Qualitative decision theory:

issues, axioms, aggregation functions, refinements, bipolarity

Didier Dubois
IRIT - Universite Paul Sabatier -
31062 Toulouse Cedex (France)
Works done with Helène Fargier, Henri Prade

July 2017

A common framework

Three similar problems :

1. *Decision under uncertainty*: Decision outcomes depend on the state of the world. The latter is ill-known.
2. *Decision with several criteria*: the worth of a decision depends on the viewpoint. Viewpoints or criteria have various importance levels
3. *Group decision*: the worth of a decision depends on the person. Persons can be more or less important.

Similar problems but, different approaches according to the traditions.

Various traditions at odds with each other

1. *Decision under uncertainty*: The use of weighted averages. Axiomatic approaches to justify it (Von Neumann, Savage...)
2. *Group decision*:In voting theory the tradition is ordinal: preference relations to be merged (impossibility theorems and how to cope with them)
3. *Decision with several criteria*: Both approaches : weighted averages (anglo-saxon) and preference relations to be merged (Continental Europe: ELECTRE, etc.)

Various working assumptions

1. *Decision under uncertainty*:
   The numerical setting with probabilities is natural for economics
   Common scale for utility, and probability.

2. *Group decision*:
   one value scale per person, or even no value scale at all
   Equity between persons.

3. *Decision with several criteria*:
   Not always easy to use the same value scale for all criteria: commensurateness issues
   Possibility of incomparability between decisions

Motivation for our work

In the setting of Artificial Intelligence (recommender systems, cognitive robotics, but other fields as well), the use of decision rules based on numerical aggregation functions is not always natural.

- What if probabilities cannot be elicited (without total ignorance about the state) ?
- What if we cannot or have no time to elicit fine-grained preferences ?
- If the state space is large you may not assess the probability and utility of each state (adopt a coarse description)
- Information systems advising persons,
  - do not ask too many questions about a user for his preferences
  - Collect numbers representing probabilities /importance levels and utilities and computing averages looks debatable.

Example: Assess the merits of a paper for a journal

Principle

Why should we use numbers that look arbitrary or hard to collect, if we can address the issue of evaluating decisions in a reasoned approach without numerical calculations?

$\rightarrow$ gain in robustness and parsimony of data.

$\rightarrow$ methods that lends themselves to a logical representation (explainability).

Two possible choices, for decision under uncertainty

- Preference and likelihood on distinct non-commensurate scales
- Preference and likelihood on commensurate scales

The first line is very restrictive in the purely ordinal setting (impossibility theorems).

The second approach (absolute) relies on a notion of certainty equivalent and can rely on a finite ordinal scale (operations max and min).
(a finite ordered set of classes)

Requirements for a qualitative approach

- **Faithfulness** to the preference data supplied by an agent : ordinal and linguistic.
- **Cognitive relevance**: short value scales
- **Good discrimination**: in agreement with Pareto dominance
- **Good decisiveness power**: minimize incomparability, total order if possible
- Should account for the **decision-maker attitude in front of uncertainty**.

These requirements may prove conflicting

WHAT DOES QUALITATIVE IMPLY ?

An aggregation operation in a qualitative scale L only involves operations min and max, which assumes:

- **Negligibility Effect:** steps in the evaluation scale are far away from each other.
  - It implies a lack of compensation between attributes.
  - $\min(5,5,5,5,3) < \min(4,4,4,4,4)$: many 4's cannot compensate for a 3.
  - a focus on the most likely states of nature, on the most important criteria.
- **Drowning effect:** There is no comparison of the number of equally satisfied attributes.
  - $\min(5,1,1,1,1) = \min(5,5,5,5,1)$, because of no counting.

A classical framework

- A set S of states of the world (or criteria, or voters)
- A set of decisions (acts f)
- A set of possible more or less favorable consequences X for decisions
- An act is a mapping from S to X : $x = f(s)$.

Basic Problem in DMU : Given partial knowledge on the state of the world, and a preference relation on consequences of acts, find a preference relation between acts (on X^S) suggesting a good decision.

Representing partial knowledge in possibility theory

Simplest form of incomplete knowledge: it is only known that $s_0 \in E \subseteq S$.

