Eliciting Hybrid Probability-Possibility Functions and Their Decision Evaluation Models

Didier Dubois DUBOIS@IRIT.FR
Romain Guillaume ROMAIN.GUILLAUME@IRIT.FR
IRIT-CNRS, Université Jean Jaurès, Toulouse, France

Agnès Rico AGNES.RICO@UNIV-LYON1.FR
Eric, Université de Lyon, France

Abstract

We focus on a decision tree model under uncertainty using so-called hybrid probability-possibility functions. They allow to handle behaviours lying between possibilistic decision making and probabilistic decision making while keeping the good properties of both approaches namely *Dynamic Consistency*, *Consequentialism* and *Tree Reduction*. We shed light on the various utility functionals in this setting. More precisely, in this paper, we investigate the question of parameterizing the compromise between possibilistic and probabilisic models in different contexts. To this end, we outline elicitation methods.

Keywords: decision under uncertainty, possibility theory, decomposable measures.

1. Introduction

In sequential decision making, a strategy is a conditional plan that assigns an action to each state where a decision has to be made. Each strategy leads to a compound lottery following Von Neumann and Morgenstern's terminology [8]. A tree represents the different scenarios. The optimal strategy is the one that minimizes a criterion whose value depends on utilities of final states and on the resulting compound lottery.

Three assumptions are instrumental to enable an optimal strategy to be computed using dynamic programming [6]:

- *Dynamic Consistency*: when following an optimal strategy and reaching a decision node, the best decision at this node is the one that had been considered so when computing this strategy, i.e. prior to applying it.
- *Consequentialism*: the best decision at each step of the problem only depends on potential consequences at this point.
- *Tree Reduction*: a compound lottery is equivalent to a simple one, assigning probabilities to final states.

The expected utility of probabilistic simple lotteries was proposed by Von Neuman and Morgernstern [8] as a decision criterion under risk. Dubois and Prade [1] proposed to use optimistic and pessimistic possibilistic criteria to evaluate the global utility of a possibilistic lottery, thus generalizing Wald maximax and maximin criteria. More recently a new hybrid decision criteria, subsuming expected utility and possibilistic criteria, was presented in [4], based on a parameterized family of capacities completely defined by a distribution of weights on the state space [2].

The aim of this paper is to deepen the understanding of this hybrid decision model with a view to elicit it with given data. This model comes down to a convex combination of a possibility distribution and a probability distribution. First we show how to retrieve, in a generally unique way, both distributions from a lottery with weights in $[0, 1]$ whose sum is at least 1. Next we want to elicit the model from a data set composed with a possibility, a probability and a decision evaluation given by an expert. In order to do this we start with a technical but useful result simplifying the expression of the hybrid criteria used for the evaluation. This equivalent expression is just based on possibility, probability and a real-valued parameter present in the hybrid model.

The paper is organised as follows. The following section presents the background and notations. Section 3 is devoted to the definition of the hybrid prob-poss capacity and its decomposition. Section 4 shows that any weight distribution whose sum is at least 1 can be interpreted as a hybrid prob-poss set function, thus providing an elicitation method for the underlying probability and possibility measures and their mixture coefficient. The aim of Section 5, which is technical, is to propose simpler equivalent expressions of criteria for the hybrid model. Section 6 outlines an elicitation method of the mixture coefficient from the knowledge of the assessed worth of prob-poss lotteries.

2. Background on Sequential Decision Problems

Sequential decision problems under uncertainty are usually modelled by decision trees [6] that rely on the following graphical model. Formally, a decision tree is a tree containing three kinds of nodes: (see Figure 1 for an example):

- A set of decision nodes (depicted by rectangles). Edges from such nodes represent decisions among which to choose.
- A set of chance nodes (depicted by circles) reached by edges stemming from decision nodes. Edges from such nodes are attached probabilities (or possibilities) and may lead to further decision nodes or to terminal nodes.
- The set $\mathcal{S}$ of terminal nodes or leaves represent states of nature; such states evaluated by a utility function: $\forall s_i \in \mathcal{S}$, $u(s_i)$ is the degree of satisfaction of eventually reaching state s_i. For the sake of simplicity we assume, without loss of generality, that only terminal nodes are attached utilities. The utility degrees belong to a totally ordered scale. The scale $[0, 1]$ is assumed in this paper.

In a decision tree, the children of any decision node form the set of chance nodes that can be reached by choosing a decision at this node. The children of any chance node form the set of possible outcomes of the previously selected decision - either a terminal node is observed (a state), or another decision node is reached (and then a new action should be chosen).

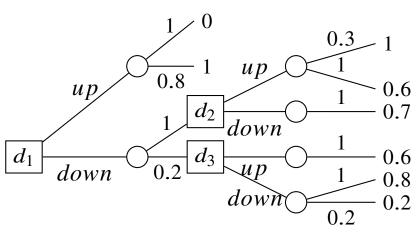

Figure 1: A possibilistic decision tree.

Solving a decision tree amounts to building a *strategy*, i.e. a function δ that associates to each decision node d_i an action (i.e., a child chance node): this is the action to be executed when decision node d_i is reached. Each possible scenario (sequence of decisions) is modelled by a path from the root to a terminal node of the tree.

The information at each chance node is fully captured by a distribution (probability or possibility) over the outcomes of the chance nodes, namely a possibility distribution π and/or a probability distribution p. When bearing on the leaf nodes, such distributions define simple lotteries on the utility degrees of states. More formally:

- A simple probabilistic lottery L^p [8] is a probability distribution p on a set of utility degrees, $\Lambda = \{\lambda_1, ..., \lambda_n\}$, where $\lambda_i = u(s_i)$. The probabilistic lotteries will be written as $L^p = (p_1/\lambda_1, ..., p_n/\lambda_n)$, with $p_i \in [0, 1], \sum_{i=1}^{n} p_i = 1$.
- A simple possibilistic lottery L^π [1] is a normalized possibility distribution π on a set of utility degrees, Λ, both being expressed in the same ordered scale. The possibilistic lotteries will be written as $L^p = (\pi_1/\lambda_1, ..., \pi_n/\lambda_n)$ with $\pi_i \in [0, 1], \max_{i=1}^{n} \pi_i = 1$.

In a simple lottery L^π (resp. L^p), the value π_i (resp. p_i) is the possibility (resp. probability) degree of reaching a state with utility λ_i. For the sake of brevity, the λ_i's such that $\pi_i = 0$ (resp. $p_i = 0$) are often omitted in the notation of a lottery (e.g., $< 1/0.8 >$ denotes the lottery that provides utility 0.8 for sure, all the other utility degrees being impossible).

A given strategy leads to a composite lottery, i.e., a nested lottery in the form of an uncertainty (probability or possibility) tree in agreement with the decision strategy selected in the decision tree. In order to respect the three properties of dynamic consistency, consequentialism, and tree reduction, the compound lottery should be equivalent to a simple one. For instance in a probability tree, the probability of a state is the sum of the probabilities of all scenarios leading to this state, the probability of a scenario being the product of all probabilities attached to the edges of the path. In the case of a possibility tree, the same property applies, changing sum into maximum and product into a more general triangular norm (e.g., minimum). This reduction property enables to compute the utility of a strategy using dynamic programming on the decision tree structure.

In the next section we recall a class of set functions introduced in [2] that generalizes both probability and possibility functions and lead to generalized forms of simple and compound lotteries, while preserving the tree reduction property allowing the equivalence between compound lotteries and simple ones, and the use of dynamic programming to compute the generalized utility of strategy. As shown in [2], these set functions are the only ones allowing the reduction of uncertainty trees to simple (generalized) lotteries.

