Kybernetika

Ivan Kramosil

Alternative definitions of conditional possibilistic measures

Kybernetika, Vol. 34 (1998), No. 2, [137]--147

Persistent URL: `http://dml.cz/dmlcz/135193`

KYBERNETIKA — VOLUME 34 (1998), NUMBER 2, PAGES 137–147

ALTERNATIVE DEFINITIONS OF CONDITIONAL POSSIBILISTIC MEASURES[1]

IVAN KRAMOSIL

The aim of this paper is to survey and discuss, very briefly, some ways how to introduce, within the framework of possibilistic measures, a notion analogous to that of conditional probability measure in probability theory. The adjective "analogous" in the last sentence is to mean that the conditional possibilistic measures should play the role of a mathematical tool to actualize one's degrees of beliefs expressed by an a priori possibilistic measure, having obtained some further information concerning the decision problem under uncertainty in question. The properties and qualities of various approaches to conditionalizing can be estimated from various points of view. Here we apply the idea according to which the properties of independence relations defined by particular conditional possibilistic measures are confronted with those satisfied by the relation of statistical (or stochastical) independence descending from the notion of conditional probability measure. For the reader's convenience the notions of conditional probability and statistical independence are recalled in the introductory chapter.

1. CONDITIONAL PROBABILITIES

In our days, probability theory represents the most developed mathematical tool for uncertainty quantification and processing, and the other mathematical models suggested for the same sakes like fuzzy sets, possibility theory, Dempster–Shafer theory, etc., can use probability theory and its approaches as a source of inspiration for their own development, but must also confront their own ideas and results with those introduced in and achieved by the probability theory. The notion of conditioning is a good example of both these aspects.

The basic notion of axiomatic (Kolmogorov) probability theory in its most abstract setting is that of *probability space.* Probability space is a triple $\langle \Omega, \mathcal{A}, P \rangle$ where Ω is a nonempty set the elements of which are called *elementary random events*, $\mathcal{A}$ is a nonempty σ-field of subsets of Ω the elements of which are called *measurable sets* or, in our context, *random events*, and P is a σ-additive *probability measure* on $\mathcal{A}$, i.e., $P : \mathcal{A} \to \langle 0, 1 \rangle$ is such that $P(\Omega) = 1$ and $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ for each sequence $\langle A_1, A_2, \ldots \rangle$ of mutually disjoint sets from $\mathcal{A}$.

[1]This paper has been supported by the grant No. A 1030803 of Grant Agency of the Academy of Sciences of the Czech Republic.

If the space Ω is finite or countable, i.e. $\Omega = \{\omega_1, \omega_2, \ldots\}$, then a probability measure P on the σ-field $\mathcal{A} = \mathcal{P}(\Omega)$ of all subsets of Ω is completely defined by the sequence $p_1 = P(\{\omega_1\})$, $p_2 = P(\{\omega_2\}), \ldots$ of probabilities ascribed to elementary random events. Supposing that each p_i quantifies a subject's degree of belief that ω_i will be the result of a random experiment (or the subject's betting rate to this result), and supposing the subject receives a reliable information that the result will belong to a (proper, to avoid trivialities) subset A of Ω, how should the subject actualize her/his probabilities $p_1, p_2, \ldots$ into a new sequence $q_1, q_2, \ldots$? At least within the framework of classical probability theory and mathematical statistics the following three demands are taken as reasonable.

(i) $q_i = 0$ for all i such that $\omega_i \in \Omega - A$;

(ii) $q_1, q_2, \ldots$ should be, again, a probability distribution on Ω, i.e. $q_i \in \langle 0, 1\rangle$ for each i and $\sum_{i=1}^{\infty} q_i = 1$ should hold;

(iii) the ratio of any two actualized probabilities of elementary random events from A should be the same as before the actualization, i.e., $q_i/q_j = p_i/p_j$ should hold for each i, j such that $\{\omega_i, \omega_j\} \subset A$ and $p_j \neq 0$.

These demands already imply that $q_i = p_i \Big| \sum_{\omega_j \in A} p_j$, if $\omega_i \in A$ and $\sum_{\omega_j \in A} p_j > 0$, $q_i = 0$ for i such that $\omega_i \in \Omega - A$, and q_i is undefined for all i, if $\sum_{\omega_j \in A} p_j = 0$. According to this idea, if $\langle \Omega, \mathcal{A}, P\rangle$ is a probability space and $B \in \mathcal{A}$ is such that $P(B) > 0$ holds, then the *conditional probability measure* $P(\cdot|B)$ on $\mathcal{A}$ is defined by $P(A|B) = P(A \cap B)/P(B)$ for each $A \in \mathcal{A}$.

As a matter of fact, this definition results as a very special case from the most general definition of *conditional expected value*. Let $X : \Omega \to R = (-\infty, \infty)$ be a *real-valued random variable* defined on a probability space $\langle \Omega, \mathcal{A}, P\rangle$, hence $\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$ holds for each Borel subset B of R (as a matter of fact, $\{\omega \in \Omega : X(\Omega) < x\} \in \mathcal{A}$ for each $x \in R$ is a sufficient condition). Let $\mathcal{B} \subset \mathcal{A}$ be a nonempty σ-field of subsets of Ω, let $P_{\mathcal{B}}$ be the restriction of P to $\mathcal{B}$. Due to the well-known Radon–Nikodym theorem there exists a $\mathcal{B}$-*measurable* function $E^{\mathcal{B}}X : \Omega \to R$ such that, for each $B \in \mathcal{B}$

$$\int_B (E^{\mathcal{B}}X)\,\mathrm{d}P_{\mathcal{B}} = \int_B X\,\mathrm{d}P. \tag{1}$$

$E^{\mathcal{B}}X$ is defined uniquely up to $\mathcal{B}$-measurable zero sets, i.e., if $E_1^{\mathcal{B}}X$ and $E_2^{\mathcal{B}}X$ both satisfy (1), then $A = \{\omega \in \Omega : E_1^{\mathcal{B}}X \neq E_2^{\mathcal{B}}X\} \in \mathcal{B}$ and $P(A) = P_{\mathcal{B}}(A) = 0$.

