Fuzzy Sets in Decision Sciences: assets and limitations.

Didier Dubois
IRIT - Université Paul Sabatier -
31062 Toulouse Cedex (France)
e-mail: dubois@irit.fr

Introduction

- The idea of applying fuzzy sets in decision sciences comes from the seminal paper of Bellman and Zadeh ”Decision-Making in a Fuzzy Environment” in Management Sciences, 1970.
- That pioneering paper makes two main points:
  1. Objective functions of criteria and constraints can be viewed as Membership functions that can be aggregated by fuzzy set conjunctions, especially the minimum
  2. Multiple-stage decision-making problems can then been stated and solved by means of dynamic programming.
- So according to this view fuzzy optimization comes down to max-min bottleneck optimization
- Taken over by Zimmermann who developed multicriteria linear optimisation in the seventies.

Two optimisation paradigms

- **Egalitarist :** Maximizing $\min_i \mu_i(x)$
  - a feasible solution is one that satisfies all constraints $\mu_i(x) \in \{0, 1\}$
  - Making the crisp environment fuzzy $\mu_i(x) \in [0, 1]$
  - μ_i defines a soft constraint.
  - Belmann and Zadeh assume constraints are soft (see literature on constraint satisfaction).
- **Utilitarist:** Maximizing $\sum_i \mu_i(x)$
  - Making fuzzy environment crisp $\mu_i(x) \in \{0, 1\}$:
  - μ_i defines an ideal goal, an objective, a wish
  - an optimal solution is one that satisfies as many objectives as possible.
- The choice of one or the other depends on the application;
- Hybrid cases: optimizing objectives under soft constraints.

Beyond Bellman and Zadeh

- Other ingredients relevant to fuzzy decision-making have been proposed
  1. Fuzzy preference relations (orderings, Zadeh, 1971) studied further by Orlowski, Roubens, Fodor, De Baets, Susanna Diaz and others
  2. Aggregation operations so as to refine the multicriteria technique of Zadeh : t-norms symmetric sums, uninorms, leximin, Sugeno and Choquet integrals etc
  3. Fuzzy interval computations so as to cope with uncertainty in numerical aggregation schemes.
  4. Fuzzy interval comparison techniques for selecting the best option in a set of alternatives with fuzzy interval ratings
  5. Linguistic variables, so as to get decision methods closer to the user cognition.

This talk takes a skeptical viewpoint on the fuzzy decision literature, so as to help laying bare what is its actual contribution.

What is a fuzzy environment ?

It is not clear that fuzzy sets have led to a new decision paradigm.

- Bottleneck optimisation and maximin decisions existed independently of fuzzy sets.
- In many cases, fuzzy sets have been added to existing techniques (fuzzy AHP methods, fuzzy weighted averages, fuzzy extensions of Electre-style MCDM methods)
- Fuzzy preference modelling is an extension of standard preference modelling and must be compared to probabilistic preference modeling.
- Traditional decision settings can accommodate possibility theory instead of probability theory (Savage act-based approach, for instance)

One cannot oppose ”fuzzy environment” to ”uncertain environment”: the former means ”using fuzzy sets”, while the latter refers to an actual decision situation.

What is the contribution of fuzzy sets to decision analysis?

In fact, it is not always the case that adding fuzzy sets to an existing method improves it in a significant way. 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: using qualitative scales instead of numerical ones, CP-nets, possibilistic logic...
- possesses a convincing rationale and a sound formal setting liable to some axiomatization (why such an aggregation method?).

Unfortunately, many submitted or even published papers seem to contain no such contribution.

what next?

Examine the state of the art on some specific topics:

- Scales used for membership grades
- Linguistic variables and fuzzy intervals in decision analysis
- Ranking fuzzy numbers and preference relations
- Aggregation operations : going qualitative.

Membership functions

- A membership function is an abstract notion, a mathematical tool. As such adding membership functions to a decision problem does not enrich its significance.
- One must always declare what a given membership function stands for:
  - A measure of *similarity* to prototypes of a linguistic concept (this is related to distance)
  - A *possibility distribution* representing our incomplete knowledge of a parameter, state of nature, etc. that we cannot control
  - A numerical encoding of a *preference* relation over feasible options, similar to a *utility or an objective function* ?
- 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
  - or in its set-theoretic view of numerical functions.

Scales

The totally ordered set of truth-values L is also an abstract construct. Assumptions must be laid bare if it is used as a value scale

- What is the meaning of the end-points ? has the mid point any meaning ?
  - (0 = BAD, 1 = NEUTRAL): *negative unipolar scale*. For instance, a possibility distribution, a measure of loss.
  - (0 = NEUTRAL, 1 = GOOD): *positive unipolar*. For instance degrees of necessity, a measure of gain.
  - (0 = BAD, 1/2 = NEUTRAL, 1 = GOOD): *bipolar scale*. For instance a degree of probability.