3. Hybrid Possibility/Probability Measures

We consider so-called hybrid π-p measures that combine probabilistic and possibilistic behaviors in the uncertainty context. More precisely we will use convex combinations of possibility and probability distributions:

$$\rho^\alpha(s) = \alpha\pi(s) + (1 - \alpha)p(s), \quad \alpha \in [0, 1]$$

where p and π satisfy the constraint $p(s) = 0$ if $\pi(s) < 1$ for all s (see [4] for more details). Clearly ρ^α is a possibility distribution if $\alpha = 1$ and a probability distribution if $\alpha = 0$. As a consequence, note that $1 \le \sum_{s\in S} \rho^\alpha(s) \le n$.

Intuitively, the decision-maker defines probabilities over the fully possible states of nature ($\pi(s) = 1$), while more or less impossible ones are taken into account ($\pi(s) < 1$). It is worth noticing that

- $\rho^\alpha(s) > \alpha$ is equivalent to $p(s) > 0$ and these conditions imply $\pi(s) = 1$.
- $\rho^\alpha(s) = \alpha$ is equivalent to $\pi(s) = 1$ and $p(s) = 0$.
- $\rho^\alpha(s) < \alpha$ is equivalent to $\pi(s) < 1$ and $p(s) = 0$.

It is then clear that we have

- $\rho^\alpha(s) = \alpha\pi(s)$ if $\rho^\alpha(s) \le \alpha$ and 1 otherwise,
- $\rho^\alpha(s) = \alpha + (1-\alpha)p(s)$ if $\rho^\alpha(s) > \alpha$ and 0 otherwise.

For instance, consider a coin. Usually the result of flipping it is either head (h) or tail (t). But a very rare, yet not fully impossible, occurrence is that the coin falls on its edge (e). So the state of affairs is $S = \{h, t, e\}$. It is natural to consider h, t as totally possible events, and e almost impossible. So $\pi(t) = \pi(h) = 1 > \pi(e) = \epsilon$. Moreover, if the coin is fair then $p(t) = p(h) = 0.5$, and $p(e) = 0$, if we consider the probability of "edge" negligible.

Hybrid distributions generate a class of decomposable capacities, which are monotonic set functions such that there is a triangular conorm of the form

$$S^\alpha(x, y) = \begin{cases} \min(1, x + y - \alpha) \text{ if } x > \alpha, y > \alpha \\ \max(x, y) \text{ otherwise,} \end{cases}$$

such that, as explained in [2], if $A \cap B = \emptyset$:

$$\rho^\alpha(A \cup B) = S^\alpha(\rho^\alpha(A), \rho^\alpha(B)) \quad (1)$$

They are actually Shafer plausibility functions [7] of the form $\rho^\alpha(A) = \alpha\Pi(A) + (1-\alpha)P(A)$, where $\Pi(A) = \max_{s\in A} \pi(s)$ is a possibility measure. It is an interesting family of plausibility measures with very low complexity since completely defined by a distribution of weights over S. Note that $S^\alpha(x, y) = \max(x, y, \min(1, x + y - \alpha))$. Indeed, if $x > \alpha, y > \alpha$ then $x + y - \alpha > \max(x, y)$. If $x \le \alpha$, then $x + y - \alpha \le y$.

We can define the conjugate of ρ^α as $\overline{\rho}^\alpha$ defined by $\overline{\rho}^\alpha(A) = 1 - \rho^\alpha(\overline{A}) = 1 - S^\alpha_{s\in\overline{A}}(\rho^\alpha(s)) = T^\alpha_{s\in\overline{A}}(1 - \rho^\alpha(s))$ where $T^\alpha(a, b) = 1 - S^\alpha(1 - a, 1 - b)$. The reader can check, letting $N(A) = 1 - \Pi(\overline{A})$ (a necessity measure), that $\overline{\rho}^\alpha(A) = \alpha N(A) + (1-\alpha)P(A)$, which is a belief function.

Example 1 *On $S = \{s_1, s_2, s_3\}$, let $\pi(s_1) = 0.4$, $\pi(s_2) = \pi(s_3) = 1$, $p(s_2) = 0.6$, $p(s_3) = 0.4$ and $\alpha = 0.5$. We can build the $\rho^{0.5}$ distribution on S from the π and p distributions: $\rho^{0.5}(\{s_1\}) = 0.2$, $\rho^{0.5}(\{s_2\}) = 0.8$ and $\rho^{0.5}(\{s_3\}) = 0.7$. We then have that $\rho^{0.5}(\{s_1, s_2\}) =$* $\max(0.2, 0.8) = 0.8$*,* $\rho^{0.5}(\{s_1, s_3\}) = \max(0.2, 0.7) = 0.7$*,* $\rho^{0.5}(\{s_2, s_3\}) = 0.8 + 0.7 - 0.5 = 1$*. This distribution defines a convex set of probability distributions. It contains the convex combinations with weight α of the probability measures, P', compatible with possibility distribution π, i.e., $P' \le \Pi$, and the probability measure P with distribution p. We can express this probability set by inequalities: $P(\{s_1, s_2, s_3\}) = 1$, $0 \le P(\{s_1\}) \le 0.2$, $0.3 \le P(\{s_2\}) \le 0.8$, $0.2 \le P(\{s_2\}) \le 0.7$, $0.3 \le P(\{s_1, s_2\}) \le 0.8$, $0.2 \le P(\{s_1, s_3\}) \le 0.7$ and $0.3 \le P(\{s_2, s_3\}) \le 1$.*

One interesting feature of this class of fuzzy measures is that it can be expressed as uncertainty trees generalizing probability trees involving lotteries, that can be reduced to distributions over the set of states of affairs, as shown in [2]. However, in order to reduce probability-possibility lotteries, an operation $*$ is needed to generalize probabilistic independence, in such a way that if A and B are disjoint sets independent of another set C, so must be $A \cup B$, i.e., we have that

$$\begin{aligned}\rho^\alpha((A \cup B) \cap C) &= S^\alpha(\rho^\alpha(A), \rho^\alpha(B)) * \rho^\alpha(C) \\ &= S^\alpha(\rho^\alpha(A) * \rho^\alpha(C), \rho^\alpha(B) * \rho^\alpha(C)).\end{aligned}$$

This distributivity property is valid only when the operation $*$ is a triangular norm of the form

$$x *_\alpha y = \begin{cases} \alpha + \frac{(x-\alpha)(y-\alpha)}{1-\alpha} \text{ if } x > \alpha, y > \alpha \\ \min(x, y) \text{ otherwise.} \end{cases}$$

provided that $S^\alpha(\rho^\alpha(A), \rho^\alpha(B)) < 1$ [5, 2]. The latter condition implies that the upper bound 1 is never trespassed in expressions of the form $\min(1, x+y-\alpha)$, so that the useful part of the conorm is of the form $S^\alpha(x, y) = \max(x, y, x + y - \alpha)$.

Denote by C^+_α the set $\{s : \rho^\alpha(s) > \alpha\} = \{s : p(s) > 0\}$. If $C^+_\alpha \ne \emptyset$, the normalization condition $\rho^\alpha(S) = 1$ enforces the condition:

$$\sum_{s\in C^+_\alpha} \rho^\alpha(s) - \alpha(card(C^+_\alpha) - 1) = 1. \quad (2)$$

Note that it can be more simply written as

$$\sum_{s\in S} \max(0, \rho^\alpha(s) - \alpha) = 1 - \alpha.$$

For instance, the distribution $\rho^{0.5}$ in Ex. 1 satisfies $0.7 - 0.5 + 0.8 - 0.5 = 1 - 0.5$, involving only s_2, s_3. If $\alpha = 0$ (ρ^0 is a probability measure), the normalization condition (2) reads $\sum_{s\in S} \rho^0(s) = 1$. If $\alpha = 1$, (ρ^1 is a possibility measure), the normalization condition $\rho^\alpha(S) = 1$ reads

$\max_{s\in S} \rho^1(s) = 1$. Note that any set-function of the form $g = \alpha\Pi + (1-\alpha)P$ for any possibility measure Π and probability measure P is a Shafer plausibility function as well, without requiring the condition $p(s) > 0$ implies $\pi(s) = 1$. The latter condition is due to the tree reduction property which enforces $g = \rho^\alpha$, a decomposable set-function with respect to a conorm S^α, with an independence operator $*_\alpha$. Only such convex mixtures of probability and possibility functions allow for the reduction of compound lotteries into simple ones, as explained in [2, 4].

