The role of fuzzy sets in decision sciences: old techniques and new directions

Didier Dubois [a]

[a]*IRIT, Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse Cedex 9*

Abstract

We try to provide a tentative assessment of the role of fuzzy sets in decision analysis. We discuss membership functions, aggregation operations, linguistic variables, fuzzy intervals and the valued preference relations they induce. The importance of the notion of bipolarity and the potential of qualitative evaluation methods are also pointed out. We take a critical standpoint on where we stand, in order to highlight the actual achievements and question what is often considered debatable by decision scientists observing the fuzzy decision analysis literature.

Key words: Decision, qualitative value scales, aggregation, linguistic variables, preference relations, fuzzy intervals, ranking methods.

1. Introduction

The idea of using fuzzy sets in decision sciences is not surprizing since decision analysis is a field where human-originated information is pervasive. The seminal paper in this area was written by Bellman and Zadeh [1] in 1970, highlighting the role of fuzzy set connectives in criteria aggregation. That pioneering paper makes three main points:

1. Membership functions can be viewed as a variant of utility functions or rescaled objective functions, and optimized as such.
2. Combining membership functions, especially using the minimum, can be one approach to criteria aggregation.
3. Multiple-stage decision-making problems based on the minimum aggregation connective can then been stated and solved by means of dynamic programming.

*Corresponding author

Email address: `dubois@irit.fr` (Didier Dubois)

This paper is based on an invited talk at the EUROFUSE Workshop Preference Modelling and Decision Analysis, Pamplona (Spain), 16/09/2009-18/09/2009.

This view was taken over by Zimmermann[2] who developed popular multicriteria linear optimisation techniques in the seventies. The idea is that constraints are soft and can be viewed as critieria. Then any linear programming problem becomes a max-min fuzzy linear programming problem.

Other ingredients of fuzzy set theory like fuzzy ordering relations, linguistic variables and fuzzy intervals have played a major role in the diffusion of fuzzy set ideas in decision sciences. We can especially point out the following:

1. Gradual or valued preference relations (stemming from Zadeh's fuzzy orderings [3]) further studied by Orlowski [4], Fodor and Roubens [5], and others [6].
2. Many other aggregation operations are used so as to refine the multicriteria aggregation technique of Belman and Zadeh: t-norms symmetric sums, uninorms, leximin, Sugeno and Choquet integrals etc. This trend is testified by three recent books (Beliakov et al. [7], Torra and Narukawa [8], Grabisch et al.[9]).
3. Fuzzy interval computations so as to cope with uncertainty in numerical aggregation schemes. Especially, extensions of the weighted average with uncertain weights [10].
4. Fuzzy interval comparison techniques enable the best option in a set of alternatives with fuzzy interval ratings to be selected [11]
5. Linguistic variables [12] are supposed to model human originated information , so as to get decision methods closer to the user cognition [13].

What has been the contribution of fuzzy sets to decision analysis? Following the terminology of the original Bellman-Zadeh paper, fuzzy decision analysis (FDA) is supposed to take place in a "fuzzy environment", in contrast with probabilistic decision analysis, taking place "under uncertainty". But, what is a fuzzy environment ? I seems that many authors take it an environment where the major source of information is linguistic, so that linguistic variables are used, which does not correspond to Bellman and Zadeh's proposal. One should nevertheless not oppose "fuzzy environment" to "uncertain environment": the former in fact often means "using fuzzy sets", while the latter refers to an actual decision situation: there is epistemic uncertainty due to missing information, not always related to linguistic imprecision.

Actually, for many decision theory specialists, it is not clear that fuzzy sets have ever led to a new decision paradigm. Indeed, one may argue that either some techniques already existed under a different terminology, or that fuzzy decision methods are fuzzifications of standard decision techniques. More precisely,

- Fuzzy optimization following Bellman and Zadeh and Zimmermann comes down to max-min bottleneck optimization. But bottleneck optimisation and maximin decisions already existed independently of fuzzy sets.

- In many cases, fuzzy sets have just been added to existing techniques (fuzzy AHP methods, fuzzy weighted averages, fuzzy extensions of Electre-style MCDM methods) with no clear benefits (especially when fuzzy information is changed into precise at the preprocessing level, which can be observed sometimes).

- Fuzzy preference modelling is an extension of standard preference modelling and must be compared to probabilistic or measurement-based preference modeling.

In fact, contrary to what is often claimed in FDA papers, it is not always the case that adding fuzzy sets to an existing method improves it in a significant way. That it does needs to be articulated by convincing arguments, based on sufficient knowledge of state-of-the-art existing techniques.

To make a real contribution one must show that the new technique

- addresses in a correct way an issue not previously handled by previous methods: e.g. criterion dependence using Choquet integral.
- proposes a new setting for expressing decision problems more in line with the information provided by users: for instance using qualitative information instead of numerical.
- possesses a convincing rationale (e.g. why such an aggregation method? why model uncertainty by a fuzzy set ?) and a sound formal setting liable to some axiomatization.

Unfortunately, it is not always clear that any such contribution appears in many proposals and published papers on FDA. This position paper takes a skeptical viewpoint on the fuzzy decision literature, so as to help laying bare what is its actual contribution. Given the large literature available there is no point to providing a complete survey. However we shall try to study various ways fuzzy sets were instilled in decision methods, and provide a tentative assessment of the cogency of such proposals.

2. Membership functions in decision-making

First we discuss the role played by membership functions in decision techniques. Then we consider the use of membership grades and linguistic terms for rating the worth of decisions.

2.1. Membership functions and truth sets in decision analysis

A membership function is an abstract notion, a mathematical tool just like a set in general. It just introduce grades in the abstract Boolean notion of set-membership. So, using the terminology of membership functions in a decision problem does not necessarily enrich its significance. In order to figure out the contribution of fuzzy sets, one must always declare what a given membership function accounts for in a given problem or context. Indeed, there is not a single semantic interpretation of membership functions. Several ones have been laid bare [14] and can be found in the literature

- A measure of similarity to prototypes of a linguistic concept (then membership degrees are related to distance); this is used when linguistic terms are modeled by membership functions, and in fuzzy clustering as well (Ruspini [15]).
- A possibility distribution [16] representing our incomplete knowledge of a parameter, state of nature, etc. that we cannot control. Possibility distributions can be numerical or qualitiative [17]. In the numerical case, such a membership function can encode a family of probability functions (see [18] for a survey)
- A numerical encoding of a preference relation over feasible options, similar to a utility or an objective function. This is really the idea in the Bellman-Zadeh paradigm of decision-making in a fuzzy environment. In decision problems, membership functions introduce grades in the traditionally Boolean notion of feasibility. A degree of feasibility differs from the degree of attainment of a goal. In the former case, a membership function models a fuzzy constraint [19, 20].

In the scope of decision under uncertainty, membership functions offer an alternative to both probability distributions and utility functions, especially when only qualitative value scales are used. But both types of membership functions should not be confused not used one for the other in problems involving both fuzzy constraints and uncertainty [20].

Then the originality of the fuzzy approach may lie

- either in its capacity to translate linguistic terms into quantitative ones in a flexible way
- or to explicitly account for the lack of information, avoiding the questionable use of unique, often uniform probability distributions [18]
- or in its set-theoretic view of numerical functions. Viewing a utility function as a fuzzy set, a wider range of aggregation operations becomes available, some of which generalize the standard weighted average, some of which generalize logical connectives.

However, not only must a membership function be interpreted in the practical context under concern, the scale in which membership degrees lie must also be well-understood and its expressive power made clear.

2.2. Truth-sets as value scales: the meaning of end-points

The often totally ordered set of truth-values, we shall denote by $(L, \geq)$, is also an abstract construct. Interpretive assumptions must be laid bare if it is used as a value scale for a decision problem: The first issue concerns the meaning of the end-points of the scale; and whether a mid-point in the scale exists or has any meaning ? Let us denote by 0 the least element in L and by 1 the greatest element. Let us define a mid-point of L as an element $e \in L$ such that

1. $\exists \lambda^-, \lambda^+ \in L, \lambda^- < e < \lambda^+$
2. There is an order-reversing bijection $n : L \to L$ such that $n(1) = 0; n(e) = e$ (n is a strong negation function)

Three kinds of scales can be considered depending on the existence and the meaning of these landmark points [21]:

- *Negative unipolar scales*: when 0 has a totally negative flavor while 1 has a neutral flavor. For instance, a possibility distribution, a measure of loss.
- *Positive unipolar scales*: 0 has a neutral flavor. For instance degrees of necessity, a measure of gain.
- *Bipolar scales*: when 1 has a totally positive flavor while 0 has a totally negative flavor. Then the scale contains a mid-point e that has has a neutral flavor and that plays the role of a boundary between positive and negative values.

For instance, the unit interval viewed as a probability scale is bipolar since 0 means impossible, 1 means certain and 1/2 indicates a balance between the probable and the improbable. A membership scale of a fuzzy set is in the principle bipolar, insofar as 1/2 represents the cross-over point between membership and non-membership. However if a membership function is used as a possibility distribution as suggested by Zadeh [16], the scale becomes negative unipolar since then while 0 means impossible, 1 means only possible, which is neutral. The dual scale of necessity degrees in possibility theory [17] is on the contrary positive unipolar since the top value of L expresses full certainty while the bottom represents full uncertainty, hence neutral. In these latter cases, the midpoint even if it exists, plays no role in the representation.

Finally, if membership grades express preference, the degree of satisfaction of a goal is ofter bipolar (in order to express satisfaction, indifference and dissatisfaction) [22]. However, another approach is to use two separate unipolar scales, one to express the degree of feasibility of a solution (it is a negative unipolar scale ranging from not feasible to feasible) one to express the attractiveness of solutions (a positive unipolar scale ranging from indifference to full satisfaction). More generally, loss functions map to a negative unipolar scale, while gain functions map to positive unipolar scales. Gains and losses can be separately handled as in cumulative prospect theory[23].

This information about landmark points in the scale captures ideas of good or bad in the absolute. A simple preference relation cannot express this kind of knowledge: ranking solutions to a decision problem from the best to the worst without making the meaning of the value scale explicit, nothing prevents the best solution found from being judged rather bad, or on the contrary the worst solution from being somewhat good.

The choice of landmark points also has strong impact on the proper choice of aggregation operations (t-norms, co-norms, uninorms..)[24]. Especially landmark points in the scale are either neutral or absorbing elements of such aggregation operations.

2.3. Truth-sets as value scales: quantitative or qualitative ?

The second issue pertains to the expressive power of grades in a scale L and has to do with its algebraic richness. One can first decide if a infinite scale makes sense or not. It clearly makes sense when representing preference about a continuous measurable attribute. Then, the reader is referred to the important literature on measurement theory (see [25] and Chapter 16 in [26]) whose aim is to represent preference relations by means of numerical value functions. According to this literature, there are three well-known kinds of continuous value scales

- *Ordinal scales*: The numerical values are defined up to a monotone increasing transformation. Only the ordering on L matters. It makes no sense to add degrees in such scales.
- *Interval scales*: The numerical values are defined up to a positive affine transformation ($\lambda \in L \mapsto a\lambda + b, a > 0$). Interestingly, in decision theory the most popular kind of value scales are interval scales. But they cannot express the idea of good and bad (they are neither bipolar nor even unipolar) since the value 0 plays no specific role, and these scales can be unbounded.
- *Ratio scales*: The numerical values are defined up to a positive linear transformation $a\lambda, a > 0$. This kind of scale is often unipolar positive as the bottom value 0 plays a major role.

Another option is to go for a *finite* scale: $L = \{0 < \lambda_1 < \lambda_2 < \cdots < \lambda_m = 1\}$ where elements of the scale are not numerical. In the following we shall speak of a *qualitative* scale. It underlies the assumption that the values of the scale are significantly distinct from one another (hence they cannot be too numerous): in particular, the value λ_i is significantly better that λ_{i-1}. This case is often neglected both in usual measurement theory and in fuzzy set theory. Yet a qualitative scale is more expressive than a simple ordering relation because of the presence of absolute landmarks that can have a positive, negative or neutral flavor.

Clearly the nature of the scale also affects the kind of aggregation function that can be used to merge degrees. An aggregation operation $*$ on an ordinal scale must satisfy an ordinal invariance property such as the following:

$$a * b > c * d \iff \varphi(a) * \varphi(b) > \varphi(c) * \varphi(d)$$

for all monotonic increasing transformations φ of an ordinal scale L. See [9], chap. 8, on this problem. Basically only operations based on maximum and minimum remain meaningful. Averaging operations make no sense on such scales.

More often than not in decision problems, people are asked to express their preferences by ticking a value on a continuous line segment. Then such values are handled as if they were genuine real numbers, computing averages or variances. This kind of technique is nearly as debatable as asking someone to explicitly provide an real number expressing preference. All we can assume is that the corresponding scale is an ordinal scale. In particular, there is a problem of commensurateness between scales used by several individuals: the same numerical value provided by two individuals may fail to bear the same meaning. On the other hand qualitative scales can better handle this problem: some landmark values can be identically understood by several individuals and may be compared across several criteria. A small qualitative scale is cognitively easier to grasp than a continuous value scale and has thus more chance to be consensual.

In summary there is a whole literature on numerical utility theory that should be exploited if fuzzy set decision researchers wish to justify the use of numerical membership grades in decision techniques. From this point of view, calling a utility function a membership function is not a contribution. Yet, f uzzy set theory offers a framework to think of aggregation connectives in a broader way than the usual weighted averaging schemes. But there is no reason to move away from the measurement tradition of standard decision analysis. Fuzzy set tools should essentially enrich it.

