Agrégation de critères interactifs au moyen de l'intégrale de Choquet discrète*

J.-L. Marichal
Département de gestion, Université de Liège
Boulevard du Rectorat 7 - B31
B-4000 Liège, Belgique
Email: jl.marichal@ulg.ac.be

Résumé

L'opérateur le plus souvent utilisé pour agréger des critères dans les problèmes d'aide multicritère à la décision est la moyenne arithmétique pondérée. Cependant, dans de nombreux problèmes pratiques, les critères interagissent, ce qui nécessite l'emploi d'un agrégateur plus général. Nous montrons que, sous des hypothèses assez naturelles, l'intégrale de Choquet discrète est un opérateur d'agrégation approprié qui généralise la moyenne arithmétique pondérée par la prise en compte de l'interaction parmi les critères. L'axiomatique qui permet de caractériser l'intégrale de Choquet est présentée en détail.

Mots Clef

Aide multicritère à la décision, critères interactifs, intégrale de Choquet.

Abstract

The most often used operator to aggregate criteria in decision making problems is the classical weighted arithmetic mean. In many problems however, the criteria considered interact, and a substitute to the weighted arithmetic mean has to be adopted. We show that, under rather natural conditions, the discrete Choquet integral is an adequate aggregation operator that extends the weighted arithmetic mean by the taking into consideration of the interaction among criteria. The axiomatic that supports the Choquet integral is presented in detail.

Keywords

Multicriteria decision making, interacting criteria, Choquet integral.

*A présenter aux "Rencontres francophones sur la logique floue et ses applications" (LFA'99), Valenciennes, France, 21–22 octobre 1999.

1 Introduction

Considérons un ensemble fini d'*alternatives* $A = \{a, b, c, \ldots\}$ et un ensemble fini de *critères* $N = \{1, \ldots, n\}$ dans un problème d'aide multicritère à la décision. A chaque alternative $a \in A$ est associé un *profil* $x^a = (x_1^a, \ldots, x_n^a) \in E^n$, où, pour tout $i \in N$, x_i^a représente le score partiel de a sur le critère i, et E est un interval réel, éventuellement non borné. Habituellement, on prend $E = [0, 1]$. Nous supposons que tous les scores partiels sont définis sur une même échelle d'intervalle, selon la théorie du mesurage.
A partir du profil d'une alternative quelconque a, on peut calculer un score global $M(x^a)$ au moyen d'un opérateur d'agrégation $M : E^n \to \mathbb{R}$ qui prend en compte les poids d'importance des critères. Une fois les scores globaux calculés, ils peuvent être utilisés pour ranger les alternatives ou pour sélectionner celle qui satisfait le mieux les critères proposés.
Il y a peu de temps encore, les opérateurs d'agrégation les plus utilisés étaient les moyennes arithmétiques pondérées, c'est-à-dire des opérateurs de la forme

$$M_\omega(x) = \sum_{i=1}^{n} \omega_i \, x_i, \tag{1}$$

avec $\sum_i \omega_i = 1$ et $\omega_i \geq 0$ pour tout $i \in N$. Malheureusement, puisque ces opérateurs sont incapables de modéliser une interaction entre les critères, ils ne peuvent être utilisés qu'en présence de critères indépendants. Ils ne sont donc pas appropriés pour l'agrégation de critères interactifs.

Pour obtenir une représentation assez souple des phénomènes complexes d'interaction, il est utile de substituer au vecteur poids ω une fonction d'ensemble sur N, non additive, permettant de définir un poids non seulement sur chaque critère, mais aussi sur chaque sous-ensemble de critères. Pour ce faire, on utilise le concept de mesure floue [25].

Une *mesure floue* (ou *capacité de Choquet*) sur N est une fonction d'ensemble $v : 2^N \to [0,1]$ croissante par rapport à l'inclusion et telle que $v(\emptyset) = 0$ et $v(N) = 1$. Pour chaque sous-ensemble S de critères, $v(S)$ est alors interprété comme le *poids* relatif à S.

Cela étant, un opérateur d'agrégation approprié, qui généralise la moyenne arithmétique pondérée, est l'intégrale de Choquet discrète, dont l'utilisation a été proposée par de nombreux auteurs (voir par ex. [6] et les références qui y sont mentionnées). Elle est donnée par

$$\mathcal{C}_v(x) := \sum_{i=1}^{n} x_{(i)} \, [v(A_{(i)}) - v(A_{(i+1)})],$$

où $(\cdot)$ indique une permutation de N telle que $x_{(1)} \leq \ldots \leq x_{(n)}$. De plus, $A_{(i)} = \{(i), \ldots, (n)\}$, et $A_{(n+1)} = \emptyset$.
Nous voyons que l'intégrale de Choquet est une expression linéaire, à un réarrangement des éléments près. De plus, elle s'identifie à la moyenne arithmétique pondérée dès que la mesure floue v est additive.

L'objet principal de ce papier est de présenter l'intégrale de Choquet comme une extension appropriée de la moyenne arithmétique pondérée pour l'agrégation de critères. Cet opérateur offre en effet une très grande souplesse tout en gardant, dans un certain sens, une forme linéaire. Bien que sa définition ne soit pas très intuitive, nous montrerons que l'intégrale de Choquet peut être caractérisée axiomatiquement au moyen de propriétés assez naturelles.
Le papier est organisé comme suit. A la section 2, nous examinons trois types de dépendance entre critères: la corrélation, l'interchangeabilité, et la dépendance préférentielle. A la section 3, nous précisons le cadre de notre étude en introduisant quelques propriétés souvent requises pour l'agrégation. A la section 4, nous présentons l'intégrale de Choquet d'une manière assez intuitive et nous proposons une caractérisation axiomatique. Finalement, à la section 5, nous introduisons les indices d'importance et d'interaction, qui permettent d'interpréter le comportement de l'agrégation.

De façon à éviter des notations trop lourdes, le cardinal des sous-ensembles $S, T, \ldots$ sera très souvent noté par les lettres minuscules correspondantes $s, t, \ldots$, sinon par la notation standard $|S|, |T|, \ldots$. De plus, nous omettrons souvent les accolades pour les singletons, en écrivant par exemple $a(i)$, $N \setminus i$ au lieu de $a(\{i\})$, $N \setminus \{i\}$. De même, pour les paires, nous écrirons souvent ij au lieu de $\{i, j\}$, comme par exemple $a(ij)$.

Pour chaque sous-ensemble $S \subseteq N$, e_S dénotera le vecteur caractéristique de S dans $\{0,1\}^n$, c'est-à-dire le vecteur de $\{0,1\}^n$ dont la i-ème composante est 1 si et seulement si $i \in S$. Nous introduisons aussi la notation

$$xSy := \sum_{i \in S} x_i \, e_i + \sum_{i \in N \setminus S} y_i \, e_i, \qquad x, y \in \mathbb{R}^n.$$

Finalement, $\wedge$ et $\vee$ dénoterons respectivement le minimum et le maximum.