4. Elicitation of a Prob-Poss Model from Given Weights

In this section we propose to use the hybrid model to interpret a distribution of weights $\rho = (\rho_1, \dots \rho_n) \in [0,1]^n$ on S with $n \geq \sum_{i\in[n]} \rho_i \geq 1$. Let us consider the case where a distribution ρ is given but it satisfies neither the normalisation condition of probability distributions nor the one for possibility distributions.

When representing a distribution of uncertainty, two usual normalization options could be investigated: normalize ρ to build a probability distribution (dividing the weights by their sum), normalize ρ to make a possibility distribution (dividing the weights by their maximum). This paper suggests that, alternatively, we can keep the weight distribution ρ and interpret it as hybrid prob-poss distribution. Namely, we show that for any such distribution of weights, there generally exists a unique parameter value α, a unique possibility distribution π and a unique probability distribution p such that $\rho_i = \alpha\pi_i + (1-\alpha)p_i$.

The latter approach looks more natural. For instance, the probabilistic and possibilistic renormalizations of two different uniform ρ and ρ' yield equal probability distributions $p = p'$ and possibility distributions $\pi = \pi'$. But their decompositions will be different.

Example 2 *Consider two distributions on $S = \{a, b\}$*

- *1: $\rho_a = \rho_b = 0.6$*
- *2: $\rho'_a = \rho'_b = 0.5$*

We can see that renormalizing these distributions in agreement with possibility or probability, the resulting two distributions 1 and 2 are the same ($p_a^1 = p_a^2 = 0.5 = p_b'^1 = p_b'^2$ and $\pi_a'^1 = \pi_a'^2 = 1 = \pi_b'^1 = \pi_b'^2$), hence no distinction between 1 and 2 can be made using this kind of transformation. Using the hybrid interpretation we easily see that:

- *Case 1: with $\alpha = 0.2$, $\pi_a = \pi_b = 1$, and $p_a = p_b = 0.5$ we can check that $\rho_a = \rho_b = 0.2 + 0.8 \cdot 0.5$, (a mixture between uniform probabilities and possibilities).*
- *Case 2: $\alpha = 0$, $\pi_a = \pi_b = 1$, and $p_a = p_b = 0.5 = \rho_a = \rho_b$, (a pure probability distribution).*

Formally, the question is: given a distribution of weights $(\rho_1, \dots \rho_n) \in [0,1]^n$ on S such that $\sum_{i=1}^n \rho_i \geq 1$, does there exist a threshold $\alpha \in [0,1]$, a possibility distribution π and a probability distribution p on S, such that $\rho = \alpha\pi + (1-\alpha)p$? If yes, is the 3-tuple (α, π, p) uniquely defined?

Let us consider n weights with m distinct ones such that $\rho_{(m)} < \dots < \rho_{(1)}$ with $m \leq n$ and $R_{(i)} = \{j | \rho_j = \rho_{(i)}\}$. Note that if we suppose $\alpha \in [\rho_{(i+1)}, \rho_{(i)})$ Then $\{s_i : \rho_i > \alpha\} = \cup_{j=1}^i R_{(j)}$, with cardinality $\sum_{j=1}^i |R_{(j)}|$. The normalization condition (2) reads

$$\sum_{j=1}^{i} |R_{(j)}| \rho_{(j)} = 1 + \alpha(\sum_{j=1}^{i} |R_{(j)}| - 1).$$

We can see that

- if $|\{i | \rho_i > 0\}| = 1$, i.e., $\rho_k = 1, \rho_i = 0$ if $i \neq k$, then α is not unique: $\forall\alpha \in [0,1]$, $\rho = \alpha\pi + (1-\alpha)p$ where $\pi_k = 1, \pi_i = 0$ if $i \neq k$, $p_k = 1, p_i = 0$ if $i \neq k$. So ρ^α is the convex mixture of a possibility measure and a Dirac measure focused on s_k.
- if $|\{i | \rho_i > 0\}| > 1$, we must find α such that the normalization condition (2) holds. It is clear that based on this condition and, if $\rho = \alpha\pi + (1-\alpha)p$, we have

  $$\alpha = \frac{\sum_{j=1}^{i} |R_{(j)}| \rho_{(j)} - 1}{\sum_{j=1}^{i} |R_{(j)}| - 1} \in [\rho_{(i+1)}, \rho_{(i)}). \quad (3)$$

  Note that if $|R_{(1)}| = 1$ and $\rho_{(1)} < 1$, we cannot choose $\alpha \in [\rho_{(2)}, \rho_{(1)}]$ since we cannot make the division in (3) and the normalisation equation reads $\rho_{(1)} - \alpha = 1 - \alpha$, which is impossible.

So given the weight distribution only, with $\rho_{(1)} < 1$, the associated parameter α computed as in (3) must belong to some interval $[\rho_{(i+1)}, \rho_{(i)})$. We can prove that such a parameter value exists and is generally unique, as shown now. Note that if $\sum_{i=1}^n \rho_i = 1$, we have a probability distribution and $\alpha = 0$ is enforced, while π is arbitrary. So we focus on the case when $\sum_{i=1}^n \rho_i > 1$.

Proposition 1 *For all n-tuples $\rho \in [0,1]^n$, such that $\sum_{i=1}^n \rho_i > 1$:*

- *if $\rho_{(1)} \neq 1$, there exists a unique value α and a unique index i, $1 \leq i \leq m$ such that: $\alpha = \frac{\sum_{j=1}^{i} |R_{(j)}| \rho_{(j)} - 1}{\sum_{j=1}^{i} |R_{(j)}| - 1} \in [\rho_{(i+1)}, \rho_{(i)})$;*
- *if $\rho_{(1)} = 1$ and $|R_{(1)}| = 1$, then any $\alpha > \rho_{(2)}$ can be chosen.*
- *If $\rho_{(1)} = 1$ and $R_{(1)}$ contains several elements then $\alpha = 1$.*

Proof We need to distinguish two cases:

- $\rho_{(1)} < 1$. The normalization condition (2) also writes $\sum_{i=1}^{n} \max(0, \rho_i - \alpha) = 1 - \alpha$. It can thus be expressed as the intersection points of the two functions $h(\alpha) = 1-\alpha$ and $f(\alpha) = \sum_{i=1}^{n} \max(0, \rho_i - \alpha)$ on interval $[0, 1]$. Notice that h and f are both continuous. h is linear decreasing on $[0, 1]$ with slope -1, $h(1) = 0, h(0) = 1$; f is piecewise linear decreasing on $[0, \rho_{(1)}]$ and $f(\alpha) = 0$ on $]\rho_{(1)}, 1]$ and $f(\alpha) = \sum_{i=1}^{n} \rho_i > 1 = h(0)$ on $[0, \rho_{(m)}]$. So, $\exists \alpha^* \in (\rho_{(m)}, \rho_{(1)}]$ such that $f(\alpha^*) = h(\alpha^*)$ and this value is unique due to the shape and positions of these functions. So there is a unique index i such that $\alpha^* \in [\rho_{(i+1)}, \rho_{(i)})$ of the form indicated in (3).
- Case $\rho_{(1)} = 1$ and $R_{(1)} = \{i^*\}$: In this case, we can assume $\alpha = 1$ since the normalization condition reads $h(1) = f(1) = 0$. The capacity is then a possibility measure $\rho = \pi$. Choosing $\alpha < \rho_{(2)}$ is impossible because $f(\alpha) \geq \rho_2 - \alpha + 1 - \alpha > 1 - \alpha$. If there is a single i^* such that $\rho_{i^*} = 1$, choosing any α with $\rho_{(2)} \leq \alpha < 1$ is possible since $f(\alpha) = 1 - \alpha$, and the decomposable capacity has distribution $\rho = \alpha\pi + (1 - \alpha)\delta_{i^*}$, where δ_i is a Dirac function on i^*.
- If there are more than one $\rho_i = 1$, only $\alpha = 1$ is possible since $f(\alpha) \geq 2(1 - \alpha)$ if $\alpha < 1$, and $f(1) = h(1)$.