  This information captures what is good or bad in the absolute. A simple preference relation cannot.

  The choice of landmark points may have strong impact on the proper choice of aggregation operations (t-norms, co-norms, uninorms...)

Scales

- What is the expressive power of a scale L ?
  - Ordinal scales : only the ordering on L matters.
  - Interval scales : a numerical scale defined up to a positive affine transformation ($a\lambda + b, a > 0$).
  - Ratio scales : a numerical scale defined up to a positive linear transformation $a\lambda$
  - (Finite) qualitative scales: $L = \{\lambda_1 < \lambda_2 < \cdots < \lambda_n\}$
- Also influences the choice of aggregation operations :
  - Ordinal invariance : $a * b > c * d$ iff $\varphi(a) * \varphi(b) > \varphi(c) * \varphi(d)$ for all monotonic increasing transformation φ of an ordinal scale L (addition forbidden).
  - Asking people to tick a value on a line segment does not yield a genuine number (ordinal scale)
  - Interval scales cannot express the idea of good and bad (neither bipolar nor even unipolar)
  - Small qualitative scales are cognitively easier to grasp and can be consensual.

Evaluations by pairs of values

- Some authors like to extend number-crunching evaluation techniques using pairs of values $(\mu, \nu) \in [0, 1]^2$ with $\mu + \nu \leq 1$, following Atanassov. This representation technique is ambiguous:
  1. Uncertainty semantics: Either it expresses less information than point values: an ill-known value $\lambda \in [\mu, 1 - \nu]$. Then it is better to use an uncertainty interval.
  2. Argumentation semantics : Or it expresses more information than point values : μ is the strength in favour of a decision, ν in disfavour of this decision.
- The standard injection $[0, 1] \to [0, 1]^2$ is not the same : $\lambda \mapsto (\lambda, 1 - \lambda)$ is the first case, $\lambda \mapsto (\lambda, 0)$ for a positive unipolar scale in the second case.
- It also affects the way information will be combined:
  1. In the first case, you need to apply interval analysis methods to see the impact of uncertainty on the global evaluation.
  2. In the second case, you may separately aggregate positive and negative information by appropriate (possibly distinct) methods.

Linguistic vs. Numerical Scales

- A qualitative scale can represent an ordered set of linguistic value labels understood as fuzzy intervals of $[0, 1]$.
- There is a natural temptation to model this information by means of a fuzzy partition on the unit interval. However:
  1. If an aggregation operation is not meaningful on the underlying numerical scale, the use of a linguistic variable does not make the extended operation more meaningful.
  2. It makes little sense to build a fuzzy partition on an abstract numerical interval which is at best an ordinal scale. It is OK if the underlying scale is a concrete measurable attribute.
  3. Some fuzzy linguistic approaches are as quantitative as a standard number-crunching method (e.g. the 2-tuple method that handles pairs (i, σ) where i denotes a label and $\sigma \in [-0.5, 0.5)$ and it encodes the number $i + \sigma \in [0, n]$.)

Linguistic vs. Numerical Scales

- Building qualitative operations on a linguistic scale from fuzzy arithmetics + linguistic approximation leads to a loss of algebraic properties (e.g. associativity) of the resulting operation on the corresponding qualitative scale.
- Using discrete t-norms other than min on finite scales comes down to using Lukasiewicz discrete t-norm, that is a truncated sum. It underlies assumptions on the meaning of qualitative scale $L = \{\lambda_0, \lambda_1, \ldots, \lambda_n\}$,
  1. L is mapped to the integers: $\lambda_0 = 0$, $\lambda_1 = 1 \ldots \lambda_n$.
  2. λ_i is i times as strong as λ_1
  3. There is a saturation effect that leads to ties when aggregating objective functions in this setting.
- So it is not really qualitative....

Fuzzy Intervals in Decision Analysis

- Fuzzy intervals have been widely used in fuzzy decision analysis 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
  - Fuzzy weighted averages using fuzzy interval weights
  - Fuzzy extensions of Saaty's Analytical hierarchical technique.
  - Fuzzy extensions of numerical or relational MCDM techniques (like TOPSIS, PROMETHEE, ELECTRE,...)
- Many of these techniques are ad hoc, partially erroneous, and lack originality.
- Unjustified use of defuzzification (if the result is precise why run a sensitivity analysis??)

