Constructing coherent models of conditional and unconditional upper probabilities

Hugo Janssen
KHLeuven – Departement REGA
Sint-Maartensstraat 55D, 3000 Leuven
hugo.janssen@pi.be
http://ippserv.rug.ac.be/∼hugo

Abstract

We investigate the coherence of a possibilistic system modelled by the joint possibility distribution function of a finite sequence of linearly ordered possibilistic variables together with the conditional possibility distribution function of any variable in this sequence, given the values assumed by all preceding variables. To ensure the coherence of this model it is necessary and sufficient that the conditional possibilities are between those calculated from the joint possibility distribution function by Dempster's rule and the natural extension rule. We then show how a coherent model of conditional and uncondional upper probabilities can be constructed, using a given coherent model and a finitely additive probability. The method used to construct this model is to form convex combinations of the finitely additive probability and the upper probabilities belonging to the given model.

Keywords. Coherence, natural extension rule, Dempster's conditioning rule, possibility measure, upper probability.

1 Introduction

We present a technique that can be used to generate coherent models of conditional and unconditional upper probabilities. Therefore, we need a coherent model with the following structure: an unconditional upper probability $\overline{P}$, defined on the power set $\wp(\Omega)$ of some finite set Ω, and conditional upper probabilities $\overline{P}(\cdot|B)$, where B is an element of a partition $\mathcal{B}$ of Ω, all defined on some field $\mathcal{A}$ on Ω. Using a finitely additive probability P on $\wp(\Omega)$, we then construct a new model as follows. The unconditional upper probability is replaced by a convex combination $\overline{P}_\alpha = \alpha P + (1-\alpha)\overline{P}$ (where $\alpha \in\,]0,1[$). Using a suitable non-empty class $\mathcal{M}$ of finitely additive probabilities, we will replace each conditional upper probability $\overline{P}(\cdot|B)$, $B \in \mathcal{A}$, by the upper envelope $\overline{P}_\alpha(\cdot|B)$ of the conditional probabilities, calculated from the convex combinations $\alpha P + (1-\alpha)Q$, $Q \in \mathcal{M}$ by Bayes' rule. We will propose two conditions that may be imposed to derive suitable formulae for these upper envelopes. Naturally the model constituted by these replacements is also coherent. In particular we will show that, when the given conditional upper probabilities are calculated from $\overline{P}$ by Dempster's rule, their replacements are also calculated from $\overline{P}_\alpha$ by the same rule.

The paper is organised as follows. In Section 2 we will discuss the notion of weak coherence for models of conditional and unconditional upper probabilities. We will furthermore explain how this rationality criterion can be formulated in terms of finitely additive probabilities. This alternative definition will be used in Section 3 to characterise the coherence of a specific model of conditional and unconditional possibilities. This characterisation improves a result derived in a previous study, initiated in order to support a behavioural interpretation of possibilistic processes [6, 10]. The concluding section presents the just-mentioned coherence preserving technique to transform models of conditional and unconditional upper probabilities (for instance the coherent models treated in Section 3).

2 Consistency criteria

If we want to give a behavioural interpretation to a system modelled by conditional and unconditional upper probabilities, it is necessary to verify whether or not the model describing the system, is coherent. Essentially, coherence is a rationality criterion ensuring that the upper probabilies are assessed in a consistent way. To make this more clear, let Ω be a possibility space, i.e., a non-empty set of possible states of the world. Unless stated otherwise, we assume that Ω a finite set. Consider an upper probability $\overline{P}\colon \mathcal{S} \to [0,1]$ on a non-empty class $\mathcal{S}$ of subsets of Ω. Under the behavioural interpretation [3, 13], $\overline{P}(A)$, $A \in \mathcal{S}$ is interpreted as a subject's infimum acceptable rate for betting against the event A[1]. This means that the subject is willing to bet against the event A at any rate $\lambda > \overline{P}(A)$,

[1]Note that all gains and losses from betting are assumed to be measured on a linear utility scale.

giving him λ units of utility when A does not occur, and $\lambda - 1$ units when A occurs. The net reward resulting from the bet at rate $\overline{P}(A)$ can also be written as:

$$G(A) = \overline{P}(A) - I_A,$$

where I_A is the indicator function of A. The behavioural interpretation of $\overline{P}$ implies that all gambles $G(A)$, $A \in \mathcal{S}$ are *marginally acceptable* to the subject, meaning that he is disposed to accept $G(A) + \delta$ for all $\delta > 0$. Moreover, a rational subject should consider positive linear combinations of acceptable gambles as acceptable [13].

To ensure that the values $\overline{P}(A)$, $A \in \mathcal{S}$ are assessed in a consistent way, $\overline{P}$ should be coherent, i.e., for any natural number n, for any non-negative real numbers $\lambda_o, \ldots, \lambda_n$ and for any events $A_o, \ldots, A_n \in \mathcal{S}$, it must hold that

$$\sup_{\omega\in\Omega}\left[\sum_{j=1}^{n}\lambda_j G(A_j)(\omega) - \lambda_o G(A_o)(\omega)\right] \geq 0. \quad (1)$$

If (1) fails and $\lambda_o = 0$, then there is some $\delta > 0$ such that $\sum_{j=1}^{n}\lambda_j[G(A_j) + \delta] \leq -\delta$. This is a 'sure loss' since there is a positive linear combination of acceptable gambles that is uniformly negative, meaning that the subject cannot avoid losing some positive amount of utility. The coherence condition (1) ensures that the subject avoids incurring a sure loss.

If (1) fails and $\lambda_o > 0$, then there is some $\delta > 0$ such that $[\overline{P}(A_o) - \delta] - I_{A_o} \geq \lambda_o^{-1}\sum_{j=1}^{n}\lambda_j[G(A_j) + \delta]$. Since all gambles $G(A_i) + \delta$, $i \in \{1, \ldots, n\}$ are acceptable to the subject, the linear combination $\lambda_o^{-1}\sum_{j=1}^{n}\lambda_j[G(A_j) + \delta]$ is acceptable too, and so is the gamble $[\overline{P}(A_o) - \delta] - I_{A_o}$, as it yields a uniformly higher gain. As a result, the subject can be induced to bet against A_o at the rate $\overline{P}(A_o) - \delta$, which is strictly smaller than his *infimum* acceptable betting rate $\overline{P}(A_o)$. Coherence rules out this type of inconsistency.

Let us now consider the behavioural interpretation of conditional upper probabilities. Let $\mathcal{B}$ be a partition of the possibility space Ω. An upper probability $\overline{P}(\cdot|B)$ defined on a class $\mathcal{S}_B$ of subsets of Ω, is called a conditional upper probability on $\mathcal{S}_B$, associated with B. Two interpretations can be given to the values taken by $\overline{P}(\cdot|B)$. Under the *contingent* interpretation $\overline{P}(A|B)$ is the marginally acceptable upper rate for betting against the event $A \in \mathcal{S}_B$ contingent on B, i.e., where the bet is called off unless the event B is observed. Under the *updating* interpretation $\overline{P}(A|B)$ is the marginally acceptable upper rate for betting against A that the subject would adopt after observing event B, provided he learns nothing more about Ω. Under both interpretations the net reward is the gamble

$$G(A|B) = I_B(\overline{P}(A|B) - I_A),$$

where I_A and I_B are the indicator functions of A and B. By the *Updating Principle* $\overline{P}(A|B)$ should have the same value under both interpretations [13]. The conditional upper probabilities $\overline{P}(\cdot|B)$, $B \in \mathcal{B}$, are said to be *separately coherent* if $\overline{P}(\cdot|B)$, $B \in \mathcal{B}$, is a coherent upper probability on $\mathcal{S}_B$ such that $\Omega \setminus B \in \mathcal{S}_B$ and $\overline{P}(\Omega \setminus B|B) = 0$.

Consider now the next model constituted by the assessments of a subject facing uncertainty:

- $\mathcal{B}_0, \ldots, \mathcal{B}_n$, $n \in \mathbb{N}$ is a finite sequence of partitions of the possibility space Ω;
- $\mathcal{A}_j$, $j \in \{0, \ldots, n\}$ is a field on Ω containing all elements of the partition $\mathcal{B}_j$, i.e., $\mathcal{B}_j \subseteq \mathcal{A}_j$;
- for all $B \in \mathcal{B}_j$,$j \in \{0, \ldots, n\}$, $\overline{P}(\cdot|B)$ is a conditional upper probability on $\mathcal{A}_j$.