2.4. From numerical to fuzzy value scales

Being aware that precise numerical techniques in decision evaluation problems are questionable, because they assume more information than can actually been supplied by individuals, many works have been published that claim to circumvent this difficulty by means of fuzzy set-related tools. The rationally often goes as follows: If a precise value in the real line provided by an expert is often ill-known, it can be more faithfully represented by an interval or a fuzzy interval. Moreover, the elements in a qualitative scale may encode linguistic value judgments, which can be modeled via linguistic variables.

2.4.1. Evaluations by pairs of values

When an individual ticks a value in a value scale or expresses a subjective opinion by means of a number x, it sounds natural to admit that this value has limited precision. The unit interval is far too refined to faithfully interpret subjective value judgments. It is tempting to use an interval $[a, b]$ in order to describe this imprecision. However, it is not clear that this approach takes into account the ordinal nature of the numerical encoding of the value judgment. It is natural to think that the width of an interval reflects the amount of imprecision of this interval. However in an ordinal scale, width of intervals make no sense: if $[a, b] = [c, d]$, in general $[\varphi(a), \varphi(b)] \neq [\varphi(c), \varphi(d)]$ for a monotonic scale transformation φ. So the use of interval-valued ratings presupposes an assumption on the nature of the value scale, which must be more expressive than an ordinal scale. It must be equipped with some kind of metric, again resorting to suitable (e.g. preference difference) measurement techniques. Moving from an interval to a fuzzy interval with a view to cope with the uncertainty of the interval boundaries, one is not better off, since on an ordinal scale, the shape of the membership function is meaningless: there is no such thing as a triangular fuzzy number on an ordinal scale.

Some authors use pairs of values $(\mu, \nu) \in [0,1]^2$ with $\mu + \nu \leq 1$ following Atanassov's convention[27]. Not only the latter encoding looks problematic in the light of the above considerations[2], this representation technique is moreover ambiguous: it is not clear that this pair of values corresponds to *more information* or *less information* than a single value[28].

1. Using an uncertainty semantics, it expresses less information than point values because it encodes an ill-known value $\lambda \in [\mu, 1-\nu]$. Then the uncertainty interval representation is more explicit. Moreover, the aggregation operations proposed by Atanassov are fully compatible with the interval-extension of pointwise aggregation operations [29].
2. Or it expresses more information that point values: then μ is the strength in favour of a decision, ν in disfavour of this decision. This is a unipolar bivariate convention that fits argumentation semantics, and departs from the Savagean utility theory tradition. However it is not clear that researchers adopting Atanassov convention refer to pioneering works, like cumulative prospect theory, adopting this kind bipolar view. This setting is even more information demanding than using single evaluations so that it does not address at all the concerned raised by the debatable richness assumptionof numerical ratings.

The proper choice of a semantics of Atanassov style value pairs affects the way information will be processed [28]:

1. The standard injection $L \to L^2$ is not the same: $\lambda \mapsto (\lambda, 1-\lambda)$ in the interval case, $\lambda \mapsto (\lambda, 0)$ for a positive unipolar scale in the bipolar case (then the genuine pairs add negative information to single positive evaluations).
2. Under the uncertainty semantics, you need to apply interval analysis methods to see the impact of uncertainty on the global evaluation (insofar as the numerical scale is meaningful).
3. Under the argumentation semantics, you may first separately aggregate positive and negative information by appropriate (possibly distinct) methods and then aggregate the results as done in CPT.

2.4.2. Linguistic vs. Numerical Scales

Quite a number of papers on FDA have been published in the last 15 years or so, with the aim of exploiting linguistic information provided by decision-makers. Namely a criterion is viewed as a mapping decisions on a finite linguistic term set forming a qualitative scale. A number of authors then consider a criterion as a linguistic variable after Zadeh [12], namely they represent fuzzy terms as fuzzy intervals, that form a fuzzy partition of the unit interval. Contrary to the

[2] Indeed, the addition $\mu+\nu$ is questionable on an ordinal scale. One may replace $\mu+\nu \leq 1$ by $\mu \leq n(\nu)$, for a strong negation on L, but then $\mu \leq n(\nu)$ implies $\varphi(\mu) \leq \varphi(n(\nu))$ while we need $\varphi(\mu) \leq n(\varphi(\nu))$.

free use of any value on a numerical scale surrounded by imprecision in the form of a fuzzy interval, the linguistic approach only uses a finite set prescribed fuzzy intervals, and the decision-maker picks ratings among them. More often than not the unit interval, taken as a value scale, is shared into overlapping intervals of equal length that form the supports of the fuzzy intervals. It corresponds here to what Zadeh calls a *granulation* of the unit interval [30]. One advantage of this framework is that criteria aggregations can be modelled by means of fuzzy if-then rules that can be processed using any fuzzy inference method (like in fuzzy control). However this approach is debatable for a number of reasons

- It is not clear that the granular scale thus used is qualitative any more since each linguistic term is viewed as a fuzzy interval on some numerical scale. If the scale is in fact ordinal, sharing this scale into fuzzy intervals with the same shape clearly makes no sense. Such linguistic scales are thus not qualitative.

- Arithmetic aggregation operations that do not make sense on the underlying numerical scale will not make more sense when applied to fuzzy intervals.

- Combining fuzzy intervals from a partition (crisp or fuzzy) generally does not yield elements of the partition. One has to resort to some form of linguistic approximation in order to construct a closed operation on the linguistic scale. But if this operation is abstracted from a numerical one, properties of the latter (for instance associativity) will be often lost (see [31]).

- If moreover one applies fuzzy control interpolation methods to build an aggregation function (using the standard Mamdani fuzzification-inference -defuzzification scheme), what is constructed is a numerical function which highly depends, for instance, on the choice of a defuzzification method.

In fact, linguistic variables proposed by Zadeh are meaningful if the underlying numerical scale corresponds to an objective measurable attribute, like height, temperature, etc. A linguistic variable on an abstract numerical scale is all the more meaningless because, prior to the membership function measurement problems that are already present on measurable attributes, the question of how to make sense of ratings on this abstract scale is to be solved first. So this trend leads to debatable techniques that are neither more meaningful nor more robust to a change of numerical encoding (of the linguistic values) than purely numerical techniques (see the last chapter of [33] for a detailed critique of this line of works.

Besides, due to the above difficulties, there have been some attempts at directly using linguistic labels, like the 2-tuple linguistic representation [32]. The 2-tuple method handles pairs (i, σ) where i denotes the rank of label $\lambda_i \in L$ in a finite qualitative scale and $\sigma \in [-0.5, 0.5)$. The purpose is to easily go from a numerical value x lying between 0 and n to a symbolic one in L by means of the integer closest to x, representing $\lambda_i \in L : x = i + \sigma \in [0, n]$. Then any numerical aggregation function $*$ can be applied the qualitative scale L: $\lambda_i * \lambda_j = \lambda_k$ where $i * j = k + \sigma$. In this view, σ is a numerical value expressing the precision of the translation from the original result to the linguistic scale.

This kind of so-called linguistic approaches are as quantitative as any standard number-crunching method It just uses a standard rounding technique as a linguistic approximation tool, for the sake of practical convenience. It no longer accounts for the imprecision of linguistic evaluations. Moreover the idea that a qualitative scale should be mapped to a sequence of adjacent integers is debatable. First one must justify the choice of equally distributed integers. Then, one must study how the ranking of decisions obtained by aggregation of partial ratings on the integer scale depends on the choice of the monotonic mapping $L \to \mathbb{N}$ encoding the linguistic values.

2.5. The two meanings of fuzzy preference relations

There is an important stream of works triggered by the book by Fodor and Roubens [5], that extend preference modelling to the gradual situation. In classical preference modelling, an outranking relation provided by a decision-maker is decomposed into strict preference, indifference and incomparability components in order to be used. Fuzzy preference relations are valued extension of relations expressing preference, i.e. of variants of ordering or preordering relations ([26], Chap. 2). There is no point discussing the current state of this literature in detail here (see Fodor and De Baets [34] for a recent survey). However most existing works in this vein develop mathematical aspects of fuzzy relations, not so much its connection to actual preference data. This is probably due to the fact that the meaning of membership grades to fuzzy relations is not so often discussed.

The notion of fuzzy ordering originated in Zadeh's early paper [3] has been improved and extensively studied in recent years (Bodenhofer et al. [6]). A fuzzy relation on a set S is just a mapping $R : S \times S \to [0, 1]$ (or to any totally ordered scale). One assumption that pervades the fuzzy relational setting is not often emphasized: a fuzzy relation makes sense only if it is meaningful to compare $R(x, y)$ to $R(z, w)$ for 4-tuples of acts (x, y, z, z), that is, in the scope of preference modelling, to decide whether x is more preferred (or not) to y in the same way as z is more preferred to w. A fuzzy preference relation should thus be viewed as the result of a measurement procedure reflecting the expected or observed properties of crisp quaternary relations $Q(x, y, z, w)$ that should be specified by the decision-maker (see Fodor[35] for preliminary investigations).

In this case one may argue that $R(x, y)$ reflects the intensity of the preference of x over y. Nevertheless, the mathematical properties of R will again be dictated by the meaning of the extreme values of the preference scale, namely when $R(x, y) = 0$ or 1. If the unit interval is viewed as a bipolar scale, then $R(x, y) = 1$ means full strict preference of x over y, and is equivalent to $R(y, x) = 0$, which expresses full negative preference, and suggests indifference be modelled by $R(x, y) = R(x, y) = 1/2$, and more generally the property

$$R(x, y) + R(y, x) = 1$$

is naturally assumed. This property generalizes completeness, and $R(x, y) > 1/2$ expresses a degree of strict preference. Antisymmetry then reads $R(x, y) = 1/2 \implies x = y$. These are tournament relations that do not fit with the usual encoding reflexive crisp relations, and exclude incomparability. Indeed, in the usual convention of the crisp case, reflexivity reads $R(x, x) = 1$, while incomparability reads $R(x, y) = R(y, x) = 0$.

In order to stick to the latter convention, the unit interval must then be viewed as a negative unipolar scale, with neutral upper end, and the preference status between x and y cannot be judged without checking the pair $(R(x, y), R(y, x))$. In this case, $R(x, y)$ evaluates weak preference, completeness means

$$\max(R(x, y), R(y, x)) = 1,$$

and indifference is when $R(x, y) = R(y, x) = 1$. On the contrary, $R(x, y) = R(y, x) = 0$ captures incomparability. In other words, this convention allows the direct extension of usual preference relations to valued ones on a unipolar scale where 1 has a neutral flavor. Conventions of usual outranking relations are retrieved when restricting to Boolean values. The expression of antisymmetry must be handled with care in connection to the underlying similarity relation on the preference scale, as shown by Bodenhofer [36].

The above type of fuzzy relations presupposes that objects to be compared are known precisely enough to allow for a precise quantification of preference intensity. However there is another possible explanation of why preference relations should be valued, and this is when the objects to be compared are ill-known even if the preference between them remains crisp. Then $R(x, y)$ reflects the *likelihood of a crisp weak preference* $x \succeq y$. Under this interpretation, some valued relations directly refer to probability. Probability of preference is naturally encoded by valued tournament relations [37], letting

$$R(x, y) = Prob(x \succ y) + \frac{1}{2} Prob(x \sim y),$$

where $x \sim y \iff x \succeq y$ and $x \succeq y$, which implies $R(x, y) + R(y, x) = 1$. Uncertainty about preference can be defined by a probability distribution P over possible preference relations, i.e. $T_i \subset S \times S$ with $x \succeq_i y \iff (x, y) \in T_i$ and $P(T_i) = p_i, i = 1, \ldots, N$. Then

$$R(x, y) = \sum_{i: x \succ_i y} p_i + \sum_{i: x \sim_i y} \frac{1}{2} p_i.$$

This comes close to the setting of voting theory, historically the first suggested framework for interpreting (what people thought could be understood as) fuzzy relations [38].

This approach also applies when the merit of alternatives x and y can be quantified on a numerical scale and represented by a probability distribution on this scale. Then, $R(x, y) = P(u(x) > u(y))$, where $u : S \to \mathbb{R}$ is a utility function. Calling such valued tournament relations fuzzy can be misleading in the probabilistic setting, unless one considers that fuzzy just means gradual. However, what is gradual here is the likelihood of preference, the latter remaining a crisp notion, as opposed to the case when shades of preference are taken into account. Only when modelling preference intensity does a valued relation fully deserve to be called fuzzy.

Other uncertainty theories can be used as well to quantify uncertain preference, including possibility theory, i.e. $R(x, y) = \Pi(x \succeq y)$ is the degree of possibility of preference. It is such that $\max(R(x, y), R(y, x)) = 1$ since $\max(\Pi(x \succeq y), \Pi(y \succeq x)) = 1$ in possibility theory. In this case again, the underlying scale for $R(x, y)$ is negative unipolar, which does correspond to similar conventions as the gradual extension of outranking relations outlined above. However, in the possibilistic uncertainty setting, $1 - R(x, y) = N(y \succ x)$ corresponds to the degree of certainty of a strict preference. This kind of valued relations is closely akin to interval orderings [39] and the comparison of fuzzy intervals discussed later on in this paper.