Fuzzy weighted averages

- The computation of fuzzy weighted averages cannot be done by means of fuzzy arithmetics: it needs to solve (simple) linear programs and apply interval analysis.
- The problem is already present with imprecise (interval weights)
  - Find $\sup\{\frac{\sum_{i=1}^n x_i \cdot w_i}{\sum_{i=1}^n w_i} : w_i \in [a_i, b_i], x_i \in [c_i, d_i]\}$ (also $\inf$)
  - Find $\sup\{\sum_{i=1}^n x_i \cdot p_i : p_i \in [\frac{a_i}{a_i + \sum_{j \neq i} b_j}, \frac{b_i}{b_i + \sum_{j \neq i} a_j}], x_i \in [c_i, d_i], \sum_{i=1}^n p_i = 1\}$ (also $\inf$)
  - The two expressions differ. Fast methods exist.
  - Extensions to Choquet integrals with fuzzy-valued set-functions more difficult as the issue of ranking intervals $[c_, d_i]$ must be addressed.

Fuzzy weight vectors

One issue is: what is a normalized vector of interval weights $([a_1, b_1], [a_2, b_2] \ldots, [a_n, b_n])$?

- It is a (fuzzy) set of possible normalized weight vectors, not a fuzzy or interval substitute to a normalised set of weights.
- Specific conditions must be satisfied if all bounds are to be reachable by weight vectors $(p_1, p_2 \ldots, p_n)$ such that $\sum_{i=1}^{n} p_i = 1$.
- Namely see the theory of probability intervals (De Campos and Moral)
  1. $\sum_{i=1}^{n} a_i \leq 1 \leq \sum_{i=1}^{n} b_i$
  2. $a_i + \sum_{j \neq i} b_j \geq 1$; $b_i + \sum_{j \neq i} a_j \leq 1$.
- Computing FWA has more to do with constraint propagation than arithmetics with fuzzy numbers.

Fuzzy AHP

- There are many papers on fuzzy extensions of this technique.
- Again, most proposals are ad hoc and erroneous.
- *Principle of AHP* : compute a normalized weight vector from pairwise comparison data
  1. Given a set of n items, provide for each pair (i, j) a value $v_{ij} \in \{2, \ldots 9\}$ if i is preferred to j, $v_{ij} = 1$ if there is indifference.
  2. Build the matrix A with coefficients $a_{ij} = v_{ij}$ if $v_{ij} \geq 1; a_{ij} = 1/a_{ij}$.
  3. Find the largest eigenvalue λ of A.
  4. If λ close enough to n, the derived weights form the eigenvector of A
- Even if widely used, this method has been criticised by MCDM scholars as being ill founded at the measurement level (Bouyssou et al., 2000).

Fuzzy AHP

- The AHP method relies on the following ideal situation
  1. An $n \times n$ consistent preference matrix A is reciprocal in the sense that $a_{ij} = 1/a_{ji}$ and product-transitive ($a_{ij} = a_{ik} \cdot a_{kj}$).
  2. Then its eigen-value is n and there exists a normal weight vector $(p_1, p_2 \ldots, p_n)$with $a_{ij} = \frac{p_i}{p_j}$
  3. But in practice, pairwise comparison data do not provide consistent matrices and are arguably imprecise.
- So many researchers (since Van Laaroven and Pedrycz, 1983) have considered fuzzy valued pairwise comparison data, using fuzzy-valued matrices $\tilde{A}$ with fuzzy intervals $\tilde{a}_{ij} = \tilde{v}_{ij}$ if $\tilde{v}_{ij} \geq 1; \tilde{a}_{ij} = 1/\tilde{a}_{ij}$.

Fuzzy AHP

- However, it is hard to extend the computation scheme of Saaty with fuzzy intervals
  1. How to properly write the reciprocal and transitivity conditions ?(e.g. $\tilde{a}_{ij} = 1/\tilde{a}_{ji}$ is not the same as $\tilde{a}_{ij} \cdot \tilde{a}_{ji} = 1$).
  2. The transitivity property cannot hold in the form $\tilde{a}_{ij} = \tilde{a}_{ik} \cdot \tilde{a}_{kj}$ for fuzzy intervals.
  3. Fuzzy eigen-values or vectors of fuzzy-valued matrices are hard to define in a rigorous way (the usual definitions make no sense and are overconstrained)
- The bottom-line issue is that the natural extension of a crisp equation $ax = b$ is not a fuzzy equation of the form $\tilde{a}\tilde{x} = \tilde{b}$.
  - The first equation refers to a constraint to be satisfied in reality
  - Fuzzy intervals $\tilde{a}, \tilde{x}, \tilde{b}$ represent knowledge about a, x, b
  - Even if $ax = b$ it is not clear why the knowledge about ax should be equated to the knowledge about b : only consistency is requested ($\tilde{a}\tilde{x} \cap \tilde{b} \neq \emptyset$).