■

We can summarize the results in this section as follows:

Theorem 2 *Let n weights $\rho_j \in [0, 1], j = 1, \ldots, n$ such that $\sum_{j=1}^{n} \rho_j \geq 1$, with $\max_{j=1}^{n} \rho_j < 1$. There exists a unique value α, an integer i_0 such that $\rho_{(i_0+1)} \leq \alpha < \rho_{(i_0)}$, a unique possibility distribution π and a unique probability distribution p such that $\rho = \alpha\pi + (1 - \alpha)p$ with*

- $\forall i \leq i_0 + 1 \; \forall j \in R_{(i)}, \pi_j = \frac{\rho_j}{\alpha}$ *and* $p_j = 0$.
- $\forall i \geq i_0 \; \forall j \in R_{(i)}, p_j = \frac{\rho_j - \alpha}{1-\alpha}$ *and* $\pi_j = 1$.

Proof We just need to prove that $\sum_{i=1}^{n} p_i = 1$.
$\sum_{i=1}^{n} p_i = \sum_{i:\rho_i > \alpha} (\frac{\rho_i}{1-\alpha} - \frac{\alpha}{1-\alpha}) = \frac{1+\alpha(|C_\alpha^+|-1)}{1-\alpha} - \frac{\alpha|C_\alpha^+|}{1-\alpha} = 1$.
■

Note that if $\max_{j=1}^{n} \rho_j = \rho_{(1)} = 1$, and $R_{(1)} = \{i^*\}$ any $\alpha \in [\rho_{(2)}, 1]$ can be chosen and we have that $\rho = \alpha\pi + (1-\alpha)p$ with $p_{i*} = 1$ and $\pi_j = \rho_j/\alpha, j \neq i^*$.

Example 3 *Consider the distribution of weights (0.5, 0.5, 0.8, 0.9). So we have $\rho_{(1)} = 0.9, \rho_{(2)} = 0.8, \rho_{(3)} = 0.5$ and $|R_{(1)}| = |R_{(2)}| = 1$ and $|R_{(3)}| = 2$. We are in the case where the value of α is one such that $\alpha = \frac{\sum_{j=1}^{i} |R_{(j)}|\rho_{(j)} - 1}{\sum_{j=1}^{i} |R_{(j)}|} \in [\rho_{(i+1)}, \rho_{(i)})$.*

- *Suppose $i = 1, \alpha \in [0.8, 0.9)$: the normalization condition reads $0.9 - \alpha = 1 - \alpha$; there is no solution.*
- *$i = 2, \alpha \in [0.5, 0.8)$: the normalization condition reads $0.9 - \alpha + 0.8 - \alpha = 1 - \alpha$; hence $\alpha = 0.7 \in [0.5, 0.8)$, which is true.*

So $\alpha = 0.7$ is the solution. So, $\pi_1 = \pi_2 = 0.5/0.7; \pi_3 = \pi_4 = 1$ and $0.8 = 0.7 + 0.3p_3$ so $p_3 = 1/3$, $0.9 = 0.7 + 0.3p_4$, so $p_4 = 2/3$.

5. Hybrid Prob-Poss Utility Functionals

In this section we show how utility functionals defined for decomposable capacities in the ρ^α family can be expressed in terms of the possibility distribution and the probability distribution that underlie ρ^α.

5.1. The Utility of Generalized Lotteries

Let us recall the new decision criteria, beyond expected utility and possibilistic integrals based on a hybrid probability-possibility function. Let us consider generalized lotteries $L^{\rho^\alpha} =< \rho_1^\alpha/\lambda_1, ..., \rho_n^\alpha/\lambda_n >$. Two utility functionals $ES^{Opt}(L^{\rho^\alpha})$ and $ES^{Pes}(L^{\rho^\alpha})$, respectively optimistic and pessimistic, are defined in [2] and explicited in [4]:

$$ES^{Opt}(L^{\rho^\alpha}) = S^\alpha_{i=1,\ldots,n} \rho_i^\alpha *_\alpha \lambda_i \tag{4}$$
$$= \begin{cases} \alpha + \frac{\sum_{i|\lambda_i > \alpha, \rho_i^\alpha > \alpha} (\lambda_i - \alpha)(\rho_i^\alpha - \alpha)}{1-\alpha} & \text{if } \exists i : \lambda_i > \alpha \text{ and } \rho_i^\alpha > \alpha \\ U^{Opt}(L^{\rho^\alpha}) \text{ otherwise} \end{cases}$$

$$ES^{Pes}(L^{\rho^\alpha}) = 1 - S^\alpha_{i=1,\ldots,n} \; \rho_i^\alpha *_\alpha (1 - \lambda_i) \tag{5}$$
$$= \begin{cases} 1 - \alpha - \frac{\sum_{i|1-\lambda_i < \alpha, \rho_i^\alpha > \alpha} (1 - \lambda_i - \alpha)(\rho_i^\alpha - \alpha)}{1-\alpha} \\ \text{ if } \exists i : \lambda_i < 1 - \alpha \text{ and } \rho_i^\alpha > \alpha \\ U^{Pes}(L^{\rho^\alpha}) \text{ otherwise.} \end{cases}$$

where U^{Pes} and U^{Opt} are respectively optimistic and pessimistic possibilistic utility functionals proposed by Dubois and Prade [1] of the form:

$$U^{Opt}(L^{\rho^\alpha}) = \max_{\lambda_i \in \Lambda} \min(\rho_i^\alpha, \lambda_i).$$

$$U^{Pes}(L^{\rho^\alpha}) = \min_{\lambda_i \in \Lambda} \max(1 - \rho_i^\alpha, \lambda_i).$$

Note that ES^{Opt} and ES^{Pes} generalize these optimistic and pessimistic possibility criteria, replacing max by S^α and min by $*_\alpha$. Moreover when $\alpha = 1$, then $ES^{Opt} = U^{Opt}$, and $ES^{Pes} = U^{Pes}$. When $\alpha = 0$, $ES^{Opt} = ES^{Pes}$ is the standard expected utility of the lottery since ρ^0 is a probability distribution.

The new hybrid possibilistic-probabilistic generalized criteria, $ES^{Opt}(L^{\rho^\alpha})$ and $ES^{Pes}(L^{\rho^\alpha})$ are very appealing because they can be conveniently applied to decision trees. Indeed, they respect the three major properties needed to that effect: (*Dynamic Consistency, Consequentialism* and *Tree Reduction*). See details in [4]. Let us recall two useful properties proved in that paper:

Considering a lottery $L^{\rho} =< \rho_1/\lambda_1, ..., \rho_n/\lambda_n >$ we define another lottery $(1-L)^{\rho}$, with utility scale upside down, by $(1-L)^{\rho} =< \rho_1/(1-\lambda_1), ..., \rho_n/(1-\lambda_n) >$. We have the following De Morgan-like duality relation

$$ES^{Pes}(L^{\rho^{\alpha}}) = 1 - ES^{Opt}(1 - L^{\rho^{\alpha}}). \quad (6)$$

Moreover the parameter α delimits zones where the pessimistic or the optimistic criteria apply:

Proposition 3 *[4] $ES^{Opt}(L^{\rho^{\alpha}}) \leq \alpha$ iff $ES^{Opt}(L^{\rho^{\alpha}}) = U^{Opt}(L^{\rho^{\alpha}})$ and $\nexists i$ s.t. $\lambda_i > \alpha$ with $p_i > 0$.*
Likewise, $ES^{Pes}(L^{\rho^{\alpha}}) \geq 1-\alpha$ iff $ES^{Pes}(L^{\rho^{\alpha}}) = U^{Pes}(L^{\rho^{\alpha}})$ and $\exists i$ s.t. $\lambda_i > \alpha$ with $p_i > 0$.

