INFORMATION IN VAGUE DATA SOURCES

Milan Mareš and Radko Mesiar

This paper deals with the concept of the "size" or "extent" of the information in the sense of measuring the improvement of our knowledge after obtaining a message. Standard approaches are based on the probabilistic parameters of the considered information source. Here we deal with situations when the unknown probabilities are subjectively or vaguely estimated. For the considered fuzzy quantities valued probabilities we introduce and discuss information theoretical concepts.[1]

1. INTRODUCTION

The information can be measured by many, completely different, methods with regard to its feature which is significant for a given problem. Here, we are not interested in the technological volume of information measured by the space needed for its storage or by the capacity of channel needed for its transmission. Our attention is focused on the change of knowledge following from the acceptance of a message. Such approach is typical for information theory and its methods.

The classical information theory formulated by C. Shannon (see, e. g., [6, 25, 28]) is based on the stochastic parameters of the source of information, namely on the probabilities of actual messages which essentially means the probabilities of particular symbols from the source alphabet. For more advanced information theoretical methods also the conditional probabilities of symbols in dependence on the foregoing n-tuple are important and their knowledge is necessary. Note that there are several alternative approaches, such as Havrda–Charvát entropy [9], or information theory proposed and developed by Kampé de Fériet [10, 11].

In practical applications, the measurement of these probabilities by means of relative frequencies and with the confidence into the effect of the large numbers laws is simple or at least realizable only in some cases. Very often, it is too complex or even in principle impossible whenever the relevant information source is not used frequently enough.

In such situations, it is not very rational to give up the information theoretical methods completely. The probabilities or conditional probabilities can be subjectively, which

---

[1]This paper is an extended version of our conference contribution [21], where a preliminary look at the discussed problems was presented.

means vaguely, estimated. In a more formal terms, the probabilities of symbols cannot be represented by crisp real numbers from $[0, 1]$ but rather by fuzzy subsets of the same interval. It means that we granulate the uncountable set of probability values into (usually) finite number of fuzzy quantities representing, for example, such vague expressions like "about 0.5", "almost 1", "close to 0", and, may be, some others.

Our paper is focused on the analysis and discussion of the adequacy of such valued probabilities to the information theoretical principles. We also discuss an alternative "fuzzy information theory" using formal tools of fuzzy set theory. Note that our approach differs from "fuzzy information theory" discussed by Vivona and Divari in [27], where the fuzziness is in the domain of information measures, as well as from the idea of "fuzzy entropy" [3, 16] measuring the fuzziness of the considered fuzzy sets.

The paper is organized as follows. In the next section, the basic concept of Shannon's entropy is briefly recalled. Section 3 brings a brief overview of fuzzy quantities and their processing. Sections $4-6$ bring our main results on information in vague data sources. Finally, some concluding remarks are added.

2. SHANNON'S INFORMATION SOURCE

In the whole paper, we denote by A a non-empty and finite set called an *alphabet*. Its elements $a, b, \ldots \in A$ are called *symbols*. Moreover, in the whole paper we denote, for any non-empty set M, by $\mathcal{P}(M)$ the class of all probability distributions over M. Let us denote by $p \in \mathcal{P}(A)$ a *probability distribution* over the alphabet, and by $p(a)$ the probability of a symbol $a \in A$. Suppose that $p(a) > 0$ for all $a \in A$.

Due to [25] or [6, 28], the *information* following from the acceptance of a symbol $a \in A$ is denoted by $I(a)$ and equal to

$$I(a) = \log_2 \frac{1}{p(a)} = -\log_2 p(a). \tag{1}$$

This information is often called the Hartley information measure [8].

The *information source* is defined by the pair (A, p) and its degree of uncertainty is characterized by the so called *entropy* which is denoted by $H(A, p)$ and defined as the expected value of information $I(\cdot)$ over the alphabet A, i.e.,

$$H(A, p) = \sum_{a \in A} p(a) \cdot I(a) = -\sum_{a \in A} p(a) \log_2 p(a). \tag{2}$$

It is a well known result that if we denote by n the number of symbols in A then the entropy $H(A, p)$ is maximal for the $\overline{p} \in \mathcal{P}(A)$ for which

$$\overline{p}(a) = \frac{1}{n} \quad \text{for all } a \in A. \tag{3}$$

One of the elementary concepts of information theory is the *aggregated information*. If $(a, b) \in A^2$, then we can introduce the aggregated information $I(a, b)$ and define it by the sum $I(a, b) = I(a) + I(b)$. The sum is a natural consequence of (1) and the properties of probability. Indeed, without going deeper into the necessary formalism of probabilities on product spaces, considering $p(a, b) = p(a)p(b)$, we get

$$I(a, b) = -\log_2 p(a, b) = -\log_2 p(a) + (-\log_2 p(b)) = I(a) + I(b).$$

The Shannon information-theoretical model is successful and effective but it is not unique. Namely, it is based on perfect knowledge of probabilities acting in relations (1) and (2). Nevertheless, even if it is possible (and, may be, useful) to suggest alternative approaches to the quantification of such qualitative concepts like information or knowledge, there exist some general structures which are to be respected to preserve the essence of the above concepts.