Two particular cases are of importance. If $\mathcal{B} = \{\emptyset, \Omega\}$, i.e., if $\mathcal{B}$ is the minimal nonempty σ-field of subsets of Ω, then $E^{\mathcal{B}}X$ is a constant value on whole the Ω identical with the usual notion of expected value of the random variable X, e.g., if $X(\omega) = x_i$ with the probability p_i, $i = 1, 2, \ldots$, then $E^{\mathcal{B}}X = \sum_{i=1}^{\infty} x_i p_i$ for $\mathcal{B} = \{\emptyset, \mathcal{B}\}$ supposing that this value is defined. If $\mathcal{B} = \{\emptyset, B, \Omega - B, \Omega\}$ for some $B \in \mathcal{A}, B \neq \emptyset$, $B \neq \Omega$, and if X is the *characteristic function* or *identifier* of some random event $A \in \mathcal{A}$, so that $X(\Omega) = 1$, if $\omega \in A$ and $X(\omega) = 0$ for $\omega \in \Omega - A$, then $E^{\mathcal{B}}(X) = P(A \cap B)/P(B)$ for $\omega \in A$ and $E^{\mathcal{B}}(X) = P((\Omega - A) \cap B)/P(B)$ for $\omega \in \Omega - A$, hence, we have arrived at the elementary definition of conditional probability (again, if $P(B) = 0$, $E^{\mathcal{B}}(X)$ is not defined). The function $E^{\mathcal{B}}X$ defined by (1) is called *conditional expected value of the random variable X given (or: with respect to) the σ-field $\mathcal{B}$.* Cf. e. g. [7] for the abstract axiomatic approach to conditional probabilities and random variables.

As far as the author knows, there is no paradigma of actualization within the framework of probability theory alternative to the conditioning model as briefly outlined above; the reason consists probably in the fact that each such an alternative approach would violate some of the principles (i)–(iii) above. Among the models violating (iii) perhaps the most interesting and worth considering is the Lewis' concept of imaging, another form of conditioning (cf. [6] and [10] for more detail). As we intend to classify various conceptions of conditional possibilistic measures among themselves as well as with the conditional probabilities through the notion of statistical (stochastical) independence generated by these measures, let us recall the definition of the last notion and its main properties.

Let X, Y, Z be real-valued random variables defined on a probability space $\langle \Omega, \mathcal{A}, P \rangle$. Random variables X, Y are called *statistically* (or *stochastically*) *independent*, if for all Borel sets B_1, $B_2 \subset (-\infty, \infty)$ the equality

$$\begin{aligned} & P(\{\omega \in \Omega : X(\omega) \in B_1, Y(\omega) \in B_2\}) \\ = \; & P(\{\omega \in \Omega : X(\omega) \in B_1\}) \, P(\{\omega \in \Omega : Y(\omega) \in B_2\}) \end{aligned} \tag{2}$$

holds. The relation of statistical (stochastical) conditional independence will be defined, for the sake of simplicity, only for the case of discrete random variables X, Y, Z taking their values in a finite or countable subset of the real line (cf. [8]), even if the most general definition in the terms of certain subalgebras of the σ-field $\mathcal{A}$ is also possible. Random variables X and Y are *statistically (stochastically) conditionally independent given the random variable Z*, if for each x, y, $z \in (-\infty, \infty)$ for which the conditional probabilities in question are defined, the equality

$$\begin{aligned} & P\left(\{\omega \in \Omega : X(\omega) = x\} | \{\omega \in \Omega : Y(\omega) = y, Z(\omega) = z\}\right) \\ = \; & P\left(\{\omega \in \Omega : X(\omega) = x\} | \{\omega \in \Omega : Y(\omega) = y\}\right) \end{aligned} \tag{3}$$

holds. Let us briefly denote this case by $I(X, Y | Z)$.

Let U be a finite set of random variables defined on $\langle \Omega, \mathcal{A}, P \rangle$, let $\mathcal{X} = \{X_1, X_2, \dots, X_n\}$, $\mathcal{Y} = \{Y_1, Y_2, \dots, Y_m\}$ and $\mathcal{Z} = \{Z_1, Z_2, \dots, Z_k\}$ be subsets of U, let us write $I(\mathcal{X}, \mathcal{Y}, | \mathcal{Z})$ if

$$I(\langle X_1, \dots, X_n \rangle, \langle Y, \dots, Y_m \rangle | \langle Z_1, \dots, Z_k \rangle) \tag{4}$$

holds for the corresponding vector variables $\langle X_1, \dots, X_n \rangle$, $\langle Y_1, \dots, Y_m \rangle$ and $\langle Z_1, \dots, Z_k \rangle$.

