1st International Symposium on Imprecise Probabilities and Their Applications, Ghent, Belgium, 29 June - 2 July 1999

Coherent Models for Discrete Possibilistic Systems

Hugo J. Janssen
Universiteit Gent
Krijgslaan 281–S9
9000 Gent, Belgium
hugo.janssen@rug.ac.be

Gert de Cooman
Universiteit Gent
Technologiepark – Zwijnaarde 9
9052 Zwijnaarde, Belgium
gert.decooman@rug.ac.be

Etienne E. Kerre
Universiteit Gent
Krijgslaan 281–S9
9000 Gent, Belgium
etienne.kerre@rug.ac.be

Abstract

We consider discrete possibilistic systems for which the available information is given by one-step transition possibilities and initial possibilities. These systems can be represented by a collection of variables satisfying a possibilistic counterpart of the Markov condition. This means that, given the values assumed by a selection of variables, the possibility that a subsequent variable assumes some value is only dependent on the value taken by the most recent variable of the selection. The one-step transition possibilities are recovered by computing the conditional possibility of any two consecutive variables. Under the behavioural interpretation as marginal betting rates against events these 'conditional' possibilities and the initial possibilities should satisfy the rationality criteria of 'avoiding sure loss' and 'coherence'. We show that this is indeed the case when the conditional possibilities are defined using Dempster's conditioning rule.

Keywords. Possibilistic Markov system, Markov condition, coherence, Dempster's conditioning rule.

1 Introduction

Possibility measures are supremum preserving set mappings. They were proposed by Zadeh [10] for modelling linguistic information in natural language. Formally, a *possibility measure* 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 $\Pi(\Omega) = \sup_{\omega\in\Omega} \pi(\omega) = 1$. The triple $(\Omega, \wp(\Omega), \Pi_\Omega)$ is called a *possibility space*.

Possibility measures can be given the behavioural interpretation of upper probabilities [8]. On this view, the value $\Pi(A)$, $A \in \wp(\Omega)$ 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 $\lambda > \Pi(A)$, 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 $\Pi(A)$ can also be written as:

$$G(A) = \Pi(A) - I_A,$$

where I_A is the indicator function of A. The behavioural interpretation of the possibility measure Π implies that all gambles $G(A)$, $A \in \wp(\Omega)$ 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 [8].

To ensure that the values $\Pi(A)$, $A \in \wp(\Omega)$ are assessed in a consistent way, Π should be coherent, i.e., for any natural number n, for any non-negative real numbers $\lambda_o, \dots, \lambda_n$ and for any events $A_o, \dots, A_n \in \wp(\Omega)$, 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) guarantees the subject avoids the incurrence of sure losses.

If (1) fails and $\lambda_o > 0$, then there is some $\delta > 0$ such that $[\Pi(A_o) - \delta] - I_{A_o} \geq \lambda_o^{-1} \sum_{j=1}^{n} \lambda_j [G(A_j) + \delta]$. Since all gambles $G(A_i) + \delta$, $i \in \{1, \dots, n\}$ 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 $[\Pi(A_o) - \delta] - I_{A_o}$,

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

as it yields a uniformly higher gain. As a result, the subject can be induced to bet against A_o at the rate $\Pi(A_o)-\delta$, which is strictly smaller than his *infimum* acceptable betting rate $\Pi(A_o)$. Coherence rules out this type of inconsistency.

For the special type of upper probabilities, namely possibility measures, we are dealing with here, coherence and avoiding sure loss both reduce to the requirement of normality. Thus a possibility measure Π is a coherent upper probability provided that it is normal [3]. We shall furthermore call a distribution π coherent when the possibility measure associated with π is coherent, or equivalently, when π is a normal distribution.

A model constituted by the assessments of a subject facing uncertainty may be more complex. In Section 2, for instance, we consider a discrete possibilistic system that is specified by one-step transition possibilities and initial possibilities. We explain that this system can be modelled by a collection of possibilistic variables. 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. We denote by Π_f the unique possibility measure on $\wp(X)$ with distribution π_f. The *joint possibility distribution function* of a finite sequence $f_o, \ldots, f_n$, $n \in \mathbb{N}$ of possibilistic variables, having basic space $(\Omega, \wp(\Omega), \Pi_\Omega)$ and sample spaces $X_o, \ldots, X_n$, is given in any element $(x_o, \ldots, x_n) \in \times_{i=o}^n X_i$ by

$$\pi_{(f_o,\ldots,f_n)}(x_o, \ldots, x_n) = \Pi_\Omega(\bigcap_{i=0}^{n} f_i^{-1}(\{x_i\})).$$

The possibility measure on $\wp(\times_{i=o}^n X_i)$ with distribution $\pi_{(f_o,\ldots,f_n)}$ is denoted by $\Pi_{(f_o,\ldots,f_n)}$. Using Dempster's conditioning rule we explain in Section 2 how the one-step transition possibilities can be recovered as, or interpreted as, the conditional possibilities of any two consecutive variables. In fact, we indicate that the variables also satisfy a possibilistic counterpart of the well-known Markov condition in the theory of stochastic Markov processes [4].