Here, we suggest heuristic formulation of those structures. Recall that the classical information theory was developed with respect to these principles:

- Information following from some message, its segment or a symbol is the larger the less expected the message (segment, symbol) is.

- Information following from two (or more) messages (their segments or symbols) is cumulated; i. e., the value of information of associated symbols is aggregated from the information connected with the components.

- The source of information can be characterized by a function of information following from particular symbols of the alphabet.

- This characteristic should reflect the organization (or, vice versa, chaotism) of the source. In other words, its value has to depend on the diversity of information following from particular symbols.

- The source in which all symbols are equally expected is to be characterized as the most chaotic one.

3. FUZZY QUANTITIES AND THEIR PROCESSING

We suppose the reader to be familiar with the basic notions and concepts of fuzzy set theory [29], such as fuzzy subset, fuzzy union, fuzzy complement, etc.

Due to [4, 17, 18] and other papers, *fuzzy quantity*, in general, is a fuzzy subset $\boldsymbol{r}$ of the set of real numbers $\mathbb{R}$ with *membership function* $\mu_{\boldsymbol{r}} : \mathbb{R} \to [0, 1]$ such that:

$$\text{There exists } x_{\boldsymbol{r}} \in \mathbb{R} \text{ for which } \mu_{\boldsymbol{r}}(x_{\boldsymbol{r}}) = 1. \tag{4}$$

$$\text{There exist } x_1, x_2 \in \mathbb{R}, \ x_1 < x_{\boldsymbol{r}} < x_2, \text{ where } \mu_{\boldsymbol{r}}(x) = 0 \text{ for } x \notin [x_1, x_2]. \tag{5}$$

$$\mu_{\boldsymbol{r}} \text{ is nondecreasing on } [x_1, x_{\boldsymbol{r}}] \text{ and nonincreasing on } [x_{\boldsymbol{r}}, x_2]. \tag{6}$$

Any real number $x_{\boldsymbol{r}}$ fulfilling (4) is called the *modal value* of $\boldsymbol{r}$.

For any non-empty set M we denote by $\mathcal{F}(M)$ the class of all its fuzzy subsets. It means that fuzzy quantities are fuzzy sets from $\mathcal{F}(\mathbb{R})$ fulfilling (4), (5) and (6). By $\mathcal{F}^*(\mathbb{R})$ we denote $\mathcal{F}^*(\mathbb{R}) = \{\boldsymbol{r} \in \mathcal{F}(\mathbb{R}) : \boldsymbol{r} \text{ fulfils } (4), (5) \text{ and } (6)\}$.

To simplify notation, for any $x \in \mathbb{R}$ we denote by $\langle x \rangle \in \mathcal{F}^*(\mathbb{R})$ the fuzzy quantity for which

$$\mu_{\langle x \rangle}(x) = 1, \quad \mu_{\langle x \rangle}(y) = 0 \quad \text{for } y \in \mathbb{R}, \ y \neq x,$$

formally, $\langle x \rangle$ can be seen as the characteristic function of the singleton $\{x\} \subset \mathbb{R}$.

The algebraic operations over fuzzy quantities are well investigated and summarized in many publications, e. g., in [1, 4, 17, 18], and related results are also in [12]. Those

operations are based on the so called *extension principle* [30]. We give here its most general form, though our next considerations deal with the original form based on the minimum as fuzzy connective.

Recall that a mapping $T\colon [0,1]^2 \to [0,1]$ is called a triangular norm whenever it is associative, commutative, nondecreasing in both arguments and 1 is its neutral element [13]. For any classical binary relation $\mathcal{R} \subset \mathbb{R}^2$, its T-based fuzzy extension $\mathcal{FR}_T\colon \mathcal{F}^*(\mathbb{R})^2 \to [0,1]$ is given by, see [24],

$$\mathcal{FR}_T(\boldsymbol{r}_1, \boldsymbol{r}_2) = \sup_{(x_1,x_2)\in\mathcal{R}} T\left(\mu_{\boldsymbol{r}_1}(x_1), \mu_{\boldsymbol{r}_2}(x_2)\right). \tag{7}$$

Formula (7) allows to introduce fuzzy extensions of binary relations over reals, standard real functions, etc. We recall some of these extensions, considering the strongest triangular norm T_M, $T_M(x,y) = \min(x,y)$.