- If $A \cap E = \emptyset$, A is impossible: $\Pi(A) = 0$, and 1 otherwise.
- If $E \subseteq A$, A is certain: $N(A) = 1 - \Pi(A^c) = 1$, and 0 otherwise.
- Characteristic properties: $\Pi(A \cup B) = \max(\Pi(A), \Pi(B)$;
  $N(A \cap B) = \min(N(A), N(B)$
- 3 situations: A can be
  - sure : $N(A) = 1$, equivalently, $\Pi(A^c) = 0$
  - impossible : $\Pi(A) = 0$, equivalently, $N(A^c) = 1$
  - unknown: $N(A) = 0 = N(A^c)$ equivalently, $\Pi(A) = \Pi(A^c) = 1$,
    ($\equiv 0 < P(A) < 1$)

Representing partial knowledge in possibility theory

More generally $(S, \leq_\pi)$ a *plausibility ranking* on S (weak order), representable by a discrete possibility distribution $\pi : S \to L$ on a finite totally ordered scale ranging from plausible to impossible.

- $A \leq_\Pi B \iff \exists w \in A, \forall w' \in B w \leq_\pi w'$
- Characteristic axiom: if $A \leq_\Pi B$ then $\forall C, A \cup C \leq_\Pi B \cup C$
- **Degree of possibility** : $\Pi(A) = \max_{s \in A} \pi(s)$
  - focus on the most normal situations, neglect others)
  - The only set-functions compatible with axiom of comparative possibility.
- **Degree of necessity**: $N(A) = n(\Pi(A^c))$ (n is the order reversing map on L)
- A is all the more certain as its complement is less plausible.

A possibility distribution is always subjective.

Elementary qualitative preference functionals

- *Ignorance and pessimism* : $W_u^-(f) = \min_{s \in S} u(f(s))$, where u is a qualitative utility function on a finite scale.
- *Ignorance and optimism* : $W_u^-(f) = \max_{s \in S} u(f(s))$.
- *Realistic pessimism* : $W_u^r(f) = \min_{s \in A^\star} u(f(s))$
  $A^\star =$ most plausible states according to a normality ordering.
- A qualitative possibility distribution π (plausibility of states, importance of criteria) on the same scale as utility.
  - *Possibilistic Optimistic* $W_{\pi,u}^+(f) = \max_{s \in S} \min(\pi(s), u(f(s)))$.
    (Prioritized max): enough that one important criterion is satisfied
  - *Possibilistic pessimistic*: $W_{\pi,u}^-(f) = \min_{s \in S} \max(n(\pi(s)), u(f(s)))$
    (Prioritized min): what you know implies what you wish
    ($\max(n(a), b)$ Kleene implication)

A general qualitative functional

The most general is Sugeno integral

$$S_{\gamma,u}(f) = \max_{A \subseteq S} \min(\gamma(A), \min_{s \in A} u(f(s)))$$

measures to what extent there is a group of important criteria that are all satisfied.

- in MCDM: $\gamma(A)$ is the importance level of the group of criteria A, and γ is inclusion-monotonic, in DMU: plausibility, belief
- The level of pessimism is governed by the set-function γ.
- Special cases:
  - *Prioritized maximum* If $\gamma = \Pi$ then $W^{+}_{\pi,u}(f) = S_{\Pi,u}(f)$
    Optimistic evaluation for DMU, redundant criteria in MCDM
  - *Prioritized minimum*: If $\gamma = N$ then $W^{-}_{\pi,u}(f) = S_{N,u}(f)$ with $N(A) = n(\max_{s \notin A} \pi(s))$ and n a scale reversal map
    Pessimistic Evaluation in DMU, positive synergy of criteria in MCDM

Rule-based Representation by Sugeno Integral in MCDM

$$S_{\gamma,u}(f) = \max_{A \subseteq S} \min(\gamma(A), \min_{s \in A} u(f(s)))$$

If we know the relative importance of some groups of criteria $\mathcal{F} = \{(A_i, \gamma_i) : i = 1, \dots k\}$ with $\max_i \gamma_i = 1$, we can

- define a monotonic set-function $\gamma(A) = \max_{A_i \subseteq A} \gamma_i$ and we compute the value of decision f as $S_{\gamma,u}(f)$.
- Or write a set of if-then rules of the form :

  If $u(f_j) \geq \theta, \forall s_j \in A_i$ and $\gamma(A_i) \geq \theta$, then $u(f) \geq \theta$ for $i = 1, \dots k$.

  (Dubois, Prade, Rico, 2014)

  This is close to some works by Greco.

Sugeno integral is a median

Sugeno integral can be equivalently written under various forms:

- Consider the ranking $u(f(s_1)) \geq \cdots \geq u(f(s_{|S|}))$ and let

$$F_i = \{s \in S | u(f(s)) \geq u(f(s_i))\} = \{s_1, s_2, \ldots s_i\}.$$

$$S_{\gamma,u}(f) = \max_{i=1}^{|S|} \min(u(f(s_i)), \gamma(F_i))$$

  $F_i$is the largest A s.t $\min_{s \in A} u(f(s)) = u(f(s_i))$.