The conditional upper probabilities $(\overline{P}(\cdot|B) \mid B \in \mathcal{B}_j)$, $j \in \{0, \ldots, n\}$ are furthermore assumed to be separately coherent. Consequently, any positive linear combination of the marginally acceptable gambles

$$G(A|B) = I_B(\overline{P}(A|B) - I_A),$$

where $(A, B) \in \mathcal{A}_j \times \mathcal{B}_j$ and where j may assume any natural number between 0 and n, is at least marginally acceptable for the subject. To guarantee the coherence of the model such a positive linear combination may not produce a sure loss for the subject. Moreover, such a gamble may not be uniformly smaller than a gamble $G(A|B) = I_B(\overline{P}(A|B) - I_A)$, $(A, B) \in \mathcal{A}_\ell \times \mathcal{B}_\ell$ and $\ell \in \{0, \ldots, n\}$, since this would force the subject to bet against A at a rate smaller than his infimum acceptable betting rate $\overline{P}(A|B)$. Formally, $(\overline{P}(\cdot|B) \mid B \in \mathcal{B}_j)$, $j \in \{0, \ldots, n\}$ is said to be *coherent*[2] if for all mappings μ_j, $j \in \{0, \ldots, n\}$ that assign to a couple $(A, B) \in \mathcal{A}_j \times \mathcal{B}_j$ a nonnegative real number $\mu_j(A, B) \in \mathbb{R}^+$, and for any couple $(A_\ell, B_\ell) \in \mathcal{A}_\ell \times \mathcal{B}_\ell$ where $\ell \in \{0, \ldots, n\}$:

$$\sup_{\omega\in\Omega}\left(\sum_{j=0}^{n}\sum_{\substack{A\in\mathcal{A}_j\\ B\in\mathcal{B}_j}}\mu_j(A, B)G(A|B) - G(A_\ell|B_\ell)\right)(\omega) \geq 0. \quad (2)$$

Let us now assume that $\mathcal{B}_o = \{\Omega\}$ and that $\mathcal{A}_o$ coincides with the power set of Ω, i.e., $\mathcal{A}_o = \wp(\Omega)$. For notational ease we will denote the unconditional upper probability $\overline{P}(\cdot|\Omega)$ on $\wp(\Omega)$ by $\overline{P}$. The coherence of the model constituted by $\overline{P}$ and $(\overline{P}(\cdot|B) \mid B \in \mathcal{B}_j)$, $j \in \{1, \ldots, n\}$), can be characterised by a class of finitely additive probabilities. Therefore, we will need the class $\mathcal{M}_1$ of finitely additive probabilities on $\wp(\Omega)$ determined as follows: $P \in \mathcal{M}_1$ iff

- $P(A) \leq \overline{P}(A)$ for all $A \in \wp(\Omega)$;

[2] This condition is equivalent to what is called 'weak coherence' in [13].

– $P(A \cap B) \leq P(B)\overline{P}(A|B)$ for all $(A, B) \in \mathcal{A}_j \times \mathcal{B}_j$ where $j \in \{1, \dots, n\}$.

Theorem 2.1 characterises the coherence of a model of unconditional and conditional upper probabilities by a non-empty class of finitely additive probabilities [6, 13, 15].

Theorem 2.1 *Assume that $\overline{P}$ and $(\overline{P}(\cdot|B) \mid B \in \mathcal{B}_j)$, $j \in \{1, \dots, n\}$ are separately coherent. Then $\overline{P}$ and $(\overline{P}(\cdot|B) \mid B \in \mathcal{B}_j)$, $j \in \{1, \dots, n\}$ are coherent iff there exists a non-empty class $\mathcal{M}$ of finitely additive probabilities on $\wp(\Omega)$ such that*

(E_1) $\overline{P}(A) = \sup\{P(A)\colon P \in \mathcal{M}\}$ *for all* $A \in \wp(\Omega)$*;*

(E_2) $\overline{P}(A|B) \geq \sup\{\frac{P(A \cap B)}{P(B)}\colon P \in \mathcal{M}$ *and* $P(B) > 0\}$ *for all* $(A, B) \in \mathcal{A}_j \times \mathcal{B}_j$, $j \in \{1, \dots, n\}$*, and the equality holds whenever* $\overline{P}(\Omega \setminus B) < 1$.

If there exists a class $\mathcal{M}$ with the above-mentioned properties, then $\mathcal{M}_1$ is the greatest class having these properties.

In the next section we will use this result to prove the coherence of a specific model of conditional and unconditional possibilities.

3 Coherent models of unconditional and conditional possibilities

In [6, 9, 10] we have investigated the coherence of models with the following structure: the joint possibility distribution function of a finite collection of *linearly* ordered possibilistic variables $f_o, \dots, f_N$, $N \in \mathbb{N} \setminus \{0\}$ together with the conditional possibility distribution functions of any variable f_{n+1}, $n \in \{0, \dots, N-1\}$, given that the preceding variables $(f_o, \dots, f_n)$ jointly assume some value $(x_o, \dots, x_n) \in X_o \times \dots \times X_n$. These models were proven to be coherent when all variables $f_o, \dots, f_N$ have a finite sample space, all possibility distribution functions are normal, and all conditional possibilities are calculated by Dempster's conditioning rule. In this section we will prove that, to ensure the coherence of these models, it is necessary and sufficient that the conditional possibilities are between those calculated by Dempster's rule and natural extension.

Before formulating our result, we will recall some basic notions and results of possibility theory. A *possibility measure* [16] Π on the power set $\wp(\Omega)$ of a non-empty set Ω is a set mapping taking elements of $\wp(\Omega)$ to values in the real unit interval $[0, 1]$, such that for any collection $(A_j \mid j \in J)$ of elements of $\wp(\Omega)$:

$$\Pi(\bigcup_{j \in J} A_j) = \sup_{j \in J} \Pi(A_j).$$

The $\Omega - [0, 1]$-mapping π defined by $\pi(\omega) = \Pi(\{\omega\})$, $\omega \in \Omega$ is called the *distribution* of Π. Obviously, Π is completely determined by π, since for any $E \in \wp(\Omega)$: $\Pi(E) = \sup_{\omega \in E} \pi(\omega)$. The possibility measure Π and its distribution π are called *normal* if $\Pi(\Omega) = \sup_{\omega \in \Omega} \pi(\omega) = 1$. It it well-known that normality of a distribution is a sufficient and a necessary requirement for the corresponding possibility measure to be coherent.

Similarly to a stochastic variable, a possibilistic variable [1, 2] has a *basic space* Ω and a *sample space* X. The available information is represented by a possibility measure Π_Ω on $(\Omega, \wp(\Omega))$. Any $\Omega - X$-mapping f is then called a possibilistic variable in X. The $X - [0, 1]$-mapping π_f, given for any $x \in X$ by $\pi_f(x) = \Pi_\Omega(f^{-1}(\{x\}))$, is called the *possibility distribution function* of the possibilistic variable f. $\pi_f(x)$ is the possibility (or upper probability) that the variable f assumes the value x. We denote by Π_f the unique possibility measure on $\wp(X)$ with distribution π_f.

Consider now a finite sequence of possibilistic variables $f_o, \dots, f_N$, $N \in \mathbb{N} \setminus \{0\}$, having basic space $(\Omega, \wp(\Omega), \Pi_\Omega)$. Assume that the variables $f_o, \dots, f_N$ have the finite sets $X_o, \dots, X_N$ as sample spaces. The possibility that these variables jointly assume some value $(x_o, \dots, x_N) \in \times_{i=o}^N X_i$, is expressed by their *joint possibility distribution* $\pi_{(f_o, \dots, f_N)}$ as follows:

$$\pi_{(f_o, \dots, f_N)}(x_o, \dots, x_N) = \Pi_\Omega(\bigcap_{i=0}^N f_i^{-1}(\{x_i\})).$$

The possibility measure on $\wp(\times_{i=0}^N X_i)$ that is uniquely determined by $\pi_{(f_o, \dots, f_N)}$, is denoted by $\Pi_{(f_o, \dots, f_N)}$.

It will be assumed that $\Pi_{(f_o, \dots, f_N)}$ is normal. This implies that $\Pi_{(f_o, \dots, f_N)}$ and all the marginals that can be derived from $\Pi_{(f_o, \dots, f_N)}$ are coherent upper probabilities. To remind, Π_{f_o} is the marginal of $\Pi_{(f_o, \dots, f_N)}$ on the space X_o, i.e., $\Pi_{f_o}(A) = \Pi_{(f_o, \dots, f_N)}(A \times X_1 \times \dots \times X_N)$ for all $A \subseteq X_o$, or equivalently, $\pi_{f_o}(x_o) = \sup_{(x_1, \dots, x_N) \in \times_{i=1}^N X_i} \pi_{(f_o, \dots, f_N)}(x_o, \dots, x_N)$. Similar formulas can be given for other, more complicated types of marginals, such as for instance $\Pi_{(f_1, f_2, f_N)}$.

