Entropy Function for Chained Lattice-Valued Possibilistic Distributions and Their Non-Interactive Products

Kramosil, Ivan
2010

Institute of Computer Science
Academy of Sciences of the Czech Republic

Entropy Function for Chained Lattice-Valued Possibilistic Distributions and Their Non-Interactive Products

Ivan Kramosil

Technical report No. 1066

March 2010

Pod Vodárenskou věží 2, 182 07 Prague 8, phone: +420 266 051 111, fax: +420 286 585 789,
e-mail:ics@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

Entropy Function for Chained Lattice-Valued Possibilistic Distributions and Their Non-Interactive Products

Ivan Kramosil

Technical report No. 1066

March 2010

Abstract:

Given a possibilistic distribution on a nonempty space Ω with possibility degrees in a chained complete lattice, lattice-valued entropy function for such distribution is defined as the expected value (in the sense of Sugeno possibilistic integral) of the lattice-valued function ascribing to each $\omega \in \Omega$ the possibilistic measure of its complement $\Omega - \{\omega\}$. This entropy function is proved to possess some properties syntactically close to those of Shannon probability entropy, even if in other aspects both the entropy functions prove qualitative differences. In particular, explicit expressions are proved for lattice-valued entropy of possibilistically independent (non-interactive, in other terms) products of lattice valued possibilistic distributions.

Keywords:
complete lattice, chained lattice, lattice-valued possibilistic distribution, possibilistic expected value, entropy function, non-interactive product of lattice-valued possibilistic distributions

1 INTRODUCTION, MOTIVATION, PRELIMINARIES

Research effort leading to the notion of possibility (or possibilistic) distributions and measures to be investigated also below originated in the famous pioneering work by L. A. Zadeh [18] introducing and analyzing the notion of real-valued fuzzy sets. For our purposes, all the mathematical, philosophical and methodological aspects and problems related to fuzzy sets may and will be omitted ([6] can be recommended as a good survey) and real-valued fuzzy subsets of a universe Ω will be identified with mappings $A : \Omega \to [0, 1]$; as a rule, only normalized fuzzy sets for which the condition $\bigvee_{\omega\in\Omega} A(\omega) = 1$ holds will be taken into consideration. Here and below, $\bigvee$ ($\bigwedge$, resp.) denotes the supremum (infimum, resp.) operation no matter whether the standard operations in $[0, 1]$ or operations defined by a (partial) ordering in some non-numerical structures are considered, this should be always easy to recognize from the context.

Given a fuzzy subset A of Ω and a crisp subset $B \subset \Omega$, L. A. Zadeh quantified, in [17], the total amount of fuzziness defined by A and contained in B by the value $\Pi(B) = \bigvee_{\omega\in B} A(\omega)$ and the resulting mapping $\Pi : \mathcal{P}(\Omega) \to [0, 1]$, induced on the power-set $\mathcal{P}(\Omega)$ of all subsets of Ω, he called the *possibility measure induced by* A (in what follows, we prefer the adjective "possibilistic" to avoid a confusion with an informal use of the word "possibility"). In order to emphasize parallel syntactical features between probabilistic and possibilistic measures the primary fuzzy set A is often denoted by π and it is called *(real-valued) possibilistic distribution* on Ω defining the possibilistic measure Π on $\mathcal{P}(\Omega)$.

For a number of formally mathematical as well as practical reasons (applications of fuzzy sets to various situations and problems from the real world around), the idea of fuzzy sets with non-numerical membership degrees, namely with membership degrees from a complete lattice, appeared in [8] as soon as in 1967 (cf. also [2], [5] or elsewhere for a more detailed discussion of the related matters).

The reader is supposed to be familiar with the notion of *partially ordered set (p.o.set* or *poset)* defined by the pair $\langle T, \leq_{\mathcal{T}}\rangle = \mathcal{T}$, where $\leq_{\mathcal{T}}$ is a reflexive, antisymmetric and transitive binary relation on T (subset of $T \times T$). For $S \subset T$ the supremum $\bigvee^{\mathcal{T}} S = \bigvee^{\mathcal{T}}_{s\in S} s$ and infimum $\bigwedge^{\mathcal{T}} S = \bigwedge^{\mathcal{T}}_{s\in S} s$, if defined, are defined uniquely in the standard way. If no misunderstanding menaces, we omit the index $\mathcal{T}$ in $\leq_{\mathcal{T}}, \bigvee^{\mathcal{T}}$, and $\bigwedge^{\mathcal{T}}$. Let $\mathcal{T} = \langle T, \leq\rangle$ be a p.o.set. As can be easily seen, in general neither $\bigvee S$ nor $\bigwedge S$ need be defined for any $S \subset T$. P.o.set $\mathcal{T} = \langle T, \leq\rangle$ is called a *lattice*, if for each $s_1, s_2 \in T, s_1 \vee s_2$ as well as $s_1 \wedge s_2$ is defined in T; it follows immediately that in this case the values $\bigvee S$ and $\bigwedge S$ are defined in T for each *finite* nonempty subset $S \subset T$. If both the values $\bigvee S$ and $\bigwedge S$ are defined in T for each $\emptyset \neq S \subset T$, the p.o.set $\mathcal{T} = \langle T, \leq\rangle$ is called *complete lattice* and just complete lattices will play the role of our key tool when quantifying and processing possibility degrees with non-numerical values. Let us note that both the structures most often used as tools for quantification, $\langle [0, 1], \leq\rangle$ and $\langle \mathcal{P}(X), \subset\rangle$, define complete lattices.

If $\mathcal{T} = \langle T, \leq\rangle$ is a complete lattice, the element $\bigvee T (\bigwedge T$, resp.) is called the *unit (element)* (the *zero (element),* resp.) of $\mathcal{T}$ and denoted by $\mathbf{1}_{\mathcal{T}}$ ($\oslash_{\mathcal{T}}$, resp.). In this case the index $\mathcal{T}$ will be saved in order to avoid a confusion with the standard notation for integers 1 and 0. For the empty subset $\emptyset$ of T we define, by convention consistent with the properties of complete lattices, $\bigvee \emptyset = \bigwedge T = \oslash_{\mathcal{T}}$, and $\bigwedge \emptyset = \bigvee T = \mathbf{1}_{\mathcal{T}}$.

Partial ordering relation $\leq$ defined on T is called *linear* (or a *chain*), if $s \leq t$ or $t \leq s$ holds for each $s, t \in T$, hence, due to the antisymmetry axiom imposed on partial ordering, if $s < t$ or $t < s$ holds for each $s \neq t$. Here and below, $s < t$ means that $s \leq t$ and $s \neq t$ holds together, instead of $s \leq t$ and $s < t$ we will write also $t \geq s$ and $t > s$. Complete lattice $\mathcal{T} = \langle T, \leq\rangle$ with linear ordering $\leq$ is called *complete chained lattice.*

Let us note that the complete lattice $\langle [0, 1], \leq\rangle$ is chained (the standard ordering $\leq$ on $[0, 1]$ is linear, but for $\langle \mathcal{P}(X), \subset\rangle$ it is not the case up to the trivial possibility when X is a singleton.

Given a complete lattice $\mathcal{T} = \langle T, \leq\rangle$ and a nonempty set Ω, a mapping $\pi : \Omega \to T$ such that $\bigvee_{\omega\in\Omega} \pi(\omega) = \mathbf{1}_{\mathcal{T}}$ holds is called $\mathcal{T}$*-(valued normalized) possibilistic distribution on* Ω *or over* Ω. As a matter of fact, it is nothing else than a lattice-valued normalized fuzzy subset of Ω as conceived by J. A. Goguen in [8]. The mapping $\Pi : \mathcal{P}(\Omega) \to T$ ascribing to each (crisp) subset $A \subset \Omega$ the value $\Pi(A) = \bigvee_{\omega\in A} \pi(\omega)$ is called the *possibilistic measure induced by* π *on* $\mathcal{P}(\Omega)$ (for $A = \emptyset$ the convention $\Pi(\emptyset) = \oslash_{\mathcal{T}}$ applies). The following immediate consequence is easy to check. If $\mathcal{T} = \langle T, \leq\rangle$

is a complete chained lattice and, if π is a $\mathcal{T}$-possibilistic distribution on a nonempty space Ω, then for each $A \subset \Omega$ either $\Pi(A) = \mathbf{1}_{\mathcal{T}}$ or $\Pi(\Omega - A) = \mathbf{1}_{\mathcal{T}}$ (or both) holds, like as it is the case for real-valued possibilistic measures taking their values in the unit interval of reals. Indeed, for each $A \subset \Omega$ the relation $\Pi(A) \vee \Pi(\Omega - A) = \mathbf{1}_{\mathcal{T}}$ holds, but for linear ordering $\leq$ on T either $t_1 \vee t_2 = t_1$ or $t_1 \vee t_2 = t_2$ holds for each $t_1, t_2 \in T$.

The following notion will be very important on our further reasonings. A $\mathcal{T}$-possibilistic distribution π on Ω is called *isolated*, if there exists $\omega_0 \in \Omega$ such that $\Pi(\Omega - \{\omega_0\}) = \bigvee_{\omega \in \Omega, \omega \neq \omega_0} \pi(\omega) < \mathbf{1}_{\mathcal{T}}$ holds. If $\mathcal{T} = \langle T, \leq \rangle$ is a complete chained lattice isolated in ω_0, then $\pi(\omega) = \mathbf{1}_{\mathcal{T}}$ if and only if $\omega = \omega_0$ (the assertion is almost obvious, cf. also the proof of Lemma 4.2, below and the discussion at the end of Section 4).

In what follows, we always assume that both the spaces Ω and T contain at least two elements in order to avoid degenerated and trivial cases.

Our aim will be, in this paper, to propose and analyze, in more detail, a $\mathcal{T}$-valued global characteristic of the total amount of uncertainty (in the sense of fuzziness) contained in a given $\mathcal{T}$-possibilistic distribution π on Ω. This characteristic should stand as close as possible to the principles on which Shannon entropy, taken as a global characteristic of the total amount of randomness contained in a probability distribution, is based. The main restriction imposed on our constructions will be that only notions and tools definable within the framework of lattice theory may be applied. These apriori methodological assumptions are discussed, in more detail, in Section 2. In Section 3, two alternative approaches to definition of the notion of entropy for possibilistic distributions, proposed and investigated by other authors, are discussed and confronted with the methodological demands from Section 2. Our variant of lattice-valued entropy function meeting the demands in question is introduced in Section 4 and its properties, mainly as far as the possibilistically independent (non-interactive, in other terms) products of $\mathcal{T}$-possibilistic distributions are concerned, are investigated in Section 5.

2 Shannon-Like Lattice-Valued Entropy Function with Strongly Reduced Ontologically Independent Assumptions

Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let Ω be a nonempty space, and let $\pi : \Omega \to T$ be a $\mathcal{T}$-(lattice-valued normalized possibilistic) distribution on Ω, so that $\bigvee_{\omega \in \Omega} \pi(\omega) = \mathbf{1}_{\mathcal{T}}(= \bigvee T)$ holds. Our aim will be, in what follows, to propose and analyze, in more detail, a mapping H ascribing to π a value $H(\pi)$ from T in a way expressing adequately the total amount of uncertainty (in the sense of vagueness or fuzziness) contained in the possibilistic distribution π. The two meta-theoretical conditions ultimately imposed on our effort will read as follows.

