$$\psi(x_1, x_2) := |x_1|\psi(1, -1) + (|x_2| - |x_1|)\psi(0, -1)$$

ce qui est l'intégrale de Choquet de $|x|$ par rapport à la fonction ν_1 définie par :

$$\begin{aligned} \nu_1(\{1, 2\}) &= \psi(1, -1) \\ \nu_1(\{2\}) &= \psi(0, -1). \end{aligned}$$

Examinons maintenant le cas général. On définit $N^+ = \{i \in N \mid x_i \geq 0\}$, $N^- = N \setminus N^+$.

Alors

$$\begin{aligned} \psi(x) &= |x_{\sigma(1)}|\psi(1_{N^+}, 1_{N^-}) + \\ &\quad \sum_{i=2}^n (|x_{\sigma(i)}| - |x_{\sigma(i-1)}|)\psi(1_{\{\sigma(i), \dots, \sigma(n)\} \cap N^+}, -1_{\{\sigma(i), \dots, \sigma(n)\} \cap N^-}) \end{aligned}$$

ce qui est l'intégrale de Choquet de $|x|$ par rapport à une fonction ν_{N^+} définie par

$$\nu_{N^+}(A) := \psi(1_{A \cap N^+}, -1_{A \cap N^-}).$$

En résumé, on arrive à la définition suivante. Soit v une bi-capacité et f une fonction à valeurs réelles définie sur N . L'intégrale de Choquet de f par rapport à v est définie par :

$$\mathcal{C}_v(f) := \mathcal{C}_{\nu_{N^+}}(|f|) \tag{23}$$

où ν_{N^+} est une fonction sur $\mathcal{P}(N)$ définie par $\nu_{N^+}(C) := v(C \cap N^+, C \cap N^-)$, et $N^+ := \{i \in N | f_i \geq 0\}$, $N^- = N \setminus N^+$.

On a par construction $\mathcal{C}_v(1_A, -1_B) = v(A, B)$ pour tout $(A, B) \in \mathcal{Q}(N)$. De plus, si v est du type CPT, avec $v(A, B) = \nu_+(A) - \nu_-(B)$, alors

$$\mathcal{C}_v(f) = \mathcal{C}_{\nu_+}(f^+) - \mathcal{C}_{\nu_-}(f^-) \quad (24)$$

Cette équation prouve que ce modèle contient effectivement comme cas particulier le modèle CPT.

Essayons maintenant de représenter les préférences du directeur sur A, B, C et D. On obtient les inégalités suivantes :

$$\begin{aligned} v(\{M, S\}, \emptyset) - v(\{M, S\}, \{L\}) &> v(\{S\}, \emptyset) \\ v(\{S\}, \{L\}) &> 0. \end{aligned}$$

Il n'y a manifestement pas de contradiction dans ce système, donc le modèle des bi-capacités peut représenter les préférences du directeur.

7 Conclusion

Nous avons tenté dans cet article de donner une vue d'ensemble unifiée des méthodes de décision multicritère, sans prétendre à l'exhaustivité. La taille de l'article n'a pas permis de pouvoir détailler les nombreux concepts introduits, et certains thèmes importants n'ont pu être abordés.

Parmi les thèmes que nous aurions souhaité pouvoir plus développer figure le concept de bi-capacité introduit récemment. Il semble qu'il puisse ouvrir des horizons nouveaux en décision multicritère, mais aussi en théorie des jeux coopératifs, où les bi-capacités sont déjà connues sous le nom de jeux bi-coopératifs [6]. Des travaux proches sont également menés par Greco *et al.* dans le domaine de la décision [26].

Enfin, un thème complètement passé sous silence est celui de la décision dans un contexte ordinal. Que faire si les données (en l'occurrence les scores sur les critères) ne sont pas exprimées sur une échelle numérique ? Il existe une recherche très active dans ce domaine, qui nécessiterait un chapitre entier, c'est pourquoi nous avons choisi d'omettre

ce sujet. On trouvera dans le chapitre de synthèse [22] une section consacrée à ce sujet, dans le même esprit que cet article.