If $\boldsymbol{r},\ \boldsymbol{s} \in \mathcal{F}^*(\mathbb{R})$ and $\circ$ is a binary operation over $\mathbb{R}$, then $\circ$ can be extended on $\mathcal{F}^*(\mathbb{R})$ and $\boldsymbol{r} \circ \boldsymbol{s}$ is a fuzzy quantity, see [30], with

$$\mu_{\boldsymbol{r}\circ\boldsymbol{s}}(z) = \sup_{\substack{x,y\in\mathbb{R} \\ z=x\circ y}} \min(\mu_{\boldsymbol{r}}(x),\ \mu_{\boldsymbol{s}}(y)), \quad z \in \mathbb{R}. \tag{8}$$

For the operations of addition $\oplus$ and product $\odot$ over fuzzy quantities, (8) gains the form

$$\mu_{\boldsymbol{r}\oplus\boldsymbol{s}}(z) = \sup_{x+y=z} \min(\mu_{\boldsymbol{r}}(x),\ \mu_{\boldsymbol{s}}(y)), \quad z \in \mathbb{R}, \tag{9}$$

$$\mu_{\boldsymbol{r}\odot\boldsymbol{s}}(z) = \sup_{x\cdot y=z} \min(\mu_{\boldsymbol{r}}(x),\ \mu_{\boldsymbol{s}}(y)), \quad z \in \mathbb{R}. \tag{10}$$

Observe that the class $\mathcal{F}^*(\mathbb{R})$ is closed under operations $\oplus$ and $\odot$, see [28].

Example 3.1. Consider fuzzy quantities $\boldsymbol{r},\ \boldsymbol{s} \in \mathcal{F}^*(\mathbb{R})$ with membership functions given by

$$\mu_{\boldsymbol{r}}(x) = \max(0, \min(x, 2-x)) = \begin{cases} x & \text{if } x \in [0,1], \\ 2-x & \text{if } x \in\ ]1,2], \\ 0 & \text{else,} \end{cases}$$

and

$$\mu_{\boldsymbol{s}}(x) = \max\left(0, \min\left(1, x, 2-x/2\right)\right) = \begin{cases} x & \text{if } x \in [0,1], \\ 1 & \text{if } x \in\ ]1,2], \\ 2-x/2 & \text{if } x \in\ ]2,4], \\ 0 & \text{else.} \end{cases}$$

Then

$$\mu_{\boldsymbol{r}\oplus\boldsymbol{s}}(z) = \max\left(0, \min\left(1, z/2, 2-z/3\right)\right) = \begin{cases} z/2 & \text{if } z \in [0,2], \\ 1 & \text{if } z \in\ ]2,3], \\ 2-z/3 & \text{if } z \in\ ]3,6], \\ 0 & \text{else,} \end{cases}$$

and

$$\mu_{\boldsymbol{r}\odot\boldsymbol{s}}(z) = \begin{cases} \sqrt{z} & \text{if } z \in [0,1], \\ 1 & \text{if } z \in\ ]1,2], \\ 2-\sqrt{z/2} & \text{if } z \in\ ]2,8], \\ 0 & \text{else.} \end{cases}$$

In the following sections, we also need the following concepts. If $r \in \mathcal{F}(\mathbb{R})$ is a fuzzy quantity, then its *opposite* fuzzy quantity $(-r) \in \mathcal{F}^*(\mathbb{R})$ and *reciprocal* fuzzy quantity $(1/r) \in \mathcal{F}^*(\mathbb{R})$ (supposing $\mu_r(0) = 0$) are defined by

$$\mu_{(-r)}(x) = \mu_r(-x), \quad \mu_{(1/r)}(x) = \mu_r(1/x) \tag{11}$$

for $x \in \mathbb{R}$, and $x \neq 0$ for the reciprocity. For $x = 0$, $\mu_{(1/r)}(x) = 0$ by definition.

There are many different suggestions how to compare and order fuzzy quantities. Their representative overview is given in [12], some of them are discussed in [4, 18] and in some other papers. Here, we use the method based on the paradigm that relations between vague objects are naturally vague. A fuzzy relation of *ordering* $\succeq$ on $\mathcal{F}^*(\mathbb{R})$ with membership function $\nu_\succeq(r, s) : \mathcal{F}^*(\mathbb{R}) \times \mathcal{F}^*(\mathbb{R}) \to [0, 1]$ is given by

$$\nu(r, s) = \sup_{x \geq y} \min(\mu_r(x), \mu_s(y)). \tag{12}$$

Finally, let $f : \mathbb{R} \to \mathbb{R}$ be a monotonous and continuous function. Then we can extend it on fuzzy arguments and fuzzy values of function f using a simple formula

$$\text{for } r \in \mathcal{F}^*(\mathbb{R}), \ f(r) \in \mathcal{F}^*(\mathbb{R}), \text{ as well, and } \mu_{f(r)}(x) = \mu_r(f(x)) \tag{13}$$

where $f : \mathcal{F}^*(\mathbb{R}) \to \mathcal{F}^*(\mathbb{R})$.

The methodology of the treatment of vagueness by means of fuzzy set theoretical tools is deeply analyzed and assessed in [31, 34, 35] and other related papers. The ideas presented here are especially significant for processing of quantitative data, uncertain "measurements", ordering the qualitative attributes into potentially quantifiable scales and other bridges between quantitative parameters and qualitative phenomena.

4. GRANULATED EXPECTATIONS OF INFORMATION SOURCES

We have already mentioned that classical information theory is based on the assumption that the probability distribution p over the alphabet A is exactly known and, consequently, the information source (A, p) is strictly defined. In many cases, this assumption appears too strong – the probabilities are not exactly known and instead of them we rather dispose with some vague expectations. Usually subjective, often rather verbal than numerical. The expectations are expressed by words like "very low", "medium", "high", or several similar formulations. In this way, the possibly rich alphabet A is parted in several clusters of symbols whose appearance in messages is expected with subjective probability characterized by the above words. Such clustering or, in another term, a granulation (see [31, 34, 35]), is formally reflected by the methods and concepts of fuzzy set (and fuzzy quantities) theory.