First, not only the value $H(\pi)$ itself, but all the values used when introducing and when analyzing the properties of the mapping H should be values from T, and the only operations and relations applied when processing these values should be the relation of partial ordering $\leq$ on T, the operations of supremum ($\bigvee$) and infimum ($\bigwedge$) induced by $\leq$ in T, and perhaps further relations and operations secondary defined on the ground of $\vee$, $\wedge$, and $\leq$ taken as the primary tools. Consequently, the complete lattice $\mathcal{T}$ will be considered at the most general level in the sense that only the conditions definable within the framework of lattice-based operations and relations as specified above may be imposed on $\mathcal{T}$ in order to investigate some particular cases of complete lattices (e.g., the assumption that $\leq$ defines a linear ordering, i.e., that $\mathcal{T} = \langle T, \leq \rangle$ is a chained lattice, can be defined within these restrictions). It is perhaps worth being noted explicitly that according to this methodological assumption, which can be taken as the *principle of elimination of all ontologically independent assumptions* from our mathematical model, *no* mappings of T into real numbers as well as *no* algebraic operations over such real-valued images or relations among them will be supposed and applied. Even when $\mathcal{T} = \langle [0,1], \leq \rangle$ is the particular case of complete lattice under consideration, no arithmetical or algebraic operations and relations on $[0,1]$ except the lattice-based ones will be admitted.

The second meta-theoretical assumption will read that the mapping H should stand as close to the Shannon probabilistic entropy function and should share as much properties possessed by the Shannon entropy (or properties in a reasonable sense close or analogous to the properties of the Shannon entropy) as it is possible within the restricted framework of mathematical models and tools remaining at our disposal when keeping in mind the principle of elimination of ontologically independent inputs as introduced one paragraph above. If is why a very brief re-calling of the basic ideas on which Shannon entropy relies seems to be useful.

Let $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ be a finite space, let p be a probability distribution on Ω, i.e., $p : \Omega \to [0,1], \sum_{i=1}^{n} p(\omega_i) = 1$ holds. Taking a random sample from p and obtaining, say, ω_{i_0} as the result, our surprise on this result (informally told) will be the greater the smaller $p(\omega_{i_0})$ is, i.e., the smaller had been our apriori expectation that just ω_{i_0} occurs. Calling a numerical value quantifying the degree of surprise by aposteriori information on the distribution p obtained when obtaining the sample value ω_{i_0} and denoting this information by a real-valued function $f : \Omega \to R^+ = [0, \infty]$, we obtain that f should be a decreasing function of $p(\omega)$ such that $f(\omega) = 0$, if $p(\omega) = 1$ (the occurrence of random event expected with the probability 1 brings no surprise and no new information); and $f(\omega)$ should tend to ∞, if $p(\omega)$ tends to 0. As a reasonable characteristic of the total amount of uncertainty (in the sense of randomness) contained in the probability distribution p on Ω the expected value of f over Ω, i.e., the value

$$H(p) = \sum_{i=1}^{n} f(\omega_i)p(\omega_i) \tag{2.1}$$

could be accepted (applying the convention that $0.\infty = 0$, if $p(\omega_i) = 0$). Considering two spaces Ω^i with probability distributions $p^i : \Omega^i \to [0,1], i = 1, 2$, setting

$$(p^1 \times p^2)(\omega^1, \omega^2) = p^1(\omega^1)p^2(\omega^2) \tag{2.2}$$

for each $\langle \omega^1, \omega^2 \rangle \in \Omega^1 \times \Omega^2$, and imposing on the function f the demand that

$$H(p^1 \times p^2) = H(p^1) + H(p^2) \tag{2.3}$$

should be valid, we arrive at the conclusion that up to a multiplicative constant, i.e., up to the base to which logarithm function will be defined, only the function $f(\omega) = \log(1/p(\omega))$ meets our demands. Taking by convention as the unit of the information quantity the information obtained when observing the result of regular $(1/2 : 1/2)$ coin tossing, we arrive at the Shannon entropy function H defined by

$$H(p) = -\sum_{i=1}^{n} (\log_2 p(\omega_i))p(\omega_i), \tag{2.4}$$

for the first time presented in 1948 ([12]). From the mathematical point of view, Shannon entropy develops and generalizes the Hartley entropy function (cf. [9] or the monograph [10]), as far as the technical aspects were concerned, inspiration came from thermodynamics and some results concerning rational message encoding.

Unfortunately, the function $\log_2(1/p(\omega))$ applied in (2.4) is too closely related to the structure and operations over the unit interval of real numbers to be translated into the lattice structures so that we have to try other decreasing function of $p(\omega)$, namely, the simple function $1 - p(\omega)$. Replacing $\log_2(1/p(\omega))$ by $1 - p(\omega)$ in (2.4), we arrive at the function H^Q defined by

$$H^Q(p) = \sum_{i=1}^{n} (1 - p(\omega_i))p(\omega_i) = 1 - \sum_{i=1}^{n} (p(\omega_i))^2, \tag{2.5}$$

referred in the surveyal paper [11]. Like as in the Shannon entropy, $H^Q(p) = 0$ (the minimum value of H^Q), iff $p(\omega) = 1$ for some (obviously just one) $\omega \in \Omega$ and $H^Q(p) = 1 - (1/n)$ (the maximum value of H^Q), iff $p(\omega_i) = 1/n$ for each $\omega_i \in \Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Contrary to Shannon entropy taking its values in $[0, \infty)$, $H^Q(p)$ takes its values in $[0, 1)$; if p^1, p^2 are probability distributions on spaces Ω_1, Ω_2, then for their statistically (stochastically) independent product $p^1 \times p^2$ defined by (2.2) the relation

$$H^Q(p^1 \times p^2) = 1 - [(1 - H^Q(p^1))(1 - H^Q(p^2))] \tag{2.6}$$

holds.

Denoting by P the probability measure induced by p on $\mathcal{P}(\Omega)$, i.e., setting $P(A) = \sum_{\omega \in A} p(\omega)$ for each $A \subset \Omega$, the expression $1 - p(\omega)$ can be written as $P(\Omega - \{\omega\})$, so that, for finite or countable space Ω, (2.5) yields that

$$H^Q(p) = \sum_{\omega \in \Omega} [P(\Omega - \{\omega\}) p(\omega)]. \tag{2.7}$$

Hence, given a complete lattice $\mathcal{T} = \langle T, \leq \rangle$, a $\mathcal{T}$-possibilistic distribution π on Ω, and replacing $\sum_{\omega \in \Omega}$ by $\bigvee_{\omega \in \Omega}$, product by infimum, and $P(\Omega - \{\omega\})$ by $\Pi(\Omega - \{\omega\})$ (Π induced by π on $\mathcal{P}(\Omega)$), we arrive at the expression

$$I(\pi) = \bigvee_{\omega \in \Omega} [\Pi(\Omega - \{\omega\}) \wedge \pi(\omega)]. \tag{2.8}$$

The value $I(\pi)$ belongs to T and can be taken as the expected value (in the sense of Sugeno integral, cf. [2] for more detail) of the nonincreasing (in the sense of the partial ordering $\leq$ on T) lattice-valued function $\Pi(\Omega - \{\omega\})$ with respect to the $\mathcal{T}$-possibilistic distribution π on Ω. Indeed, if $\pi(\omega_1) \leq \pi(\omega_2)$ holds for $\omega_1, \omega_2 \in \Omega$, then

$$\begin{aligned} \Pi(\Omega - \{\omega_1\}) &= \Pi(\Omega - \{\omega_1, \omega_2\}) \vee \pi(\omega_2) \geq \\ &\geq \Pi(\Omega - \{\omega_1, \omega_2\}) \vee \pi(\omega_1) = \Pi(\Omega - \{\omega_2\}) \end{aligned} \tag{2.9}$$

easily follows.

To conclude this section, we may quote that the mapping $I(\pi)$, defined by (2.8), meets the demands imposed above on lattice-valued entropy function over lattice-valued possibilistic distributions. Indeed, no ontologically independent assumptions are imposed and applied and the pattern of Shannon entropy is followed by the closest path compatible with the narrow framework of lattice theory apparatus. So, the mapping $I(\pi)$ deserves, at least in the author's opinion, to be considered as a promising candidate on the role of lattice-valued entropy function defined over $\mathcal{T}$-possibilistic distributions and, consequently, to be analyzed in more detail, from this point of view. Nevertheless, before doing so, let us briefly mention two alternative approaches to the definition of entropy function over possibilistic distributions, first of all comparing their methodological and meta-theoretical basic principles with those introduced and defended above.

The following fact should be clear from what has been already told, but let us recall it explicitly. The introduction of classical Shannon entropy $H(p)$(2.4) and the function $H^Q(p)$(2.5), both of them mapping probability distributions p into real numbers, does not violate our assumption to avoid from consideration notions not embeddable into the framework of lattice theory. Indeed, both the functions H and H^Q serve just as an intuitive and informal inspiration, when proposing the lattice-valued function $I(\pi)$(2.8), but neither H nor H^Q are used as inputs taking part in the mathematical definition of the function $I(\pi)$ and they are not applied when proving properties of $I(\pi)$. Moreover, neither H nor H^* are subjects of our investigations and both these functions could be eliminated from the text without menacing the formal correctness; just the intuition behind would be perhaps less obvious.

3 Two Alternative Approaches to Entropy Functions over Possibilistic Distributions

In [3], De Luca and Termini investigate standard real-valued fuzzy sets defined by a mapping $f : \Omega \rightarrow [0, 1]$ for a finite space $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$. They try to analyze, in which sense and degree Shannon entropy function could be used as a reasonable mathematical tool in order to quantify the total amount of uncertainty contained in the fuzzy set f, in spite of qualitative differences between the uncertainty in the sense of randomness (for which Shannon entropy was fitted) and the uncertainty in the sense of vagueness and fuzziness. As the first approximation, the authors propose the function

$$H(f) = -K \sum_{i=1}^{N} f(\omega_i) ln f(\omega_i) \tag{3.1}$$

where K is a positive constant and ln denotes logarithm to the base e. However, this function does not meet the intuitively reasonable demand according to which $H(f)$ should take its maximum value when $f(\omega_i) = 1/2$ for each $\omega_i \in \Omega$. In order to satisfy this demand, De Luca and Termini propose to define the entropy function by the mapping

$$d(f) = H(f) + H(1-f), \tag{3.2}$$

where $1-f$ is the fuzzy subset of Ω defined by $(1-f)(\omega_i) = 1 - f(\omega_i)$ for each $\omega_i \in \Omega$. The authors propose an alternative nonprobabilistic interpretation of this entropy (illustrated by an example) and prove some non-trivial mathematical results dealing with the mapping d, which may be of interest and use for specialists analyzing or applying standard real-valued fuzzy sets. However, from the point of view of our intentions, the notions introduced and results achieved in [3] are too closely related to the properties of real numbers, and to wide spectrum of operations by which these numbers can be processed, to be immediately applicable for non-numerical fuzzy sets without some additionally defined projection of these uncertainty degrees into the real line.