3. Fuzzy connectives for decision evaluation in the qualitative setting

Fuzzy sets connectives have triggered a considerable development of aggregation operators for decision evaluation [7, 9, 8]. It was the pioneering Bellman-Zadeh's paper that popularized a non-compensatory operation (the minimum), in place of averaging, for aggregation processes in multi-objective problems. Yet, this mode of aggregation had been already extensively used since the 1940's in non-cooperative game theory, and more recently in bottleneck optimisation. However, Bellman-Zadeh's proposal has sometimes been misunderstood, as to its actual role. In fact this non-compensatory approach pioneered later developments in the literature on soft constraint satisfaction methods [20]. In particular, the non-compensatory maxmin approach and its refinements stands in opposition to the traditional optimisation literature where constraints are crisp.as well as to the systematic use of averages and their extensions for aggregating criteria [**? ?**]. The framework of aggregation operations takes a very general view that subsumes the two traditions. Rather than providing another survey of aggregation operations that are already documented in the above-cited recent books, we discuss the issue of qualitative approaches that can be developed in the light of current developments so as to overcome the above critique of linguistic scales.

3.1. Aggregation operations: qualitative or quantitative

The nature of the value scale employed for rating the worth of decisions dictates whether an aggregation operation is legitimate or not. Should we use a qualitative or a quantitative approach? There are pros and cons. We are faced with a modeling dilemma.

Using quantitative scales, we dispose of a very rich framework:

- we can account for very refined aggregation attitudes, especially trade-off, compensation and dependence between criteria

- A very fine-grained ranking of alternatives can be obtained.
- The aggregation technique can be learned from data.
- However, numerical preference data are not typically what decision-makers provide.

Qualitative approaches (ordinal or qualitative scales) may look more anthropomorphic. Indeed, contrary to what classical decision theory suggests, people can make decisions in the face of several criteria, sometimes without numerical utility nor criteria importance assessments (see the works by Gigerenzer [40], for instance). More precisely, in a qualitative setting:

- We are closer to the information human can actually supply
- We can nevertheless model preference dependence structures (see the recent literature on CP-nets [41])
- Making small qualitative scales commensurate with each other is easier.
- But the choice of aggregation operations is very limited (it ranges from impossibility theorems in ordinal case, to only min and max and their combination in the qualitative case).
- Finite value scales induce a strong lack of discrimination: the set of potential decisions can be sorted into as many groups of indifferent alternatives as the number of levels in the absolute scale. It is well-known people make little sense of refined absolute value scales (not more 7 levels).
- It is not clear how to handle bipolar information (pros and cons) ?

In fact there are discrete t-norms other than minimum on finite scales. The main alternative is Lukasiewicz discrete t-norm, that is a truncated sum. This choice underlies assumptions on the meaning of a qualitative scale $L = \{\lambda_0, \lambda_1, \ldots, \lambda_n\}$,

1. Like with 2-tuple method, L is mapped to the integers: $\lambda_i = i$. In particular, λ_i is understood as being i times stronger λ_1
2. There is a saturation effect that create counterintuitive ties when aggregating objective functions in this setting.

So this approach is not really qualitative, and not very attractive altogether at the practical level. In fact it is important to better lay bare the meaning of a qualitative value scale, and point out the assumptions motivating the restriction of aggregation operations to operations min and max. Using a qualitative scale, two effects can be observed

1. **Negligibility Effect:** Steps in the evaluation scale are far away from each other. It implies a strong focus on the most likely states of nature, on the most important criteria. This is what implies a lack of compensation between attributes. For instance, aggregating five ratings by the minimum, $\min(5,5,5,5,1) < \min(2,2,2,2,2)$: many 5's cannot compensate for a 1 and beat several 2's.

2. **Drowning effect:** There is no comparison of the number of equally satisfied attributes. The rating vector $(5, 5, 5, 5, 1)$ is worth the same as $(1, 1, 1, 1, 1)$ if compared by means of the min operation. It means that we refrain from counting.

It is clear that focusing on important criteria is something expected from human behavior [40]. However the equivalence between $(5, 5, 5, 5, 1)$ and $(1, 1, 1, 1, 1)$ is much more debatable and conflicts with the intuition, be it because the latter Pareto-dominates the former. The main idea to improve the efficiency of qualitative aggregation operations is to preserve the negligibility effect, but allow for counting. Note that if we build a preference relation on the set of alternatives rated on an absolute scale L on the basis of pairwise comparison made by the decision-maker, one may get chains of strictly preferred alternatives with length $m >\mid L \mid$. So, humans discriminate better on pairwise comparisons than using absolute value scales.

3.2. Refinements of qualitative aggregation operations

Let V be a set of alternatives, and assume a unique finite value scale L for rating n criteria, L being small enough to ensure commensurability. At one extreme, one may consider the smallest possible value scale $L = \{0, 1\}$ So each alternative is modelled by a Boolean vector $\vec{u} = (u_1, u_2, \ldots, u_n) \in \{0, 1\}^n$. Let $\succeq$ denote the overall preference relation over $\{0, 1\}^n$, supposed to be a weak order. Suppose without loss of generality that criteria are ranked in the order of their relative importance (criterion i as at least as important as criterion $i+1$.) Three principles for a qualitative aggregation operations should be respected for the aggregation to be rational

1. **Focus Effect**: If an alternative satisfies the most important criterion where the ratings differ then it should be preferred. Formally it reads as follows: given two vectors of ratings $\vec{u}$ and $\vec{v}$, if $u_i = v_i, i = 1, \ldots, k-1$, and $u_k = 1, v_k = 0$, where criterion k is strictly more important than criterion $k+1$, then $\vec{u} \succ \vec{v}$
2. **Compatibility with strict Pareto-dominance** (CSPD): If an alternative satisfies only a subset of criteria satisfied by another then the latter should be preferred.
3. **Restricted compensation**: If an alternative satisfies a number of equally important criteria greater than the number of criteria of the same importance satisfied by another alternative, on the most important criteria where some ratings differ, the former alternative should be preferred.

Strict Pareto-Dominance is defined for any value scale as $\vec{u} >_P \vec{v}$ iff $\forall i = 1, \ldots n, u_i \geq v_i$ and $\exists j, u_j > v_j$. Then the CSPD principle reads:

$$\vec{u} >_P \vec{v} \text{ implies } \vec{u} \succ \vec{v}.$$

Clearly, the basic aggregation operations min and max violate Pareto-Dominance. Indeed we may have $\min_{i=1,\ldots,n} u_i = \min_{i=1,\ldots,n} v_i$ while $\vec{u} >_P \vec{v}$. In fact, there is no strictly increasing function $f : L^n \to L$. So any aggregation function on a finite scale will violate strict Pareto-Dominance. But just applying the latter to L^n, the obtained partial order on V contains chains $\vec{v_1} >_P \vec{v_2} >_P \cdots >_P \vec{v_m}$ much longer that the numbers of elements in the value scale. Given that we take the negligibility effect for granted, the approach to mend these basic operations is thus not to change, but to refine them. Two known methods recover Pareto-dominance by refining the min-ordering (see [42] for a bibliography):

- Discrimin: $\vec{u} >_{dmin} \vec{v}$ iff $\min_{i:u_i \neq v_i} u_i = \min_{i:u_i \neq v_i} v_i$
- Leximin: Rank $\vec{u}$ and $\vec{v}$ in increasing order: let $\vec{u^\sigma} = (u_{\sigma(1)} \leq u_{\sigma(2)} \leq \ldots, u_{\sigma(n)})$ and $\vec{v^\tau} = (v_{\tau(1)} \leq v_{\tau(2)} \leq \ldots, v_{\tau(n)}) \in L^n$, then $\vec{u} >_{lmin} \vec{v}$ iff $\exists k, \forall i < k, u_{\sigma(i)} = v_{\sigma(j)}$ and $u_{\sigma(k)} > v_{\sigma(k)}$

The Discrimin method deletes vector positions that bear equal values in $\vec{u}$ and $\vec{v}$ prior to comparing the remaining components. The leximin method is similar but it cancel pairs of equal entries, one from each vector, regardless of their positions. Similar refinements of the maximum operation, say Discrimax and leximax can be defined.

Clearly, $\vec{u} >_P \vec{v}$ implies $\vec{u} >_{dmin} \vec{v}$ which implies $\vec{u} >_{lmin} \vec{v}$. So by constructing a preference relation that refines a qualitative aggregation operation, we recover a good behavior of the aggregation process without needing a more refined absolute scale.

The minimum and the maximum aggregation operations can be extended so as to account for criteria importance. Consider a weight distribution $\vec{\pi}$ that evaluates the priority of criteria, with $\max \pi_i = 1$. Consider the order-reversing map $\nu(\lambda_i) = \lambda_{m-i}$ on a scale with $m+1$ steps. The following extensions of the minimum and the maximum are now well-known:

- **Prioritized Maximum**: $P\max(\vec{u}) = \max_{i=1,\ldots n} \min(\pi_i, u_i)$
  Here $P\max(\vec{u})$ is high as soon as there is an important criterion with high satisfaction rating.
- **Prioritized Minimum**: $P\min(\vec{u}) = \min_{i=1,\ldots n} \max(\nu(\pi_i), u_i)$
  Here $P\min(\vec{u})$ is high as soon as all important important criterion get high satisfaction ratings.
- **Sugeno Integral**: $S_{\gamma,u}(f) = \max_{\lambda_i \in L} \ \min(\lambda_i, \gamma(U_{\lambda_i}))$
  where $U_{\lambda_i} = \{i, u_i \geq \lambda_i\}$ and $\gamma : 2^S \mapsto L$ ranks groups of criteria.

In the last aggregation scheme, $\gamma(A)$ is the priority degree of the group of criteria $A \subseteq \{1, \ldots, n\}$. It is a capacity, i.e. If $A \subseteq B$ then $\gamma(A) \leq \gamma(B)$. When γ is a possibility (resp. necessity) measure, i.e. $\gamma(A) = \max_{i \in A} \gamma(\{i\})$ (resp. $\min_{i \notin A} \nu(\gamma(\{i\}))$ then the Prioritized Maximum $P\max$ (resp. Minimum $P\min$) operation is retrieved. These operations have been used in decision under uncertainty (as a substitute to expected utility [44]) and for criteria aggregation with finite value scales [43]

The leximin and leximax operations can be extended in order to refine $P\max$ and $P\min$. The idea is as follows (Fargier, Sabbadin [45]). Given a totally ordered set $(\Omega, \trianglerighteq)$ the leximin and leximax relations $>_{lmin}$ and $>_{lmax}$ compare vectors in Ω^n, based on comparing values using the relation $\trianglerighteq$. Call these techniques Leximax($\trianglerighteq$), Leximin($\trianglerighteq$). For the leximin and leximax comparison of vectors of utility we use $\Omega = L$ and $\trianglerighteq = \geq$.

Iterating this scheme allows for a comparison of matrices with entries in L. Namely let $H = [h_{i,j}]$ be such a matrix, $\Omega = L^n$. The Leximax($\succeq_{lmin}$) relation (where $\trianglerighteq = \succeq_{lmin}$) can be used for comparing the rows $H_{j\cdot}$ of the matrix:

$$F \succeq_{lmax(lmin)} G \Leftrightarrow \begin{cases} \forall j, F_{(j)\cdot} \sim_{lmin} G_{(j)\cdot} \text{ or} \\ \exists i \text{ t.q. } \forall j > i, F_{(j)\cdot} \sim_{lmin} G_{(j)\cdot} \text{ and } F_{(i)\cdot} \succ_{lmin} G_{(i)\cdot} \end{cases}$$

where $H_{(i)\cdot} = i^{th}$ row of H, reordered increasingly. It takes the minimum on elements inside the rows, and the leximax across rows. To compute this ordering, we must shuffle the entries of each matrix so that to rank values on each line in increasing order, and rows top-down in decreasing lexicographic order. Then compare the two matrices lexicographically, first the top rows, then, if equal, the second top rows, etc..

Example: Consider the comparison of the two matrices.

$F =$

7	**3**	4	8	5
6	**3**	7	4	9
5	6	**3**	7	7

$G =$

8	**3**	**3**	5	9
3	7	**3**	8	4
7	**3**	8	5	5

It is clear that $\max_i \ \min_j \ f_{i,j} = \max_i \ \min_j \ g_{i,j}$. Reordering increasingly inside lines:

$F' =$

3	4	5	7	8
3	4	6	7	9
3	5	6	7	7

$G' =$

3	3	5	8	9
3	3	4	7	8
3	5	5	7	8

Then rows are rearranged reordered top down in the sense of leximax. It is clear that (see bold face entries):

3	5	**6**	7	7
3	4	6	7	9
3	4	5	7	8

$\succ_{lmax(lmin)}$

3	5	**5**	7	8
3	3	5	8	9
3	3	4	7	8

The Leximax($\succeq_{lmin}$) relation is a (very discriminative) complete and transitive relation. Two matrices are equally preferred if and only if they have the same coefficients up to a reshuffling (possibly in different positions). It refines the maximin comparison of matrices based on computing $\max_i \ \min_j \ h_{i,j}$. Likewise we can define Leximin($\succeq_{lmin}$), Leximax($\succeq_{lmax}$); Leximin($\succeq_{lmax}$).

These notions are instrumental to refine the prioritized maximum and minimum. In the prioritized case, alternatives $\vec{u}$ are encoded in the form of $n \times 2$ matrices $F^u = [f_{ij}]$ with $f_{i1} = \pi_i$ and $f_{i2} = u_i, i = 1, \ldots, n$: It is then clear that $P\max(\vec{u}) = \max_{i=1,n} \min_{j=1,2} f_{ij}$. Hence $P\max$ is refined by the *Leximax*(*Leximin*($\geq$)) procedure:

$$P\max(\vec{u}) > P\max(\vec{v}) \implies F^u \succ_{lmax(\succeq lmin)} F^v.$$

The prioritized minimum can be similarly refined applying Leximin($\succeq_{lmax}$) to matrices $F^u = [f_{ij}]$ with $f_{i1} = n(\pi_i)$ and $f_{i2} = u_i$. It is easy to verify that the Leximin($\succeq_{lmax}$) and Leximax($\succeq_{lmin}$) obey the three principles of focus effect, strict Pareto-dominance, and restricted compensation.

