Kybernetika

Ivan Kramosil

Decision-making under uncertainty processed by lattice-valued possibilistic measures

Kybernetika, Vol. 42 (2006), No. 6, 629--646

Persistent URL: http://dml.cz/dmlcz/135741

KYBERNETIKA — VOLUME 42 (2006), NUMBER 6, PAGES 629 – 646

DECISION–MAKING UNDER UNCERTAINTY PROCESSED BY LATTICE–VALUED POSSIBILISTIC MEASURES

Ivan Kramosil

The notion and theory of statistical decision functions are re-considered and modified to the case when the uncertainties in question are quantified and processed using lattice-valued possibilistic measures, so emphasizing rather the qualitative than the quantitative properties of the resulting possibilistic decision functions. Possibilistic variants of both the minimax (the worst-case) and the Bayesian optimization principles are introduced and analyzed.

Keywords: decision making under uncertainty, complete lattice, lattice-valued possibilistic measures, possibilistic decision function, minimax and Bayesian optimization

AMS Subject Classification: 28E10, 28E99

1. INTRODUCTION, MOTIVATION, PRELIMINARIES

In this paper, we take an inspiration and motivation in the theory of statistical decision functions developed on the grounds of Kolmogorov's axiomatic probability theory (see, e. g. [8] or [10]). Consider a system $\mathcal{S}$ the actual internal state of which is s_0. The value of s_0 is neither known nor immediately observable by a subject who is to control the system $\mathcal{S}$, just a set S in which s_0 is included is known. The subject takes a decision d from a fixed set D of possible decisions. The consequences resulting this act depend only on d and the actual state s_0 and they are supposed to be quantified by a non-negative real number $\lambda(s,d)$ taken as the loss suffered when s is the actual state and d is applied. Hence, a decision d is optimal w.r.t. s, if the inequality $\lambda(s,d_0) \leq \lambda(s,d_1)$ holds for each $d_1 \in D$.

However, up to the trivial cases the actual states are not known or directly observable, the only what is at the subject's disposal are empirical data taking their values in a set E. We suppose that the way in which the subject takes his/her decisions can be described by a decision function $\delta : E \to D$. Combining our notation together, if s is the actual state, e is the empirical value being at the subject disposal and δ is the decision function which he/she applies, the suffered loss is $\lambda(s,\delta(e)) \in [0,\infty)$. A decision function $\delta_0 : E \to D$ is uniformly optimal (uniformly the best one), if for every $\delta : E \to D$, every $s \in S$, and every $e \in E$ the inequality $\lambda(s,\delta_0(e)) \leq \lambda(s,\delta(e))$ is valid. However, up to very elementary cases when the actual state of the system can be identified from the empirical data, such an optimal decision function does not exist.

Let us analyze, in more detail, the case when the phenomenon of randomness enters our model, so giving arise the notion of statistical decision functions. The reader is supposed to be familiar with the notion of *probability space* $\langle \Omega, \mathcal{A}, P\rangle$. Given a nonempty set Z and a σ-field $\mathcal{Z}$ of subsets of Z, a $\langle Z, \mathcal{Z}\rangle$-measurable mapping $X : \Omega \to Z$ is called a *$\langle Z, \mathcal{Z}\rangle$-valued random variable* defined on the probability space $\langle \Omega, \mathcal{A}, P\rangle$. If Z is finite and $\mathcal{Z}$ is not specified, we suppose that $\mathcal{Z} = \mathcal{P}(Z)$. Hence, we will suppose that the empirical value being at the subject's disposal, is the value taken by an E-valued random variable $\eta : \Omega \to E$. Accepting, below, the so called bayesian approach, we will suppose that also the actual internal state s_0 of the system in question is the value taken by an S-valued random variable $\sigma : \Omega \to S$. For the sake of simplicity we suppose that E, S, D are finite sets, so that the conventions on the corresponding σ-fields apply. Under this setting, the decision function δ converts into a D-valued random variable $\delta^* : \Omega \to D$ such that $\delta^*(\omega) = \delta(\eta(\omega))$ for every $\omega \in \Omega$, and the loss function λ converts into a real-valued random variable $\lambda^* : \Omega \to R = (-\infty, \infty)$, setting $\lambda^*(\omega) = \lambda(\sigma(\omega), \delta(\eta(\omega)))$ for every $\omega \in \Omega$.

In general, there does not exist a decision function $\delta_0 : E \to D$ such that the inequality $\lambda(\sigma(\omega), \delta_0(\eta(\omega))) \leq \lambda(\sigma(\omega), \delta(\eta(\omega)))$ would hold for every $\delta : E \to D$ and every $\omega \in \Omega$. However, the bayesian approach enables to define, under some conditions, the expected value $E\lambda^* = \int_{-\infty}^{\infty} \lambda^*(\cdot)\, \mathrm{d}P$ of the random variable λ^*. A decision function $\delta_0 : E \to D$ is *optimal* in the bayesian sense w.r.t. the apriori random variable σ, if

$$E(\lambda(\delta(\cdot), \delta_0(\eta(\cdot)))) = \inf\{E(\lambda(\sigma(\cdot), \delta(\eta(\cdot)))) : \delta : E \to D\} \tag{1.1}$$

holds. Given $\varepsilon > 0$, a decision function $\delta_{0,\varepsilon} : E \to D$ is *ε-optimal* in the bayesian sense w.r.t. σ, if for each $\delta : E \to D$ the inequality

$$E(\lambda(\delta(\cdot), \delta_{0,\varepsilon}(\eta(\cdot)))) < E(\lambda(\delta(\cdot), \delta(\eta(\cdot)))) + \varepsilon \tag{1.2}$$

is the case. A decision function δ_0 satisfying (1.1) need not, in general, exist, but a decision function $\delta_{0,\varepsilon}$ satisfying (1.2) obviously exists for each $\varepsilon > 0$. Cf. [2] or [12] for more detail.

Keeping the idea that degrees of randomness are defined by sizes of (some) subsets of the space Ω of all elementary random events under consideration, let us abandon the assumption that the values ascribed to these sets are real numbers, and let us take into consideration also non-numerical values from a set equipped by a structure weaker than the structures definable over the unit interval of real numbers. E.g., degrees of randomness from a partially ordered set or lattice-valued degrees may be taken into consideration, with the operation of addition replaced by that of supremum, definable in partially ordered sets and lattices. Pursuing this way of reasoning in more detail, we arrive at the notion of lattice-valued possibilistic measure.

Our aim will be, in this paper, to investigate whether the idea of statistical decision function, very briefly sketched above, can be modified to the case when probability measure is replaced by a lattice-valued possibilistic measure. To be able to do so at a formalized level, the reader is supposed to be familiar with some more or less elementary notions of lattice theory and their relations (cf. [1] or [6], e. g.), let us list them very briefly.

Let us begin with the notion of *partially ordered set (poset)* $\mathcal{T} = \langle T, \leq_{\mathcal{T}} \rangle$ (or simply $\leq$, if no misunderstanding menaces). By $t_1 \vee t_2$ or $\bigvee S$ ($t_1 \wedge t_2, \bigwedge S$, resp.) we denote the supremum (infimum, resp.) operation defined in $\mathcal{T}$ in the standard way. The values $\bigvee S$ and/or $\bigwedge S$ need not be defined in general, but if it is the case, they are defined uniquely. A pose $\mathcal{T} = \langle T, \leq_{\mathcal{T}} \rangle$ is called a *complete lattice*, if for each $S \subset T$ the values $\bigvee S$ and $\bigwedge S$ are defined (for $S = \emptyset$, the conventions $\bigvee S = \oslash_{\mathcal{T}} = \bigwedge T$ and $\bigwedge S = \mathbf{1}_{\mathcal{T}} = \bigvee T$ apply). $\mathbf{1}_{\mathcal{T}}$ is the unit element of $\mathcal{T}$ and $\oslash_{\mathcal{T}}$ is the zero element of $\mathcal{T}$, only the nontrivial cases when $\oslash_{\mathcal{T}} < \mathbf{1}_{\mathcal{T}}$ holds are considered below ($t_1 < t_2$ means that $t_1 \leq t_2$ and $t_1 \neq t_2$ hold).

Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice. A mapping $\tau : T \times T \to T$ is called *triangular seminorm* (*t-seminorm)* on $\mathcal{T}$, if the conditions

(i) $\tau(t, \mathbf{1}_{\mathcal{T}}) = \tau(\mathbf{1}_{\mathcal{T}}, t) = t$ for each $t \in T$ (boundary conditions), and

(ii) for each $s_1, s_2, t_1, t_2 \in T$ such that $s_1 \leq t_1$ and $s_2 \leq t_2$ hold, the inequality $\tau(s_1, s_2) \leq \tau(t_1, t_2)$ holds (isotonicity)

are valid. If τ is, moreover, commutative and associative, it is called *triangular norm* (*t-norm)* on $\mathcal{T}$.