If we want to give a behavioural interpretation to the initial possibilities and the one-step transition possibilities, it is mandatory that we verify whether or not the models these possibilities are used to construct are coherent. It has been proven by Walley and De Cooman [9] that the unconditional joint possibility distribution function $\pi_{(f_o,f_1)}$ of any two possibilistic variables f_o and f_1 having finite sample spaces X_o and X_1 together with the conditional possibility distribution functions $\pi_{f_1|f_o}(\cdot \mid x_o)$, $x_o \in X_o$ — calculated by some conditioning rule — 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 $\pi_{f_1|f_o}(x_1 \mid x_o)$, $(x_o, x_1) \in X_o \times X_1$ satisfy

$${}_{\mathrm{DE}}\pi_{f_1|f_o}(x_1 \mid x_o) \leq \pi_{f_1|f_o}(x_1 \mid x_o) \leq {}_{\mathrm{NE}}\pi_{f_1|f_o}(x_1 \mid x_o) \tag{2}$$

whenever $\pi_{f_o}(x_o) > 0$. According to (2) the foregoing model is coherent if and only if the conditional possibilities are intermediate between those calculated by Dempster's rule and natural extension. Recall that for any couple of elements $(x_o, x_1) \in X_o \times X_1$, Dempster's conditioning rule yields the following value for the conditional possibility ${}_{\mathrm{DE}}\pi_{f_1|f_o}(x_1 \mid x_o)$:

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

where the least committal, or most conservative, value is taken for ${}_{\mathrm{DE}}\pi_{f_1|f_o}(x_1 \mid x_o)$ when $\pi_{f_o}(x_o) = 0$. We shall denote by ${}_{\mathrm{DE}}\Pi_{f_1|f_o}(\cdot \mid x_o)$ the possibility measure on $\wp(X_1)$ that is associated with the distribution ${}_{\mathrm{DE}}\pi_{f_1|f_o}(\cdot \mid x_o)$. As the natural extension rule for conditioning has no further role in this paper, we refer to [9] for its explicit definition.

In Section 3 we investigate the coherence of models with the following more general structure: the joint possibility distribution function of a finite collection of linearly ordered possibilistic variables $f_o, \ldots, f_N$, $N \in \mathbb{N} \setminus \{0\}$ together with the conditional possibility distribution functions of any variable f_{n+1}, $n \in \{0, \ldots, N-1\}$, given that the preceding variables $(f_o, \ldots, f_n)$ jointly assume some value $(x_o, \ldots, x_n) \in X_o \times \cdots \times X_n$. Here as well, all variables are assumed to have a finite sample space. We show that coherence is guaranteed if we additionally require that all conditional possibilities should be determined using Dempster's conditioning rule.

From this result we may conclude that a behavioural interpretation of the model in terms of possibilistic variables that we proposed in order to represent the previously introduced discrete possibilistic system makes sense, provided that all initial possibilities are normal and that the one-step transition possibilities are computed using Dempster's conditioning rule.

2 Discrete Possibilistic Systems

Suppose that we are dealing with a discrete possibilistic system having the set of all natural numbers $\mathbb{N}$ as its time set. $\mathbb{N}$ is taken to be ordered by the usual linear ordering $\leq$ of natural numbers.

Assume that we have the following information about the system:

- X_n, $n \in \mathbb{N}$ is the set of all possible states for the system at time n;

- *initial possibilities* $\overline{q}$, i.e., a $X_o - [0,1]$-mapping $\overline{q}$ such that $\overline{q}(x)$ is the possibility that the system is in state $x \in X_o$ at time 0;

- a $X_n \times X_{n+1} - [0,1]$-mapping ${}_n\overline{\mathbb{P}}$, $n \in \mathbb{N}$ such that, for any couple $(x,y) \in X_n \times X_{n+1}$, ${}_n\overline{\mathbb{P}}(x,y)$ denotes the *one-step transition possibility* from state x at time n to state y at time $n+1$, and that is normalized as follows:

$$\sup_{y \in X_{n+1}} {}_n\overline{\mathbb{P}}(x,y) = 1, \quad \forall x \in X_n.$$

Consequently, the partial mapping ${}_n\overline{\mathbb{P}}(x,\cdot)$ is the distribution of a unique, normal possibility measure on $(X_{n+1}, \wp(X_{n+1}))$ for every element $x \in X_n$ where $n \in \mathbb{N}$. The mapping $\overline{q}$ can be viewed as the distribution of a unique possibility measure $\overline{Q}$ on $(X_o, \wp(X_o))$.