As mentioned before, the model we are investigating here, contains also the conditional possibility distribution functions $\pi_{f_{n+1}|(f_o, \dots, f_n)}(\cdot \mid x_o, \dots, x_n)$ of each variable f_{n+1}, $n \in \{0, \dots, N-1\}$, given the values $(x_o, \dots, x_n) \in \times_{i=0}^n X_i$ assumed by all preceding variables. These distribution functions may be calculated by some conditioning rule from $\pi_{(f_o, \dots, f_N)}$, but they may also be assessed independently of the joint possibility distribution function. Similarly to what we did before for $\Pi_{(f_o, \dots, f_N)}$ we require that all conditional possibility distribution functions should be normal. This ensures that the corresponding possibility measures $\Pi_{f_{n+1}|(f_o, \dots, f_n)}(\cdot \mid x_o, \dots, x_n)$, $(x_o, \dots, x_n) \in \times_{i=0}^n X_i$

and $n \in \{0, \dots, N-1\}$, are coherent upper probabilities. For convenience we will use the (conditional and uncon-ditonal) possibility distribution functions to indicate our model.

For the case where $N = 1$, it has been proven by Walley and De Cooman [15] that the (unconditional) joint possibility distribution function $\pi_{(f_o,f_1)}$ of any two possibilistic variables f_o and f_1 with finite sample spaces X_o and X_1 together with the conditional possibility distribution functions $\pi_{f_1|f_o}(\cdot \mid x_o)$, $x_o \in X_o$ of f_1, given that f_o assumes some value $x_o \in X_o$, constitute a coherent model if and only if the conditional possibilities are between those calculated by Dempster's rule and natural extension, or equivalently, the conditional possibilities $\pi_{f_1|f_o}(x_1 \mid x_o)$, $(x_o, x_1) \in X_o \times X_1$ satisfy

$$ {}_{\text{DE}}\pi_{f_1|f_o}(x_1 \mid x_o) \leq \pi_{f_1|f_o}(x_1 \mid x_o) \leq {}_{\text{NE}}\pi_{f_1|f_o}(x_1 \mid x_o) \tag{3} $$

whenever $\pi_{f_o}(x_o) > 0$. In (3), we have used the notation ${}_{\text{DE}}\pi_{f_1|f_o}(x_1 \mid x_o)$ for the conditional possibility calculated by *Dempster's rule*: for any couple of elements $(x_o, x_1) \in X_o \times X_1$,

$$ {}_{\text{DE}}\pi_{f_1|f_o}(x_1 \mid x_o) = \begin{cases} \dfrac{\pi_{(f_o,f_1)}(x_o, x_1)}{\pi_{f_o}(x_o)} & \text{if } \pi_{f_o}(x_o) > 0 \\ 1 & \text{if } \pi_{f_o}(x_o) = 0, \end{cases} $$

where the least committal value is taken for ${}_{\text{DE}}\pi_{f_1|f_o}(x_1 \mid x_o)$ when $\pi_{f_o}(x_o) = 0$. The notation ${}_{\text{NE}}\pi_{f_1|f_o}(x_1 \mid x_o)$ is used for the conditional possibility calculated by the *natural extension rule* for conditioning. To remind, for any couple $(x_o, x_1) \in X_o \times X_1$, we have the following expression: if $\pi_{(f_o,f_1)}(x_o, x_1) \leq \beta_{f_o}(x_o) < 1$, then

$$ {}_{\text{NE}}\pi_{f_1|f_o}(x_1 \mid x_o) = \frac{\pi_{(f_o,f_1)}(x_o, x_1)}{\pi_{(f_o,f_1)}(x_o, x_1) + 1 - \beta_{f_o}(x_o)}; $$

if $\beta_{f_o}(x_o) < \pi_{(f_o,f_1)}(x_o, x_1)$, then

$$ {}_{\text{NE}}\pi_{f_1|f_o}(x_1 \mid x_o) = \pi_{(f_o,f_1)}(x_o, x_1); $$

and ${}_{\text{NE}}\pi_{f_1|f_o}(x_1 \mid x_o) = 1$ when $\beta_{f_o}(x_o) = 1$, where

$$ \begin{aligned} \beta_{f_o}(x_o) &= \Pi_{f_o}(X_o \setminus \{x_o\}) \\ &= \max\{\pi_{f_o}(x) \colon x \in X_o, x \neq x_o\}. \end{aligned} $$

A discussion of the background of this conditioning rule can be found in [15].

According to (3) the above-mentioned model is coherent.

In [6, 9, 10] we have generalised Walley and De Cooman's result to a finite number of possibilistic variables $f_o, \dots, f_N$, $N \in \mathbb{N} \setminus \{0\}$, provided that all conditional possibilities are calculated by Dempster's conditioning rule.

Theorem 3.1 *The model constituted by the distributions* $\pi_{(f_o,\dots,f_N)}$ *and* ${}_{\text{DE}}\pi_{f_{n+1}|(f_o,\dots,f_n)}(\cdot \mid x)$, $x \in \times_{i=0}^{n} X_i$, $n \in \{0, \dots, N-1\}$ *is coherent.*

We will now motivate that the coherence of the general model is guaranteed iff the conditional possibilities are between those calculated by Dempster's rule and natural extension rule. This means that the characterisation derived by Walley and De Cooman for the case of two variables, also holds for the general case.

To see this, suppose that the distributions $\pi_{(f_o,\dots,f_N)}$ and $\pi_{f_{n+1}|(f_o,\dots,f_n)}(\cdot \mid x)$, $x \in \times_{i=0}^{n} X_i$, $n \in \{0, \dots, N-1\}$, constitute a coherent model. Then, for any $n \in \{0, \dots, N-1\}$, it follows that the marginal distribution function $\pi_{(f_o,\dots,f_{n+1})}$ and $\pi_{f_{n+1}|(f_o,\dots,f_n)}(\cdot \mid x)$, $x \in \times_{i=0}^{n} X_i$, also constitute a coherent model. According to Walley's and De Cooman's result [15] the conditional possibility $\pi_{f_{n+1}|(f_o,\dots,f_n)}(y \mid x)$, $x \in \times_{i=0}^{n} X_i$ and $y \in X_{n+1}$, should be between ${}_{\text{DE}}\pi_{f_{n+1}|(f_o,\dots,f_n)}(y \mid x)$ and ${}_{\text{NE}}\pi_{f_{n+1}|(f_o,\dots,f_n)}(y \mid x)$, provided that $\pi_{(f_o,\dots,f_n)}(x_o, \dots, x_n) > 0$. The following theorem tells us that the converse assertion also holds.

Theorem 3.2 *The model constituted by the distributions* $\pi_{(f_o,\dots,f_N)}$ *and* $\pi_{f_{n+1}|(f_o,\dots,f_n)}(\cdot \mid x)$, $x \in \times_{i=0}^{n} X_i$, $n \in \{0, \dots, N-1\}$ *is coherent iff the conditional possibilities* $\pi_{f_{n+1}|(f_o,\dots,f_n)}(y \mid x)$, $(x, y) \in \times_{i=0}^{n} X_i \times X_{n+1}$ *and* $n \in \{0, \dots, N-1\}$ *are, analogously to* (3)*, between those calculated by Dempster's rule and natural extension.*

Proof: The proof of this result is rather long and technical. Therefore, we just give an outline of the main steps. A more detailed similar proof for the special case where the conditional possibilities are defined by Dempster's rule can be found in [6]. For the coherence of the model it is necessary and sufficient that there is a non-empty class $\mathcal{M}$ of finitely additive probability measures on $\wp(\times_{i=0}^{N} X_i)$ such that

(i) $\Pi_{(f_o,\dots,f_N)}$ is the upper envelope of $\mathcal{M}$, i.e., for all $A \in \wp(\times_{i=0}^{N} X_i)$:

$$ \Pi_{(f_o,\dots,f_N)}(A) = \sup\{P(A) \colon P \in \mathcal{M}\}; $$

(ii) for all $n \in \{0, \dots, N-1\}$, for all $x \in \times_{i=0}^{n} X_i$, and for all $A \in \wp(X_{n+1})$:

$$ \Pi_{f_{n+1}|(f_o,\dots,f_n)}(A \mid x) \geq \sup\{\frac{P(C_{x,A})}{P(C_x)} \colon P \in \mathcal{M}_x\} \tag{4} $$

where

$$ \begin{aligned} C_x &= \{(y_o, \dots, y_N) \in \mathop{\times}_{i=0}^{N} X_i \colon (y_o, \dots, y_n) = x\} \\ C_A &= \{(y_o, \dots, y_N) \in \mathop{\times}_{i=0}^{N} X_i \colon y_{n+1} \in A\} \\ C_{x,A} &= C_x \cap C_A \\ \mathcal{M}_x &= \{P \in \mathcal{M} \colon P(C_x) > 0\}, \end{aligned} $$

and the equality in (4) holds whenever $\Pi_{(f_o,\ldots,f_N)}(\times_{i=0}^N X_i \setminus C_x) < 1$.

For the construction of a suitable collection of finitely additive probability measures we need the following mappings.

