KYBERNETIKA — VOLUME 47 (2011), NUMBER 1, PAGES 144–163

INFORMATION MEASURES AND UNCERTAINTY OF PARTICULAR SYMBOLS

Milan Mareš

The measurement of information emitted by sources with uncertainty of random type is known and investigated in many works. This paper aims to contribute to analogous treatment of information connected with messages from other uncertain sources, influenced by not only random but also some other types of uncertainty, namely with imprecision and vagueness. The main sections are devoted to the characterization and quantitative representation of such uncertainties and measures of information produced by sources of the considered type.

Keywords: information measure, uncertainty, randomness, vagueness, imprecision, information source, alphabet, message

Classification: 94A20, 94A15, 94D05, 28E10

1. INTRODUCTION

The common intuitive feeling ranks information and knowledge among phenomena too subtle to be measured and quantitatively processed. In spite of it, there exists a relatively long tradition of successful attempts to compare the informational values of symbols, messages or even their entire sources. One of the first information measures was suggested by R. A. Fisher in [7], others were published in the seminal works of rising cybernetics in [27, 28], and the most successful attempt was done by Shannon and Weaver in [26]. Their model of information source and information transmission via a communication channel has established a pattern, further developed by numerous authors (just for illustration, see [6, 29]). It is, usually quite explicitly, observable even in the modern construction of the measures of fuzziness (see, e. g., [1, 2, 4, 10, 15, 20, 25, 30], and less explicitly, also in [13, 19] or even in [32]).

The information is, at least in the frame of information theory, inseparable of uncertainty and chaos as their complementary concept. The Shannon model of information is connected with the randomness and probability theory. This approach to the uncertainty is especially effective as the stability of random events allows to apply the statistical estimations of probabilities. A very significant attribute of the Shannon probabilistic model of information is the additivity of probabilistic measure as a set function.

The pattern of the probabilistic model of the information has proved to be so effective that it motivated significant generalizations and modifications. Let us recollect here, at least, the contribution of Jean-Maria Kampé de Fériet, e. g. his works [11] and[12].

Most of the later alternative models of information measures are focused on the vagueness as further specific type of uncertainty, and its mathematical models deal with fuzzy sets and further fuzzy objects derived from them (for fuzziness and its derivatives see, e. g., [5, 16, 17, 31]). The main general differences between probabilistic tools and fuzzy set functions are discussed in [14], and partly in [21], including the properties of fuzzy set functions, among which the most typical is their monotonicity substituting the additivity of probabilistic measures. One of the main goals of this paper is to suggest a fuzzy information concept respecting that monotonicity.

The attempts to the fuzzy information published in the referred literature (especially in [2, 4, 10, 15]) and other papers is effective and offers valuable results, but it includes some elements in which the additive view on the uncertainty prevails the monotonicity. Moreover, the referred papers are focused on the fuzzy entropy as a measure of vagueness of a fuzzy set, and they do not deal with the information theoretical concepts like symbol, message, information included in a symbol, and some others. From such point of view, the information acquisition is not a process of the acceptance of new knowledge symbol by symbol, but some fixed attribute of a fuzzy set. The suggestion of a fuzzy set theoretical model of the information source represents further of the main goals of this paper.

The following sections are devoted to the suggestion and a brief discussion of an information source contaminated by uncertainty. This uncertainty is defined as general as possible to include its important specific types like randomness, imprecision or vagueness. Let us note that it cannot be directly used for the description of uncertainty of the type of granulation formally described by rough sets (cf. [24]).

The following chapters are organized as follows. The next Chapter 2 is devoted to the elementary model of information source and its uncertainty, including three special cases. Chapter 3 deals with the information connected with particular symbols from the alphabet of the source, and with the information included in the finite sequences of symbols – the messages. Brief Chapter 4 offers some interpretation of the model.

The motivation of the suggested model is connected with non-traditional applications of the information theoretical models and concepts. There exist situations, connected rather with vagueness than with randomness which can appear, e. g., in the model of human or social choice, decision-making, strategic or cooperative collisions of interest, etc. (see, e. g., [8, 9, 18, 23]) in which the uncertainty and information play essential role, but the Shannon's model of information transmission is not adequate to them. The models suggested here are intended to offer some elementary concepts and results as tools for the processing of uncertain information even in the mentioned sort of models.

2. HEURISTIC PRINCIPLES OF INFORMATION

In the seminal work by Shannon and Weaver [26], the measurement of information was based on the phenomenon of randomness with probability as the main tool for its formal processing.

The theory of fuzzy sets and fuzzy phenomena has pointed at the fact that there are other sources of uncertainty than only the randomness, and that they deserve their own information measures, starting from the analysis of the information offered by them. The papers dealing with this problem, like [2, 4, 15, 20, 30], focus their attention on more advanced topic, namely on the analysis of the entropy – like complex characteristics of the information source as a closed unit.

Hence, the aim of this paper is to suggest and describe, analogously to the Shannon's probabilistic model, the measure of information transmitted by a single element of a message produced by the source contaminated by uncertainty of various type. We turn our attention on three such sources – we briefly recollect the concept suggested by Shannon for sources with *random uncertainty* formally described by probabilistic tools, then we suggest the information measure for sources with *vague uncertainty*, described by fuzzy set theoretical concepts and, finally, we suggest the information measure for the sources with *crisp imprecise uncertainty* represented by crisp intervals.

We formulate, first of all heuristically, the basic principles that would be respected by any of, in principle so different, information measures. They can be summarized in the following verbal postulated.

(A) The value of information transmitted by a single symbol depends exclusively on the uncertainty connected with the appearance of the symbol.

(B) The information value increases with the decreasing degree of uncertainty.

(C) The information measure is cummulative, i. e., the total information transmitted by two symbols cannot be smaller than any of the individual informations values of those two symbols.

(D) The information measure is non-negative.

(E) The information transmitted by deterministically sure symbol vanishes.

3. FORMAL MODEL

First of all, we introduce a few symbols used in the remaining parts of this paper. Let M be a non-empty set. Then

$$2^M \text{ is the class of all (crisp) subsets of } M, \tag{1}$$

$$\mathcal{P}(M) \text{ is the class of all probability distributions over } M, \tag{2}$$

$$\mathcal{F}(M) \text{ is the class of all fuzzy subsets of } M. \tag{3}$$

3.1. Information source with uncertainty

Let us consider a non-empty and discrete set A, called an *alphabet*. Its elements $a, b, c, \ldots \in A$ are called *symbols*, and a sequence (finite or infinite) of symbols is called a *message*. By A^*, where

$$A^* = A \cup (A \times A) \cup (A \times A \times A) \cup \ldots, \tag{4}$$

we denote the class of all possible finite messages.