Using this information we want to determine a consistent collection of distributions giving the possibility that the system visits a finite number of states $x_o, \dots, x_n$, $n \in \mathbb{N}$ at the corresponding times $0, \dots, n$. We are furthermore interested in determining the k-step transition possibilities of the system, where $k \geq 2$. We shall first give a number of formulae for these possibilities, and then show how the formulae can be justified.

We define the k-step transition possibility ${}_n\overline{\mathbb{P}}^{(k)}(x,y)$ from state $x \in X_n$ at time n to state $y \in X_{n+k}$ at time $n+k$ as:

$${}_n\overline{\mathbb{P}}^{(k)}(x,y) = \sup_{\substack{(z_n,\dots,z_{n+k}) \in \times_{i=n}^{n+k} X_i \\ z_n = x, z_{n+k} = y}} \prod_{j=n}^{n+k-1} {}_j\overline{\mathbb{P}}(z_j, z_{j+1}). \tag{3}$$

For $k = 1$ the above formula naturally simplifies to ${}_n\overline{\mathbb{P}}^{(1)}(x,y) = {}_n\overline{\mathbb{P}}(x,y)$. In a similar way we define the possibility $\pi_{\{0,\dots,n\}}(x_o,\dots,x_n)$ that some 'joint state' $x = (x_o,\dots,x_n)$, $n \in \mathbb{N}$ is assumed by the system at the corresponding times $0, \dots, n$ as:

$$\pi_{\{0,\dots,n\}}(x) = \begin{cases} \overline{q}(x_o) \prod_{j=0}^{n-1} {}_j\overline{\mathbb{P}}(x_j, x_{j+1}) & \text{if } n \geq 1 \\ \overline{q}(x) & \text{if } n = 0. \end{cases} \tag{4}$$

Obviously, $\pi_{\{0,\dots,n\}}$ can be considered as the distribution of a possibility measure on $(\times_{i=o}^{n} X_i, \wp(\times_{i=o}^{n} X_i))$.

By invoking our possibilistic Daniell-Kolmogorov theorem [6, 7], it is possible to construct a possibility space $(\Omega, \mathcal{R}_\Omega, \Pi_\Omega)$ and a family of possibilistic variables $(f_n \mid n \in \mathbb{N})$ with basic space $(\Omega, \mathcal{R}_\Omega, \Pi_\Omega)$, for which the corresponding sample spaces are given by X_n, $n \in \mathbb{N}$, such that

$$\pi_{(f_o,\dots,f_n)} = \pi_{\{0,\dots,n\}}, \quad \forall n \in \mathbb{N}. \tag{5}$$

To establish this, the following choices can be made:

- $(\Omega, \mathcal{R}_\Omega) = (\times_{i=o}^{+\infty} X_i, \wp(\times_{i=o}^{+\infty} X_i))$;
- for Π_Ω take the possibility measure $\Pi_\mathbb{N}$ on $(\times_{i=o}^{+\infty} X_i, \wp(\times_{i=o}^{+\infty} X_i))$ with distribution $\pi_\mathbb{N}$ whose value in a sequence $x = (x_o, \dots, x_n, \dots) \in \times_{i=o}^{+\infty} X_i$ is given by

$$\pi_\mathbb{N}(x) = \inf_{n \in \mathbb{N}} \pi_{\{0,\dots,n\}}(x_o,\dots,x_n)$$

Note that $\pi_\mathbb{N}$ is the pointwise greatest (least committal or most conservative) distribution on $\times_{i=o}^{+\infty} X_i$ whose marginal on $\times_{i=o}^{n} X_i$ is $\pi_{\{0,\dots,n\}}$ for all $n \in \mathbb{N}$.

We now give a justification for the formulae (3)–(5). Any collection $(f_n \mid n \in \mathbb{N})$ of possibilistic variables representing the information $\pi_{\{0,\dots,n\}}$, $n \in \mathbb{N}$ as expressed by (3)–(5) satisfies a possibilistic analogon of the Markov condition [4, 5]. To establish this result, we use Dempster's conditioning rule.