2 Que signifie "critères interactifs" ?

Dans beaucoup d'applications pratiques, les critères de décision présentent une certaine interaction. Cependant, le problème de la modélisation d'une telle interaction reste une question difficile, souvent négligée. Le fait que les phénomènes d'interaction existent dans les situations réelles est accepté par tout le monde, mais le manque d'outils appropriés pour les modéliser pousse le praticien à supposer que ses critères sont indépendants et exhaustifs. Ce comportement vient notamment de l'absence d'une définition précise de l'interaction.
Les phénomènes d'interaction parmi les critères peuvent être très complexes et difficiles à identifier. Dans cette section, nous présentons trois types de dépendance: la corrélation, l'interchangeabilité, et la dépendance préférentielle.

2.1 Corrélation

La corrélation est probablement le type de dépendance le mieux connu et le plus intuitif, voir par ex. Roy [20, Sect. 10.3]. Deux critères $i, j \in N$ sont positivement corrélés si on peut observer une corrélation positive entre les scores partiels relatifs à i et ceux relatifs à j.
Par exemple, considérons le problème de l'évaluation d'étudiants par rapport à trois cours de mathématiques (critères): statistique, probabilité, et algèbre. Les deux premiers critères sont clairement corrélés puisque, habituellement, les étudiants qui sont bons en statistique sont également bons en probabilité, et vice versa. Ainsi, ces deux critères présentent un certain degré de redondance.
Supposons qu'une moyenne arithmétique pondérée soit utilisée pour évaluer les étudiants et supposons aussi que le troisième critère soit plus important que les deux premiers, de tel sorte que les poids soient par exemple 0.3, 0.3, 0.4, respectivement. Puisque les deux premiers critères se recouvrent quelque peu, l'évaluation globale sera surestimée (resp. sousestimée) pour les étudiants qui sont bons (resp. mauvais) en statistique et/ou en probabilité.

Ce phénomène indésirable peut être facilement surmonté en utilisant une mesure floue convenable v et l'intégrale de Choquet $\mathcal{C}_v$. Une corrélation positive entre les critères i et j doit alors être modélisée par l'inégalité suivante:

$$v(ij) < v(i) + v(j),$$

qui exprime une *interaction négative* ou une *synergie négative* entre i et j. Plus exactement, si i et j sont positivement corrélés alors la contribution marginale de j à chaque combinaison de critères contenant i est strictement inférieure à la contribution marginale de j à cette même combinaison mais où i est exclu, c'est-à-dire,

$$v(T\cup ij) - v(T\cup i) < v(T\cup j) - v(T), \qquad T \subseteq N\setminus ij.$$

En cas d'égalité, les critères i et j ne sont pas corrélés.

Supposons maintenant que i et j soient négativement corrélés: des scores partiels élevés sur i entrainent souvent des scores partiels bas sur j, et vice versa. Dans ce cas, la satisfaction simultanée des deux critères est plutôt rare et les alternatives qui présentent un tel profil de satisfaction devraient être favorisées (par exemple, des étudiants bons à la fois en droit et en algèbre). Ainsi, une corrélation négative entre les critères i et j doit être modélisée par l'inégalité suivante:

$$v(ij) > v(i) + v(j),$$

qui exprime une *interaction positive* ou une *synergie positive* entre i et j. Ces deux critères présentent alors un certain degré d'opposition ou de complémentarité. Ici encore, une modélisation correcte de la corrélation entre i et j nécessite la prise en compte des autres combinaisons. Nous écrivons alors

$$v(T\cup ij) - v(T\cup i) > v(T\cup j) - v(T), \qquad T \subseteq N\setminus ij.$$

2.2 Interchangeabilité

Un autre type de dépendance est celui de l'interchangeabilité entre les critères. Considérons encore deux critères $i, j \in N$, et supposons que le décideur demande que la satisfaction d'un seul critère produise presque le même effet que la satisfaction des deux. Par exemple, il est important que les étudiants soient bons dans les matières scientifiques *ou* littéraires. S'ils sont bons dans les deux directions, c'est à peine plus satisfaisant.

Bien sûr, un tel comportement ne peut être exprimé par une moyenne arithmétique pondérée. Ici, l'importance de la paire $\{i, j\}$ est proche de l'importance des critères i et j, même en présence des autres critères. Cette condition peut être facilement exprimée par une mesure floue v telle que

$$v(T) < \left\{\begin{matrix} v(T\cup i) \\ v(T\cup j) \end{matrix}\right\} \approx v(T\cup ij), \qquad T \subseteq N\setminus ij,$$

où "$\approx$" signifie "approximativement égal". Dans ce cas, nous observons que les critères i et j sont presque *substitutifs* ou *interchangeables*. Dans le cas extrême de l'égalité, ils peuvent même être confondus.

Alternativement, le décideur peut demander que la satisfaction d'un seul critère produise très peu d'effet par rapport à la satisfaction des deux. Nous parlons alors de *complémentarité*, qui est modélisée par une mesure floue v telle que

$$v(T) \approx \left\{\begin{matrix} v(T\cup i) \\ v(T\cup j) \end{matrix}\right\} < v(T\cup ij), \qquad T \subseteq N\setminus ij.$$

Notons que, contrairement aux phénomènes de corrélation, l'interchangeabilité et la complémentarité entre critères ne peuvent être détectées en observant la table des scores. Elles représentent simplement l'opinion du décideur sur l'importance relative des critères, indépendamment des scores partiels obtenus par les alternatives sur ces critères.

2.3 Dépendance préférentielle

Le dernier type de dépendance que nous présentons est la *dépendance préférentielle*, et son opposée, l'*indépendance préférentielle*, bien connues en théorie de l'utilité multiattributs (MAUT), voir par ex. [5, 10, 26].

Supposons que les préférences sur A du décideur soient connues et exprimées par un préordre total $\succeq$. Par l'identification naturelle des alternatives avec leurs profils dans E^n, cette relation de préférence peut être considérée comme une relation de préférence sur E^n.

Définition 2.1 Le sous-ensemble S de critères est dit *préférentiellement indépendant* de $N\setminus S$ si, pour tout $x, x', y, z \in E^n$, nous avons

$$xSy \succeq x'Sy \quad \Leftrightarrow \quad xSz \succeq x'Sz.$$

L'ensemble tout entier N des critères est dit *mutuellement préférentiellement indépendant* si S est préférentiellement indépendant de $N\setminus S$ pour chaque $S \subseteq N$.

En termes simples, la préférence de xSy sur $x'Sy$ n'est pas influencée par la partie commune y. Le problème suivant [14] montre qu'il peut être naturel pour un sous-ensemble d'être préférentiellement indépendant de son complémentaire.