Each symbol is connected with some uncertainty regarding its frequency in messages, the exactness of its meaning, its precision or its expectedness. It means that there exist several formal representations of particular types of uncertainty. Many of them can have the theoretical background characterized in [14] by means of the apparatus of the set functions theory.

In general, let us consider a mapping $u : A \to R$, such that $u(a) \geq 0$ for all $a \in A$, and called the *uncertainty measure*. Then we call the pair

$$(A, u) \tag{5}$$

the *source* of uncertain information.

Let us extend the uncertainty measure u on the entire class A^* and define the *extended* uncertainty measure $u^* : A^* \to R$, such that for any $n = 1, 2, \ldots$, $a^* = (a_1, a_2, \ldots, a_n) \in A^n \subset A^*$

$$u^*(a^*) \geq 0, \tag{6}$$
$$\text{if } \ a^* = (a),\ a \in A \ \text{ then } \ u^*(a^*) = u(a), \tag{7}$$
$$u^*(a^*) \leq \min\left(u(a_1), \ldots, u(a_n)\right). \tag{8}$$

The previous conditions characterize the general uncertainty measures. The next condition is not necessary but it simplifies eventual interpretation of the source concept, when being senseful. In the case of one signal, it is trivial.

$$\begin{array}{l}\text{If } a^* = (a_1, \ldots, a_n) \in A^n,\ b^* = (b_1, \ldots, b_n) \in A^n \ \text{ and } \ u^*(a_i) \geq u^*(b_i) \\ \text{for all } \ i = 1, 2, \ldots, n, \ \text{ and some } \ n = 1, 2, \ldots \ \text{ then} \\ \qquad\qquad u^*(a^*) \geq u^*(b^*).\end{array} \tag{9}$$

Remark 1. Let a^*, b^*, $c^* \in A^*$, $a^* = (a_1, \ldots, a_m) \in A^m$, $b^* = (b_1, \ldots, b_n) \in A^n$, and let $c^* = (c_1, \ldots, c_{m+n}) \in A^{m+n}$ be such that

$$c_i = a_i \text{ for } i = 1, \ldots, m, \qquad c_i = b_{i+m} \text{ for } i = m+1, \ldots, m+n.$$

Then (8) immediately implies that

$$u^*(c^*) \leq \min\left(u(a^*),\, u^*(b^*)\right).$$

The general concepts presented above can be illustrated by the following more specific examples.

3.1.a Probabilistic information source

In this case, A is a general alphabet and we suppose that it is finite in order to simplify the notations.

The uncertainty measure $u_P : A \to [0,1]$ is a probability distribution $u_P \in \mathcal{P}(A)$. Such information sources are investigated in the classical papers on information theory starting by [26] and in many fundamental works like [6, 29]. Also the Fisher's statistical information concept [7] is developed on this ground. Let us briefly recollect that in this model

$$0 \leq u_P(a) \leq 1 \quad \text{for all } a \in A, \quad \sum\nolimits_{a \in A} u_P(a) = 1.$$

The probability distribution u_P can be extended on the class of all finite sequences of symbols A^* by means of the conditional probabilities. Let $a_1, a_2, \dots, a_n$ be symbols from A, and let us for every m-tuple, $m = 1, 2, \dots, n-1$, $u(a_m | a_1, \dots, a_{m-1})$ be the conditional probability of a_m under the condition that the ordered m-tuple $a_1, a_2, \dots, a_{m-1}, a_m$ of symbols was emitted. Then the extended probability distribution u_P^* over A^* is defined for any $a^* = (a_1, \dots, a_n) \in A^n \subset A^*$ by

$$u_P^*(a^*) = u_P(a_1) \cdot u_P(a_2 \,|\, a_1) \cdot \dots \cdot u_P(a_n \,|\, a_1, \dots, a_{n-1}), \tag{10}$$

where the previous notation was preserved. If the symbols in A are independent then the conditional probabilities $u_P(a_m | a_1, \dots, a_{m-1})$ are equal to $u_P(a_m)$ for any $(a_1, \dots, a_{m-1}, a_m)$, and (10) turns into

$$u_P^*(a^*) = u_P(a_1) \cdot u_P(a_2) \cdot \dots \cdot u_P(a_n). \tag{11}$$

Lemma 1. Probabilistic information source (A^*, u_P^*) fulfils conditions (6), (7), (8), and it fulfils (9) if (11) is fulfilled.

Proof. The statement follows from (10), eventually (11), and from the fact that $0 \leq u_P^*(a) \leq 1$, immediately. □

3.1.b Interval represented imprecision

Here, we consider the uncertain information generated by imprecise measuring of some physical, technical or generally natural quantity. Such measuring, however careful it is, cannot be absolutely accurate, and the unknown measured quantity $\alpha \in R$ is approximated by a (closed) real interval $a = [x_a, y_a] \subset R$. This interval represents the information generated by means of one measurement. If one unknown quantity α is measured repeatedly or by different methods, the sequence of the results can be considered for a message mediating information about the value of the quantity.

More formally. If $\alpha \in R$ is the measured quantity, then each measurement results into an interval $[x^{(\alpha)}, y^{(\alpha)}]$. In this sense, the alphabet A_α is defined as the set

$$A_\alpha = \{[x, y] \subset R : x < y,\ \alpha \in [x, y]\}. \tag{12}$$

If the measurement of the quantity $\alpha \in R$ are n-times repeated (by different observers, different devices or under different conditions) then the sequence of imprecise results

$$([x_1, y_1],\ [x_2, y_2], \ldots, [x_n, y_n]) \in A^n \subset A^*$$

forms a message $[x, y]^* \in A^*$ informing about the value of quantity α.

Remark 2. Evidently, $u_\alpha \in [x_i, y_i]$ for all $i = 1, 2, \ldots, n$, as follows from (12).