Obviously, for each complete lattice $\mathcal{T} = \langle T, \leq \rangle$ infimum is a t-norm on $\mathcal{T}$ and for each t-norm τ on $\mathcal{T}$ the relation $\tau(s,t) \leq s \wedge t$ holds. The inequality $\bigvee_{s \in A} \tau(t,s) \leq \tau(t, \bigvee A)$ also holds for each t-norm τ on $\mathcal{T}$ and each $A \subset T$, but the equality need not hold in general. A t-norm τ on T is called *distributive,* if the equality $\bigvee_{s \in A} \tau(t,s) = \tau(t, \bigvee A)$ holds for each $t \in T$ and $A \subset T$.

Cf. [1], [6], or [11] for more detail on posets, lattices, Boolean algebras and related notions.

2. LATTICE–VALUED POSSIBILISTIC MEASURES AND VARIABLES

Given a nonempty set Ω, the reader is supposed to be familiar with the notion of field and σ-field of subsets of Ω. A nonempty system $\mathcal{A} \subset \mathcal{P}(\Omega)$ is called *complete field* (in [3] the term *ample field* is proposed), if for each $A \in \mathcal{A}$ and each $\mathcal{A}_0 \subset \mathcal{A}$ the sets $\Omega - A$ and $\bigcup \mathcal{A}_0$ are in $\mathcal{A}$.

Definition 2.1. Let Ω be a nonempty set, let $\mathcal{A} \subset \mathcal{P}(\Omega)$ be a complete field (hence, Ω and $\emptyset$ are in $\mathcal{A}$), let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice. A mapping $\Pi : \mathcal{A} \to T$ is called a $\mathcal{T}$-*(valued) possibilistic measure on* $\mathcal{A}$, if $\Pi(\Omega) = \mathbf{1}_{\mathcal{T}}$, $\Pi(\emptyset) = \oslash_{\mathcal{T}}$ and, for every $A, B \in \mathcal{A}$, $\Pi(A \cup B) = \Pi(A) \vee \Pi(B)$. A $\mathcal{T}$-possibilistic measure Π on $\mathcal{A}$ is *complete*, if for each $\emptyset \neq \mathcal{A}_0 \subset \mathcal{A}$ the relation

$$\Pi\left(\bigcup \mathcal{A}_0\right) = \bigvee \{\Pi(A) : A \in \mathcal{A}_0\} \tag{2.1}$$

holds. The triple $\langle \Omega, \mathcal{A}, \Pi \rangle$ is called $\mathcal{T}$-*possibilistic space.*

Evidently, $\mathcal{T}$-possibilistic measures are particular cases of the so called *$\mathcal{T}$-monotone* (or *$\mathcal{T}$-fuzzy) measures on $\mathcal{A}$.*

Definition 2.2. Given nonempty sets Ω and Z and nonempty complete fields $\mathcal{A} \subset \mathcal{P}(\Omega)$ and $\mathcal{Z} \subset \mathcal{P}(Z)$, a mapping $X : \Omega \to Z$ is *$\mathcal{Z}$-(valued) possibilistic variable*, if it is $(\mathcal{A}, \mathcal{Z})$-measurable, i. e., if for each $B \in \mathcal{Z}$ its inverse image $X^{-1}(B)$ is in $\mathcal{A}$.

In what follows, we will investigate $\langle T, \mathcal{Z}_T \rangle$-valued possibilistic variables and we will need and suppose that for each $t \in T$ the set $\{s \in T : s \geq t\}$ is in $\mathcal{Z}_T$. However, it can be easily proved that this is the case just when $\mathcal{Z}_T = \mathcal{P}(T)$. Consequently, we will suppose, in what follows and if not stated explicitly otherwise, that $\mathcal{Z}_T = \mathcal{P}(T)$ and such possibilistic variables will be denoted as T-valued.

Definition 2.3. Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let $\langle \Omega, \mathcal{A}, \Pi \rangle$ be a $\mathcal{T}$-possibilistic space, let X be a T-possibilistic variable, let τ be a t-norm on T. The *expected value E_τ X of X w.r.t.* τ is defined by

$$E_\tau X = \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : X(\omega) \geq t\})] = \oint X d\Pi, \tag{2.2}$$

as this value is nothing else than the so called *Sugeno integral* of the T-valued function X taken w.r.t. $\mathcal{T}$-possibilistic measure Π and the t-norm τ. Under the conditions imposed above on the notions occurring in (2.1), the value $E_\tau X$, belonging to T, is always defined.

Given $A \in \mathcal{A}$, let $X_A(\omega) = \mathbf{1}_{\mathcal{T}}$, if $\omega \in A, X_A(\omega) = \oslash_{\mathcal{T}}$, if $\omega \in \Omega - A$. Then, for each t-norm τ on T, $E_\tau X_A \leq E_\wedge X_A = \Pi(A)$ holds ($\wedge$ is the minimum).

Cf. [3], [4] and [5] for more detail on possibilistic measures.

3. POSSIBILISTIC DECISION FUNCTIONS

In this section our aim will to reconsider the problem of decision making under uncertainty as sketched in Section 1, but this time with the phenomenon of uncertainty (randomness) formally described and processed, and with the degrees of uncertainty quantified, by the tools offered by lattice-valued possibilistic measures and variables. Hence, let the symbols S (set of states), D (set of decisions and E (set of empirical values) keep their meanings and the intuition behind as introduced above. Decision function δ is, again, a mapping which takes E into D. Given a complete lattice $\mathcal{T} = \langle T, \leq \rangle$, a $\mathcal{T}$-valued loss function λ is a mapping which takes the Cartesian product $S \times D$ into T. So, the loss suffered when s is the actual state, e the empirical value and δ the decision function applied can be denoted by $\lambda(s, \delta(e))$.

In order to implement the possibilistically processed phenomenon of uncertainty into our model, let us fix a $\mathcal{T}$-possibilistic space $\langle \Omega, \mathcal{A}, \Pi \rangle$ and suppose that (i) the actual internal state s is the value taken by an S-valued possibilistic variable defined on $\langle \Omega, \mathcal{A}, \Pi \rangle$. In both the cases we suppose that the complete fields over S and E are the power-sets $\mathcal{P}(S)$ and $\mathcal{P}(E)$.

Under these notations and assumptions the loss suffered when $\omega \in \Omega$ is the elementary random event can be defined by the value $\lambda(\sigma(\omega), \delta(\eta(\omega)))$ from T; let us prove that it is a T-valued possibilistic variable (let us recall that we take $\mathcal{Z}_T = \mathcal{P}(T)$). Indeed, given $d \in D$ and $t \in T$ and setting $\delta^{-1}(d) = \{e \in E : \delta(e) = d\}$, $R(t) = \{\langle s, d\rangle \in S \times D : \lambda(s, d) = t\}$, we obtain that

$$\begin{aligned} & \{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega))) = t\} \\ = \; & \{\omega \in \Omega : \langle \sigma(\omega), \delta(\eta(\omega))\rangle \in R(t)\} \\ = \; & \bigcup_{\langle s,d\rangle \in R(t)} \{\omega \in \Omega : \sigma(\omega) = s, \delta(\eta(\omega)) = d\} \\ = \; & \bigcup_{\langle s,d\rangle \in R(t)} (\{\omega \in \Omega : \sigma(\omega) = s\} \cap \bigcup_{e \in \delta^{-1}(d)} \{\omega \in \Omega : \eta(\omega) = e\}). \end{aligned} \tag{3.1}$$

As the sets $\{\omega \in \Omega : \sigma(\omega) = s\}$ and $\{\omega \in \Omega : \eta(\omega) = e\}$ are in $\mathcal{A}$ for each $s \in S$ and $e \in E$ and $\mathcal{A}$ is a complete field, the set $\{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega))) = t\}$ is also in $\mathcal{A}$. Hence, for each $B \subset T$, the set

$$\{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega))) \in B\} = \bigcup_{t \in B} \{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega))) = t\} \tag{3.2}$$

is in $\mathcal{A}$, so that $\lambda(\sigma(\omega), \delta(\eta(\omega)))$ is a T-possibilistic variable defined on the $\mathcal{T}$-possibilistic space $\langle \Omega, \mathcal{A}, \Pi\rangle$ and taking $\langle \Omega, \mathcal{A}\rangle$ into $\langle T, \mathcal{P}(T)\rangle$.

Hence, we can define the expected value of the T-valued loss function $\lambda(\sigma(\cdot), \delta(\eta(\cdot)))$ (abbreviated by λ^*, if σ, δ and η are fixed) w.r.to a t-norm τ on T, setting

$$E_\tau \lambda^* = E\tau\lambda(\sigma(\cdot), \delta(\eta(\cdot))) = \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega))) \geq t\})] \tag{3.3}$$

and using this value as a global characteristic of the quality of the decision function δ. Due to the conditions imposed on $\mathcal{T} = \langle T, \leq\rangle$ this value is always defined, however, contrary to the case of real-valued loss functions, some pairs $\langle \delta_1, \delta_2\rangle$ of decision functions may be incomparable w.r.t to corresponding expected values. As $\mathcal{T}$ is complete lattice, the value

$$E_\tau^{\inf} \lambda^* = \bigwedge_{\delta : E \to T} \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega))) \geq t\})] \tag{3.4}$$

is defined in T, so that, for each $t_0 > E_\tau^{\inf} \lambda^*$, there exists $\delta_0 : E \to D$ such that the inequality $E_\tau \lambda(\sigma(\cdot), \delta_0(\eta(\cdot))) < t_0$ holds. However, in general, a decision function δ_1 satisfying $E_\tau \lambda(\sigma(\cdot), \delta_1(\eta(\cdot))) = E_\tau \lambda^*$ need not exist, as the following simple example demonstrates.