De Luca and Termini go on in their effort to apply Shannon entropy function to fuzzy sets in [4]. The mathematical model proposed and analyzed here is rather abstract to be described as briefly as in [3], but the intuition behind may be as follows. Given $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ as before, not only one fuzzy set f, but an M-tuple of fuzzy sets $f_1, f_2, \ldots, f_M$ on Ω is defined, each of them taking Ω into $[0,1]$. E.g., each f_i may be taken as fuzzy description of a property which is possessed, in the degree $f_j(\omega_i)$, by an object $\omega_i \in \Omega$ (the size of the object ω_i is large in degree $f_1(\omega_i)$, the weight of ω_i is high in degree $f_2(\omega_i)$, its color is blue in degree $f_3(\omega_i), \ldots$). The sequence $\langle f_1(\omega_i), f_2(\omega_i), \ldots, f_M(\omega_i)\rangle$ of numbers from $[0,1]$ (or its subsequence, if the classification of certain objects with respect to certain properties is not known) is taken as vector-valued membership degree $f(\omega_i)$ ascribed to $\omega_i \in \Omega$, hence it is a mapping taking Ω into the set of all m-tuples, $m \leq M$, of real numbers from $[0,1]$. The standard linear ordering $\leq$ on $[0,1]$, according to which for each $j \leq M$ and each $i_1, i_2 \leq N$ such that $f_j(\omega_{i_1})$ and $f_j(\omega_{i_2})$ are defined either $f_j(\omega_{i_1}) \leq f_j(\omega_{i_2})$ or the inverse inequality holds, obviously does not induce a linear ordering on the space of vectors $\langle f_1(\omega, f_2(\omega), \ldots, f_M(\omega)\rangle$ and their sub-vectors, but under some intuitive and acceptable conditions investigated in [4] the space of values possibly taken by f defines a complete lattice with respect to the partial ordering $\leq^*$ such that

$$f(\omega_1) = \langle f_1(\omega_1), f_2(\omega_1), \ldots, f_M(\omega_1) \leq^* \langle f_1(\omega_2), f_2(\omega_2), \ldots, f_M(\omega_2)\rangle = f(\omega_2) \tag{3.3}$$

holds iff $f_j(\omega_1) \leq f_j(\omega_2)$ is valid for each $\omega_1, \omega_2 \in \Omega$ for which both the values $f_j(\omega_1)$ and $f_j(\omega_2)$ are defined. The vector-valued entropy d of the lattice-valued mapping f is then defined by

$$d(f) = \langle d_1(f_1), d_2(f_2), \ldots, d_M(f_M)\rangle \tag{3.4}$$

where each d_j defines a real-valued entropy from a class of Shannon-like real-valued entropy functions, possibly different for different j's, but each of then defined like (3.2). Hence, the entropy function d takes its values in the set of finite sequences of real numbers partially ordered by $\leq^*$; under certain conditions analyzed in [4] also this partially ordered structure defines a complete lattice.

The approach presented in [4] offers a number of interesting ideas and valuable results which may be of use when applying the notion of fuzziness when solving some problems from the surrounding us real world. However, the complete lattices in which membership degrees $f_j(\omega_i)$ and entropy function $d(f)$ take their values are particular and sophistically proposed complete lattices over vectors and matrices of real numbers, not a free input into our reasonings and constructions ("Let $\mathcal{T} = \langle T, \leq\rangle$ be *a* complete lattice...") as demanded by our methodological assumptions imposed on lattice-valued fuzzy sets and entropy functions in Section 2.

A qualitatively different approach to the quantification of the total amount of fuzziness contained in a fuzzy set is presented by R. R. Yager in [14] (for real-valued fuzzy sets) and in [15] (for the lattice-valued ones). His basic idea reads that a fuzzy set $f : \Omega \to [0,1]$ is the more "fuzzy" the smaller is the difference between f and its fuzzy complement $1-f$ defined by $(1-f)(\omega)=1-f(\omega)$ for each $\omega \in \Omega$. Consequently, the most "fuzzy" is the fuzzy set $f_{1/2}(\omega) = 1/2$ for each $\omega \in \Omega$, which is identical with its complement, and fuzzy set f_2 is "at least as fuzzy" as fuzzy set f_1 ($f_1 \preceq f_2$, in symbols), if

$$f_1(\omega) \wedge (1-f_1)(\omega) \leq f_2(\omega) \wedge (1-f_2)(\omega) \tag{3.5}$$

($\wedge$ stands for the standard minimum in $[0,1]$) holds for each $\omega \in \Omega$. Indeed, $f \preceq f_{1/2}$ holds for each $\omega \in \Omega$.

As $\preceq$ obviously does not define a linear ordering on the space of real-valued fuzzy subsets of Ω, various metrics and distance functions are introduced and examined in more detail in order to obtain a linear ordering of fuzzy subsets over Ω (e.g., expected value of the difference $|f(\omega) - f_{1/2}(\omega)|$ with respect to a measure on $\mathcal{P}(\Omega)$ or some more sophisticated functions).

Given a distance function δ taking pairs of real-valued fuzzy sets into $[0,1]$, the value $1-\delta(f, f_{1/2})$ could be taken as a real-valued quantification of the total amount of fuzziness contained in f. However, such a mapping is not applicable to lattice-valued fuzzy sets and, moreover, some ontologically independent inputs are necessary (the mapping δ, the complement function $1-.,\dots$) and may be chosen in various ways. Hence, our methodological conditions imposed on lattice-valued entropy functions are, again, violated.

In the case of real-valued fuzzy subsets of Ω the relation $f_1 \preceq f_2$, defined by (3.5), can be expressed also in the form that f_2 and $1-f_2$ "lie between" f_1 and $1-f_1$ in the sense that, for each $\omega \in \Omega$, both the values $f_2(\omega) \wedge (1-f_2)(\omega)$ and $1-(f_2(\omega) \wedge (1-f_2)(\omega))$ lie between the values $f_1(\omega) \wedge (1-f_1)(\omega)$ and $1-(f_1(\omega) \wedge (1-f_1)(\omega))$, indeed, $f_{1/2} = 1 - f_{1/2} \equiv 1/2$ lies between $f_1(\omega) \wedge (1-f_1(\omega))$ and its complement for each f_1.

In [1], G. Birkhoff defines the notion of betweenness (the ternary relation "to lie between") for lattice $\langle T, \leq \rangle$ in this way: given $a, b, c \in T$, b lies between a and c, if the relation

$$(a \wedge b) \vee (b \wedge c) = b = (a \vee b) \wedge (b \vee c) \tag{3.6}$$

holds. For distributive lattices (3.6) reduces to

$$(a \vee c) \wedge b = b = (a \wedge c) \vee b. \tag{3.7}$$

In order to be able to apply this relation to lattice-valued fuzzy sets following the pattern just outlined for real-valued fuzzy sets, a lattice-valued modification of the complement function $1-f(.)$ is necessary. R. R. Yager in [15] enriches the lattice structure in which fuzzy sets take their values by a new unary operation of negation which is defined as ontologically independent input into the lattice structure; for $x \in T$ the negation of x is denoted by x^C. Axiomatically imposed demands on the operation of negation are that of involution and order reversing, under some simplifying conditions these demands reduce to De Morgan rules. Moreover, for crisp fuzzy sets (i.e., for fuzzy sets taking only $\emptyset_{\mathcal{T}}$ or $\mathbf{1}_{\mathcal{T}}$ as membership degrees) negation reduces to standard set complement operation.

Using the negation operation, the relation $f_1 \preceq f_2$ (f_2 is at least as fuzzy as f_1) can be extended also to lattice-valued fuzzy sets in order to define the case when the values $f_2(\omega) \wedge (f_2(\omega))^C$ and $[f_2(\omega) \wedge (f_2(\omega))^C]^C$ lie between $f_1(\omega) \wedge (f_1(\omega))^C$ and $[f_1(\omega) \wedge (f_1(\omega))^C]^C$ for each $\omega \in \Omega$. As in other cases briefly reviewed in this section a number of interesting and valuable results concerning the relation $\preceq$ between lattice-valued fuzzy sets are presented and proved, but their methodological basis does not meet our demands introduced in Section 2 at least in two points: an ontologically independent input in the form of operation of negation is necessary and the total amount of fuzziness contained in a fuzzy set is not quantified by a single value from the lattice under consideration, but fuzzy sets are compared directly, hence, there is not single value which would characterize the fuzzy set in question in the sense of its entropy value.

4 Entropy Function over Complete Chained Lattices

Definition 4.1 Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let π be a $\mathcal{T}$-possibilistic distribution on a nonempty space Ω. *$\mathcal{T}$-(valued) entropy (function)* $I(\pi)$ is defined by

$$I(\pi) = \bigvee_{\omega \in \Omega} (\Pi(\Omega - \{\omega\}) \wedge \pi(\omega)). \tag{4.1}$$

As may be easily observed, $\mathcal{T}$,-entropy $I(\pi)$ takes its minimum value $\oslash_{\mathcal{T}}$, if $\pi(\omega_0) = \mathbf{1}_{\mathcal{T}}$ for some $\omega_0 \in \Omega$ and $\pi(\omega) = \oslash_{\mathcal{T}}$ for each $\omega \in \Omega, \omega \neq \omega_0$. $I(\pi)$ takes its maximum value $\mathbf{1}_{\mathcal{T}}$, if π is not isolated (e.g., if there are $\omega_1 \neq \omega_2$ such that $\pi(\omega_1) = \pi(\omega_2) = \mathbf{1}_{\mathcal{T}}$, cf. Lemma 4.1. Let π_1, π_2 be $\mathcal{T}$-possibilistic distributions on Ω such that $\pi_1(\omega) \leq \pi_2(\omega)$ holds for each $\omega \in \Omega$, then $\Pi_1(A) \leq \Pi_2(A)$ holds for each $A \subset \Omega$ and, consequently, the relation $I(\pi_1) \leq I(\pi_2)$ is valid. Let us note that if p_1 and p_2 are probability distributions on finite or countable space Ω, the relation $p_1(\omega) \leq p_2(\omega)$ for each $\omega \in \Omega$ may be the case only when the probability distributions p_1 and p_2 are identical.

Lemma 4.1 Let $\mathcal{T} = \langle T, \leq \rangle$ be complete chained lattice (i.e., the partial ordering $\leq$ is linear), let Ω, π and $I(\pi)$ be as in Definition 4.1. Then $I(\pi) < \mathbf{1}_{\mathcal{T}}$ holds if and only if π is isolated in some $\omega_0 \in \Omega$ and if this is the case, then $I(\pi) = \Pi(\Omega - \{\omega_0\})$.