The uncertainty measure $u_\alpha : A \to R$ is defined as a difference

$$u_\alpha([x, y]) = y - x, \tag{13}$$

and its extension on A^* is defined by means of an intersection

$$\left[x^{(n)}, y^{(n)}\right] = [x_1, y_1] \cap [x_2, y_2] \cap \cdots \cap [x_n, y_n],$$

as the difference

$$u^*_\alpha\left([x^{(n)}, y^{(n)}]\right) = y^{(n)} - x^{(n)}. \tag{14}$$

Lemma 2. Preserving the previous notations,

$$u^*_\alpha\left([x^{(n)}, y^{(n)}]\right) = \min\left(y_1, y_2, \ldots, y_n\right) - \max\left(x_1, x_2, \ldots, x_n\right).$$

P r o o f. Due to Remark 2, for all $i = 1, 2, \ldots, n$, $y_i \geq \alpha$ and $x_i \leq \alpha$, and

$$x^{(n)} \leq \alpha, \quad y^{(n)} \geq \alpha, \quad x^{(n)} \geq x_i, \quad y^{(n)} \leq y_i \quad \text{for all } i = 1, \ldots, n.$$

The statement follows from these inequalities, immediately. □

Lemma 3. Uncertainty measure u^*_α fulfils conditions $(6), (7), (8)$.

P r o o f. Property (6) follows from (13), properties (7) and (8) follow from Lemma 2, immediately. □

Condition (9) is not generally fulfilled but it is valid in a rather weakened form.

Lemma 4. Let $[x^{[n]}, y^{[n]}] = ([x_1, y_1], \ldots, [x_n, y_n]) \in A^*_\alpha$, $[s^{[n]}, t^{[n]}] = ([s_1, t_1], \ldots$ $\ldots, [s_n, t_n]) \in A^*_\alpha$, and let for each $i = 1, 2, \ldots, n$, $[x_i, y_i] \subset [s_i, t_i]$. Then

$$u^*_\alpha\left([x^{[n]}, y^{[n]}]\right) \leq u^*_\alpha\left([s^{[n]}, t^{[n]}]\right).$$

P r o o f. Under the assumptions of this lemma, $x^{[n]} \geq s^{[n]}$ and $y^{[n]} \leq t^{[n]}$ if the notation used in (14) is preserved. Then (14) implies the inequality. □

3.1.c Fuzzy information sources

The last example of uncertain information source will be the one in which the generated information is vague. It means that the emitted signals are well (or relatively well) identified but their interpretation, the real content of the data represented by them, is deformed by subjectivity of imprecise understanding (cf. [16, 17, 19, 31, 32], e. g.). Such sources are, in fact, intensively investigated in many papers like [1, 2, 4, 13, 15, 30] where the main attention of this investigations is focused on the entire sources and their entropy-like characteristics.

The alphabet A of a fuzzy information source is a general alphabet. The uncertainty measure u_F is a fuzzy subset of A, i. e., $u_F \in \mathcal{F}(A)$, and we use the symbol u_F for its membership function, as well. Hence $u_F(a)$, for an $a \in A$, denotes the possibility with which the symbol a will be (or was) emitted by the source.

If $a^* = (a_1, \dots, a_n) \in A^n \subset A^*$, then we define the value u_F by means of

$$u_F^*(a^*) = \min\left(u_F(a_1), \dots, u_F(a_n)\right). \tag{15}$$

Remark 3. Obviously, function u_F displays the properties of membership function, i. e., it identifies a fuzzy subset of A^n.

Even in this case the general properties of the uncertainty measures are fulfilled.

Lemma 5. Fuzzy information source (A^*, u_F^*) fulfils properties (6), (7), (8), (9).

Proof. The validity of (6) follows from the definitoric properties, immediately, (7), (8) and (9) are immediate consequences of (15). □

3.2. General properties of uncertainty measures

The presentation of three illustrative examples motivates a few general comments and formal results. The main comment regards the choice of the above examples. Two of them, the probabilistic and the fuzzy information sources represent the uncertainty which is described by some distribution of its measures over the alphabet as a basic universum. The alphabet itself need not be limited by any essential formal assumptions, even if its finiteness simplifies the formalism of the model. In both cases the uncertainty measure represents something what can be called “density of uncertainty” and related to some non-negative set functions (generalized measures, by [14]). One of them is additive, another one is monotonous. The remaining example is focused on the uncertain information source with a strictly specified alphabet (each unknown measured quantity α has its own alphabet), and particular signals, themselves are subsets of the set of possible results of measurements. These imprecise measurements mean a different point of view on the uncertainty (cf., [21], too), and evidently extend the scale of abstract information sources.

The following statements formalize the consequences of combinations of several uncertainty measures. We consider a general non-empty and finite alphabet A.

Lemma 6. Let (A, u_1), (A, u_2) be uncertain information sources, let u_1^*, u_2^* be their extensions fulfilling (6), (7), (8), and let (A, u) be an information source. Let u^* be a mapping, $u^* : A^* \to R$. Then the following statements are true if their assumptions are fulfilled for any $a \in A$, $a^* \in A^*$.

(a) If $u(a) = u_1(a) + u_2(a)$, $u^*(a^*) = u_1^*(a^*) + u_2^*(a^*)$ then u^* fulfils (6), (7), (8).

(b) If $r \in R$, $r > 0$, and $u(a) = r \cdot u_1(a)$, $u^*(a^*) = r \cdot u_1^*(a^*)$ then u^* fulfils (6), (7), (8).

(c) If $u(a) = u_1(a) \cdot u_2(a)$, $u^*(a^*) = u_1^*(a^*) \cdot u_2^*(a^*)$, then u^* fulfils (6), (7), (8).

(d) If $u(a) = \min(u_1(a), u_2(a))$, $u^*(a^*) = \min(u_1^*(a^*), u_2^*(a^*))$, then u^* fulfils (6), (7), (8).

P r o o f. The validity of (6) for u^* follows from its validity for u_1^* and u_2^* and from the operations defining u^* in (a), (b), (c), (d), immediately. Analogously, (7) is an immediate consequence of the assumptions of this lemma and of its validity for u_1^* and u_2^*. Let us turn our attention to (8).