Let us note that even in the case of very simple, e. g., binary, alphabet A containing exactly two symbols, the granulation has its sense. Most of information-theoretical methods deal with conditional probabilities (and, consequently, conditional information, etc.) conditioned by an n-tuple of foregoing symbols in the message. In such case, the frequencies and probabilities of ordered $(n + 1)$ tuples are to be considered, and their number rapidly grow.

As we have excluded the existence of symbols with zero probability (cf. Section 2), the probabilities of particular symbols gain crisp (and unknown) probabilities from $(0, 1]$. Their vague estimations, called here the *subjective expectation* of symbols (or their ordered n-tuples) are formally represented by fuzzy quantities with possible values in $(0, 1]$. More formally, let us denote

$$\mathcal{F}^*((0, 1]) = \{\boldsymbol{r} \in \mathcal{F}^*(\mathbb{R}) \mid \mu_{\boldsymbol{r}}(x) = 0 \text{ for all } x \notin (0, 1]\}. \tag{14}$$

Then the *expectation* of a symbol $a \in A$ is a fuzzy quantity

$$\boldsymbol{p}(a) \in \mathcal{F}^*((0, 1]). \tag{15}$$

This expectation of $a \in A$ specifies the probabilities with which a can be expected and the possibility degrees with which they can be expected.

It could be useful to point at the formal similarity between the above concept of expectations, and the formal apparatus of triangular norms and conorms. This relation is not the topic of this contribution (cf. [13, 2]).

Similarly, the above definitions resemble the concept of ultra-fuzzy set (see [33]). Even this relation, however close it can be, is not investigated in this paper.

Using the above procedure, we can define the *fuzzy source* of information as a pair $(A, (\boldsymbol{p}(a))_{a \in A})$ whose components are specified above. To simplify notation, we use the symbol $\boldsymbol{p}$ instead of $(\boldsymbol{p}(a))_{a \in A}$ when it is possible. Then $(A, \boldsymbol{p})$ denotes the fuzzy source. The main purpose of the following subsections is to verify and discuss some approaches to the formal representation of the information produced by that fuzzy source.

5. SHANNON'S ENTROPY OF FUZZY SOURCES

Shannon's entropy of a classical information source (A, p) given by (2) can be seen as a function of n variables, $n = \text{card} A$, as well as an $(n+1)$-dimensional real relation. In the first case one can apply the fuzzy extension of real functions as discussed in Section 3. This approach was considered in our preliminary contribution [21].

Theorem 5.1. Let us consider a fuzzy source $(A, \boldsymbol{p})$. Then the fuzzy entropy

$$\boldsymbol{H}(A, \boldsymbol{p}) = \sum_{a \in A}^{\oplus} \boldsymbol{p}(a) \odot \log_2(1/\boldsymbol{p}(a)),$$

with the membership function $\mu_{\boldsymbol{H}(A, \boldsymbol{p})}$ given by

$$\mu_{\boldsymbol{H}(A, \boldsymbol{p})}(x) = \sup \left\{ \min\{\mu_{\boldsymbol{p}(a)}(z_a) \mid a \in A\} \mid z_a \in (0, 1], \sum_{a \in A} z_a \log_2(1/z_a) = x \right\}$$

is a fuzzy quantity belonging to $\mathcal{F}^*((0, \infty))$.

Note that $\mathcal{F}^*((0, \infty))$ is formed by fuzzy quantities from $\mathcal{F}^*(\mathbb{R})$ with support in $(0, \infty)$.

The previous result can be formulated in a rather stronger way. Let us consider fuzzy source $(A, \boldsymbol{p})$ with granulated expectations of symbols. It means that usually

several symbols $a_1, a_2, \ldots, a_m \in A$ are expected with the same fuzzy expectation $\boldsymbol{p}(a_1) = \boldsymbol{p}(a_2) = \cdots = \boldsymbol{p}(a_m)$.

Let us denote by n the number of symbols in A, and for every symbol $a_i \in A$ and fuzzy probability $\boldsymbol{p}(a_i) \in \mathcal{F}^*((0,1])$ we denote by $x_{a_i} \in (0,1]$ the modal value of $\boldsymbol{p}(a_i)$, $i = 1, 2, \ldots, n$ (let us note that there may be more than one modal value for every $\boldsymbol{p}(a_i)$).

Let $\boldsymbol{p}(a_i)$, $i = 1, \ldots, n$, be fuzzy expectations, and let there exist modal values x_{a_i} of $\boldsymbol{p}(a_i)$, $i = 1, \ldots, n$, respectively, such that

$$\sum_{i=1}^{n} x_{a_i} = 1.$$

Then we say that the fuzzy source $(A, \boldsymbol{p})$ is a *fuzzy extension* of the crisp source $(A, (x_{a_i})_{i=1}^{n})$. As follows from the results summarized in [17, 18], the properties of modal values characterize even the fuzzy expectations.