Let $\langle \Omega, \mathcal{A}, \Pi\rangle$ be a possibilistic space, let $S = D = \{s_1, s_2\}$, let $E = \{e\}$ be the degenerated observational space, let $\mathcal{T} = \langle T, \leq\rangle$ be a complete lattice, let $\lambda(s, d) = \oslash_{\mathcal{T}}$, if $s = d$, $\lambda(s, d) = \mathbf{1}_{\mathcal{T}}$, if $s \neq d$, let $\Pi(\{\omega \in \Omega : \sigma(\omega) = s_i\}) = t_i > \oslash_{\mathcal{T}}$ for both $i = 1, 2$, let $t_1 \wedge t_2 = \oslash_{\mathcal{T}}$. Such a possibilistic measure can be easily obtained, e. g., take $\mathcal{T} = \langle \mathcal{P}(\Omega), \subset\rangle$ and define Π as the identity on the complete field $\mathcal{P}(\Omega)$. Then

$$\begin{aligned} & \Pi(\{\omega \in \Omega : \sigma(\omega) = s_1\}) \wedge \Pi(\{\omega \in \Omega : \sigma(\omega) = s_2\}) \\ & = \{\omega \in \Omega : \sigma(\omega) = s_1\} \cap \{\omega \in \Omega : \sigma(\omega) = s_2\} = \emptyset = \oslash_{\mathcal{T}}, \end{aligned} \tag{3.5}$$

even if the sets $\{\omega \in \Omega, \sigma(\omega) = s_i\}$ are nonempty for both $i = 1, 2$.

Only the two decision functions are possible: $\delta_{s_1}(e) = s_1$ and $\delta_{s_2}(e) = s_2$. As $\eta(\omega) = e$ for each $\omega \in \Omega$, we obtain that

$$\begin{aligned} \lambda(\sigma(\omega), \delta_{s_1}(\eta(\omega))) &= \lambda(\sigma(\omega), s_1) = \emptyset_{\mathcal{T}}, \text{ if } \sigma(\omega) = s_1, \\ \lambda(\sigma(\omega), \delta_{s_1}(\eta(\omega))) &= \lambda(\sigma(\omega), s_{\cdot 1}) = \mathbf{1}_{\mathcal{T}}, \text{ if } \sigma(\omega) = s_2, \end{aligned} \tag{3.6}$$

and dually for δ_{s_2}. So, given a t-norm τ on T,

$$\begin{aligned} & E_\tau \lambda(\sigma(\cdot), \delta_{s_1}(\eta(\cdot))) = \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : \lambda(\sigma(\omega), \delta_{s_1}(\eta(\omega))) \geq t\})] \\ \leq\ & \bigvee_{t \in T} [t \wedge \Pi(\{\omega \in \Omega : \lambda(\sigma(\omega), \delta_{s_1}(\eta(\omega))) \geq t\})] \\ =\ & \bigvee_{t \in T} [t \wedge \Pi(\{\omega \in \Omega : \lambda(\sigma(\omega), s_1) \geq t\})] \\ =\ & \mathbf{1}_{\mathcal{T}} \wedge \Pi(\{\omega \in \Omega : \sigma(\omega) = s_2\}) = \mathbf{1}_{\mathcal{T}} \wedge t_2 = t_2. \end{aligned} \tag{3.7}$$

Analogously, $E_\tau \lambda(\sigma(\cdot), \delta_{s_2}(\eta(\cdot))) \leq t_1$ holds, so that the relation

$$E_\tau^{\inf} \lambda^* = (E_\tau \lambda(\sigma(\cdot), \delta_{s_1}(\eta(\cdot)))) \wedge (E_\tau \lambda(\sigma(\cdot), \delta_{s_2}(\eta(\cdot)))) = t_1 \wedge t_2 = \emptyset_{\mathcal{T}} \tag{3.8}$$

follows, but the value $E_\tau^{\inf} \lambda^*$ is reachable neither by δ_{s_1} nor by δ_{s_2}.

4. CLASSIFICATION OF POSSIBILISTIC DECISION FUNCTIONS BASED ON THE MINIMAX PRINCIPLE

One way how to introduce the minimax principle into our reasoning reads as follows. Instead of the loss $\lambda(s, d)$ suffered when s is the actual state and d is the decision we consider its "pessimistic" approximation from above, supposing that the loss $\bigvee_{s \in S} \lambda(s, d)$ is suffered, hence, we define the loss function $\hat{\lambda} : D \to T$ by $\hat{\lambda}(d) = \bigvee_{s \in S} \lambda(s, d)$. So $\hat{\lambda}(\eta(\omega))$ is the loss suffered when $\eta(\omega)$ is the empirical value under consideration, it defines a T-valued possibilistic variable on $\langle \Omega, \mathcal{A}, \Pi \rangle$ and the expected value $E_\tau \hat{\lambda}(\delta(\eta(\cdot)))$ of this variable, denoted by $\chi^{MM}(\delta)$ (MM for minimax) can serve as a T-valued degree of quality of the decision function δ.

Another criterion of quality of the decision function δ obeying the minimax principle may be like this. Given $s \in S$, take the expected value of the loss function $\lambda(s, \delta(\eta(\cdot)))$ and set $\chi^{mm}(\delta) = \bigvee_{s \in S} E_\tau \lambda(s, \delta(\eta(\cdot)))$. Denoting by $\chi_\sigma^B(\delta)$ (B for Bayes) the value $E_\tau \lambda^*$ defined by (3.3), the following relation can be proved.

Theorem 4.1. Let S, D and E be as in Section 3, let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let τ be a t-norm on T, let $\lambda : S \times D \to T$ be a T-valued loss function, let $\langle \Omega, \mathcal{A}, \Pi \rangle$ be a $\mathcal{T}$-possibilistic space. Then, for each possibilistic variables $\sigma : \Omega \to S$, $\eta : \Omega \to E$, and each decision function $\delta : E \to D$ the relation $\chi_\sigma^B(\delta) \leq \chi^{mm}(\delta) \leq \chi^{MM}(\delta)$ holds.

When proving this assertion, the following lemma will be of use.

Lemma 4.1. Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let $\langle \Omega, \mathcal{A}, \Pi \rangle$ be a $\mathcal{T}$-possibilistic space, let $\mathcal{F}$ be a nonempty set of T-valued possibilistic variables on $\langle \Omega, \mathcal{A}, \Pi \rangle$, let τ be a t-norm on T. Then the relation $\bigvee_{f \in \mathcal{F}} (E_\tau f) \leq E_\tau(\bigvee \mathcal{F})$ is valid, where $(\bigvee \mathcal{F})(\omega) = \bigvee_{f \in \mathcal{F}} f(\omega)$ for each $\omega \in \Omega$. If there exists, for every $\omega \in \Omega$, at most one $f \in \mathcal{F}$ such that $f(\omega) > \oslash_{\mathcal{T}}$ holds, then the equality $\bigvee_{f \in \mathcal{F}} (E_\tau f) = E_\tau(\bigvee \mathcal{F})$ holds.

P r o o f. For each $f \in \mathcal{F}$ and each $t \in T$ the inequality

$$\Pi(\{\omega \in \Omega : f(\omega) \geq t\}) \leq \Pi(\{\omega \in \Omega : \left(\bigvee \mathcal{F}\right)(\omega) \geq t\}) \tag{4.1}$$

is obvious, hence, the inequalities $E_\tau f \leq E_\tau(\bigvee \mathcal{F})$ and $\bigvee_{f \in \mathcal{F}} (E_\tau f) \leq E_\tau(\bigvee \mathcal{F})$ immediately follow. Let there exist, for every $\omega \in \Omega$, at most one $f \in \mathcal{F}$ such that $f(\omega) > \oslash_{\mathcal{T}}$ is the case. Then, for each $t \in T$, $t > \oslash_{\mathcal{T}}$,

$$\{\omega \in \Omega : \left(\bigvee \mathcal{F}\right)(\omega) \geq t\} = \bigcup_{f \in \mathcal{F}} \{\omega \in \Omega : f(\omega) \geq t\}, \tag{4.2}$$

so that

$$\Pi(\{\omega \in \Omega : \left(\bigvee \mathcal{F}\right)(\omega) \geq t\}) = \bigvee_{f \in \mathcal{F}} \Pi(\{\omega \in \Omega : f(\omega) \geq t\}) \tag{4.3}$$

follows. Moreover, for each $t > \oslash_{\mathcal{T}}$, if $\{\omega \in \Omega : (\bigvee \mathcal{F})(\omega) \geq t\} \neq \emptyset$, then there exists just one $f_t \in \mathcal{F}$ such that

$$\{\omega \in \Omega : \left(\bigvee \mathcal{F}\right)(\omega) \geq t\} = \{\omega \in \Omega : f_t(\omega) \geq t\}. \tag{4.4}$$