If (a) is fulfilled, then

$$\begin{aligned} u^*(a^*) &= u_1^*(a^*) + u_2^*(a^*) \\ &\le \min(u_1(a_1), \ldots, u_1(a_n)) + \min(u_2(a_1), \ldots, u_2(a_n)) \\ &\le \min(u_1(a_1) + u_2(a_1), \ldots, u_1(a_n) + u_2(a_n)) \\ &= \min(u(a_1), \ldots, u(a_n)) . \end{aligned}$$

Analogously, if assumptions of (b) are fulfilled then

$$\begin{aligned} u^*(a^*) &= r \cdot u_1^*(a^*) \ \le \ r \cdot \min(u_1(a_1), \ldots, u_1(a_n)) \\ &= r \cdot \min(r \cdot u_1(a_1), \ldots, r \cdot u_1(a_n)) \ = \ \min(u(a_1), \ldots, u(a_n)) . \end{aligned}$$

If assumptions of (c) are fulfilled then

$$\begin{aligned} u^*(a^*) &= u_1^*(a^*) \cdot u_2^*(a^*) \ \le \ \min(u_1(a_1), \ldots, u_1(a_n)) \cdot \min(u_2(a_1), \ldots, u_2(a_n)) \\ &\le \min(u_1(a_1) \cdot u_2(a_1), \ldots, u_1(a_n) \cdot u_2(a_n)) \ = \ \min(u(a_1), \ldots, u(a_n)) . \end{aligned}$$

And if assumptions of (d) are fulfilled then

$$\begin{aligned} u^*(a^*) &= \min(u_1^*(a^*), u_2^*(a^*)) \\ &\le \min\big(\min\big(u_1(a_1), \ldots, u_1(a_n)\big), \min\big(u_2(a_1), \ldots, u_2(a_n)\big)\big) \\ &\le \min\big(\min\big(u_1(a_1), u_2(a_1)\big), \ldots, \min\big(u_1(a_n), u_2(a_n)\big)\big) \\ &= \min(u(a_1), \ldots, u(a_n)) . \end{aligned}$$

$\square$

3.3. Information measure for sources with uncertainty

Having introduced the concept of the source of signals contaminated by uncertainty, we can suggest the measure of information mediated by such signals. Information contained in each particular symbol appears to be the elementary starting concept for the total characteristics of the entire source and its structure. The most famous of such total characteristics is the information entropy. For the probabilistic sources, it was suggested and analyzed by Shannon and Weaver (cf. [26] where its introduction starts from the information mediated by particular signals).

Its analogy for fuzzy information sources was analyzed in [2, 4, 15, 20, 29, 30] where the information of the signal is treated rather as an implicite concept, and the main attention is focused on integrating entropy-like structures.

Let (A, u) be an uncertain information source with alphabet A and uncertainty measure u. Let A^* be the set of finite messages and u^* be the extension of u on A^*, as formulated in (4) and in properties (6), (7), (8).

If $I : A^* \to R$ is a mapping such that

$$I(a^*) \geq 0, \tag{16}$$

$$\text{if } a^*,\, b^* \in A^*,\ u^*(a^*) \geq u^*(b^*), \text{ then } I(a^*) \leq I(b^*), \tag{17}$$

then we say that I is an *information measure* on (A, u).

Remark 4. If $a^* = (a_1, \dots, a_n) \in A^n$, and if $b^* = (a_1, \dots, a_n,\, a_{n+1}) \in A^{n+1}$ then (8) and Remark 1 imply that $I(a^*) \leq I(b^*)$.

Remark 5. Keeping notations of (9), if (9) is fulfilled then $I(a^*) \leq I(b^*)$.

Let us test the adequacy of the suggested concept of information measure to our intuitive (or also traditional) expectations regarding the three examples of sources analyzed in Section 3.1.

3.3.a Probabilistic information source

The information measure for probabilistic source (A, u_P), where $u_P \in \mathcal{P}(A)$, was suggested in [26], and its properties are well known. Let us recollect that

$$I_P(a) = -\log u(a) \quad \text{for } a \in A, \tag{18}$$

where, usually, the logarithmic function is supposed to be the binary one, $\log_2$. Using (10) and (11), it is easy to see that for any $a^* = (a_1, \dots, a_n)$

$$I_P(a^*) = I_P(a_1) + I_P(a_2 \,|\, a_1) + \cdots + I_P(a_n \,|\, a_1, \dots, a_{n-1}) \tag{19}$$

or, in the case of independence,

$$I_P(a^*) = I_P(a_1) + I_P(a_2) + \cdots + I_P(a_n). \tag{20}$$

The definition of the uncertainty measure u_P as a probability distribution over A, immediately implies the following statements.

Remark 6. Conditions (16) and (17) are fulfilled for the Shannon information measure (18) and its extension (19), as follows from Lemma 1, namely from (6) and (7).

Remark 7. If $u_P(a) = 1$ for some $a \in A$ then $I_P(a) = 0$, and a does not represent any information.

The previous two simple statements illustrate the fact, well known from [26] and other works developing its approach to the information. Namely, the classical Shannon information measure is a natural and adequate representation of information connected with particular signals of probabilistic information source.

Remark 8. Evidently, if the probability $u_P(a)$ for some $a \in A$ vanishes then the respective information value $I_P(a)$ groves to infinity. The same does the information mediated by any message $a^* \in A^*$ including a as one of its components.

3.3.b Interval represented imprecision

The result of imprecise measurement of a real-valued quantity $\alpha \in R$, as described in Subsection 3.1.b, offers some information about its value. In our case, the result of a measurement is an interval $[x, y]$ about which we assume that $\alpha \in [x, y]$, and its uncertainty is measured by its length, $u_\alpha([x, y]) = y - x$.

Then, it is quite natural to define the information obtained by such single measurement by

$$I_\alpha([x, y]) = u_\alpha([x, y])^{-1} = 1/(y - x). \tag{21}$$

Using (14) we define for $\alpha^* = ([x_1, y_1], \ldots, [x_n, y_n])$

$$I_\alpha(a^*) = u^*_\alpha\left([x_1, y_1] \cap \cdots \cap [x_n, y_n]\right). \tag{22}$$

The following statement is an immediate consequence of (21) and Lemma 3.

Remark 9. Note that, due to (12) $y > x$. This inequality need not be the truth for $u^*_\alpha(a^*)$ defined by (14). If $[x_1, y_1] \cap \cdots \cap [x_n, y_n] = \{\alpha\}$ then $I_\alpha(a^*)$ increases to infinity.

Lemma 7. Information measure I_α defined by (21)and (22) fulfills conditions (16) and (17).

Remark 10. Let $[x, y] \in A_\alpha$, $[s, t] \in A_\alpha$, and let $[x, y] \subset [s, t]$. Then, evidently, $I_\alpha([x, y]) \geq I_\alpha([s, t])$.

The previous remark can be extended in the sense of Lemma 4.