Exemple 2.1 Considérons le problème de ranger des restaurants sur base de leur aptitude à préparer trois plats: cuisses de grenouille (CG), steak tartare (ST), et palourdes farcies (PF). Quatre restaurants a, b, c, d ont été évalués par un guide des restaurants français comme suit (les évaluations sont exprimées sur une échelle de 0 à 20):

restaurant	CG	ST	PF
a	18	15	19
b	15	18	19
c	18	15	11
d	15	18	11

Il est demandé au décideur de s'exprimer en donnant un rangement sur $A = \{a, b, c, d\}$. Evidemment, les préférences $a \succ c$ et $b \succ d$ sont immédiatement suggérées. Le décideur se rend alors compte que les autres comparaisons ne sont pas aussi évidentes car les profils associés s'entrelacent. Il propose alors le raisonnement suivant: lorsqu'un restaurant est réputé pour ses palourdes farcies, on s'attend à ce qu'il soit mieux coté pour son aptitude à préparer les cuisses de grenouille que le steak tartare, donc $a \succ b$. Par contre, lorsqu'un restaurant prépare mal les palourdes farcies, on s'attend à ce qu'il soit plus habile à préparer le steak tartare que les cuisses de grenouille, et donc $d \succ c$. Ainsi, les deux premiers critères ne sont pas préférentiellement indépendant du troisième.

Supposons maintenant l'existence d'un opérateur d'agrégation $M : E^n \to \mathbb{R}$ qui représente $\succeq$, c'est-à-dire tel que

$$a \succ b \quad \Leftrightarrow \quad M(x^a) > M(x^b),$$

pour tout $a, b \in A$. Un tel opérateur est appelé une fonction d'utilité en MAUT.

On sait [5, 22] que l'indépendance préférentielle mutuelle parmi les critères est une condition nécessaire (mais pas suffisante) pour que l'opérateur d'agrégation M soit additif, c'est-à-dire de la forme:

$$M(x) = \sum_{i=1}^{n} u_i(x_i),$$

où les fonctions $u_i : E \to \mathbb{R}$ sont définies à une transformation linéaire positive près.

En d'autres termes, si certains critères sont préférentiellement dépendants des autres, alors aucun opérateur d'agrégation additif ne peut modéliser les préférences du décideur. En particulier, ceci exclut l'utilisation de la moyenne arithmétique pondérée.

Notons que, dans l'exemple 2.1, l'intégrale de Choquet est capable de représenter les préférences exprimées par le décideur, voir [14] pour plus de détails.

3 Hypothèses préliminaires

Pour motiver l'utilisation de l'intégrale de Choquet en tant qu'opérateur d'agrégation, nous adopterons une approche axiomatique, basée sur quelques propriétés bien connues: *croissance*, *idempotence*, et *stabilité par rapport aux mêmes échelles d'intervalle*. Ces propriétés peuvent être désirables, voire requises dans beaucoup de situations pratiques.

3.1 Croissance

Définition 3.1 (In) $M : E^n \to \mathbb{R}$ est croissant (sur chaque argument) si, pour tout $x, x' \in E^n$, nous avons

$$x_i \le x'_i \quad \forall i \in N \quad \Rightarrow \quad M(x) \le M(x').$$

Un opérateur d'agrégation croissant présente une réponse non négative à tout accroissement sur un argument. En d'autres termes, accroître un score partiel ne peut faire décroître le résultat.

3.2 Idempotence

Au départ, l'intégrale de Choquet est l'intégrale d'une fonction réelle par rapport à une measure floue, par analogie à l'intégrale de Lebesgue qui est définie par rapport à une mesure ordinaire (c'est-à-dire additive). Comme l'intégrale d'une fonction représente d'une certaine façon sa valeur moyenne, l'intégrale de Choquet discrète peut être vue comme un opérateur de type moyenne.

Cauchy [3] considérait en 1821 la moyenne de n variables indépendantes $x_1, \ldots, x_n$ comme une fonction $M(x_1, \ldots, x_n)$ qui doit être interne à l'ensemble des valeurs x_i:

$$\min x_i \le M(x_1, \ldots, x_n) \le \max x_n. \tag{2}$$

De telles moyennes vérifient trivialement la propriété d'*idempotence*, à savoir, si tous les x_i sont identiques, $M(x_1, \ldots, x_n)$ restitue la valeur commune.

Définition 3.2 (Id) $M : E^n \to \mathbb{R}$ est idempotent si

$$M(x, \ldots, x) = x, \qquad x \in E.$$

Cette propriété semble assez naturelle. D'ailleurs, on peut facilement voir que, pour les opérateurs croissants, elle est équivalente à la propriété de Cauchy (2), et les deux propriétés sont considérées par tous les statisticiens comme étant nécessaires pour définir des moyennes et des paramètres de position.

3.3 Stabilité par rapport aux mêmes échelles d'intervalles

Supposons que chaque critère $i \in N$ est une *échelle d'intervalle* (voir par ex. [11]), c'est-à-dire un homomorphisme $x_i : A \to \mathbb{R}$ défini à une transformation linéaire positive près, $\phi_i(x_i) = r_i\, x_i + s_i$ avec $r_i > 0$. Par exemple, les cotes obtenues par des étudiants dans un cours définissent souvent une échelle d'intervalle dans ce sens qu'il est possible d'exprimer ces cotes sur une échelle $[0, 20]$ ou sur une échelle $[-1, 1]$ tout en représentant la même information.

Il est clair que l'agrégation des scores partiels d'une alternative donnée sur tous les critères n'a aucun sens si les critères ne représentent pas la même échelle. Par exemple, supposons qu'un étudiant ait passé deux examens: statistique et algèbre (les cotes obtenues sont données sur une échelle de 0 à 20). Les examinateurs correspondants ont observé que l'étudiant est excellent, si bien que le premier lui donne 20 en statistique, et le second, qui est moins tolérant, lui donne 18 en algèbre. Dans ce cas, les examinateurs ont utilisé des échelles d'intervalle différentes, bornées supérieurement par 20 et 18 respectivement.

D'un point de vue théorique, on sait que l'agrégation de valeurs définies sur des échelles d'intervalle indépendantes conduit à une agrégation de type dictatorial. En effet, en supposant que l'ensemble des valeurs agrégées définisse aussi une échelle d'intervalle, un opérateur d'agrégation approprié $M : \mathbb{R}^n \to \mathbb{R}$ devrait satisfaire l'équation fonctionnelle suivante (voir [1, 2])

$$M(r_1\, x_1 + s_1, \ldots, r_n\, x_n + s_n) = R(r, s)\, M(x) + S(r, s)$$

pour tout $x, s \in \mathbb{R}^n$, $r \in]0, +\infty[^n$, $R(r, s) > 0$, et $S(r, s) \in \mathbb{R}$. Or, les solutions de cette équation sont de la forme (voir [2, cas #11])

$$M(x) = a\, x_j + b \qquad (x \in \mathbb{R}^n),$$

où $j \in N$ et $a, b \in \mathbb{R}$. Sous l'hypothèse **Id**, ces solutions deviennent simplement $M(x) = x_j$ avec $j \in N$.