The following, obvious relation then holds between the transition possibilities and the conditional possibilities, formed with the possibilistic variables in the collection $(f_n \mid n \in \mathbb{N})$. Consider two natural numbers n and $k \neq 0$ and let $(x,y) \in X_n \times X_{n+k}$, then it follows from (3)–(5) that

$${}_{\mathrm{DE}}\pi_{f_{n+k}|f_n}(y \mid x) = {}_n\overline{\mathbb{P}}^{(k)}(x,y) \quad \text{if } \pi_{f_n}(x) > 0. \tag{6}$$

The possibilistic variables $(f_n \mid n \in \mathbb{N})$ are furthermore conditionally independent in the following way. Consider a finite subset $\{n_i \mid i \in \{1,\dots,k\}\}$ of the time set $\mathbb{N}$ such that $k \in \mathbb{N} \setminus \{0\}$ and $n_1 < \dots < n_k$. Let $n \in \mathbb{N}$ such that $n_k < n$. If $x = (x_{n_1},\dots,x_{n_k}) \in \times_{i=1}^{k} X_{n_i}$ and $y \in X_n$, then

$${}_{\mathrm{DE}}\pi_{f_n|(f_{n_1},\dots,f_{n_k})}(y \mid x) = {}_{\mathrm{DE}}\pi_{f_n|f_{n_k}}(y \mid x_{n_k}), \tag{M}$$

provided that $\pi_{(f_{n_1},\dots,f_{n_k})}(x) > 0$. Condition (M) can be regarded as a possibilistic analogon of the Markov condition [4]. In [5] we used condition (M) as a starting point for the development of a formal, measure-theoretic account of possibilistic Markov families (processes), i.e., families of possibilistic variables satisfying property (M).

Families of possibilistic variables satisfying condition (M) also satisfy an analogon of the Chapman-Kolmogorov equation [4].

3 Consistency Criteria for Unconditional and Conditional Possibilities

In the behavioural theory of imprecise probabilities two rationality criteria have a central part: avoiding sure loss and coherence. If we want to give a behavioural interpretation to the previously introduced initial possibilities, transition possibilities, etc., it is mandatory that we verify whether or not these criteria are satisfied for the models in terms of possibilistic variables constructed from these possibilities. This is the problem discussed in the present section.

Consider a finite collection of possibilistic variables f_o, $\dots$, f_N, $N \in \mathbb{N} \setminus \{0\}$. Let X_o, $\dots$, X_N be their corresponding sets of possible values. Assume that all the sets X_o, $\dots$, X_N are finite. For notational ease we denote the Cartesian product $\times_{i=o}^{N} X_i$ by $\mathcal{X}$.

Let $\pi_{(f_o,\dots,f_N)}$ be the joint possibility distribution function of the variables f_o, $\dots$, f_N, and let $\Pi_{(f_o,\dots,f_N)}$ be the possibility measure on $\wp(\mathcal{X})$ generated by $\pi_{(f_o,\dots,f_N)}$, that is,

$$\Pi_{(f_o,\dots,f_N)}(A) = \max_{x \in A} \pi_{(f_o,\dots,f_N)}(x), \quad \forall A \in \wp(\mathcal{X}).$$

It will be assumed that $\Pi_{(f_o,\dots,f_N)}$ is normal. Consequently, $\Pi_{(f_o,\dots,f_N)}$ and all marginals that can be derived from $\Pi_{(f_o,\dots,f_N)}$ are coherent upper probabilities. For any subset A of $\mathcal{X}$ the value $\Pi_{(f_o,\dots,f_N)}(A)$ may be interpreted as a subject's marginally acceptable upper rate for betting against A. The net reward resulting from such bet is given by

$$G(A) = \Pi_{(f_o,\dots,f_N)}(A) - I_A,$$

where I_A is the indicator function of A.

Consider now two non-empty, disjoint subsets of $\{0,\dots,N\}$ specified as follows:

$$\{n_i \mid i \in \{1,\dots,k\}\}$$

where $k \in \mathbb{N} \setminus \{0\}$ such that $0 \leq n_1 < \dots < n_k \leq N$, and

$$\{m_j \mid j \in \{1,\dots,\ell\}\}$$

where $\ell \in \mathbb{N} \setminus \{0\}$ such that $0 \leq m_1 < \dots < m_\ell \leq N$. The possibility distribution function of $g = (f_{n_1},\dots,f_{n_k})$ is given by the marginal $\pi_{(f_{n_1},\dots,f_{n_k})}$ of $\pi_{(f_o,\dots,f_N)}$. Similarly, the possibility distribution function of $h = (f_{m_1},\dots,f_{m_\ell})$ is given by the marginal $\pi_{(f_{m_1},\dots,f_{m_\ell})}$ of $\pi_{(f_o,\dots,f_N)}$. For notational ease we denote the Cartesian products $\times_{i=1}^{k} X_{n_i}$ and $\times_{j=1}^{\ell} X_{m_j}$ by $\mathcal{X}_g$ and $\mathcal{X}_h$, since they are the domains of the possibility distribution functions of g and h.