For any element $n \in \{0,\ldots,N-1\}$, let $g_{n,n+1}$ be the $\times_{i=0}^n X_i - \times_{i=0}^{n+1} X_i$-mapping that assigns to an element $x = (x_o,\ldots,x_n)$ of $\times_{i=0}^n X_i$ an element $y = (y_o,\ldots,y_{n+1})$ of $\times_{i=0}^{n+1} X_i$ such that

$$\begin{cases} y_i & = x_i, \quad i \in \{0,\ldots,n\} \\ \pi_{(f_o,\ldots,f_{n+1})}(y) & = \pi_{(f_o,\ldots,f_n)}(x). \end{cases} \tag{5}$$

Note that such a map $g_{n,n+1}$ always exists since X_{n+1} is a finite set. Let us furthermore denote by $g_{n,n}$ the identical permutation of $\times_{i=0}^n X_i$, $n \in \{0,\ldots,N\}$. Using the equation

$$g_{k,n} = g_{n-1,n} \circ g_{k,n-1},$$

where $(k,n) \in \mathbb{N}^2$ such that $0 \le k < n \le N$, a class of mappings $g_{k,n}$, $(k,n) \in \mathbb{N}^2$ such that $0 \le k \le n \le N$ can be generated recursively, and in such a way that

- for any couple $(k,n) \in \mathbb{N}^2$ such that $0 \le k \le n \le N$ and for any element $x = (x_o,\ldots,x_k) \in \times_{i=0}^k X_i$:

$$\begin{cases} g_{k,n}(x)_i & = x_i, \quad i \in \{0,\ldots,k\} \\ \pi_{(f_o,\ldots,f_n)}(g_{k,n}(x)) & = \pi_{(f_o,\ldots,f_k)}(x); \end{cases}$$

- for any triple $(k,\ell,n) \in \mathbb{N}^3$ such that $0 \le k \le \ell \le n$:

$$g_{k,n} = g_{\ell,n} \circ g_{k,\ell}.$$

Since π_{f_o} is a normal possibility distribution function on a finite set, we may choose an element c_o of X_o such that $\pi_{f_o}(c_o) = 1$. We denote by e_o an element of X_o such that $e_o \ne c_o$ and

$$\begin{aligned}\pi_{f_o}(e_o) &= \max\{\pi_{f_o}(x) : x \in X_o, x \ne c_o\} \\ &= \beta_{f_o}(c_o).\end{aligned}$$

Finally, we need the following sequences of real numbers. For any element $x = (x_o,\ldots,x_N) \in \times_{i=0}^N X_i$ such that $\pi_{(f_o\ldots,f_N)}(x) > 0$, we may define recursively a sequence of elements in $]0,1]$ as follows:

$$\lambda_{N,N}^x = \pi_{(f_o,\ldots,f_N)}(x)$$

and

$$\lambda_{k,N}^x = \frac{\lambda_{k+1,N}}{\pi_{f_{k+1}|f_o,\ldots,f_k}(x_{k+1} \mid (x_o,\ldots,x_k))}$$

for all $k \in \mathbb{N}$ such that $0 \le k < N$. This sequence has the following properties. If $k \in \mathbb{N}$ such that $0 \le k \le N$, then

$$\lambda_{k,N}^x \le \pi_{(f_o,\ldots,f_k)}(x_o,\ldots,x_k).$$

For any couple of natural numbers (k,l) such that $0 \le k < l \le N$ we have:

$$\begin{aligned}\pi_{(f_o,\ldots,f_k)}(x_o,\ldots,x_k) - \pi_{(f_o,\ldots,f_l)}(x_o,\ldots,x_l)& \\ \ge \lambda_{k,N}^x - \lambda_{l,N}^x&.\end{aligned}$$

Using the sequence $\lambda_{k,N}^x$, $k \in \{0,\ldots,N\}$, we may now assign to x a finitely additive probability measure P_x on $\wp(\times_{i=0}^N X_i)$ that is uniquely determined by the following conditions:

- $P_x(\{x\}) = \lambda_{N,N}^x$;
- for all natural numbers $k \in \{0,\ldots,N-1\}$ such that
$$\lambda_{k,N}^x - \lambda_{k+1,N}^x > 0$$
let $P_x(\{g_{k,n}(x_0,\ldots,x_k)\}) = \lambda_{k,N}^x - \lambda_{k+1,N}^x$;
- $P_x(\{g_{0,N}(c_o)\}) = 1 - \lambda_{0,N}^x$ if $x_o \ne c_o$;
- $P_x(\{g_{0,N}(e_o)\}) = 1 - \lambda_{0,N}^x$ if $x_o = c_o$, and
- $P_x(\{y\}) = 0$ in all other elements y of $\times_{i=0}^N X_i$.

It can be shown that the class

$$\mathcal{M} = \{P_x : x \in \times_{i=0}^N X_i \text{ and } \pi_{(f_o,\ldots,f_N)}(x) > 0\}$$

has properties (i) and (ii). As mentioned before, this ensures the coherence of the model we are dealing with. $\square$

Remark 3.3 When the possibilistic variables $f_o,\ldots,f_N$ additionally satisfy the Markov condition, i.e., for all $x = (x_o,\ldots,x_n) \in \times_{i=0}^n X_i$ and $y \in X_{n+1}$ where $n \in \{0,\ldots,N-1\}$:

$$_{\mathrm{DE}}\pi_{f_{n+1}|(f_o,\ldots,f_n)}(y \mid x) = {}_{\mathrm{DE}}\pi_{f_{n+1}|f_n}(y \mid x_n)$$

whenever $\pi_{(f_o,\ldots,f_n)}(x) > 0$, then $\pi_{(f_o,\ldots,f_N)}$ and $_{\mathrm{DE}}\pi_{f_{n+k}|f_n}(\cdot \mid x)$, $x \in X_n$, $(n,k) \in \mathbb{N} \times \mathbb{N} \setminus \{0\}$ such that $0 \le n < n+k \le N$, constitute a coherent model [6, 9, 10]. From this result we may conclude that a behavioural interpretation of possibilistic Markov processes [5, 6, 7, 8, 11] is meaningful, provided that all conditional possibilities are calculated by Dempster's rule. $\triangleleft$

4 Constructing coherent models by convex combinations

Mixtures of possibility distributions and probability distributions are more flexible than possibility or probability distributions. Due to the limitations imposed by possibility measures these models were suggested in [4, 14] as replacements to express second-order uncertainty, especially because they include both possibility and probability distributions as extreme cases. In this section we will show how a coherent model of conditional and unconditional upper probabilities can be constructed from an existing coherent model and a finitely additive probability. The method proposed here is to form convex combinations of the given unconditional upper probabilities and the given finitely additive probability. We will indicate how the conditional upper probabilities can be defined, so that constructed model preserves the coherence property. Naturally this method can be used to derive coherent models from the coherent models of conditional and unconditional possibilities, mentioned in Section 3 and in [15].

Consider a coherent model that consists of an unconditional upper probability $\overline{P}$ on the power set $\wp(\Omega)$ of a finite set Ω (the possibility space), and conditional upper probabilities $\overline{P}(\cdot|B)$, where B is an element of a partition $\mathcal{B}$ of Ω, all defined on a field $\mathcal{A}$ on Ω such that $\mathcal{B} \subseteq \mathcal{A}$. We will furthermore denote by $\underline{P}$ the lower probability on $\wp(\Omega)$ conjugate to $\overline{P}$, i.e., the lower probability given by:

$$\underline{P}(E) = 1 - \overline{P}(\Omega \setminus E), \quad \forall E \in \wp(\Omega).$$

As explained in Section 2 (see Theorem 2.1) we may characterise the coherence of this model by a non-empty class $\mathcal{M}$ of finitely additive probabilities, all defined on the power set of Ω, and such that

- (E_1) $\overline{P}$ is the upper envelope of $\mathcal{M}$, i.e., $\overline{P}(A) = \sup\{Q(A)\colon Q \in \mathcal{M}\}$ for all $A \in \wp(\Omega)$;
- (E_2) $\overline{P}(A|B) \geq \phi_{\mathcal{M}}(A|B)$ for all $(A, B) \in \mathcal{A} \times \mathcal{B}$, and the equality holds whenever $\overline{P}(\Omega \setminus B) < 1$.

For notational ease, we have used the notation $\phi_{\mathcal{M}}(A|B)$, $(A, B) \in \mathcal{A} \times \mathcal{B}$, in the formulation of condition (E_2) to denote the value

$$\phi_{\mathcal{M}}(A|B) = \sup\{\frac{Q(A \cap B)}{Q(B)}\colon Q \in \mathcal{M} \text{and} Q(B) > 0\}. \tag{6}$$

This means that $\phi_{\mathcal{M}}(\cdot|B)$ is the upper envelope of the conditional probabilities, calculated for a conditioning event $B \in \mathcal{B}$ by applying Bayes' rule to all elements Q of $\mathcal{M}$ such that $Q(B) > 0$.