Fuzzy AHP : A constraint-based view

- *Fuzzy pairwise preference data should be used as flexible constraints on non-fuzzy coefficients, that define a fuzzy set of fully consistent preference matrices.*
  - The fuzzy matrix $\tilde{A}$ has entries $\tilde{a}_{ij} = \tilde{v}_{ij}$ or $1/\tilde{v}_{ji}$ and $\tilde{a}_{ii} = 1$.
  - A given normal weight vector $\vec{p} = (p_1, p_2 \ldots, p_n)$ satisfies the fuzzy preference matrix $\tilde{A}$ to degree

$$\mu(\vec{p}) = \min_{i,j} \mu_{ij}(\frac{p_i}{p_j}) = \min_{i \text{ preferred to } j} \mu_{\tilde{v}_{ji}}(\frac{p_i}{p_j})$$

    where μ_{ij} is the membership function of $\tilde{a}_{ij}$.
  - The degree of consistency of the preference data is $Cons(\tilde{A}) = \sup \mu(\vec{p})$
  - The best induced weight vectors are the Pareto maximal elements among $\{\vec{p}, \mu(\vec{p}) = Cons(\tilde{A})\}$.

Ranking fuzzy intervals

- There is a huge literature on the issue of ranking fuzzy intervals.
- Again the use of ad hoc methods is very common.
- There is a lack of first principles for devising well-founded techniques. However
  1. Wang and Kerre 's classification (2001) based on axioms like
     - $A \sim A$,
     - If $A \cap B = \emptyset$ then $A > B$ or $B > A$.
     - $A > B$ implies $A + C > B + C$
  2. Dubois Kerre et al. (2000) classification into
     (a) scalar indices (based on defuzzification)
     (b) goal-based indices : computing the degree of attainment of a fuzzy goal by each fuzzy interval
     (c) relational indices : computing to what extent a fuzzy interval dominates another.

Ranking fuzzy intervals

- *Another approach is to exploit links between fuzzy intervals and other settings: possibility, probability theories, interval orders.*
- There are methods for **comparing intervals**
  1. *Interval orders* : $[a, b] >_{IO} [c, d]$ iff $a > d$ (Fishburn)
  2. *Interval extension of the usual ordering* : $[a, b] \geq_C [c, d]$ iff $a \geq c$ and $b \geq d$
  3. *Subjective* (pessimistic/ optimistic Hurwicz) : $[a, b] \geq_\lambda [c, d]$ iff $\lambda a + (1 - \lambda)b \geq \lambda c + (1 - \lambda)d$.
- There are methods for comparing **random variables**
  1. *1st Order Stochastic Dominance* : $x \geq y$ iff $\forall \theta, P(x \geq \theta) \geq P(y \geq \theta)$
  2. *Comparing expected utilites* of x and y (= comparing expectations).
  3. *Stochastic preference relations* : Define $R(x, y) = P(x \geq y)$ and exploit it (e.g. $x > y$ iff $R(x, y) > \alpha > 0.5$).

Preference relations

- There is an important stream of works that extend preference modelling (decomposition of an outranking relation into strict preference, indifference and incomparability) to the gradual situation.
- What the fuzzy relational setting presupposes:
  - It makes sense to compare $R(x, y)$ to $R(z, w)$ (e.g. x is more preferred to y in the same way as z is more preferred to w).
  - The degree $R(x, y)$ must capture a clear intuition : intensity of gradual preference ? probability of crisp preference, possibility of crisp preference.
- The modelled situation enforces some basic properties of the preference scale (bipolar or not)
  - Probability of preference: $R(x, y) + R(y, x) = 1$ expresses completeness, and no room for incomparability.
  - Possibility of preference : $\max(R(x, y), R(y, x)) = 1$ expresses completeness.

Ranking fuzzy intervals

- *The choice of a method can be dictated by the point of view on what is a fuzzy interval* :

  Fuzzy intervals can be viewed as

  1. Intervals bounded by gradual numbers
  2. Ordinal possibility distributions
  3. One point-coverage functions of nested random intervals
  4. Families of probability functions

- According to the point of view methods for comparing intervals and probabilities should be extended, possibly conjointly and thus define well-founded ranking techniques for fuzzy intervals.