Proof: For each $\omega_0 \in \Omega$, (4.1) yields that

$$I(\pi) = [\Pi(\Omega - \{\omega_0\}) \wedge \pi(\omega_0)] \vee \bigvee_{\omega \in \Omega, \omega \neq \omega_0} [\Pi(\Omega - \{\omega\}) \wedge \pi(\omega)]. \tag{4.2}$$

As $\leq$ is linear on T and π is isolated in $\omega_0, \pi(\omega_0) = \mathbf{1}_{\mathcal{T}}$ follows. Indeed, if $\pi(\omega_0) \leq \Pi(\Omega - \{\omega_0\})$ were the case, then $\bigvee_{\omega \in \Omega} \pi(\omega) < \mathbf{1}_{\mathcal{T}}$ would follow, but this contradicts the definition of π. Hence, $\pi(\omega_0) > \Pi(\Omega - \{\omega_0\})$ and $\Pi(\Omega) = \mathbf{1}_{\mathcal{T}} = \Pi(\Omega - \{\omega_0\}) \vee \pi(\omega_0) = \pi(\omega_0)$ results. Hence, as $\Omega - \{\omega\}$ contains ω_0 for each $\omega \neq \omega_0$, the relation $\Pi(\Omega - \{\omega\}) = \mathbf{1}_{\mathcal{T}}$ for each $\omega \neq \omega_0$ follows. Consequently, (4.2) yields that

$$\begin{aligned} I(\pi) &= (\Pi(\Omega - \{\omega_0\}) \wedge \mathbf{1}_{\mathcal{T}}) \vee \bigvee_{\omega \in \Omega, \omega \neq \omega_0} (\mathbf{1}_{\mathcal{T}} \wedge \pi(\omega)) = \\ &= \Pi\left(\Omega - \{\omega_0\}\right) \vee \bigvee_{\omega \in \Omega, \omega \neq \omega_0} \pi(\omega) = \Pi(\Omega - \{\omega_0\}) < \mathbf{1}_{\mathcal{T}}. \end{aligned} \tag{4.3}$$

If π is not isolated, then $\Pi(\Omega - \{\omega\}) = \mathbf{1}_{\mathcal{T}}$ holds for each $\omega \in \Omega$, so that

$$I(\pi) = \bigvee_{\omega \in \Omega} (\Pi(\Omega - \{\omega\}) \wedge \pi(\omega)) = \bigvee_{\omega \in \Omega} \pi(\omega) = \mathbf{1}_{\mathcal{T}}. \tag{4.4}$$

The assertion is proved. □

For possibilistically independent (non-interactive, in other terms) product of two possibilistic distributions over the Cartesian product of their support spaces the generalized version of the assertion of Lemma 4.1 reads as follows.

Lemma 4.2 Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete chained lattice. For both $i = 1, 2$, let Ω_i be a nonempty space, let π_i be an isolated $\mathcal{T}$-possibilistic distribution on Ω_i, and let $I(\pi_i)$ be defined by (4.1). Let $\pi_{12} : \Omega_1 \times \Omega_2 \to T$ be defined by $\pi_{12}(\omega_1, \omega_2) = \pi_1(\omega_1) \wedge \pi_2(\omega_2)$ for each $\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2$, let $I(\pi_{12})$ be defined by (4.1) with Ω replaced by $\Omega_1 \times \Omega_2$, hence,

$$I(\pi_{12}) = \bigvee_{\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2} (\Pi_{12}((\Omega_1 \times \Omega_2) - \{\langle \omega_1, \omega_2 \rangle\}) \wedge \pi_{12}(\omega_1, \omega_2)). \tag{4.5}$$

Then π_{12} defines an isolated $\mathcal{T}$-possibilistic distribution on $\Omega_1 \times \Omega_2$ and the relation

$$I(\pi_{12}) = I(\pi_1) \vee I(\pi_2) < \mathbf{1}_{\mathcal{T}} \tag{4.6}$$

holds.

Proof: As π_1 and π_2 are isolated and $\leq$ is linear, there exist uniquely defined elements $\omega_1^0 \in \Omega_1$ and $\omega_2^0 \in \Omega_2$ such that $\pi_1(\omega_1^0) = \pi_2(\omega_2^0) = \mathbf{1}_{\mathcal{T}}$, so that $\pi_{12}(\omega_1^0, \omega_2^0) = \mathbf{1}_{\mathcal{T}} \wedge \mathbf{1}_{\mathcal{T}} = \mathbf{1}_{\mathcal{T}}$. Hence,

$$\bigvee_{\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2} \pi_{12}(\omega_1, \omega_2) = \mathbf{1}_{\mathcal{T}} \tag{4.7}$$

and π_{12} defines a $\mathcal{T}$-possibilistic distribution on $\Omega_1 \times \Omega_2$. Moreover,

$$\begin{aligned}
& \Pi((\Omega_1 \times \Omega_2) - \{\langle \omega_1^0, \omega_2^0 \rangle\}) = \\
= & \bigvee_{\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2, \langle \omega_1, \omega_2 \rangle \neq \langle \omega_1^0, \omega_2^0 \rangle} (\pi_1(\omega_1) \wedge \pi_2(\omega_2)) \\
= & \bigvee_{\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2, \omega_1 \neq \omega_1^0, \omega_2 \neq \omega_2^0} (\pi_1(\omega_1) \wedge \pi_2(\omega_2)) \vee \\
\vee & \bigvee_{\omega_2 \in \Omega_0, \omega_2 \neq \omega_2^0} (\pi_1(\omega_1^0) \wedge \pi_2(\omega_2)) \vee \\
\vee & \bigvee_{\omega_1 \in \Omega_1, \omega_1 \neq \omega_1^0} (\pi_1(\omega_1) \wedge \pi_2(\omega_2^0)).
\end{aligned} \tag{4.8}$$

For the third line in (4.8) we obtain that

$$\begin{aligned}
& \bigvee_{\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2, \omega_1 \neq \omega_1^0, \omega_2 \neq \omega_2^0} (\pi_1(\omega_1) \wedge \pi_2(\omega_2)) \leq \\
\leq & \left(\bigvee_{\omega_1 \in \Omega_1, \omega_1 \neq \omega_1^0} \pi_1(\omega_1) \right) \wedge \left(\bigvee_{\omega_2 \in \Omega_2, \omega_2 \neq \omega_2^0} \pi_2(\omega_2) \right) = \\
= & (\Pi_1(\Omega_1 - \{\omega_1^0\})) \wedge (\pi_2(\Omega_2 - \{\omega_2^0\}))
\end{aligned} \tag{4.9}$$

holds, for the fourth line in (4.8) we obtain that

$$\begin{aligned}
& \bigvee_{\omega_2 \in \Omega_2, \omega_2 \neq \omega_2^0} (\pi_1(\omega_1^0) \wedge \pi_2(\omega_2)) = \\
= & \bigvee_{\omega_2 \in \Omega_2, \omega_2 \neq \omega_2^0} (\mathbf{1}_{\mathcal{T}} \wedge \pi_2(\omega_2)) = \Pi_2(\Omega_2 - \{\omega_2^0\}),
\end{aligned} \tag{4.10}$$

and analogously for the fifth line in (4.8)

$$\bigvee_{\omega_1 \in \Omega_1, \omega_1 \neq \omega_1^0} (\pi_1(\omega_1) \wedge \pi_2(\omega_2^0)) = \Pi_1(\Omega_1 - \{\omega_1^0\}). \tag{4.11}$$

Consequently, the relation

$$\begin{aligned}
& I(\pi_{12}) = \Pi_{12}((\Omega_1 \times \Omega_2) - \{\langle \omega_1^0, \omega_2^0 \rangle\}) = \\
= & ((\Pi_1(\Omega_1 - \{\omega_1^0\})) \wedge (\Pi_2(\Omega_2 - \{\omega_2^0\}))) \vee \Pi_1(\Omega_1 - \{\omega_1^0\}) \vee \Pi_2(\Omega_2 - \{\omega_2^0\}) = \\
= & \Pi_1(\Omega_1 - \{\omega_1^0\}) \vee \Pi_2(\Omega_2 - \{\omega_2^0\}) < \mathbf{1}_{\mathcal{T}}
\end{aligned} \tag{4.12}$$

follows. Indeed, $\Pi_i(\Omega_i - \{\omega_i^0\}) < \mathbf{1}_{\mathcal{T}}$ holds for both $i = 1, 2$ and, as $\leq$ is a linear ordering on T, $t_1 \vee t_2 = t_1$ or $t_1 \vee t_2 = t_2$ holds for each $t_1, t_2 \in T$. The assertion is proved. □

It is perhaps worth being noted explicitly that the condition according to which *both* the particular $\mathcal{T}$-possibilistic distributions π, on Ω_1 and π_2 on Ω_2 must be isolated in order to obtain an isolated $\mathcal{T}$-possibilistic distribution π_{12} on $\Omega_1 \times \Omega_2$ is substantial. Indeed, let only π_2 be isolated, so that $\Pi_1(\Omega_1 - \{\omega_1\}) = \mathbf{1}_{\mathcal{T}}$ holds for each $\omega_1 \in \Omega_1$ and $\Pi_2(\Omega_2 - \{\omega_2^0\}) < \mathbf{1}_{\mathcal{T}}$ holds for uniquely defined $\omega_2^0 \in \Omega_2$.

Given $t \in T, t < \mathbf{1}_{\mathcal{T}}$, there exists $\omega_A \in \Omega_1$ such that $t < \pi_1(\omega_1)$ holds (as $\leq$ is linear, this follows immediately from the condition that $\Pi_1(\Omega_1) = \bigvee_{\omega_1 \in \Omega_1} \pi_1(\omega_1) = \mathbf{1}_{\mathcal{T}}$). If there were only one $\omega_A \in \Omega_1$ such that $t < \pi_1(\omega_A)$ were the case, we would obtain that $\Pi_1(\Omega_1 - \{\omega_A\}) \leq t < \mathbf{1}_{\mathcal{T}}$ holds, but this relation contradicts the assumption that π_1 is not isolated. Consequently, for each $t < \mathbf{1}_{\mathcal{T}}$ there exists *different* elements $\omega_A, \omega_B \in \Omega_1$ such that $\pi_1(\omega_A) > t$ and $\pi_1(\omega_B) > t$ holds.

Take $\omega_2^0 \in \Omega_2$, so that $\pi_2(\omega_2^0) = \mathbf{1}_{\mathcal{T}}$ and

$$\pi_{12}(\omega_1, \omega_2^0) = \pi_1(\omega_1) \wedge \pi_2(\omega_2^0) = \pi_1(\omega_1) \wedge \mathbf{1}_{\mathcal{T}} = \pi_1(\omega_1) > t \tag{4.13}$$

holds for both $\omega_1 = \omega_A$ and $\omega_1 = \omega_B$. As $\langle \omega_A, \omega_2^0 \rangle$ and $\langle \omega_B, \omega_2^0 \rangle$ are *different* elements of the Cartesian product $\Omega_1 \times \Omega_2$, for each $\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2$ at least one of these pairs is in $(\Omega_1 \times \Omega_2) - \{\langle \omega_1, \omega_2 \rangle\}$, so that

$$\Pi_{12}((\Omega_1 \times \Omega_2) - \{\langle \omega_1, \omega_2 \rangle\}) \geq \pi_{12}(\omega_A, \omega_2^0) \wedge \pi_{12}(\Omega_B, \omega_2^0) > t \tag{4.14}$$

follows. As $t < \mathbf{1}_{\mathcal{T}}$ may be chosen arbitrarily, we obtain that

$$\bigvee_{\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2} [\Pi_{12}((\Omega_1 \times \Omega_2) - \{\langle \omega_1, \omega_2 \rangle\}) \wedge \pi_{12}(\omega_1, \omega_2)] = \mathbf{1}_{\mathcal{T}} = I(\pi_{12}) \tag{4.15}$$

holds. So, π_{12} is not an isolated $\mathcal{T}$-possibilistic distribution on $\Omega_1 \times \Omega_2$ and, as $I(\pi_1) = \mathbf{1}_{\mathcal{T}}$ according to (4.4), the relation

$$I(\pi_{12}) = I(\pi_1) \vee I(\pi_2) = \mathbf{1}_{\mathcal{T}} \vee I(\pi_2) = \mathbf{1}_{\mathcal{T}} \tag{4.16}$$

holds trivially.