Among the properties possessed by the ternary relation $I(\mathcal{X}, \mathcal{Y} | \mathcal{Z})$ we shall pick up the five following ones (here and below we follow [1])

(A1)	Symmetry:	$(\forall \mathcal{X},\mathcal{Y},\mathcal{Z}\subset U)\,(I(\mathcal{X},\mathcal{Y}\vert\mathcal{Z})\Rightarrow I(\mathcal{Y},\mathcal{X}\vert\mathcal{Z}))$
(A2)	Decomposition:	$(\forall \mathcal{X},\mathcal{Y},\mathcal{W},\mathcal{Z}\subset U)\,(I(\mathcal{X},\mathcal{Y}\cup\mathcal{W}\vert\mathcal{Z})\Rightarrow I(\mathcal{X},\mathcal{Y}\vert\mathcal{Z}))$
(A3)	Weak Union:	$(\forall \mathcal{X},\mathcal{Y},\mathcal{W},\mathcal{Z}\subset U)\,(I(\mathcal{X},\mathcal{Y}\cup\mathcal{W}\vert\mathcal{Z})\Rightarrow I(\mathcal{X},\mathcal{W}\vert\mathcal{Z}\cup\mathcal{Y}))$
(A4)	Contraction:	$(\forall \mathcal{X},\mathcal{Y},\mathcal{W},\mathcal{Z}\subset U)\,((I(\mathcal{X},\mathcal{Y}\vert\mathcal{Z}))$ and $I(\mathcal{X},\mathcal{W}\vert\mathcal{Z}\cup\mathcal{Y})\Rightarrow I(\mathcal{X},\mathcal{Y}\cup\mathcal{W}\vert\mathcal{Z}))$
(A5)	Intersection:	$(\forall \mathcal{X},\mathcal{Y},\mathcal{W},\mathcal{Z}\subset U)\,((I(\mathcal{X},\mathcal{Y}\vert\mathcal{Z}\cup\mathcal{W})$ and $I(\mathcal{X},\mathcal{W}\vert\mathcal{Z}\cup\mathcal{Y}))\Rightarrow I(\mathcal{X},\mathcal{Y}\cup\mathcal{W}\vert\mathcal{Z}))$.

Fact 1. Let U be a finite set of random variables defined on a probability space $\langle\Omega,\mathcal{A},P\rangle$ and taking values in finite subsets of $R=(-\infty,\infty)$. Let $I\subset\mathcal{P}(U)\times\mathcal{P}(U)\times\mathcal{P}(U)$ be the ternary relation defined by (4). Then I satisfies (A1)–(A4). If all the random variables take every of their possible values with a positive probability, then I satisfies also (A5) (cf. [8]).

In the next chapter we shall define several conceptions of conditional possibilistic measures and we shall survey which of the properties (A1)–(A5) are fulfilled by the corresponding independence relations.

2. CONDITIONAL POSSIBILISTIC MEASURES

Let Ω be a nonempty set. A mapping $\pi:\Omega\to\langle 0,1\rangle$ is called *possibilistic* (or: *possibility*) *distribution* over Ω, and the mapping $\Pi:\mathcal{P}(\Omega)\to\langle 0,1\rangle$ (here $\mathcal{P}(\Omega)=\{A: A\subset\Omega\}$) defined by $\Pi(A)=\sup\{\pi(\omega):\omega\in A\}$ for each $A\subset\Omega$ is called the *possibilistic* (or: *possibility*) measure induced by π on Ω. Obviously, (i) $\Pi(A\cup B)=\max\{\Pi(A),\Pi(B)\}$ and (ii) $\Pi(\bigcup_{A\in\mathcal{A}}A)=\sup\{\Pi(A):A\in\mathcal{A}\}$ hold for each A, $B\subset\Omega$ and for each nonempty system $\mathcal{A}$ of subsets of Ω (by convention, $\Pi(\emptyset)=0$ for the empty set $\emptyset$). An alternative approach defines possibility measures directly as mappings $\Pi:\mathcal{P}(\Omega)\to\langle 0,1\rangle$ satisfying (i) for all A, $B\subset\Omega$. Then we can define possibilistic distribution π on Ω by $\pi(\omega)=\Pi(\{\omega\})$ for all $\omega\in\Omega$. If Ω is finite, we arrive back at Π, as $\sup\{\pi(\omega):\omega\in A\}=\Pi(A)$ now follows trivially for all $A\subset\Omega$. If Ω is infinite, the direct definition of possibilistic measure can be modified either by a restriction of its domain from $\mathcal{P}(\Omega)$ to an appropriate subsystem $\mathcal{S}\subset\mathcal{P}(\Omega)$ (a nonempty σ-field, say), or by replacing the condition (i) by (ii) holding for all nonempty subsystems $\mathcal{A}\subset\mathcal{P}(\Omega)$ (or $\mathcal{A}\subset\mathcal{S}$) with $\mathrm{card}(\mathcal{A})\le c$ for some given cardinal number c (often $c=\aleph_0$). Possibilistic measures were introduced by Zadeh in [11], cf. also [2] as a good introduction into the domain. In what follows we shall suppose that Ω is finite and that $\Pi(\Omega)=\sup\{\pi(\omega):\omega\in\Omega\}=1$, hence, there exists $\omega\in\Omega$ such that $\pi(\omega)=\Pi(\{\omega\})=1$.