Remark 4.1

- Note that $\phi_{\mathcal{M}}(A|B) = 0$ whenever $\overline{P}(B) = 0$.
- For any couple $(A, B) \in \mathcal{A} \times \mathcal{B}$:

  $$\phi_{\mathcal{M}}(A|B) > 0 \quad \text{iff} \quad \overline{P}(A \cap B) > 0.$$

- Consider a couple $(A, B) \in \mathcal{A} \times \mathcal{B}$. Taking into account (E_1) we have the following equality

  $$\overline{P}(A \cap B) = \sup\{Q(A \cap B)\colon Q \in \mathcal{M} \text{ and } Q(B) > 0\}. \tag{7}$$

  In case that $\overline{P}(B) = 0$, it is obvious that $\overline{P}(A \cap B) = 0$, hence $\overline{P}(B)\phi_{\mathcal{M}}(A|B) \geq \overline{P}(A \cap B)$. Let us now consider the case where $\overline{P}(B) > 0$. Consider a finitely additive probability $Q \in \mathcal{M}$ such that $Q(B) > 0$.Since $\phi_{\mathcal{M}}(A|B) \geq \frac{Q(A \cap B)}{Q(B)}$ by definition, we obtain, using (E_1) again, that

  $$\overline{P}(B)\phi_{\mathcal{M}}(A|B) \geq Q(B)\phi_{\mathcal{M}}(A|B) \quad \geq Q(A \cap B).$$

  Since the inequality above holds for all $Q \in \mathcal{M}$ such that $Q(B) > 0$, it follows from equality (7) that $\overline{P}(B)\phi_{\mathcal{M}}(A|B) \geq \overline{P}(A \cap B)$. This gives us the following lower bound for $\phi_{\mathcal{M}}(A|B)$:

  $$\frac{\overline{P}(A \cap B)}{\overline{P}(B)} \leq \phi_{\mathcal{M}}(A|B), \quad \text{if } \overline{P}(B) > 0. \tag{8}$$

- Since $\overline{P}$ is a coherent upper probability on $\wp(\Omega)$ satisfying (E_1), we have for any couple $(A, B) \in \mathcal{A} \times \mathcal{B}$ the following upper bound[3] for $\phi_{\mathcal{M}}(A|B)$:

  $$\phi_{\mathcal{M}}(A|B) \leq \frac{\overline{P}(A \cap B)}{\overline{P}(A \cap B) + \underline{P}(B \setminus A)}, \tag{9}$$

  provided that $\overline{P}(A \cap B) + \underline{P}(B \setminus A) > 0$. To see this, consider a finitely additive probability $Q \in \mathcal{M}$ such that $Q(B) > 0$. Since $\frac{Q(A \cap B)}{Q(B)} = \frac{Q(A \cap B)}{Q(A \cap B) + Q(B \setminus A)}$, $Q(A \cap B) \leq \overline{P}(A \cap B)$ and $Q(B \setminus A) \geq \underline{P}(B \setminus A)$, we find that

  $$\frac{Q(A \cap B)}{Q(B)} \leq \frac{\overline{P}(A \cap B)}{\overline{P}(A \cap B) + \underline{P}(B \setminus A)}.$$

  Taking into account (6) we obtain inequality (9).

$\triangleleft$

Let $\overline{P}_\alpha = \alpha P + (1 - \alpha)\overline{P}$ (where $\alpha \in]0, 1[$) be a convex combination of a finitely additive probability P on $\wp(\Omega)$ and the coherent upper probability $\overline{P}$. Various rules [15] can be used to calculate conditional upper probabilities from $\overline{P}_\alpha$. We will investigate here the coherence of the model constituted by $\overline{P}_\alpha$ and conditional upper probabilities $\overline{P}_\alpha(\cdot|B)$, $B \in \mathcal{B}$, on $\mathcal{A}$, defined as follows. Consider a couple $(A, B) \in \mathcal{A} \times \mathcal{B}$. If $P(B) = 0$, then

$$\overline{P}_\alpha(A|B) = \overline{P}(A|B). \tag{α_1}$$

If $P(B) > 0$, we define

$$\overline{P}_\alpha(A|B) = \frac{\alpha P(A \cap B) + (1 - \alpha)\overline{P}(A \cap B)}{\alpha P(B) + (1 - \alpha)\frac{\overline{P}(A \cap B)}{\phi_{\mathcal{M}}(A|B)}}, \tag{α_2}$$

provided that $\phi_{\mathcal{M}}(A|B) > 0$, otherwise

$$\overline{P}_\alpha(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{α_3}$$

Note that $\Omega \setminus B \in \mathcal{A}$ and $\overline{P}_\alpha(\Omega \setminus B|B) = 0$.

[3]See also Theorem 6.4.6 in [13] for the case where $\underline{P}(B) > 0$.

Remark 4.2 Imagine that Dempster's rule is used to calculate the conditional upper probabilities in the given model from $\overline{P}$. Then, the conditional upper probability ${}_{\text{DE}}\overline{P}(A|B)$, $(A, B) \in \mathcal{A} \times \mathcal{B}$ is explicitly given by

$$ {}_{\text{DE}}\overline{P}(A|B) = \begin{cases} \frac{\overline{P}(A \cap B)}{\overline{P}(B)} & \text{if } \overline{P}(B) > 0; \\ 1 & \text{if } \overline{P}(B) = 0, A \cap B \neq \emptyset; \\ 0 & \text{if } \overline{P}(B) = 0, A \cap B = \emptyset. \end{cases} $$

Taking into account inequality (8) in Remark 4.1, and condition (E_2) we find:

$$ \phi_{\mathcal{M}}(A|B) = {}_{\text{DE}}\overline{P}(A|B), \quad \text{whenever } \overline{P}(B) > 0. $$

This allows us to express $\overline{P}_\alpha(A|B)$, $(A, B) \in \mathcal{A} \times \mathcal{B}$ as follows:

$$ \overline{P}_\alpha(A|B) = \begin{cases} {}_{\text{DE}}\overline{P}(A|B) & \text{if } P(B) = 0; \\ \frac{\alpha P(A \cap B) + (1-\alpha)\overline{P}(A \cap B)}{\alpha P(B) + (1-\alpha)\overline{P}(B)} & \text{if } P(B) > 0, \end{cases} $$

or equivalently, $\overline{P}_\alpha(A|B)$ coincides with the conditional upper probability ${}_{\text{DE}}\overline{P}_\alpha(A|B)$, derived from $\overline{P}_\alpha$ by Dempster's rule, i.e.,

$$ \overline{P}_\alpha(A|B) = {}_{\text{DE}}\overline{P}_\alpha(A|B), $$

since $\overline{P}_\alpha(B) = 0$ iff $P(B) = 0 = \overline{P}(B)$.

Remark that ${}_{\text{DE}}\overline{P}_\alpha(A|B)$ may be expressed as a convex combination of the conditional upper probabilities ${}_{\text{DE}}P(A \mid B)$ and ${}_{\text{DE}}\overline{P}(A|B)$, derived from P and $\overline{P}$ using Dempster's rule, namely:

$$ {}_{\text{DE}}\overline{P}_\alpha(A|B) = \alpha \frac{P(B)}{\overline{P}_\alpha(B)} {}_{\text{DE}}P(A \mid B) + (1-\alpha) \frac{\overline{P}(B)}{\overline{P}_\alpha(B)} {}_{\text{DE}}\overline{P}(A|B). $$

From this representation it is immediately clear that the conditional upper probabilities $\overline{P}_\alpha(\cdot|B)$, $B \in \mathcal{B}$ are separately coherent[4]. $\triangleleft$

To ensure the coherence of the model constituted by $\overline{P}_\alpha$ and $(\overline{P}_\alpha(\cdot|B) \mid B \in \mathcal{B})$, where $\alpha \in]0, 1[$, we will now investigate whether or not the non-empty class $\mathcal{M}_\alpha$ of finitely additive probabilities on $\wp(\Omega)$, given by

$$ \mathcal{M}_\alpha = \alpha P + (1-\alpha)\mathcal{M} = \{\alpha P + (1-\alpha)Q : Q \in \mathcal{M}\}, $$

has properties (E_1) and (E_2). Then, if we can prove that all conditional upper probabilities $\overline{P}_\alpha(\cdot|B)$, $B \in \mathcal{B}$ are separately coherent, the coherence of the new model will follow directly from Theorem 2.1. Conditions (E_1) and (E_2) may be formulated as follows.

[4]This follows immediately from the Convexity Theorem (Theorem 2.6.4) in [13].