Lemma 5.2. If $(A, \boldsymbol{p})$ is a fuzzy extension of $(A, (x_a)_{a \in A})$ for some crisp values x_a, $a \in A$, and if we denote by $x_{\boldsymbol{H}}$ the crisp entropy of $(A, (x_a)_{a \in A})$, given by

$$x_{\boldsymbol{H}} = H\left(A, (x_a)_{a \in A}\right),$$

then $x_{\boldsymbol{H}}$ is a modal value of $\boldsymbol{H}(A, \boldsymbol{p})$.

Theorem 5.3. Let $(A, \overline{\boldsymbol{p}})$ be a fuzzy source, let for any $a \in A$, $\overline{x}_a$ be a modal value of $\overline{\boldsymbol{p}}(a)$, and let $\overline{x}_a = 1/n$ for all $a \in A$. Let, moreover, $(A, \boldsymbol{p})$ be a fuzzy source with modal values x_a of $\boldsymbol{p}(a)$, $a \in A$, such that

$$\sum_{a \in A} x_a = 1.$$

Then

$$\nu_{\succeq}\left(\boldsymbol{H}(A, \overline{\boldsymbol{p}}), \boldsymbol{H}(A, \boldsymbol{p})\right) = 1.$$

The previous statements show that the direct and relatively mechanical substitution of crisp probabilities by fuzzy expectations is possible. Moreover, if the fuzzy expectations extend some crisp source (it means that they respect its probabilistic structure) then the result preserves the sense of the main information-theoretical principles.

Nevertheless, the processing of fuzzy data often cumulates the vagueness of data. Even quite "narrow" supports of the membership functions on the input (i. e., in fuzzy expectations) usually generate "wide" supports of fuzzy information and, consequently, also of fuzzy entropy. This leads to the growth of vagueness of the relevant fuzzy quantities which is not desirable.

To reduce the growth of vagueness fuzzy processing suffers from, in several papers the constraint fuzzy arithmetic was proposed, see, e. g. [7, 15]. In our case this approach corresponds to the relational look on Shannon's entropy. Based on the results from [7, 15], the next result can be shown.

Theorem 5.4. Let $(A, \boldsymbol{p})$ be a fuzzy source. Then the constraint fuzzy entropy $c\boldsymbol{H}(A, \boldsymbol{p})$ characterized by the membership function $\mu_{c\boldsymbol{H}(A, \boldsymbol{p})}$ given by

$$\mu_{c\boldsymbol{H}(A, \boldsymbol{p})}(x) = \sup\left\{\min\{\mu_{\boldsymbol{p}(a)}(z_a) | a \in A\} | z_a \in (0,1], \sum_{a \in A} z_a = 1, \sum_{a \in A} z_a \log_2(1/z_a) = x\right\}$$

is a fuzzy quantity belonging to $\mathcal{F}^*((0, \infty))$.

Evidently, Lemma 5.2 can be applied to the constrained fuzzy entropy, too.

Example 5.5. Let $A = \{a, b\}$ be an alphabet with two symbols and let the fuzzy source $(A, \boldsymbol{p})$ be given by $\boldsymbol{p}(a) = [1/5, 2/5]$, $\boldsymbol{p}(b) = [3/5, 4/5]$, i. e.,

$$\mu_{\boldsymbol{p}(a)}(x) = \left\{ \begin{array}{ll} 1 & \text{if } x \in [1/5, 2/5], \\ 0 & \text{otherwise}, \end{array} \right.$$

and similarly $\mu_{\boldsymbol{p}(b)}$. Then

$$\boldsymbol{H}(A, \boldsymbol{p}) = \left[\log_2 5 - \frac{8}{5} \frac{1}{e} \log_2 e + \frac{3}{5} \log_2 \frac{5}{3} \right] = [0.722, 0.973]$$

and

$$c\boldsymbol{H}(A, \boldsymbol{p}) = \left[\log_2 5 - \frac{8}{5}, \log_2 5 - \frac{1}{5} \log_2 108 \right] = [0.722, 0.971] \subset \boldsymbol{H}(A, \boldsymbol{p}).$$

6. PROCESSING OF FUZZY INFORMATION SOURCES

Fuzzy set theory aims to describe and analyze another type of uncertainty than probability theory. It can motivate the endeavour to suggest not only an alternative interpretation of the model of fuzzy source but also alternative approaches to its processing. Here, we briefly mention one of them.

Let us consider a fuzzy source $(A, \boldsymbol{p})$, described above, where $\boldsymbol{p}(a)$ is *fuzzy expectation* of $a \in A$, and with membership function $\mu_{\boldsymbol{p}(a)} \in \mathcal{F}^*((0, 1])$.