Let Ω_1, Ω_2 be two nonempty sets, let ω denote elements of Ω_1, let η denote elements of Ω_2, and let $\pi:\Omega_1\times\Omega_2\to\langle 0,1\rangle$ be a possibilistic distribution on the Cartesian product $\Omega_1\times\Omega_2$. Let us define the *marginal possibilistic distributions* π_1 on Ω_1 and π_2 on Ω_2 by

$$\pi_1(\omega)=\max\{\pi(\omega,\eta):\eta\in\Omega_2\},\quad \pi_2(\eta)=\max\{\pi(\omega,\eta):\omega\in\Omega_1\}. \tag{5}$$

The Dempster-rule based idea of conditional possibilistic measure analogous to that one suggested by Shafer in [9]:

$$\pi_d(\omega|\eta) = \frac{\pi(\omega,\eta)}{\pi_2(\eta)} = \frac{\pi(\omega,\eta)}{\max\{\pi(\omega,\eta) : \omega \in \Omega_1\}}, \quad \text{if } \pi_2(\eta) > 0. \tag{6}$$

Here and below we suppose that $U = \{x_1, x_2, \dots, x_n\}$, where each x_i is a variable taking its values in a finite set Ω_i. Let π be an n-dimensional possibilistic distribution on the Cartesian product $\times_{i=1}^n \Omega_i$. If $X \subset U$, then $\boldsymbol{x}$ is a value from $\times_j \Omega_{i_j}$, $x_{i_j} \in X$, similarly for $Y \subset U$ and $\boldsymbol{y}$, and $Z \subset U$ and $\boldsymbol{z}$.

Fact 2. Let X, Y, Z be three disjoint subsets of U, let $I(X,Y|Z)$ mean that $\pi_d(\boldsymbol{x}|\boldsymbol{y}\,\boldsymbol{z}) = \pi_d(\boldsymbol{x}|\boldsymbol{z})$ for all values $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ such that $\pi(\boldsymbol{y}\,\boldsymbol{z}) \neq 0$. Then (A1)–(A4) hold. If π is strictly positive, i.e. $\pi(\boldsymbol{u}) > 0$ for all $\boldsymbol{u} \in \times_{i=1}^n \Omega_i$, then (A5) holds as well.

It is possible to argue that this definition is too strict as it requests the equality between the two conditional possibility measures (the same objection can be imposed also to the notion of statistical independence). An alternative idea reads: the supplementary knowledge of the value $\boldsymbol{y}$ cannot improve our knowledge of $\boldsymbol{x}$ given $\boldsymbol{z}$, but can *deteriorate it*, i.e., some information can be lost. In symbols

$$I(X,Y|Z) \Leftrightarrow_{\mathrm{df}} \pi_d(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) \geq \pi_d(\boldsymbol{x}|\boldsymbol{z}) \tag{7}$$

for all $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ such that $\pi(\boldsymbol{y},\boldsymbol{z}) \neq 0$. It is perhaps worth being explicitly noticed here, that "large" value of Π means a "small" knowledge as far as the "true" or "actual" element of Ω is concerned.

Fact 3. Let $I(X,Y|Z)$ be defined by (7). Then (A1), (A2) and (A4) hold, however, (A3) (Weak Union) and (A5) (Intersection) are not fulfilled. It is perhaps worth being explicitly noticed here, that "large" value of Π means a "small" knowledge as far as the "true" or "actual" element of Ω is concerned.

A modification of π_d, suggested in [1] and denoted by π_{d_c}, reads as follows.

$$\pi_{d_c}(\boldsymbol{x}|\boldsymbol{y}) = \begin{cases} \pi(\boldsymbol{x}), \text{ if } \pi(\boldsymbol{x},\boldsymbol{y}) \geq \pi(\boldsymbol{x})\,\pi(\boldsymbol{y}) \text{ holds for all } \boldsymbol{x}, \\ \pi_d(\boldsymbol{x}|\boldsymbol{y}), \text{ if there exists } \boldsymbol{x}' \text{ such that } \pi(\boldsymbol{x}',\boldsymbol{y}) < \pi(\boldsymbol{x}')\,\pi(\boldsymbol{y}). \end{cases} \tag{8}$$

Fact 4. Let $I(X,Y|Z)$ be defined by

$$I(X,Y|Z) \Leftrightarrow_{\mathrm{df}} \pi_{d_c}(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) = \pi_{d_c}(\boldsymbol{x}|\boldsymbol{z}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}. \tag{9}$$

Then I satisfies the properties (A2)–(A5) (also the distributions which are not strictly positive satisfy (A5)), but (A1) (Symmetry) does not hold.

Another approach is to define X and Y as conditionally independent given Z, if the possibility distributions $\pi_d(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z})$ and $\pi_d(\boldsymbol{x}|\boldsymbol{z})$ are similar or close to each other in a sense. So, if $\simeq$ is a binary relation in the set of possibility distributions defined on X, we could define $I(X,Y|Z)$ in this way

$$I(X,Y|Z) \Leftrightarrow_{\mathrm{df}} \pi_d(\cdot|\boldsymbol{y},\boldsymbol{z}) \simeq \pi_d(\cdot|\boldsymbol{z}) \quad \text{for all } \boldsymbol{y},\ \boldsymbol{z} \text{ such that } \pi(\boldsymbol{y},\boldsymbol{z}) \neq 0. \tag{10}$$

The three following particular cases of the similarity relationship $\simeq$ are worth considering.

