The application of fuzzy integrals in Multicriteria Decision Making

Michel GRABISCH
Thomson-CSF, Central Research Laboratory
Domaine de Corbeville, 91404 Orsay cedex, France

Abstract This paper presents a synthesis on the application of fuzzy integral as an innovative tool for criteria aggregation in decision problems. The main point is that fuzzy integrals are able to model interaction between criteria in a flexible way. The methodology has been elaborated mainly in Japan, and has been applied there successfully in various fields such as design, reliability, evaluation of goods, etc. It seems however that this technique is still very little known in Europe. It is one of the aim of this review to disseminate this emerging technology in many industrial fields.

Keywords: multicriteria decision making, aggregation operator, fuzzy measure, fuzzy integral, interactive criteria

1 Introduction: history and motivations

The theory of multicriteria decision making has witnessed the birth of many new paradigms in this second half of the twentieth century, especially in Europe, where several research groups have proposed alternatives to the well known and sovereign multiattribute utility model: the outranking procedure, represented by the successive ELECTRE methods, is one of the most typical example of this recent tendency, PROMETHEE, ORESTE should also be mentionned, as well as all the research undertaken in the fuzzy logic community [3].

Nevertheless, a common feature to all these approaches in multicriteria decision making, including the classical multiattribute utility (MAUT) model, is that we need somewhere a fundamental operation which is *aggregation*. In MAUT, we aggregate uni-dimensional utility functions into a single global utility function combining all the criteria, while in ELECTRE, we have to aggregate preference relations on pair of alternatives.

It can be said that, within all the above mentionned paradigms, the effort, or the innovation, has been much more on the modeling of what could be called the choice procedure, rather than on the aggregation step. In fact, there are yet few available —well-rooted in the theory— tools for aggregation, and the most common aggregation tool which is used today is still the weighted arithmetic mean (or weighted sum), with all its well-known drawbacks.

However, some twenty years ago in Japan was proposed the concept of fuzzy integral by Sugeno [22, 23], which in the discrete case is merely a kind of distorded mean. Although this was followed by a rather mathematically oriented research, far from application concerns, some Japanese researchers, including Sugeno himself, thought that fuzzy integrals could be applied to multicriteria evaluation: since 1985, papers have been published on wood quality evaluation [11], analysis of public attitude towards the use of nuclear energy [20], evaluation of printed color images [27], design of speakers [10], analysis of human reliability [28], etc. These applications as well as others will be presented briefly in section 5.

The distinguishing feature of fuzzy integral is that it is able to represent a certain kind of interaction between criteria, ranging from redundancy (negative interaction) to synergy (positive interaction). To our knowledge, there is almost no well established method to deal with interacting criteria, and usually people tend to avoid the problem by constructing independent (or supposed to be so) criteria. This innovative feature was without any doubt the reason of its success in various fields of application.

An analysis of the above mentionned works shows a rather intuitive and *ad hoc* approach, at least on a multicriteria decision making point of view: the justification of the methodology with respect to some axiomatic system for aggregation, its relation with existing techniques, in particular MAUT, the properties for aggregation shared by fuzzy integrals, were not addressed. Moreover, the precise definition of what was understood by “interaction”, “support” or “synergy” was lacking. It is only recently that these issues have been —partially— solved, essentially by Murofushi *et al.* [15, 16, 18], and the author of this review [6, 7]. Sections 2 and 3 will present the main results.

However, the richness of fuzzy integrals has to be paid by the complexity of the model, since the number of coefficients involved in the fuzzy integral model grows exponentially with the number of criteria to be aggregated. The main difficulty is to identify all these coefficients, either by some learning data, or by questionnaire, or both. It is well known that even the identification of the weights in a weighted arithmetic mean model is not a trivial job [14]. Coming back to the fuzzy integral, we think that this problem is not yet solved in a fully satisfactory way, but some operational solutions already exist. They will be presented in section 4.

This review aims at presenting to the European Operational Research Community the methodology of fuzzy integrals for solving multicriteria decision problems, since it seems that this innovative technique has not yet spread out of the Japanese archipelago.

2 Multicriteria decision making by fuzzy integral

In this section we briefly present the multicriteria aggregation problem, its requirements and usual solutions, and introduce fuzzy integrals as a new tool for aggregation. The reader should consult [7] for a more detailed presentation of this section.

2.1 The multicriteria decision making problem

We choose the MAUT framework for ease of presentation, although this is not limitative. We suppose to have a set of *alternatives* or *acts* $\Omega = \{\omega_1, \omega_2, \ldots, \omega_p\}$, among which the decision maker must choose. Each act is associated with a vector $x = [x_1 \cdots x_n]^T$ of *attributes* or *criteria*, which represents what the decision maker will get if he chooses this alternative; for this reason x is called the *consequence* or the *result* in MAUT. Let X be the set of consequences. The preferences over X of the decision maker are expressed by a weak order relation $\succeq$. Then the basic idea of MAUT is to construct a *utility function* $u : X \longrightarrow \mathbb{R}$ so that the property

$$x \succ y \Leftrightarrow u(x) > u(y) \tag{1}$$

always holds. Clearly u is a n dimensional function. An easy way to construct u is to consider unidimensional utility functions u_i on each attribute, and then to *aggregate* them by a suitable operator $\mathcal{H}$.

$$u(x_1, \ldots, x_n) = \mathcal{H}(u_1(x_1), \ldots, u_n(x_n)) \tag{2}$$

“Suitable” means that the function u so constructed should verify (1). One of the simplest solution is the arithmetic sum:

$$u(x_1, \ldots, x_n) = \sum_{i=1}^{n} u_i(x_i) \tag{3}$$

or its weighted version. Such a u is called additive utility, and a lot of work has been done in order to find conditions on the preference relation $\succeq$ such that an additive utility function satisfying (1) exists . Our aim is to present an alternative solution for $\mathcal{H}$: the fuzzy integral. Before this, we try to list all the practical requirement on $\mathcal{H}$.