For all $y \in \mathcal{X}_h$ write $\pi_{g|h}(\cdot \mid y)$ for the conditional possibility distribution function of g given that h assumes the value y — calculated by some conditioning rule from information contained in $\pi_{(f_o,\dots,f_N)}$. Let $\Pi_{g|h}(\cdot \mid y)$ be the possibility measure on $\wp(\mathcal{X}_g)$ generated by $\pi_{g|h}(\cdot \mid y)$, i.e., for all $B \in \wp(\mathcal{X}_g)$:

$$\Pi_{g|h}(B \mid y) = \max\{\pi_{g|h}(x \mid y) \mid x \in B\}. \tag{7}$$

Consider a subset B of $\mathcal{X}_g$ and an element $y = (y_{m_1},\dots,y_{m_\ell})$ of $\mathcal{X}_h$. Two interpretations may be given to $\Pi_{g|h}(B \mid y)$ [8]. Under the *updating* interpretation $\Pi_{g|h}(B \mid y)$ is the marginally acceptable upper rate for betting against B that a subject would adopt after learning that $h = y$. Under the *contingent* interpretation $\Pi_{g|h}(B \mid y)$ is the marginally acceptable upper rate for betting against B contingent on $h = y$, i.e., the betting is called off unless $h = y$. Under both interpretations the net reward is the gamble

$$G_{g|h}(B \mid y) = I_{C_y}[\Pi_{g|h}(B \mid y) - I_{C_B}]$$

on $\mathcal{X}$, where

$$C_y = \{(x_o,\dots,x_N) \in \mathcal{X} \colon (x_{m_1},\dots,x_{m_\ell}) = y\};$$
$$C_B = \{(x_o,\dots,x_N) \in \mathcal{X} \colon (x_{n_1},\dots,x_{n_k}) \in B\},$$

and I_{C_y} and I_{C_B} are the indicator functions of C_y and C_B. By the *Updating Principle* [8], $\Pi_{g|h}(B \mid y)$ should have the same value under both interpretations.

Similarly to what we did before for $\Pi_{(f_o,\dots,f_N)}$, we now require that all conditional possibility distribution functions $\pi_{g|h}(\cdot \mid y)$, $y \in \mathcal{X}_h$ should be normal. On the updating interpretation this requirement ensures that the new possibility distribution function $\pi_{g|h}(\cdot \mid y)$ the subject would adopt if he learned only that h assumes the value y avoids sure loss. Actually, as we already explained in the Introduction, normality of a distribution is a sufficient and a necessary requirement for the corresponding possibility measure to avoid sure loss, and to be coherent. Let us now interpret $\Pi_{g|h}(B \mid y)$, $(B,y) \in \wp(\mathcal{X}_g) \times \mathcal{X}_h$, as contingent conditional possibilities. Suppose that $\pi_{g|h}(\cdot \mid y)$ is not normal for some value $y \in \mathcal{X}_h$. Then any bet $I_{C_y}[\mu - I_{\mathcal{X}}]$ against the sure event $\mathcal{X}_g$ contingent on $h = y$, at a rate μ such that $\Pi_{g|h}(\mathcal{X}_g \mid y) < \mu < 1$, is acceptable. When h assumes the value y, such bet produces a sure loss of $1-\mu$, and otherwise it is called off. To avoid the acceptance of such bets, we have to require again that all distributions $\pi_{g|h}(\cdot \mid y)$, $y \in \mathcal{X}_h$ should be normal.

Assume that (g_1,h_1), $\dots$, (g_s,h_s) where $s \in \mathbb{N} \setminus \{0\}$ are couples of possibilistic variables that are determined in a similar way as the possibilistic variables (g,h) above. We have already argued why we want the possibility measures $\Pi_{(f_o,\dots,f_N)}$ and $\Pi_{g_r|h_r}(\cdot \mid y)$, $y \in \mathcal{X}_{h_r}$, $r \in \{1,\dots,s\}$ — or the distributions $\pi_{(f_o,\dots,f_N)}$ and $\pi_{g_r|h_r}(\cdot \mid y)$, $y \in \mathcal{X}_{h_r}$, $r \in \{1,\dots,s\}$ — to be normal: this guarantees that considered separately, these models avoid sure loss and are coherent. We now introduce additional rationality requirements to be imposed on $\pi_{(f_o,\dots,f_N)}$ and $\pi_{g_r|h_r}(\cdot \mid y)$, $y \in \mathcal{X}_{h_r}$, $r \in \{1,\dots,s\}$, which guarantee the *mutual* consistency of these distributions [8].