(E_1) $\overline{P}_\alpha$ is the upper envelope of $\mathcal{M}_\alpha$, i.e., $\overline{P}_\alpha(A) = \sup\{R(A) : R \in \mathcal{M}_\alpha\}$ for all $A \in \wp(\Omega)$;

(E_2) For all $(A, B) \in \mathcal{A} \times \mathcal{B}$

$$ \overline{P}_\alpha(A|B) \geq \phi_{\mathcal{M}_\alpha}(A|B) = \sup\{\frac{R(A \cap B)}{R(B)} : R \in \mathcal{M}_\alpha, R(B) > 0\} $$

and $\overline{P}_\alpha(A|B)$ is equal to $\phi_{\mathcal{M}_\alpha}(A|B)$ whenever $\overline{P}_\alpha(\Omega \setminus B) < 1$.

It is obvious that $\mathcal{M}_\alpha$ satisfies (E_1). To see this, note that $\mathcal{M}$ is a non-empty class of finitely additive probabilities on $\wp(\Omega)$ satisfying (E_1), i.e., $\overline{P}(A) = \sup\{Q(A) : Q \in \mathcal{M}\}$. Since every element of $\mathcal{M}_\alpha$ is a convex combination of P and an element of $\mathcal{M}$, it follows that $\mathcal{M}_\alpha$ is also a non-empty class of finitely additive probabilities on $\wp(\Omega)$ satisfying (E_1).

The following conditions are sufficient to prove (E_2):

(S_1) if $(A, B) \in \mathcal{A} \times \mathcal{B}$, then

$$ P(A \cap B) \leq \phi_{\mathcal{M}}(A|B)P(B); $$

(S_2) if $(A, B) \in \mathcal{A} \times \mathcal{B}$ such that $\overline{P}(A \cap B) > 0$, then there exists for any real number $\delta > 0$ a finitely additive probability $Q \in \mathcal{M}$ with the following properties:

$$ \begin{cases} \overline{P}(A \cap B) - \delta < Q(A \cap B); \\ \phi_{\mathcal{M}}(A|B) - \delta < \frac{Q(A \cap B)}{Q(B)}. \end{cases} $$

Both conditions are formulated in terms of the upper envelopes $\phi_{\mathcal{M}}(\cdot|B)$, $B \in \mathcal{B}$. The first condition imposes that the conditional probability $P(\cdot|B)$, obtained by conditioning P on $B \in \mathcal{B}$, should be dominated by $\phi_{\mathcal{M}}(\cdot|B)$, provided that $P(B) > 0$. From (S_1) it follows that $P(A \cap B) = 0$ whenever $\phi_{\mathcal{M}}(A|B) = 0$. If $P(B) > 0$ it follows from (α_3) that $\overline{P}_\alpha(A|B) = 0$. By the second condition $\mathcal{M}$ should allow to construct estimates for $\overline{P}(A \cap B)$ and $\phi_{\mathcal{M}}(A|B)$, $(A, B) \in \mathcal{A} \times \mathcal{B}$, based on a single element of $\mathcal{M}$, which are equally accurate as wanted.

Let us now consider two special cases in which condition (S_2) always hold.

Remark 4.3 Condition (S_2) is always fulfilled when all conditional upper probabilities in the given model are calculated from $\overline{P}$ by Dempster's rule. Consider therefore a couple $(A, B) \in \mathcal{A} \times \mathcal{B}$ such that $\overline{P}(A \cap B) > 0$. Let $\delta > 0$ such that $\delta < \min(\phi_{\mathcal{M}}(A|B), \overline{P}(A \cap B))$. From Remark 4.2 we know that $\phi_{\mathcal{M}}(A|B) = {}_{\text{DE}}\overline{P}(A|B)$. Since $\overline{P}$ is by assumption the upper envelope of $\mathcal{M}$, we can always find an element $Q \in \mathcal{M}$ such that $\overline{P}(A \cap B) - \delta\overline{P}(B) < Q(A \cap B)$, giving us the inequalities $\overline{P}(A \cap B) - \delta < Q(A \cap B)$ and ${}_{\text{DE}}\overline{P}(A|B) - \delta < \frac{Q(A \cap B)}{Q(B)}$. $\triangleleft$

Remark 4.4 Assume that $\overline{P}$ is a 2-alternating coherent upper probability on $\wp(\Omega)$. Suppose that $\mathcal{M}$ coincides with $\mathcal{M}(\overline{P})$, i.e. the class of all finitely additive probabilities that are dominated on $\wp(\Omega)$ by $\overline{P}$. Consider furthermore a couple $(A, B) \in \mathcal{A} \times \mathcal{B}$. It is well-known that there exists a finitely additive probability $Q \in \mathcal{M}(\overline{P})$ such that $Q(A \cap B) = \overline{P}(A \cap B)$ and $Q(B \setminus A) = \underline{P}(B \setminus A)$ (see for instance [12]). Taking into account (6) and (9) it follows that $\phi_{\mathcal{M}}(A|B) = \frac{Q(A\cap B)}{Q(B)} = \frac{\overline{P}(A\cap B)}{\overline{P}(A\cap B)+\underline{P}(B\setminus A)}$, provided that $\overline{P}(A \cap B) + \underline{P}(B \setminus A) > 0$. It can now easily be verified that (S_2) holds. ◁

The following result will be helpful to derive property (E_2) for $\mathcal{M}_\alpha$.

Lemma 4.5 *Let $\alpha \in \,]0,1[$ and let p_1, p_2, u_1, u_2, v_1, v_2 be real numbers such that $\alpha p_2 + (1-\alpha)u_2 > 0$ and $\alpha p_2 + (1-\alpha)v_2 > 0$. Assume that*

(L_o) $u_1 v_2 \le u_2 v_1$,

and that at least one of the following conditions holds:

(L_1) $p_2(v_1 - u_1) - p_1(v_2 - u_2) \ge 0$;

(L_2) $p_2 v_1 - p_1 v_2 \ge 0$, $u_1 \le v_1$, $v_1 > 0$, $v_2 > 0$ *and* $\alpha p_1 + (1-\alpha)v_1 \ge 0$.

Then

$$\frac{\alpha p_1 + (1-\alpha)u_1}{\alpha p_2 + (1-\alpha)u_2} \le \frac{\alpha p_1 + (1-\alpha)v_1}{\alpha p_2 + (1-\alpha)v_2}. \qquad (10)$$

Proof: Since $\alpha p_2 + (1-\alpha)u_2 > 0$ and $\alpha p_2 + (1-\alpha)v_2 > 0$, we may rewrite (10) as follows:

$$(\alpha p_1 + (1-\alpha)u_1)(\alpha p_2 + (1-\alpha)v_2) \le (\alpha p_1 + (1-\alpha)v_1)(\alpha p_2 + (1-\alpha)u_2).$$

By the assumption $\alpha \in \,]0,1[$ the foregoing inequality is equivalent to

$$\alpha p_2(v_1 - u_1) - \alpha p_1(v_2 - u_2) + (1-\alpha)(v_1 u_2 - u_1 v_2) \ge 0. \qquad (11)$$

The inequality above is satisfied when conditions (L_o) and (L_1) hold. Let us now verify (11) for the case that conditions (L_o) and (L_2) hold. Note that

$$\begin{aligned} v_1 u_2 - u_1 v_2 &= v_1 u_2 - v_1 v_2 + v_1 v_2 - u_1 v_2 \\ &= v_1(u_2 - v_2) + v_2(v_1 - u_1) \\ &= v_2(v_1 - u_1) - v_1(v_2 - u_2). \end{aligned}$$

Since both v_1 and v_2 are strictly positive real numbers, we may rewrite (11) as follows:

$$(\alpha\frac{p_2}{v_2} + (1-\alpha))v_2(v_1 - u_1) \ge (\alpha\frac{p_1}{v_1} + (1-\alpha))v_1(v_2 - u_2). \qquad (12)$$

From condition (L_o) it follows that

$$v_2(v_1 - u_1) = v_2 v_1 - v_2 u_1 \ge v_1 v_2 - v_1 u_2 = v_1(v_2 - u_2). \qquad (13)$$

Since $p_2 v_1 \ge p_1 v_2$, $v1 > 0$ and $v_2 > 0$ by condition (L_2), it follows that $\frac{p_2}{v_2} \ge \frac{p_1}{v_1}$, hence

$$\alpha\frac{p_2}{v_2} + (1-\alpha) \ge \alpha\frac{p_1}{v_1} + (1-\alpha). \qquad (14)$$

Note that the right-hand side in the foregoing inequality is non-negative. Inequality (12) follows directly from $v_2(v_1 - u_1) \ge 0$, (13) and (14). □

Let us now prove that $\mathcal{M}_\alpha$ satisfies (E_2) given that (S_1) and (S_2) hold. Consider therefore a couple of events $(A, B) \in \mathcal{A} \times \mathcal{B}$. Let us firstly consider the case where $P(B) = 0$. Consider an element R of $\mathcal{M}_\alpha$. By definition there exists an element Q of $\mathcal{M}$ such that $R = \alpha P + (1-\alpha)Q$. Assume that $R(B) > 0$. Note that this equivalent to $Q(B) > 0$. From the assumption $P(B) = 0$ it follows that $P(A \cap B) = 0$, and

$$\begin{aligned} \frac{R(A\cap B)}{R(B)} &= \frac{\alpha P(A\cap B) + (1-\alpha)Q(A\cap B)}{\alpha P(B) + (1-\alpha)Q(B)} \\ &= \frac{Q(A\cap B)}{Q(B)}. \end{aligned} \qquad (15)$$

Since $\mathcal{M}$ satisfies (E_2) with respect to the given model, we obtain that

$$\frac{R(A\cap B)}{R(B)} \le \overline{P}(A|B) = \overline{P}_\alpha(A|B).$$

Since $P(\Omega \setminus B) = 1 - P(B) = 1$ and $\alpha \in \,]0,1[$, it follows that $\overline{P}_\alpha(\Omega \setminus B) < 1$ iff $\overline{P}(\Omega \setminus B) < 1$. Consequently, we obtain the equality

$$\phi_{\mathcal{M}_\alpha}(A|B) = \overline{P}_\alpha(A|B),$$

provided that $\overline{P}_\alpha(\Omega \setminus B) < 1$. This follows immediately from (15) and the assumption that $\mathcal{M}$ satisfies (E_2) with respect to the given model.