- $S_{\gamma,u}(f)$ is the median of $2|S| - 1$ values :

$$\{u(f(s_i)), i = 1 \ldots |S|\} \cup \{\gamma(\{s_1, s_2, \ldots s_i\}) : i = 1 \ldots |S| - 1\}$$

Example: the h-index.

$$S_{\gamma,u}(f) = \max_{i=1}^{|S|} \min(u(f(s_i)), \gamma(F_i))$$

- S = set of publications, $L = \mathbb{N}$
- i = number of citations
- $u(f(s)) =$ number of citations of publication s by author f
- $F_i = \{s_1, s_2, \ldots s_i\}$ the set of publications cited at least i times
- $\gamma(\{s_1, s_2, \ldots s_i\}) = i$.

Median decomposability (Marichal)

For $c \in L$ and $i \in [1, 2, \ldots n]$, and any vector $\mathbf{v} \in L^n$, let
$\mathbf{v}_i^c = (v_1, \ldots, v_{i-1}, c, v_{i+1}, \ldots, v_n)$.

A function $\psi \colon L^n \to L$ is *median decomposable* if for each index i

$$\psi(\mathbf{v}) = median\big(\, \psi(v_i^0)\, ,\, v_i\, ,\, \psi(v_i^1)\, \big), \qquad \text{for every } \mathbf{v} \in L^n.$$

Fact Every median decomposable function is order-preserving.
$\psi(v_1, v_2) = median(v_1, c, v_2)$ is an associative function
Sugeno integral is an idempotent lattice polynomial using $\min, \max$

Theorem (Marichal): A function $p \colon L^n \to L$ is

1. a lattice polynomial function **iff** it is median decomposable.
2. a Sugeno integral **iff** it is idempotent and median decomposable.

Sugeno integral viewed as a lower optimistic possibilistic integral

Let $\gamma : 2^S \to L$ be a monotonic set-function. **The possibilistic core of γ is :**

$$\mathcal{C}_q(\gamma) = \{\pi : \Pi(A) \geq \gamma(A), \forall A \subseteq S\}.$$

Never empty, it has least elements forming the set $\mathcal{C}_q^*(\gamma) : \gamma(A) = \min_{\pi \in \mathcal{C}*_q(\gamma)} \Pi(A)$

A Sugeno integral is a lower qualitative optimistic possibilistic functional and an upper qualitative pessimistic one

Proposition

- $\mathcal{S}_\gamma(f) = \min_{\pi \in \mathcal{C}*_q(\gamma)} \mathcal{S}_\Pi(f) = \min_{\pi \in \mathcal{C}_q^*(\gamma)} W^+_{\pi,u}(f)$.
- $\mathcal{S}_\gamma(f) = \max_{\pi \in \mathcal{C}*_q(\gamma^c)} \mathcal{S}_N(f) = \max_{\pi \in \mathcal{C}*_q(\gamma^c)} W^-_{\pi,u}(f)$,
  with $\gamma^c(A) = n(\gamma(A^c))$ and $N = \Pi^c$.

If $\mathcal{C}_q^*(\gamma)$ is small, the exponential nature of Sugeno integral disappears.

Main representation results

Looking for theoretical foundations similar to classical utility theory

- IJCAI 1995: Axiomatisation of possibilistic criteria a la Von Neumann (given π) (with Prade).
- UAI98, EJOR 2001: Axiomatisation of possibilistic criteria and Sugeno integral a la Savage (Thesis Sabbadin)
- JACM 2002, AIJ2003: Axiomatisation à la Savage for the plausible dominance rule (with H. Fargier, P. Perny)
- Ecsqaru07 Conditional possibilistic functionals (with B. Vantaggi)
- ECAI 2012 multiscale Sugeno integral (with M. Couceiro)

The 5 Savage Postulates

The reference in terms of rationality : Savage.

- Totally ordered set of acts (weak order) including
  - constant acts : $f(s) = x \in X, \forall s \in S$.
  - binary acts : $xAy(s) = x$ if $s \in A, y < x$ otherwise (gain if A occurs, loss otherwise).
- **Sure thing principle**: preference among two acts not affected by states where acts share the same consequence.
- Preference between constant acts is stable when conditioning on non-impossible states.
- Preference between binary acts is independent of the involved consequences
- There are two strictly ordered acts

In the finite setting these postulates are not enough to jusfify probability!!!