En conséquence, nous supposerons que tous les scores partiels sont donnés selon une même échelle d'intervalle, de telle sorte que tout score partiel sur un critère puisse être comparé à tout autre sur un autre critère. Les critères sont alors dit *commensurables*, et l'opérateur M doit vérifier une équation moins restrictive, à savoir

$$M(r\, x_1 + s, \ldots, r\, x_n + s) = R(r, s)\, M(x) + S(r, s)$$

pour tout $x \in \mathbb{R}^n$, $r > 0$, $s \in \mathbb{R}$, $R(r, s) > 0$, et $S(r, s) \in \mathbb{R}$. De plus, si on ajoute **Id**, on aura nécessairement $R(r, s) = r$ et $S(r, s) = s$ pour tout $r > 0$ et $s \in \mathbb{R}$. En effet, pour tout $x \in \mathbb{R}$, on a

$$\begin{aligned} r\, x + s &= M(r\, x + s, \ldots, r\, x + s) \\ &= R(r, s)\, M(x, \ldots, x) + S(r, s) \\ &= R(r, s)\, x + S(r, s). \end{aligned}$$

Définition 3.3 (SPL) $M : \mathbb{R}^n \to \mathbb{R}$ est stable pour les mêmes transformations linéaires positives si

$$M(r\, x_1 + s, \ldots, r\, x_n + s) = r\, M(x) + s$$

pour tout $x \in \mathbb{R}^n$, $r > 0$, $s \in \mathbb{R}$.

Nous voyons clairement que tout opérateur $M : \mathbb{R}^n \to \mathbb{R}$ vérifiant **SPL** satisfait **Id** également. De plus, nous avons le résultat suivant.

Proposition 3.1 *Tout opérateur d'agrégation $M : \mathbb{R}^n \to \mathbb{R}$ vérifiant* **SPL** *est complètement défini par sa restriction à $[0, 1]^n$.*

Ainsi, si le profil x est exprimé dans un intervalle $[\alpha, \beta]^n$, alors son représentant x' dans $[0, 1]^n$, défini par

$$x'_i = \frac{x_i - \alpha}{\beta - \alpha}, \qquad i \in N,$$

a pour score global $M(x')$ et, par **SPL**, on a

$$M(x) = (\beta - \alpha)\, M(x') + \alpha.$$

4 L'intégrale de Choquet

Dans cette section, nous nous proposons de présenter l'intégrale de Choquet d'une manière intuitive. De plus, nous proposons une caractérisation axiomatique pour motiver l'utilisation de cet opérateur dans les applications.

4.1 L'utilisation des mesures floues

Comme nous l'avons mentionné dans les deux premières sections, les mesures floues sont capables de modéliser la dépendance entre les critères dans beaucoup de situations, quelque soit la nature de la dépendance. En fait, ces mesures floues ont été proposées par Sugeno en 1974 [25] pour généraliser les mesures additives. Il semble maintenant acquis que l'additivité des fonctions d'ensemble n'est pas toujours une propriété requise dans les situations réelles, particulièrement en présence de raisonnements humains. Pour pouvoir exprimer la subjectivité humaine, Sugeno a proposé de remplacer la propriété d'additivité par une propriété plus faible: la croissance, et il a appelé ces mesures croissantes non additives *mesures floues*. Rappelons-en la définition dans le cas discret.

Définition 4.1 Une *mesure floue* sur N est une fonction d'ensemble $v : 2^N \to [0, 1]$ vérifiant les conditions suivantes:

i) $v(\emptyset) = 0, v(N) = 1,$

ii) $S \subseteq T \;\Rightarrow\; v(S) \leq v(T).$

Dans la suite, l'ensemble de toutes les mesures floues sur N sera dénoté $\mathcal{F}_N$.
Pour tout $S \subseteq N$, $v(S)$ peut être interprété comme le poids d'importance de la combinaison S de critères, ou mieux encore, son importance ou pouvoir de prendre seule la décision (sans les autres critères). La croissance signifie alors que l'importance d'une combinaison ne peut diminuer lorsqu'on lui ajoute un élément. Evidemment, $v(N)$ est maximal, égal à 1 par convention.
Nous supposerons toujours que les poids sont des valeurs numériques définies sur une échelle cardinale. En particulier, des expressions comme $v(S) + v(T)$ ou $v(T \cup i) - v(T)$ peuvent toujours être interprétées.
Une mesure floue $v \in \mathcal{F}_N$ est dite *additive* si $v(S \cup T) = v(S) + v(T)$ chaque fois que $S \cap T = \emptyset$. Dans ce cas, il suffit de définir les n coefficients (poids) $v(1), \ldots, v(n)$ pour définir la mesure complètement.
Lorsque la mesure floue n'est pas additive, certains critères interagissent. En revenant à l'exemple de la section 2.1, nous voyons que, puisque statistique (St) et probabilité (Pr) se recouvrent, le poids des deux cours pris ensemble devrait être inférieur à la somme des poids pris séparément:

$$v(\mathrm{St}, \mathrm{Pr}) < v(\mathrm{St}) + v(\mathrm{Pr}). \tag{3}$$

4.2 Définition et approche intuitive

Le concept d'intégrale de Choquet a été d'abord introduit en théorie des capacités [4]. Son utilisation comme intégrale (floue) par rapport à une mesure floue a ensuite été proposée par Murofushi et Sugeno [17, 18]. Puisque cette intégrale est considérée ici comme un opérateur d'agrégation, nous adopterons une notation de type “connecteur” au lieu d'utiliser la notation classique des intégrales.

Définition 4.2 Soit $v \in \mathcal{F}_N$. L'*intégrale de Choquet* de $x : N \to \mathbb{R}$ par rapport à v est défini par

$$\mathcal{C}_v(x) := \sum_{i=1}^{n} x_{(i)}\,[v(A_{(i)}) - v(A_{(i+1)})], \tag{4}$$

où $(\cdot)$ représente une permutation telle que $x_{(1)} \leq \ldots \leq x_{(n)}$. De plus, $A_{(i)} = \{(i), \ldots, (n)\}$, et $A_{(n+1)} = \emptyset$.