First of all, $\pi_{(f_o,\dots,f_N)}$ and $\pi_{g_r|h_r}(\cdot \mid y)$, $y \in \mathcal{X}_{h_r}$, $r \in \{1,\dots,s\}$ *avoid sure loss* if for all non-negative functions λ on $\wp(\mathcal{X})$ and for all non-negative functions μ_r, $r \in \{1,\dots,s\}$ on $\wp(\mathcal{X}_{g_r}) \times \mathcal{X}_{h_r}$, there is an element $x \in \mathcal{X}$ such that

$$\sum_{A \in \wp(\mathcal{X})} \lambda(A)[\Pi_{(f_o,\dots,f_N)}(A) - I_A(x)] + \sum_{r=1}^{s} \sum_{\substack{B \in \wp(\mathcal{X}_{g_r}) \\ y \in \mathcal{X}_{h_r}}} \mu_r(B,y) G_{g_r|h_r}(B \mid y)(x) \geq 0. \quad (8)$$

Secondly, $\pi_{(f_o,\dots,f_N)}$ and $\pi_{g_r|h_r}(\cdot \mid y)$, $y \in \mathcal{X}_{h_r}$, $r \in \{1,\dots,s\}$ are *coherent* if for all non-negative functions λ on $\wp(X)$ and for all non-negative functions μ_r, $r \in \{1,\dots,s\}$ on $\wp(\mathcal{X}_{g_r}) \times \mathcal{X}_{h_r}$:

(i) for all $C \in \wp(\mathcal{X})$, there is an element $x \in \mathcal{X}$ such that

$$\sum_{A \in \wp(\mathcal{X})} \lambda(A)[\Pi_{(f_o,\dots,f_N)}(A) - I_A(x)] + \sum_{r=1}^{s} \sum_{\substack{B \in \wp(\mathcal{X}_{g_r}) \\ y \in \mathcal{X}_{h_r}}} \mu_r(B,y) G_{g_r|h_r}(B \mid y)(x) \geq \Pi_{(f_o,\dots,f_N)}(C) - I_C(x); \quad (9)$$

(ii) for all $(D,z) \in \wp(\mathcal{X}_{g_t}) \times \mathcal{X}_{h_t}$ where $t \in \{1,\dots,s\}$, there is an element $x \in \mathcal{X}$ such that

$$\sum_{A \in \wp(\mathcal{X})} \lambda(A)[\Pi_{(f_o,\dots,f_N)}(A) - I_A(x)] + \sum_{r=1}^{s} \sum_{\substack{B \in \wp(\mathcal{X}_{g_r}) \\ y \in \mathcal{X}_{h_r}}} \mu_r(B,y) G_{g_r|h_r}(B \mid y)(x) \geq G_{g_t|h_t}(D \mid z)(x). \quad (10)$$

For the case of two variables, i.e., $N = 1$, Walley and De Cooman have proven the following result [9].

Theorem 3.1. *Suppose that the conditioning rule satisfies for all $(x_o, x_1) \in X_o \times X_1$ the following condition:*

$$\pi_{(f_o,f_1)}(x_o,x_1) = 1 \Rightarrow \pi_{f_1|f_o}(x_1 \mid x_o) = 1. \quad (11)$$

Then $\pi_{(f_o,f_1)}$ and $\pi_{f_1|f_o}(\cdot \mid x_o)$, $x_o \in X_o$ avoid sure loss. Provided the rule satisfies (11) *and the analogous condition with x_o and x_1 interchanged, $\pi_{(f_o,f_1)}$, $\pi_{f_1|f_o}(\cdot \mid x_o)$, $x_o \in X_o$, and $\pi_{f_o|f_1}(\cdot \mid x_1)$, $x_1 \in X_1$ avoid sure loss.*

Moreover, $\pi_{(f_o,f_1)}$ and $\pi_{f_1|f_o}(\cdot \mid x_o)$, $x_o \in X_o$ are coherent if and only if the conditional possibilities $\pi_{f_1|f_o}(x_1 \mid x_o)$, $(x_o,x_1) \in X_o \times X_1$ satisfy

$$_{\mathrm{DE}}\pi_{f_1|f_o}(x_1 \mid x_o) \leq \pi_{f_1|f_o}(x_1 \mid x_o) \leq {_{\mathrm{NE}}\pi_{f_1|f_o}}(x_1 \mid x_o) \quad (12)$$

whenever $\pi_{f_o}(x_o) > 0$.

As already explained in the Introduction, in (12) $_{\mathrm{NE}}\pi_{f_1|f_o}(x_1 \mid x_o)$ denotes the value that is produced by the technique of natural extension for the conditional possibility that f_1 assumes the value x_1 given that f_o is equal to x_o.