Ordinality without commensurateness

Suppose the preference between acts does not depend on the absolute worth of their consequences:

Definition: (f, g) is said to be **ordinally equivalent** to (f', g') if and only if for each state, the consequence of f is preferred to the one of g iff the consequence de f' is preferred to the one of g': $\forall s$, $f(s)$ better than $g(s)$ iff $f'(s)$ better than $g'(s)$.

Ordinal invariance axiom : If (f, g) is ordinally equivalent to (f', g') then $f \succ g$ iff $f' \succ g'$

This axiom + the 5 of Savage (except for transitivity of non-strict preference between acts) justify the likely dominance rule :

$$\exists \mathcal{F}, \forall \pi \in \mathcal{F}, f \succ g \iff \Pi(f > g) > \Pi(g > f)$$

where $\mathcal{F}$ is a set of possibility distributions π.

Ordinality without commensurateness

Likely dominance rule

$$\exists \mathcal{F}, \forall \pi \in \mathcal{F}, f \succ g \iff \Pi(f > g) > \Pi(g > f)$$

where $\mathcal{F}$ is a set of possibility distributions π.

- This decision rule does not require commensurateness
- It is a qualitative acyclic variant of Condorcet rule
- The underlying uncertainty representation is nothing but possibilistic.
- but it displays much incomparability.

Qualitativeness with commensurateness

- If acts are supposedly totally ordered according to global preference $\succeq$, then there exists a common value scale for utility and uncertainty.
- This scale is just the linearly ordered quotient set $L = X^S / \sim$, where $\sim$ is the indifference relation between acts (an equivalence relation).
- The utility function : $u(x)$ is the equivalence class of the constant act $f_x(s) = x, \forall s \in S$
- The likelihood value $\gamma(A)$ is the equivalence class of the binary act $1A0$: sup X if A occurs and inf X if not.
- The notion of certainty equivalent $u(z) = u(xAy)$ of a risky act xAy exists (at least we can compare z with xAy).

Axioms for possibilistic decision theory

- **Three Savage axioms** :
  - $\succeq$ a weak order
  - $x \succeq y$ implies $xAf \succeq yAf$ on non impossible events
  - $f \succ g$ for at least two acts
- **Conjunctive Dominance** : $f \succ g$ and $h \succ g$ imply $f \wedge h \succ g$ (CD)
- **Disjunctive Dominance** : $f \succ g$ and $f \succ h$ imply $f \succ h \vee g$ (DD)
- **Restricted Conjunctive Dominance**: CD where $h = x$ is a constant act (RCD)
- **Restricted Disjunctive Dominance**: RD where $h = x$ is a constant act (RDD)

$f \wedge h$ (resp. $f \vee h$) is an act obtained by taking for each s the worst (resp. best) among consequences $f(s)$ and $g(s)$.

Representation results for qualitative utility functionals

- The (pessimistic) prioritized min $W^-_{\pi,u}(f) = \min_{s \in S} \max(n(\pi(s)), u(f(s)))$ is obtained using restricted disjunctive dominance (RDD) and general conjunctive dominance (CD).
- The optimistic prioritized maximum $W^+_{\pi,u}(f) = \max_{s \in S} \min(\pi(s), u(f(s)))$ is obtained using restricted conjunctive dominance (RCD) and general disjunctive dominance (DD).
- Sugeno integral $S_{\gamma,u}(f) = \max_{A \subseteq S} \min(\gamma(A), \min_{s \in A} u(f(s)))$ is obtained using restricted conjonctive (RCD) and disjunctive dominance (RDD).
- For Sugeno integral, alternatively take the three axioms of Savage plus **non-compensation** :

  $u(xAy)$ is equal to, either $u(x)$, or $u(y)$, or yet $\gamma(A) = 1A0$

  (actually, their median).

Discussion

One may legitimately object to such criteria on the real line: let f, x, y be acts such that

- f : bet 1000 euros on A against nothing, with certainty equivalent $u(x) = 400$ euros
- y : get 400 $-\epsilon$ euros ; **then** $f \succ y : x \succ y$**.**
- so $f \wedge x$: bet 400 euros on A against nothing

RCD implies $f \wedge x \succ y$, which means that betting 400 euros on A against nothing is always better than receiving 400 $-\epsilon$ euros however small ϵ is.

Response: The value scale being qualitative means a big jump between consecutive values: if $x \succ y$ the utility of y is significantly smaller than the one of x (no epsilons....). Moreover qualitative decision rules make sense for one-shot decisions only.

Decision-theoretic approach to Sugeno integral: MCDM

There are distinct attribute ranges $X_1, \ldots X_n$, one for each attribute (state) s_i.