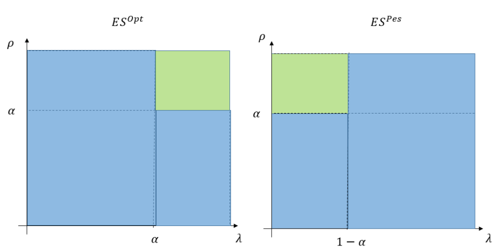

Figure 2: $ES^{Opt}(L^{\rho^{\alpha}})$ and $ES^{Pes}(L^{\rho^{\alpha}})$

In other words, as explained in [4], the criterion $ES^{Opt}(L^{\rho^{\alpha}})$ is optimistic and possibilistic (= $U^{Opt}(L^{\rho^{\alpha}})$) so long as entries (utilities or plausibilities) are below the threshold α (distribution included in blue area on Figure 2). Otherwise, we get a re-scaled expected value over states with plausibilities and utilities greater than α (see green area in Figure 2, left).

Likewise, with $ES^{Pes}(L^{\rho^{\alpha}})$, we get a re-scaled expected value over states with utility less than $1-\alpha$ and with high enough plausibility i.e. greater than α. We get the pessimistic possibilistic criterion $U^{Pes}(L^{\rho^{\alpha}})$ otherwise (with either high utilities or low plausibilities); see the green area in Figure 2 right side.

5.2. Improving the Expression of the Generalized Utility Functionals

This section is devoted to new technical results in order to obtain expressions for $ES^{Pes}(L^{\rho^{\alpha}})$ and $ES^{Opt}(L^{\rho^{\alpha}})$ that explicitly use the possibility and probability distributions underlying a distribution ρ. We have shown in the previous section that (α, π, p) are generally unique such that $\rho = \alpha\pi + (1-\alpha)p$ and we provided explicit expressions for (α, π, p). Let $C^+_{\alpha} = \{i : \rho^{\alpha}_i > \alpha\} = \{i : p_i > 0\}$ and $\Lambda^+_{\alpha} = \{i : \lambda_i > \alpha\}$. Note that $P(C^+_{\alpha} \cap \Lambda^+_{\alpha}) = P(\Lambda^+_{\alpha})$.

Proposition 4

$$ES^{Opt}(L^{\rho^{\alpha}}) = \begin{cases} \max_i \min(\alpha\pi_i, \lambda_i) \ \textit{if}\ C^+_{\alpha} \cap \Lambda^+_{\alpha} = \emptyset, \\ \alpha + \sum_{i \in \Lambda^+_{\alpha}} p_i(\lambda_i - \alpha) \ \textit{otherwise.} \end{cases} \quad (7)$$

Proof We can start from the expression $ES^{Opt}(L^{\rho^{\alpha}}) = S^{\alpha}_{i=1,...,n} \rho^{\alpha}_i *_{\alpha} \lambda_i$. Recall that $\rho^{\alpha}_i = \begin{cases} \alpha + (1-\alpha)p_i \text{ if } i \in C^+_{\alpha}, \\ \alpha\pi_i \text{ otherwise.} \end{cases}$.
Then it is clear that

- If $i \notin C^+_{\alpha} \cap \Lambda^+_{\alpha}$, then $\rho^{\alpha}_i *_{\alpha} \lambda_i = \min(\alpha\pi_i, \lambda_i)$;
- If $i \in C^+_{\alpha} \cap \Lambda^+_{\alpha}$, then $\rho^{\alpha}_i *_{\alpha} \lambda_i = \alpha + \frac{(\alpha+(1-\alpha)p_i-\alpha)(\lambda_i-\alpha)}{1-\alpha}$ $= \alpha + p_i(\lambda_i - \alpha)$.

It follows that if $C^+_{\alpha} \cap \Lambda^+_{\alpha} = \emptyset$, then $S^{\alpha} = \max$ and $ES^{Opt}(L^{\rho^{\alpha}}) = \max_i \min(\alpha\pi_i, \lambda_i)$. If $C^+_{\alpha} \cap \Lambda^+_{\alpha} \neq \emptyset$, then: $ES^{Opt}(L^{\rho^{\alpha}}) = \sum_{i \in C^+_{\alpha} \cap \Lambda^+_{\alpha}} \alpha + p_i(\lambda_i - \alpha) - \alpha(|C^+_{\alpha} \cap \Lambda^+_{\alpha}| - 1)$ $= \alpha + \sum_{i \in \Lambda^+_{\alpha}} p_i(\lambda_i - \alpha)$. ■

When there are $\lambda_i > \alpha$ with $p_i > 0$ (probable situations with sufficient utility), the optimistic hybrid utility function takes the form

$$ES^{Opt}(L^{\rho^{\alpha}}) = \alpha(1 - P(\Lambda^+_{\alpha})) + P(\Lambda^+_{\alpha})E[\lambda|\Lambda^+_{\alpha}]$$

where $E[\lambda|\Lambda^+_{\alpha}] = \sum_{i=1}^{n} \lambda_i P(\{i\}|\Lambda^+_{\alpha}) = \frac{\sum_{i \in \Lambda^+_{\alpha}} \lambda_i p_i}{P(\Lambda^+_{\alpha})}$. In other words, in this expression, the factor α reflects in an optimistic way the utility of the least satisfying situations. Indeed:
$\alpha + \sum_{i \in \Lambda^+_{\alpha}} p_i(\lambda_i - \alpha) = \alpha + \sum_{i \in \Lambda^+_{\alpha}} \lambda_i p_i - \alpha p_i$
$= \alpha - \sum_{i \in \Lambda^+_{\alpha}} \alpha p_i + \sum_{i \in \Lambda^+_{\alpha}} \lambda_i p_i = \alpha(1 - P(\Lambda^+_{\alpha})) + \sum_{i \in \Lambda^+_{\alpha}} \lambda_i p_i$

So $ES^{Opt}(L^{\rho^{\alpha}})$ is a weighted average between the optimistic evaluation of the least attractive situations, and the conditional utility of the most attractive ones, weighted by their respective probabilities.

Using semi duality relation (6), we obtain the following result for the pessimistic hybrid utility functional $ES^{Pes}(L^{\rho^{\alpha}})$, letting $\overline{\Lambda}_{\alpha} = \{i : 1 - \lambda_i > \alpha\}$.

Proposition 5 *If $\rho = \alpha\pi + (1-\alpha)p$, then*

$$ES^{Pes}(L^{\rho^{\alpha}}) = \begin{cases} \min_i \max(1 - \alpha\pi_i, \lambda_i) \ \textit{if}\ C^+_{\alpha} \cap \overline{\Lambda}_{\alpha} = \emptyset \\ 1 - \alpha - \sum_{i \in \Lambda^+_{\alpha}} p_i(1 - \lambda_i - \alpha) \ \textit{otherwise.} \end{cases} \quad (8)$$

Similarly to the optimistic case, when there exist probable situations with sufficiently poor utility values, the pessimistic hybrid utility functional takes the form

$$ES^{Pes}(L^{\rho^{\alpha}}) = (1-\alpha)(1 - P(\overline{\Lambda}_{\alpha})) + P(\overline{\Lambda}_{\alpha})E[\lambda|\overline{\Lambda}_{\alpha}].$$

We can see from Propositions 5 and 4 that when the possibilistic criteria are used, they only depend on the possibility distribution and the parameter α. And when the conditional expected utility is used it only depends on the probability distribution and on α. So α controls the switch from possibility-driven decision to probability-driven decision.