Lemma 4.3 Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete chained lattice, let Ω be a nonempty space, let π be a $\mathcal{T}$-possibilistic distribution on Ω. Let $\Omega^* = \{\Omega_i : i \in \mathcal{J}\}$ be a decomposition of Ω, i.e., $\Omega_i \cap \Omega_j = \emptyset$ for each $i, j \in \mathcal{J}, i \neq j$, and $\bigcup_{i \in \mathcal{J}} \Omega_i = \Omega$, let $\pi^*(i) = \Pi(\Omega_i)$ for each $i \in \mathcal{J}$. Then the mapping $\pi^* : \mathcal{J} \to T$ defines a $\mathcal{T}$-possibilistic distribution on $\mathcal{J}$ such that, for $I(\pi)$ and $I(\pi^*)$ defined by (4.1), the inequality $I(\pi^*) \leq I(\pi)$ holds.

Remark 1 The $\mathcal{T}$-possibilistic distribution π^* on $\mathcal{J}$ can be obviously taken as $\mathcal{T}$-possibilistic distribution on Ω^*, simply identifying each $\Omega_i \in \Omega^*$ with its index $i \in \mathcal{J}$.

Proof: If π is not isolated, then $I(\pi) = \mathbf{1}_{\mathcal{T}}$ and the assertion holds trivially. If π is isolated, then there exists uniquely defined $\omega_0 \in \Omega$ such that $\Pi(\Omega - \{\omega_0\}) < \mathbf{1}_{\mathcal{T}}$ and, consequently, $\pi(\omega_0) = \mathbf{1}_{\mathcal{T}}$ holds. Let $i_0 \in \mathcal{J}$ be the uniquely defined $i \in \mathcal{J}$ such that $\omega_0 \in \Omega_i$ is the case. Then we obtain

$$\begin{aligned} I(\pi^*) &= \bigvee_{i \in \mathcal{J}} (\Pi^*(\mathcal{J} - \{i\}) \wedge \pi^*(i)) = \bigvee_{\Omega_i \in \Omega^*} (\Pi(\Omega - \Omega_i) \wedge \Pi(\Omega_i)) = \\ &= [\Pi(\Omega - \Omega_{i_0}) \wedge \Pi(\Omega_{i_0})] \vee \bigvee_{\Omega_i \in \Omega^*, \Omega_i \neq \Omega_{i_0}} (\Pi(\Omega - \Omega_i) \wedge \Pi(\Omega_i)). \end{aligned} \tag{4.17}$$

As $\omega_0 \in \Omega_{i_0}$ and, consequently, $\omega_0 \in \Omega - \Omega_i$ for each $i \in \mathcal{J}, i \neq i_0$, holds, we obtain that $\Pi(\Omega_{i_0}) = \Pi(\Omega - \Omega_i) = \mathbf{1}_{\mathcal{T}}$ for each $i \in \mathcal{J}, i \neq i_0$, follows. Hence,

$$\begin{aligned} I(\pi^*) &= \Pi(\Omega-\Omega_{i_0}) \vee \bigvee_{\Omega_i\in\Omega^*,\Omega_i\neq\Omega_{i_0}} \Pi(\Omega_i) = \\ &= \Pi(\Omega-\Omega_{i_0}) \vee \Pi\left(\bigcup_{\Omega_i\neq\Omega_{i_0}} \Omega_i\right) = \Pi(\Omega-\Omega_{i_0}) \vee \Pi(\Omega-\Omega_{i_0}) = \Pi(\Omega-\Omega_{i_0}) \\ &\le \Pi(\Omega-\{\omega_0\}) = I(\pi) \end{aligned} \tag{4.18}$$

holds, as $\omega_0 \in \Omega_{i_0}$ implies that $\Omega-\Omega_{i_0} \subset \Omega-\{\omega_0\}$ and $\Pi(\Omega-\Omega_{i_0}) \le \Pi(\Omega-\{\omega_0\})$ holds. The assertion is proved. □

An immediate corollary of Lemma 4.3 reads that if the $\mathcal{T}$-possibilistic distribution π on Ω is isolated, i.e., if $I(\pi) < \mathbf{1}_\mathcal{T}$ holds, then the $\mathcal{T}$-possibilistic distribution π^* induced by π on (the parametric set $\mathcal{J}$ of indices of sets from) a decomposition Ω^* of Ω is also isolated, i.e., $I(\pi^*) < \mathbf{1}_\mathcal{T}$ follows. On the other side, π^* may be isolated even when π is not isolated, e.g., if there are $\omega_1,\omega_2 \in \Omega, \omega_1 \neq \omega_2$, such that $\pi(\omega_1)=\pi(\omega_2)=\mathbf{1}_\mathcal{T}$. Indeed, if the decomposition Ω^* of Ω is such that there exists $\Omega_{i_0} \subset \Omega$, $\Omega_{i_0} \in \Omega^*$, with the property that there exists $t \in T, t < \mathbf{1}_\mathcal{T}$, such that $\pi(\omega) \le t$ holds for each $\omega \in \Omega - \Omega_0$ (hence, all ω's with $\pi(\omega_= \mathbf{1}_\mathcal{T}$ are contained in Ω_{i_0}), then $\pi^*(i_0)=\Pi(\Omega_{i_0})=\mathbf{1}_\mathcal{T}, \pi^*(i) \le \Pi(\Omega-\Omega_{i_0}) \le t$ for each $i \in \mathcal{J}, i \neq i_0$, holds, hence, $\Pi^*(\mathcal{J}-\{i_0\}) = \Pi(\Omega-\Omega_{i_0}) \le t < \mathbf{1}_\mathcal{T}$ follows, so that π^* is an isolated $\mathcal{T}$-possibilistic distribution on $\mathcal{J}$ (on Ω^*, resp.).

Lemma 4.4 Let $\mathcal{T}=\langle T,\le\rangle$ be a complete chained lattice, let Ω_1,Ω_2 be nonempty spaces, let $\pi_{12}: \Omega_1\times\Omega_2 \to T$ be a $\mathcal{T}$-possiblistic distribution on $\Omega_1\times\Omega_2$, i.e., $\bigvee_{\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2} \pi_{12}(\omega_1,\omega_2)=\mathbf{1}_\mathcal{T}$ holds. For both $i=1,2$, let π_i^* denote the marginal $\mathcal{T}$-possiblistic distribution on Ω_i induced by π_{12}, so that

$$\pi_1^*(\omega_1) = \bigvee_{\omega_2\in\Omega_2} \pi_{12}(\omega_1,\omega_2), \pi_2^*(\omega_2) = \bigvee_{\omega_1\in\Omega_1} \pi_{12}(\omega_1,\omega_2) \tag{4.19}$$

holds for each $\omega_1\in\Omega_1, \omega_2\in\Omega_2$. Let $\pi_1^*\times\pi_2^*$ be the possibilistically independent product of marginal $\mathcal{T}$-possibilistic distributions π_1^*,π_2^*, so that, for each $\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2$, the relation

$$(\pi_1^*\times\pi_2^*)(\omega_1,\omega_2) = \pi_1^*(\omega_1)\wedge\pi_2^*(\omega_2) \tag{4.20}$$

holds, let the $\mathcal{T}$-entropy $I(\pi)$ be defined by (4.1) for $\pi=\pi_{12},\pi_1^*,\pi_2^*$ and $\pi_1^*\times\pi_2^*$. Then the inequality

$$I(\pi_{12}) \le I(\pi_1^*\times\pi_2^*) = I(\pi_1^*)\vee I(\pi_2^*) \tag{4.21}$$

is valid.

Proof: Obviously,

$$\bigvee_{\omega_1\in\Omega_1} \pi_1^*(\omega_1) = \bigvee_{\omega_1\in\Omega_1}\left(\bigvee_{\omega_2\in\Omega_2} \pi_{12}(\omega_1,\omega_2)\right) = \bigvee_{\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2} \pi_{12}(\omega_1,\omega_2) = \mathbf{1}_\mathcal{T} \tag{4.22}$$

and similarly for $\pi_2^*(\omega_2)$, so that, for both $i=1,2, \pi_i^*$ defines a $\mathcal{T}$-possibilistic distribution on Ω_i. For each $\omega_1\in\Omega_1$, the inequality

$$\pi_{12}(\omega_1,\omega_2) \le \bigvee_{\omega_2\in\Omega_2} \pi_{12}(\omega_1,\omega_2) = \pi_1^*(\omega_1) \tag{4.23}$$

holds, analogously we obtain the inequality $\pi_{12}(\omega_1,\omega_2) \le \pi_2^*(\omega_2)$. Hence, the inequality

$$\pi_{12}(\omega_1,\omega_2) \le \pi_1^*(\omega_1)\wedge\pi_2^*(\omega_2) = (\pi_1^*\times\pi_2^*)(\omega_1,\omega_2) \tag{4.24}$$

is valid for each $\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2$. As proved at the beginning of this section, in this case the relation

$$I(\pi_{12}) \le I(\pi_1^*\times\pi_2^*) = I(\pi_1^*)\vee I(\pi_2^*) \text{ (due to Lemma 4.2)} \tag{4.25}$$

follows and the assertion is proved. □

5 Some Informal Reasonings Concerning the $\mathcal{T}$-Valued Entropy Function $I(\pi)$

Before going on with our investigations of the properties of the lattice-valued entropy function $I(\pi)$ defined by (4.1), let us briefly re-consider our meta-theoretical demands imposed on this function in Section 2 above. As the demand to avoid ontologically independent inputs from our model is concerned, we may take this assumption as having been met. Neither in the Definition 4.1 nor in the statements of Lemmata 4.1 to 4.4 and their proofs, any assumptions and tools going beyond the framework of lattice theory have been applied and the only specification imposed on the notion of general complete lattice $\mathcal{T} = \langle T, \leq \rangle$, namely the assumption that the ordering $\leq$ on T is linear (so that $\mathcal{T}$ is chained) may be and has been defined using just the apparatus of lattice theory.

As far as the promise to follows the pattern of Shannon probabilistic entropy function as closely as possible, the assertions of Lemmata 4.2–4.4 can be seen as possibilistic modifications of well-known and important properties possessed by the Shannon entropy function H (up to the syntactical difference when replacing $\sum$ by $\bigvee$, as could be more or less expected). Indeed, let Ω_1, Ω_2 be finite spaces, let p^1, p^2 be probabilistic distributions on Ω_1 and Ω_2, and let p_{12} be their statistically (stochastically) independent product on $\Omega_1 \times \Omega_2$, so that $p_{12}(\omega_1, \omega_2) = p_1(\omega_1)p_2(\omega_2)$, then $H(p_{12}) = H(p_1) + H(p_2)$ (compare with Lemma 4.2). Given a general probability distribution on $\Omega_1 \times \Omega_2$, i.e., $p_{12} : \Omega_1 \times \Omega_2 \rightarrow [0, 1]$, $\sum_{\langle \omega_1, \omega_2 \rangle \in \Omega_1 \times \Omega_2} p_{12}(\omega_1, \omega_2) = 1$, and denoting by $p_1^*(p_2^*$, resp.) the marginal distribution induced by p_{12} on $\Omega_1(\Omega_2$, resp.), the inequality $H(p_{12}) \leq H(p_1^*) + H(p_2^*)$ holds (compare with Lemma 4.4). Moreover, given a probability distribution p on Ω, given a decomposition $\Omega^* \subset \mathcal{P}(\Omega)$ of Ω and setting $p^*(\Omega_0) = \sum_{\omega \in \Omega_0} p(\omega)$ for any $\Omega_0 \in \Omega^*$, the inequality $H(p^*) \leq H(p)$ holds, as in Lemma 4.3 for the lattice-valued possibilistic distributions π and π^*.