Now we consider a **preference relation**, a weak order on $\succeq$ on $\mathbf{X} = X_1 \times \cdots \times X_n$.

And we try to find a representation of $\succeq$ in terms of a **Sugeno utility functional** :

$$f(\mathbf{x}) = \mathcal{S}_\mu\big(u_1(x_1), \ldots, u_n(x_n)\big),$$

where each local $u_i \colon X_i \to L = \mathbf{X}/\sim$ is order-preserving, and $\mathcal{S}_\mu$ is a Sugeno integral.

Theorem (ECAI 2012): A function $f \colon \mathbf{X} \to L$ is a Sugeno utility functional if and only if f is pseudo-median decomposable, that is, for each $k \in [n]$ there is a local utility function $u_k \colon X_k \to Y$ such that

$$\psi(\mathbf{x}) = \text{median}(\psi(\mathbf{x}_k^0), u_k(x_k), \psi(\mathbf{x}_k^1)) \quad (1)$$

Sugeno utility functionals are more general

Two criteria, with $X_1 = X_2 = X = \{1, 2, 3\} = L$ with $3 > 2 > 1$, and a preference relation $\preceq$ on X^2 with rank function:

$$[(3,3)] = \{(3,3), (2,3)\},$$
$$[(2,2)] = \{(3,2), (3,1), (2,1), (1,3), (2,2)\},$$
$$[(1,1)] = \{(1,1), (1,2)\}.$$

This relation does not satisfy (RDD) : $(1,3) \vee (2,2) = (2,3)$
nor (RCD) : $(1,3) \wedge (2,2) = (1,2)$.
It cannot be represented by a Sugeno integral with a single utility function.
However, $\preceq$ is represented by the Sugeno utility functional

$q(x_1, x_2) = (2 \wedge u_1(x_1)) \vee (2 \wedge u_2(x_2)) \vee (3 \wedge u_1(x_1) \wedge u_2(x_2)),$
with $u_1(3) = u_2(3) = 3$, $u_1(1) = u_2(1) = 1$, $u_1(2) = 3$, $u_2(2) = 1$.

Merits of decision evaluation methods

	Cognitive Relevance	Discrimination	Decisiveness	Attitude
Expected Utility	No	Yes	Yes	Yes
Likely Dominance	Yes	No	No	No
Sugeno Integrals	Yes	No	Yes	Yes

Qualitative approach : Possibility to encode the evaluation process in a logical format

- A base of goals $G(f)$ and a knowledge base K
- Pessimistic possibilistic criterion: choose f such that $K \vdash G(f)$.

Weaknesses and repairs

Qualitative absolute decision rules lack discrimination (drowning effect).

If the value scale has n steps, acts or decisions fall in n classes: many ties

For instance: two different decisions can be judged equivalent % max or min even if one Pareto-dominates the other.

IDEA: *Refine criteria and break the ties*

- Known refinements of minimum and maximum : leximin and leximax are powerful to rank vectors of ratings
- One can encode leximin and leximax by sums of super-increasing numbers.

Leximin (resp. leximax) refinement of the min (resp. max) criterion

Case of total ignorance: compare utility vectors $\vec{f} = (f_1, \ldots, f_n)$ where $f_i = u(f(s_i))$

$W_u^-(f) \;=\; \min_{i=1\ldots n}\; f_i \qquad\qquad W_u^+(f) \;=\; \max_{i=1\ldots n}\; f_i$

For any $\vec{f} \in L^N$, let $f_{(k)}$ be the k-th *largest* element of $\vec{f}$ (i.e. $f_{(1)} \geq \ldots \geq f_{(n)}$): reordered vector $\vec{f}^{\uparrow}$.

Leximax: $\vec{f} \succ_{lmax} \vec{g} \Leftrightarrow f_{(i^*)} > g_{(i^*)}$ with $i^* = \max\{i : f_{(i)} \neq g_{(i)}\}$
(Anti-lexicographic ordering between $\vec{f}^{\uparrow}$ and $\vec{g}^{\uparrow}$).