The same kind of trick can be applied to refine the ordering induced by Sugeno integrals. But the choice of the matrix encoding alternatives depend on the form chosen for expressing this aggregation operation. For instance, using the form proposed above, you can choose $f_{i1} = \lambda_i, f_{i2} = \gamma(U_{\lambda_i})$. This choice is not the best one, as shown in [46]. Moreover, part of the lack of discrimination is due to the capacity γ itself that estimates the importance of groups of criteria. In order to refine the capacity γ one idea is to generalize the leximax refinement of possibility measures [47]. To this end, it is useful to consider the so-called "Qualitative" Moebius transform [48] of γ:

$$\begin{aligned} \gamma^{\#}(A) &= \gamma(A) \text{ if } \gamma(A) > \max_{B \subsetneq A} \gamma(B) \\ &= 0 \ \text{ otherwise.} \end{aligned}$$

It is such that $\gamma(A) = \max_{E \subseteq A} \gamma_{\#}(E)$, that is $\gamma^{\#}$ contain the minimal amount of information to recover γ. It is clear that if γ is a possibility measure, then $\gamma_{\#}(E) > 0$ only if E is a singleton, i.e. $\gamma_{\#}$ is a possibility distribution. The leximax refinement of $\gamma_{\#}$ is then just obtained by comparing for two groups of criteria A and B the sets $\{\gamma_{\#}(E), E \subseteq A\}$ and $\{\gamma_{\#}(E), E \subseteq B\}$ using leximax. More details on these issues can be found in [46, 49]

An interesting question is to define counterparts of discrimin and leximin procedures to any (monotonic) aggregation operation $f : L^2 \to L$ on qualitative scales. The question also makes sense for continuous scales (see[50]).

3.3. Numerical encoding of qualitative aggregation functions

Additive encoding of the leximax and leximin procedures exist for a long time when the number of alternatives to be compared or the evaluation scale is finite. (such an encoding is not possible in the continuous case). A mapping $\phi : L \to [0, 1]$, where $L = \{0 < \lambda_1 < \lambda_2 < \cdots < \lambda_m = 1\}$, is said to be n-super-increasing if and only if $\phi(\lambda_i) > n\phi(\lambda_{i-1}), \forall i = 2, ..m$. We also assume $\phi(0) = 0$ and $\phi(\lambda_m) = 1$. Mapping ϕ is also called *big-stepped*. It is clear that for any such mapping,

$$\max_{i=1,\dots n} u_i > \max_{i=1,\dots n} v_i \text{ implies } \sum_{i=1,\dots n} \phi(u_i) > \sum_{i=1,\dots n} \phi(v_i)$$

e.g. $\phi(\lambda_i) = k^{i-m}$ for $k > n$ achieves this goal.The worst case is when $\max(0, 0, \dots, 0, \lambda_i) > \max(\lambda_{i-1}, \dots, \lambda_{i-1})$.
This is a numerical representation of the leximax ordering:

Property: $\vec{u} \succ_{lmax} \vec{v}$ if and only if $\sum_{i=1,\dots n} \phi(u_i) > \sum_{i=1,\dots n} \phi(v_i)$.

Now consider the big-stepped mapping $\psi(\lambda_i) = \frac{1-k^{-i}}{1-k^{-m}}, k > n$. Again it holds that:

$$\min_{i=1,\dots n} u_i > \min_{i=1,\dots n} v_i \text{ implies } \sum_{i=1,\dots n} \psi(u_i) > \sum_{i=1,\dots n} \psi(v_i)$$

And it offers a numerical representation of the leximin ordering.

Property: $\vec{u} \succ_{lmin} \vec{v}$ if and only if $\sum_{i=1,\dots n} \psi(u_i) > \sum_{i=1,\dots n} \psi(v_i)$.

These representation results have been extended to the above refinements of the prioritized maximum and minimum [45] by means of weighted averages involving super increasing sequences of numerical values. Namely, there exists a weighted average, say $AV_+(\vec{u})$, representing $\succeq_{lmax(\succeq_l min)}$ and thus refining $P\max$. Namely consider a super increasing transformation χ of the scale L such that:

$$\max_i \min(\pi_i, u_i) > \max_i \min(\pi_i, v_i) \implies \sum_{i=1,\dots n} \chi(\pi_i)\cdot\chi(u_i) > \sum_{i=1,\dots n} \chi(\pi_i)\cdot\chi(u_i).$$

The worst case is when:
$\max(\min(\lambda_j, \lambda_j), 0, \dots, 0) > \max(\min(\lambda_j, \lambda_{j-1}), \min(1_L, \lambda_{j-1}), \dots, \min(1_L, \lambda_{j-1}))$
Hence the sufficient condition:

$$\forall j \in \{0, \dots, m-1\}, \chi(\lambda_j)^2 > (n+1)\chi(\lambda_{j-1}) \cdot \chi(1_L)$$

The following result then holds:

$\vec{u} \succ_{lmax(lmin)} \vec{v}$ if and only if $\sum_{i=1,\dots n} \chi(u_i)\chi(\pi_i) > \sum_{i=1,\dots n} \chi(v_i)\chi(\pi)$.

The values $\chi(\pi_i), i = 1, \dots, n$ can be normalized in such a way as to satisfy $\sum_{i=1}^{n} \chi(\pi_i) = 1$ so that we do use a weighted average to represent $\succ_{lmax(lmin)}$. In order to refine Sugeno integral by mapping it to a numerical scale. The idea is to use a Choquet integral. However, in order to get a minimally redundant expression of Sugeno integral, it must be put in the form

$$S_\gamma(\vec{u}) = \max_{A\subseteq N} \min(\gamma^{\#}(A), \min_{i\in A} u_i)$$

where $\gamma^{\#}(A)$ is the above defined qualitative Moebius transform. We can use use a super-increasing transformation of $\gamma_{\#}$ into a mass function $m_{\#} : 2^S \mapsto [0,1] : m_{\#}(E) = \chi(\gamma_{\#}(E))$ in the sense of Shafer [51], such that $\sum_{E\subseteq \mathcal{C}} m_{\#}(E) = 1$. The above leximax refinement of the ranking induced by γ can then be represented by means of the belief function $Bel(A) = \sum_{E\subseteq A} m_{\#}(E)$ When γ is a possibility measure, the refining belief function is a probability measure. The Sugeno integral can then be refined by the Choquet integral of the form (see [46, 49] for details):

$$Ch_{\#}^{lsug}(\vec{u}) = \sum_{A\subseteq S} m_{\#}(A) \cdot \min_{s\in A} \chi(u_i).$$

The lessons drawn from this line of study is that the discrimination power of qualitative aggregation methods (which in some sense are the off-springs of the Bellman-Zadeh decision framework) can be drastically increased by lexicographic refinement techniques that respect the qualitative nature of the preference information as well as the focus effect on most important issues observed in human decision behavior. Moreover these refinement techniques bring us back to standard numerical aggregation methods, that, through the use of super-increasing transformations, are robust (because qualitative in essence) contrary to many number crunching-preference aggregation methods.

3.4. Bipolarity in qualitative evaluation processes

Cumulative Prospect Theory, due to Tversky & Kahneman [23] was motivated by the empirical finding that people, when making decisions, do not perceive the worth of gains and losses in the same way. This approach assesses the importance of positive affects and negative affects of decisions *separately*, by means of two monotonic set functions $g^+(A^+)$, $g^-(A^-)$, which respectively evaluate the importance of the set of criteria A^+ where the alternatives a score favorably and for the set of criteria A^-, where they score unfavorably. For instance, one can separately compute the expected utility of the losses and of the gains, using different utility functions. Then they suggest to compute the so-called net predisposition $N(a) = g^+(A^+) - g^-(A^-)$ of each decision a in order to rank-order them in terms of preference. This view is at odds with classical decision theory where there is no distinction between gains and losses. Couched in terms of fuzzy sets, the CPT approach, which relies on the idea of bipolar information, is akin to a form of independence between membership and non-membership grades in the spirit of Atanassov [27]. However, decision works inspired by the misleadingly called intuitionistic fuzzy sets never make the connection with CPT.

This line of thought, was recently extended to the case where positive and negative affects are not independent (see [22]):

- Using bi-capacities on a bipolar scale in the form of functions $N(a) = g(A^+, A^-)$ monotonic in the first place, antimonotonic with the second one.
- So-called bipolar capacities $N(a) = (g^+(A^+, A^-), g^-(A^+, A^-))$ living on bivariate unipolar scales, keeping the positive and the negative evaluations separate.

An interesting question is then how to evaluate decisions from *qualitative* bipolar information, namely how to extend the min and max rules if there are both positive and negative arguments? This question was recently discussed by the author with colleagues [52, 53]. In the proposed simplified setting, a set of Boolean criteria (C) is used, each of which has a polarity $p = +$ (positive) or $-$ (negative). Such criterial are called affects. For instance, when buying a house, the presence of a garden is a positive affect; the location of the house in a noisy or dangerous environment is a negative affect. Each affect is supposed to possess an importance level in a qualitative scale L. The focus effect is assumed in the sense that the order of magnitude of the importance of a group A of affects with a prescribed polarity is the one of the most important affect, in the group ($\gamma(A) = \max_{x \in A} \pi_i$ is a possibility measure).

Preference between two alternatives a and b is then achieved by comparing the pairs $(\Pi(A^-), \Pi(A^+))$ and $(\Pi(B^-), \Pi(B^+))$, evaluating separately positive and negative affects of a (respectively $A^+ = \{i, p(i) = +, u_i = 1\}$ and $A^- = \{i, p(i) = -, u_i = 1\}$), based on the relative importance of these affects. The most natural rule that comes to mind is called the Bipolar Possibility Relation The principle at work (that plays the role of computing the net predisposition in the qualitative setting) is:
Comparability of positive and negative affects: When comparing Boolean vectors $\vec{u}$ and $\vec{v}$, a negative affect for $\vec{u}$ (resp. against $\vec{v}$) is a positive argument pro $\vec{v}$ (resp. pro $\vec{u}$).
Along with the way group of affects are weighted, the decision-maker is supposed to focus on the most important affect regardless of its polarity. The following decision rule follows [53]:

$$a \succeq^{Biposs} b \iff \max(\Pi(A^+), \Pi(B^-)) \geq \max(\Pi(B^+), \Pi(A^-))$$

The relation $\succeq^{Biposs}$ is complete, but only its strict part is transitive. This relation collapses to the Bellman-Zadeh minimum aggregation rule if all affects are negative and to the maximum rule if all affects are positive (which also has something to do with Atanassov connectives).This is similar to the CPT approach where: $a > b \iff g^+(A^+) + g^-(B^-) > g^-(B^+) + g^-(A^-)$, the possibilistic rule being obtained by changing + into max. This decision rule is sound and cognitively plausible but it is too rough as it creates too many indifference situations.

Refinements of this decision rule can build on top of an idea originally due to B. Franklin: canceling arguments of equal importance for $\vec{u}$ or against $\vec{v}$, by arguments for b or against a until we find a difference on each side. This leads to a complete and transitive refinement of $\succeq^{Biposs}$. Let $A^+_\lambda = \{i \in A^+, \pi_i = \lambda\}$ be the arguments for a with strength λ. (resp. A^-_λ the arguments against a with strength λ.).

$$a \succeq^{Lexi} b \iff \exists \lambda \in L \text{ such that} \begin{cases} (\forall \beta > \lambda, & |A^+_\beta| - |A^-_\beta| = |B^+_\beta| - |B^-_\beta|) \\ \text{and} & (|A^+_\lambda| - |A^-_\lambda| > |B^+_\lambda| - |B^-_\lambda|) \end{cases}$$

This decision rule checks the positive and negative affects for each element of a pair (a, b) of alternatives top-down in terms of importance. It stops at the maximal level of importance when there are more arguments for one than for the other (using the comparability postulate). This rule generalizes Gigerenzer's "take the best" heuristic [40] and can be encoded in the cumulative prospect theory framework using super-increasing transformations. It has been empirically tested, and proves to be the one people often use when making decisions according to several criteria [52].

4. Uncertainty handling in decision evaluation using fuzzy intervals

Fuzzy intervals have been widely used in FDA so as to account for the fact that many evaluations are imprecise. In many cases, it comes down to applying the extension principle to existing evaluation tools: weighted averages or expected utilities using fuzzy interval weights [10], fuzzy extensions of Saaty's Analytical Hierarchical Process[54], and other numerical or relational MCDM techniques (like TOPSIS, PROMETHEE, ELECTRE,..[55]). Many of these techniques, in their original formulations are ad hoc, or even debatable (see [33] for a critique of many of them). So their fuzzy-valued extensions are often liable of the same defects.

The problem with fuzzy set methods extending existing ones is that more often than not the proposed handling of fuzzy intervals is itself ad hoc or so approximate that the benefit of the fuzzification is lost. Moreover the thrust of the fuzzy interval analysis is to provide information about the uncertainty pervading the results of the decision process. Some authors make an unjustified use of defuzzification that erases all information of this type. For instance the decision-maker is asked for some figures in the form of fuzzy intervals so as to account for the difficulty to provide precise ratings, and then these ratings are defuzzified right away. Or the fuzzy intervals are propagated through the decision process but the final results are defuzzified in order to make the final decision ranking step easier. In such procedures it is not clear why fuzzy intervals were used in the first stand. The uncertainty pervading the ratings should play a role in the final decision-making process, namely to warn the decision maker when information is not sufficient for justifying a clear ranking of alternatives.