On the other side, Lemma 4.1 proves the properties in which lattice-valued entropy $I(\pi)$ defined by (4.1) qualitatively differs from the Shannon probabilistic entropy. Before all, the entropy $I(\pi)$ is normalized, hence, its values are majorized by the unit element $\mathbf{1}_\mathcal{T}$ of the complete chained lattice under consideration. This fact enables to compare, at least partially, two $\mathcal{T}$-possibilistic distributions as far as the proximity of these entropy values to the maximum value is concerned, however, this maximum value is not too sensitive to separate among wide variety of cases when the value $I(\pi) = \mathbf{1}_\mathcal{T}$ may occur. E.g., as shown in Lemma 4.2, if $\pi(\omega_1) = \pi(\omega_2) = \mathbf{1}_\mathcal{T}$ holds for different $\omega_1, \omega_2 \in \Omega$, then $I(\pi) = \mathbf{1}_\mathcal{T}$ no matter whether ω_1, ω_2 are the only elements in Ω with this property or whether, e.g., $\pi(\omega) = \mathbf{1}_\mathcal{T}$ for each $\omega \in \Omega$. However, from a certain point of view this consequence is quite intuitive. Indeed, let the case $\pi(\omega) = \mathbf{1}_\mathcal{T}$ describe the situation when we have no argument weakening our apriori expectation that ω is identical with the actual value ω_0 of a hidden parameter under consideration (the actual elementary random event, "state of world", ω_0 in the term of random variables). If there are *two* different elements $\omega_1, \omega_2 \in \Omega$ such that $\pi(\omega_1) = \pi(\omega_2) = \mathbf{1}_\mathcal{T}$, there is no reason based on the $\mathcal{T}$-possibilistic distribution π in question for which just one element, say ω_1, should be picked up as the most favorable candidate to (or estimation of) the actual value $\omega_0 \in \Omega$. Picking up, in spite of this fact, just one element (ω_1 or ω_2) as the value ω_0, the degree of uncertainty admitted, when taking this decision, can be taken as the maximum one, i.e., in the terms of $\mathcal{T}$-valued entropy function $I(\pi)$, the case $I(\pi) = \mathbf{1}_\mathcal{T}$ occurs.

Nevertheless, even in the case of non-isolated $\mathcal{T}$-possibilistic distribution π on Ω a reasonable and useful information for decision making can be obtained when replacing Ω by a factor-space $\Omega^* = \Omega/\approx$ and π by π^* on Ω^* as in Lemma 4.3. E.g., supposing that elements of Ω are grouped into classes of Ω^* according to their "colours" in such a way that each ω's with $\pi(\omega) = \mathbf{1}_\mathcal{T}$ are "blue" and $\pi(\omega) \leq t_0 < \mathbf{1}_\mathcal{T}$ holds for each ω's which are "not blue", we obtain an isolated $\mathcal{T}$-possibilistic distribution π^* on Ω^*. Hence, we can make the conclusion as follows: there are no arguments based on π to pick up just one $\omega \in \Omega$ as the best candidate for ω_0, but there are some arguments based on π^* (hence, on π) according to which the most favorable candidate on ω_0 is a "blue" element of Ω.

Another way of reasoning would bring us to the following refinement of the definition of the $\mathcal{T}$-valued entropy function $I(\pi)$. Namely, let π be a $\mathcal{T}$-possibilistic distribution on Ω, let $\Omega_0 = \{\omega \in \Omega : \pi(\omega) = \mathbf{1}_\mathcal{T}\}$. For each $\omega_0 \in \Omega_0$ define the modification π^{ω_0} of π in this way: $\pi^{\omega_0}(\omega) = \pi(\omega)$, if $\omega = \omega_0$ or if $\pi(\omega) < \mathbf{1}_\mathcal{T}$ holds, $\pi^{\omega_0}(\omega) = \oslash_\mathcal{T}$ otherwise, i.e., if $\pi(\omega) = \mathbf{1}_\mathcal{T}$, but $\omega \neq \omega_0$. Set

$$\hat{I}(\pi) = I(\pi^{\omega_0}). \tag{5.1}$$

As a matter of fact, the value $I(\pi^{\omega_0})$ is the same for each $\omega_0 \in \Omega_0$. Obviously, if π is isolated and $\mathcal{T}$ is chained, then $\hat{I}(\pi) = I(\pi)$, but if $I(\pi^{\omega_0}) < \mathbf{1}_{\mathcal{T}}$ holds for each $\omega_0 \in \Omega_0$, i.e., if each π^{ω_0} is isolated, then $\hat{I}(\pi) < \mathbf{1}_{\mathcal{T}}$ follows even when Ω_0 contains more than one element. However, let us postpone a more detailed analysis of the modified entropy function $\hat{I}(\pi)$ till another occasion.

6 General Possibilistically Independent Products of Lattice-Valued Possibilistic Distributions

Let us focus our attention, in this section, to a generalization of the notions, constructions and results leading to Lemma 4.2 above to the case of possibilistically independent (or noninteractive, in other terms) products of nonempty systems of $\mathcal{T}$-possibilistic distributions without any apriori restrictions as far as the cardinality of such systems (i.e., the number of $\mathcal{T}$-possibilistic distribution combined together) is concerned.

Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let $\mathcal{J}$ be a nonempty parametric set (in order to avoid degenerated cases we assume, in what follows, that $\mathcal{J}$ contains at least two elements). For each $i \in \mathcal{J}$, let Ω_i be a nonempty space and $\pi_i : \Omega \to T$ a $\mathcal{T}$-possibilistic distribution so that, for each $i \in \mathcal{J}$, $\bigvee_{\omega \in \Omega_i} \pi_i(\omega) = \mathbf{1}_{\mathcal{T}}$. Without any loss of generality we may and will suppose that $\Omega_i = \Omega$ for each $i \in \mathcal{J}$. Indeed, if this is not the case, we set $\Omega = \bigcup_{i \in \mathcal{J}} \Omega_i$ and we extend each π_i from Ω_i to Ω, setting $\pi_i(\omega) = \oslash_{\mathcal{T}}$ for each $\omega \in \Omega - \Omega_i$. Let $\Omega^{\mathcal{J}}$ denote the set of all mappings $\omega^{\mathcal{J}}$, each of them taking $\mathcal{J}$ into Ω in such a way that, for each $i \in \mathcal{J}$, $\omega^{\mathcal{J}}(i)$ is in Ω_i. Taking as an illustration the most simple case investigated in Lemma 4.2, we have $\mathcal{J} = \{1, 2\}, \Omega^{\mathcal{J}}$ reduces to the set of all mappings of $\{1, 2\}$ into Ω, i.e., into the set of all ordered pairs $\langle \omega_1, \omega_2 \rangle$ of elements of Ω, most often denoted by $\Omega_1 \times \Omega_2$ (or $\Omega \times \Omega$ or Ω^2, if $\Omega_1 = \Omega_2 = \Omega$), and each $\omega^{\mathcal{J}}$ is nothing else than a particular ordered pair of this kind.

Set, for each $\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}$,

$$\pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \bigwedge_{i \in \mathcal{J}} \pi_i(\omega^{\mathcal{J}}(i)) \tag{6.1}$$

The mapping $\pi^{\mathcal{J}} \to T$ is called the *possibilistically independent (or noninteractive) product of $\mathcal{T}$-possibilistic distributions π_i over the parametric space $\mathcal{J}$.*

Lemma 6.1 Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete chained lattice, let $\Omega, \mathcal{J}, \Omega^{\mathcal{J}}, \pi_i : i \in \mathcal{J}$ be as above. Then $\pi^{\mathcal{J}}$ defines a $\mathcal{T}$-possibilistic distribution on $\Omega^{\mathcal{J}}$, i.e., the relation

$$\bigvee_{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \mathbf{1}_{\mathcal{T}} \tag{6.2}$$

holds.

Proof: Let us consider, separately, the two following cases. First, let $\mathcal{T} = \langle T, \leq \rangle$ be such that there exists $t_0 \in T, t_0 < \mathbf{1}_{\mathcal{T}}$ with this property: $t \leq t_0 < \mathbf{1}_{\mathcal{T}}$ holds for each $t \in T, t < \mathbf{1}_{\mathcal{T}}$ (e.g., $T = [0, x] \cup \{1\}, x < 1$, and $\leq$ is the standard linear ordering in $[0, 1]$). As $\bigvee_{\omega \in \Omega} \pi_i(\omega) = \mathbf{1}_{\mathcal{T}}$ is valid for each $i \in \mathcal{J}$, there exists, for each $i \in \mathcal{J}$, $\omega_i^* \in \Omega$ such that $\pi_i(\omega_i^*) = \mathbf{1}_{\mathcal{T}}$, as the ordering $\leq$ is linear. Taking $\omega^{\mathcal{J},*} \in \Omega^{\mathcal{J}}$ such that $\omega^{\mathcal{J},*}(i) = \omega_i^*$ for each $i \in \mathcal{J}$, we obtain immediately that

$$\bigvee_{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \pi^{\mathcal{J}}(\omega^{\mathcal{J},*}) = \bigwedge_{i \in \mathcal{J}} \pi_i(\omega^{\mathcal{J},*}(i)) = \bigwedge_{i \in \mathcal{J}} \pi_i(\omega_i^*) = \mathbf{1}_{\mathcal{T}} \tag{6.3}$$

and (6.2) is proved for this case.

Otherwise, i.e., if such t_0 does not exist, we obtain that $\bigvee\{t \in T : t < \mathbf{1}_{\mathcal{T}}\} = \mathbf{1}_{\mathcal{T}}$ is the case, consequently, as $\leq$ is linear on T, for each $t_1 \in T, t_1 < \mathbf{1}_{\mathcal{T}}$, and for each $i \in \mathcal{J}$ there exists $\omega_i^* \in \Omega$ such that $\pi_i(\omega_i^*) \geq t_i$ holds. Setting $\omega^{\mathcal{J}}(i) = \omega_i^*$ for each $i \in \mathcal{J}$, we obtain that

$$\pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \bigwedge_{i\in\mathcal{J}} \pi_i(\omega^{\mathcal{J}}(i)) = \bigwedge_{i\in\mathcal{J}} \pi_i(\omega_i^*) \geq t_i \tag{6.4}$$

holds, hence,

$$\bigvee_{\omega^{\mathcal{J}}\in\Omega^{\mathcal{J}}} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \bigvee\{t_i \in T : t_i < \mathbf{1}_{\mathcal{T}}\} = \mathbf{1}_{\mathcal{T}} \tag{6.5}$$

follows, and this relation completes the proof of the assertion. □

Let us note that Lemma 6.1 does not hold in general (in the sense that the mapping $\pi^{\mathcal{J}} : \Omega^{\mathcal{J}} \to T$ does not define a $\mathcal{T}$-possibilistic distribution on $\Omega^{\mathcal{J}}$) supposing that the complete lattice $\mathcal{T} = \langle T, \leq\rangle$ is not chained (i.e., the partial ordering $\leq$ is not linear) and the parametric space $\mathcal{J}$ is infinite. For the reader's convenience a counter-example can be found in the Appendix section below.