Property : $W_u^+(f) > W_u^+(g) \implies \vec{f} \succ_{lmax} \vec{g}$

Leximin: $\vec{f} \succeq_{lmin} \vec{g} \Leftrightarrow f_{(i_*)} > g_{(i_*)}$ with $i_* = \min\{i : f_{(i)} \neq g_{(i)}\}$
(lexicographic ordering between $\vec{f}^{\uparrow}$ and $\vec{g}^{\uparrow}$)

Property : $W_u^-(f) > W_u^-(g) \implies \vec{f} \succ_{lmin} \vec{g}$

Additive encoding of the leximax and leximin procedures

L and S finite: simulate the leximax by a sum of utilities

$\phi : L \mapsto L$ such that:

$$\max_{i=1,\dots n} f_i > \max_{i=1,\dots n} g_i \text{ implies } \sum_{i=1,\dots n} \phi(f_i) > \sum_{i=1,\dots n} \phi(g_i) \tag{2}$$

e.g. $\phi(\lambda_i) = k^i$ for $k > n$ achieves this goal.

$$f \succ_{lmax} g \text{ if and only if } \sum_{i=1,\dots n} \phi(f_i) > \sum_{i=1,\dots n} \phi(g_i). \tag{3}$$

A similar encoding for the leximin procedure : use e.g. the big-stepped mapping $\psi(\lambda_i) = 1 - k^{-i}, k > n$

(such encoding is not possible in the continuous case)

Refinement of max-min possibilistic criteria

Main ideas

- Embed leximin inside leximax and conversely: fine-grained ranking of matrices
- Project the ordinal scale (for utility + possibility) inside positive reals, images forming a super-increasing sequence.

Results (Fargier et Sabbadin IJCAI 2003; AIJ 2005)

- we can refine criteria $W^-_{\pi,u}(f)$ and $W^+_{\pi,u}(f)$ by weighted averages that simulate $leximin(leximax)$
- Possibility distributions are mapped to big-stepped probabilities
- Utility values are super-increasing
- for $W^-_{\pi,u}(f)$(pessimistic) the utility function is concave, for $W^-_{\pi,u}(f)$(optimistic) the utility function is convex.

Leximax($\trianglerighteq$), Leximin($\trianglerighteq$) Fargier, Sabbadin(2003)

$Leximax(\trianglerighteq)$ defined from a totally ordered set $(\Omega, \trianglerighteq)$ and compares vectors of Ω^n.

- **Classical leximax**: comparison of vectors of utility: $\Omega = L$ and $\trianglerighteq = \geq$
- **Comparison of matrices** $H = [h_{i,j}]$: $\Omega = L^n$ and $\trianglerighteq = \succeq_{lmin}$ for comparing the rows $H_{j\cdot}$ of the matrix.

IDEA: Shuffle each matrix so as to rank entries on each line in **increasing** order, and then rows top-down in **decreasing** lexicographic order. Matrix H becomes H^* with entries $[h_{(i),(j)}]$. Then compare the two matrices lexicographically, first the top rows (via leximin), then if equal the second top rows, etc...

$F \succ_{lmax(lmin)} G \Leftrightarrow \exists i$ s.t. $\forall j < i, F_{(j)\cdot} \sim_{lmin} G_{(j)\cdot}$ and $F_{(i)\cdot} \succ_{lmin} G_{(i)\cdot}$

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

It is a very discriminative complete and transitive relation.

Lexi-refinement of $W^+_{\pi,u}$

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

If act f is encoded as a $n \times 2$ matrix $F = [f_{ij}]$ with $f_{i1} = \pi(s_i)$ and $f_{i2} = u(f(s_i)), i = 1, \ldots, n$:

$$W^+_{\pi,u}(f) = \max_{i=1,n} \min_{j=1,2} f_{ij}$$

$\succeq_{W^+_{\pi,u}}$ be refined by a $Leximax(Leximin(\geq))$ procedure :

$$W^+_{\pi,u}(f) > W^+_{\pi,u}(g) \implies F \succ_{lmax(\succeq lmin)} G$$

EU-refinement of the optimistic utility

Claim :There exists an expected utility $EU_+(f)$ representing $\succeq_{lmax(\succeq lmin)}$ and thus refining $W^+_{\pi,u}$

Define a transformation of the scale L :

$$\chi^*(\lambda_i) = \frac{K}{Card(S)^{2^{i+1}}}, i = 1, m-1 \qquad \chi^*(\lambda_m) = 0.$$

We can normalize χ^* such that $\chi^*(\pi)$ is a probability distribution. Then define :

$$EU_+(f) = \sum_{i=1,\dots n} \chi^*(\pi(s_i)) \cdot \chi^*(u(f(s_i)))$$

Theorem:

$$EU_+(f) \geq EU_+(g) \iff F \succeq_{lmax(\succeq lmin)} G$$

CONSEQUENCES

- $EU_+(f)$ is an expected utility that refines $W^+_{\pi,u}$
- $\succeq_{lmax(\succeq_{lmin})}$ satisfies Savage's axioms 1 to 5 and can be represented by an expected utility !
- $p = \chi^*(\pi)$ is big-stepped ($p(s) > P(\{s', p(s') < p(s)\})$) and ordinally equivalent to π.
- P refines both the possibility and the necessity measures induced by π (leximax refinement of Π).
- $\chi^*(u(\cdot))$ is convex (optimism), ordinally equivalent to u and big-stepped ($\chi^*(\lambda_i) > Card(S)\chi^*(\lambda_{i+1})$)
- If π is uniform it reduces to the leximax comparison of vectors.