6. Elicitation from Global Ratings of Lotteries

In this section, we focus on a method to obtain the value of the parameter α of a hybrid prob-poss function in a sequential decision problem, knowing the probability and the possibility distribution, on the one hand, and the utility of the corresponding decision.

At each chance node x of the decision tree, a distribution ρ_x of weights is assumed over the decision or leaf nodes in $Succ(x)$. Since $\rho_x = \alpha\pi_x + (1-\alpha)p_x$, there is a subset $Succ^+(x)$ of probable successors ($p_x(d) > 0, \pi_x(d) = 1$) and a subset $Succ(x) \setminus Succ^+(x)$ of more or less impossible successors ($p_x(d) = 0, \pi_x(d) < 1$). So we have two nested decision trees for the same decision problem: one with a probability distribution on each set $Succ^+(x)$, one with possibility distribution on $Succ(x)$.

When a strategy is chosen, i.e, one decision for each decision node, we are left with a generalized probability tree, where probabilities are replaced by ρ^α coefficients. The properties of the hybrid prob-poss functions are such that such composite generalized lotteries can be reduced to a weight distribution on the leaves (final states) of the decision trees providing that all ρ^α distributions present at each chance node share the same value of α (see [3] for details). This is reasonable if all the data about the decision tree is provided by the same decision-maker. Comparing the utilities of strategies come down to comparing generalized utilities of ρ^α distributions on final states. We suppose that the decision-maker can provide probabilities and possibilities on the decision tree, and can also assess the utilities of the various strategies. However it seems difficult to ask for the value α.

We thus start from a dataset, each item of which is made of both distributions, namely the probability and the possibility ones and the utility evaluation given by decision maker. In this case we look for a unique value α across data items, that models the behaviour of the decision maker regarding the trade-off between possibilistic and probabilistic decision models.

The dataset is a set of tuple (π^j, p^j, β^j) $j \in J = \{1, \ldots, m\}$ where π^j is a possibility distribution, p^j is a probability distribution j is a strategy, and β^j is the global evaluation given by an expert.

We want to identify α such that $ES^{Opt}(L^{\rho^\alpha})$ represents the given dataset, i.e., $ES^{Opt}(L^{\rho^\alpha_j}) = \beta^j$. Without loss of generality, we assume that, for utilities of states, we have $\lambda_i < \lambda_{i+1}, \forall i \in [n-1]$. Before giving the results on the feasibility of determining α, we give some facts on the link between ES^{Opt}, α and β^j.

Lemma 6 *Consider a lottery L^{ρ^α} with $\rho^\alpha = (p, \pi, \alpha)$. If $\exists ! i^*$ $p_{i^*} = 1$ then*

- *for all $\alpha < \lambda_{i^*}$, $ES^{Opt}(L^{\rho^\alpha}) = \lambda_{i^*}$.*
- *for $\alpha \geq \lambda_{i^*}$, $ES^{Opt}(L^{\rho^\alpha}) \in [\lambda_{i^*}, U^{OPT}(L^\pi)]$.*

Proof First it is worth noticing that the condition $\exists ! i^*$ $p_{i^*} = 1$ is equivalent to $|\{i | p_i > 0\}| = |\{i^*\}| = 1$.

If $\alpha < \lambda_{i^*}$ then $ES^{Opt}(L^{\rho^\alpha}) = \alpha + p_{i^*}(\lambda_{i^*} - \alpha) = \lambda_{i^*}$.

If $\alpha \geq \lambda_{i^*}$ then $ES^{Opt}(L^{\rho^\alpha}) = \max_i \min(\alpha\pi_i, \lambda_i) \geq \min(\alpha \times 1, \lambda_{i^*}) = \lambda_{i^*}$, since $p_{i^*} > 0$ implies $\pi_{i^*} = 1$ by construction. So, $ES^{Opt}(L^{\rho^\alpha}) = \max(\lambda_{i^*}, \max_{i > i^*} \min(\alpha\pi_i, \lambda_i)) \in [\lambda_{i^*}, \max_i \min(\pi_i, \lambda_i)]$. ■

As a consequence, if a piece of data (π, p, β) where p is a Dirac function on s_{i*} is represented by $ES^{Opt}(L^{\rho^\alpha})$, we must have that $\beta \in [\lambda_{i^*}, ES^{Opt}(L^\pi)]$.

Lemma 7 *Consider a lottery L^{ρ^α} with $\rho^\alpha = (p, \pi, \alpha)$. If $|\{i | p_i > 0\}| > 1$ with $\lambda_{iM} = \max_i(\lambda_i | p_i > 0)$, and $\alpha_1 \geq \lambda_{iM} > \alpha_2$, then $ES^{Opt}(L^{\rho^{\alpha_1}}) \neq ES^{Opt}(L^{\rho^{\alpha_2}})$.*

Proof

If $\alpha < \lambda_{iM}$ then $ES^{Opt}(L^{\rho^\alpha}) = \alpha + \sum_{i:\lambda_i > \alpha} p_i(\lambda_i - \alpha) < \alpha + \lambda_{iM} - \alpha = \lambda_{iM}$.

If $\alpha \geq \lambda_{iM}$ then $ES^{Opt}(L^{\rho^\alpha}) = \max_i \min(\alpha\pi_i, \lambda_i) \geq \min(\alpha \times 1, \lambda_{iM}) = \lambda_{iM}$. ■

The above result shows that the best utility value among probable outcomes serves as a threshold separating additive and maxitive behaviors of the prop-poss optimistic utility functional.

Note that in the probabilistic area, $ES^{Opt}(L^{\rho^\alpha})$ ranges between $\sum_{i=1}^n p_i\lambda_i$ ($\alpha = 0$) and 1 ($\alpha = 1$); in the possibilistic area, $ES^{Opt}(L^{\rho^\alpha})$ ranges between $\max_{i=1}^n \min(\pi_i, \lambda_i)$ ($\alpha = 1$) and 0 ($\alpha = 0$). The function $\alpha \in [0, 1] \mapsto ES^{Opt}(L^{\rho^\alpha})$ is non-decreasing. Moreover, we can show that

$$\max_{i=1}^n \min(\pi_i, \lambda_i) \geq \lambda_{iM} \geq \sum_{i=1}^n p_i\lambda_i.$$

The second inequality is obvious. The first is due to the fact that $\pi_{iM} = 1$.

More specifically, the following property holds for ES^{Opt}:

Proposition 8
For $\alpha = \lambda_{i^M}$, $ES^{Opt}(L^{\rho^\alpha}) = \lambda_{i^M}$.