Références

- [1] A. Appriou, A. Ayoun, S. Benferhat, P. Besnard, L. Cholvy, R. Cooke, F. Cuppens, D. Dubois, H. Fargier, M. Grabisch, R. Kruse, J. Lang, S. Moral, H. Prade, A. Safiotti, P. Smets, and C. Sossai. Fusion : general concepts and characteristics. *Int. J. of Intelligent Systems*, 16 :1107–1134, 2001.
- [2] K. Arrow. *Social choice et individual values*. Wiley, 2nd edition, 1963.
- [3] C.A. Bana e Costa and J.C. Vansnick. A theoretical framework for Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH). In *Proc. XIth Int. Conf. on MultiCriteria Decision Making*, pages 15–24, Coimbra, Portugal, August 1994.
- [4] C.A. Bana e Costa and J.C. Vansnick. Applications of the MACBETH approach in the framework of an additive aggregation model. *J. of Multicriteria Decision Analysis*, 6 :107–114, 1997.
- [5] C.A. Bana e Costa and J.C. Vansnick. The MACBETH approach : basic ideas, software and an application. In N. Meskens and M. Roubens, editors, *Advances in Decision Analysis*, pages 131–157. Kluwer Academic Publishers, 1999.
- [6] J.M. Bilbao, J.R. Fernandez, A. Jiménez Losada, and E. Lebrón. Bicooperative games. In J.M. Bilbao, editor, *Cooperative games on combinatorial structures*. Kluwer Acad. Publ., 2000.
- [7] J.T. Cacioppo, W.L. Gardner, and G.G. Berntson. Beyond bipolar conceptualizations and measures : the case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1(1) :3–25, 1997.
- [8] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5 :131–295, 1953.
- [9] D. Denneberg. *Non-Additive Measure and Integral*. Kluwer Academic, 1994.
- [10] D. Denneberg and M. Grabisch. Interaction transform of set functions over a finite set. *Information Sciences*, 121 :149–170, 1999.

- [11] I. Dragan. Personal communication, 1997.
- [12] D. Dubois and H. Prade. Weighted minimum and maximum operations in fuzzy set theory. *Information Sciences*, 39 :205–210, 1986.
- [13] J.C. Fodor and M. Roubens. *Fuzzy Preference Modelling and Multi-Criteria Decision Aid*. Kluwer Academic Publisher, 1994.
- [14] M. Grabisch. Alternative representations of discrete fuzzy measures for decision making. *Int. J. of Uncertainty, Fuzziness, and Knowledge Based Systems*, 5 :587–607, 1997.
- [15] M. Grabisch. k -order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92 :167–189, 1997.
- [16] M. Grabisch. A graphical interpretation of the Choquet integral. *IEEE Tr. on Fuzzy Systems*, 8 :627–631, 2000.
- [17] M. Grabisch. The interaction and Möbius representations of fuzzy measures on finite spaces, k -additive measures : a survey. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals — Theory and Applications*, pages 70–93. Physica Verlag, 2000.
- [18] M. Grabisch. Set function over finite sets : transformations and integrals. In E. Pap, editor, *Handbook of Measure Theory*, pages 1381–1401. Elsevier Science Publ., 2002.
- [19] M. Grabisch. The Choquet integral as a linear interpolator. In *10th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2004)*, Perugia, Italy, July 2004.
- [20] M. Grabisch and Ch. Labreuche. Bi-capacities. In *Joint Int. Conf. on Soft Computing and Intelligent Systems and 3d Int. Symp. on Advanced Intelligent Systems*, Tsukuba, Japan, October 2002.
- [21] M. Grabisch and Ch. Labreuche. Bi-capacities for decision making on bipolar scales. In *EUROFUSE Workshop on Informations Systems*, pages 185–190, Varenna, Italy, September 2002.
- [22] M. Grabisch and Ch. Labreuche. Fuzzy measures and integrals in MCDA. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis*. Kluwer Academic Publishers, to appear.