Refining Sugeno integral

- Prioritized min and max satisfy a weak form of the sure thing principle
- Sugeno integral severely violates it (order reversal).

So, cannot refine Sugeno integral by a weighted average

- Sugeno integral is minitive and maxitive for comonotonic acts.
- It is the qualitative counterpart of Choquet integral (additive for such acts)
- *Hence refine Sugeno by Choquet*

3 causes causes for drowning with Sugeno integral

1. maximum
2. minimum
3. The monotonic set-function

Results (IJAR, 2009, Ecsqaru09)

One uses a non-redundant form of Sugeno integral:

$$S_\gamma(f) = \max_{A \subseteq S : \gamma^\#(A) > 0} \min(\gamma^\#(A), \min_{s \in A} u(f(s))$$

with $\gamma^\# =$ minimal information to build γ : $\gamma(A) = \max_{E \subset A} \gamma^\#(E)$.

1. It is possible to turn Sugeno integral into a Choquet integral that refines it by turning fonction $\gamma^\#$ into a (super-increasing) mass function in the sense of Dempster-Shafer:

   $$C_{\mu_\gamma}(f) = \sum_{A \subseteq S} m_{\gamma^\#}(A) \cdot \min_{s \in A} \tau(u(f(s)))$$

   with $m_{\gamma^\#}(A) = \tau(\gamma^\#(A)) \in \mathbb{R}$

2. A lexicographic refinement of the set-function γ.

3. It uses the transformation of maxmin into a weighted arithmetic mean already met.

Limitation of preference relations in terms of expressive power

A preference relation having good properties ranks alternatives in terms of relative merit.

But, the intrinsic merit of each alternative is then lost.

Namely

- In some cases, the best choice may be an alternative that the decision-maker does not really like.
- In some other cases even the worst alternative is agreeable.

The notion of **polarity** adds expressive power to qualitative preference representations.

This notion is absent from Savage decision theory as it uses an interval scale for utilities.

Deciding with good and bad affects

People make choices by checking the good sides and the bad sides of alternatives separately. Then they choose according to whether the good or the bad sides are stronger.

Example : choosing a house

- Positive arguments : presence of a garden, nice piece of architecture
- Negative arguments : desperately high sale price, seismic area, close to an airport landing area.

Cognitive psychologists made experiments confirming the *bipolar* nature of decision evaluation by humans : good affects and bad affects are processed in different parts of the brain

Accounting for the positive/negative distinction requires the use of a bipolar evaluation scale

For instance : Kahneman -Tverski Cumulative prospect theory: numerical additive, and bipolar.

Qualitative framework and notations

1. A set D of potential decisions $a, b, c, \ldots$
2. A set of X of qualitative criteria, viewed as attributes whose domain is the bipolar scale $\{-, 0, +\}$. So we get
   - If $x(a) = +$, then x is an argument for d
   - If $x(a) = -$ then x is an argument against d
   - If $x(a) = 0$ then x does not matter for d
3. A totally ordered scale L expressing the relative importance $\pi(x)$ of criteria

Let $A = \{x, x(a) \neq 0\}$

List the pros ($A^+ = \{x, x(a) = +\}$) and the cons ($A^- = \{x, x(d) = -\}$) of a.

Compare a and b = compare the pairs (A^-, A^+) and (B^-, B^+)

A simplified framework

We assume that the set of criteria X can be partitioned into affects of various polarities

$$X^+ \text{ (positive args)}, X^- \text{ (negative args) and } X^0 \text{ (indifferent args)}$$

- $x \in X^+ \iff \forall d \in D, x(d) \in \{0, +\}$ (positive affects)
- $x \in X^- \iff \forall d \in D, x(d) \in \{0, -\}$ (negative affects)

So each affect (attribute) $x \in X$ is Boolean (presence vs. absence) but has

- **a polarity**: the presence of x is either good or bad, its absence is neutral
- **an importance** $\pi(x) \in L$ with top 1_L (full importance) and bottom 0_L (no importance). For $x \in X^0, \pi(x) = 0_L$.