Proof Indeed, $\lambda_{i^M} + \sum_{i:\lambda_i > \lambda_{i^M}} p_i(\lambda_i - \lambda_{i^M}) = \lambda_{i^M} + 0$; likewise $\max_i \min(\lambda_{i^M}\pi_i, \lambda_i) = \max_{i \geq i^M} \min(\lambda_{i^M}\pi_i, \lambda_i) = \min(\lambda_{i^M} \times 1, \lambda_{i^M})$ since for $i > i^M$, $\min(\lambda_{i^M}\pi_i, \lambda_i) \leq \lambda_{i^M}$. ■

So, the value λ_{i^M} serves as a break-point between the two forms of $ES^{Opt}(L^{\rho^\alpha})$. More precisely:

Proposition 9 *For a given piece of data (π, p, β) such that $|\{i|p_i > 0\}| > 1$ suppose there is a value α such that $ES^{Opt}(L^{\rho^\alpha})$ represents (π, p, β) (i.e. $ES^{Opt}(L^{\rho^\alpha}) = \beta$ with $\rho^\alpha = \alpha\pi + (1-\alpha)p$).*

$$\begin{cases} \text{if } \lambda_{i^M} > \beta, \text{ then } \beta = \alpha + \sum_{i:\lambda_i > \alpha} p_i(\lambda_i - \alpha) > \alpha \\ \text{if } \lambda_{i^M} \leq \beta, \text{ then } \beta = \max_i \min(\alpha\pi_i, \lambda_i) \leq \alpha. \end{cases}$$

Proof First, note that if $\Lambda^+_\alpha = \{i : \lambda_i > \alpha\} \neq \emptyset$, $ES^{Opt}(L^{\rho^\alpha}) = \alpha + \sum_{i:\lambda_i > \alpha} p_i(\lambda_i - \alpha) < \lambda_{i^M}$. Indeed, $\alpha + \sum_{i:\lambda_i > \alpha} p_i(\lambda_i - \alpha) < \alpha + \sum_{i:\lambda_i > \alpha} p_i(\lambda_{i^M} - \alpha) = \alpha(1 - P(\Lambda^+_\alpha)) + \lambda_{i^M} P(\Lambda^+_\alpha) < \lambda_{i^M}$.

Besides, if $\Lambda^+_\alpha = \{i : \lambda_i > \alpha\} = \emptyset$, $ES^{Opt}(L^{\rho^\alpha}) = \max_i \min(\alpha\pi_i, \lambda_i) \geq \min(\alpha, \lambda_{i^M}) = \lambda_{i^M}$. So,

- If $\lambda_{i^M} > \beta$ then $\beta = ES^{Opt}(L^{\rho^\alpha}) = \alpha + \sum_{i:\lambda_i > \alpha} p_i(\lambda_i - \alpha), \forall \alpha \in [0, \lambda_{i^M}[$ and clearly $\beta > \alpha$.
- If $\lambda_{i^M} \leq \beta$ and $|\{i|p_i > 0\}| > 1$, due to Lemma 7, we know that $\beta = ES^{Opt}(L^{\rho^\alpha}) = \max_i \min(\alpha\pi_i, \lambda_i) \leq \alpha, \forall \alpha \in [\lambda_{i^M}, 1]$.

■

From Proposition 9 we can reduce the range of possible values of α and identify if a data item is in the possibilistic area or in the probabilistic area. In the following we will distinguish between the possibilistic and probabilistic cases. In the possibilistic case, Proposition 10 shows that α is unique only if the evaluation β is not the utility of a state. Otherwise, the evaluation is coherent if the best states are not the most possible ones.

First, we need a result simplifying the expression of $ES^{Opt}(L^{\rho^\alpha}) = \max_i \min(\alpha\pi_i, \lambda_i)$ in the possibilistic area. Not all terms $\min(\alpha\pi_i, \lambda_i)$ are useful in the computation of $ES^{Opt}(L^{\rho^\alpha})$. If i and j are such that $\min(\alpha\pi_i, \lambda_i) \geq \min(\alpha\pi_j, \lambda_j), \forall \alpha \in [0, 1]$, the latter term is dominated and can be deleted from the expression.

Let ND be the set of indices of non dominated terms. We have that $ES^{Opt}(L^{\rho^\alpha}) = \max_{i \in ND} \min(\alpha\pi_i, \lambda_i)$.

Suppose only k dominating terms remain in the non-redundant formulation of $ES^{Opt}(L^{\rho^\alpha})$, with indices $\ell_1, \ldots \ell_k$. It is clear that $\lambda_{\ell_1} = \lambda_{i^M} < \lambda_{\ell_2} \cdots < \lambda_{\ell_k}$. Moreover, due to non-redundancy, we must have that $\pi_{\ell_1} = \pi_{i^M} = 1 > \pi_{\ell_2} \cdots > \pi_{\ell_k}$. Indeed if there were an equality between two π_{ℓ_j}'s, one of the corresponding terms would be redundant.

We can now solve the equation $\beta = ES^{Opt}(L^{\rho^\alpha})$ for α in the possibilistic area.

Proposition 10 *Suppose a piece of data (π, p, β) with $|\{i|p_i > 0\}| > 1$ is represented by $ES^{Opt}(L^{\rho^\alpha})$ with $\beta \in [\lambda_{i^M}, \max_i \min(\pi_i, \lambda_i)]$. Then there is an index $j > 1$ such that $\lambda_{\ell_{j-1}} < \beta \leq \lambda_{\ell_j}$ and the value α is such that*

- *$\alpha = \frac{\beta}{\pi_{\lambda_{\ell_j}}}$ if $\beta > \lambda_{\ell_{j-1}}$;*
- *$\alpha \in [\frac{\beta}{\pi_{\ell_j}}, \frac{\beta}{\pi_{\ell_{j+1}}}]$ if $\beta = \lambda_{\ell_j}$.*

Proof We have seen that $= ES^{Opt}(L^{\rho^\alpha}) = \max_{i \in ND} \min(\alpha\pi_i, \lambda_i) = \max_{j=1}^k \min(\alpha\pi_{\ell_j}, \lambda_{\ell_j})$. The function $ES^{Opt}(L^{\rho^\alpha})$ is continuous, piecewise linear non-decreasing, and may have flat sections (when $ES^{Opt}(L^{\rho^\alpha}) = \lambda_i$) and increasing sections (when $ES^{Opt}(L^{\rho^\alpha}) = \alpha\pi_i$). Fixing $\beta \in [\lambda_{i^M}, \max_i \min(\pi_i, \lambda_i)]$, there is an index j such that $\beta = \min(\alpha\pi_{\ell_j}, \lambda_{\ell_j})$. We consider two cases:

- $\beta = \alpha\pi_{\ell_j} < \lambda_{\ell_j}$. Then $\alpha = \frac{\beta}{\pi_{\ell_j}}$. This value ranges in $[\frac{\lambda_{\ell_{j-1}}}{\pi_{\ell_j}}, \frac{\lambda_{\ell_j}}{\pi_{\ell_j}}[$ when β ranges in $[\lambda_{\ell_{j-1}}, \lambda_{\ell_j}[$.
- $\beta = \lambda_{\ell_j} \leq \alpha\pi_{\ell_j}$; then $\alpha \geq \frac{\lambda_{\ell_j}}{\pi_{\ell_j}}$. However, $\alpha\pi_{\ell_{j+1}}$ cannot overpass λ_{ℓ_j}; otherwise we get $\beta = \min(\alpha\pi_{\ell_{j+1}}, \lambda_{\ell_{j+1}})$, contrary to the assumption (remember that $\pi_{\ell_{j+1}} < \pi_{\ell_j}$). So, if $\beta = \lambda_{\ell_j}$,α can be any value in $[\frac{\lambda_{\ell_j}}{\pi_{\ell_j}}, \min(1, \frac{\lambda_{\ell_j}}{\pi_{\ell_{+1}}})]$.

Note that $\frac{\lambda_{\ell_j}}{\pi_{\ell_j}} < \frac{\lambda_{\ell_{j+1}}}{\pi_{\ell_{j+1}}}, \forall j = 1, \ldots, k-1$. And $\frac{\lambda_{\ell_j}}{\pi_{\ell_j}} \leq 1$, $j = 1, \ldots, k-1$. For if $\frac{\lambda_{\ell_j}}{\pi_{\ell_j}} > 1$ then $\lambda_{\ell_j} > \pi_{\ell_j} \geq \pi_{\ell_j}\alpha$. So the value of ES^{Opt} cannot reach λ_{ℓ_j}, (hence not $\lambda_{\ell'_j}$, $j' > j$, which is contradictory with the non-redundancy assumption, for $j \leq k-1$. For $j = k$ it means that $ES^{Opt} = \pi_{\ell_j}$ when $\alpha = 1$.