Ranking fuzzy intervals as intervals of gradual numbers

- A gradual number $\tilde{r}$ 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 A with cuts $[a_\alpha, b_\alpha]$: $\alpha \mapsto r_\alpha = \frac{b_\alpha - a_\alpha}{2}$.
- Ranking gradual numbers : $\tilde{r} \geq \tilde{s}$ iff $\forall \alpha, r_\alpha \geq s_\alpha$.
- A fuzzy interval is an interval of gradual numbers (picking one elements in each cut) bounded by a_α and b_α.
- **Retrieving ranking methods**:
  1. *Interval extension* of $>$: $A \geq_C B$ iff $\inf A_\alpha \geq \inf B_\alpha$ and $\sup A_\alpha \geq \sup B_\alpha$.
  2. *Subjective approach* (A. Gonzalez etc.): $A \geq_\lambda B$ iff $\lambda \inf A_\alpha + (1-\lambda) \sup A_\alpha \geq \lambda \inf B_\alpha + (1-\lambda) \sup B_\alpha$.
  3. *Comparing expectations* (Fortemps and Roubens, ...) $A \geq_\lambda B$ iff $\int_0^1 (\lambda \inf A_\alpha + (1-\lambda) \sup A_\alpha) d\alpha \geq \int_0^1 (\lambda \inf B_\alpha + (1-\lambda) \sup B_\alpha) d\alpha$.

Ranking fuzzy intervals as ordinal possibility distributions

- **Idea** : Ill-known quantities x and y with ordinal possibility distributions $\pi_x = \mu_A$ and $\pi_y = \mu_B$

- *Probabilistic indices are turned into possibility indices.*

1. Interval canonical extension of stochastic dominance : $A \geq_C B$ iff $\forall \theta, \Pi(x \geq \theta) \geq \Pi(y \geq \theta)$ and $N(x \geq \theta) \geq N(y \geq \theta)$.
   It comes down to $A \geq_C B$ iff $\tilde{\max}(A, B) = A$ (or $\tilde{\min}(A, B) = B$)

2. Compute the possibility and the necessity of reaching a fuzzy goal G using possibility and necessity of fuzzy events.

3. Compute valued preference relations:
   $R(x, y) = N(x \geq y) = 1 - \sup_{v > u} \min(\pi_x(u), \pi_y(v))$.
   It extends interval orderings since $N(x \geq y) = 1 - \inf\{\alpha : A_\alpha >_{IO} B_\alpha\}$.

Ranking fuzzy intervals as numerical possibility distributions

- A fuzzy interval A is viewed as the nested random set $([0, 1], \text{Lebesgue measure } \ell) \to \mathbb{R} : \alpha \mapsto A_\alpha$. Then $\pi_x(u) = \ell(\{\alpha, u \in A_\alpha\})$ where $A_\alpha = [a_\alpha, b_\beta]$.
- More generally intervals limited by two random variables $\dot{x} \leq \ddot{x}$ with disjoint support. Then $\pi_x(u) = Prob(\dot{x} \leq u \leq \ddot{x})$, where $[\dot{x}, \ddot{x}] = [a_\alpha, b_\beta]$.
- You can probabilize interval ranking methods (Chanas)
  1. Random interval order: $R_{IO}(A, B) = Prob([\dot{x}, \ddot{x}] \geq_{IO} [\dot{y}, \ddot{y}])$
  2. Canonical interval extension of $>$: $R_C(A, B) = Prob(\dot{x} \geq \dot{y}$ and $\ddot{x} \geq \ddot{y})$
  3. Subjective approach: $R_\lambda(A, B) = Prob(\lambda\dot{x} + (1 - \lambda)\ddot{x} \geq \lambda\dot{y} + (1 - \lambda)\ddot{y})$

Ranking fuzzy intervals as families of probability measures

- A fuzzy interval A is viewed as the set of probabilities $\mathcal{P}_A = \{P : \forall S, P(S) \leq \Pi_x(S)\}$
- A possibility (resp. necessity) measure is a coherent upper (resp. lower) probability in the sense of Walley.
- Now one can extend probabilistic methods:
  1. 1st Order Stochastic Dominance : $x \geq y$ iff $\forall \theta, \Pi(x \geq \theta) \geq \Pi(y \geq \theta)$ and $N(x \geq \theta) \geq N(y \geq \theta)$ again...
  2. Comparing upper and lower expected utilites of x and y (e.g. comparing mean intervals $[\int_0^1 \inf A_\alpha d\alpha, \int_0^1 \sup A_\alpha d\alpha]$).
  3. Construct interval- valued preference relations : $\begin{cases} R^*(x,y) = P^*(x \geq y), \\ R_*(x,y) = P_*(x \geq y). \end{cases}$

     and exploit them.

Aggregation operations: qualitative or quantitative

When using a given value scale, its nature dictates what is or not a legitimate aggregation operation. We are faced with a modeling dilemma

- *Using quantitative scales,*
  - we can account for very refined aggregation attitudes, especially trade-off, compensation and dependence between criteria
  - supply a very fine-grained ranking.
  - learn the aggregation operator from data.
  - *but numerical preference data are not typically what humans provide*

Aggregation operations: qualitative or quantitative

When using a given value scale, its nature dictates what is or not a legitimate aggregation operation. We are faced with a modeling dilemma

- *Using qualitative approaches (ordinal or qualitative scales)*
  - We are closer to what human can supply
  - We can model preference dependence structures (CP-nets)
  - But we are *less expressive* in terms of aggregation operations (from impossibility theorems in ordinal case, to only $\min$ and $\max$ in the qualitative case).
  - In the case of finite value scales : *strong lack of discrimination*.
  - what about *bipolar information* (pros and cons) ?
- In fact people make little sense of refined absolute value scales (not more 7 levels). But if we build a preference relation on a set V of alternatives by pairwise comparison, one may get chains $v_1 < v_2 < \cdots < v_m$ with $m > 7$.