In this section we discuss two examples where fuzzy intervals have been extensively used, but where some intrinsic technical or computational difficulties need to be properly addressed in decision evaluation techniques: fuzzy weighted averages and fuzzy AHP methods.

4.1. Fuzzy weighted averages

An obvious way of introducing fuzzy sets in classical aggregation techniques is to assume that local evaluations are ill-known and represented by fuzzy intervals. The question is then how to aggregate fuzzy evaluations and how to provide a ranking of alternatives. Fuzzy weighted averages is a good example of such a problem, that dates back to Baas and Kwakernaak [59]. There have been numerous papers on fuzzy weighted averages since then (see [56] for a bibliography before 2000 and [10] for a recent survey). The key technique involved here is fuzzy interval analysis [56]. Two important points need to be stressed, that are not always acknowledged

- Computing arithmetic expressions with fuzzy intervals cannot be always be done by means of plain fuzzy arithmetics (that is, combining partial results obtained by means of the four operations).
- Before considering a fuzzy interval approach, one must understand how to solve the problem with plain intervals.

The first point is actually a corollary of the second one. In particular the name "fuzzy number" seems to have misled many authors who seem not to realize that what is often called a fuzzy number is a generalized (gradual) interval. For instance some authors have tried to equip fuzzy addition with a group structure, which is already lost for intervals. Many expect the defuzzification of fuzzy numbers to yield a precise number, while stripping a fuzzy set from its fuzziness naturally yields a set. These points are discussed at length in [57] where a genuine fuzzy extension of a number (a gradual number) is suggested, such that a fuzzy interval is a standard interval of such gradual numbers. In this paper we systematically use the name *fuzzy interval* to remove all ambiguity.

As a consequence, fuzzy interval analysis inherits all difficulties encountered when computing with intervals [58]. In particular if a given ill-known quantity appears several times in some arithmetic expression, one must be very careful to observe that substituting this quantity with the same interval in several places does not preserve the constraint stating that behind this interval lies the same quantity. For instance, if $x \in [a,b]$, then $[a,b] - [a,b] \neq 0$, while $x - x = 0$ regardless of the value of x. More generally interval analysis, hence fuzzy interval analysis requires an optimization problem to be solved. Computing fuzzy weighted averages has more to do with constraint propagation than with the arithmetics of fuzzy numbers.

The difficulty is already present with imprecise (interval weights). The problem of computing interval-valued averages can be posed in two ways:

1. Maximise and minimize $\frac{\sum_{i=1}^{n} x_i \cdot w_i}{\sum_{i=1}^{n} w_i}$ under the constraints: $w_i \in [a_i, b_i] \subset [0, +\infty), x_i \in [c_i, d_i], i = 1, \ldots, n$;
2. Maximise and minimize $\sum_{i=1}^{n} x_i \cdot p_i$ under the constraints: $p_i \in [u_i, v_i] \subseteq [0, 1], x_i \in [c_i, d_i], i = 1, \ldots, n, \sum_{i=1}^{n} p_i = 1$.

The two problems yield different results if the same intervals $[u_i, v_i] = [a_i, b_i]$ are used. In both cases, the maximal (resp. minimal) solution is attained for $x_i = d_i$ (resp. $x_i = c_i$). In the second problem, one issue is: what does normalization mean when only intervals $([u_1, v_1], [u_2, v_2] \ldots, [u_n, v_n])$ are available ? It is clear that the sum of these intervals is an interval, hence never equal to 1. One way out is to view a vector of interval weights as a set of standard normalized vectors $\vec{p} = (p_1, p_2 \ldots, p_n)$.

Specific conditions must be satisfied if all bounds are to be reachable by such weight vectors, that is $\forall i = 1 \ldots n, \exists (p_1, p_2 \ldots, p_n) \in \prod_{i=1}^{n} [u_i, v_i]$ such that $p_i = u_i$ and another such vector such that $p_i = v_i$. The question is completely solved by De Campos et al. [61] in the case of imprecise probability weights. Necessary and sufficient conditions are (the second one implies the first one)

1. $\sum_{i=1}^{n} u_i \leq 1 \leq \sum_{i=1}^{n} v_i$ (non-emptiness)
2. $u_i + \sum_{j \neq i} v_j \geq 1$; $v_i + \sum_{j \neq i} u_j \leq 1$ (attainability).

Attainability is a form of arc-consistency in the terminology of constraint satisfaction. Updating the bounds of the intervals $[u_i, v_i]$ so that this condition be satisfied can be viewed as a form of normalization.

Fast methods exist for the second problem, with explicit expressions whose computation is linear in the number of arguments [60]. Now the first problem can be connected to the second one as follows:

$$\{\frac{\sum_{i=1}^n x_i \cdot w_i}{\sum_{i=1}^n w_i}, w_i \in [a_i, b_i]\} = \{\sum_{i=1}^n x_i \cdot p_i, p_i = \frac{w_i}{\sum_{i=1}^n w_i}, w_i \in [a_i, b_i]\}$$
$$= \{\sum_{i=1}^n x_i \cdot p_i, p_i \in [\frac{a_i}{a_i + \sum_{j\neq i} b_j}, \frac{b_i}{b_i + \sum_{j\neq i} a_j}], \sum_{i=1}^n p_i = 1\}$$

That these bounds are attainable for the normalized vectors $(p_1, p_2 \ldots, p_n)$ is established in [62]. Note that the range

$$\{\frac{\sum_{i=1}^n x_i \cdot q_i}{\sum_{i=1}^n q_i}, q_i \in [\frac{a_i}{a_i + \sum_{j\neq i} b_j}, \frac{b_i}{b_i + \sum_{j\neq i} a_j}]\}$$

strictly contains the former one since there is no longer any constraint on $\sum_{i=1}^n q_i$.

In the case of fuzzy weight vectors, the above reasoning must be applied to α-cuts of the fuzzy intervals involved. It should be clear that a fuzzy weight vector $\vec{\tilde{p}} = (\tilde{p}_1, \tilde{p}_2 \ldots, \tilde{p}_n)$ should be actually viewed as a fuzzy set of normalized weight vectors $\vec{p}$, where attainability conditions are met for all interval weight vectors $((\tilde{p}_1)_\alpha, (\tilde{p}_2)_\alpha \ldots, (\tilde{p}_n)_\alpha)$ formed by cuts. The degree of membership of $\vec{p}$ in $\vec{\tilde{p}}$ is equal to $\min_{i=1}^n \mu_{\tilde{p}_i}(p_i)$.

The above problems are considered by scholars dealing with type 2 fuzzy sets for calculating (fuzzy-valued) centroids for instance, but these authors do not seem to rely on existing results in fuzzy interval analysis and sometimes reinvent their own methods. Extensions of such calculations to Choquet integrals with fuzzy-valued importance weights for groups of criteria more difficult, as the issue of ranking intervals $[c_i, d_i]$, let alone fuzzy intervals, must be addressed within the calculation [63].

4.2. Fuzzy extensions of Analytical Hierarchy Process

A number of multicriteria decision-making methods have been extended so as to deal with fuzzy data. Here we confine ourself to the case of Saaty's Analytical Hierarchy Process [64]. Numerous fuzzy versions of Saaty's methods have been proposed (since Van Laaroven and Pedrycz, [65], see the bibliography in [56]). Many such proposals seem to pose and solve questionable fuzzy equations, as we shall argue.

The principle of the AHP method relies on the following ideal situation

- The pairwise relative preference of n items (alternatives, criteria) is modelled by a $n \times n$ *consistent* preference matrix A, where each coefficient a_{ij} is supposed to reflect how many times item i is preferred to item j.
- A consistent preference matrix is one that is reciprocal in the sense that $\forall i, j, a_{ij} = 1/a_{ji}$ and product-transitive ($\forall i, j, k, a_{ij} = a_{ik} \cdot a_{kj}$).

- Then its largest eigenvalue is $\lambda = n$ and there exists a corresponding normalized weight vector $p = (p_1, p_2 \ldots, p_n)$ with $\forall i, j, a_{ij} = \frac{p_i}{p_j}$

Even if widely used, this method is controversial and has been criticised by MCDM scholars as being ill-founded at the measurement level, and having paradoxical lacks of intuitive invariance properties (the scale used is absolute, with no degree of freedom, see [33], for instance). Moreover in practice, pairwise comparison data do not provide consistent matrices. Typically, the decision-maker provides, for each pair (i, j), a value $v_{ij} \in \{2, \ldots 9\}$ if i is preferred to j v_{ij} times, $v_{ij} = 1$ if there is indifference. The matrix A with coefficients $a_{ij} = v_{ij}, a_{ji} = 1/a_{ij}$ if $v_{ij} \geq 1$ is then built. Generally product transitivity is not empirically achieved. A preference matrix $A$is considered all the more consistent as the largest eigenvalue of A is close to n. Then the derived normalized weights form the eigenvector of A for this eigenvalue.

Asking for precise values v_{ij} is debatable, because these coefficients are arguably imprecisely known. So many researchers have considered fuzzy valued pairwise comparison data. The fuzzification of Saaty's AHP has consisted to extend the computation scheme of Saaty with fuzzy intervals. However this task turned out to be difficult for several reasons.

- Replacing a consistent preference matrix by a fuzzy-valued preference matrix loses the properties of the former. The reciprocal condition $\tilde{a}_{ij} = 1/\tilde{a}_{ji}$ no longer implies $\tilde{a}_{ij} \cdot \tilde{a}_{ji} = 1$, nor can the product transitivity property hold in the form $\tilde{a}_{ij} = \tilde{a}_{ik} \cdot \tilde{a}_{kj}$ when $\tilde{a}_{ij}$ are fuzzy intervals.

- Fuzzy eigen-values or vectors of fuzzy-valued matrices are hard to define in a rigorous way: writing the the usual equation $Ap = \lambda p$ replacing vector and matrix entries by fuzzy intervals leads to overconstrained equations.

- It is tempting to solve the the problems with each interval-matrix defined from α-cuts $(\tilde{a}_{ij})_\alpha = [\underline{a_{ij}}_\alpha, \overline{a_{ij}}_\alpha]$ of the fuzzy coefficients, as done by Csutora and Buckley [66] for instance. These authors suggest to solve the eigenvalue problem for the two extreme preference matrices with respective coefficients $\underline{a_{ij}}_\alpha$ and $\overline{a_{ij}}_\alpha$, with the view to find an interval-valued eigenvalue. However these boundary matrices are not even reciprocal since if $\tilde{a}_{ij} = 1/\tilde{a}_{ji}$, $\underline{a_{ij}}_\alpha = 1/\overline{a_{ji}}_\alpha$, not $1/\underline{a_{ji}}_\alpha$. So the meaning of normalized weights computed from these boundary matrices is totally unclear.

The bottom-line is that the natural extension of a simple crisp equation $ax = b$ (let alone an eigenvalue problem) is not necessarily a fuzzy equation of the form $\tilde{a}\tilde{x} = \tilde{b}$ where fuzzy intervals are substituted to real numbers and equality of membership function on each side is required.

- The first equation $ax = b$ refers to a constraint to be satisfied by a model referring to a certain reality.

- But fuzzy intervals $\tilde{a}, \tilde{x}, \tilde{b}$ only represent *knowledge* about actual values a, x, b

- Even if $ax = b$ is taken for granted, it is not clear why the knowledge about ax should be equated to the knowledge about b ($\tilde{a}$, and $\tilde{b}$ may derive from independent sources). The objective constraint $ax = b$ only enforces a consistency condition $\tilde{a}\tilde{x} \cap \tilde{b} \neq \emptyset$.

- If indeed a fuzzy set $\tilde{x}$ is found such that $\tilde{a}\tilde{x} = \tilde{b}$, it does not follows that the actual quantities a, x, b verify $ax = b$. Moreover, equation $\tilde{a}\tilde{x} = \tilde{b}$ may fail to have solutions.

Recently, Ramik and Korviny [67] proposed to compute the degree of consistency of a fuzzy preference matrix $\tilde{A}$ whose entries are triangular fuzzy intervals $\tilde{a}_{ij}$ and $\tilde{a}_{ji}$ with respective supports and modes:

- for $\tilde{a}_{ij}$: $[\underline{a_{ij}}, \overline{a_{ij}}]$ and a_{ij}^M;

- for $\tilde{a}_{ji}$: $[1/\overline{a_{ij}}, 1/\underline{a_{ij}}]$ and $1/a_{ij}^M$

Their consistency degree is construed as the minimal distance between $\tilde{A}$ and a fuzzy consistent matrix $\tilde{X}$ understood as a so-called ratio matrix with coefficients of the form $\tilde{x}_{ij} = \frac{\tilde{x_i}}{\tilde{x}_J}$, where $\tilde{x_i}$ is a triangular fuzzy interval with mode x_i^M, and the normalisation condition $\sum_{i=1}^{n} x_i^M = 1$ is assumed. The distance $d(\tilde{A}, \tilde{X})$ used is a scalar distance between the vectors formed by the three parameters of the triangular fuzzy intervals, chosen such that the solution of the problem can be analytically obtained in agreement with the geometric mean method of computation of weights, already used in Van Laaroven and Pedrycz [65]. Fuzzy weights $\tilde{x}_i$ are thus obtained.

Let alone the fact that the formulation of the problem is partially ad hoc due to triangular approximation of inverses of triangular fuzzy intervals, and due to the choice of the normalisation condition (see the previous subsection), this approach also suffers from the above epistemological flaw consisting is considering a fuzzy interval as a simple substitute to a precise number, whereby a direct extension of the standard method consists just in replacing numbers by fuzzy intervals and running a similar computation as in the precise case. In particular the distance $d(\tilde{A}, \tilde{X})$ arguably evaluates an *informational* proximity between epistemic states (states of knowledge) about preference matrices, and says little about the scalar distance between the underlying precise ones.