(i) The qualitative, ordinal and ordering preserving relation:
$$\pi \simeq \pi' \Leftrightarrow_{\mathrm{df}} (\forall\, \boldsymbol{x},\boldsymbol{x}')\,[\pi(\boldsymbol{x}) < \pi(\boldsymbol{x}') \Leftrightarrow \pi'(\boldsymbol{x}) < \pi'(\boldsymbol{x}')] \tag{11}$$

(ii) distance-based relation: let $m \in \mathcal{N}^+$ and $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_{m-1} < \alpha_m = 1$ be given, let $I_k = \langle \alpha_{k-1}, \alpha_k)$ for $k = 1,2,\ldots,m-1$, let $I_m = \langle \alpha_{m-1}, \alpha_m \rangle$. Let
$$\pi \simeq \pi' \Leftrightarrow_{\mathrm{df}} (\forall\, \boldsymbol{x})\,(\exists\, k,\ 1 \leq k \leq m)\,(\pi(\boldsymbol{x}) \in I_k,\ \pi'(\boldsymbol{x}) \in I_k). \tag{12}$$

(iii) α-cuts-based relation: two possibility distributions are taken as similar, if they possibly differ only in the values smaller than a threshold $\alpha_0 \in \langle 0,1 \rangle$. In symbols,
$$\pi \simeq \pi' \Leftrightarrow_{\mathrm{df}} (\forall\, \alpha \geq \alpha_0)\,(\forall\, \boldsymbol{x})\,(\pi(\boldsymbol{x}) \geq \alpha \Leftrightarrow \pi'(\boldsymbol{x}) \geq \alpha). \tag{13}$$

Fact 5. Let I be defined by (10) with $\simeq$ defined by one of the relations (11), (12) or (13). Then (A2) – (A4) hold, (A5) holds for strictly positive distributions, but (A1) does not hold.

The problem can be formulated also in an "inverse" way: which properties the relation $\simeq$ must possess in order to be sure that the resulting (i. e., by (10) defined) independence relation I possesses a (or some) given property (properties) from (A1) – (A5)? This way of reasoning is briefly discussed in [1], but we shall not go here into detail.

E. Hisdal in [3] proposes an alternative definition of conditional possibility measure π_h, which can be seen as "qualitative", if compared with π_d. Under the same notation as used in (5) and (6), set

$$\pi_h(\boldsymbol{x}|\boldsymbol{y}) = \begin{cases} \pi(\boldsymbol{x},\boldsymbol{y}), & \text{if } \pi(\boldsymbol{x},\boldsymbol{y}) < \pi(\boldsymbol{y}), \\ 1, & \text{if } \pi(\boldsymbol{x},\boldsymbol{y}) = \pi(\boldsymbol{y}), \text{ both for all } \boldsymbol{x}. \end{cases} \tag{14}$$

Fact 6. Let $I(X,Y|Z)$ be defined by

$$I(X,Y|Z) \Leftrightarrow_{\mathrm{df}} \pi_h(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) = \pi_h(\boldsymbol{x}|\boldsymbol{z}) \quad \text{for all } \boldsymbol{x},\ \boldsymbol{y},\ \boldsymbol{z}. \tag{15}$$

Then (A2) – (A5) hold, but (A1) does not hold. Setting

$$I'(X,Y|Z) \Leftrightarrow_{\mathrm{df}} I(X,Y|Z) \text{ and } I(Y,X|Z), \tag{16}$$

we obtain that the relation I' satisfies (A1) – (A5).

(Let us remark that the notion of independence defined by I' is rather restrictive, e. g., $I'(X,Y|Z)$ implies either $(\forall\, \boldsymbol{x})\,(\pi(\boldsymbol{x}) = 1)$ or $(\forall\, \boldsymbol{y})\,(\pi(\boldsymbol{y}) = 1)$).

Using the same idea as in the case of π_d above, we can weaken the definition (15), setting.

$$I(X,Y|Z) \Leftrightarrow_{\mathrm{df}} \pi_h(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) \geq \pi_h(\boldsymbol{x}|\boldsymbol{z}) \quad \text{for all } \boldsymbol{x},\ \boldsymbol{y},\ \boldsymbol{z}. \tag{17}$$

Fact 7. Let $I(X,Y|Z)$ be defined by (17). Then (A1)–(A4) hold, but (A5) does not hold. Moreover, the definition (17) turns to be equivalent to

$$I(X,Y|Z) \Leftrightarrow \pi(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = \pi(\boldsymbol{x},\boldsymbol{z}) \wedge \pi(\boldsymbol{y},\boldsymbol{z}) \quad \text{for all } \boldsymbol{x},\ \boldsymbol{y},\ \boldsymbol{z}. \tag{18}$$

If we consider the particular case of $Z = \emptyset$, (18) implies that

$$I(X,Y|\emptyset) \Leftrightarrow_{\mathrm{df}} \pi(\boldsymbol{x},\boldsymbol{y}) = \pi(\boldsymbol{x}) \wedge \pi(\boldsymbol{y}) \quad \text{for all } \boldsymbol{x},\ \boldsymbol{y}, \tag{19}$$

and this is nothing else than the notion of non-interactivity for possibility measures and fuzzy sets.

As in the case of π_d investigated above, we can try to modify the definition of $I(X,Y|Z)$, given by (15), by replacing the strict equality in (15) by a similarity relation $\approx$. So, using the fact that both the definitions (15) and (18) are equivalent, we can define

$$I(X,Y|Z) \Leftrightarrow_{\mathrm{df}} \pi(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \approx \pi(\boldsymbol{x},\boldsymbol{y}) \wedge \pi(\boldsymbol{x},\boldsymbol{z}). \tag{20}$$

Fact 8. Let $I(X,Y|Z)$ be defined by (20), let $\approx$ be an equivalence relation compatible with the marginalization and combination of possibility distributions using the minimum as the combination operator. Then (A1)–(A4) hold, but (A5) does not hold, in general.