2.2 Requirements on the aggregation operator

In fact,the only requirement is that equation (1) is always satisfied, but this is of little use in practice. We give here a tentative list of what should be the properties of $\mathcal{H}$ for a proper aggregation of criteria. At this level, we restrict no more to the MAUT model. We replace $u_i(x_i)$ by a more anonymous notation a_i, which could represent as well a utility function for attribute i, a degree of satisfaction with respect to criterion i, or a degree of preference for a given pair of alternatives with respect to criterion i. We will consider without loss of generality that a_i ranges in $[0,1]$.

1. mathematical properties
   - properties of extremal values
     $$\mathcal{H}(0,\ldots,0)=0, \quad \mathcal{H}(1,\ldots,1)=1$$
   - idempotence
     $$\mathcal{H}(a,a,\ldots,a)=a, \forall a \in [0,1]$$
   - continuity
   - monotonicity (usually non decreasingness) with respect to each argument
   - commutativity (or neutrality) can be required if the criteria are indifferent. However, this is more natural in voting procedure than in multicriteria decision making.
   - decomposability
     $$\mathcal{H}^{(n)}(a_1,a_k,a_{k+1},\ldots,a_n)=\mathcal{H}^{(n)}(a,\ldots,a,a_{k+1},\ldots,a_n)$$
     where $a=\mathcal{H}^{(k)}(a_1,\ldots,a_k)$, for all $(a_1,\ldots,a_n)$. The superscript (n) indicates the number of arguments of $\mathcal{H}$.
   - stability under the same positive linear transformation (SPL):
     $$\mathcal{H}(ra_1+t,\ldots,ra_n+t)=r\mathcal{H}(a_1,\ldots,a_n)+t, \forall r>0, \forall t \in \mathbb{R}.$$
     This property says that changing the scale does not change the result. It is essential in utility theory since the u_i are defined up to a positive linear transformation. Thus the global utility u must keep this property.

   There are many other properties: see a good survey indicating all relations between them in [3]. Note that monotonicity and idempotence imply that $\mathcal{H}$ lies between *min* and *max*: this restricts considerably the field of suitable operators. Associativity can be required, but this is conflicting with idempotence: the only associative and idempotent operators are the medians (see below).

2. behavioral properties
   - possibility of expressing weights of importance on criteria if this is necessary.
   - possibility of expressing the behaviour of the decision maker. Two decision makers with same monodimensional utility functions u_i, same weights on criteria, could still have different preferences, because their behaviour in decision making are different. Typically, behaviour can range from perfect tolerance to total intolerance. Tolerant decision makers can accept that only *some* criteria (at least one) are met (this corresponds to a disjunctive behaviour, whose extreme example is *max*). On the other hand, intolerant decision makers demand that *all* criteria have to be equally met (conjunctive behaviour, whose extreme example is *min*).
   - possibility of expressing a compensatory effect, or an interaction between criteria. Compensation exists if a bad score on one criterion can be compensated by a good score on another criterion. Other possible interactions between criteria are *redundancy* (two criteria are redundant if they express more or less the same thing) and support or synergy (two criteria with little importance when taken separately, become very important when considered jointly). We will develop this point in section 3.
   - possibility of an easy semantical interpretation, i.e. being able to relate the values of parameters defining $\mathcal{H}$ to the behaviour implied by $\mathcal{H}$.

2.3 Common aggregation operators

We present common solutions for the aggregation step. All these operators are idempotent, continuous, and monotonically non decreasing (thus ranging between *min* and *max*), since it can be said that these are the irreducible requirements for any problem of aggregation. We call them *averaging operators* in the sequel.

quasi-arithmetic means. As we said above, the easiest way to aggregate is the simple arithmetic mean $1/n \sum_i a_i$. Many other means exist, such as geometric, harmonic means, etc. In fact all these common means belong to the family of quasi-arithmetic means, defined as follows.

$$M_f(a_1, \ldots, a_n) = f^{-1}\left[\frac{1}{n}\sum_{i=1}^{n} f(a_i)\right] \tag{4}$$

where f is any continuous strictly monotonic function. This family has been characterized by Kolmogoroff [12], as being the class of all decomposable commutative averaging operators. It is easy to extend the definition by introducing weights of importance $w_1, \ldots, w_n$ on criteria, with the constraint $\sum_{i=1}^{n} w_i = 1$.

$$M^f_{w_1,\ldots,w_n}(a_1, \ldots, a_n) = f^{-1}\left[\sum_{i=1}^{n} w_i f(a_i)\right] \tag{5}$$

These are still decomposable averaging operators. In general,they are not stable under linear transformation (except for $f = Id$), and consider criteria as non interacting.

median. The median is a typical ordinal operator, i.e. taking into account not the values a_i themselves but only their ordering. It is defined as the middle value of the ordered list.