Theorem 6.1 Let $\mathcal{T}, \Omega, \mathcal{J}$, and $\pi_i, i \in \mathcal{J}$ be as in Lemma 6.1, let $\mathcal{T}$-entropy function $I(\pi)$ be defined by (4.1) for each $\pi_i, i \in \mathcal{J}$, for $\pi^{\mathcal{J}}$ and for each $\pi^{\mathcal{J}-\{i\}}, i \in \mathcal{J}'$. Then the relation

$$I(\pi^{\mathcal{J}}) = \bigvee_{i\in\mathcal{J}} I(\pi_i) \tag{6.6}$$

holds.

Proof: Take $i \in \mathcal{J}$ and set $\Omega_1 = \Omega, \Omega_2 = \Omega^{\mathcal{J}-\{i\}}, \pi_1 = \pi_i$ and $\pi_2 = \pi^{\mathcal{J}-\{i\}}$. As $\mathcal{T}$ is chained, the mappings $\pi^{\mathcal{J}}$ as well as $\pi^{\mathcal{J}-\{i\}}, i \in \mathcal{J}$, define $\mathcal{T}$-possibilistic distributions on Ω_1 and $\Omega^{\mathcal{J}-\{i\}}$. We obtain that for each $\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}$

$$\begin{aligned}\pi^{\mathcal{J}}(\omega^{\mathcal{J}}) &= \bigwedge_{j\in\mathcal{J}} \pi_j(\omega^{\mathcal{J}}(j)) = \pi_i(\omega^{\mathcal{J}}(i)) \wedge \bigwedge_{j\in\mathcal{J}-\{i\}} \pi_j(\omega^{\mathcal{J}}(j)) = \\ &= \pi_i(\omega^{\mathcal{J}}(i)) \wedge \pi^{\mathcal{J}-\{i\}}(\omega^{\mathcal{J}-\{i\}}) = \pi_1(\omega^{\mathcal{J}}(i)) \wedge \pi_2(\omega^{\mathcal{J}-\{i\}}),\end{aligned} \tag{6.7}$$

where $\omega^{\mathcal{J}-\{i\}}$ denotes the restriction of $\omega^{\mathcal{J}}$ to $\mathcal{J} - \{i\}$, so that $\omega^{\mathcal{J}-\{i\}}(j) = \omega^{\mathcal{J}}(j)$ for every $j \in \mathcal{J}$, $j \neq i$. Applying Lemma 4.2 to π_1 on Ω_1 and π_2 on Ω_2, we obtain that the inequality

$$I(\pi^{\mathcal{J}}) = I(\pi_1) \vee I(\pi_2) \geq I(\pi_1) = I(\pi_i) \tag{6.8}$$

follows for each $i \in \mathcal{J}$, consequently, the inequality

$$I(\pi^{\mathcal{J}}) \geq \bigvee_{i\in\mathcal{J}} I(\pi_i) \tag{6.9}$$

is valid. So, if there exists $i_0 \in \mathcal{J}$ such that π_{i_0} is not isolated, then $I(\pi_{i_0}) = \mathbf{1}_{\mathcal{T}} = I(\pi^{\mathcal{J}})$ holds and (6.6) is valid trivially. The same is the case when each π_i is isolated, i.e., $I(\pi_i) < \mathbf{1}_{\mathcal{T}}$ holds for each $i \in \mathcal{J}$, but $\bigvee_{i\in\mathcal{J}} I(\pi_i) = \mathbf{1}_{\mathcal{T}}$ is valid. Again, $I(\pi^{\mathcal{J}}) = \mathbf{1}_{\mathcal{T}}$ follows and (6.6) is proved.

The only case which still remains to be analyzed reads that $\bigvee_{i\in\mathcal{J}} I(\pi_i) = t_0 < \mathbf{1}_{\mathcal{T}}$ holds. The inequality $I(\pi^{\mathcal{J}}) \geq t_0$ then follows from (6.9), let us suppose, in order to arrive at a contradiction, that

$$I(\pi^{\mathcal{J}}) = \bigvee_{\omega^{\mathcal{J}}\in\Omega^{\mathcal{J}}} [\Pi^{\mathcal{J}}(\Omega^{\mathcal{J}} - \{\omega^{\mathcal{J}}\}) \wedge \pi^{\mathcal{J}}(\omega^{\mathcal{J}})] > t_0 \tag{6.10}$$

holds. As $\leq$ is a linear ordering on T, (6.10) yields that three exists $\omega^{\mathcal{J},0} \in \Omega^{\mathcal{J}}$ such that

$$\Pi^{\mathcal{J}}(\Omega^{\mathcal{J}} - \{\omega^{\mathcal{J},0}\}) \wedge \pi^{\mathcal{J}}(\omega^{\mathcal{J},0}) > t_0 \tag{6.11}$$

holds. Applying the assumption of linearity of $\leq$ once more we obtain that the inequalities

$$\Pi^{\mathcal{J}}(\Omega^{\mathcal{J}} - \{\omega^{\mathcal{J},0}\}) = \bigvee_{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}, \omega^{\mathcal{J}} \neq \omega^{\mathcal{J},0}} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}) > t_0 \tag{6.12}$$

and $\pi^{\mathcal{J}}(\omega^{\mathcal{J},0}) > t_0$ hold simultaneously. As

$$\pi^{\mathcal{J}}(\omega^{\mathcal{J},0}) = \bigwedge_{i \in \mathcal{J}} \pi_i(\omega^{\mathcal{J},0}(i)) \tag{6.13}$$

is valid due to (6.1), the inequality $\pi_i(\omega^{\mathcal{J},0}(i)) > t_0$ must be valid for each $i \in \mathcal{J}$. On the other side, as $\leq$ is linear, (6.12) yields that there exists $\omega^{\mathcal{J},1} \in \Omega^{\mathcal{J}} - \{\omega^{\mathcal{J},0}\}$, i.e., $\omega^{\mathcal{J},1} \neq \omega^{\mathcal{J},0}$, such that $\pi^{\mathcal{J}}(\omega^{\mathcal{J},1}) > t_0$ holds. Applying (6.13) to $\pi^{\mathcal{J}}(\omega^{\mathcal{J},1})$ we obtain that $\pi_i(\omega^{\mathcal{J},1}(i)) > t_0$ holds for each $i \in \mathcal{J}$. However, as $\omega^{\mathcal{J},1} \neq \omega^{\mathcal{J},0}$, there exists $i_0 \in \mathcal{J}$ such that $\omega^{\mathcal{J},0}(i_0) \neq \omega^{\mathcal{J},1}(i_o)$, consequently, the inequalities $\pi_{i_0}(\omega^{\mathcal{J},0}(i_o)) > t_0$ and $\pi_{i_0}(\omega^{\mathcal{J},1}(i_o)) > t_0$ hold simultaneously and $\omega^{\mathcal{J},1}(i_0) \in \Omega - \{\omega^{\mathcal{J},0}(i_0)\}$ is the case. Hence, the inequalities $\Pi_{i_0}(\Omega - \{\omega^{\mathcal{J},0}(i_0)\}) > t_0$ and $\pi_{i_0}(\omega^{\mathcal{J},0}(i_0)) > t_0$ hold simultaneously, so that the relation

$$\begin{aligned} & t_0 < \Pi_{i_0}(\Omega - \{\omega^{\mathcal{J},0}(i_0)\}) \wedge \pi_{i_0}(\omega^{\mathcal{J},0}(i_0)) \leq \\ \leq \ & \bigvee_{\omega \in \Omega} (\Pi_{i_0}(\Omega - \{\omega\}) \wedge \pi_{i_0}(\omega)) = I(\pi_{i_0}) \leq \bigvee_{i \in I} I(\pi_i) = t_0 \end{aligned} \tag{6.14}$$

follows – a contradiction. Consequently, also in the case when $\bigvee_{i \in I} I(\pi_i) < \mathbf{1}_{\mathcal{T}}$ holds, the relation (6.6) is valid and the proof of our assertion is completed. □

The result of Lemma 4.4 can be extended from general $\mathcal{T}$-possibilistic distributions π_{12} over Cartesian product $\Omega_1 \times \Omega_2$ to the case of general $\mathcal{T}$-possibilistic distributions – $\pi^{\mathcal{J}}$ over product space $\Omega^{\mathcal{J}}$ in the following way.

Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete chained lattice, let Ω and $\mathcal{J}$ be spaces as in Lemma 6.1, let $\pi^{\mathcal{J}} : \Omega^{\mathcal{J}} \to T$ be a $\mathcal{T}$-possibilistic distribution on $\Omega^{\mathcal{J}}$, so that the relation $\bigvee_{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \mathbf{1}_{\mathcal{T}}$ holds. Given $i \in \mathcal{J}$ and $\omega \in \Omega$, set

$$\pi_i^*(\omega) = \bigvee_{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}, \omega^{\mathcal{J}}(i) = \omega} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}), \tag{6.15}$$

so that the mapping $\pi_i^* : \Omega \to T$ obviously defines the marginal $\mathcal{T}$-possibilistic distribution induced by $\pi^{\mathcal{J}}$ and by $i \in \mathcal{J}$ on Ω (taken as the i-th dimension of $\Omega^{\mathcal{J}}$).

Given a fixed $\omega^{\mathcal{J},0} \in \Omega^{\mathcal{J}}$, set $\omega_i = \omega^{\mathcal{J},0}(i)$ for each $i \in \mathcal{J}$. So, the element $\omega^{\mathcal{J},0}$ of $\Omega^{\mathcal{J}}$ belongs to the set $\{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}} : \omega^{\mathcal{J}}(i) = \omega_i\}$ for each $i \in \mathcal{J}$. So,

$$\begin{aligned} & \pi^{\mathcal{J}}(\omega^{\mathcal{J},0}) \leq \Pi^{\mathcal{J}}(\{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}} : \omega^{\mathcal{J}}(i) = \omega_i\}) = \bigvee_{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}, \omega^{\mathcal{J}}(i) = \omega_i} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}) \\ = \ & \pi_i^*(\omega_i) = \pi_i^*(\omega^{\mathcal{J},0}(i)) \end{aligned} \tag{6.16}$$

holds for each $i \in \mathcal{J}$ using the notation (6.15). As (6.16) is valid for each $\omega^{\mathcal{J},0} \in \Omega^{\mathcal{J}}$ and each $i \in \mathcal{J}$, we obtain that the relation

$$\pi^{\mathcal{J}}(\omega^{\mathcal{J}}) \leq \bigwedge_{i \in \mathcal{J}} \pi_i^*(\omega^{\mathcal{J}}(i)) = (\mathbb{X}_{i \in \mathcal{J}} \pi_i^*)(\omega^{\mathcal{J}}) \tag{6.17}$$

follows, where $\mathbb{X}_{i \in \mathcal{J}} \pi_i^*$ denotes the possibilistically independent product of the $\mathcal{T}$-possibilistic distributions π_i^* on $\Omega^{\mathcal{J}}, i \in \mathcal{J}$. Consequently, as in the proof of Lemma 4.4 and applying Theorem 6.1 (formula (6.6)), we obtain that the inequality

$$I(\pi^{\mathcal{J}}) \leq I(\mathbb{X}_{i \in \mathcal{J}} \pi_i^*) = \bigvee_{i \in \mathcal{J}} I(\pi_i^*) \tag{6.18}$$

extending (4.20) from $\mathcal{J} = \{1, 2\}$ to any $\mathcal{J} \neq \emptyset$ is valid.