Walley and De Cooman's result can be generalised for a finite number of possibilistic variables $f_o,\dots,f_N$, $N \in \mathbb{N}\setminus\{0\}$, provided that Dempster's conditioning rule is used to compute the conditional possibilities.

Theorem 3.2. *The distributions $\pi_{(f_o,\dots,f_N)}$ and $_{\mathrm{DE}}\pi_{f_{n+1}|(f_o,\dots,f_n)}(\cdot \mid x)$, $x \in \times_{i=o}^{n} X_i$, $n \in \{0,\dots,N-1\}$ are coherent. Moreover, if the possibilistic variables $f_o,\dots,f_N$ satisfy the Markov condition, i.e., for all $x = (x_o,\dots,x_n) \in \times_{i=o}^{n+1} X_i$ and $y \in X_{n+1}$ where $n \in \{0,\dots,N-1\}$:*

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

whenever $\pi_{(f_o,\dots,f_n)}(x) > 0$, then $\pi_{(f_o,\dots,f_N)}$ 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 $0 \leq n < n+k \leq N$ are coherent.

Sketch of the proof. For the coherence of the model formed by $\pi_{(f_o,\dots,f_N)}$ and $_{\mathrm{DE}}\pi_{f_{n+1}|(f_o,\dots,f_n)}(\cdot \mid x)$, $x \in \times_{i=o}^{n} X_i$, $n \in \{0,\dots,N-1\}$ it is necessary and sufficient that there is a non-empty class $\mathcal{M}$ of finitely additive probability measures on $\wp(\mathcal{X})$ such that

(a) $\Pi_{(f_o,\dots,f_N)}$ is the upper envelope of $\mathcal{M}$, i.e., for all $A \in \wp(\mathcal{X})$:

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

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

$$_{\mathrm{DE}}\Pi_{f_{n+1}|(f_o,\dots,f_n)}(A \mid x) \geq \sup\{\frac{P(C_{x,A})}{P(C_x)} \mid P \in \mathcal{M}_x\} \quad (13)$$

where

$$\begin{aligned} C_x &= \{(y_o,\dots,y_N) \in \mathcal{X} : (y_o,\dots,y_n) = x\}; \\ C_A &= \{(y_o,\dots,y_N) \in \mathcal{X} : y_{n+1} \in A\}; \\ C_{x,A} &= C_x \cap C_A; \\ \mathcal{M}_x &= \{P \mid P \in \mathcal{M} \text{ such that } P(C_x) > 0\}, \end{aligned}$$

and the equality in (13) holds whenever $\Pi_{(f_o,\dots,f_N)}(\mathcal{X} \setminus C_x) < 1$.

For the model formed by $\pi_{(f_o,\dots,f_N)}$ 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 $0 \leq n < n+k \leq N$, a sufficient and necessary condition for coherence now lies in the existence of a non-empty class $\mathcal{M}$ of finitely additive probability measures on $\wp(\mathcal{X})$ such that:

(a) $\Pi_{(f_o,\ldots,f_N)}$ is the upper envelope of $\mathcal{M}$, i.e., for all $A \in \wp(\mathcal{X})$:

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

(b') for all $(n,k) \in \mathbb{N}\times\mathbb{N}\setminus\{0\}$ such that $0 \leq n < n+k \leq N$, for all $x \in X_n$, for all $A \in \wp(X_{n+k})$:

$$_{\text{DE}}\Pi_{f_{n+k}|f_n}(A \mid x) \geq \sup\{\frac{P(C_{x,A})}{P(C_x)} \mid P \in \mathcal{M}_x\} \tag{14}$$

where

$$\begin{aligned} C_x &= \{(y_o,\ldots,y_N) \in \mathcal{X} \colon y_n = x\};\\ C_A &= \{(y_o,\ldots,y_N) \in \mathcal{X} \colon y_{n+k} \in A\};\\ C_{x,A} &= C_x \cap C_A;\\ \mathcal{M}_x &= \{P \mid P \in \mathcal{M} \text{ such that } P(C_x) > 0\}, \end{aligned}$$

and the equality in (14) holds whenever $\Pi_{(f_o,\ldots,f_N)}(\mathcal{X} \setminus C_x) < 1$.

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

$$\begin{cases} y_i & = x_i, \quad i \in \{0,\ldots,n\}\\ \pi_{(f_o,\ldots,f_{n+1})}(y) & = \pi_{(f_o,\ldots,f_n)}(x). \end{cases} \tag{15}$$

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

- for any couple $(k,n) \in \mathbb{N}^2$ such that $0 \leq k \leq n \leq N$ 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 $(k,l,n) \in \mathbb{N}^3$ such that $0 \leq k \leq l \leq n$:

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

Finally, consider an element c_o of X_o such that $\pi_{f_o}(c_o) = 1$. Such an element can always be choosen since π_{f_o} is assumed to be a normal possibility distribution function on a finite set.