Hence, the inequalities

$$\tau[t, \Pi(\{\omega \in \Omega : \left(\bigvee \mathcal{F}\right)(\omega) \geq t\})] \leq \bigvee_{f \in \mathcal{F}} \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : f(\omega) \geq t\})] = \bigvee_{f \in \mathcal{F}} (E_\tau f) \tag{4.5}$$

and

$$E_\tau\left(\bigvee \mathcal{F}\right) = \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : \left(\bigvee \mathcal{F}\right)(\omega) \geq t\})] \leq \bigvee_{f \in \mathcal{F}} (E_\tau f) \tag{4.6}$$

follow, what completes the proof of Lemma 4.1. □

P r o o f o f T h e o r e m 4.1. Set, for each $s \in S$ and $\omega \in \Omega$, $\hat{\lambda}_s(\omega) = \lambda(s, \delta(\eta(\omega)))$, if $\sigma(\omega) = s$, $\hat{\lambda}_s(\omega) = \oslash_{\mathcal{T}}$, if $\sigma(\omega) \neq s$. Then, for each $\omega \in \Omega$, the relations

$$\hat{\lambda}_s(\omega) \leq \lambda(s, \delta(\eta(\omega))), \; \lambda(\sigma(\omega), \delta(\eta(\omega))) = \bigvee_{s \in S} \hat{\lambda}_s(\omega) \tag{4.7}$$

are valid and for each $\omega \in \Omega$ there exists at most one $s \in S$ such that $\hat{\lambda}_s(\omega) > \oslash_{\mathcal{T}}$ holds. Applying Lemma 4.1 to $\mathcal{F} = \{\hat{\lambda}_s : s \in S\}$, we obtain that $\bigvee_{s \in S} (E_{\mathcal{T}} \hat{\lambda}_s(\cdot)) =$

$E_\tau\left(\bigvee_{s\in S}\hat{\lambda}_s(\cdot)\right)$, so that the relation

$$\begin{aligned}\chi_\sigma^B(\delta) &= E_\tau\lambda(\sigma(\cdot),\delta(\eta(\cdot))) = E_\tau\left(\bigvee_{s\in S}\hat{\lambda}_s(\cdot)\right) = \bigvee_{s\in S}(E_\tau\hat{\lambda}_s(\cdot)) \\ &\le \bigvee_{s\in S}(E_\tau\lambda(s,\delta(\eta(\cdot))) = \chi^{mm}(\delta) \qquad (4.8)\end{aligned}$$

easily follows. Applying Lemma 4.1 again, now to $\mathcal{F} = \{\rho(s,\delta(\eta(\cdot))) : s \in S\}$, we obtain that $\chi^{mm}(\delta) \le E_\tau\left(\bigvee_{s\in S}\lambda(s,\delta(\eta(\cdot)))\right) = \chi^{MM}(\delta)$ holds and Theorem 4.1 is proved. □

The equality $\chi^{mm}(\delta) = \chi^{MM}(\delta)$ does not hold in general. Indeed, let $S = \{s_1, s_2\}$, $D = \{d_1, d_2\}$, $E = \{e_1, e_2\}$, let $\delta(e_i) = d_i$ for both $i = 1, 2$. Set $\Pi(\{\omega \in \Omega : \eta(\omega) = e_i\}) = t_i$ for both $i = 1, 2$ and suppose that $\oslash_\mathcal{T} < t_1, t_2 < \mathbf{1}_\mathcal{T}, t_1 \wedge t_2 = \oslash_\mathcal{T}$ holds ($t_1 \vee t_2 = \mathbf{1}_\mathcal{T}$ easily follows). Let $\lambda(s_1, d_1) = \lambda(s_2, d_2) = t_2, \lambda(s_1, d_2) = \lambda(s_2, d_1) = t_1$. Then, for each $\omega \in \Omega, \lambda(s_1,\delta(\eta(\omega))) \vee \lambda(s_2,\delta(\eta(\omega))) = t_1 \vee t_2 = \mathbf{1}_\mathcal{T} = \chi^{MM}(\delta)$. However,

$$\begin{aligned}&\{\omega \in \Omega : \lambda(s_i,\delta(\eta(\omega))) = t_2\} = \{\omega \in \Omega : \eta(\omega) = e_i\}, \\ &\{\omega \in \Omega : \lambda(s_i,\delta(\eta(\omega))) = t_1\} = \{\omega \in \Omega : \eta(\omega) \neq e_i\}, \qquad (4.9)\end{aligned}$$

consequently,

$$\begin{aligned}& E_\tau\lambda(s_1,\delta(\eta(\cdot))) = \tau[t_1, \Pi(\{\omega \in \Omega : \eta(\omega) = e_2\})] \vee \tau[t_2, \Pi(\{\omega \in \Omega : \eta(\omega) = e_1\})] \\ = \; & \tau[t_1, t_2] \vee \tau[t_2, t_1] \le (t_1 \wedge t_2) \vee (t_2 \wedge t_1) = \oslash_\mathcal{T} \qquad (4.10)\end{aligned}$$

The proof that $E_\tau\lambda(s_2,\delta(\eta(\cdot))) = \oslash_\mathcal{T}$ is quite analogous, so that

$$\chi^{mm} = (E_\tau\lambda(s_1,\delta(\eta(\cdot)))) \vee (E_\tau\lambda(s_2,\delta(\eta(\cdot)))) = \oslash_\mathcal{T} < \mathbf{1}_\mathcal{T} = \chi^{MM}(\delta). \qquad (4.11)$$

Theorem 4.2. Let the notations and conditions of Theorem 4.1 hold, let the loss function λ take only the values $\oslash_\mathcal{T}$ or $\mathbf{1}_\mathcal{T}$. Then, for each decision function $\delta : E \to D, \chi^{mm}(\delta) = \chi^{MM}(\delta)$.

Proof. Obviously, under the restrictions given, for each $s \in S, E_\tau\lambda(s,\delta(\eta(\cdot))) = \tau[\mathbf{1}_\mathcal{T}, \Pi(\{\omega \in \Omega : \lambda(s,\delta(\eta(\omega))) = \mathbf{1}_\mathcal{T}\})] = \Pi(\{\omega \in \Omega : \lambda(s,\delta(\eta(\omega))) = \mathbf{1}_\mathcal{T}\})$. So,

$$\begin{aligned}& \chi^{mm}(\delta) = \bigvee_{s\in S} E_\tau\lambda(s,\delta(\eta(\cdot))) = \bigvee_{s\in S}\Pi(\{\omega \in \Omega : \lambda(s,\delta(\eta(\omega))) = \mathbf{1}_\mathcal{T}\}) \\ = \; & \Pi\left(\bigcup_{s\in S}\{\omega \in \Omega : \lambda(s,\delta(\eta(\omega))) = \mathbf{1}_\mathcal{T}\}\right) = \Pi(\{\omega \in \Omega : \bigvee_{s\in S}\lambda(s,\delta(\eta(\omega))) = \mathbf{1}_\mathcal{T}\}) \\ = \; & \tau[\mathbf{1}_\mathcal{T}, \Pi(\{\omega \in \Omega : \bigvee_{s\in S}\lambda(s,\delta(\eta(\omega))) = \mathbf{1}_\mathcal{T}\}) \\ = \; & \bigvee_{t\in T}\tau[t, \Pi(\{\omega \in \Omega : \bigvee_{s\in S}\lambda(s,\delta(\eta(\omega))) \ge t\})]\end{aligned}$$

$$= E_{\tau}\left(\bigvee_{s\in S}\lambda(s,\delta(\eta(\cdot)))\right) = \chi^{MM}(\delta), \tag{4.12}$$

as $\bigvee_{s\in S}\lambda(s,\cdot)$ is also a mapping which takes $S\times D$ into $\{\oslash_{\mathcal{T}},\mathbf{1}_{\mathcal{T}}\}$. □

Corollary 4.1. Let the notations and conditions of Theorem 4.1 hold, let S contain at least two elements, let $S = D$, let $\lambda(s,d) = \oslash_{\mathcal{T}}$, if $s = d, \lambda(s,d) = \mathbf{1}_{\mathcal{T}}$, if $s \neq d$. Then, for each $\delta : E \to D, \chi^{mm}(\delta) = \chi^{MM}(\delta) = \mathbf{1}_{\mathcal{T}}$.