Par exemple, si $x_3 \leq x_1 \leq x_2$, nous avons

$$\begin{aligned} \mathcal{C}_v(x_1, x_2, x_3) \quad = \quad & x_3\,[v(3,1,2) - v(1,2)] \\ & + x_1\,[v(1,2) - v(2)] \\ & + x_2\,v(2). \end{aligned}$$

L'intégrale de Choquet présente un certain lien avec celle de Lebesgue (moyenne arithmétique pondérée) puisque les deux coïncident lorsque la mesure est additive:

$$\mathcal{C}_v(x) = \sum_{i=1}^{n} v(i)\,x_i, \qquad x \in \mathbb{R}^n.$$

Dans ce sens, l'intégrale de Choquet est une généralisation de celle de Lebesgue.
Dans l'annexe, nous établissons une connexion entre l'intégrale de Choquet et l'extension de Lovász d'une fonction pseudo-booléenne, un concept utilisé en optimisation combinatoire [24]. Une telle connexion permet d'obtenir une interprétation géométrique du graphe de l'intégrale de Choquet.

Passons à présent à une présentation intuitive de l'intégrale de Choquet. Etant donné $v \in \mathcal{F}_N$, nous recherchons en fait un opérateur d'agrégation convenable $M_v : \mathbb{R}^n \to \mathbb{R}$ qui généralise la moyenne arithmétique pondérée dans le sens qu'il s'identifie à cette dernière dès que v est additif.
D'abord, nous observons facilement que toute mesure floue $v \in \mathcal{F}_N$ peut s'exprimer d'une manière unique comme:

$$v(S) = \sum_{T \subseteq S} a(T), \qquad S \subseteq N, \tag{5}$$

où $a(T) \in \mathbb{R}$ pour chaque $T \subseteq N$. En combinatoire, a, vu comme une fonction d'ensemble sur N, est appelé la *transformée de Möbius* de v (voir par ex. Rota [19]). Elle est donnée par

$$a(S) = \sum_{T \subseteq S} (-1)^{s-t}\,v(T), \qquad S \subseteq N,$$

où $s = |S|$ et $t = |T|$.
Par exemple, on a $a(\emptyset) = 0$ et $a(i) = v(i)$ pour tout $i \in N$. Pour une paire de critères $i, j \in N$, $a(ij)$ représente la différence entre le poids de la paire $\{i, j\}$ et la somme des poids de i et de j:

$$a(ij) = v(ij) - [v(i) + v(j)].$$

Cette différence est positive (resp. négative) en cas d'interaction positive (resp. négative). Elle est nulle si i et j s'additionnent sans présenter d'interaction. Ainsi, $a(ij)$ reflète quelque peu le degré d'interaction entre i et j. En fait, une définition appropriée de l'interaction a été proposée. Nous la présenterons à la section 5.
Nous pouvons aussi observer que si v est additif alors $a(S) = 0$ pour tous les sous-ensembles $S \subseteq N$ tels que $s \geq 2$. Dans ce cas, l'opérateur d'agrégation

doit nécessairement être la moyenne arithmétique pondérée:

$$M_v(x) = \sum_{i\in N} v(i)\, x_i = \sum_{i\in N} a(i)\, x_i.$$

Lorsque v n'est pas additif, nous devons prendre en compte l'interaction parmi les critères. Nous pouvons alors partir de la moyenne arithmétique pondérée, qui est une expression linéaire, et lui ajouter des termes du "second ordre" faisant intervenir les coefficients correcteurs $a(ij)$, ensuite des termes du troisième ordre, etc. On obtient alors

$$M_v(x) = \sum_{i\in N} a(i)\, x_i + \sum_{\{i,j\}\subseteq N} a(ij)\, [x_i \wedge x_j] + \ldots$$

c'est-à-dire,

$$M_v(x) = \sum_{T\subseteq N} a(T) \bigwedge_{i\in T} x_i. \tag{6}$$

Cette expression n'est rien d'autre que celle de l'intégrale de Choquet $\mathcal{C}_v$ en termes de la représentation de Möbius, voir l'annexe pour plus de détails. Notons que, dans la partie non linéaire, nous avons utilisé l'opération minimum au lieu du produit pour s'assurer que l'opérateur vérifie la propriété **SPL**.

4.3 Caractérisation axiomatique

Bien que les intégrales de Choquet soient devenues populaires en aide multicritère à la décision, il existe dans la littérature très peu de caractérisations axiomatiques de cette famille. La plus représentative est celle de Schmeidler [21], qui utilise le concept d'additivité comonotone. Malheureusement, une telle caractérisation n'est pas très attrayante dans le contexte de l'aide multicritère à la décision.

Nous présentons ici une caractérisation de la classe des intégrales de Choquet au moyen de quatre propriétés. Cette caractérisation peut être trouvée dans la thèse de doctorat de l'auteur [13, Sect. 6.1].

Définition 4.3 (LM) Les opérateurs $M_v : \mathbb{R}^n \to \mathbb{R}$ ($v \in \mathcal{F}_N$) sont *linéaires par rapport à la mesure floue* s'il existe 2^n fonctions $f_T : \mathbb{R}^n \to \mathbb{R}$ ($T \subseteq N$) telles que

$$M_v = \sum_{T\subseteq N} v(T)\, f_T, \qquad v \in \mathcal{F}_N.$$

Cette définition est motivée par l'observation suivante. Nous savons que l'opérateur $M_v(x)$ que nous voulons caractériser n'est pas linéaire par rapport à son argument $x \in \mathbb{R}^n$. Néanmoins, nous pouvons lui demander d'être linéaire au moins par rapport à la mesure floue v, et ceci, de manière à garder le modèle d'agrégation aussi simple que possible. C'est une hypothèse assez naturelle.

Puisque les formules de conversion entre v et a sont linéaires, M_v vérifie **LM** si et seulement s'il existe 2^n fonctions $g_T : \mathbb{R}^n \to \mathbb{R}$ ($T \subseteq N$) telles que

$$M_v = \sum_{T\subseteq N} a(T)\, g_T, \qquad v \in \mathcal{F}_N. \tag{7}$$

Définissons $v_T \in \mathcal{F}_N$ par $v_T(S) = 1$ si et seulement si $S \supseteq T$, et 0 sinon. En théorie des jeux, v_T est appelé le *jeu unanime* pour T. La représentation de Möbius de v_T est donnée par $a_T(S) = 1$ si et seulement si $S = T$, et 0 sinon. De (7) il découle immédiatement que

$$g_T = M_{v_T}, \qquad T \subseteq N.$$

Pour identifier M_v à l'intégrale de Choquet $\mathcal{C}_v$, nous devons encore imposer que (voir (6))

$$M_{v_T}(x) = \bigwedge_{i\in T} x_i, \qquad x \in \mathbb{R}^n.$$

Ceci peut se faire en utilisant **In** et **SPL**, mais aussi en donnant une définition appropriée des poids des sous-ensembles de critères. Une telle définition n'a pas encore été précisée jusqu'ici. En fait, l'interprétation de $v(S)$ comme le poids d'importance du sous-ensemble S devrait apparaître clairement dans la définition de l'intégrale de Choquet.