Instead of viewing fuzzy interval preference matrices as fuzzy *substitutes* to precise ones, one may on the contrary acknowledge fuzzy pairwise preference data as *imprecise knowledge about regular preference information.* The fuzzy interval preference matrix is then seen as constraining an ill-known precise*consistent* comparison matrix. Inconsistencies in comparison data are thus explicitly explained by the imprecise nature of human-originated information. Such a constraint-based view of fuzzy AHP has been explained by Ohnishi and colleagues [68].

Namely consider a fuzzy matrix $\tilde{A}$ with entries $\tilde{a}_{ij} = 1/\tilde{a}_{ji}$ and $\tilde{a}_{ii} = 1$ and denote by μ_{ij} the membership function of $\tilde{a}_{ij}$. The relevance of a consistent preference matrix A to the user preference data described by $\tilde{A}$ can be evaluated without approximation as

$$\mu_{\tilde{A}}(A) = \min_{i,j:i<j} \mu_{ij}(a_{ij}).$$

A given normal weight vector $\vec{p} = (p_1, p_2 \ldots, p_n)$ satisfies the fuzzy preference matrix $\tilde{A}$ to degree $\mu(\vec{p}) = \min_{i<j} \mu_{ij}(\frac{p_i}{p_j})$. The degree of consistency of the preference data is

$$Cons(\tilde{A}) = \sup \mu(\vec{p}) = \sup_{\vec{p}:a_{ii}=\frac{p_i}{p_j}} \mu_{\tilde{A}}(A).$$

The best induced weight vectors are the Pareto maximal elements among $\{\vec{p}, \mu(\vec{p}) = Cons(\tilde{A})$. The reader is referred to [68] for details. It is interesting to contrast this methodology, where the problem can be posed without approximation, and the otherwise elegant one by Ramik and Korviny [67]. In their method , the fuzzy matrix $\tilde{X}$ can be viewed as an approximation of the imprecise information matrix $\tilde{A}$ such that $Cons(\tilde{X}) = 1$ in the sense of Ohnishi et al. (the core matrix X^M is consistent in the sense of Saaty by construction). But the second approach seems to be more respectful of the original imprecise preference data, that can be directly exploited.

The constrained-based approach does not solve the difficulties linked to the critiques addressed to Saaty's method, but its interpretation is much clearer than its direct fuzzification, in the sense that it does not require a new theory of fuzzy eigenvalues, nor does it pose fuzzy interval equations with debatable meanings. It makes sense if it is taken for granted that human preference can be ideally modelled by consistent preferences matrices in the sense of Saaty. The approach outlined above only tries to cope with the problem of inconsistency of human judgments by acknowledging their lack of precision.

5. Comparing Fuzzy intervals: a Constructive Setting

There is an enormous literature on fuzzy interval ranking methods, but very few attempts at proposing a rational approach to the definition of ranking criteria. This section tries to suggest one possible approach towards a systematic classification of ranking indices and fuzzy relations induced by the comparison of fuzzy intervals. The issue of ranking objects rated by fuzzy intervals should be discussed in the perspective of decision under uncertainty. The connection between the ranking of fuzzy intervals and fuzzy preference relations will be outlined. There are many ranking methods surveyed elsewhere [11, 56]. Here we suggest a unifying principle: such ranking methods should be directly based on probabilistic notions of dominance (see Chapter 8 in [26]) or their possibilistic counterparts on the one hand, and interval orders [71] on the other hand.

While many ranking methods have been proposed (and still are), most of the time they are derived on an ad hoc basis: an often clever index is proposed, sometimes for triangular fuzzy intervals only, and its merits are tested on a few examples. Systematic comparison studies are not so numerous (except for Bortolan and Degani [69], Lee [70], for instance). Moreover these comparison are based on intuitive feelings of what a good ranking method should do, tested on a few examples and conterexamples. There is a lack of first principles for devising well-founded techniques. However Wang and Kerre [11] proposed an interesting set of axioms that any preference relation $\succeq$ between fuzzy intervals $\tilde{a}, \tilde{b}$.. should satisfy. For instance

- *Reflexivity* : $\tilde{a} \succeq \tilde{a}$,
- *Certainty of dominance*: If $\tilde{a} \cap \tilde{b} = \emptyset$ then $\tilde{a} \succ \tilde{b}$ or $\tilde{b} \succ \tilde{a}$.
- *Consistency with fuzzy interval addition*: $\tilde{a} \succeq \tilde{b}$ implies $\tilde{a} + \tilde{c} \succeq \tilde{b} + \tilde{c}$

Dubois, Kerre et al. [56] also classified existing methods distinguishing between

1. *scalar indices* (based on defuzzification understood by the replacement of a fuzzy interval by a representative number)
2. *goal-based indices*: computing the degree of attainment of a fuzzy goal by each fuzzy interval, which bears some similarity to the expected utility approach, understanding a fuzzy goal as a utility function
3. *relational indices*: based on computing to what extent a fuzzy interval dominates another. In this case, some methods are based on metric considerations (those based on computing possibility and necessity of dominance), and others are based on computing areas limited by the membership functions of the fuzzy intervals to be compared.

Another view of the comparison of fuzzy intervals can exploit links between fuzzy intervals and other settings: possibility, probability theories and interval orders. This kind of idea can be found in some previous papers in the literature, but it has never been systematically explored. Yet, it might provide a systematic way to produce comparison indices and classifying them. In this section, we outline such a research program.

5.1. Four views of fuzzy intervals

A fuzzy interval $\tilde{a}$, like any fuzzy set is defined by a membership function $\mu_{\tilde{a}}$. This fuzzy set is normalized ($\exists x \in \mathbb{R}, \mu_{\tilde{a}}(x) = 1$ and its cuts are bounded closed intervals. Let us denote by $[\underline{a}, \overline{a}]$ its core and $[a_*, a^*]$ its support.

Like any fuzzy set it needs to be cast inside an interpretive setting in order to be usefully exploited. To our knowledge there are four existing view of fuzzy intervals understood as a representation of uncertainty

1. *Metric possibility distributions*: in this case the membership function of $\tilde{a}$ is viewed as a possibility distribution π_x, following the suggestion of Zadeh[16]. A fuzzy interval represent gradual incomplete information about some ill-known precise quantity x: $\pi_x(r)$ is all the greater as $r \in \mathbb{R}$ is close to a totally possible value. Moreover, we consider the interval $[0, 1]$ as a similarity scale and the membership function as a rescaling of the distance between elements on the real line.

2. *One point-coverage functions of nested random intervals*: In this case, a fuzzy interval $\tilde{a}$ is induced by the Lebesgue measure ℓ on $[0,1]$, and the cut multi-mapping with range in the set closed intervals of the real line $\mathcal{I}(\mathbb{R})$:

$$[0,1] \to \mathbb{R} : \alpha \mapsto \tilde{a}_\alpha \in \mathcal{I}(\mathbb{R}).$$

   Then $\pi_x(u) = \ell(\{\alpha, x \in \tilde{a}_\alpha\})$. This view comes from the fact that a numerical necessity measure is a special case of belief functions and a possibility distribution s a one-point coverage a random set [72]. In this case, the membership function of $\tilde{a}$ is viewed as the contour function of a consonant continuous belief function. This line has been followed by Dubois and Prade [73] and Chanas and colleagues [74]. Chanas [74] also envisaged a more general framework in the form of random intervals limited by two random variables $\dot{x} \leq \ddot{x}$ with disjoint support, such that $\pi_x(u) = Prob(\dot{x} \leq u \leq \ddot{x})$. Then the nestedness property is lost and the possibility distribution thus obtained is no longer equivalent to the knowledge of the two random variables: they lead to a belief function on the real line in the sense of Smets [75]
3. *Families of probability functions*: As formally a possibility measure is a special case of belief function, and a belief function is a special case of (coherent) lower probability in the sense of Walley [76], a fuzzy interval $\tilde{a}$ also encodes a special family of probability measures

$$\mathcal{P}_{\tilde{a}} = \{P : P(\tilde{a}_\alpha) \geq 1 - \alpha, \alpha \in [0,1]\}.$$

   Namely it can be shown that for fuzzy intervals $\tilde{a}$, $\Pi(A) = \sup\{P(A), P \in \mathcal{P}_{\tilde{a}}\}$ for all measurable subsets A of the real line. This approach proposed by Dubois and Prade [77] was studied by De Cooman and Aeyels [78]. It is clear that it allows probabilistic inequalities (like Chebychev's, see [18]) to be interpreted in terms of fuzzy intervals.
4. *Intervals bounded by gradual numbers*: This is a more recent view advocated by Fortin et al.[57]. The idea is to view a fuzzy interval as a regular interval of functions. To this end, the interval component of the fuzzy interval must be disentangled from its fuzzy (or gradual) component. A gradual number $\check{r}$ is a mapping from the positive unit interval to the reals: $\alpha \in (0,1] \mapsto r_\alpha \in \mathbb{R}$. For instance, the mid-point of a fuzzy interval $\tilde{a}$ with cuts $[\overline{a}_\alpha, \underline{a}_\alpha]$ is a gradual number $\breve{a}_\alpha = \frac{\underline{a}_\alpha + \overline{a}_\alpha}{2}$. It is clear that a fuzzy interval can be viewed as an interval of gradual numbers lower bounded by $\underline{a}_\alpha$ and upper-bounded by $\overline{a}_\alpha$. Gradual numbers inside the fuzzy interval $\tilde{a}$ can be generated by selecting a number inside each cut of $\tilde{a}$. In fact, they are the selection functions of the cut-mapping. Although the idea of using a pair of functions to represent a fuzzy interval is not new (e.g. the so-called *L-R fuzzy numbers*, that enable closed forms of arithmetic operations on fuzzy intervals to be derived in terms of inverses of shape functions L and R[56]), the key novelty here is to treat a fuzzy interval as a regular one.

What is clear from the above classification is that ranking fuzzy intervals should be related to techniques for ranking intervals or for ranking random quantities. There are well-known methods for comparing intervals, namely

1. *Interval orders*: $[a, b] >_{IO} [c, d]$ iff $b > c$ (Fishburn [71]). Note that, interpreting intervals as possibility distributions $[a, b]$ and $[c, d]$, respectively restricting ill-known quantities x and y, the statement $[a, b] >_{IO} [c, d]$ can be interpreted as by means of the necessity degree as $N(x > y) = 1$, given that $(x, y) \in [a, b] \times [c, d]$.
2. *Interval lattice extension of the usual ordering*: If we extend the maximum and minimum operations on the real line to intervals, it yields

   $$\max([a, b], [c, d]) = \{z = \max(x, y)x \in [a, b], y \in [c, d]\} = [\max(a, c), \max(b, d)]$$

   and likewise $\min([a, b], [c, d]) = [\min(a, c), \min(b, d)]$. The set of closed intervals equipped with min and max forms a lattice, and the canonical ordering in this lattice is

   $$[a, b] \geq_{lat} [c, d] \iff \max([a, b], [c, d]) = [a, b]$$
   $$\iff \min([a, b], [c, d]) = [c, d] \iff a \geq c \text{ and } b \geq d.$$

3. *Subjective approach*: this is Hurwicz criterion that uses a coefficient of optimism $\lambda \in [0, 1]$ for describing the attitude of the decision-maker. Intervals can be compared via a selection of precise substitutes to intervals:

   $$[a, b] \geq_\lambda [c, d] \iff \lambda a + (1 - \lambda)b \geq \lambda c + (1 - \lambda)d.$$

It is clear that $[a, b] >_{IO} [c, d]$ implies $[a, b] \geq_{lat} [c, d]$, which is equivalent to $\forall \lambda \in [0, 1], [a, b] \geq_\lambda [c, d]$.

There are also well-known methods for comparing random variables

1. *1st Order Stochastic Dominance*: $x \geq_{SD} y$ iff $\forall \theta, P(x \geq \theta) \geq P(y \geq \theta)$ ([26] Chap. 8)
2. *Probabilistic preference relations*: $R(x, y) = P(x \geq y)$ and exploit them (e.g. $x > y$ iff $R(x, y) > \alpha > 0.5$) [37].
3. *Scalar substitutes* : Comparing x and y by their expectations, more generally the expectation of their utilities $u(x)$ and $u(y)$.

For independent random variables, $P(x \geq y) = 1$ implies $x \geq_{SD} y$, which is equivalent to $\int u(t)d_xP(t) \geq \int u(t)dP_t(x), \forall$ monotonic increasing utility functions u. For monotonically related random variables with joint distribution function $\min(F_x(r), F_y(r'))$ it is clear that $P(x \geq y) = 1$ corresponds to 1st order stochastic dominance.

5.2. Constructing fuzzy interval ranking methods

According to the chosen interpretation of fuzzy intervals, the above methods for comparing intervals and probabilities can be extended, possibly conjointly and thus define well-founded ranking techniques for fuzzy intervals. So doing, a number of existing ranking methods can be retrieved, and make sense in a particular setting.