We now proceed with the case where $P(B) > 0$ and $\phi_{\mathcal{M}}(A|B) = 0$. Consider an element R of $\mathcal{M}_\alpha$. By definition there exists an element Q of $\mathcal{M}$ such that $R = \alpha P + (1-\alpha)Q$. Note that $R(B) > 0$, independently of the element Q chosen from $\mathcal{M}$. Assume that $Q(B) = 0$. Then $Q(A \cap B) = 0$ and

$$\frac{R(A\cap B)}{R(B)} = \frac{P(A\cap B)}{P(B)} = \overline{P}_\alpha(A|B).$$

Suppose that $Q(B) > 0$. From (6) and the assumption $\phi_{\mathcal{M}}(A|B) = 0$ it follows that $Q(A \cap B) = 0$. Hence

$$\begin{aligned} \frac{R(A\cap B)}{R(B)} &= \frac{\alpha P(A\cap B)}{\alpha P(B) + (1-\alpha)Q(B)} \\ &\le \frac{P(A\cap B)}{P(B)} = \overline{P}_\alpha(A|B). \end{aligned}$$

In both cases we have that

$$\phi_{\mathcal{M}_\alpha}(A|B) = \frac{\alpha P(A\cap B)}{\alpha P(B) + (1-\alpha)\underline{P}(B)} \le \overline{P}_\alpha(A|B),$$

where $\underline{P}$ is the lower probability on $\wp(\Omega)$, conjugate to $\overline{P}$. From condition (S_1) we know that $\frac{P(A\cap B)}{P(B)} \le \phi_{\mathcal{M}}(A|B) = 0$, hence $P(A\cap B) = 0$. This implies that

$$\phi_{\mathcal{M}_\alpha}(A|B) = 0 = \frac{P(A\cap B)}{P(B)} = \overline{P}_\alpha(A|B).$$

Finally, assume that $P(B) > 0$ and $\phi_{\mathcal{M}}(A|B) > 0$. From Remark 4.1 we know that $\overline{P}(A\cap B) > 0$. Consider then an element $R \in \mathcal{M}_\alpha$. By definition there exists an element $Q \in \mathcal{M}$ such that $R = \alpha P + (1-\alpha)Q$. Note that $R(B) > 0$, since $P(B) > 0$. Then

$$Q(A\cap B)\frac{\overline{P}(A\cap B)}{\phi_{\mathcal{M}}(A|B)} \le Q(B)\overline{P}(A\cap B).$$

When $Q(B) > 0$, the inequality above follows immediately from the definition of $\phi_{\mathcal{M}}(A|B)$, the assumption that $\phi_{\mathcal{M}}(A|B) > 0$, and $Q \in \mathcal{M}$. If $Q(B) = 0$, then this inequality is actually an equality, since $Q(A\cap B) = 0$, due to the monotonicity of Q. Similarly, we have that

$$P(A\cap B)\frac{\overline{P}(A\cap B)}{\phi_{\mathcal{M}}(A|B)} \le P(B)\overline{P}(A\cap B).$$

The inequality now follows from (S_1) and the assumption that $\phi_{\mathcal{M}}(A|B) > 0$. Since $\overline{P}$ is the upper envelope of $\mathcal{M}$ ($\mathcal{M}$ satisfies (E_1)) and $Q \in \mathcal{M}$, we have $Q(A\cap B) \le \overline{P}(A\cap B)$. Using $\overline{P}(A\cap B) > 0$, $\phi_{\mathcal{M}}(A|B) > 0$, and Lemma 4.5 we obtain that

$$\begin{aligned}\frac{R(A\cap B)}{R(B)} &= \frac{\alpha P(A\cap B) + (1-\alpha)Q(A\cap B)}{\alpha P(B) + (1-\alpha)Q(B)}\\ &\le \frac{\alpha P(A\cap B) + (1-\alpha)\overline{P}(A\cap B)}{\alpha P(B) + (1-\alpha)\frac{\overline{P}(A\cap B)}{\phi_{\mathcal{M}}(A|B)}}\\ &= \overline{P}_\alpha(A|B).\end{aligned}$$

Consequently, we find that

$$\overline{P}_\alpha(A|B) \ge \phi_{\mathcal{M}_\alpha}(A|B).$$

Note that $\overline{P}_\alpha(B) > 0$ (this is a direct consequence of the assumption that $P(B) > 0$). To ensure that $\mathcal{M}_\alpha$ satisfies (E_2), we are obliged to prove that the equality holds in the expression above. Essentially, this will follow from condition (S_2). Consider a positive real number $\delta > 0$ such that $\delta < \min(\overline{P}(A\cap B), \phi_{\mathcal{M}}(A|B))$. Then, condition (S_2) guarantees the existence of an element $Q \in \mathcal{M}$ such that

$$\begin{cases}\phi_{\mathcal{M}}(A|B) - \delta < \frac{Q(A\cap B)}{Q(B)}\\ \overline{P}(A\cap B) - \delta < Q(A\cap B)\end{cases}$$

Assume that $P(A\cap B) > 0$. Without loss of generality we may assume that $\delta < \frac{P(A\cap B)}{P(B)}$. Using Lemma 4.5 and (S_1) we obtain that

$$\begin{aligned}&\frac{\alpha(P(A\cap B) - \delta P(B)) + (1-\alpha)(\overline{P}(A\cap B) - \delta)}{\alpha P(B) + (1-\alpha)\frac{\overline{P}A\cap B - \delta}{\phi_{\mathcal{M}}(A|B) - \delta}}\\ &\le \frac{\alpha(P(A\cap B) - \delta P(B)) + (1-\alpha)Q(A\cap B)}{\alpha P(B) + (1-\alpha)Q(B)}\\ &\le \frac{\alpha P(A\cap B) + (1-\alpha)Q(A\cap B)}{\alpha P(B) + (1-\alpha)Q(B)}\\ &\le \phi_{\mathcal{M}_\alpha}(A|B)\end{aligned}$$

Since the inequality above holds for all positive real numbers δ such that $\delta < \min(\frac{P(A\cap B)}{P(B)}, \phi_{\mathcal{M}}(A|B), \overline{P}(A\cap B))$, we obtain that

$$\begin{aligned}\overline{P}_\alpha(A|B) &= \frac{\alpha P(A\cap B) + (1-\alpha)(\overline{P}(A\cap B)}{\alpha P(B) + (1-\alpha)\frac{\overline{P}A\cap B}{\phi_{\mathcal{M}}(A|B)}}\\ &\le \phi_{\mathcal{M}_\alpha}(A|B).\end{aligned}$$

The case where $P(A\cap B) = 0$ can be proven in a similar way.

Let us recapitulate the results we have proven so far. The class of finitely additive probabilities $\mathcal{M}_\alpha$ has properties (E_1) and (E_2) with respect to the new model. As pointed out by Theorem 2.1 we only have to verify now that $\overline{P}_\alpha$ and the conditional upper probabilities $\overline{P}_\alpha(\cdot|B)$, $B \in \mathcal{B}$ are separately coherent. The following considerations may be used to prove this.