To each element $x = (x_o,\ldots,x_N) \in \mathcal{X}$ we may assign a finitely additive probability measure P_x on $\wp(\mathcal{X})$ that is uniquely determined by the following conditions:

- $P_x(\{x\}) = \pi_{(f_o,\ldots,f_N)}(x)$;
- for all natural numbers $k \in \{0,\ldots,N-1\}$ such that $\pi_{(f_o,\ldots,f_k)}(x_o,\ldots,x_k) > \pi_{(f_o,\ldots,f_{k+1})}(x_o,\ldots,x_{k+1})$:

$$\begin{aligned} &P_x(\{g_{k,n}(x_0,\ldots,x_k)\})\\ &\quad= \pi_{(f_o,\ldots,f_k)}(x_o,\ldots,x_k)\\ &\quad\quad - \pi_{(f_o,\ldots,f_{k+1})}(x_o,\ldots,x_{k+1}); \end{aligned}$$

- $P_x(\{g_{0,N}(c_o)\}) = 1 - \pi_{f_o}(x_o)$ if $\pi_{f_o}(x_o) < 1$;
- $P_x(\{y\}) = 0$ in any other element y of $\mathcal{X}$.

It can be shown that the class $\mathcal{M} = \{P_x \mid x \in \mathcal{X}\}$ has properties (a) and (b). Consequently, $\pi_{(f_o,\ldots,f_N)}$ and $_{\text{DE}}\pi_{f_{n+1}|(f_o,\ldots,f_n)}(\cdot \mid x)$, $x \in \times_{i=o}^{n} X_i$, $n \in \{0,\ldots,N-1\}$ constitute a coherent model. When the variables $f_o,\ldots,f_N$ have the Markov property (M'), then $\mathcal{M}$ also satisfies (b'). To see this, take into consideration that, for any element $x = (x_o,\ldots,x_n) \in \times_{i=0}^{n} X_i$, the $n+2$-th component of $g_{n,n+1}(x)$ only depends on the $n+1$-th component x_n of the given element x. This follows from the normality of all conditional possibility distribution functions and the following formula that can now be written down for the joint possibility distribution function of the variables $f_o,\ldots,f_n$, $n \in \{0,\ldots,N\}$:

$$\pi_{(f_o,\ldots,f_n)}(x) = \pi_{f_o}(x_o) \prod_{j=o}^{n-1} \pi_{f_{j+1}|f_j}(x_{j+1} \mid x_j)$$

where $x \in \times_{i=0}^{n} X_i$. $\square$

4 Conclusion

The results in this paper point towards two interesting conclusions. First of all, they show by means of a concrete example that it is possible to work with imprecise probabilities in modelling Markov processes. Secondly, they indicate that the Dempster conditioning rule is of special importance in a specifically possibilistic context, for two reasons: (i) as Theorems 3.1 and 3.2 indicate, it is the most specific (or least conservative, or most committal) conditioning rule that is coherent in the context of possibilistic Markov processes; and (ii) it is very easy to work with, which makes a possibilistic theory of Markov processes computationally tractable.

Acknowledgements

Gert de Cooman is a Postdoctoral Fellow of the Fund for Scientific Research – Flanders (Belgium) (FWO). He would like to thank the FWO for partially funding the research reported in this paper.

The authors are grateful to Peter Walley and two anonymous referees, whose numerous suggestions helped us make this paper more readable.

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. 27 pages, accepted for publication in *Information Sciences*, 1998.

[4] J. L. Doob. *Stochastic Processes.* John Wiley & Sons, New York, 1967.

[5] H. J. Janssen, G. de Cooman, and E. E. Kerre. First results for a mathematical theory of possibilistic Markov processes. In *Proceedings of IPMU '96 (Information Processing and Management of Uncertainty in Knowledge Based Systems, Granada, Spain, July 1-5, 1996)*, volume III, pages 1425–1431, 1996.

[6] H. J. Janssen, G. de Cooman, and E. E. Kerre. Some remarks on stationary possibilistic processes. In *Prooceedings of the 3rd International FLINS Workshop (Antwerp, Belgium, September 14-16, 1998)*, pages 52–60, 1998.

[7] H. J. Janssen, G. de Cooman, and E. E. Kerre. A Daniell-Kolmogorov theorem for supremum preserving upper probabilities. *Fuzzy Sets and Systems*, 102:429–444, 1999.

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

[9] P. Walley and G. de Cooman. Coherence of rules for defining conditional possibilities. 29 pages, accepted for publication in: *International Journal of Approximate Reasoning*, 1998.

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