$$\mathrm{med}(a_1, \ldots, a_n) = \begin{cases} a_{(\frac{n+1}{2})}, & \text{if } n \text{ is odd} \\ \frac{1}{2}(a_{(\frac{n}{2})} + a_{(\frac{n}{2}+1)}), & \text{if } n \text{ is even.} \end{cases} \tag{6}$$

where parentheses () around the index show that elements have been arranged in increasing order, i.e. $a_{(1)} \leq a_{(2)} \leq \cdots \leq a_{(n)}$. Weights of importance on criteria can be modelled by duplicating the corresponding elements in the list. *Order statistics*, which give as output the kth value of the ordered list, are a generalisation of the median.

weighted minimum and maximum. They have been introduced in [2] in the framework of possibility theory, and are a generalisation of *min* and *max*. They are defined as follows ($\wedge$ and $\vee$ denote *min* and *max* respectively).

$$\mathrm{wmin}_{w_1,\ldots,w_n}(a_1, \ldots, a_n) = \bigwedge_{i=1}^{n} [(1 - w_i) \vee a_i] \tag{7}$$

$$\mathrm{wmax}_{w_1,\ldots,w_n}(a_1, \ldots, a_n) = \bigvee_{i=1}^{n} [w_i \wedge a_i] \tag{8}$$

where weights are normalized so that $\vee_{i=1}^{n} w_i = 1$.

ordered weighted averaging operators (OWA). They have been introduced by Yager [29, 30].

$$\mathrm{OWA}_{w_1,\ldots,w_n}(a_1, \ldots, a_n) = \sum_{i=1}^{n} w_i a_{(i)} \tag{9}$$

where $\sum_{i=1}^{n} w_i = 1$, and $a_{(1)} \leq a_{(2)} \leq \cdots \leq a_{(n)}$ as above. Taking all weights equal to $1/n$ leads to the arithmetic mean, while the *min* operator (resp. *max*, the median, order statistics) can be recovered by taking all weights equal to 0 except $w_1 = 1$ (resp. $w_n = 1$, $w_{\frac{n+1}{2}} = 1$, $w_k = 1$). They are stable under positive linear transformation and commutative. The arithmetic mean is the only common point between OWA and weighted sums: in fact, these are very different operators, which can be said to be "orthogonal" in an intuitive sense. Their principal interest lies in the fact that they can express vague quantifiers, as for example: *"at least some criteria must be met"*, which can be modeled by the following weights: $w_1 = 0$, $w_2 = 0$, $w_3 = 0.8$, $w_4 = 0.2$, $w_5 = 0$, when $n = 5$.

Many examples from above come from the field of fuzzy logic. The reader interested in more details, analysis, comparisons and classifications of family of operators can consult references [1, 3, 30].

All these operators present some drawbacks. On the one hand, large families do not possess all desirable properties (e.g. quasi-arithmetic means are not stable under positive linear transformation), and on the other hand, small families seem to be too restrictive (arithmetic sums, OWA, etc). Moreover, no one is able to model in some understandable way an interaction between criteria.

We introduce in the next section fuzzy integrals, which constitute a new large family of aggregation operators, without these drawbacks.

2.4 Fuzzy measures and integrals

For space limitation, we will restrict definitions and mathematical complexity to the strict minimum. In particular, we consider discrete spaces only. More complete definitions can be found in [9, 17].

Let us denote by $X = \{x_1, \ldots, x_n\}$ the set of criteria, and by $\mathcal{P}(X)$ the power set of X, i.e. the set of all subsets of X.

Definition 1 *A* fuzzy measure *on the set X of criteria is a set function $\mu : \mathcal{P}(X) \longrightarrow [0,1]$, satisfying the following axioms.*

(i) $\mu(\emptyset) = 0$, $\mu(X) = 1$.

(ii) $A \subset B \subset X$ *implies* $\mu(A) \leq \mu(B)$.

In this context, $\mu(A)$ represents the weight of importance of the set of criteria A. Thus, in addition to the usual weights on criteria taken separately, weights on any combination of criteria are also defined.

A fuzzy measure is said to be *additive* if $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $A \cap B = \emptyset$, *superadditive* (resp. *subadditive*) if $\mu(A \cup B) \geq \mu(A) + \mu(B)$ (resp. $\mu(A \cup B) \leq \mu(A) + \mu(B)$) whenever $A \cap B = \emptyset$. Note that if a fuzzy measure is additive, then it suffices to define the n coefficients (weights) $\mu(\{x_1\}), \ldots, \mu(\{x_n\})$ to define entirely the measure. In general, one needs to define the 2^n coefficients corresponding to the 2^n subsets of X.

We introduce now the concept of fuzzy integrals. As we consider fuzzy integrals as operators on $[0,1]^n$, we restrict the definitions to $[0,1]$-valued functions. Moreover, we will adopt a connective-like notation instead of the usual integral form.

Definition 2 *Let μ be a fuzzy measure on X. The* Sugeno integral *of a function $f : X \longrightarrow [0,1]$ with respect to μ is defined by :*

$$\mathcal{S}_\mu(f(x_1), \ldots, f(x_n)) := \bigvee_{i=1}^{n} (f(x_{(i)}) \wedge \mu(A_{(i)})) \tag{10}$$

where $\cdot_{(i)}$ indicates that the indices have been permuted so that $0 \leq f(x_{(1)}) \leq \cdots \leq f(x_{(n)}) \leq 1$, and $A_{(i)} := \{x_{(i)}, \ldots, x_{(n)}\}$.

Another definition was proposed later by Murofushi and Sugeno [17], using a concept introduced by Choquet in capacity theory.