P r o o f. As the conditions of Theorem 4.2 hold, only the relation $\chi^{mm}(\delta) = \mathbf{1}_{\mathcal{T}}$ remains to be proved. (4.12) yields that

$$\begin{aligned}\chi^{mm}(\delta) = \Pi\left(\bigcup_{s\in S}\{\omega\in\Omega : \lambda(s,\delta(\eta(\omega))) = \mathbf{1}_{\mathcal{T}}\}\right)&\\ = \quad \Pi\left(\bigcup_{s\in S}\{\omega\in\Omega : \delta(\eta(\omega)) \neq s\}\right)&\\ = \quad \Pi\left(\Omega - \bigcap_{s\in S}\{\omega\in\Omega : \delta(\eta(\omega)) = s\}\right) = \Pi(\Omega) = \mathbf{1}_{\mathcal{T}},&\end{aligned} \tag{4.13}$$

as S contains at least two elements and for no $\omega\in\Omega$ $\delta(\eta(\omega))$ can take more than one value from S. □

Let us analyse, in more detail, the case when $\chi^{mm}(\delta) < \mathbf{1}_{\mathcal{T}}$ holds, i. e., applying (4.12), when $\chi^{mm}(\delta) = \Pi(\Omega - \bigcap_{s\in S}\{\omega\in\Omega : \lambda(s,\delta(\eta(\omega))) = \oslash_{\mathcal{T}}\}) < \mathbf{1}_{\mathcal{T}}$ is the case. Hence, setting $t_0 = \chi^{mm}(\delta)$, the inequality $\Pi(\{\omega\in\Omega : \lambda(s,\delta(\eta(\omega))) = \mathbf{1}_{\mathcal{T}}\}) \leq t_0 < \mathbf{1}_{\mathcal{T}}$ is valid for each $s\in S$. Setting $t_1 = \Pi(\bigcap_{s\in S}\{\omega\in\Omega : \lambda(s,\delta(\eta(\omega))) = \oslash_{\mathcal{T}}\})$, we obtain that, for each $s\in S, \oslash_{\mathcal{T}} < t_1 \leq \Pi(\{\omega\in\Omega : \lambda(s,\delta(\eta(\omega))) = \oslash_{\mathcal{T}}\})$ holds. If the complete lattice $\mathcal{T} = \langle T,\leq\rangle$ defines a standard linear ordering on T ($t_1 \leq t_2$ or $t_2 \leq t_1$ for each $t_1,t_2\in T$}, then $\Pi(\Omega - A) = \mathbf{1}_{\mathcal{T}}$ easily follows for each $A\in\mathcal{A}$ such that $\Pi(A) < \mathbf{1}_{\mathcal{T}}$ holds. In this particular case, $t_1 = \Pi(\bigcap_{s\in S}\{\omega\in\Omega : \lambda(s,\delta(\eta(\omega))) = \oslash_{\mathcal{T}}\}) = \mathbf{1}_{\mathcal{T}}$ holds for each $s\in S$.

5. POSSIBILISTIC DECISION FUNCTIONS FOR STATE IDENTIFICATION UNDER BAYESIAN CLASSIFICATION

In this section, we will go on in analyzing the most simple lattice-valued possibilistic functions related to the identification of the actual state of the system under investigation. Hence, we suppose that $S = D$ and that the $\{\oslash_{\mathcal{T}},\mathbf{1}_{\mathcal{T}}\}$ loss function is applies, i. e., $\lambda(s,d) = \oslash_{\mathcal{T}}$, if $s = d, \lambda(s,d) = \mathbf{1}_{\mathcal{T}}$ otherwise. We also suppose that the actual state is defined by the value of an S-valued possibilistic variable σ defined on the fixed possibilistic space $\langle\Omega,\mathcal{A},\Pi\rangle$. Under these conditions,

$$\chi^{B}_{\sigma}(\delta) = E_{\tau}\lambda^{*} = \tau[\oslash_{\mathcal{T}},\Pi(\{\omega\in\Omega : \lambda(\sigma(\omega),\delta(\eta(\omega))) \geq \oslash_{\mathcal{T}}\})]$$

$$
\begin{aligned}
&\vee \quad \tau[\mathbf{1}_{\mathcal{T}}, \Pi(\{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega))) \geq \mathbf{1}_{\mathcal{T}}\})] \\
&= \quad \tau[\oslash_{\mathcal{T}}, \mathbf{1}_{\mathcal{T}}] \vee \Pi(\{\omega \in \Omega : \lambda(\sigma(\omega), \delta(\eta(\omega)) = \mathbf{1}_{\mathcal{T}}\}) \\
&= \quad \Pi(\{\omega \in \Omega : \sigma(\omega) \neq \delta(\eta(\omega))\}). \qquad (5.1)
\end{aligned}
$$

Hence, the expected loss is defined by the possibilistically quantified size of the set of those elementary random events $\omega \in \Omega$ for which the decision function δ fails.

The following attributes will be related only to the particular case of possibilistic decision functions as specified in the introduction of this section.

Decision function $\delta : E \to D$ is called *optimal* in $e \in E$, if

$$\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = \delta(e)\}) = \bigvee_{s \in S} \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s\}) \qquad (5.2)$$

holds. Decision function δ is *(uniformly) optimal on* E, if (5.2) holds for every $e \in E$. Decision function δ is *weakly optimal in* $e \in E$, if there is no $s \in S$ with the property

$$\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s\}) > \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = \delta(e)\}) \qquad (5.3)$$

Decision function δ is *(uniformly) weakly optimal on* E, if it is weakly optimal in every $e \in E$.

If the state space S is finite, then there always exists a decision function $\delta : E \to D(= S)$ which is uniformly weakly optimal on E. Indeed, take $A = \{\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s\}) : s \in S\}) \subset T$, denote by A^e the subset of all elements of A which are not dominated by other element of A w.r.t $\leq$ on $\mathcal{T}$, and set

$$S^e = \{s \in S : \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s\}) \in A^e\}. \qquad (5.4)$$

As S is finite, the sets $A^e \subset T$ and $S^e \subset S$ are nonempty for each $e \in E$, hence, a value $\delta(e) \in S^e$ can be chosen. The resulting mapping $\delta : E \to S$ obviously defines a uniformly weakly optimal decision function on E.

If $e \in E$ is such that the set A^e contains at least two elements, there is no decision function optimal in e. Indeed, let $t_1, t_2 \in A^e$ be different, hence, incomparable w.r.t. $\leq$ elements, let, for both $i = 1, 2, s_i \in S^e$ be such that $\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s_i\}) = t_i$. Suppose, in order to arrive at contradiction, that $\delta : E \to S$ is optimal in e. Then the inequality $t_i < \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = \delta(e)\})$ must be valid for both $i = 1, 2$, but this contradicts the assumption that both t_1, t_2 are non-dominated elements in the set of values taken by Π for the given $e \in E$ and for s ranging over S.

Theorem 5.1. Let $\mathcal{T} = \langle T, \leq \rangle$, $\tau : T \times T \to T$, $\langle \Omega, \mathcal{A}, \Pi \rangle, S, D, \sigma : \Omega \to S, E,$ $\eta : \Omega \to E, \lambda : S \times D \to T$ keep their meaning standard in this text with the following further conditions imposed: $S = D$ is finite, $\lambda(s, d) = \oslash_{\mathcal{T}}$, if $s = d, \lambda(s, d) = \mathbf{1}_{\mathcal{T}}$, if $s \neq d$, and $\leq$ defines a linear ordering on T. Let $\delta_{opt} : E \to D(= S)$ be such that $\delta_{opt}(e) \in S^e$ holds for each $e \in E$, where S^e is defined by (5.4). Then δ_{opt} is a uniformly on E optimal decision function, moreover, (5.2) holds for every $e \in E$ and the relation $\chi_\sigma^B(\delta_0) \geq \chi_\sigma^B(\delta_{opt})$ is valid for $\chi_\sigma^B(\delta)$ defined by (5.1).

P r o o f. Using the same way of reasoning as above (below (5.4)), we obtain that each set S^e is nonempty, so that the mapping δ_{opt} satisfying the conditions can be defined. It follows immediately that each such δ_{opt} is uniformly optimal on E. The set A^e is singleton for each $e \in E$, but there may exist more $s \in S$ with the same value $\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s\})$.