Again as above, we can replace π_h by π_{h_c} defined in this way.

$$\pi_{h_c}(\boldsymbol{x}|\boldsymbol{y}) = \begin{cases} \pi(\boldsymbol{x}), & \text{if } \pi_h(\boldsymbol{x}|\boldsymbol{y}) \geq \pi(\boldsymbol{x}) \text{ for all } \boldsymbol{x}, \\ \pi_h(\boldsymbol{x}|\boldsymbol{y}), & \text{if there exists } \boldsymbol{x}' \text{ such that } \pi_h(\boldsymbol{x}'|\boldsymbol{y}) < \pi(\boldsymbol{x}'). \end{cases} \tag{21}$$

Now, set

$$I(X,Y|Z) \Leftrightarrow_{\mathrm{df}} \pi_{h_c}(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) = \pi_{h_c}(\boldsymbol{x}|\boldsymbol{z}). \tag{22}$$

Fact 9. Let $I(X,Y|Z)$ be defined by (22). Then (A2)–(A4) hold, but (A1) (Symmetry) and (A5) do not hold, in general.

Let us remark that the way of conditioning defined by (14) or (21) is based only on the comparison of the corresponding possibility degrees, no computations like in (6) being possible. Hence, we can replace the $\langle 0,1\rangle$ interval as the space of possibility values by a finite set $\mathcal{L} = \{L_0, L_1, \ldots, L_m\}$ of some non-numerical objects ordered by a total ordering relation $\preceq$ in such a way that $L_0 \preceq L_1 \preceq L_2 \preceq \cdots \preceq L_m$ holds. If $\Pi : \mathcal{P}(D_x) \to \mathcal{L}$ is such that $\Pi(D_x) = L_m$ and $\Pi(A \cup B) = \bigvee_{\preceq}(\Pi(A), \Pi(B))$ for all A, $B \subset D_x$, where $\bigvee_{\preceq}$ is the supremum operation generated by $\preceq$ in $\mathcal{L}$, then conditioning and independence can be defined as in the case of π_h and the same results are valid. In what follows, we shall investigate a particular case when the space $\langle 0,1\rangle$ of values is kept, but the usual total ordering in $\langle 0,1\rangle$ is replaced by a particular partial one.

3. NONSTANDARD ORDERING IN THE UNIT INTERVAL

Let $d \geq 2$ be an integer, let $N_d = \{0,1,2,\ldots,d-1\} \subset \mathcal{N} = \{0,1,2,\ldots\}$, let $\delta(d) = \langle 0,1\rangle \to N_d^\infty = \{\langle x_1, x_2, \ldots\rangle : x_i \in D_n\}$ be a mapping ascribing to each real number $x \in \langle 0,1\rangle$ (one of) its d-adic decomposition(s), i.e., decomposition(s) to the base d, $\delta(d)(x) = \langle x_i^{\delta(d)}\rangle_{i=1}^\infty = \langle x_1^{\delta(d)}, x_2^{\delta(d)}, \ldots\rangle$. Hence $\delta(d)(x)$ is a sequence of integers from N_d such that $\sum_{i=1}^\infty x_i^{\delta(d)} d^{-i} = x$. Decadic ($d = 10$) and binary ($d = 2$) bases are the best known ones, below we shall consider the case when $d = 3$. We shall define $x \leq_{\delta(d)} y$ for $x, y \in \langle 0,1\rangle$, if $x_i^{\delta(d)} \leq y_i^{\delta(d)}$ holds for each $i \in \mathcal{N}^+ = \mathcal{N} - \{0\}$. Given a nonempty subset $X \subset \langle 0,1\rangle$ we define $\bigvee_{x\in X}^{\delta(d)} x$ and $\bigwedge_{x\in X}^{\delta(d)} x$ ($\bigvee^{\delta(d)} X$ and $\bigwedge^{\delta(d)} X$, abbreviately), by

$$\bigvee\nolimits_{x\in X}^{\delta(d)} x = \left\langle \sup\left\{x_i^{\delta(d)} : x \in X\right\}\right\rangle_{i=1}^{\infty}, \tag{23}$$

$$\bigwedge\nolimits_{x\in X}^{\delta(d)} x = \left\langle \inf\left\{x_i^{\delta(d)} : x \in X\right\}\right\rangle_{i=1}^{\infty}. \tag{24}$$

Obviously, $\leq_{\delta(d)}$ is a partial ordering in $\langle 0,1\rangle$ and $\bigvee^{\delta(d)}$ ($\bigwedge^{\delta(d)}$, resp.) is the supremum (the infimum, resp.) operation generated in $\langle 0,1\rangle$ by this partial ordering. It is also clear that $x \leq_{\delta(d)} y$ implies $x \leq y$ in the usual sense but not, in general, vice versa, and that the case when $x \vee^{\delta(d)} y$ $(= \bigvee^{\delta(d)}\{x,y\})$ is greater than x but also than y can easily happen. Given $\emptyset \neq X \subset \langle 0,1\rangle$, define $\sum_{x\in X}^{\delta(d)} x$ (or $\sum^{\delta(d)} X$) by $\sum_{i=1}^\infty (\sum_{x\in X} x_i^{\delta(d)})\, d^{-i}$, supposing that the sequence $\left\langle \sum_{x\in X} x_i^{\delta(d)}\right\rangle_{i=1}^\infty$ is in N_d^∞, i.e., supposing that $\sum_{x\in X} x_i^{\delta(d)} \leq d-1$ holds for each $i \in \mathcal{N}^+$, $\sum^{\delta(d)} X$ being undefined otherwise. For each $i \in N_d$ let

$$w_i^{\delta(d)}(x) = \lim_{n\to\infty} (1/n)\,\mathrm{card}\{j \in \mathcal{N}^+, j \leq n, x_j^{\delta(d)} = i\} \tag{25}$$

be the limit value of the relative frequence of occurrences of the digit i in finite initial segments of the d-adic decomposition of x defined by $\delta(d)$ supposing that this limit value exists, $w_i^{\delta(d)}(x)$ being undefined otherwise.