Definition 3 *Let μ be a fuzzy measure on X. The* Choquet integral *of a function $f : X \longrightarrow [0,1]$ with respect to μ is defined by*

$$\mathcal{C}_\mu(f(x_1), \ldots, f(x_n)) := \sum_{i=1}^{n} (f(x_{(i)}) - f(x_{(i-1)}))\mu(A_{(i)}) \tag{11}$$

with the same notations as above, and $f(x_{(0)}) = 0$.

Sugeno and Choquet integrals are essentially different in nature, since the former is based on non linear operators (*min* and *max*), and the latter on usual linear operators. Both compute a kind of distorted average of $f(x_1), \ldots, f(x_n)$. More general definitions exist but will not be considered here (see [9]).

2.5 Properties of fuzzy integrals for aggregation

We present here the main properties of fuzzy integrals for aggregation, in order to show their suitability for multicriteria decision problems. A more exhaustive study of these properties, including proofs will be found in [7, 9].

Property 1 *The Sugeno and Choquet integral are idempotent, continuous, monotonically non decreasing operators.*

This implies that fuzzy integrals are always comprised between *min* and *max*.

Property 2 *A Choquet integral with respect to an additive measure μ coincides with a weighted arithmetic mean, whose weights w_i are $\mu(\{x_i\})$.*

As a consequence, only the Choquet integral with respect to an additive measure is decomposable.

Property 3 *The Choquet integral is stable under positive linear transformations.*

The Sugeno integral does not share this property, but satisfies a similar property with *min* and *max* replacing product and sum. In this sense, it can be said that the Choquet integal is suitable for cardinal aggregation (where numbers have a real meaning), while the Sugeno integral seems to be more suitable for ordinal aggregation (where only order makes sense). This, however, has to be further investigated.

Property 4 *Any OWA operator with weights $w_1, \ldots, w_n$ is a Choquet integral, whose fuzzy measure μ is defined by*

$$\mu(A) = \sum_{j=0}^{i-1} w_{n-j}, \quad \forall A \textit{ such that } |A| = i \tag{12}$$

where $|A|$ denotes the cardinal of A. Reciprocally, any commutative Choquet integral is such that $\mu(A)$ depends only on $|A|$, and coincides with an OWA operator.

The fact that the Choquet integral encompasses both the weighted arithmetic sums and OWA operators, which have been said to be "orthogonal" in an intuitive sense, shows the strong expressive power of it, since we can mix arbitrarily the two kinds of operators.

Property 5 *The Sugeno and the Choquet integral contains all order statistics, thus in particular,* min, max *and the median.*

This means that they can range freely from the most tolerant behaviour (*max*) to the most intolerant (*min*).

Property 6 *Weighted minimum and weighted maximum are particular cases of Sugeno integral. The fuzzy measure corresponding to $\mathrm{wmax}_{w_1,\ldots,w_n}$ is defined by*

$$\mu(A) = \bigvee_{x_i \in A} \mu(\{x_i\}), \tag{13}$$

while for $\mathrm{wmin}_{w_1,\ldots,w_n}$

$$\mu(A) = 1 - \bigvee_{x_i \notin A} \mu(\{x_i\}). \tag{14}$$

In the above,$\mu(\{x_i\}) = w_i$.

Figure 1 gives a summary of all set relations between various aggregation operators and fuzzy integrals. In this section, we did not mention the properties of fuzzy integrals related to the representation of interaction between criteria. This the aim of the next section.

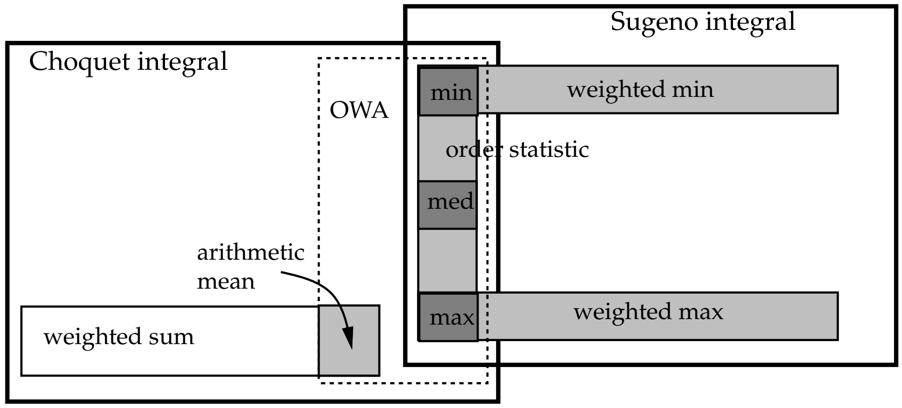

Figure 1: Set relations between various aggregation operators and fuzzy integrals

3 Importance of criteria and interaction between them

It has been said above that the main interest of fuzzy integrals lies in the fact that they can represent interaction between criteria. This is due to the fact that a weight of importance is attributed to *every* subset of criteria. Let us take a simple example when $n = 3$ in order to illustrate what is understood by interaction and how they can be modelled by fuzzy integrals.