WHAT DOES QUALITATIVE MEAN ?

A qualitative aggregation operation only involves operations min and max on a qualitative scale L

- **Negligibility Effect:** steps in the evaluation scale are far away from each other.
  - a focus on the most likely states of nature, on the most important criteria.
  - It implies a lack of compensation between attributes.
  - $\max(5,1,1,1,1) > \max(4,4,4,4,4)$: many 4's cannot compensate for a 5.
- **Drowning effect:** There is no comparison of the number of equally satisfied attributes.
  - $\max(5,1,1,1,1) = \max(5,5,5,5,5)$, because of no counting.

The main idea to improve the efficiency of qualitative criteria is to preserve the negligibility effect, but allow for counting.

Refinements of qualitative aggregation operations

- On qualitative scales the basic aggregation operations are $\min$ and $\max$.
- But comparing vectors of evaluations by $\min$ or $\max$ aggregation operations does not permit to satisfy a natural property: Pareto-Dominance
  - Let $\vec{u} = (u_1, u_2, \ldots, u_n), \vec{v} = (v_1, v_2, \ldots, v_n) \in L^n$
  - $\vec{u} >_P \vec{v}$ iff $\forall i = 1, \ldots n, u_i \geq v_i$ and $\exists j, u_j > v_j$
  - Pareto-dominance of aggregation $f : L^n \to L$: $\vec{u} >_P \vec{v}$ implies $f(\vec{u}) > f(\vec{v})$.
  - One 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 $f : L^n \to L$.....

Refinements of qualitative aggregation operations

- Two known methods to recover Pareto-dominance by refining the min-ordering:
  - **Discrimin**: $\vec{u} >_{dmin} \vec{v}$ iff $\min_{i:u_i \neq v_i} u_i = \min_{i:u_i \neq v_i} v_i$
    delete components that bear equal values in $\vec{u}$ and $\vec{v}$
  - **Leximin** : Rank $\vec{u}$ and $\vec{v}$ in increasing order :
    $\vec{u^\sigma} = (u_{\sigma(1)} \leq u_{\sigma(2)} \leq \ldots, u_{\sigma(n)}); \vec{v^\tau} = (v_{\tau(1)} \leq v_{\tau(2)} \leq \ldots, v_{\tau(n)}) \in L^n$
    $\vec{u} >_{lmin} \vec{v}$ iff $\exists k, \forall i < k, u_{\sigma(i)} = v_{\sigma(j)}$ and $u_{\sigma(k)} > v_{\sigma(k)}$
- $\vec{u} >_P \vec{v}$ 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.

Additive encoding of the leximax and leximin procedures

(L has $m+1$ elements; such encoding is not possible in the continuous case)

- Consider a *big-stepped* mapping $\phi : L \mapsto \mathbb{R}$ such that: $\phi(\lambda_i) > n\phi(\lambda_{i-1})$. Then:

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

  e.g. $\phi(\lambda_i) = k^i$ for $k > n$ achieves this goal. It captures the worst case when $\max(0, 0, \dots, 0, \lambda_i) > \max(\lambda_{i-1}, \dots, \lambda_{i-1})$

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

- Consider the big-stepped mapping $\psi(\lambda_i) = 1 - k^{-i}, k > n$

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

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

The weighted qualitative aggregation operations

Consider a weight distribution $\vec{\pi}$ that evaluates the importance of dimensions, and $\max \pi_i = 1$.

Let $\nu(\lambda_i) = \lambda_{m-i}$ on a scale with $m+1$ steps.

Prioritized Maximum: $P\max(\vec{u}) = \max_{i=1,\dots n} \min(\pi_i, u_i)$

- $P\max(\vec{u})$ is high as soon as there is a totally important dimension with high rating

Prioritized Minimum: $P\min(\vec{u}) = \min_{i=1,\dots n} \max(\nu(\pi_i), u_i)$

- $P\min(\vec{u})$ is high as soon as all important dimensions get high rating.

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 dimensions ($\gamma(A)$ = degree of importance of $A \subseteq \{1, \dots, n\}$).

It is a capacity : If $A \subseteq B$ then $\gamma(A) \leq \gamma(B)$.