7 Conclusions

The results obtained in this paper belong to the group of results containing assertions already known, or even weaker than the already known results, but proved under weaker, or at least incomparable with the original ones, assumptions. These results are far not so popular as those ones explicitly picking up new results, as everything what may be picked up (e.g., assertions of theorems) in our case is from the first sight already known or even trivial, but the weakening of the conditions necessary in order to prove these results is hidden behind the first look horizon.

We have developed lattice-valued variant of the notion of Shannon entropy and we have proved some results perhaps weaker than those achievable when implementing some tools from real-valued entropy functions (as integral parts, not only as informal inspiration into the model in question). However, our results have been obtained without any help of tools borrowed from outside the lattice theory. We have focused our attention to the lattice-valued entropy function for possibilistically independent (non-interactive, in other terms) product of lattice-valued possibilistic distributions and we have proved, for such products, a result syntactically rather similar to that well-known one valid for Shannon entropy of the product of statistically independent probability distributions (just the roles of summation and supremum symbols are interchanged). For possibilistic distributions taking values in complete chained lattice the result holds for products of any system of lattice-valued possibilistic distributions contrary to the fact that for possilibistic distributions taking values in general complete lattices (i.e., not necessary chained, the assertion in question is valid only for products of finite systems of lattice-valued possibilistic distributions.

As far as a possible further theoretical research in the field of lattice-valued entropy functions is concerned, at least the three following directions seem to be interesting and promising. In Section 5 above we mentioned two ways how to overcome the rather restricted sensitivity of the lattice-valued entropy $I(\pi)$ when various cases of lattice-valued possibilistic distributions π with $I(\pi) = \mathbf{1}_{\mathcal{T}}$ are to be distinguished. One remedy consists in replacing the original space Ω, on which π is defined, by a factor-space $\Omega/\approx$, or perhaps by more factor-spaces $\Omega/\approx_i$, induced by different equivalence relations $\approx_i$ on Ω. Another way is to replace the entropy function $I(\pi)$ by a more sophisticated entropy function $\hat{I}(\pi)$ defined by (5.1).

The following approach is perhaps also promising. Let us take the supremum and infimum operators on T not as secondary tools defined by the primary partial (or linear, in particular cases) ordering $\leq$ on T, but rather as triangular norm and corresponding conorm defined as primary operations on T. In this case we may replace this t-norm and conorm by another t-norm and conorm and define a $\mathcal{T}$-valued entropy function as the Sugeno integral of the function $\Pi(\Omega - \{\omega\}) \wedge \pi(\omega)$, but this time with Π defined by the conorm and $\wedge$ replaced by the t-norm in question.

In every case, however, all these proposals would deserve a long and detailed further research effort to be seriously analyzed. Let us hope that we will have an occasion to participate at this research effort in future.

8 Appendix

As a matter of fact, for lattice-valued possibilistic distributions π_i with i ranging over an infinite parametric set $\mathcal{J}$ the mapping $\pi^{\mathcal{J}} : \Omega^{\mathcal{J}} \to T$, defined by (6.1), need not define a $\mathcal{T}$-possibilistic distribution on $\Omega^{\mathcal{J}}$, as the following counter-example demonstrates.

Let $\mathcal{J} = \mathcal{N}^+ = \{1, 2, \dots\}$, let $\Omega = [0,1] - Q$, let $\mathcal{T} = \langle \mathcal{P}(Q), \subset \rangle$, where Q denotes the set of all rational numbers from $[0,1]$, hence, Ω is the set of all irrational numbers from $[0,1]$ and $\mathcal{T}$ trivially defines a complete lattice. For every $n \in \mathcal{N}^+$ and every $j \in \mathcal{N}^+$, $j \leq n$, let

$$R_{j,n} = [(j-1)/n, j/n] = \{x \in [0,1] : (j-1)/n \leq x \leq j/n\}, \tag{8.1}$$

let $Q_{j,n} = R_{j,n} \cap Q$. Obviously, for each $n \in \mathcal{N}^+, \bigcup_{j=1}^{n} Q_{j,n} = Q = \mathbf{1}_{\mathcal{T}}$, and even if the sets $R_{j,n}, j = 1, 2, \dots, n$, are not mutually disjoint, all the points common for two $R_{j,n}$'s are rational numbers. So, for each irrational $\omega \in [0,1]$ and each $n \in \mathcal{N}^+$ there exists just one $1 \leq j \leq n$ such that $\omega \in R_{j,n}$ holds, let us denote this j by $j(\omega, n)$. On the other side, for each irrational $\omega \in [0,1]$ and each rational

$\alpha \in Q, |\alpha - \omega| > 0$ holds, so that, for $n > 1/|\alpha - \omega|, \alpha \in Q_{j(\omega,n),n}$ cannot hold. Moreover, for each $n \in \mathcal{N}^+$ and each $j \leq n$ there exists an irrational real number $\omega \in R_{j,n}$, hence, there exists $\omega \in \Omega$ such that $j(\omega, n) = j$.

Given $n \in \mathcal{N}^+$, let us define the mapping $\pi_n : \Omega \to T(= \mathcal{P}(Q))$ setting $\pi_n(\omega) = Q_{j(\omega,n),n} \in T$ for every $\omega \in \Omega = [0,1] - Q$. Due to the fact quoted two lines above we obtain that $\bigcup_{\omega \in \Omega} \pi_n(\omega) = Q = \mathbf{1}_{\mathcal{T}}$ holds for each $n \in \mathcal{N}^+$, hence, each π_n defines a $\mathcal{T}$-valued possibilistic distribution on Ω. Applying (6.1) we obtain that, for every $\omega^{\mathcal{J}} = \langle \omega_1, \omega_2, \dots \rangle \in \Omega^{\mathcal{J}}$, $\pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \bigcap_{n=1}^{\infty} \pi_n(\omega_n)$.

Let $\{a, b\}, 0 \leq a \leq b \leq 1$, be an open, semi-open, or closed subinterval of $[0, 1]$, i.e., $\{a, b\} = (a, b), [a, b), (a, b]$ or $[a, b]$. For each $\{a, b\}$ set $l(\{a, b\}) = b - a = l(\{a, b\} \cap Q)$. As can be easily verified, for each infinite sequence $\{a_i, b_i\}_{i=1}^{\infty}$ of intervals the relations $\bigcap_{i=1}^{\infty}(\{a_i, b_i\} \cap Q) = (\bigcap_{i=1}^{\infty}\{a_i, b_i\}) \cap Q$ and $l(\bigcap_{i=1}^{\infty}\{a_i, b_i\}) \cap Q) \leq \inf\{l(\{a_i, b_i\} \cap Q : i = 1, 2, \dots\}$ are valid. Applying these relations to $\pi^{\mathcal{J}}(\omega^{\mathcal{J}})$ and $\pi_i(\omega_i)$, $i = 1, 2, \dots$, we obtain that $l(\pi^{\mathcal{J}}(\omega^{\mathcal{J}})) = \inf\{l(\pi_i(\omega_i)) : i = 1, 2, \dots\} = 0$. Moreover, due to the properties of $\pi_i(\omega)$ as defined and analyzed above, the only real number contained in $\bigcap_{n=1}^{\infty} R_{j(\omega,n),n} = \bigcap_{n=1}^{\infty}[(j(\omega, n) - 1)/n, j(\omega, n)/n]$ is the irrational number ω, so that

$$\pi^I(\omega^I) = \bigcap_{n=1}^{\infty} \pi_n(\omega_n) = \bigcap_{n=1}^{\infty} Q_{j(\omega,n),n} =$$

$$= \left(\bigcap_{n=1}^{\infty} [(j(\omega, n) - 1)/n, j(\omega, n)/n]\right) \cap Q = \{\omega\} \cap Q = \emptyset = \oslash_{\mathcal{T}}. \tag{8.2}$$

Informally, the only nonempty interval in $[0, 1]$ which is of the length 0 and which contains the irrational number $\omega \in [0, 1]$ is the singleton $\{\omega\}$ which does not contain any rational number, so that $\{\omega\} \cap Q = \emptyset = \oslash_{\mathcal{T}}$. Hence, $\bigvee_{\omega^{\mathcal{J}} \in \Omega^{\mathcal{J}}} \pi^{\mathcal{J}}(\omega^{\mathcal{J}}) = \oslash_{\mathcal{T}}$, so that the possibilistically independent product of $\mathcal{T}$-possibilistic distributions π_i on Ω, with i ranging over $\mathcal{N}^+ = \{1, 2, \dots\}$, does not define a $\mathcal{T}$-possibilistic distribution on $\Omega^{\mathcal{J}} = \mathbb{X}_{i=1}^{\infty}\Omega_i, \Omega_i = \Omega$ for every $i = 1, 2, \dots$

Hence, the situation with independent products of lattice-valued possibilistic distributions qualitatively differs when shifting our attention from finite to infinite parametric set $\mathcal{J}$. Let us recall the similar problem in probability theory, when infinite products of probability distributions need not define a probability distribution on the Cartesian product in question – indeed, considering an infinite sequence of independent and identically distributed coin tosses with $0 < P(Head) < 1$, each sequence in $\{Head, Tail\}^{\infty}$ occurs with the probability 0.

Acknowledgement

This work was partially supported by the grant No. IAA100300503 of GA AS CR and by the Institutional Research Plan AV0Z 10300504 “Computer Science for the Information Society: Models, Algorithms, Applications”.

Bibliography

[1] G. Birkhoff: Lattice Theory. Third edition. American Mathematical Society, Providence, Rhode Island, 1967.

[2] G. De Cooman: Possibility theory I, II, III. International Journal of General Systems *25* (1997), 291-323, 325-351, 353-371.

[3] A. De Luca and S. Termini: A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. Information and Control *20* (1972), 301-312.

[4] A. De Luca and S. Termini: Entropy of $\mathcal{L}$-fuzzy sets. Information and Control *24* (1974), 55-73.

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

[6] 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.

[7] R. Faure and E. Heurgon: Structures Ordonnées et Algèbres de Boole. Gauthier-Villars, Paris, 1971.

[8] J. A. Goguen: $\mathcal{L}$-fuzzy sets. Journal of Mathematical Analysis and Applications *18* (1967), 145-174.

[9] R. V. L. Hartley: Transmission of information. The Bell Systems Technical Journal *7* (1928), 535-563.

[10] G. J. Klir and M. J. Wierman: Uncertainty-Based Information. Physica-Verlag, Heidelberg-New Youk, 1999.

[11] D. Morales, L. Pardo, and I. Vajda: Uncertainty of discrete stochastic systems: General theory and statistical inference. IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans *26* (1996), 681-697.

[12] C. E. Shannon: The mathematical theory of communication. The Bell Systems Technical Journal *27* (1948), 379-423, 623-656.

[13] R. Sikorski: Boolean Algebras. Second edition. Springer-Verlag, Berlin–Göttingen–Heidelberg–New York, 1964.

[14] R. R. Yager: On the measure of fuzziness and negation, Part I: Membership in the unit interval. International Journal of General Systems *5* (1979), 221-229.

[15] R. R. Yager: On the measure of fuzziness and negation, Part II: Lattices. Information and Control *44* (1980), 236-260.

[16] R. R. Yager: Entropy and specifity in a mathematical theory of evidence. International Journal of General Systems *9* (1983), 249-260.

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

[18] L. A. Zadeh: Fuzzy sets. Information and Control *8* (1965), 338-353.