Consider the problem of evaluation of students in high school with respect to three subjects: mathematics, physics and literature. Usually, this is done by a simple weighted sum, whose weights are the coefficients of importance of the different subjects. Suppose that the school is more scientifically than literary oriented, so that weights could be for example 3, 3 and 2 respectively. Then the weighted sum will give the following results for the three students A, B, and C (marks are given on a scale from 0 to 20):

	Mathematics	Physics	Literature	Global evaluation (weighted sum)
Student A	18	16	10	15.25
Student B	10	12	18	12.75
Student C	14	15	15	14.62

If the school want to favor well equilibrated students without weak points, the above ranking is not fully satisfactory, since student A has a severe weakness in literature, but has been considered better than student C, which has no weakpoint. The reason is that too much importance is given to mathematics and physics, which are in a sense *redundant*, since usually, students good at mathematics are also good at physics (and vice versa), so that the evaluation is overestimated (resp. underestimated) for students good (resp. bad) at mathematics and/or physics. We solve this problem by using a suitable fuzzy measure μ and the Choquet integral.

1. since scientific subjects are more important than literature, we put the following weights on subjects taken individually.

   $\mu(\{\text{mathematics}\}) = \mu(\{\text{physics}\}) = 0.45$
   $\mu(\{\text{literature}\}) = 0.3$

   Note that the initial ratio of weights (3,3,2) is kept unchanged.

2. since mathematics and physics are redundant, the weight attributed to the set {mathematics,physics} should be less than the sum of the weights of mathematics and physics.

   $\mu(\{\text{mathematics,physics}\}) = 0.5 < 0.45 + 0.45$

3. since we must favor students equally good at scientific subjects *and* literature (and this is rather uncommon), the weight attributed to the set {literature,mathematics} should be greater than the sum of individual weights (the same for physics and literature).

$\mu(\{\text{literature,mathematics}\}) = 0.9 > 0.45 + 0.3$
$\mu(\{\text{literature,physics}\}) = 0.9 > 0.45 + 0.3$

4. $\mu(\emptyset) = 0, \mu(\{\text{mathematics,physics,literature}\}) = 1$ by definition.

Applying this fuzzy measure to the three students A, B and C, we get the following result.

	Mathematics	Physics	Literature	Global evaluation (Choquet integral)
Student A	18	16	10	13.9
Student B	10	12	18	13.6
Student C	14	15	15	14.9

One can see that we get the expected result.

The above example shows how it is easy to translate requirements of the decision maker into coefficients of the fuzzy measure. The idea is that superadditivity of the fuzzy measure implies synergy between criteria, and subadditivity implies redundancy. However, things are much more complicated when n is greater than 3, since we must consider sets of three and more criteria, which becomes difficult to grasp. Moreover, we can remark the following.

- the global importance of a criterion, say x_j is not solely determined by the value $\mu(\{x_j\})$, but also by all $\mu(A)$ such that $x_j \in A$. But how to extract from these values the contribution of x_j alone?
- in the same spirit, the interaction between x_i and x_j is not only determined by the difference $\mu(\{x_i, x_j\}) - \mu(\{x_i\}) - \mu(\{x_j\})$ but also by all the coefficients $\mu(A)$ such that $\{x_i, x_j\} \subset A$. Then, how to compute a degree of interaction which is meaningful?

Murofushi *et al.* [15, 16] have proposed a solution, based on game theory for the importance of criteria, and on multiattribute utility theory for defining an interaction index between two criteria.

Definition 4 *Let μ be a fuzzy measure on $X = \{x_1, \dots, x_n\}$. The* importance index *or* Shapley value *of criterion x_i with respect to μ is defined by:*

$$\Lambda(x_i) = \sum_{A \subset X - \{x_i\}} \gamma_X(A)[\mu(A \cup \{x_i\}) - \mu(A)] \qquad (15)$$

with $\gamma_X(A) = \frac{(|X|-|A|-1)! \cdot |A|!}{|X|!}$ where $|A|$ indicates the cardinal of A, and 0!=1 as usual.

This definition has been proposed by Shapley in cooperative game theory [21]. These values have the property that $\sum_{i=1}^{n} \Lambda(x_i) = 1$. It is convenient to scale these values by a factor n, so that an importance index greater than 1 indicates an attribute more important than the average.

Definition 5 *The* average interaction index *between two criteria x_i and x_j with respect to a fuzzy measure μ is defined by:*

$$I(\mu)(x_i, x_j) = \sum_{A \subset X - \{i,j\}} \xi_X(A) \cdot I(\mu)(x_i, x_j | A)$$

with $\xi_X(A) = \frac{(|X|-|A|-2)! \cdot |A|!}{(|X|-1)!}$, and $I(\mu)(x_i, x_j|A) = \mu(A \cup \{x_i, x_j\}) - \mu(A \cup \{x_i\}) - \mu(A \cup \{x_j\}) + \mu(A)$.

The interaction index ranges in $[-1, 1]$, and is negative in case of redundancy, positive in case of synergy. Let us give as illustration the Shapley value and interaction indexes for the example of the high school (M, P, L stand for mathematics, physics and literature resp.).

criteria	M	P	L
scaled Shapley value	0.876	0.876	1.251

pair of criteria	M,P	M,L	P,L
interaction index	-0.45	0.100	0.100

Numerical values are in accordance with the interpretation given above.

4 Identification of fuzzy measures

We address now the difficult problem of the identification of fuzzy measures, that is, how to determine the 2^n coefficient of a fuzzy measure μ, taking into account the monotonicity relations between the coefficients. The interested reader should consult [9] for a full explanation of this topic. There are essentially three approaches.