Leximax($\trianglerighteq$), Leximin($\trianglerighteq$) (Fargier, Sabbadin(2005))

$Leximax(\trianglerighteq)$ compares vectors in Ω^n using a totally ordered set $(\Omega, \trianglerighteq)$.

Classical leximax: comparison of vectors of utility: $\Omega = L$ and $\trianglerighteq = \geq$

Comparison of matrices using lmax(lmin) $H = [h_{i,j}] : \Omega = L^n$ and $\trianglerighteq = \succeq_{lmin}$ for comparing the rows $H_{j\cdot}$ of the matrix.

IDEA: Shuffle each matrix so that to rank entries on each line in increasing order, and then 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...

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

where $H_{(i)\cdot} = i^{th}$ row of H w.r.t. $\succeq_{lmin}$.

It is a (very discriminative) complete and transitive relation.

Take the minimum on elements of the lines, and the maximum across lines

$A =$

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

$B =$

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

$$\max_i \ \min_j \ a_{i,j} = \max_i \ \min_j \ a_{i,j}$$

Reordered in increasing order inside lines:

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

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

Take the minimum on elements of the lines, and the maximum across lines

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

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

Lines reordered top down:

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

Lexi-refinement of the prioritized maximum

The $Leximax(Leximin(\geq))$ procedure refines the ranking of matrices according to $\max_i \ \min_j \ h_{i,j}$

If alternative $\vec{u}$ is encoded as a $n \times 2$ matrix $F^u = [f_{ij}]$ with $f_{i1} = \pi_i$ and $f_{i2} = u_i), i = 1, \ldots, n$:

$$P\max(\vec{u}) = \max_{i=1,n} \min_{j=1,2} f_{ij}$$

$P\max$ is refined by a $Leximax(Leximin(\geq))$ procedure :

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

Weighted average refinement of the prioritized maximum

Claim :There exists a weighted average $AV_+(\vec{u})$ representing $\succeq_{lmax(\succeq_l min)}$ and thus refining $P\max$. Define a transformation χ of the scale L such that :

$\max_i \min(\pi_i, u_i) > \max_i \min(\pi_i, v_i)$ implies

$\sum_{i=1,\dots n} \chi(\pi_i) \cdot \chi(u_i) > \sum_{i=1,\dots n} \chi(\pi_i) \cdot \chi(u_i)$

Worst case: $\max(\min(\lambda_j, \lambda_j), 0, \dots, 0) >$
$\max(\min(\lambda_j, \lambda_{j-1}), \min(1_L, \lambda_{j-1}), \dots, \min(1_L, \lambda_{j-1}))$

Sufficient condition:

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

Refining Sugeno integral

- The first idea is to apply the refinement of $P\max$ to Sugeno integral and to refine the capacity γ that estimates the importance of criteria.
- The second idea is to use a Choquet integral to refine a Sugeno integral.

Put the aggregation operation in the form

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

where $\gamma^{\#}(A)$ is a “Qualitative” Moebius transform :
$\gamma^{\#}(A) = \gamma(A)$ if $\gamma(A) > \max_{B \subsetneq A} \gamma(B)$ and 0 otherwise.

Leximax refinement of a Sugeno integral

- The capacity γ is such that $\gamma(A) = \max_{E \subseteq A} \gamma_{\#}(E)$
- We can use use a super-increasing transformation of $\gamma_{\#}$ into a mass function $m_{\#} : 2^S \mapsto [0, 1] : m_{\#}(E) = \chi(\gamma_{\#}(E))$
- Similarly, we can refine the ordering provided by the capacity by applying the leximax refinement to the vector of $\gamma_{\#}(E)$'s.
- This refined ordering is representable 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 $E_{\#}^{lsug}(\vec{u}) = \sum_{A \subseteq S} m_{\#}(A) \cdot \min_{s \in A} \chi(u_i)$

Bipolar choice and ranking: a basic framework

1. A set of n dimensions, viewed as arguments, whose domain is t
   (a) The positive scale $\{0, +\}$ for positive arguments
   (b) The negative scale $\{0, -\}$ for negative arguments
2. A set of potential decisions evaluated in terms of pro and con arguments $\vec{u}, \vec{v}, \ldots$
3. A totally ordered scale L expressing the relative importance of arguments

So we get

- If $u_i = +$, then i is an argument for $\vec{u}$
- If $u_i = -$ then i is an argument against $\vec{u}$
- If $u_i = 0$ then i does not matter for $\vec{u}$

Let $U = \{i, u_i \neq 0\}$ the arguments that matter for $\vec{u}$

List the pros ($U^+ = \{i, u_i = 1\}$) and the cons ($U^- = \{i, u_i = -1\}$) of $\vec{u}$.