For every $a \in A$, we define *fuzzy knowledge mediated by* a, denoted by $\boldsymbol{K}(a) \in \mathcal{F}(\mathbb{R})$, with membership function $\mu_{\boldsymbol{K}(a)}$ such that

$$\mu_{\boldsymbol{K}(a)}(x) = \mu_{\boldsymbol{p}(a)}(1 - x), \quad x \in \mathbb{R}. \tag{16}$$

Remark 6.1.   (i) Evidently, $\boldsymbol{K}(a) \in \mathcal{F}^*([0, 1])$, as follows from (16) and (15).

(ii) For any fuzzy negation $N : [0, 1] \to [0, 1]$, $N(1) = 0$, $N(0) = 1$ and $N(x) \le N(y)$ whenever $x \ge y$, one can define N-fuzzy knowledge mediated by a, denoted by $N\boldsymbol{K}(a) \in \mathcal{F}(\mathbb{R})$, with membership function $\mu_{N\boldsymbol{K}(a)}$ given by

$$\mu_{N\boldsymbol{K}(a)}(x) = \left\{ \begin{array}{ll} \mu_{\boldsymbol{p}(a)}(N(x)) & \text{if } x \in [0, 1], \\ 0 & \text{otherwise}. \end{array} \right.$$

Evidently, $N\boldsymbol{K}(a) \in \mathcal{F}^*([0, 1])$. As in the case of triangular norms, where we have considered the original Zadeh's approach based on $T_M = \min$ only, also in the case of fuzzy knowledge we will only consider the original fuzzy negation $N_Z : [0, 1] \to [0, 1]$ given by $N_Z(x) = 1 - x$. Obviously, $N_Z\boldsymbol{K}(a) = \boldsymbol{K}(a)$ for all $a \in A$.

If a, $b \in A$ then we can define the aggregated fuzzy knowledge $\boldsymbol{K}(a,b)$ with membership function $\mu_{\boldsymbol{K}(a,b)}$ given by

$$\mu_{\boldsymbol{K}(a,b)}(x) = \max\left(\mu_{\boldsymbol{K}(a)}(x),\, \mu_{\boldsymbol{K}(b)}(x)\right), \quad x \in \mathbb{R}. \tag{17}$$

It means that $\boldsymbol{K}(a,b) = \boldsymbol{K}(a) \cup \boldsymbol{K}(b)$ in the usual fuzzy set theoretical sense (cf., [29]).

Evidently, the fuzzy knowledge is a fuzzy counterpart of the probabilistic information $I(a)$ and its fuzzy modification, fuzzy expectation $\boldsymbol{I}(a)$. This measure of vague information reflects the fuzzy set theoretical methodology and it is based on monotonicity instead of additivity.

Lemma 6.2. Let a, $b \in A$. Then $\nu_{\succeq}(\boldsymbol{p}(a), \boldsymbol{p}(b)) = \nu_{\succeq}(\boldsymbol{K}(b), \boldsymbol{K}(a))$.

P r o o f . The statement follows immediately from (16) and (12). $\qquad\square$

The modified concept of entropy, reflecting the typical fuzzy style, would be rather more difficult. Let us discuss two possible approaches.

The first possibility is to define the analogy of entropy for fuzzy sources as fuzzy quantity. We call it a *fuzzy disorganization* of fuzzy source $(A, \boldsymbol{p})$ and denote $\boldsymbol{D}(A, \boldsymbol{p}) \in \mathcal{F}([0,1])$ with membership function $\mu_{\boldsymbol{D}} : [0,1] \to [0,1]$, where

$$\mu_{\boldsymbol{D}}(x) = \min\left\{\mu_{\boldsymbol{K}(a)}(x) \mid a \in A\right\}, \quad x \in [0,1]. \tag{18}$$

Theorem 6.3. Let $(A, \overline{\boldsymbol{p}})$ be a fuzzy source with fuzzy expectations $\overline{\boldsymbol{p}}(a)$, $a \in A$, and let $\overline{x}_a \in (0,1)$ be modal values, $a \in A$, such that

$$\sum_{a \in A} \overline{x}_a = 1.$$

Let, moreover, $\overline{x}_a = 1/n$. Then there exists a modal value $x_{\boldsymbol{D}} \in [0,1]$ of $\overline{\boldsymbol{D}}(A, \overline{\boldsymbol{p}})$, and for each fuzzy disorganization $\boldsymbol{D}(A, \boldsymbol{p})$ of some fuzzy source $(A, \boldsymbol{p})$ with membership function $\mu_{\boldsymbol{D}}$ and each $x \in [0,1]$ it holds

$$\mu_{\overline{\boldsymbol{D}}}(\overline{x}_{\boldsymbol{D}}) = 1 \geq \mu_{\boldsymbol{D}}(x).$$

P r o o f . If the modal values of all fuzzy expectations $\boldsymbol{p}(a)$, $a \in A$, are mutually equal then $\mu_a(1/n) = 1$ for all $a \in A$, and, due to (16) and (18), $\mu_{\overline{\boldsymbol{D}}}(\overline{x}_{\boldsymbol{D}}) = 1$ for $\overline{x}_{\boldsymbol{D}} = (n-1)/n$. As for any fuzzy disorganization $\boldsymbol{D}$, $\mu_{\boldsymbol{D}}(x) \in [0,1]$ for each $x \in [0,1]$, the statement is true. $\qquad\square$