4.1 Identification based on the semantics

This approach is more or less the one taken in example of section 3. It consists of guessing the coefficients of μ, on the basis of semantical considerations. This could be the following.

importance of criteria. This can be properly done by use of the Shapley value, or by the value of $\mu(\{x_j\})$ alone (but see section 3).

interaction between criteria. The interaction index (def. 4) is suitable for this. For $n \leq 3$, the sign of $\mu(\{x_i, x_j\}) - \mu(\{x_i\}) - \mu(\{x_j\})$ should be sufficient, although not equivalent to the interaction index.

symmetric criteria. Two criteria x_i, x_j are symmetric if they can be exchanged without changing the aggregation mode. This is the case of mathematics and physics in the previous example. Then, $\mu(A \cup \{x_i\}) = \mu(A \cup \{x_j\}), \forall A \subset X - \{x_i, x_j\}$. This reduces the number of coefficients.

veto effect. A criterion x_j is said to be a *veto* for the decision problem if the aggregation operator can be decomposed as follows.

$$\mathcal{H}(a_1, \dots, a_j, \dots, a_n) = a_j \wedge \mathcal{G}(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n) \quad (16)$$

This means that a bad score on criterion j will lead to a bad global score, whatever the degree of satisfaction of the other criteria. A Choquet or a Sugeno integral is able to model this. It suffices to take a fuzzy measure μ such that $\mu(A) = 0, \forall A \subset X - \{x_j\}$. In this case, it remains 2^{n-1} coefficients to be determined. The definition can be extended easily to subsets of criteria which are veto. Each veto criterion divides roughly by 2 the number of coefficients to be determined.

pass effect. Similarly, a criterion is said to be a *pass* if the aggregation operator can be decomposed as follows.

$$\mathcal{H}(a_1, \dots, a_j, \dots, a_n) = a_j \vee \mathcal{G}(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n) \quad (17)$$

Here, a good score on x_j will lead to a global good score, whatever the values of the remaining a_i. This can be also achieved by a Choquet or a Sugeno integral, by taking a fuzzy measure so that $\mu(A) = 1$, $\forall A \ni x_j$. Again in this case, the number of remaining coefficients is 2^{n-1}. The definition can be extended to any subset of criteria. Moreover, it is possible to mix veto and pass effects.

Although it is possible to reduce the number of coefficients by considering symmetry, veto or pass effect, this approach is practicable only for low values of n, and above all, if one has at his disposal an expert or decision maker who is able to tell the relative importance of criteria, and the kind of interaction between them, if any. This could be the case, in application of design of new products, where the marketing “defines” what should be the ideal product, in terms of aggregation of criteria, in the same spirit of what has been done in the high school example, where the “ideal student” was defined.

As one can see, this approach is far from being a well established methodology, and involves a non negligible amount of experience. Clearly, further research is needed in order to clarify all this. Note also that, due to the fact that Choquet and Sugeno integrals have different properties and behaviours, one or the other may be more appropriate for certain classes of applications.

4.2 Identification based on learning data

Considering the fuzzy integral model as a system, one can identify its parameters by minimizing an error criterion, provided learning data are available. Suppose that (z_k, y_k), $k = 1, \ldots, l$ are learning data where $z_k = [z_{k1} \ \cdots \ z_{kn}]^t$ is a n dimensional input vector, containing the degrees of satisfaction of object k with respect to criteria 1 to n (or the degrees of preference between two objects if outranking methods are used), and y_k is the global evaluation of object k. Then, one can try to identify the best fuzzy measure μ so that the squared error criterion is minimized.

$$E^2 = \sum_{k=1}^{l} (\mathcal{C}_\mu(z_{k1}, \ldots, z_{kn}) - y_k)^2 \tag{18}$$

It can be shown [9] that (18) can be put under a quadratic program form, that is

minimize $\frac{1}{2}\mathbf{u}^t\mathbf{D}\mathbf{u} + \mathbf{c}^t\mathbf{u}$
under the constraint $\mathbf{A}\mathbf{u} + \mathbf{b} \geq \mathbf{0}$

where $\mathbf{u}$ is a $(2^n - 2)$ dimensional vector containing all the coefficients of the fuzzy measure μ (except $\mu(\emptyset)$ and $\mu(X)$ which are fixed), $\mathbf{D}$ is a $(2^n - 2)$ dimensional square matrix, $\mathbf{c}$ a $(2^n - 2)$ dimensional vector, $\mathbf{A}$ a $n(2^{n-1} - 1) \times (2^n - 2)$ matrix, and $\mathbf{b}$ a $n(2^{n-1} - 1)$ dimensional vector. This program has a unique solution, and can be solved by the Lemke method.

If there are too few learning data (there must be at least $n!/[(n/2)!]^2$ data: see [9] for details), matrices may be ill-conditioned. Moreover, the constraint matrix $\mathbf{A}$ is a sparse matrix, and become sparser as n grows, causing bad behaviour of the algorithm. For all these reasons, including memory problems, time of convergence, the solution given by a quadratic program is not always reliable in practical situations.

Some authors have proposed alternatives to quadratic programming under the form of "heuristic" algorithms taking advantage of the peculiar structure of fuzzy measures, but often suboptimal. It seems that the best one in terms of performance, time and memory is the one proposed recently by the author [8]. The basic idea is that, in the absence of any information, the most reasonable way of aggregation is the arithmetic mean (provided the problem is cardinal), thus a Choquet integral with respect to an additive equidistributed fuzzy measure. Any input of information tends to move away the fuzzy measure from this equilibrium point. This means that, in case of few data, coefficients of the fuzzy measure which are not concerned with the data are kept as near as possible to the equilibrium point. Thus, in this algorithm, there is no problem of having too few data.