Compare $\vec{u}$ and $\vec{v}$ = comparing the pairs (U^-, U^+) and (V^-, V^+)

Cumulative prospect theory and extensions

Cumulative Prospect Theory for multiaspect decision (Tversky & Kahneman, 1992):

- Measure the importance of positive affects and negative affects of decisions *separately*, by two monotonic set functions $\sigma^+(U^+)$, $\sigma^-(U^-)$
- Compute the net predisposition $N(a) = \sigma^+(U^+) - \sigma^-(U^-)$.
- Rank the decisions

If positive and negative affects are not independent :

- use bi-capacities on a bipolar scale: $N(a) = g(U^+, U^-)$ increasing with the former, decreasing with the latter (Grabisch-Labreuche)
- use bipolar capacities $N(a) = (g^+(U^+, U^-), g^-(U^+, U^-))$ on bivariate unipolar scales (Greco-Slowinski)

How to evaluate decisions from *qualitative* bipolar information ?

Contrary to what classical decision theory suggests, people can make decisions in the face of several criteria **without numerical utility nor criteria importance assessments** (works by Gigerenzer, for instance): How to extend the $\min$ and $\max$ rules if there are both positive and negative arguments?

Qualitativeness assumption (focalisation) : 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.

$$\Pi(U) = \max_{x \in U} \pi_i$$

Express preference between $\vec{u}$ and $\vec{v}$ in terms the pairs ($\Pi(U^-)$, $\Pi(U^+)$) and ($\Pi(V^-)$, $\Pi(V^+)$) of positive and negative affects, based on the relative importance of these affects.

1. basic decision rules, complete and incomplete, usually weakly discriminative
2. refinements obeying a form of independence between affects
3. further (lexicographic) refinements counting affects of equal importance

The Bipolar Possibility Relation

Principle at work: *Comparability*: When comparing $\vec{u}$ and $\vec{v}$, any argument against $\vec{u}$ (resp. against $\vec{v}$) is an argument pro $\vec{v}$ (resp. pro $\vec{u}$).

The agent focuses on the most important argument regardless of its polarity.

$$\vec{u} \succeq^{Biposs} \vec{v} \iff \max(\Pi(U^+), \Pi(V^-)) \geq \max(\Pi(V^+), \Pi(U^-))$$

- $\succeq^{Biposs}$ is complete, but only its strict part is transitive.
- This relation collapses to the maximin rule if all arguments are negative and to the maximax rule if all arguments are positive.
- Similar to CPT: $\vec{u} > \vec{v} \iff \sigma^+(U^+) + \sigma^-(V^-) > \sigma^-(V^+) + \sigma^-(U^-)$. Here, we change $+$ into $\max$.

This relation is sound and cognitively plausible but it is too rough (too many indifference situations).

The full lexi-bipolar rule

Idea (originally B. Franklin's): Canceling arguments of equal importance for $\vec{u}$ or against $\vec{v}$, by arguments for $\vec{v}$ or against $\vec{u}$ until we find a difference on each side.

An even more decisive procedure, cancelling conflicts: A complete and transitive refinement of $\succeq^{Bilexi}$

Let $U^+_\lambda = \{i \in U^+, \pi_i = \lambda\}$ be the arguments for $\vec{u}$ with strength λ. (resp. U^-_λ the arguments against $\vec{u}$ with strength λ.).

$$\vec{u} \succeq^{Lexi} \vec{v} \iff \exists \lambda \in L \text{ such that } \begin{cases} (\forall \beta > \lambda, & |U^+_\beta| - |U^-_\beta| & = & |V^+_\beta| - |V^-_\beta|) \\ \text{and} & (|U^+_\lambda| - |U^-_\lambda| & > & |V^+_\lambda| - |V^-_\lambda|) \end{cases}$$

This decision rule

- Generalizes Gigerenzer's ”take the best” heuristic and can be encoded in the cumulative prospect theory framework
- Was empirically tested, and proves to be the one people use when making decisions according to several criteria.

Conclusion

- 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 getting qualitative.
  3. The right question is : how to encode qualitative techniques on numerical scales.
- There is a strong need to develop original approaches to multicriteria decision analysis that are not a rehashing of existing techniques with ad hoc fuzzy 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

- Refine any qualitative aggregation function using discri-schemes or lexi-schemes
- Computational methods for finding discrimin-leximin solutions to fuzzy optimization problems.
- Solving decision problems with fuzzy bipolar information
  - With pros and cons
  - multicriteria optimisation (positive preference) under fuzzy constraints (graded feasibility)
- Behavioral axiomatization of aggregation operations in MCDM, decision under uncertainty and fuzzy voting methods
- Principled general framework for ranking fuzzy intervals.
- Semantics of fuzzy preference modeling (probabilistic, probabilistic, distance-based,....) and their learning from data.
- Fuzzy choice functions