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.

$$\Pi(A) = \max_{x \in A} \pi(x)$$

A decision rule : The Bipolar Possibility Relation

Principle at work: *Comparability of affects across polarities*

When comparing a and b, any argument against a (resp. against b) is an argument pro b (resp. pro a).

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

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

- $\succeq^{Biposs}$ is complete, but only its strict part is transitive.
- $A \succeq^{Pareto} B \Rightarrow A \succeq^{Biposs} B$
- Both relations collapse to the maximin rule if all arguments are negative and to the maximax rule if all arguments are positive.

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

The full lexi-bipolar rule

A very discriminant procedure, counting and cancelling conflicts:

Definition

$$A \succeq^{Lexi} B \iff \exists i \in L \text{ such that } \begin{cases} \forall j > i, & |A_j^+| + |B_j^-| = |A_j^-| + |B_j^+| \\ \text{and} & |A_i^+| + |B_i^-| > |A_i^-| + |B_i^+| \end{cases}$$

- A complete and transitive refinement of $\succeq^{Biposs}$
- In agreement with Cumulative Prospect Theory: There exist two capacities σ^+ and σ^- such that:

$$A \succeq^{Lexi} B \Leftrightarrow \sigma^+(A^+) - \sigma^-(A^-) \geq \sigma^+(B^+) - \sigma^-(B^-)$$

Where do we stand on bipolar criteria ?

- These criteria are axiomatised (article JAIR, 2008)
- Empirically tested on human subjects with J.-F. Bonnefon (article in Theory and Decision, 2008).
- The experimental study indicates people frequently use the lexicographic bipolar decision rule.
- Perspective: extend the approach to non-Boolean qualitative bipolar criteria with more-general importance weights (bipolar Sugeno integrals)

Perspectives

1. Refinement of Sugeno integral : two possible refinements to be unified
2. Axiomatisation of refined max-min criteria.
3. Non-commensurate decision rules : approaches without the assumption of independence of irrelevant alternative (Perny...)
4. Variants of Sugeno integral (using residuated structures)
5. Extension of bipolar criteria to Sugeno integral
6. Identification of qualitative set-functions (with Prade, Rico, Couceiro)
7. Qualitative conditional criteria: toward dynamic decision evaluation

References

D. Dubois, H. Prade, M. Roubens, R. Sabbadin, J.-L. Marichal. The use of the discrete Sugeno integral in decision-making: a survey. Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems, 9, 539-561, 2001.

D. Dubois, H. Prade. R. Sabbadin. Decision theoretic foundations of qualitative possibility theory. European Journal of Operational Research, 128, 459-478, 2001.

D. Dubois, H. Fargier, and P. Perny. Qualitative models for decision under uncertainty: an axiomatic approach. Artificial Intelligence, 148:219-260, 2003 (Corrigendum 171:361-362, 2007).

D. Dubois, H. Fargier, P. Perny, and H. Prade. A characterization of generalized concordance rules in multicriteria decision making. Int. J. of Intell.Systems, 18:751-774, 2003.

H. Fargier and R. Sabbadin. Qualitative decision under uncertainty: Back to expected utility. Artificial Intelligence, 164:245-280, 2005.

J.-F. Bonnefon, D. Dubois, H. Fargier, S. Leblois, Qualitative heuristics for balancing the pros and the cons, Theory and Decision, 65:71-95 (2008).

D. Dubois, H. Fargier, J.-F. Bonnefon. On the qualitative comparison of decisions having positive and negative features. J. AI Research, 32, 385-417, 2008.

D. Dubois, H. Fargier, , Making Discrete Sugeno Integrals More Discriminant, International Journal of Approximate Reasoning, 50, 2009, 880-898

D. Dubois, H. Fargier, H. Prade, R. Sabbadin. A survey of qualitative decision rules under uncertainty. Decision-making Process- Concepts and Methods. D. Bouyssou, et al. (Eds.), ISTE London & Wiley, Chap. 11, p. 435-473, 2009.

M. Couceiro, D. Dubois, H. Prade, T. Waldhauser. Decision-Making with Sugeno Integrals - Bridging the Gap Between Multicriteria Evaluation and Decision Under Uncertainty. Order 33(3): 517-535 (2016)

D. Dubois, H. Prade, A. Rico: Residuated variants of Sugeno integrals: Towards new weighting schemes for qualitative aggregation methods. Inf. Sci. 329: 765-781 (2016)

D. Dubois, A. Rico, New axiomatisations of discrete quantitative and qualitative possibilistic integrals, To appear in Fuzzy Sets and Systems