■

Now we consider a data item that we know lies in the probabilistic area. Proposition 11 shows that the evaluation β is coherent with ES^{Opt} if it lies between the expected utility and the best utlity value λ_i^M with positive probability. And in this case α is unique.

Proposition 11 *For a given piece of data (π, p, β) with $|\{i|p_i > 0\}| > 1$, represented by $ES^{Opt}(L^{\rho^\alpha})$ with $\beta \in$*

	$i:$ λ_i	1 0.01	2 0.3	3 0.5	4 0.8	5 1	β
$j=1$	π^1	0.2	0.6	1	1	0	0.8
	p^1	0	0	0	1	0	
$j=2$	π^2	1	1	0.5	0.5	0	0.3
	p^2	0.4	0.6	0	0	0	
$j=3$	π^3	0	1	0.5	1	1	0.82
	p^3	0	0.1	0	0.6	0.3	
$j=4$	π^4	1	1	1	1	1	0.51
	p^4	0.2	0.3	0.3	0.2	0	

Table 1: Data

$[\sum_i \pi_i\lambda_i, \lambda_{iM}]$, *if* $\lambda_{iM} > \beta \geq \sum_{i=1}^n \lambda_i p_i$ *then there exists a unique* α *such that* $ES^{OPt}(L^{\rho^\alpha}) = \beta$. *It is of the form:*

$$\alpha = \frac{\beta - P(\Lambda_\alpha^+)E[\lambda|\Lambda_\alpha^+]}{(1-P(\Lambda_\alpha^+))} = \frac{\beta - \sum_{j=i+1}^n p_j\lambda_j}{(1-\sum_{j=i+1}^n p_j)} \in [\lambda_i, \lambda_{i+1}[. \tag{9}$$

Proof The proof is in the same style as the one of Proposition 1. The equation to be solved can be put in the form:

$$\beta - \alpha = \sum_{i=1}^n p_i \max(0, \lambda_i - \alpha).$$

It can thus be expressed as the intersection points of the two functions $h(\alpha) = \beta - \alpha$ and $f(\alpha) = \sum_{i=1}^n p_i \max(0, \lambda_i - \alpha)$ on interval $[0,1]$. Notice that h and f are both continuous. h is linear decreasing on $[0,1]$ with slope -1, $h(0) = \beta, h(\beta) = 0$; f is piecewise linear decreasing on $[0, \lambda_{iM}]$ and $f(\alpha) = 0$ on $]\lambda_{iM}, 1]$ and $f(0) = \sum_{i=1}^n p_i$. From condition $\lambda_{iM} > \beta \geq \sum_{i=1}^n \lambda_i p_i$, $f(0) \leq \beta = h(0)$ and $f(\beta) \geq 0 = h(\beta)$. So $\exists \alpha^* \in [0, \lambda_{iM}]$ such that $f(\alpha^*) = h(\alpha^*)$ and this value is unique due to the shape and positions of these functions. So there is a unique index i such that $\alpha^* \in [\lambda_i, \lambda_{i+1})$ of the form indicated in (9). ■

To conclude this analysis, we can answer the question whether the decision maker judgment can be represented by $ES^{OPt}(L^{\rho^\alpha})$ whatever the example but α is unique only if we have at least one example in the expected utility part of $ES^{OPt}(L^{\rho^\alpha})$ or at least one example using the possibility part of $ES^{OPt}(L^{\rho^\alpha})$ with $\beta_j \in]\lambda_{i^*}\lambda_{i^*+1}[$. Below is an example of elicitation.

Example 4 *Let us consider the problem with 4 data items on 5 states* $\{1,\ldots,5\}$ *represented in table 1. The intervals of possible* α*'s applying proposition 9 are presented in table 2.*

- *Piece of data* $j=1$*: There is only one* i *such that* $|i|p_i = 1| = 1$ *so according to the Lemma 6 all* α *in* $[0,1]$ *are possible.*
- *Piece of data* $j=2$*:* $\lambda_{iM} = 0.3 = \beta$*. So,* $\alpha = 0.3$ *is a solution. In the possibilistic area,* $E^{opt} = \max(\min(\alpha, 0.3), \min(0.5\alpha, 0.5), \min(0.5\alpha, 0.8))$. *The second term is redundant. We can use proposition 9.*
  We have $\beta = \lambda_2$ *so* $\alpha \geq \frac{\lambda_2}{\pi_2} = \frac{0.3}{1} = 0.3$ *and* $\alpha \leq \frac{0.3}{0.5} = 0.6$*. So the possible values of* α *are in* $[0.3, 0.6]$.

	$\underline{\alpha}$	$\overline{\alpha}$
$j=1$	0	1
$j=2$	0.3	0.6
$j=3$	0	0.82
$j=4$	0	0.51
α	0.3	0.51

Table 2: Candidate values of α after Prop. 9

- *Piece of data* $j=3$*:* $\lambda_{iM} = 1 > \beta = 0.82$*. So, we are in the probabilistic case. We know from Proposition 9 that* $\alpha < 0.82$*. To be more precise, we need to test several intervals.*
  - *We start with interval* $\alpha \in [0.5, 0.8[$*, we have from equ. (9):* $\alpha = \frac{0.82-0.6\times0.8-0.3\times1}{1-0.6-0.3} = 0.4 \notin [0.5, 0.8[$
  - *so we check the interval* $\alpha \in [0.3, 0.5]$*, the equation remains the same since* $p_3 = 0$*. We obtain* $\alpha = 0.4 \in [0.3, 0.5]$.
- *Piece of data* $j=4$*:* $\lambda_{iM} = 0.8 > \beta = 0.51$*. This is a probabilistic case. We know from proposition 9 that* $\alpha < 0.51$*. The reader can check that again* $\alpha = 0.4$ *since* $\alpha = \frac{0.51-0.3\times0.5-0.2\times0.8}{1-0.3-0.2} = 0.4$.

  So, the decision-maker is consistent across all four examples: $\alpha = 0.4$ *is a valid choice for the 4 items.*

7. Conclusion

We have pursued the study of a joint extension of possibilistic and probabilistic utility functionals, that preserves good dynamic properties in decision trees. The capacity at work is a special case of belief functions that is a convex mixture between a probability and a possibility function. We have shown that any weight distribution on states whose sum is at least one can be interpreted as determining in a unique way such a capacity. With such a model, we can distinguish between normal states with positive probability, and abnormal ones that have zero probability, but are more or less impossible. We have given explicit expressions of generalized utility functionals based on this capacity. We also provided first steps toward the identification of the mixture coefficient α, based on human estimation of the worth of poss-prob lotteries. Future works may focus on the elicitation of α from a preference ordering of lotteries, rather than human-originated numerical values.

References

[1] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. In *Proc. IJCAI*, volume 95, pages 1924–1930, 1995.

[2] D. Dubois, E. Pap, and H. Prade. Hybrid probabilistic-possibilistic mixtures and utility functions. In J. Fodor, B. De Baets, and P. Perny, editors, *Preferences and Decisions under Incomplete Knowledge*, pages 51–73. Springer, 2000.

[3] D. Dubois, Hélène Fargier, and Vincent Galvagnon. On latest starting times and floats in activity networks with ill-known durations. *European Journal of Operational Research*, 147:266–280, 2003.

[4] Didier Dubois, Hélène Fargier, Romain Guillaume, and Agnès Rico. Sequential decision-making under uncertainty using hybrid probability-possibility functions. In Vicenç Torra and Yasuo Narukawa, editors, *Modeling Decisions for Artificial Intelligence*, pages 54–66, Cham, 2021. Springer International Publishing.

[5] E P Klement, R Mesiar, and E Pap. *Triangular Norms*. Kluwer Academic, Dordrecht, 2000.

[6] H. Raiffa. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Addison-Wesley, 1968.

[7] Glenn Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.

[8] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.