En vertu de la proposition 3.1, nous pouvons faire l'hypothèse que tous les scores partiels sont donnés dans $[0, 1]$. Pour la moyenne arithmétique pondérée (1), nous avons trivialement

$$M_\omega(e_i) = \omega_i, \qquad i \in N,$$

ce qui montre que le poids ω_i du critère i est en fait l'évaluation globale de l'alternative qui satisfait pleinement le critère i et qui ne satisfait pas les autres du tout. Plus généralement, nous avons

$$M_\omega(e_S) = \sum_{i\in S} \omega_i, \qquad S \subseteq N,$$

ce qui indique que le poids d'un groupe de critères indépendants est défini comme l'évaluation globale de l'alternative qui satisfait complètement les critères S et qui ne satisfait en rien les autres critères. En adaptant cette observation au cas des critères dépendants, nous proposons la définition suivante.

Définition 4.4 (PW) Soit $v \in \mathcal{F}_N$. $M_v : \mathbb{R}^n \to \mathbb{R}$ est *proprement pondéré par* v si $M_v(e_S) = v(S)$ pour tout $S \subseteq N$.

En revenant à l'exemple de la section 2.1, nous définissons le poids de la paire "statistique et probabilité" par l'évaluation globale d'un étudiant qui présenterait le profil $(1, 1, 0)$ dans $[0, 1]^3$:

$$M_v(1, 1, 0) = v(\mathrm{St}, \mathrm{Pr}).$$

Si les critères étaient non corrélés, nous aurions

$$M_v(1, 1, 0) = v(\mathrm{St}) + v(\mathrm{Pr}),$$

ce qui montre bien que (3) est le résultat d'une corrélation positive.
Notons que, dans certaines applications pratiques, des alternatives qui présentent de tels profils binaires sont quelque peu imaginaires, voire impensables. Dans ce cas, une définition équivalente et plus intuitive consiste à définir $v(S)$ comme le score global le plus bas qu'on est prêt à accorder à une alternative qui satisfait pleinement les critères S, sans connaissance préalable des scores sur les autres critères. Formellement, nous posons

$$v(S) := \min_{x\in[0,1]^n} M_v(e_N S x).$$

Bien sûr, l'équivalence entre les deux définitions découle immédiatement de **In**.
Notons cependant que l'utilisation de cette définition alternative peut conduire à certains paradoxes. Par exemple, si nous savons que statistique et probabilité sont fortement corrélés, nous aurons tendance à donner un score global plus important si l'étudiant présente le profil $(?, 1, 1)$, dans lequel le premier score est inconnu, que s'il présente le profil $(0, 1, 1)$.

Nous pouvons maintenant énoncer la caractérisation de la classe de toutes les intégrales de Choquet à n arguments:

Théorème 4.1 *Les opérateurs* $M_v : \mathbb{R}^n \to \mathbb{R}$ ($v \in \mathcal{F}_N$) *vérifient* **LM**, **In**, **SPL**, **PW** *si et seulement si* $M_v = \mathcal{C}_v$ *pour tout* $v \in \mathcal{F}_N$.

5 Analyse du comportement de l'agrégation

Maintenant que nous avons à notre disposition un agrégateur approprié, une question importante se pose. Comment interpréter le comportement de l'intégrale de Choquet ou celui de sa mesure floue associée? Bien sûr, la signification des valeurs $v(T)$ n'est pas toujours claire pour le décideur. Ces valeurs ne fournissent pas immédiatement l'importance globale des critères, ni même le degré d'interaction parmi eux. En fait, au départ d'une mesure floue donnée, il est possible de construire des indices (ou paramètres) qui permettent d'interpréter le comportement de la mesure floue. Ces indices constituent une sorte de *carte d'identité* de la mesure floue. Dans cette section, nous présentons deux types d'indices: l'importance et l'interaction. D'autres indices, tels que la tolerance et la dispersion, ont été proposés et étudiés par l'auteur dans [13].

5.1 Indices d'importance

L'importance globale d'un critère $i \in N$ dans un problème de décision n'est pas déterminé uniquement par les nombres $v(i)$, mais aussi par tous les $v(T)$ tels que $i \in T$. En effet, nous pouvons avoir $v(i) = 0$, suggérant que l'élément i est sans importance, mais il peut arriver que pour beaucoup de sous-emsembles $T \subseteq N$, le nombre $v(T \cup i)$ soit beaucoup plus grand que $v(T)$, suggérant que i est en fait un élément très important dans la décision.
Shapley [23] a proposé en 1953 la definition d'un coefficient d'importance, basée sur un ensemble d'axiomes raisonnables. L'*indice d'importance* ou *valeur de Shapley* du critère i par rapport à v est défini par:

$$\phi(v, i) := \sum_{T\subseteq N\setminus i} \frac{(n-t-1)!\, t!}{n!} \,[v(T \cup i) - v(T)]. \quad (8)$$

La valeur de Shapley est un concept fondamental en théorie des jeux qui exprime un indice de pouvoir. Elle peut être interprétée comme une valeur moyenne pondérée de la contribution marginale $v(T \cup i) - v(T)$ de l'élément i seul dans toutes les combinaisons. Cette observation est plus évidente encore lorsqu'on réécrit l'indice comme ceci:

$$\phi(v, i) = \frac{1}{n} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T\subseteq N\setminus i\\ |T|=t}} [v(T \cup i) - v(T)].$$

Ainsi, la valeur moyenne de $v(T \cup i) - v(T)$ est d'abord calculée sur les sous-ensembles de même taille t et ensuite sur toutes les tailles possibles.
L'utilisation de la valeur de Shapley en aide multicritère à la décision a été proposée en 1992 par Murofushi [15]. Il faut noter qu'une propriété fondamentale de la valeur de Shapley est

$$\sum_{i=1}^{n} \phi(v, i) = 1.$$

Notons aussi que, lorsque v est additif, nous avons clairement $v(T \cup i) - v(T) = v(i)$ pour tout $i \in N$ et tout $T \subseteq N \setminus i$, et donc

$$\phi(v, i) = v(i), \quad i \in N. \quad (9)$$

Si v n'est pas additif alors certain critères sont dépendant et l'égalité (9) n'est généralement plus valable. Ceci montre qu'il est raisonnable de définir un coefficient d'importance global pour chaque critère.
En termes de la représentation de Möbius, la valeur de Shapley a une forme très simple [23]:

$$\phi(v,i) = \sum_{T \ni i} \frac{1}{t}\, a(T). \tag{10}$$

5.2 Indices d'interaction

Un autre concept intéressant est celui de l'*interaction* parmi les critères. Nous avons vu que, lorsque la mesure floue n'est pas additive, certains critères interagissent. Bien sûr, il serait intéressant d'évaluer le degré d'interaction parmi n'importe quel sous-ensemble de critères.
Considérons tout d'abord une paire de critères $\{i,j\} \subseteq N$. Il peut arriver que $v(i)$ et $v(j)$ soient petits et qu'en même temps $v(ij)$ soit grand. Clairement, le nombre $\phi(v,i)$ mesure simplement la contribution moyenne que le critère i apporte à toutes les combinaisons possibles, mais il n'explique pas pourquoi le critère i peut avoir une grande importance. En d'autres termes, il ne donne aucune information sur les phénomènes d'interaction existant parmi les critères.
Nous avons vu à la section 2.1 que selon que la correlation entre i et j est ≤ 0 ou ≥ 0, l'expression

$$v(T \cup ij) - v(T \cup i) - v(T \cup j) + v(T)$$

est ≥ 0 ou ≤ 0 pour tout $T \subseteq N \setminus ij$, respectivement. Nous appellerons cette expression l'interaction marginale entre i et j, conditionné à la présence des éléments de la combinaison $T \subseteq N \setminus ij$. Cela étant, un indice d'interaction pour $\{i,j\}$ est donné par une valeur moyenne de cette interaction marginale. Murofushi and Soneda [16] ont proposé en 1993 de calculer cette valeur moyenne de la même façon que pour la valeur de Shapley. En posant

$$(\Delta_{ij}\, v)(T) := v(T \cup ij) - v(T \cup i) - v(T \cup j) + v(T),$$

l'*indice d'interaction* des critères i et j relatif à v est alors défini par

$$I(v,ij) = \sum_{T \subseteq N \setminus ij} \frac{(n-t-2)!\, t!}{(n-1)!}\, (\Delta_{ij}\, v)(T).$$

Nous voyons immédiatement que cet indice est négatif dès que i et j sont positivement corrélés ou interchangeables. De même, il est positif lorsque i et j sont négativement corrélés ou complémentaires. De plus, il a été démontré dans [7] que $I(v,ij) \in [-1,1]$ pour tout $i,j \in N$.

L'indice d'interaction parmi une combinaison S de critères a été introduit par Grabisch [7] comme une extension naturelle du cas $s=2$. L'*indice d'interaction* de S ($s \geq 2$) relatif à v, est défini par

$$I(v,S) := \sum_{T \subseteq N \setminus S} \frac{(n-t-s)!\, t!}{(n-s+1)!}\, (\Delta_S\, v)(T),$$

où nous avons posé

$$(\Delta_S\, v)(T) := \sum_{L \subseteq S} (-1)^{s-l} v(L \cup T).$$

En termes de la représentation de Möbius, cet indice s'écrit [7]

$$I(v,S) = \sum_{T \supseteq S} \frac{1}{t-s+1}\, a(T), \qquad S \subseteq N. \tag{11}$$

Vu comme une fonction d'ensemble, il coïncide sur les singletons avec la valeur de Shapley (8).
Il a été démontré dans [8] que la transformation (11) est inversible et son inverse s'écrit:

$$a(S) = \sum_{T \supseteq S} B_{t-s}\, I(v,T), \qquad S \subseteq N, \tag{12}$$

où B_n est le n-ième nombre de Bernoulli, c'est-à-dire le n-ième élément de la suite numérique $\{B_n\}_{n \in \mathbb{N}}$ définie récursivement par

$$\begin{cases} B_0 = 1, \\ \sum\limits_{k=0}^{n} \binom{n+1}{k} B_k = 0, \qquad n \in \mathbb{N} \setminus \{0\}. \end{cases}$$

Annexe : Fonctions pseudo-booléennes et extensions de Lovász

Toute fonction d'ensemble $v : 2^N \to \mathbb{R}$ peut être assimilée sans ambiguïté à une fonction pseudo-booléenne $f : \{0,1\}^n \to \mathbb{R}$. Les formules de correspondance s'écrivent

$$f(x) = \sum_{T \subseteq N} v(T) \prod_{i \in T} x_i \prod_{i \notin T} (1-x_i), \qquad x \in \{0,1\}^n,$$

et $v(S) = f(e_S)$ pour tout $S \subseteq N$.
En particulier, toute fonction pseudo-booléenne qui correspond à une mesure floue est croissante sur chaque variable et vérifie les conditions limites: $f(e_\emptyset) = 0$ et $f(e_N) = 1$.

Hammer et Rudeanu [9] ont montré que toute fonction pseudo-booléenne peut s'écrire de manière unique sous la forme d'un polynôme multilinéaire à n variables:

$$f(x) = \sum_{T \subseteq N} a(T) \prod_{i \in T} x_i, \qquad x \in \{0,1\}^n,$$

avec $a(T) \in \mathbb{R}$ pour tout $T \subseteq N$. Ces coefficients $a(T)$ correspondent à la transformée de Möbius de v, voir (5).

Lovász [12, Sect. 3] a observé que tout $x \in (\mathbb{R}^+)^n \setminus \{e_\emptyset\}$ peut être écrit de manière unique sous la forme

$$x = \sum_{i=1}^{k} \lambda_i \, e_{S_i} \tag{13}$$

où $\lambda_1, \ldots, \lambda_k > 0$ et $\emptyset \neq S_1 \subsetneq \cdots \subsetneq S_k \subseteq N$. Donc, toute fonction $f : \{0,1\}^n \to \mathbb{R}$ avec $f(e_\emptyset) = 0$ peut être étendue à $\hat{f} : (\mathbb{R}^+)^n \to \mathbb{R}$, par $\hat{f}(e_\emptyset) = 0$ et

$$\hat{f}(x) = \sum_{i=1}^{k} \lambda_i \, f(e_{S_i}) \quad \big(x = \sum_{i=1}^{k} \lambda_i \, e_{S_i} \in (\mathbb{R}^+)^n \setminus \{e_\emptyset\}\big);$$

en effet, $\hat{f}$ est bien défini (cf. unicité de (13)) et $\hat{f}(x) = f(x)$ pour tout $x \in \{0,1\}^n$. La fonction $\hat{f}$ est appelée l'*extension de Lovász* de f.
L'extension de Lovász d'une fonction arbitraire $f : \{0,1\}^n \to \mathbb{R}$ est alors définie par

$$\hat{f}(x) = f(e_\emptyset) + \hat{f}_0(x), \qquad x \in (\mathbb{R}^+)^n,$$

où $\hat{f}_0$ est l'extension de Lovász de $f_0 = f - f(e_\emptyset)$.