Lemma 8. Let $([x_i, y_i])_{i=1,\dots,n}$, $([s_i, t_i])_{i=1,\dots,n}$ be messages from A_α^*, for some n. Let $[x_i, y_i] \subset [s_i, t_i]$ for all $i = 1, 2, \dots, n$. Then

$$I_\alpha\left([x_i, y_i]_{i=1,\dots,n}\right) \geq I_\alpha\left([s_i, t_i]_{i=1,\dots,n}\right).$$

Proof. The statement follows from Remark 9, Lemma 4 and from (17), immediately. □

The repetitive measurements of one quantity $\alpha \in R$, and their aggregation by means of intersection (cf. (14)) effectively increases the information about the unknown real value of α.

Lemma 9. Let us denote by

$$[x_1, y_1],\ [x_2, y_2], \dots, [x_n, y_n], \dots \tag{23}$$

a sequence of intervals from A_α, and by

$$\left[x^{(1)}, y^{(1)}\right],\ \left[x^{(2)}, y^{(2)}\right], \dots, \left[x^{(n)}, y^{(n)}\right], \dots \tag{24}$$

the sequence of partial intersections,

$$\left[x^{(1)}, y^{(1)}\right] = [x_1, y_1], \quad \left[x^{(k)}, y^{(k)}\right] = \left[x^{(k-1)}, y^{(k-1)}\right] \cap [x_k, y_k] \text{ for } k = 1, 2, \dots$$

Then

(a) The sequence of information values $(I_\alpha([x^{(k)}, y^{(k)}]))_{k=1,2,\dots}$ is not decreasing.

(b) If the limit of sequence (24) is the one-element set $\{\alpha\}$ then the sequence $(I_\alpha([x^{(k)}, y^{(k)}]))_{k=1,2,\dots}$ is increasing to infinity.

(c) If we, vice versa, denote by $([X^{(j)}, Y^{(j)}])_{j=1,2,\dots}$ the sequence of partial unions $\left[X^{(1)}, Y^{(1)}\right] = [x_1, y_1]$, $\left[X^{(j)}, Y^{(j)}\right] = \left[X^{(j-1)}, Y^{(j-1)}\right] \cup (x_j, y_j)$ for $j = 1, 2, \dots$, then the sequence of information values

$$\left(I_\alpha([X^{(j)}, Y^{(j)}])\right)_{j=1,2,\dots} \tag{25}$$

is not increasing and it has a non-negative limit.

Proof. The statements easily follow from assumptions. Evidently $\alpha \in [x_i, y_i] \cap [x^{(i)}, y^{(i)}] \cap [X^{(i)}, Y^{(i)}]$ for all $i = 1, 2, \dots$ Hence, all considered sets are non-empty closed intervals or one-element sets. It means that definition (13) can be used for them, as well, and due to (21) $I_\alpha([x, y]) > 0$ for $x \neq y$, and statement (a) is proven. Definitions (21) and (24) mean that $I_\alpha([(x^{(n)}), y^{(n)}]) \to \infty$ if

$$\lim_{n\to\infty} x^{(n)} = \lim_{n\to\infty} y^{(n)} = \alpha,$$

hence statement (b) is obviously true. Finally, (25) is an increasing sequence of closed intervals containing α. Due to (21), sequence $I_\alpha([X^{(n)}, Y^{(n)}])$ for $n \to \infty$, is decreasing and limited by 0 from below (cf. (a)). This implies statement (c). □

3.3.c Fuzzy information sources

Let us consider, now, the fuzzy information source (A, u_F) defined in Subsection 3.1.c, where u_F is a membership function of a fuzzy subset of the alphabet A, and its extension u_F^* on A^* is introduced by (15). Such fuzzy information sources are carefully analyzed by a wide class of works, some of which are referred here, as well. These works deal with a total view on fuzzy sources as compact objects, and the analysis of informational content of particular symbols (or its measure) does not represent the essential object of attention.

Nevertheless, the papers mentioned above, deal with some implicite concept of the information of single symbols. Namely, the fuzzy entropy dealt by them, is a very close analogy of the probabilistic source entropy suggested in [26]. The Shannon entropy H_P is defined as a mean value of probabilistic informations $I_P(a)$ for $a \in A$, i. e.

$$H_P(A, u_P) = \sum_{a\in A} p(a) \cdot I_P(a) = -\sum_{a\in A} p(a) \cdot \log_2 p(a), \tag{26}$$

where $-\log_2 p_a$ is the information transmitted by the symbol a (cf. (18) in Subsection 3.3.a). Analogously to this probabilistic entropy, its fuzzy counterpart is usually defined as a value formally similar to the mean value,

$$H_F(A, u_F) = -\sum_{a\in A} u_F(a) \cdot \log_2 u_F(a). \tag{27}$$

Formula (27) implies the conclusion that its authors consider

$$I_F(a) = -\log_2 u_F(a) \tag{28}$$

for the (implicitly introduced) measure of information contained by symbol a of the information source (A, u_F).

The above discussion shows that we may consider value (28) for a correct definition of *fuzzy information* of symbol a. This fuzzy information has the properties demanded for an information measure and it can be extended on the set A^* in several ways. Here, respecting the analogy with the probabilistic case, we may define, for $a^* = (a_1, a_2, \ldots, a_n) \in A^n$,

$$\begin{aligned} I_F(a^*) &= I_F(a_1) + I_F(a_2) + \cdots + I_F(a_n) \\ &= -\log_2 \left(u_F(a_1) \cdot u_F(a_2) \cdot \cdots \cdot u_F(a_n)\right). \end{aligned} \tag{29}$$

Lemma 10. The fuzzy information measure I_F defined by (28) and (29) fulfils conditions (6) and (7).

P r o o f. The statement follows from the elementary properties of logarithmic function, immediately. □

Remark 11. If $u_F(a) = 1$ for some $a \in A$ then $I_F(a) = 0$.

Lemma 11. If $u_F^*(a^*) = 1$, where u_F^* is defined by (15) $I_F(a^*)$ is defined by (29) then

$$I_F(a^*) = 0.$$

Proof. If $a^* = (a_1, a_2, \ldots, a_n) \in A^n \subset A^*$, if $u_F^*(a^*) = 1$ and u_F^* is defined by (15) then obviously $u_F(a_i) = 1$ for all $i = 1, 2, \ldots, n$. Hence $I_F(a_i) = 0$ as follows from Remark 10 and, consequently, $I_F(a^*) = 0$ as follows from (29). □

Remark 12. If, generally, $I_F(a) = -\log_2 u_F(a)$ and if $a_1, a_2, \ldots, a_n, \ldots$ is a sequence of signals such that

$$\lim_{i\to\infty} u_F(a_i) \to 0$$

then

$$\lim_{i\to\infty} I_F(a_i) = \infty,$$

as follows from (28) if we define $I_F(a) = -\log_2 u_F(a)$.