Applying (5.1) we obtain that

$$\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = \delta_{opt}(e)\}) = \bigvee_{s \in S} \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s\})$$

$$= \Pi\left(\bigcup_{s \in S} \{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = s\}\right) = \Pi(\{\omega \in \Omega : \eta(\omega) = e\}). \tag{5.5}$$

Let $\delta_0 : E \to D(= S)$ and $e \in E$ be such that $\delta_0(e)$ is *not* in S^e. Consequently,

$$\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = \delta_0(e)\}) < \Pi(\{\omega \in \Omega : \eta(\omega) = e\}) \tag{5.6}$$

follows from (5.5). However, as Π is a $\mathcal{T}$-possibilistic measure on $\mathcal{A}$, also

$$\begin{aligned} \Pi(\{\omega \in \Omega : \eta(\omega) = e\}) &= \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) = \delta_0(e)\}) \\ &\vee \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_0(e)\}) \end{aligned} \tag{5.7}$$

holds. As $\leq$ defines a linear ordering on T, we obtain that

$$\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_0(e)\}) < \Pi(\{\omega \in \Omega : \eta(\omega) = e\}) \tag{5.8}$$

follows for each δ_0 such that $\delta_0(e)$ is not in S^e. Consequently, for each $\delta_0 : E \to D$ $(= S)$ we obtain that

$$\Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_{opt}(e)\}) \leq \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_0(e)\}). \tag{5.9}$$

Combining (5.9) together for different e's, we obtain that for each $\delta_0 : E \to D(= S)$,

$$\begin{aligned} \chi_\sigma^B(\delta_0) &= \Pi(\{\omega \in \Omega : \sigma(\omega) \neq \delta_0(\eta(\omega))\}) \\ &= \Pi\left(\bigcup_{e \in E} \{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_0(\eta(\omega))\}\right) \\ &= \bigvee_{e \in E} \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_0(e)\}) \\ &\geq \bigvee_{e \in E} \Pi(\{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_{opt}(e)\}) \\ &= \Pi\left(\bigcup_{e \in E} \{\omega \in \Omega : \eta(\omega) = e, \sigma(\omega) \neq \delta_{opt}(e)\}\right) \\ &= \Pi(\{\omega \in \Omega : \sigma(\omega) \neq \delta_{opt}(\eta(\omega))\}) = \chi_\sigma^B(\delta_{opt}). \end{aligned} \tag{5.10}$$

The assertion is proved. □

6. ROBUSTNESS OF POSSIBILISTIC DECISION FUNCTIONS OVER LATTICE–VALUED POSSIBILISTIC MEASURES AND LOSS FUNCTIONS

There are numerous decision problems under uncertainty when the demand of robustness is quite intuitive and legitimate. This is to say that a "small enough" change of the values taken by the input functions $(\sigma, \eta, \delta, \lambda)$ of the model in question, or these changes being "rather rarely occurring" result in a "rather small" change of the expected value of the loss function under consideration. The following assertion, simplifying the operations with the expected values of $\mathcal{T}$-valued possibilistic variables may be of use in what follows (cf. [3]).

Lemma 6.1. Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let $\langle \Omega, \mathcal{A}, \Pi \rangle$ be a possibilistic space with complete possibilistic measure Π, let τ be a completely distributive t-norm on T, so that $\tau(t, \bigvee_{s \in A} s) = \bigvee_{s \in A} \tau(t, s)$ holds for each $t \in T$ and each $\emptyset \neq A \subset T$, let $f : \Omega \to T$ be such that $\{\omega \in \Omega : f(\omega) = t\} \in \mathcal{A}$ holds for each $t \in T$. Then the relation

$$E_\tau f(\cdot) = \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : f(\omega) = t\})] \tag{6.1}$$

is valid.

P r o o f. Using the definitions and elementary properties of expected values and t-norm we obtain the inequality

$$E_\tau f(\cdot) \geq \bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : f(\omega) = t\})] \tag{6.2}$$

Let $t, t_1 \in T$, let $t \leq t_1$, then the inequalities

$$\begin{aligned} & \tau[t, \Pi(\{\omega \in \Omega : f(\omega) = t_1\})] \leq \tau[t_1, \Pi(\{\omega \in \Omega : f(\omega) = t_1\})] \\ \leq & \bigvee_{t_1 \in T} \tau[t_1, \Pi(\{\omega \in \Omega : f(\omega) = t_1\})] \end{aligned} \tag{6.3}$$

and

$$\begin{aligned} & \tau\left[t, \bigvee_{t_1 \geq t} \Pi(\{\omega \in \Omega : f(\omega) = t_1\})\right] = \tau[t, \Pi(\{\omega \in \Omega : f(\omega) \geq t_1\})] \\ \leq & \bigvee_{t_1 \in T} \tau[t_1, \Pi(\{\omega \in \Omega : f(\omega) = t_1\})] \end{aligned} \tag{6.4}$$

are valid for each $t \in T$. Hence, the inequality

$$\bigvee_{t \in T} \tau[t, \Pi(\{\omega \in \Omega : f(\omega) \geq t\})] = E_\tau f(\cdot) \leq \bigvee_{t_1 \in T} \tau[t_1, \Pi(\{\omega \in \Omega : f(\omega) = t_1\})] \tag{6.5}$$

easily follows; combining (6.2) and (6.5) we complete the proof. $\square$

Within the framework of statistical decision functions the values of loss functions and their expected values are real numbers, so that the differences between the qualities of two decision functions can be defined by the absolute value of the difference of the two real numbers in question. In the case of values from the complete lattice let us take an inspiration in the idea of set-theoretic operation of symmetric difference which can be still generalized using the operation of residuation and the notion of residuum. Hence, given a complete lattice $\mathcal{T} = \langle T, \leq \rangle$ and a t-norm τ on T, for each $t \in T$ the value

$$t^{\tau,c} = \bigvee \{s \in T : \tau(s,t) = \emptyset\} \tag{6.6}$$

is defined (c stands for "complement", the role of which $t^{\tau,c}$ is to play) and called *τ-residuum* of t w.r.t. $\mathcal{T}$ (cf. [3] or [7] for a more detailed investigation of the residuation). As can be easily seen, for each t-norm τ on T the relations $\emptyset_{\mathcal{T}}^{\tau,c} = \mathbf{1}_{\mathcal{T}}$ and $\mathbf{1}_{\mathcal{T}}^{\tau,c} = \emptyset_{\mathcal{T}}$ hold, moreover, if τ is completely distributive, then $\tau(t, t^{\tau,c}) = \emptyset_{\mathcal{T}}$ holds for each $t \in T$, as

$$\begin{aligned} \tau(t, t^{\tau,c}) &= \tau\left(t, \bigvee \{s \in T : \tau(s,t) = \emptyset_{\mathcal{T}}\}\right) \\ &= \bigvee \{\tau(t,s) : s \in T, \tau(s,t) = \emptyset_{\mathcal{T}}\} = \bigvee \{\emptyset_{\mathcal{T}}\} = \emptyset_{\mathcal{T}}. \end{aligned} \tag{6.7}$$

In order to simplify our reasoning let us limit ourselves to the case of the greatest (w.r.t. $\leq$) t norm defined by the infimum $\wedge$ on T, consequently, the index τ in $t^{\tau,c}$ will be omitted. We will also suppose that $\wedge$ is a completely distributive t-norm on T. It does not follow, in general, that $t \vee t^c = \mathbf{1}_{\mathcal{T}}$ for each $t \in T$. Indeed, take $\mathcal{T} = \langle [0,1], \leq \rangle$, where $\leq$ denotes the standard linear ordering on $[0,1]$, then obviously $x^c = 0$, if $x > 0$, and $0^c = 1$ holds, hence, $x \wedge x^c = 0$ for each $x \in [0,1]$, but $x \vee x^c = 0 \vee x = x < 1$ for each $0 < x < 1$.

So, keeping in mind the idea of symmetric difference, set $\Delta(s,t) = (s \wedge t^c) \vee (t \wedge s^c)$ for each $s, t \in T$. As proved in [9], the relations $\Delta(t,t) = \emptyset_{\mathcal{T}}$ (reflexivity), $\Delta(s,t) = \Delta(t,s)$ (symmetry), and $\Delta(s,t) \leq \Delta(s,u) \vee \Delta(u,t)$ (triangular inequality in the lattice sense) are valid for each $s, t, u \in T$. Let us note that the condition $t^c \wedge t \equiv \emptyset_{\mathcal{T}}$ is substantial when proving these relations.

Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice such that $\wedge$ is completely distributive, let $\langle \Omega, \mathcal{A}, \Pi \rangle$ be a possibilistic space with a complete Π, let $f_1, f_2 : \Omega \to T$ be mappings such that, for both $i = 1,2$ and each $t \in T$, $\{\omega \in \Omega : f_i(\omega) = t\} \in \mathcal{A}$ holds, consequently, also $\{\omega \in \Omega : f_i(\omega) \in A\} \in \mathcal{A}$ is the case for both $i = 1,2$ and each $A \subset T$. Set

$$D_1(f_1, f_2) = E_{\wedge} \Delta(f_1(\cdot), f_2(\cdot)) d\Pi = \bigvee_{t \in T} [t \wedge \Pi(\{\omega \in \Omega : \Delta(f_1(\omega), f_2(\omega)) \geq t\})], \tag{6.8}$$

$$D_2(f_1, f_2) = \Delta(E_{\wedge} f_1(\cdot)\, d\Pi, E_{\wedge} f_2(\cdot)\, d\Pi). \tag{6.9}$$

Lemma 6.2. Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice such that $t \wedge t^c \equiv \emptyset_{\mathcal{T}}$ and $t \vee t^c \equiv \mathbf{1}_{\mathcal{T}}$ holds. Then $(s \wedge t)^c = s^c \vee t^c$ holds for each $s, t \in T$.