4.3 Combining semantics and learning data

Obviously, the combination of semantical considerations, which are able to reduce the complexity and provide guidelines, with learning data should lead to more efficient algorithms. An attempt in this direction is given by Yoneda *et al.* [31] (see also [9]). The basic ideas are the following.

- objective: minimize the distance to the additive equidistributed fuzzy measure,
- constraints: the usual constraints implied by the monotonicity of fuzzy measures, and constraints coming from semantical considerations. These could be about relative importance of criteria, redundancy, support between criteria, etc.

The constraints being linear, this leads to a quadratic program as before, which can be solved by the Lemke method.

5 A survey of applications

We now proceed to give some practical examples, along with a brief explanation. Full details can be found in the references, but many are in japanese. The reader can find a detailed presentation in english of these applications in [9].

Prediction of wood strength (Ishii and Sugeno, 1985) [11]. The problem consists to predict the ultimate bending strength of wood beams on the basis of some measurements, which are radius of curvature measured in different points under a uniform bending moment, density, grain angle and moisture content. Experienced inspectors are able to predict quite accurately the wood strength from these measurements. Ishii and Sugeno applied the Sugeno integral to this 4 criteria evaluation problem, and found a significant improvement compared to a linear model (i.e. weighted average). The identification of μ was performed by a heuristic learning algorithm of Ishii and Sugeno, and used 30 data for each type of wood.

evaluation of printed color images (Tanaka and Sugeno, 1988) [26]. It is the first application of Choquet integral. 15 criteria were used to qualify proofs made from an original (color reversal film), such as contrast, transparency, feeling of volume, rendering of details, etc... and pairwise comparison of proofs were performed. As the number of criteria is high, a factor analysis was performed in order to group the criteria into three main factors. A two levels model is then obtained, in which Choquet integral was used, in both levels. Because of the linear character of the Choquet integral, quadratic programming and relaxation procedure was used to identify the fuzzy measure, by minimization of the squared error between data and the model.

Design of speakers (Mitsubishi Electric, 1991) [10]. This application concerns the design of a small size audio speaker. 7 criteria were chosen, which are impression of lightness, of neatness, of precision, of the material, morphological characteristics, basic element, "plus alpha". The method used the Choquet integral with respect to a possibility and a necessity measure. Possibility and necessity measures, denoted Π, N respectively, are such that $\Pi(A \cup B) = \Pi(A) \vee \Pi(B)$, $N(A \cap B) = N(A) \wedge N(B)$, and can be generated by a so-called *possibility density* defined by $\pi(x_j) = \Pi(\{x_j\})$. The identification of the possibility density was performed by a possibilistic version of the AHP (Analytical Hierarchy Process) of Saaty. Choquet integral with respect to a possibility measure behaves like a disjunctive operator, so that it favours speakers with high originality but with possible weakpoints, while the use of a necessity measure, which leads to a conjunctive behaviour, tends to favour well balanced speakers without weak points, intended to the average consumer.

Human reliability analysis (Mitsubishi Research Inst., 1992) [28]. Here, a human operator controls the coolant flow rate supplied to a water cooling tank, and this rather complex task is known to have a significant error operation rate. In order to improve the reliability, three options were proposed: (1) introduce a reliable automated control system so that the operator should just set the target flow rate, (2) introduce a plant simulator to train the operator, (3) employ an assistant operator. The human reliability of each option was evaluated by Sugeno integral, using 5 criteria (here called performance shaping factors) related to the situation recognition stage, the action judgment stage and the manipulation stage. The fuzzy measure was identified through a questionnaire and simplifying assumptions on the fuzzy measure.

Modelling of public attitude towards use of nuclear energy [20]. Formally, this is the converse of the problem of evaluation: we have *one* object and n persons $p_1, \dots, p_n$ evaluating it with respect to m criteria. The global evaluation of the objet by person p_i is again obtained by a fuzzy integral:

$$h(p_i) = \mathcal{S}_\mu(h_1(p_i), \dots, h_m(p_i)) \tag{19}$$

where $h_i(p_i)$ denotes individual evaluation. Here μ models the evaluation process of the persons concerning a particular object (the use of nuclear energy), in other words, a public attitude for a particular question. In the experiment, subjects were asked to give their opinion about 30 criteria (such as: improves standart of living, is harmful to future generation, provides a cheap energy source, leads to accidents which affect large number of people, etc...), under a triplet form (evaluation: from bad to good, belief: from unlikely to likely, importance: from unimportant to important), and then an overall judgment, from unfavourable to favourable. The data came from three countries (Japan, Philippines and Germany), and the data of each country were divided into the PRO group (positive favourability), and the CON group (negative favourability), thus leading to six distinct groups. Each group was then modelled by a different fuzzy measure. As in the example of color prints evaluation, a factor analysis was first performed to group the criteria into six factors, leading to a hierarchical model. Sugeno integral was used in both levels, and the identification of the fuzzy measure was obtained by the learning algorithm of Ishii and Sugeno. The opinion of each group can be then analyzed using a semantical interpretation of the fuzzy measures.

Other applications. There are many other examples which can be cited, all developed in Japan. Among these, we can find

- an evaluation of the taste of rice and coffee [13] by Satake Eng. Recently, their data have been further processed and performance improved by Sugeno *et al.* [24],
- a dignostic system for plant, build by Fuji Electric [5, 4],
- an evaluation of living environment [19].

Let us mention finally an application of Choquet integral to time series modeling [25].