3.4. Transformations of information measure

Let us consider an alphabet A, the set of messages A^*, and an m-tuple of uncertainty measures $u^{(1)}, u^{(2)}, \ldots, u^{(m)}$. In other words, we consider an m-tuple of information sources $(A, u^{(1)})$, $(A, u^{(2)}), \ldots, (A, u^{(m)})$, where $m = 1, 2, \ldots$ The equality between some uncertainty measures is not excluded, hence, the relation $u^{(i)}(a) = u^{(j)}(a)$ for all $a \in A$ and some i, $j \in \{1, 2, \ldots, m\}$ is admissible.

Let us consider an m-tuple of information measures $I^{(1)}, I^{(2)}, \ldots, I^{(m)}$ of the sources $(A, u^{(1)}), \ldots, (A, u^{(m)})$, respectively, too. The correspondence of sources need not be one-to-one in the sense that even if $u^{(i)}(a) = u^{(j)}(a)$ for all $a \in A$ and all $a^* \in A^*$, the inequality $I^{(i)}(a^*) \neq I^{(j)}(a^*)$ for some $a^* \in A^*$ is admissible, as well.

The topic to be answered in this subsection regards the conditions under which some combination of information measures over one alphabet A, with properties formulated in preceding paragraphs, preserves the general properties of information measures, i. e., conditions (16) and (17).

Theorem 1. Let A be an alphabet and $(A, u^{(1)}), (A, u^{(2)}), \ldots, (A, u^{(m)})$ be information sources with information measures $I^{(1)}, I^{(2)}, \ldots, I^{(m)}$, respectively, for $m = 1, 2, \ldots$. Let $f : R^m \to R$ be a real-valued function of m variables. Let

$$f(0, 0, \ldots, 0) = 0 \tag{30}$$

f is non-decreasing in all variables, i. e., if $x = (x_1, \ldots, x_m) \in R^m$, (31)
$y = (y_1, \ldots, y_m) \in R^m$ and $y_i \geq x_i$ for all $i \in \{1, 2, \ldots, m\}$, then $f(y) \geq f(x)$.

Then the mapping $u^* : A^* \to R$ such that for any $a^* \in A^*$

$$u^*(a^*) = f\left(u^{*(1)}(a^*), \ldots, u^{*(m)}(a^*)\right)$$

is an uncertainty measure over the alphabet A, and the mapping $I : A^* \to R$ such that for any $a^* \in A^*$

$$I(a^*) = f\left(I^{(1)}(a^*), \ldots, I^{(m)}(a^*)\right)$$

is an information measure of the source (A^*, u^*).

P r o o f. To prove the first statement, it is necessary to verify the validity of (6), (7), (8) for $u^*(a^*)$ defined by means of f. The validity of (6) for all $u^{*(i)}$, $i = 1, 2, \ldots, m$, in combination with (30) means that $u^*(0) = 0$, and in combination with (31) it means that $u^*(a^*) \geq 0$ for all $a^* \in A^*$. The validity of (7) for any $a^* = (a) \in A \subset A^*$ follows from its validity for all $u^{*(i)}$, immediately. Finally, (8) is fulfilled for all $u^{*(i)}$, $i = 1, 2, \ldots, m$, and hence, by (31)

$$\begin{aligned}
u^*(a^*) &= f\left(u^{*(1)}(a^*), \ldots, u^{*(m)}(a^*)\right) \\
&\leq f\left(\min(u^{*(1)}(a_1), \ldots, u^{*(1)}(a_n)), \ldots, \min(u^{*(m)}(a_1), \ldots, u^{*(m)}(a_n))\right) \\
&= \min\left(f(u^{*(m)}(a_1), \ldots, u^{*(m)}(a_n))), \ldots, f(u^{*(m)}(a_1, \ldots, u^{*(m)}(a_n)))\right) \\
&= \min\left(u(a_1), \ldots, u(a_n)\right).
\end{aligned}$$

The proof of the second statement is based on the verification of validity of conditions (16) and (17).

If for some $a^* \in A^*$, $I(a^*) = f(I^{(1)}(a^*), \ldots, I^{(n)}(a^*))$ then by (30) and (31) $I(a^*) = 0$ if $I^{(j)}(a^*) = 0$ for all $j = 1, 2, \ldots, m$, and $I(a^*) \geq I^{(j)}(a^*) = 0$ if $I^{(j)}(a^*) > 0$ for some j. Finally, if for some a^*, $b^* \in A^*$ and for some $j \in \{1, 2, \ldots, m\}$ $u^{*(j)}(a^*) \geq u^{*(j)}(b^*)$ then, by (7), $I^{(j)}(a^*) \leq I^{(j)}(b^*)$. If the above conclusion is true for all $j = 1, 2, \ldots, m$, then by (31), also

$$f\left(I^{(1)}(a^*), \ldots, I^{(m)}(a^*)\right) \leq f\left(I^{(1)}(b^*), \ldots, I^{(m)}(b^*)\right),$$

and, consequently, $I(a^*) \leq I(b^*)$ as follows from (31). □

Corollary. The above Theorem 1 immediately implies that: If (A, u) is an information source, if $I^{(1)}$ and $I^{(2)}$ are two information measures on (A, u), and $r > 0$ is a real number, then the mapping $I : A^* \to R$ such that for all $a^* \in A^*$

$$\begin{aligned}
I(a^*) &= r \cdot I^{(1)}(a^*), \quad \text{or} \\
I(a^*) &= I^{(1)}(a^*) + I^{(2)}(a^*), \quad \text{or} \\
I(a^*) &= I^{(1)}(a^*) \cdot I^{(2)}(a^*), \quad \text{or} \\
I(a^*) &= \max\left(I^{(1)}(a^*), I^{(2)}(a^*)\right), \quad \text{or} \\
I(a^*) &= \min\left(I^{(1)}(a^*), I^{(2)}(a^*)\right)
\end{aligned}$$

are information measures on (A^*, u^*).