P r o o f. Let $t_1 \leq t_2$, then for each $s \in T, s \wedge t_2 = \oslash_{\mathcal{T}}$ implies that $s \wedge t_1 = \oslash_{\mathcal{T}}$, and $t_1^c \geq t_2^c$ follows. So, for each $s, t \in T, (s \wedge t)^c \geq s^c \vee t^c$ holds for each complete lattice. If, moreover, $t \wedge t^c \equiv \oslash_{\mathcal{T}}$ and $t \vee t^c \equiv \mathbf{1}_{\mathcal{T}}$, holds, then

$$\begin{aligned} (s \wedge t)^c &= (s \wedge t)^c \wedge \mathbf{1}_{\mathcal{T}} = (s \wedge t)^c \wedge (s^c \vee s) = ((s \wedge t)^c \wedge s^c) \vee ((s \wedge t)^c \wedge s) \\ &= s^c \wedge ((s \wedge t)^c \wedge s), \end{aligned} \tag{6.10}$$

moreover, $((s \wedge t)^c \wedge s) \wedge t = (s \wedge t) = \oslash_{\mathcal{T}}$, so that $(s \wedge t)^c \wedge s \leq t^c$ follows. So, $(s \wedge t)^c \leq s^c \vee t^c$ follows and the assertion is proved. □

Theorem 6.1. Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice such that $t \wedge t^c \equiv \oslash_{\mathcal{T}}$ and $t \vee t^c \equiv \mathbf{1}_{\mathcal{T}}$ holds, let Π be a complete $\mathcal{T}$-valued possibilistic measure on the power-set $\mathcal{P}(\Omega)$, let $f_1, f_2 : \Omega \to T$, let $D_1(f_1, f_2)$ and $D_2(f_1, f_2)$ be defined by (6.8) and (6.9). Then the relation $D_2(f_1, f_2) \leq D_1(f_1, f_2)$ holds.

P r o o f. Setting $\pi(\omega) = \Pi(\{\omega\})$ for each $\omega \in \Omega$ and applying (6.1), we obtain that, for each $f : \Omega \to T$,

$$E_{\wedge} f(\cdot) = \bigvee_{t \in T} [t \wedge \bigvee \{\pi(\omega) : f(\omega) = t\}] = \bigvee_{\omega \in \Omega} (f(\omega) \wedge \pi(\omega)). \tag{6.11}$$

Hence, Lemma 6.2 yields that

$$\begin{aligned} & D_2(f_1, f_2) = \Delta(E_{\wedge} f_1 d\Pi, E_{\wedge} f_2 d\Pi) = \\ = \; & ((E_{\wedge} f_1 d\Pi) \wedge (E_{\wedge} f_2 d\Pi)^c) \vee ((E_{\wedge} f_2 d\Pi) \wedge (E_{\wedge} f_1 d\Pi)^c) \\ = \; & \left[\left(\bigvee_{\omega \in \Omega} (f_1(\omega) \wedge \pi(\omega)) \right) \wedge \left(\bigvee_{\omega \in \Omega} (f_2(\omega) \wedge \pi(\omega)) \right)^c \right] \\ \vee \; & \left[\left(\bigvee_{\omega \in \Omega} (f_2(\omega) \wedge \pi(\omega)) \right) \wedge \left(\bigvee_{\omega \in \Omega} (f_1(\omega) \wedge \pi(\omega)) \right)^c \right] \\ \leq \; & \left[\left(\bigvee_{\omega \in \Omega} (f_1(\omega) \wedge \pi(\omega)) \right) \wedge (f_2(\omega) \wedge \pi(\omega))^c \right] \\ \vee \; & \left[\left(\bigvee_{\omega \in \Omega} (f_2(\omega) \wedge \pi(\omega)) \right) \wedge (f_1(\omega) \wedge \pi(\omega))^c \right] \\ = \; & \left[\left(\bigvee_{\omega \in \Omega} (f_1(\omega) \wedge \pi(\omega)) \right) \wedge (f_2(\omega)^c \vee (\pi(\omega))^c) \right] \\ \vee \; & \left[\left(\bigvee_{\omega \in \Omega} (f_2(\omega) \wedge \pi(\omega)) \right) \wedge ((f_1(\omega)^c \vee (\pi(\omega))^c) \right] \\ = \; & \left[\bigvee_{\omega \in \Omega} (f_1(\omega) \wedge \pi(\omega) \wedge (f_2(\omega))^c) \right] \vee \left[\bigvee_{\omega \in \Omega} (f_1(\omega) \wedge \pi(\omega) \wedge (\pi(\omega))^c) \right] \end{aligned}$$

$$
\begin{aligned}
&\vee \left[\bigvee_{\omega\in\Omega} (f_2(\omega)\wedge\pi(\omega)\wedge(f_1(\omega))^c)\right] \vee \left[\bigvee_{\omega\in\Omega} (f_2(\omega)\wedge\pi(\omega)\wedge(\pi(\omega))^c)\right] \\
&= \left[\bigvee_{\omega\in\Omega} (f_1(\omega)\wedge(f_2(\omega))^c\wedge\pi(\omega))\right] \vee \oslash_{\mathcal{T}} \\
&\vee \left[\bigvee_{\omega\in\Omega} (f_2(\omega)\wedge(f_1(\omega))^c\wedge\pi(\omega))\right] \vee \oslash_{\mathcal{T}} \\
&\leq \bigvee_{\omega\in\Omega} [((((f_1(\omega))\wedge(f_2(\omega))^c)\vee(((f_2(\omega))\wedge(f_1(\omega))^c))\wedge\pi(\omega))] \\
&= \bigvee_{\omega\in\Omega} [(\Delta(f_1(\omega),f_2(\omega)))\wedge\pi(\omega)] = E_\wedge(\Delta(f_1(\cdot),f_2(\cdot)))\,d\Pi = D_1(f_1,f_2).
\end{aligned}
\tag{6.12}
$$

The assertion is proved. □

The equality $D_1(f_1,f_2)=D_2(f_1,f_2)$ does not hold in general. Indeed, let $\Omega=\{\omega_1,\omega_2\}, f_1(\omega_1)=f_2(\omega_2)=\oslash_{\mathcal{T}},\ f_1(\omega_2)=f_2(\omega_1)=\mathbf{1}_{\mathcal{T}}$, let $\Pi(\emptyset)=\oslash_{\mathcal{T}}$, $\Pi(A)=\mathbf{1}_{\mathcal{T}}$ for each $\emptyset\neq A\subset\Omega$, so that $\pi(\omega_1)=\pi(\omega_2)=\mathbf{1}_{\mathcal{T}}$. Then

$$
\begin{aligned}
E_\wedge f_1(\cdot)\,d\Pi &= (f_1(\omega_1)\wedge\pi(\omega_1))\vee(f_1(\omega_2)\wedge\pi(\omega_2)) = f_1(\omega_1)\vee f_1(\omega_2) \\
&= \mathbf{1}_{\mathcal{T}} = f_2(\omega_1)\vee f_2(\omega_2) = E_\wedge f_2(\cdot)d\Pi,
\end{aligned}
\tag{6.13}
$$

so that $D_2(f_1,f_2)=\Delta(E_\wedge f_1(\cdot)d\Pi, E_\wedge f_2(\cdot)d\Pi)=\Delta(\mathbf{1}_{\mathcal{T}},\mathbf{1}_{\mathcal{T}})=\oslash_{\mathcal{T}}$. But,

$$
\Delta(f_1(\omega_1),f_2(\omega_1))=\Delta(\oslash_{\mathcal{T}},\mathbf{1}_{\mathcal{T}})=\mathbf{1}_{\mathcal{T}}=\Delta(f_1(\omega_2),f_2(\omega_2)),
\tag{6.14}
$$

hence,

$$
\Delta(f_1,f_2)=E_\wedge\Delta(f_1(\cdot),f_2(\cdot))\,d\Pi=\mathbf{1}_{\mathcal{T}}>\oslash_{\mathcal{T}}=D_2(f_1,f_2).
\tag{6.15}
$$

Theorem 6.2. Let $\mathcal{T}=\langle T,\leq\rangle$ be a complete lattice such that $t^c\wedge t\equiv\oslash_{\mathcal{T}}$ holds, let $\langle\Omega,\mathcal{A},\Pi\rangle$ be a $\mathcal{T}$-possibility space with Π complete, let $f_1,f_2:\Omega\to T$ be such that, for both $i=1,2$ and for each $t\in T,\{\omega\in\Omega: f_i(\omega)=t\}\in\mathcal{A}$ holds, let $A=\{\omega\in\Omega: f_1(\omega)\neq f_2(\omega)\}\in\mathcal{A}$, let $I_A(\omega)=\mathbf{1}_{\mathcal{T}}$, if $\omega\in A$, $I_A(\omega)=\oslash_{\mathcal{T}}$, if $\omega\in\Omega-A$. Then

$$
D_1(f_1,f_2)=E_\wedge(I_A(\cdot)\wedge\Delta(f_1(\cdot),f_2(\cdot)))\,d\Pi\leq\left[\bigvee_{\omega\in A}\Delta(f_1(\omega),f_2(\omega))\right]\wedge\Pi(A).
\tag{6.16}
$$