6 Conclusion

This review has shown the richness —but also the complexity— of this new tool for aggregation, the fuzzy integral. We have given the state-of-the-art on practical aspects of its implementation in real applications, and given many examples, some being industrial ones, for illustration.

We hope that this will encourage people in Europe to use this promising technique in multicriteria decision making.

References

[1] D. Dubois and H. Prade. A review of fuzzy set aggregation connectives. *Information Sciences*, 36:85–121, 1985.

[2] D. Dubois and H. Prade. Weighted minimum and maximum operations in fuzzy set theory. *Information Sciences*, 39:205–210, 1986.

[3] J.C. Fodor and M. Roubens. *Fuzzy Preference Modelling and Multi-Criteria Decision Aid*. Kluwer Academic Publisher, 1994.

[4] K. Goto, Y. Toriyama, and O. Itoh. Intelligent alarm method by fuzzy measure and its application to plant abnormality prediction. In *Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symp.*, pages 395–400, Yokohama, Japan, March 1995.

[5] K. Goto, Y. Toriyama, and O. Itoh. A method of state synthesis evaluation for plant based on fuzzy memory-based reasoning and fuzzy measure. *J. of Japan Society for Fuzzy Theory and Systems*, 7(2):390–401, 1995.

[6] M. Grabisch. Characterization of fuzzy integrals viewed as aggregation operators. In *3rd IEEE Conf. on Fuzzy Systems*, pages 1927–1932, Orlando, USA, June 1994.

[7] M. Grabisch. Fuzzy integral in multicriteria decision making. *Fuzzy Sets & Systems*, 69:279–298, 1995.

[8] M. Grabisch. A new algorithm for identifying fuzzy measures and its application to pattern recognition. In *Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symposium*, pages 145–150, Yokohama, Japan, March 1995.

[9] M. Grabisch, H.T. Nguyen, and E.A. Walker. *Fundamentals of Uncertainty Calculi, with Applications to Fuzzy Inference*. Kluwer Academic, 1995.

[10] K. Inoue and T. Anzai. A study on the industrial design evaluation based upon non-additive measures. In *7th Fuzzy System Symp.*, pages 521–524, Nagoya, Japan, June 1991. In Japanese.

[11] K. Ishii and M. Sugeno. A model of human evaluation process using fuzzy measure,. *Int. J. Man-Machine Studies*, 22:19–38, 1985.

[12] A. Kolmogoroff. Sur la notion de moyenne. *Atti delle Reale Accademia Nazionale dei Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez.*, 12:323–343, 1930.

[13] M. Matsuda and T. Kameoka. Application of fuzzy measure, fuzzy integral and neural network to the system which estimate taste by using industrial analysis. In *2nd Int. Conf. on Fuzzy Systems and Neural Networks*, pages 601–606, Iizuka, Japan, July 1992.

[14] V. Mousseau. Analyse et classification de la littérature traitant de l'importance relative des critères en aide multicritère à la décision. *Recherche Opérationnelle/Operations Research*, 26(4):367–389, 1992.

[15] 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.

[16] 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.

[17] T. Murofushi and M. Sugeno. A theory of fuzzy measures. Representation, the Choquet integral and null sets. *J. Math. Anal. Appl.*, 159(2):532–549, 1991.

[18] T. Murofushi and M. Sugeno. Non-additivity of fuzzy measures representing preferential dependence. In *2nd Int. Conf. on Fuzzy Systems and Neural Networks*, pages 617–620, Iizuka, Japan, July 1992.

[19] Y. Nakamori and N. Iwamoto. Analysis of environmental evaluation structure by the Choquet integral model. In *Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symp.*, pages 695–700, Yokohama, Japan, March 1995.

[20] T. Onisawa, M. Sugeno, Y. Nishiwaki, H. Kawai, and Y. Harima. Fuzzy measure analysis of public attitude towards the use of nuclear energy. *Fuzzy Sets & Systems*, 20:259–289, 1986.

[21] 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.

[22] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.

[23] M. Sugeno. Fuzzy measures and fuzzy integrals — a survey. In Gupta, Saridis, and Gaines, editors, *Fuzzy Automata and Decision Processes*, pages 89–102. 1977.

[24] M. Sugeno and S.H. Kwon. A clusterwise regression-type model for subjective evaluation. *J. of Japan Society for Fuzzy Theory and Systems*, 7(2):291–310, 1995.

[25] M. Sugeno and S.H. Kwon. A new approach to time series modeling with fuzzy measures and the Choquet integral. In *Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symp.*, pages 799–804, Yokohama, Japan, March 1995.

[26] K. Tanaka and M. Sugeno. A study on subjective evaluations of color printing images. In *4th Fuzzy System Symposium*, pages 229–234, Tokyo, Japan, May 1988. In Japanese.

[27] K. Tanaka and M. Sugeno. A study on subjective evaluation of color printing images. *Int. J. of Approximate Reasoning*, 5:213–222, 1991.

[28] T. Washio, H. Takahashi, and M. Kitamura. A method for supporting decision making on plant operation based on human reliability analysis by fuzzy integral. In *2nd Int. Conf. on Fuzzy Logic and Neural Networks*, pages 841–845, Iizuka, Japan, July 1992.

[29] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. Systems, Man & Cybern.*, 18:183–190, 1988.

[30] R.R. Yager. Connectives and quantifiers in fuzzy sets. *Fuzzy Sets & Systems*, 40:39–75, 1991.

[31] M. Yoneda, S. Fukami, and M. Grabisch. Interactive determination of a utility function represented by a fuzzy integral. *Information Sciences*, 71:43–64, 1993.