- [23] M. Grabisch, Ch. Labreuche, and J.C. Vansnick. On the extension of pseudo-Boolean functions for the aggregation of interacting bipolar criteria. *Eur. J. of Operational Research*, 148 :28–47, 2003.
- [24] M. Grabisch, S.A. Orlovski, and R.R. Yager. Fuzzy aggregation of numerical preferences. In R. Slowiński, editor, *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, The Handbooks of Fuzzy Sets Series, D. Dubois and H. Prade (eds), pages 31–68. Kluwer Academic, 1998.
- [25] M. Grabisch and P. Perny. Agrégation multicritère. In C. Marsala B. Bouchon-Meunier, editor, *Logique floue, principes, aide à la décision*, pages 81–120. Hermès, 2003.
- [26] S. Greco, B. Matarazzo, and R. Słowiński. Bipolar Sugeno and Choquet integrals. In *EUROFUSE Workshop on Informations Systems*, Varenna, Italy, September 2002.
- [27] O. Hudry. Votes et paradoxes : les élections ne sont pas monotones! *Math. et Sci. Humaines / Mathematics and Social Sciences*, 41(163) :9–39, 2003.
- [28] R.L. Keeney and H. Raiffa. *Decision with Multiple Objectives*. Wiley, New York, 1976.
- [29] D.H. Krantz, R.D. Luce, P. Suppes, and A. Tversky. *Foundations of measurement*, volume 1 : Additive and Polynomial Representations. Academic Press, 1971.
- [30] Ch. Labreuche and M. Grabisch. The Choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets and Systems*, 137 :11–26, 2003.
- [31] J.L. Marichal. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Tr. on Fuzzy Systems*, 8(6) :800–807, 2000.
- [32] P. Miranda, M. Grabisch, and P. Gil. p -symmetric fuzzy measures. *Int. J. of Uncertainty, Fuzziness, and Knowledge-Based Systems*, 10 (Suppl.) :105–123, 2002.
- [33] T. Murofushi. A technique for reading fuzzy measures (I) : the Shapley value with respect to a fuzzy measure. In *2nd Fuzzy Workshop*, pages 39–48, Nagaoka, Japan, October 1992. In Japanese.
- [34] T. Murofushi and S. Soneda. Techniques for reading fuzzy measures (III) : interaction index. In *9th Fuzzy System Symposium*, pages 693–696, Sapporo, Japan, May 1993. In Japanese.

- [35] T. Murofushi and M. Sugeno. An interpretation of fuzzy measure and the Choquet integral as an integral with respect to a fuzzy measure. *Fuzzy Sets & Systems*, 29 :201–227, 1989.
- [36] T. Murofushi and M. Sugeno. Some quantities represented by the Choquet integral. *Fuzzy Sets & Systems*, 56 :229–235, 1993.
- [37] C.E. Osgood, G.J. Suci, and P.H. Tannenbaum. *The measurement of meaning*. University of Illinois Press, Urbana, IL, 1957.
- [38] E. Peters and P. Slovic. Affective asynchrony and the measurement of the affective attitude component. Working paper.
- [39] M. Pirlot and Ph. Vincke. *Semiororders — Properties, Representations, Applications*. Kluwer Academic Publishers, 1997.
- [40] J.C. Pomerol and S. Barba-Romero. *Multicriterion decision in management : principles and practice*. Kluwer Academic Publishers, 2000.
- [41] R. Radner. Satisficing. *J. of Math. Economics*, 2 :253–262, 1975.
- [42] F.S. Roberts. *Measurement Theory*. Addison-Wesley, 1979.
- [43] M. Roubens and Ph. Vincke. *Preference Modelling*, volume 250 of *Lecture Notes in Economics and Mathematical Sciences*. Springer Verlag, 1985.
- [44] B. Roy. Classement et choix en présence de points de vue multiples (la méthode ELECTRE). *R.I.R.O.*, 2 :57–75, 1968.
- [45] B. Roy. How outranking relations helps multiple criteria decision making. In J.L. Cochrane and M. Zeleny, editors, *Multiple Criteria Decision Making*, pages 179–201. University of South California Press, Columbia, 1973.
- [46] B. Roy. Electre III : un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples. *Cahiers du Centre d'Etude de Recherche Opérationnelle*, 20(1) :32–43, 1978.
- [47] B. Roy and D. Bouyssou. *Aide multicritère à la décision : Méthodes et Cas*. Economica, 1994.
- [48] L.S. Shapley. A value for n -person games. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Games, Vol. II*, number 28 in *Annals of Mathematics Studies*, pages 307–317. Princeton University Press, 1953.

- [49] H. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63(2) :129–138, 1956.
- [50] H.A. Simon. Theories of bounded rationality. In Peter Earl, editor, *The Legacy of Herbert Simon in Economic Analysis*, volume 1. Edward Elgar Publishing Ltd, 2001.
- [51] P. Slovic, M. Finucane, E. Peters, and D.G. MacGregor. The affect heuristic. In T. Gilovitch, D. Griffin, and D. Kahneman, editors, *Heuristics and biases : the psychology of intuitive judgment*, pages 397–420. Cambridge University Press, 2002.
- [52] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.
- [53] A. Tversky and D. Kahneman. Advances in prospect theory : cumulative representation of uncertainty. *J. of Risk and Uncertainty*, 5 :297–323, 1992.
- [54] J. Šipoš. Integral with respect to a pre-measure. *Math. Slovaca*, 29 :141–155, 1979.
- [55] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. Systems, Man & Cybern.*, 18 :183–190, 1988.