The alternative possibility is to define the measure of disorganization as a crisp quantity. If $(A, \boldsymbol{p})$ is a fuzzy source, $\mu_{\boldsymbol{D}}$ is given by (18), then we denote by $d(A, \boldsymbol{p}) \in [0,1]$ and call *crisp disorganization* of $(A, \boldsymbol{p})$ the number

$$d(A, \boldsymbol{p}) = \sup\left\{\mu_{\boldsymbol{D}}(x) \mid x \in (0,1)\right\}. \tag{19}$$

Theorem 6.4. Under the notations and assumptions of Theorem 6.3,

$$d\left(A, \overline{\boldsymbol{p}}\right) = 1 \geq d(A, \boldsymbol{p})$$

for any fuzzy source $(A, \boldsymbol{p})$.

P r o o f .   The statement follows immediately from (19) and Theorem 6.3.   □

Above, we have suggested two ways to quite consequent fuzzifications of the basic information-theoretical concepts. The first one of them appears a bit more adequate to the vagueness of the input expectations.

Nevertheless, both models are based, namely, on the vague estimation of the particular probabilities of symbols, and, consequently, on the granulation of the expectations. The fuzzy information, fuzzy knowledge, fuzzy disorganization, and in some sense also the crisp disorganization, reflect rather the fuzziness of particular symbols of the alphabet, than the fuzziness of the entire alphabet and of the choice of its symbols. Let us discuss this problem.

We still consider a finite and non-empty alphabet A, and instead of (crisp) probability p we consider a fuzzy subset Q of A, with membership function $\pi_Q : A \to [0, 1]$ such that $\pi_Q(a) = 1$ for some $a \in A$. Observe that π_Q can be seen as a possibility distribution [32] and it defines a possibility measure π on A given by

$$\pi(B) = \max\left\{\pi_Q(a) \mid a \in B\right\}, \quad B \subseteq A.$$

In our notation, $Q \in \mathcal{F}(A)$. The fuzzy set Q represents our vague (mostly subjective) knowledge of the *possibilities*, with which particular symbols from the alphabet may appear in messages produced by the source.

Then, it is quite acceptable to define the *information* transmitted by symbols as a fuzzy subset $I(Q)$ of A, with membership function $\rho_Q : A \to [0, 1]$, such that

$$I(Q) = \overline{Q}, \quad \text{i.e.,} \quad \rho_Q(a) = 1 - \pi_Q(a), \quad a \in A. \tag{20}$$

If $(a, b) \in A^2$, based on possibility theory on product spaces, see [5], the *aggregated possibility* of that pair can be defined as

$$\pi_Q(a, b) = \min\left(\pi_Q(a), \pi_Q(b)\right), \tag{21}$$

which is a possibilistic counterpart of the probability approach $p(a, b) = p(a)p(b)$ considered in Section 2.

Remark 6.5. Relations (20) and (21) imply that aggregated information transmitted by the pair (a, b) is

$$\rho_Q(a, b) = \max\left(\rho_Q(a), \rho_Q(b)\right).$$

Finally, the possibilistic entropy $H(A, Q)$ can be then defined as

$$H(A, Q) = \max\left\{\min\left(\pi_Q(a), \rho_Q(a)\right) \mid a \in A\right\} = \max\left\{\min\left(\pi_Q(a), 1 - \pi_Q(a)\right) \mid a \in A\right\}. \tag{22}$$

Note that the Shannon entropy (2) can be seen as the Lebesgue integral $H(A, p) = \int_A I \, \mathrm{d}\, p$, while the possibilistic entropy (22) corresponds to the Sugeno integral [26], $H(A, Q) = Su - \int_A \rho_Q \, \mathrm{d}\pi$.

7. CONCLUSIVE REMARKS

The essential properties of vagueness rather differ from those of randomness. It means that also their models, fuzzy set theory and probability theory, are to be processed in different ways. Nevertheless, comparing Sections 2 and 4-6 from the point of view of Section 3, we may see that some general principles are valid in all presented models.

These general principles reflect some abstract models of uncertainty processing and related aggregation operations. There exist models related to fuzziness and probability dealing with triangular norms and conorms as the relevant aggregation operators (e. g., [13] or [2]), see the previous sections. Their properties with regard to the quantitative measurement of information were not analyzed yet and some discussion of them keeps an open topic.

(Received April 24, 2012)

R E F E R E N C E S

[1] J. C. Bezdek, ed.: Analysis of Fuzzy Information. CRC–Press, Boca Raton 1988.

[2] T. Calvo, G. Mayor, and R. Mesiar, eds.: Aggregation Operators. New Trends and Applications, Physica–Verlag, Heidelberg 2002.

[3] A. De Luca and S. Termini: A definition of a non probabilistic entropy in the setting of fuzzy set theory. Inform. and Control *20* (1972), 301–312.

[4] D. Dubois, E. Kerre, R. Mesiar, and H. Prade: Fuzzy interval analysis. In: Fundamentals of Fuzzy Sets, Kluwer, Dordrecht 2000, pp. 483–581.