Let d and $\delta(d)$ be as above, let $\emptyset \neq D \subset N_d$ be a nonempty proper subset of N_d, let $x, y \in \langle 0,1\rangle$ be two reals. The real number z denoted by $[x|y]_{\langle \delta(d), D\rangle}$ and called *the result of conditioning of (the real number) x by (the real number) y with respect to the parameters $\delta(d)$ and D*, or briefly *x conditioned by y* w.r.to $\langle \delta(d), D\rangle$, is defined as follows. If the set $\{j \in \mathcal{N}^+ : y_j^{\delta(d)} \in D\}$ is finite, z is undefined. If this set is infinite, denote its elements by $i_1, i_2, \ldots$ in such a way that $i_j < i_{j+1}$ holds for each $j \in \mathcal{N}^+$. Set $z_j = x_{i_j}^{\delta(d)}$ for each $j \in \mathcal{N}^+$ and define $z = [x|y]_{\langle \delta(d), D\rangle} = \sum_{j=1}^\infty z_j\, d^{-j}$. Hence, $[x|y]_{\langle \delta(d), D\rangle}$ is defined by its d-adic decomposition obtained as the subsequence of the d-adic decomposition $\delta(d)(x)$ of x resulting when choosing just those indices for which the corresponding d-adic digit of the decomposition $\delta(d)(y)$ of y is in D.

4. NONSTANDARD CONDITIONAL POSSIBILISTIC MEASURES

Definition 1. Let Ω be a nonempty set, let $\mathcal{S} \subset \mathcal{P}(\Omega)$ be a system of subsets of Ω such that $\{\emptyset, \Omega\} \subset \mathcal{S}$. Let $d \geq 2$ be an integer, let $\delta(d)$ be a mapping ascribing to each $x \in \langle 0, 1\rangle$ (one of) its d-adic decomposition(s). Let $c \geq 2$ be a cardinal number. *A c-complete nonstandard $\delta(d)$-possibilistic measure defined on $\mathcal{S}$* is a mapping $\Pi : \mathcal{S} \to \langle 0, 1\rangle$ possessing these properties: $\Pi(\emptyset) = 0$, $\Pi(\Omega) = 1$, and if $\emptyset \neq \mathcal{X} \subset \mathcal{S}$ is such that $\text{card}(\mathcal{X}) \leq c$ and $\bigcup \mathcal{X} \in \mathcal{S}$ hold, then $\Pi(\bigcup \mathcal{X}) = \bigvee_{A \in \mathcal{X}}^{\delta(d)} \Pi(A)$. Let $\emptyset \neq D \subset N_d$ be a proper subset of N_d. *The c-complete nonstandard $\langle \delta(d),\ D\rangle$-conditional possibilistic measure defined by the measure* Π is a partial mapping $\Pi_{\text{ns}} : \mathcal{S} \times \mathcal{S} \to \langle 0, 1\rangle$ such that, for each $A,\ B \in \mathcal{S}$, $\Pi_{\text{ns}}(A, B) = [\Pi(A) \mid \Pi(B)]_{\langle \delta(d), D\rangle}$ supposing that the last value is defined, i.e., supposing that $\text{card}\{j \in \mathcal{N}^+ : \Pi(B)_j^{\delta(d)} \in D\} = \infty$. We shall write $\Pi_{\text{ns}}(A|B)$ instead of $\Pi_{\text{ns}}(A, B)$.

Let us investigate a particular case when the nonstandard operations and relations defined above take a more intuitive and Boolean-like nature. Let $d = 3$, let $\delta(3)$ be such that $\langle x_i^{\delta(3)}\rangle_{i=1}^{\infty}$ does not contain the digit 1, if such a decomposition of the real number x exists. Consequently, $\delta(3)$ is a one-to-one mapping taking the so called *Cantor set* $\mathcal{C}$ (Cantor subset of the unit interval) into the space $\{0, 1\}^{\infty}$ of all infinite binary sequences consisting of the digits 0 and 2. For the sake of simplicity we shall write $*$ (asterisk) instead of $\delta(3)$.

Definition 2. Let Ω be a nonempty set, let $\mathcal{S}$ be a nonempty σ-field of subsets of Ω. An $\aleph_0$-complete nonstandard $*$-possibilistic measure defined on $\mathcal{S}$ and taking its values in $\mathcal{C}$ is called a *Cantor σ-complete nonstandard possibilistic measure on* $\mathcal{S}$. This measure is called *σ-additive*, if for each sequence $\langle A_1, A_2, \ldots\rangle$ of mutually disjoint sets from $\mathcal{S}$ the nonstandard sum $\sum_{i=1}^{*\,\infty} \Pi(A_i)$ is defined, if this is the case, then obviously $\sum_{i=1}^{*\,\infty} \Pi(A_i) = \bigvee_{i=1}^{*\,\infty} \Pi(A_i) = \Pi\left(\bigcup_{i=1}^{\infty} A_i\right)$. When defining the corresponding conditional measure Π_{ns} generated by Π we shall take $D = \{2\}$ (obviously, $N_3 = \{0, 1, 2\}$).

The following fact demonstrates an interesting relation between conditional nonstandard possibilistic measures and the usual conditional probabilities. Its proof can be obtained by a more or less routine generalization of the proof for the case of Cantor possibilistic measure (i.e., $d = 3,\ D = \{2\}$) presented in [4] or [5].