P r o o f. If $\omega\in\Omega-A$, then $\Delta(f_1(\omega),f_2(\omega))=\oslash_{\mathcal{T}}$ obviously holds, if $\omega\in A$, then $I_A(\omega)=\mathbf{1}_{\mathcal{T}}$ and $I_A(\omega)\wedge\Delta(f_1(\omega),f_2(\omega))=\Delta(f_1(\omega),f_2(\omega))$ follows. Hence,

$$
\begin{aligned}
D_1(f_1,f_2) &= E_\wedge(I_A(\cdot)\wedge\Delta(f_1(\cdot),f_2(\cdot)))\,d\Pi\leq E_\wedge I_A(\cdot)\,d\Pi \\
&= \bigvee_{t\in T}[t\wedge\Pi(\{\omega\in\Omega: I_A(\omega)\geq t\})]=\Pi(A)
\end{aligned}
\tag{6.17}
$$

obviously holds and also the inequality $D_1(f_1, f_2) \le \bigvee_{\omega \in A} \Delta(f_1(\omega), f_2(\omega))$ is obvious, so that the assertion is proved. $\square$

In the particular case of Bayesian decision making we can define, for two decision functions δ_1, δ_2.

$$D_1^*(\delta_1, \delta_2) = E_\wedge \Delta(\lambda(\sigma(\cdot), \delta_1(\eta(\cdot))), \lambda(\sigma(\cdot), \delta_2(\eta(\cdot))))\, d\Pi, \tag{6.18}$$

$$D_2^*(\delta_1, \delta_2) = \Delta(\chi_\sigma^B(\delta_1), \chi_\sigma^B(\delta_2)), \tag{6.19}$$

where $\chi_\sigma^B(\delta_i) = E_\wedge \lambda(\sigma(\cdot), \delta_i(\eta(\cdot)))\, d\Pi$ for both $i = 1, 2$. It is just a matter of routine (left to the reader) to rewrite Theorems 6.1 and 6.2 for the criteria $D_i^*(\delta_1, \delta_2)$ instead of $D_i(f_1, f_2)$.

Informally told, under our possibilistic setting, the obtained possibilistic decision functions are robust in the sense that if the losses suffered when applying different decision functions δ_1, δ_2 differ only rarely, or when the differences between the corresponding losses are small, also the qualities of decision functions δ_1, δ_2 do not differ too much from each other. As a matter of fact, in the case of possibilistic decision functions the robustness w.r.t. differences of suffered losses in rarely occurring cases is still more strong than as claimed by Theorems 6.1 and 6.2 applied to $D_i(\delta_1, \delta_2), i = 1, 2$, as this example demonstrates.

Let the notations and conditions introduced in Theorems 6.1, 6.2, (6.18) and (6.19) hold, let δ_1, δ_2 be such that the losses suffered when δ_1 or δ_2 applied differ only when s_1 is the actual state, hence, let us suppose that

$$A = \{\omega \in \Omega : \lambda(\sigma(\omega), \delta_1(\eta(\omega))) \neq \lambda(\sigma(\omega), \delta_2(\eta(\omega)))\} \subset \{\omega \in \Omega : \sigma(\omega) = s_1\}) \tag{6.20}$$

and $A \in \mathcal{A}$ holds. Consider the case when the inequality

$$\Pi(A) \le \lambda(\sigma(\omega), \delta_1(\eta(\omega))) \wedge \lambda(\sigma(\omega), \delta_2(\eta(\omega))) \tag{6.21}$$

holds for each $\omega \in A$. Then $\chi_\sigma^B(\delta_1) = \chi_\sigma^B(\delta_2)$.

Indeed, for both $i = 1, 2$ we obtain that

$$\chi_\sigma^B(\delta_i) = \bigvee_{\omega \in A} [\lambda(\sigma(\omega), \delta_i(\eta(\omega))) \wedge \pi(\omega)] \vee \bigvee_{\omega \in \Omega - A} [\lambda(\sigma(\omega), \delta_i(\eta(\omega))) \wedge \pi(\omega)]. \tag{6.22}$$

As the loss function λ is supposed to take just the values $\oslash_{\mathcal{T}}$ or $\mathbf{1}_{\mathcal{T}}$, for each $\omega \in A$ and for both $i = 1, 2$ the relation $\pi(\omega) = \Pi(\{\omega\}) \le \Pi(A) \le \lambda(\sigma(\omega), \delta_i(\eta(\omega))) = \mathbf{1}_{\mathcal{T}}$ holds, so that $\lambda(\sigma(\omega), \delta_i(\eta(\omega))) \wedge \pi(\omega) = \pi(\omega)$ for both $i = 1, 2$, $\lambda(\sigma(\omega), \delta_i(\eta(\omega))) \wedge \pi(\omega) = \pi(\omega)$ holds for each $\omega \in A$. Consequently, (6.22) yields that

$$\chi_\sigma^B(\delta_i) = \left[\bigvee_{\omega \in \Omega - A} (\lambda(\sigma(\omega), \delta_i(\eta(\omega))) \wedge \pi(\omega)) \right] \vee \Pi(A) \tag{6.23}$$

holds for both $i = 1, 2$. As the values $\lambda(\sigma(\omega), \delta_i(\eta(\omega)))$ are the same for both $i = 1, 2$, if $\omega \in \Omega - A$, then the identity $\chi_\sigma^B(\delta_1) = \chi_\sigma^B(\delta_2)$ follows.

7. CONCLUSIONS

We have submitted an attempt to re-write the model of statistical decision making under uncertainty for the case when the underlying uncertainty is quantified and processed using lattice-valued possibilistic measures, so emphasizing rather the qualitative than the quantitative aspects of the degrees of uncertainty under consideration. A complete lattice, chosen as an appropriate structure over these uncertainty degrees, is perhaps the most specific mathematical structure still covering the two most often used structures for quantification and processing of sizes: the unit interval of real numbers with their standard linear ordering, and the complete Boolean algebra with the corresponding partial ordering (as a matter of fact, a power-set over a nonempty space partially ordered by the set inclusion). We have introduced the possibilistic modifications of the two classical criteria used in order to define and quantify the quantities of procedures for decision making under uncertainty: the minimax (the worst-case) principle and the possibilistic variant of the Bayes principle. For both the approaches we have stated and proved the most elementary properties of the possibilistic decision functions under consideration.

Among the possible directions for further investigation let us mention explicitly just the following ones: (i) to apply the general model from above to a particular decision problem under uncertainty, (ii) to consider richer and more powerful structures for uncertainty degrees, but also to check which of the results achieved above remain to be valid in weaker structures (lattices, lower or upper semilattices, posets, ...), (iii) to analyze the possibilistic variant of the well-known Laplace principle, if any exists, (iv) some relations between belief functions (Dempster–Shafer theory) and possibilistic decision functions seem to be worth being analyzed.

The author is much indebted to an anonymous reviewer whose comments, remarks, and suggestions have helped to improve remarkably the final version of this paper.

ACKNOWLEDGEMENT

This work was sponsored by the grant No. IAA 100300503 of the Grant Agency of Academy of Sciences of the Czech Republic. The work was partly supported by the Institutional Research Plan AV0Z 10300504 "Computer Science for the Information Society: Models, Algorithms, Applications".

(Received February 2, 2005.)

REFERENCES

[1] G. Birkhoff: Lattice Theory. Third edition. Amer. Math. Society, Providence, RI 1967.
[2] D. Blackwell and N. A. Girshick: Theory of Games and Statistical Decisions. Wiley, New York, 1954.
[3] G. De Cooman: Possibility theory I, II, III. Internat. J. Gen. Systems *25* (1997), 4, 291–323, 325–351, 353–371.
[4] D. Dubois and H. Prade: Théorie des Possibilités – Applications à la Représentation des Connaissances en Informatique. Mason, Paris 1985.

[5] D. Dubois, H. Nguyen, and H. Prade: Possibility theory, probability theory and fuzzy sets: misunderstandings, bridges and gaps. In: The Handbook of Fuzzy Sets Series (D. Dubois and H. Prade, eds.), Kluwer Academic Publishers, Boston, 2000, pp. 343–438.
[6] R. Faure and E. Heurgon: Structures Ordonnées et Algèbres de Boole. Gauthier-Villars, Paris 1971.
[7] P. Hájek: Metamathematics of Fuzzy Logic. Kluwer Academic Publishers, Boston 1998.
[8] P. R. Halmos: Measure Theory. D. van Nostrand, New York – Toronto – London 1950.
[9] I. Kramosil: Extensions of partial lattice-valued possibility measures. Neural Network World *13* (2003), 4, 361–384.
[10] M. Loève: Probability Theory. D. van Nostrand, New York – Toronto – London 1960.
[11] R. Sikorski: Boolean Algebras. Second edition. Springer-Verlag, Berlin – Göttingen – Heidelberg – New York 1964.
[12] A. Wald: Statistical Decision Functions. Wiley, New York 1950.

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