5.2.1. Metric approach

If fuzzy intervals are viewed as mere possibility distributions $\pi_x = \mu_{\tilde{a}}$ and $\pi_y = \mu_{\tilde{b}}$, it is natural to exploit counterparts to probabilistic ranking techniques, turning probability measures into possibility and necessity measures. One gets methods that are well-known:

1. *Interval lattice extension of stochastic dominance*: Clearly there are two cumulative distributions one can derive from a fuzzy interval:
   - the upper distribution $F^*(\theta) = \Pi(x \leq \theta) = \mu_{\tilde{a}}(\theta)$ if $\theta \leq \underline{a}$ and 1 otherwise;
   - the lower distribution $F_*(\theta) = N(x \leq \theta) = 1 - \mu_{\tilde{a}}(\theta)$ if $\theta \geq \overline{a}$ and 0 otherwise;

   Then, we can

   - either combine stochastic dominance and interval ordering:

     $$\tilde{a} \geq_{IO} \tilde{b} \iff \forall \theta, N(x \geq \theta) \geq \Pi(y \geq \theta),$$

     which is a very demanding criterion that basically requires that $\overline{b} \leq \underline{a}$ and $1 - \mu_{\tilde{a}}(\theta) \geq \mu_{\tilde{b}}(\theta), \forall \theta \in [\overline{b}, \underline{a}]$
   - or combine stochastic dominance and the lattice interval ordering:

     $$\tilde{a} \geq_{lat} \tilde{b} \iff \forall \theta, \Pi(x \geq \theta) \geq \Pi(y \geq \theta) \text{ and } N(x \geq \theta) \geq N(y \geq \theta).$$

     It comes down to comparing cuts of $\tilde{a}$ and $\tilde{b}$ using the lattice interval ordering or yet the well-known comparison method via the extended minimum or maximum $\tilde{a} \geq_{\tilde{c}} \tilde{b}$ iff $\widetilde{\max}(\tilde{a}, \tilde{b}) = \tilde{a}$ (or $\widetilde{\min}(\tilde{a}, \tilde{b}) = \tilde{b}$)
2. *Counterparts of expected utility* : compute the possibility and the necessity of reaching a fuzzy goal G using possibility and necessity of fuzzy events. In this case, the membership function μ_G represents a preference profile that stands for a utility function, and special cases of Sugeno integrals can be computed:
   - The degree of possibility of reaching the goal:

     $$\Pi_{\tilde{a}}(G) = \sup_{\theta} \min(\mu_{\tilde{a}}(\theta), \mu_G(\theta)).$$

   - The degree of necessity of reaching the goal:

     $$N_{\tilde{a}}(G) = \inf_{\theta} \max(1 - \mu_{\tilde{a}}(\theta), \mu_G(\theta)).$$

   These criteria are possibilistic counterparts of expected utility functions (optimistic and pessimistic, respectively). They have been axiomatized as such by Dubois et al. [44]. When μ_G is an increasing function, one can compare fuzzy intervals $\mu_{\tilde{a}}$ and $\mu_{\tilde{b}}$ by comparing pairs $(N_{\tilde{a}}(G), \Pi_{\tilde{a}}(G))$ and $(N_{\tilde{b}}(G), \Pi_{\tilde{b}}(G))$ using interval ordering techniques. This approach systematizes the one of Chen [79].

3. *Possibilistic valued relations*: compute valued preference relations obtained as the degrees of possibility or necessity that x, restricted by $\tilde{a}$, is greater than y restricted by $\tilde{b}$. For instance the index of certainty of dominance $R(x,y) = N(x \geq y) = 1 - \sup_{v>u} \min(\pi_x(u), \pi_y(v))$. This line of thought goes back the seventies [59] and was systematized by Dubois and Prade [80]. It extends interval orderings [39] since $N(x \geq y) = 1 - \inf\{\alpha : \tilde{a}_\alpha >_{IO} \tilde{b}_\alpha\}$.

5.2.2. Random interval approach

One may wish to probabilize interval ranking methods, interpreting a fuzzy interval as a nested random interval. For instance, one may use the valued relation approach to comparing random numbers, extended to intervals:

1. The random interval order yields a valued relation of the form : $R_{IO}(\tilde{a}, \tilde{b}) = Prob(\tilde{a}_\alpha \geq_{IO} \tilde{b}_\beta)$; this kind of approach has been especially proposed by Chanas and colleagues [81, 82]. The randomized form of the canonical lattice interval extension of the usual order of reals $>$ reads: $R_C(\tilde{a}, \tilde{b}) = Prob(\tilde{a}_\alpha \geq_{lat} \tilde{b}_\beta)$; both expressions presuppose some assumption be made regarding the dependence structure between the parameters α and β viewed as random variables on the unit interval.
2. The probabilistic version of the subjective approach leads to the following valued relation that depends on the coefficient of optimism: $R_\lambda(\tilde{a}, \tilde{b}) = Prob(\lambda \underline{a}_\alpha + (1-\lambda)\overline{a}_\alpha \geq \lambda \underline{a}_\alpha + (1-\lambda)\overline{b}_\alpha)$

One may also generalize stochastic dominance to random intervals. To this end, we must notice that $Prob(\underline{a}_\alpha \leq \theta) = \Pi_{\tilde{a}}(x \leq \theta)$ and $Prob(\overline{a}_\alpha \leq \theta) = N_{\tilde{a}}(x \leq \theta)$. Hence we get the same approach as in the ordinal case when we compare any among $Prob(\underline{a}_\alpha \leq \theta)$ or $Prob(\overline{a}_\alpha \leq \theta)$ to any of $Prob(\underline{b}_\alpha \leq \theta)$ or $Prob(\overline{b}_\alpha \leq \theta)$. One gets a special case of stochastic dominance between belief functions studied by Denoeux [83].

Finally one may also compute the average interval using the Aumann integral: $E(\tilde{a}) = \int_0^1 \tilde{a}_\alpha$ [87], and compare $E(\tilde{a})$ and $E(\tilde{b})$ using interval comparison methods. For instance, the Hurwicz method then coincides with the subjective approach of Campos and Gonzales [84] and subsumes Yager's [85] and Fortemps and Roubens [86] techniques.

5.2.3. Imprecise probability approach

Viewing fuzzy intervals as families of probability measures yields techniques close to the random set approach

- The extension of 1st Order Stochastic Dominance to fuzzy intervals remains the same since $\Pi_{\tilde{a}}(x \leq \theta)$ (resp. $N_{\tilde{a}}(x \leq \theta)$ is also the upper (resp. lower) probability of the event "$x \leq \theta$" in the sense of the probability family $\mathcal{P}_\alpha$.
- Comparing upper and lower expected values of x and y, namely $E^*(\tilde{a}) = \int_0^1 \overline{a}_\alpha d\alpha$ and $E_*(\tilde{a}) = \int_0^1 \underline{a}_\alpha d\alpha$ comes down to comparing mean intervals since $E(\tilde{a}) = [E_*(\tilde{a}), E^*(\tilde{a})]$ [73].

- One may also construct interval- valued preference relations obtained as upper and lower probabilities of dominance: $\begin{cases} R^*(x,y) = P^*(x \geq y), \\ R_*(x,y) = P_*(x \geq y). \end{cases}$ and exploit them. A different interval-valued quantity that is relevant in this context is $[E_*(\tilde{a} - \tilde{b}), E^*(\tilde{a} - \tilde{b})]$ to be compared to 0. In imprecise probability theory, comparing lower expectations of x and y is not equivalent to comparing the lower expectation to $x - y$ to 0, generally.

5.2.4. Gradual number approach

Viewing fuzzy intervals as intervals of gradual numbers, we first need a method for comparing gradual numbers: again three natural techniques come to mind. They extend the comparison of random variables to some extent, because the inverse of a cumulative distribution function is a special case of gradual number:

1. *Levelwise comparison*: $\breve{r} \geq \breve{s}$ iff $\forall \alpha, r_\alpha \geq s_\alpha$. It is clear that this definition reduces to 1st order stochastic dominance when the gradual number is the inverse of a distribution function.
2. *Area comparison method*:
$$\breve{r} >^S \breve{s} \Longleftrightarrow \int_0^1 \max(0, r^\alpha - s^\alpha) d\alpha > \int_0^1 \max(0, s^\alpha - r^\alpha) d\alpha.$$
3. *Comparing defuzzified values*: the natural way of defuzzifying a gradual number is to compute the number $m(\breve{r}) = \int_0^1 r^\alpha d\alpha$. This expression reduces to standard expectation using inverses of distribution functions. And clearly $\breve{r} >^S \breve{s} \Longleftrightarrow m(\breve{r}) > m(\breve{s})$.

The connection between gradual numbers and the dominance index $P(x > y)$ is worth exploring. In fact, under suitable dependence assumptions between, $P(x > y)$ is related to the Lebesgue measure $\ell(\{\alpha, r_\alpha \geq s_\alpha\})$.

On this ground one can compare fuzzy intervals $\tilde{a}$ and $\tilde{b}$, viewed as genuine intervals of functions $[\underline{\breve{a}}, \overline{\breve{a}}]$ and $[\underline{\breve{b}}, \overline{\breve{b}}]$ limited by gradual numbers, where $\underline{\breve{a}}_\alpha = \underline{a}_\alpha$ and $\overline{\breve{a}}_\alpha = \overline{a}_\alpha$, so that $\tilde{a} = \{\breve{r} : \underline{\breve{a}} \leq \breve{r} \leq \overline{\breve{a}}.\}$ One retrieves some ranking methods already found by the above previous approaches:

- Lattice interval extension of $>$: $\tilde{a} \geq_{lat} \tilde{b}$ iff $\underline{\tilde{a}} \geq \underline{\tilde{b}}$ and $\overline{\tilde{a}} \geq \overline{\tilde{b}}$.
- Stochastic dominance with subjective approach : $\tilde{a} \geq_\lambda \tilde{b}$ iff $\lambda \underline{\breve{a}} + (1-\lambda)\overline{\breve{a}} \geq \lambda \underline{\breve{b}} + (1-\lambda)\overline{\breve{b}}$.
- Subjective approach by comparing expectations: $\tilde{a} \geq_\lambda \tilde{b}$ iff $\int_0^1 (\lambda \underline{a}_\alpha + (1-\lambda)\overline{a}_\alpha) d\alpha \geq \int_0^1 (\lambda \underline{b}_\alpha + (1-\lambda)\overline{b}_\alpha) d\alpha$.

Note that $\tilde{a} \geq_{IO} \tilde{b}$ reads $\underline{\breve{a}} \geq \overline{\breve{b}}$, which is equivalent to the comparison of the interval supports of the corresponding fuzzy numbers.

This typology only aims at emphasizing the impact of attaching a interpretation to fuzzy numbers on the search for ranking methods. It can serve as a tool for constructing well-founded ranking methods for fuzzy intervals and to study their properties.

6. Conclusion

The use of fuzzy sets in decision analysis remains somewhat debatable so long as proposals for using fuzzy intervals, fuzzy preference relations, linguistic value scales are not better positioned in the stream of current research in measurement theory [25] and decision sciences [26, 55]. There does not seem to exist a niche for an isolated theory of fuzzy decision-making. However the use of fuzzy sets may focus the attention of scholars of traditional decision theory on some issues they were not otherwise considered, like going beyond averaging for criteria aggregation, the idea of gradual preference in relational approaches, a refined handling of incomplete information, and a well-founded basis for qualitative approaches to evaluation.

Several messages are the result of the above analysis of the literature:

- Fuzzy set theory offers a bridge between numerical approaches and qualitative approaches to decision analysis, but:
    1. The use of linguistic variables encoded by fuzzy intervals does not always make a numerical method more qualitative or meaningful.
    2. Replacing numerical values by fuzzy intervals rather corresponds to a kind of sensitivity analysis, not to a move toward the qualitative.
    3. The right question is: how to faithfully encode qualitative techniques on numerical scales, rather than using linguistic terms to extend already existing ad hoc numerical techniques.
- There is a strong need to develop original fuzzy set-based approaches to multicriteria decision analysis that are not a rehashing of existing techniques with ad hoc fuzzy interval computations.
- Fuzzy set theory and its mathematical environment (aggregation operations, graded preference modeling, and fuzzy interval analysis) provide a general framework to pose decision problems in a more open-minded way, towards a unification of existing techniques.

Open questions remain, such as:

- Refine any qualitative aggregation function using discri-schemes or lexi-schemes
- Computational methods for finding discrimin-leximin solutions to fuzzy optimization problems.
- Devise a behavioral axiomatization of new aggregation operations in the scope of MCDM, decision under uncertainty and fuzzy voting methods

- Develop a general axiomatic framework for ranking fuzzy intervals based on first principles.
- Study the impact of semantics of fuzzy preference relations (probabilistic, probabilistic, distance-based,..) on how they should be exploited for ranking purposes
- Provide a unified framework for fuzzy choice functions.

References

[1] R. E. Bellman and L. A. Zadeh L.A (1970) Decision making in a fuzzy environment, *Management Science*, 17, B141-B164.

[2] H. -J. Zimmermann (1978), Fuzzy programming and linear programming with several objective functions *Fuzzy Sets and Systems*,1: 45-55.

[3] L.A. Zadeh (1971), Similarity relations and fuzzy orderings, Information Sciences, 3:177-200.

[4] S. Orlovski (1978). Decision-making with a fuzzy preference relation, Fuzzy Sets and Systems, 1:155-168.

[5] J. Fodor and M. Roubens (1994) *Fuzzy preference modelling and multicriteria decision support. Kluwer Academic Publishers.*

[6] U. Bodenhofer, B. De Baets, J. C. Fodor (2007) A compendium of fuzzy weak orders: Representations and constructions. Fuzzy Sets and Systems 158(8): 811-829

[7] G. Beliakov, A. Pradera, T. Calvo (2008) *Aggregation Functions: A Guide for Practitioners* Studies in Fuzziness and Soft Computing, Vol. 221, Springer.

[8] V. Torra, Y. Narukawa (2007) *Modeling Decisions: Information Fusion and Aggregation Operators*, 978-3-540-68789-4, Springer.