[5] D. Dubois and H. Prade: Possibility Theory. An Approach to Computerized Processing of Uncertainty. Plenum Press, New York 1988.

[6] A. Feinstein: Foundations of Information Theory. McGraw–Hill, New York 1957.

[7] R. E. Giachetti and R. E. Young: A parametric representation of fuzzy numbers and their arithmetic operators. Fuzzy Sets and Systems *91* (1997), 185–202.

[8] R. V. L. Hartley: Transmission of information. Bell System Techn. J. *7* (1928), 3, 535–563.

[9] J. Havrda and F. Charvát: Quantification method of classifications processes. Concept of structural a-entropy. Kybernetika *3* (1967), 1, 30–35.

[10] J. Kampé de Fériet: Théories de l'information. Springer, Berlin 1974.

[11] J. Kampé de Fériet and B. Forte: Information et probabilité. C. R. Acad. Sci. Paris, Sér. A *265* 1967, 110–114; 142–146; 350–353.

[12] E. E. Kerre and X. Wang: Reasonable properties for the ordering of fuzzy quantities. Part I., Part II. Fuzzy Sets and Systems *118* (2001), 375–385; 387–405.

[13] E. P. Klement, R. Mesiar, and E. Pap: Triangular Norms. Kluwer, Dordrecht 2000.

[14] G. J. Klir and T. A. Folger: Fuzzy Sets, Uncertainty and Information. Prentice Hall, Englewood Cliffs 1988.

[15] G. J. Klir: Fuzzy arithmetic with requisite constraints. Fuzzy Sets and Systems *91* (1997), 2, 165–175.

[16] A. Kolesárová and D. Vivona: Entropy of T-sums and T-products of $L - R$ fuzzy numbers. Kybernetika *37* (2001), 2, 127–145.

[17] M. Mareš: Computation Over Fuzzy Quantities. CRC–Press, Boca Raton 1994.

[18] M. Mareš: Weak arithmetics of fuzzy numbers. Fuzzy Sets and Systems *91* (1997), 143–154.

[19] M. Mareš: Compenzational vagueness. In: Proc. EUSFLAT'2007 (M. Štepnička, V. Novák, and U. Bodenhofer, eds.), Ostrava 2007, pp. 179–184.

[20] M. Mareš and R. Mesiar: Verbally generated fuzzy quantities. In: Aggregation Operators. New Trends and Applications(T. Calvo, G. Mayor, and R. Mesiar (eds.), Physica–Verlag, Heidelberg 2002, pp. 291–353.

[21] M. Mareš and R. Mesiar: Information in granulated data sources. In: Proc. Internat. Conf. on Soft Computing, Computing with Words and Perceptions in Systems (W. Pedrycz, R. Aliev, Mo Jamshidi, and B. Turksen, eds.), Antalya 2007.

[22] R. Mesiar: Triangular norms-based addition of fuzzy intervals. Fuzzy Sets and Systems *91* (1997), 73–78.

[23] R. Mesiar and S. Saminger: Domination of ordered weighted averaging operators over *t*-norms. Soft Computing *8* (2004), 562–570.

[24] H. T. Nguyen: A note on the extension principle for fuzzy sets. J. Math. Anal. Appl. *64* (1978), 369–380.

[25] C. Shannon and, W. Weaver: A mathematical theory of communication. Bell System Techn. J. *27* (1948), 379–423; 623–653.

[26] M. Sugeno: Theory of Fuzzy Integrals and Its Applications. PhD. Thesis, Tokyo Institute of Technology, 1974.

[27] D. Vivona and M. Divari: On a conditional information for fuzzy sets. In: Proc. AGOP'2005, Lugano 2005, pp. 147-149.

[28] K. Winkelbauer: Communication channels with finite past history. In: Trans. Second Prague Conference, Statistical Decision Functions and Random Processes, Prague 1959. Publishing House of the Czechoslovak Academy of Sci., Prague 1960, pp. 685–831.

[29] L. A. Zadeh: Fuzzy sets. Inform. and Control *8* (1965), 338–353.

[30] L. A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning. Inform. Sci. (1975), Part I: *8*, 199–249; Part II: *8*, 301–357; Part III: *9*, 43–80.

[31] L. A. Zadeh: Fuzzy logic = Computing with words. IEEE Trans. Fuzzy Systems *2* (1977), 103–111.

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

[33] L. A. Zadeh: Yes, no and relatively. Chemtech (1987), June: 340–344; July: 406–410.

[34] L. A. Zadeh: From computing with numbers to computing with words – from manipulation of measurements to manipulation of perception. IEEE Trans. Circuits Systems *45* (1999), 105–119.

[35] L. A. Zadeh: The concept of cointensive precisation – A key to mechanization of natural language understanding. In: Proc. IPMU, Paris 2006, pp. 13–15.

Milan Mareš, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
   e-mail:

Radko Mesiar, Department of Mathematics and Descriptive Geometry, SvF, Slovak University of Technology, Radlinského 11, 813 68 Bratislava, Slovak Republic and Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic.
   e-mail: mesiar@math.sk