Theorem 2. Let $a^* = (a_1, \dots, a_n) \in A^n \subset A^*$, where (A, u) is an information source. Let $I(a^*)$ and $I(a_1), \dots, I(a_n)$ be information measures of $a^*, a_1, \dots, a_n$, respectively. Then

$$I(a^*) \geq \max\left(I(a_1), \dots, I(a_2), \dots, I(a_n)\right).$$

P r o o f . The statement follows from (8) and (17). By (8)

$$u(a^*) \leq u(a_i) \quad \text{for all } i = 1, \dots, n.$$

Condition (17) means that

$$I(a^*) \geq I(a_i) \quad \text{for all } i = 1, \dots, n,$$

which proves the statement. □

The previous theorem shows that the information measure I defined in this section is cummulative, as demanded by postulate C of Section 2, and by condition (8).

4. MORE ABOUT FUZZY INFORMATION MEASURES

As mentioned in Subsection 3.3.c, there already exists a significant set of information sources (A, u_F), where A is an alphabet and u_F is a membership function of a fuzzy subset of A, with values characterizing the possibility with which some signal produced by the considered fuzzy source, can be interpreted as $a \in A$.

4.1. The probabilistic patterns in fuzzy information measure

To recollect the formal definitions of fuzzy information, we can refer Subsection 3.3.c, namely formulas (27) and (28). Especially (28), and its extension on the class A^*

$$I_F(a^*) = -\log_2 u_F(a^*), \tag{32}$$

which is a natural consequence (or rather premise) of (27). This approach to the measurement of vague information emitted by fuzzy source is formally correct, as well as the attention paid to it in the referred literature.

Nevertheless, there exist several features of that approach deserving some discussion. It is a very close paraphrase of the Shannon probabilistic model, in spite of the essential difference between the randomness and vagueness. The next special comments to the fuzzy information measure I_F treated in 3.3.c, and following from, e. g. [2, 4, 15, 20, 30], summarize its main contraversory components.

— The fuzzy uncertainty and fuzzy information are, in the referred theory, dealt as implicitly additive, mathematical objects. This additivity can be recognized in the fuzzy entropy (27) and it was reflected also in formula (29). It is motivated by the successful pattern of the probabilistic model, in spite of the fact that the fuzzy set theoretical concepts display rather monotonicity than additivity (see, e. g., [13, 14, 31, 32]).

— The logarithmic function used in (27) and, consequently, (18), is natural and unavoidable in the probabilistic model. Thanks to it, the multiplicative probabilities in (10) or (11), regarding also the probabilistic uncertainty measure u_P, are transformed into additive information measure $I_P : A^* \to R$. If we, in the case of fuzzy information measure I_F, do not insist on the additivity, the application of logarithms is correct but rather redundant.

— The binary logarithm usually (not always) used in the definition of probabilistic information measure is comfortable for practical handling the informational properties of sources with binary alphabet $A = \{0, 1\}$, nevertheless, is not necessary even in the case of fuzzy information sources.

Let us suggest and discuss some alternative models of fuzzy information motivated by the previous, rather heuristic, comments.

4.2. Alternative fuzzy information source

Let us consider an information source (A, u_F) with fuzzy uncertainty measure $u_F \in \mathcal{F}(A)$. It can be extended on $\mathcal{F}(A^*)$ by means of (15), i.e.,

$$u_F(a^*) = \min\left(u_F(a_1), \dots, u_F(a_n)\right)$$

for any $a^* = (a_1, \dots, a_n) \in A^n \subset A^*$.

Definitoric relation (15) represents, in fact, the first step to the alternative approach to fuzzy information, based on the paradigm of monotonicity of fuzzy measures and, generally, fuzzy operations. Let us define the *monotonous fuzzy information* $I_M : A^* \to R$ by means of

$$I_M(a^*) = 1 - u_F(a^*), \quad \text{for } a^* \in A^*. \tag{33}$$

Lemma 12. If $a^* = (a_1, \dots, a_n) \in A^n$ then

$$I_M(a^*) = \max\left(I_M(a_1), \dots, I_M(a_n)\right).$$

P r o o f. The statement follows from (15) and (33), as

$$\begin{aligned} I_M(a^*) &= 1 - u_F(a^*) \;=\; 1 - \min\left(u_F(a_1), \dots, u_F(a_n)\right) \\ &= \max\left(1 - u_F(a_1), \dots, 1 - u_F(a_n)\right) \;=\; \max\left(I_M(a_1), \dots, I_M(a_n)\right). \ \square \end{aligned}$$

Remark 13. $I_M(a^*) \in [0, 1]$ as follows from the assumption that $u_F \in \mathcal{F}(A)$, and from (15) and (33).

Theorem 3. The monotonous fuzzy information is an information measure fulfilling (16) and (17).

P r o o f. $I_M(a^*) \geq 0$ as follows from Remark 13. Condition (17) follows from (33), immediately. $\square$

4.3. Interpretations

The alternative concept of fuzzy information related to particular symbols and their finite sequences, suggested in this section, can be interpreted in the following way.

Meanwhile the classical probabilistic information can be interpreted as a consequence of randomness in the emission of symbols, the fuzzy information represents rather the vagueness connected with the phenomena of their acquisition and perception. There exist at least two types of situations in which the fuzzy approach to uncertain information can be effective – both of them are connected with subjective estimation of possibilities of symbols.

The first one of them represents an alternative to the (subjective or objective) probability of symbols produced by an uncertain source. The construction of a probability distribution is based on the knowledge of massive real data or on a multilateral analysis of personal preferences and attitudes. Both such procedures assume relative stability of input data and especially of the situation represented by them which can be partly substituted by fuzzy set theoretical tools.

The second situation in which the application of fuzzy information appears natural, regards the interpretation of already emitted and accepted but vaguely cognizable symbol or message. For example, written historical artefacts, heavily noised telecommunicated messages, remote sensing under complicated meteorological conditions, and similar events. The uncertainty is not generated by randomness, but rather by vagueness, and the approach characterized by (33) is not only formally correct but also adequate to the problem.

Anyhow – the monotonicity paradigma accepted by fuzzy set theoretical models and formally represented by the application of maxima and minima in processing fuzzy set theoretical models, is more adequate and natural for the construction of mathematical models including vague components. It regards the vague information sources and measurement of their uncertainty.