Fact 10. Let Ω be a nonempty set, let $\mathcal{S}_0$ be a nonempty field (algebra) of subsets of Ω, let Π be a 2-complete (i.e., finitely complete) nonstandard $*$-possibilistic measure defined on $\mathcal{S}$ and such that $\Pi(A \cap B) = \bigwedge^* \{\Pi(A), \Pi(B)\}$ ($= \Pi(A) \wedge^* \Pi(B)$ in a more common notation) holds for each $A,\ B \in \mathcal{S}_0$. Let $D = \{2\}$, let Π take values in the Cantor set $\mathcal{C}$, let $A,\ B \in \mathcal{S}_0$ be such that $w_2(\Pi(B))$, $w_2(\Pi(A \cap B))$ and $w_2(\Pi_{\text{ns}}(A|B))$ are defined. Then

$$w_2(\Pi(A \cap B)) = w_2(\Pi_{\text{ns}}(A|B))\, w_2(\Pi(B)). \tag{26}$$

Let us recall that $w_2(\Pi(A))$ is the limit value of the relative frequences of occurrences of the digit 2 in the ternary decomposition of the real number $\Pi(A)$ (supposed to be in $\mathcal{C}$). Denoting by $P(C)$ the value of $w_2(\Pi(C))$ for each $C \in \mathcal{S}_0$ for which it is defined, and by $P(A|B)$ the value $w_2(\Pi_{\rm ns}(A|B))$, (26) turns into the well-known equality $P(A \cap B) = P(A|B)\, P(B)$, which can be found in any elementary textbook of probability theory as the definition of the conditional probability $P(A|B)$.

The relations between Cantor nonstandard possibilistic measures and probabilistic measures go much further as the following statement shows (cf. [4] and [5] for more detail).

Fact 11. Let Ω be a nonempty set, let $\mathcal{S}$ be a nonempty σ-field of subsets of Ω, let Π be a Cantor σ-complete and σ-additive nonstandard possibilistic measure defined on $\mathcal{S}$. Then Π is also a classical σ-additive and extensional probability measure on $\mathcal{S}$. The class $\mathcal{S}_0 = \{A \in \mathcal{S} : P(A) = w_2(\Pi(A)) \text{ is defined}\}$ is a nonempty subclass of $\mathcal{S}$ closed with respect to complements and with respect to finite unions of disjoint sets and P is a finitely additive (but not necessarily σ-additive!) probability measure on $\mathcal{S}_0$ such that $P(A|B) = w_2(\Pi_{\rm ns})(A|B)$ holds for all A, $B \in \mathcal{S}_0$ for which all the three real numbers occurring in (26) are defined.

Let us formulate the following open problem the solution of which will be postponed till another occasion. Let Ω be a nonempty set, let $\mathcal{S} = \mathcal{P}(\Omega)$ be the σ-field of all subsets of Ω, and let Π be a Cantor σ-complete nonstandard possibilistic measure defined on $\mathcal{S}$. Let U, X, Y, Z, $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ denote the same objects as above, let $\Pi(\boldsymbol{x}|\boldsymbol{y}) = [\Pi(\boldsymbol{x}) \mid \Pi(\boldsymbol{y})]_*$
($= [\Pi(\boldsymbol{x}) \mid \Pi(\boldsymbol{y})]_{\langle \delta(3),\{2\}\rangle}$ by definition), let $\Pi(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) = [\Pi(\boldsymbol{x}|\boldsymbol{y})|\Pi(\boldsymbol{z})]$. Set, as above

$$I(X,Y|Z) \Leftrightarrow_{\rm df} \Pi(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) = \Pi(\boldsymbol{x}|\boldsymbol{z}) \tag{27}$$

for all $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ for which the values $\Pi(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z})$ and $\Pi(\boldsymbol{x},\boldsymbol{z})$ are defined. Which of the properties (A1)–(A5) are satisfied by $I(X,Y|Z)$ defined in this way?

(Received August 7, 1996.)

REFERENCES

[1] L. M. de Campos, J. F. Heuter and S. Moral: Possibilistic independence. In: Proceedings of EUFIT 95 (Third European Congress on Intelligent Techniques and Soft Computing), Verlag Mainz and Wissenschaftsverlag, Aachen 1995, vol. 1, pp. 69–73.

[2] D. Dubois and H. Prade: Théorie des Possibilités – Applications à la Représentation de Connaissances en Informatique. Mason, Paris 1985.

[3] E. Hisdal: Conditional possibilities, independence and noninteraction. Fuzzy Sets and Systems *1* (1978), 283–297.

[4] I. Kramosil: Extensional processing of probability measures. Internat. J. Gen. Systems *22* (1994), 159–170.

[5] I. Kramosil: An axiomatic approach to extensional probability measures. In: Proceedings of the European Conference Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Fribourg (Lecture Notes in Artificial Intelligence 946.) Springer–Verlag, Berlin 1995, pp. 267–276.

[6] D. Lewis: Probabilities of conditionals and conditional probabilities. Philos. Review *85* (1976), 297–315.

[7] M. Loève: Probability Theory. D. van Nostrand, New York 1955.

[8] J. Pearl: Probabilistic Reasoning in Intelligent Systems – Networks of Plausible Inference. Morgan and Kaufmann, San Matteo, California 1988.

[9] G. Shafer: A Mathematical Theory of Evidence. Princeton Univ. Press, Princeton, New Jersey 1976.

[10] Ph. Smets: About updating. In: Uncertainty in Artificial Intelligence 91 (D'Ambrosio, Ph. Smets and P. P. Bonissone, eds.), Morgan Kaufman, Sao Matteo, California 1991, pp. 378–385.

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

RNDr. Ivan Kramosil, DrSc., Institute of Computer Science – Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 182 07 Praha 8. Czech Republic.
e-mail: kramosil@uivt.cas.cz