- $\overline{P}_\alpha$ is a convex combination of the coherent upper probability $\overline{P}$ and the (coherent) finitely additive probability P, both defined on $\wp(\Omega)$. Consequently, $\overline{P}_\alpha$ is a coherent upper probability on $\wp(\Omega)$.
- Consider an element B of the partition $\mathcal{B}$. If $P(B) = 0$, then $\overline{P}_\alpha(\cdot|B)$ is defined as the coherent upper probability $\overline{P}(\cdot|B)$. Suppose now that $P(B) > 0$. This implies that $P(\Omega\setminus B) = 1 - P(B) < 1$. Taking into account that $\alpha \in]0,1[$, we find:

  $$\begin{aligned}\overline{P}_\alpha(\Omega\setminus B) &= \alpha P(\Omega\setminus B) + (1-\alpha)\overline{P}(\Omega\setminus B)\\ &< \alpha + (1-\alpha)\overline{P}(\Omega\setminus B) \le 1.\end{aligned}$$

  Note that $R(B) > 0$ for all $R \in \mathcal{M}_\alpha$, due to $\alpha \in]0,1[$ and $P(B) > 0$. Since $\mathcal{M}_\alpha$ satisfies (E_2), we find for any $A \in \mathcal{A}$:

  $$\overline{P}_\alpha(A|B) = sup\{\frac{R(A\cap B)}{R(B)} : R \in \mathcal{M}_\alpha\}.$$

  Using $\mathcal{M}_\alpha \ne \emptyset$ we obtain that $\overline{P}_\alpha(\cdot|B)$ is coherent.

Taking into account Remark 4.2, Remark 4.3, and Theorem 2.1 we obtain the following result.

Theorem 4.6 *Consider some coherent model specified by an upper probability $\overline{P}$ on the power set of a finite set Ω and conditional upper probabilities $\overline{P}(\cdot|B)$, where B is an element of a partition $\mathcal{B}$ of Ω, all defined on a field $\mathcal{A}$ on Ω such that $\mathcal{B} \subseteq \mathcal{A}$. Assume that $\mathcal{M}$ is a non-empty class of finitely additive probabilities on $\wp(\Omega)$ satisfying (E_1) and (E_2) with respect to this model. Let $\phi_{\mathcal{M}}(A|B)$, $(A, B) \in \mathcal{A} \times \mathcal{B}$ be given by*

$$\phi_{\mathcal{M}}(A|B) = \sup\{\frac{Q(A \cap B)}{Q(B)} : Q \in \mathcal{M} \text{ and } Q(B) > 0\},$$

and let P be a finitely additive probability on $\wp(\Omega)$, satisfying (S_1) and (S_2). If $\alpha \in]0, 1[$, then the model constituted by the upper probability $\overline{P}_\alpha = \alpha P + (1 - \alpha)\overline{P}$ and the conditional upper probabilities $\overline{P}_\alpha(\cdot|B)$, $B \in \mathcal{B}$, all defined on $\mathcal{S}$ according to (α_1), (α_2) and (α_3), is coherent. When the conditional upper probabilities of the given model are calculated from $\overline{P}$ by Dempster's rule, then the conditional upper probabilities of the new model are calculated from $\overline{P}_\alpha$ by Dempster's rule.

It should be emphasised that the above theorem can be generalised to more general models, as presented in Section 2. The technique to generate coherent models, pointed out in this section, may be applied to the models of conditional and unconditional possibilities, that were proven to be coherent in Section 3 (see Theorem 3.2 and Remark 3.3, see also [15]). This means that mixtures of possibility measures and finitely additive probabilities may be conditioned coherently by various rules.

When the unconditional upper probability $\overline{P}$ of the given coherent model is 2-alternating, condition (S_1) does not have to imposed in order to obtain a coherent model. To make this more clear, let us reconsider the situation we have been dealing with in Remark 4.4. This means: $\overline{P}$ is 2-alternating coherent upper probability on $\wp(\Omega)$, $\mathcal{M} = \mathcal{M}(\overline{P})$ and condition (S_2) holds. Note that $\overline{P}_\alpha$, as a convex combination of P and $\overline{P}$, is also a 2-alternating coherent upper probability. The lower probability conjugate to $\overline{P}_\alpha$ is equal to $\underline{P}_\alpha = \alpha P + (1 - \alpha)\underline{P}$. Consider $(A, B) \in \mathcal{A} \times \mathcal{B}$. In a similar way as we did in Remark 4.4 we can prove, using (9), that

$$\phi_{\mathcal{M}_\alpha}(A|B) = \frac{\overline{P}_\alpha(A \cap B)}{\overline{P}_\alpha(A \cap B) + \underline{P}_\alpha(B \setminus A)},$$

provided that $\alpha P(B) + (1-\alpha)(\overline{P}(A \cap B) + \underline{P}(B \setminus A)) = \overline{P}_\alpha(A \cap B) + \underline{P}_\alpha(B \setminus A) > 0$. We may now define $\overline{P}_\alpha(\cdot|B)$, $B \in \mathcal{B}$ as follows: $\overline{P}_\alpha(\cdot|B)$ is equal to $\phi_{\mathcal{M}_\alpha}(\cdot|B)$ whenever $\underline{P}_\alpha(B) = \alpha P(B) + (1-\alpha)\underline{P}(B) > 0$, otherwise $\overline{P}_\alpha(\cdot|B)$ is equal to the vacuous conditional upper probability on $\mathcal{A}$. Obviously, the model constituted by these conditional upper probabilities and $\overline{P}_\alpha$ is coherent (on ground of Theorem 2.1 and $\mathcal{M}_\alpha$). When $B \in \mathcal{B}$ such that $P(B) > 0$, then $\overline{P}_\alpha(\cdot|B)$ is not necessarily defined as stated in (α_2), unless condition (S_1) holds.

5 Conclusion

We have characterised the coherence of a model of conditional and unconditional possibilities. We have shown how a coherent model of conditional and unconditional upper probabilities can be constructed by forming convex combinations of a finitely additive probability P and a non-empty class of finitely additive probilities $\mathcal{M}$ that is assumed to ensure the coherence of an existing model of conditional and unconditional upper probabilities. In order to derive tractable expressions for the imprecise probabilities of the new model (in terms of the imprecise probabilities of the existing model) two conditions on P and $\mathcal{M}$ were proposed. For the case where the unconditional upper probabilities of the given model are 2-alternating and the conditional upper probabilities are determined by the natural (or regular) extension rule, P can (dependent on $\mathcal{M}$) be taken arbitrarily. The results in this paper indicate that mixtures of possibility measures and finitely additive probabilities can be conditioned coherently by various rules.

Acknowledgements

I would like to thank the referees whose comments and suggestions have helped me to improve this paper. I would like to express my deep gratitude to Etienne E. Kerre and Gert de Cooman for their support and encouragement. I also wish to thank Anne Evenepoel for her support.

References

[1] G. de Cooman. *Evaluatieverzamelingen en -afbeeldingen - Een ordetheoretische benadering van vaagheid en onzekerheid [Evaluation Sets and Mappings - An Order-Theoretic Approach to Vagueness and Uncertainty].* PhD thesis, Universiteit Gent, Ghent, 1993.

[2] G. de Cooman. Possibility theory I: the measure- and integral-theoretic groundwork. *International Journal of General Systems*, 25:291–323, 1997.

[3] G. de Cooman and D. Aeyels. Supremum preserving upper probabilities. *Information Sciences*, 118:173–212, 1999.

[4] G. de Cooman and P. Walley. An imprecise hierarchical model for behaviour under uncertainty. 34 pages, submitted for publication in Theory and Decision, 1999.

[5] H. J. Janssen. Classification of states for a possibilistic Markov process. In *Proceedings of the Seventh European Congress on Intelligent Techniques & Soft Computing*, pages 302 – 303 (abstract), Aachen, Germany, 1999. Paper on CDROM.

[6] H. J. Janssen. *Een ordetheoretische en behavioristische studie van possibilistische processen [An Order-Theoretic and Behavioural Study of Possibilistic Processes].* PhD thesis, Universiteit Gent, Ghent, 1999.

[7] H. J. Janssen, G. de Cooman, and E. E. Kerre. First results for a mathematical theory of possibilistic Markov processes. In *Proceedings of the Sixth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, volume 3, pages 1425–1431, Granada, 1996.

[8] H. J. Janssen, G. de Cooman, and E. E. Kerre. Some remarks on stationary possibilistic processes. In *Proceedings of the Third International FLINS Workshop*, pages 52–60, Antwerp, Belgium, 1998. World Scientific.

[9] H. J. Janssen, G. de Cooman, and E. E. Kerre. Coherent models for discrete possibilistic systems. In *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications (ISIPTA '99)*, pages 189 – 195, Ghent, Belgium, 1999.

[10] H. J. Janssen, G. de Cooman, and E. E. Kerre. Coherence of dempster's conditioning rule in discrete possibilistic markov models. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8:241–252, 2000.

[11] Cliff A. Joslyn. *Possibilistic Processes for Complex Systems Modeling*. PhD thesis, State University of New York at Binghamton, 1994.

[12] P. Walley. Coherent lower (and upper) probabilities. Technical report, University of Warwick, Coventry (U.K.), 1981. Statistics Research Report 22.

[13] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[14] P. Walley. Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems*, 26:337 – 383, 1997.

[15] P. Walley and G. de Cooman. Coherence of rules for defining conditional possibilities. *International Journal of Approximate Reasoning*, 21:63–107, 1998.

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