5. CONCLUSIVE REMARKS

The previous chapters were devoted to the description and analysis of information inherited in the uncertain phenomena, from the point of view of information-theoretical models. The attention was focused on the type of uncertainty which is called vagueness and formally represented by fuzzy sets. It is natural – the randomness, represented by the probabilistic concepts, is well known and deeply studied by the classical Shannon information theory [26], and the imprecision treated by the interval calculus is partly quite simple, partly can be transferred in the frame of fuzzy sets, too.

Nevertheless, there exist models and mathematical tools, usually connected with fuzzy sets but offering essential generalization and extension of their methods and concept. It regards, especially, triangular norms, aggregation operations and copulas, and some other related topics (see, e. g., [3, 5, 22] and many other works), where namely the theory of aggregation operators and closely related theory of copulas offer very promising effects. Especially the concept of aggregation operators appears very useful for application in the data and information processing in information sciences.

Finally, it is worth mentioning one field of study in which an effective handling of information and its measure can be significant. The information theory was originally developed for the analysis of information transmission under regular and relatively stable conditions with random noise and constant properties of the technical transmission channels. The probabilistic information theory offers optimal tools by means of which we are able to cope that problem.

But the uncertainty and information play a crucial role also in another type of human activity, namely in the decision-making and strategic behaviour (cf. [8, 9, 23] or also [18] and many other works). Here, the typical information and knowledge is vague, subjective and imprecise, its parameters are not stabilized, and its interpretation is often rather chaotic. All these properties practical exclude, or at least limit, the application of probabilistic information theoretical methods, and justify the use of alternative models of information.

ACKNOWLEDGEMENT

This work was partially supported by Grant Agency of Czech Republic, grant No. 402/08/0618, and by the Ministry of Education, Youth and Sports of the Czech Republic, namely via the Centre of Applied Research No. 1M0572, “DAR”, and via the Research Plan of Academy of Sciences ÚTIA, No. AVOZ 10750506.

The author wishes to express his deepest thanks to reviewers and editors of the Kybernetika journal for their helpful criticism and very valuable comments to the manuscript. This suggestions have significantly contributed to the final form of the paper.

(Received March 16, 2010)

REFERENCES

[1] P. Benvenuti: L’Opera Scientifica. Universitá La Sapienza, Roma 2004.

[2] P. Benvenuti, D. Vivona, and M. Divari: Order relations for fuzzy sets and entropy measure. In: New Trends in Fuzzy Systems (D. Mancini, M. Squillante, and A. Ventre, eds.). World Scientific 1998, pp. 224–232.

[3] T. Calvo, G. Mayor, and R. Mesiar (eds.): Aggregation Operators. Physica-Verlag, Heidelberg 2002.

[4] 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.

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

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

[7] R. A. Fisher: Statistical Methods for Research Workers. Olivier and Boyd, Edinburgh 1932.

[8] H. Gintis: Game Theory Evolving. Princeton Univ. Press, Princenton 2009.

[9] Ein-Ya Gura and M. B. Maschler: Insight Into Game Theory. Cabridge Univ. Press, Cambridge 2009.

[10] W.-L. Hung and M-S. Yang: Fuzzy entropy on intuicionistic fuzzy sets. Internat. J. Intelligent Systems *21* (2006), 443–451.

[11] J.-M. Kampé de Fériet: La théorie general de l'information et la mesure subjective de l'information. In: Lecture Notes in Math. *398*, Springer-Verlag, Heidelberg 1974, pp. 1–35.

[12] J.-M. Kampé de Fériet and B. Forte: Information et probabilité. C. R. Acad. Sci. Paris *265* (1967), 110–114, 142–146, 350–353.

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

[14] G. J. Klir and Zhenyuan Wang: Generalized Measure Theory. Springer, Berlin 2009.

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

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

[17] M. Mareš: Weak arithmetics of fuzzy numbers. Fuzzy Sets and Systems *91* (1997), 2, 143–154.

[18] M. Mareš: Fuzzy components of cooperative market. In: Perception-Based Data Mining and Decision Making in Economics and Finance. (I. Batyrshin and J. Kacprzyk, eds.), Physica-Verlag, Heidelberg 2007.

[19] M. Mareš and R. Mesiar: Information in granulated data source. In: Proc. ICSCCW-2007 (W. Pedrycz, R. Aliev, Mo. Jamshidi, and B. Turksen, eds.), b-Quadrat Verlag, Antalya 2007, pp. 185–194.

[20] Ming Quing and Tian-Rui Li: Some properties and new formulae of fuzzy entropy. In: Proc. 2004 IEEE Internat. Conf. on Networking, Sensing and Control, Vol. I, pp. 401–406.

[21] L. Narens: Abstract Measurement Theory. MIT Press, London 1985.

[22] R. B. Nelsen: An Introduction to Copulas. Springer, New York 2006.

[23] J. von Neumann and O. Morgenstern: Theory of Games and Economic Behaviour. Princeton Univ. Press, Princeton 1944.

[24] Z. Pawlak: Rough Sets. Kluwer, Dordrecht 1991.

[25] A. Rényi: On measures of entropy and information. In: Proc. Fourth Berkeley Symp. on Math. Statistics and Probability 1961, Vol. 1, pp. 547–561.

[26] C. Shannon and W. Weaver: A mathematical theory of communication. Bell Systems Technical Journal *27* (1948), 379–423, 623–653.

[27] N. Wiener: Cybernetics or Control and Communication in the Animal and the Machine. Cambridge Univ. Press, Cambridge 1948.

[28] N. Wiener, A. Rosenblueth, and J. Bigelow: Behaviour, purpose and teleology. Philosophy of Science *10* (1943), 1, 18–24.

[29] K. Winkelbauer: Communication channels with finite past history. In: Trans. Second Prague Conference on Information Theory etc., Publ. House of ČSAV, Prague 1960.

[30] Yao Min and Zhang Sen: Generalized fuzzy entropy and its applications. In: Proc. 4th Internat. Conf. on Signal Processing 1998, Vol. 2, pp. 1197–1200.

[31] L. A. Zadeh: Fuzzy sets. Inform. and Control *8* (1965), 3, 338–353.

[32] L. A. Zadeh: From computing with numbers to computing with words. IEEE Trans. Circuits and Systems *45* (1999), 105–109.

Milan Mareš, Institute of Information Theory and Automation — Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8 and Faculty of Sciences, Institute of Applied Information Sciences, South Bohemian University, Branišovská 31, 370 05 České Budějovice. Czech Republic.
e-mail: mares@utia.cas.cz