Le cube $[0,1]^n$ peut être subdivisé en $n!$ simplexes $\mathcal{B}_\pi$ de la forme

$$\mathcal{B}_\pi := \{x \in [0,1]^n \,|\, x_{\pi(1)} \leq \cdots \leq x_{\pi(n)}\}, \qquad \pi \in \Pi,$$

où Π est l'ensemble des permutations de N.
Il est bien connu [24] que $\mathcal{B}_\pi$ est un polytope ayant pour sommets $\varepsilon_i^\pi = e_{\{\pi(i),\ldots,\pi(n)\}}$ $(i = 1, \ldots, n+1)$. Singer [24, Sect. 2] a montré que $\hat{f}$ est défini sur chaque cône $\mathcal{K}_\pi = \{\lambda \, \mathcal{B}_\pi \,|\, \lambda \geq 0\}$ comme l'unique fonction affine qui coïncide avec f aux $n+1$ sommets de $\mathcal{B}_\pi$. Plus formellement, $\hat{f}$ peut s'écrire

$$\hat{f}(x) = f(e_\emptyset) + \sum_{i=1}^{n} x_{\pi(i)} \, [f(\varepsilon_i^\pi) - f(\varepsilon_{i+1}^\pi)], \qquad x \in \mathcal{K}_\pi. \tag{14}$$

On a alors le résultat suivant.

Proposition A.1 *L'extension de Lovász de* $f : \{0,1\}^n \to \mathbb{R}$ *est donnée par*

$$\hat{f}(x) = \sum_{T \subseteq N} a(T) \bigwedge_{i \in T} x_i, \qquad x \in (\mathbb{R}^+)^n, \tag{15}$$

où les coefficients $a(T)$ *sont ceux de l'expression multilinéaire de* f

Soit $v \in \mathcal{F}_N$. D'après (14), on voit immédiatement que l'intégrale de Choquet $\mathcal{C}_v$, définie sur $(\mathbb{R}^+)^n$, n'est rien d'autre que l'extension de Lovász de la fonction pseudo-booléenne f qui représente v:

$$\mathcal{C}_v = \hat{f} \quad \text{on } (\mathbb{R}^+)^n.$$

De plus, puisque l'expression dans (15) vérifie **SPL** sur $\mathbb{R}^n$, nous savons en vertu de la proposition 3.1 qu'elle correspond à l'intégrale de Choquet.

Proposition A.2 *Toute intégrale de Choquet* $\mathcal{C}_v : \mathbb{R}^n \to \mathbb{R}$ *peut s'écrire*

$$\mathcal{C}_v(x) = \sum_{T \subseteq N} a(T) \bigwedge_{i \in T} x_i, \qquad x \in \mathbb{R}^n,$$

où a *est la représentation de Möbius de* v.

Ainsi, l'intégrale de Choquet $\mathcal{C}_v$ est une fonction affine par morceaux sur $\mathbb{R}^n$, qui étend la fonction pseudo-booléenne représentant v:

$$\mathcal{C}_v(e_S) = v(S), \qquad S \subseteq N.$$

References

[1] J. Aczél and F.S. Roberts, On the possible merging functions, *Math. Social Sciences* **17** (1989) 205–243.

[2] J. Aczél, F.S. Roberts and Z. Rosenbaum, On scientific laws without dimensional constants, *Journal of Math. Analysis and Appl.* **119** (1986) 389–416.

[3] A.L. Cauchy, *Cours d'analyse de l'Ecole Royale Polytechnique, Vol. I. Analyse algébrique*, (Debure, Paris, 1821).

[4] G. Choquet, Theory of capacities, *Annales de l'Institut Fourier* **5** (1953) 131–295.

[5] P.C. Fishburn, *Utility theory for decision making*, (New York, Wiley, 1970).

[6] M. Grabisch, The application of fuzzy integrals in multicriteria decision making, *European Journal of Operational Research*, **89** (1996) 445-456.

[7] M. Grabisch, k-order additive discrete fuzzy measures and their representation, *Fuzzy Sets and Systems* **92** (1997) 167–189.

[8] M. Grabisch, J.-L. Marichal and M. Roubens, Equivalent representations of a set function with applications to game theory and multicriteria decision making, Preprint 9801, GEMME, Faculty of Economics, University of Liège, Belgium, 1998; also in: Preprint 98.002, Institute of Mathematics, University of Liège, Belgium, 1998.

[9] P.L. Hammer and S. Rudeanu, *Boolean methods in operations research and related areas*, (Springer, Berlin, 1968).

[10] R.L. Keeney and H. Raiffa, *Decision with multiple objectives.* (Wiley, New York, 1976).

[11] D.H. Krantz, R.D. Luce, P. Suppes and A. Tversky, *Foundations of measurement*, volume I: *Additive and polynomial representations*, (Academic Press, San Diego, 1971).

[12] L. Lovász, Submodular function and convexity. In: *Mathematical programming. The state of the art.* Bonn 1982, Eds. A. Bachem, M. Grötschel, B. Korte. (Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1983), 235–257.

[13] J.-L. Marichal, *Aggregation operators for multicriteria decision aid*, Ph.D. Thesis, University of Liège, Liège, Belgium, 1998.

[14] J.-L. Marichal and M. Roubens, Determination of weights of interacting criteria from a reference set, *European Journal of Operational Research*, to appear.

[15] T. Murofushi, A technique for reading fuzzy measures (I): the Shapley value with respect to a fuzzy measure, in: *2nd Fuzzy Workshop*, pp. 39–48, Nagaoka, Japan, October 1992. In Japanese.

[16] T. Murofushi and S. Soneda, Techniques for reading fuzzy measures (III): interaction index, in: *9th Fuzzy System Symposium*, pp. 693–696, Sapporo, Japan, May 1993. In Japanese.

[17] 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 and Systems* **29** (1989) 201–227.

[18] T. Murofushi and M. Sugeno, A theory of fuzzy measures. Representation, the Choquet integral and null sets, *Journal of Mathematical Analysis and Applications* 159/2 (1991) 532–549.

[19] G.C. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **2** (1964) 340–368.

[20] B. Roy, *Méthodologie multicritère d'aide à la décision*, (Economica, 1985).

[21] D. Schmeidler, Integral representation without additivity, *Proc. Amer. Math. Soc.* **97** (1986) 255–261.

[22] D. Scott and P. Suppes, Foundational aspects of theories of measurement, *J. Symbolic Logic* **23** (1958) 113–128.

[23] L.S. Shapley, A value for n-person games. In: H.W. Kuhn and A.W. Tucker (eds.), *Contributions to the Theory of Games*, Vol. II, Annals of Mathematics Studies, 28, (Princeton University Press, Princeton, NJ, 1953), 307–317.

[24] I. Singer, Extensions of functions of 0-1 variables and applications to combinatorial optimization, *Numerical Functional Analysis and Optimization* 7(1) (1984-85) 23–62.

[25] M. Sugeno, Theory of fuzzy integrals and its applications, Ph.D. Thesis, Tokyo Institute of Technology, Tokyo, 1974.

[26] P. Wakker, *Additive representations of preferences. A new foundation of decision analysis*, (Kluwer Academic Publishers, 1989).