[9] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap (2009) Aggregation Functions, Cambridge University Press.

[10] Y.-M. Wang, Y. Luo, Generalised fuzzy weighted mean and its applications International Journal of General Systems, 38(2009) 533 Ð 546

[11] X. Wang and E. Kerre (2001) Reasonable properties for the ordering of fuzzy quantities (2 parts), *Fuzzy Sets and Systems*, 118, 375-406.

[12] Zadeh L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning, Information Sciences, Part 1: 8, 199-249, Part 2: 8, 301-357, Part 3: 9, 43-80.

[13] F. Herrera, E. Herrera-Viedma, L. Martinez (2000) A fusion approach for managing multi-granularity linguistic term sets in decision making Fuzzy Sets and Systems,114: 43-58

[14] D. Dubois, H. Prade (1997). The three semantics of fuzzy sets, *Fuzzy Sets and Systems*, 90: 141-150.

[15] E. H. Ruspini: On the semantics of fuzzy logic. Int. J. Approx. Reasoning 5(1): 45-88 (1991)

[16] Zadeh L.A. (1978). Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems, 1, 3-28.

[17] D. Dubois, H. Prade, Possibility theory: qualitative and quantitative aspects. In: P. Smets, Ed., Handbook on Defeasible Reasoning and Uncertainty Management Systems Ñ Volume 1: Quantified Representation of Uncertainty and Imprecision. Kluwer Academic Publ., Dordrecht, The Netherlands, 169-226, 1998.

[18] D. Dubois: Possibility theory and statistical reasoning. Computational Statistics & Data Analysis 51(1): 47-69 (2006)

[19] Zadeh L.A. (1975). Calculus of fuzzy restrictions, Fuzzy Sets and their Applications to Cognitive and Decision Processes (Zadeh L.A., Fu K.S., Tanaka K. and Shimura M., eds., Academic Press, New York, 1-39.

[20] D. Dubois, H. Fargier, H. Prade (1996) Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty. *Applied Intelligence*, 6: 287-309.

[21] D. Dubois, H. Prade: An introduction to bipolar representations of information and preference. Int. J. Intell. Syst. 23(8): 866-877 (2008)

[22] M. Grabisch, S. Greco, M. Pirlot: Bipolar and bivariate models in multicriteria decision analysis: Descriptive and constructive approaches. Int. J. Intell. Syst. 23(9): 930-969 (2008)

[23] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty, 5:297Ð323, 1992.

[24] D. Dubois, H. Prade (2004) On the use of aggregation operations in information fusion processes. Fuzzy Sets and Systems, 142,143-161

[25] Kranz D. H,. Luce R. D., Suppes P., Tverski A. (1971) Foundations of Measurement, Vol. 1 Additive and Polynomial representations, Academic Press, New-York, 1971

[26] D. Bouyssou, D. Dubois, M. Pirlot, H. Prade (2009), *Decision-Making Process*, ISTE, London, UK & Wiley, Hoboken, N.J. USA

[27] K. Atanassov, Intuitionistic fuzzy sets, Fuzzy Sets & Syst., 20 (1986), 87-96.

[28] D. Dubois, H. Prade (2010) Gradualness, Uncertainty and Bipolarity: Making Sense of Fuzzy Sets, Fuzzy Sets and Systems,, to appear.

[29] G. Deschrijver, E. Kerre, On the relationship between some extensions of fuzzy set theory, Fuzzy Sets and Systems, 133 (2004), 227-235

[30] Zadeh L.A.Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic Fuzzy Sets and Systems, 90(1997): 111-127

[31] D. Dubois, H. Prade Fuzzy arithmetic in qualitative reasoning, In: Modeling and Control of Systems in Engineering, Quantum Mechanisms, Economics and Biosciences (A. Blaquière, ed.), Lecture Notes in Control and Information Sciences, Springer Verlag, Berlin, 1989, 457-467.

[32] F. Herrera, L. Martinez (2001) A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making. *IEEE Transactions on Systems, Man, and Cybernetics*, Part B 31(2): 227-234.

[33] D. Bouyssou, T. Marchant, M. Pirlot, P. Perny, A.Tsoukias, P. Vincke (2000) *Evaluation Models: a Critical Perspective*. Kluwer Acad. Pub. Boston.

[34] J. C. Fodor, B. De Baets: Fuzzy Preference Modelling: Fundamentals and Recent Advances. Fuzzy Sets and Their Extensions: Representation, Aggregation and Models (H. Bustince, F. Herrera, J. Montero , Eds.) Studies in Fuzziness and Soft Computing Vol. 220 Springer 2008,: 207-217.

[35] J. Fodor (2009) Some Remarks on Difference Measurement, Proc. 7th Int. Symposium on Applied Intelligence and Informatics, Budapest, 11-14

[36] U. Bodenhofer, Representations and constructions of similarity-based fuzzy orderings Fuzzy Sets and Systems, 137(2003): 113-136

[37] B. De Baets, H. De Meyer, Transitivity frameworks for reciprocal relations: cycle-transitivity versus FG-transitivity, Fuzzy Sets and Systems,152 (2005) 249-270.

[38] J.M. Blin A. B. Whinston Fuzzy sets and social choice, J. Cybernetics, 3(4) (1973) 17-22.

[39] M. Roubens, P.Vincke P. (1988). Fuzzy possibility graphs and their application to ranking fuzzy numbers, Non-Conventional Preference Relations in Decision Making (Kacprzyk J. and. Roubens M., eds.), Springer-Verlag, 119-128.

[40] Gigerenzer, G., Todd, P. M., & the ABC group. (1999). Simple heuristics that make us smart. New York: Oxford University Press.

[41] Boutilier, C., Brafman, R. I., Domshlak, C., Hoos, H. H., Poole, D. (2004). CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. J. Artif. Intell. Res. (JAIR), 21, 135-191.

[42] D. Dubois, P. Fortemps . Computing improved optimal solutions to max-min flexible constraint satisfaction problems. European Journal of Operational Research, V. 118, p. 95–126, 1999

[43] A. Rico, M. Grabisch, C. Labreuche, A. Chateauneuf (2005) Preference modeling on totally ordered sets by the Sugeno integral. *Discrete Applied Mathematics* 147(1): 113-124.

[44] D. Dubois, H. Prade. R. Sabbadin (2001) Decision theoretic foundations of qualitative possibility theory. *European Journal of Operational Research*, 128: 459-478.

[45] H. Fargier and R. Sabbadin (2005) Qualitative decision under uncertainty: Back to expected utility. *Artificial Intelligence*, 164:245–280.

[46] D. Dubois and H. Fargier (2009) Making Sugeno integrals more discriminant. *Int. J. Approximate Reasoning*, 50: 880-898.

[47] D. Dubois, H. Fargier, and H. Prade. Possibilistic likelihood relations. In Proceedings of IPMUÕ98, pages 1196Ð1203. EDK, 1998.

[48] M. Grabisch. The Moebius transform on symmetric ordered structures and its application to capacities on finite sets. Discrete Mathematics, 287:17Ð34, 2004.

[49] D. Dubois, H. Fargier,. Capacity refinements and their application to qualitative decision evaluation European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2009), Verona (Italy), Claudio Sossai , Gaetano Chemello (Eds.), Springer, LNAI 5590, p. 311-322, 2009.

[50] D. Kyselovà, D. Dubois, M. Komorníkovà, R. Mesiar. Refining aggregation operator-based orderings in multifactorial evaluation - Part I: Continous scales. IEEE Transactions on Fuzzy Systems, IEEE, Special issue on Aggregation operators, Vol. 15 N. 6, p. 1100-1106, 2007.

[51] G. Shafer. A Mathematical Theory of Evidence. Princeton University Press, Princeton, 1976.

[52] J.-F. Bonnefon, D. Dubois, H. Fargier, S. Leblois (2008) Qualitative heuristics for balancing the pros and the cons, *Theory and Decision*, 65:71Ð95.

[53] D. Dubois, H. Fargier, J.-F. Bonnefon (2008) On the qualitative comparison of decisions having positive and negative features. *Journal of Artificial Intelligence Research*, 32: 385-417.

[54] Lootsma F. (1997), Fuzzy Logic for Planning and Decision-Making, Kluwer Acad. Publ., Dordrecht, 1997

[55] J. Figueira, S. Greco, M. Ehrgott, M. (Eds.) Multiple Criteria Decision Analysis:State of the Art, International Series in Operations Research & Management Science, Vol. 78, Springer Verlag, Boston, 2005.

[56] D. Dubois, E. Kerre, R. Mesiar, H. Prade. Fuzzy interval analysis. In: Fundamentals of Fuzzy Sets, Dubois,D. Prade,H., Eds: Kluwer, Boston, Mass, The Handbooks of Fuzzy Sets Series, 483-581, 2000.

[57] J. Fortin D. Dubois H. Fargier, Gradual Numbers and Their Application to Fuzzy Interval Analysis, IEEE trans. Fuzzy Systems, 16, 388-402, 2008

[58] Moore R. (1979). Methods and Applications of Interval Analysis, SIAM Studies on Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia.

[59] S. M. Baas and H. Kwakernaak (1977). Rating and ranking of multiple aspect alternatives using fuzzy sets, Automatica, 13, 47-58.

[60] Dubois D. and Prade H. (1981). Additions of interactive fuzzy numbers, IEEE Trans. Automatic Control, 26, 926-936.

[61] L. M. de Campos, J.F. Huete, and S. Moral. Probability intervals : a tool for uncertain reasoning. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2:167- 196, 1994.

[62] Ying-Ming Wang, Taha M.S. Elhag On the normalization of interval and fuzzy weights Fuzzy Sets and Systems, 157(2006) 2456-2471

[63] P. Meyer, M. Roubens On the use of the Choquet integral with fuzzy numbers in multiple criteria decision support Fuzzy Sets and Systems, Volume 157, (2006) 927-938.

[64] Saaty T. L. (1980). The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation, McGraw-Hill, New York.

[65] Van Laaroven P. J. M. and Pedrycz W. (1983). A fuzzy extension of Saaty's priority theory, Fuzzy Sets and Systems, 11, 229-241

[66] R. Csutora, J. Buckley, Fuzzy hierarchical analysis: the Lambda-Max method Fuzzy Sets and Systems, 120 (2001)181-195.

[67] J. Ramík, P. Korviny, Inconsistency of pair-wise comparison matrix with fuzzy elements based on geometric mean Fuzzy Sets and Systems, 161(2010):1604-1613.

[68] S. Ohnishi, D. Dubois, H. Prade T. Yamanoi (2008) A Fuzzy Constraint-Based Approach to the Analytic Hierarchy Process. In: *Uncertainty and Intelligent Information Systems* (B. Bouchon-Meunier et al. Eds), World Scientific, Singapore, 217-228.

[69] Bortolan G. and Degani R. (1985). A review of some methods for ranking fuzzy subsets, Fuzzy Sets and Systems, 15, 1-19.

[70] Chang P. T. and Lee E. S. (1994). Fuzzy arithmetics and comparison of fuzzy numbers, Fuzzy Optimization: Recent Advances (Delgado M., Kacprzyk J., Verdegay J. L., Vila M. A., eds), Physica-Verlag, Heidelberg, Germany, 69-81.

[71] P. Fishburn, Interval Orderings, Wiley, New-York (1987)

[72] Goodman I.R. (1982). Fuzzy sets as equivalence classes of random sets, Fuzzy Sets and Possibility Theory (Yager R.R., Ed.), Pergamon Press, Oxford, 327-342.

[73] Dubois D. and Prade H. (1987). The mean value of a fuzzy number, Fuzzy Sets and Systems, 24, 279-300.

[74] Chanas S. and Nowakowski M. (1988). Single value simulation of fuzzy variable, Fuzzy Sets and Systems, 25, 43-57

[75] P. Smets. Belief functions on real numbers. International Journal of Approximate Reasoning, 40:181-223, 2005.

[76] P. Walley. Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, 1991.

[77] D. Dubois and H. Prade. When upper probabilities are possibility measures. Fuzzy Sets and Systems, 49:65-74, 1992.

[78] G. De Cooman and D. Aeyels. Supremum-preserving upper probabilities. Information Sciences, 118:173-212, 1999.

[79] Chen S. H. (1985). Ranking fuzzy numbers with maximizing sets and minimizing sets, Fuzzy Sets and Systems, 17, 113-129.

[80] Dubois D. and Prade H. (1983b). Ranking fuzzy numbers in the setting of possibility theory, Information Sciences, 30, 183-224.

[81] Chanas, S., Delgado, M., Verdegay, J.L., Vila, M. A.: Ranking fuzzy real intervals in the setting of random sets. Information Sciences, 69, 201-217 (1993)

[82] Chanas, S., Zielinski, P.: Ranking fuzzy real intervals in the setting of random sets: further results. Information Sciences, 117, 191-200 (1999)

[83] Denoeux,T. : Extending stochastic order to belief functions on the real line. Information Sciences, 179, 1362-1376 (2009)

[84] L. de Campos, A. Gonzalez, A subjective approach for ranking fuzzy numbers. Fuzzy sets and Systems 29 :145-153, 1989.

[85] R.R. Yager (1981). A procedure for ordering fuzzy subsets of the unit interval, Information Sciences, 24, 143-161.

[86] P. Fortemps, P. M. Roubens (1996). Ranking and defuzzification methods based on area compensation, Fuzzy Sets and Systems, 82, 319-330.

[87] Y. Ogura, S.M Li, D. A. Ralescu : Set Defuzzification and Choquet Integral. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems 9 : 1-12, 2001.