KYBERNETIKA — VOLUME 38 (2002), NUMBER 6, PAGES 709–716

DUAL MEANING OF VERBAL QUANTITIES*

MILAN MAREŠ AND RADKO MESIAR

The aim of the paper is to summarize and interpret some ideas regarding effective processing of vague data. The main contribution of the submitted approach consists in respecting the fact that vague data can be decomposed into two parts. The numerical one, describing the quantitative value of such data, and the semantic one characterizing the qualitative structure of the vagueness included into them.

This partition of vague verbal data leads to a significant simplification of their practical processing, namely if complicated computations are considered. This regards also the methods and concepts of descriptive statistics and their handling in the case in which vague data appear on the input.

Particular parts of the alternative method of vague data processing which is discussed here were in a more detailed way investigated in some of the referred paper. The purpose of this paper is to give a brief explanation of the method in its complete connections with the practical processing of vagueness in quantitative environment.

1. VERBAL QUANTITY

The main notion of the presented model of quantitative uncertainty is the one of the verbal quantity. By it we mean any verbal expression characterizing the quantitative phenomena, like "approximately 8", "about 10", "several", and also "a bit more than 20", "approximately divisible by 100", "something between 10 and 20" and others.

The theory of fuzzy sets offers relatively effective tools for modelling such verbal quantities by means of fuzzy quantities (see [4, 5]) as fuzzy subsets of the real line R. This representation is illustrative but further processing of such fuzzy quantities, namely if they become members of rather complicated formulas, displays some discrepancies. Many of them are connected with the classical method of fuzzy quantities processing, called *the extension principle* (cf. [1, 10] and also [4]).

Just to remember the notations, let $a_1, a_2, \dots, a_n$ be fuzzy quantities with membership functions $\mu_1, \mu_2, \dots, \mu_n$, and let f be a real-valued function of n real vari-

*The research summarized in this paper was partly supported by the Grant Agency of the Czech Republic, grant No. 201/99/0032 and Grant Agency of the Academy of Sciences of the Czech Republic, grants No. A1075905 and K 1075601, and grant of the Slovak Scientific Grant Agency VEGA 1/8331/01.
Presented as a contributed lecture at the International Conference "International Data Analysis" (IDA'2000) held in Innsbruck, on September 19–22, 2000.

ables, $f(x_1, x_2, \ldots, x_n) \in R$, $(x_1, \ldots, x_m) \in \operatorname{Dom} f \subset R^n$. Then this function can be extended to the class of fuzzy quantities with values in the same class. The membership function $\mu_b : R \to [0,1]$ of the fuzzy quantity $b = f(a_1, a_2, \ldots, a_n)$ can be defined by

$$\mu_b(y) = \sup \Big[\min(\mu_1(x_1), \mu_2(x_2), \ldots, \mu_n(x_n)) : \tag{1}$$
$$(x_1, x_2, \ldots, x_n) \in \operatorname{Dom} f,\ y = f(x_1, x_2, \ldots, x_n) \Big], \text{ with convention } \sup \emptyset = 0.$$

For the simple arithmetic operations of addition and multiplication, $a \oplus b$ and $a \odot b$, respectively, where a, b are fuzzy quantities with μ_a, μ_b,

$$\begin{aligned}
\mu_{a \oplus b}(y) &= \sup_{x \in R} [\min(\mu_a(x), \mu_b(y - x))], \quad y \in R, && (2)\\
\mu_{a \odot b}(y) &= \sup_{\substack{x \in R \\ x \neq 0}} [\min(\mu_a(x), \mu_b(y/x))], \quad y \in R,\ y \neq 0, && (3)\\
\mu_{a \odot b}(0) &= \max(\mu_a(0), \mu_b(0)).
\end{aligned}$$

Analogously, for a fuzzy quantity a and crisp real number $r \in R$ the product $r \cdot a$ is a fuzzy quantity with membership function

$$\begin{aligned}
\mu_{r \cdot a}(y) &= \mu_a(y/r) && \text{for } r \neq 0,\\
&= 1 && \text{for } r = 0,\ y = 0, \qquad (4)\\
&= 0 && \text{for } r = 0,\ y \neq 0.
\end{aligned}$$

In the following sections, we will use also the following notation. If $r \in R$ is a real number then by $\langle r \rangle$ we denote the fuzzy quantity with possibilities condensed into the value r, i.e. $\mu_{\langle r \rangle}(r) = 1$, $\mu_{\langle r \rangle}(x) = 0$ for $x \neq r$. The properties of the above operations are summarized in [1, 4, 5].

The computation using extension principle has a logical and well understandable structure analogous to the classical convolution. Anyhow, it displays also some difficulties. Namely, the algorithms for its numerical realization are not simple. What is even more significant, the obtained results sometimes do not correspond with our intuitive expectation and also with practical experience regarding the manipulation with vague data. In the case of summation or of multiplication of "higher" possible values the extent of the possible values of the results enormously grows. Moreover, the product of fuzzy quantities with integer (or, generally, discrete) possible values may display hardly interpretable gaps in the set of its possible values.

For all these reasons it appears rational to look for an alternative approach to the modelling vague (verbal) data which could better reflect their structure and which could also simplify their processing in practical algorithms.

2. DECOMPOSITION OF VERBAL QUANTITIES

The model discussed in this paper is based on a rather modified idea of the verbal quantity. It still respects the representation of verbal quantitative data by fuzzy quantities but focuses the attention to the process of their generation from the verbal expressions. In this sense the following paragraphs summarize the main steps made in some of the referred papers, namely in [6, 7, 8, 9] and [2].

The crucial idea of the presented model is the one of certain dualism of the verbal quantitative data. They include two components – the quantitative part and the semantic description of the uncertainty connected with it. The first component is represented by a crisp real number (in the following text we call it the "*crisp core*" of the considered verbal variable), the second one is represented by a real valued function (we call it the "*shape*" of the verbal variable), which can be interpreted as a normalized form of the membership function describing the distribution of uncertainty being present in the verbal formulation. For example, verbal quantity

$$A = \text{"approximately 8"}$$

evidently includes numerical deterministic component – crisp core – $x_A = 8$ and a semantic component – shape – which we denote by φ_A and which describes the uncertainty hidden in the word "approximately".

We suppose that any *simple shape* φ_A fulfils the following properties

$$\varphi_A(0) = 1, \tag{5}$$

$$\varphi_A(x) \text{ is non-decreasing for } x < 0 \text{ and non-increasing for } x > 0. \tag{6}$$

There exist also more complex verbal quantities, we call them "composed" which represent a combination of the simple ones. For example

B — "something between 10 and 20"

which can be decomposed into

"something more than 10" AND "something less than 20"

$C =$ "approximately integer multiple of 10"

which can be decomposed into

"approximately 10" OR "approximately 20" OR "approximately 30" OR ....

Their shapes, we call them *composed shapes*, need not fulfil conditions (5) and (6) but they can be decomposed, by means of simple operations, into the simple shapes.

There exist also "anonymous" verbal variables called also linguistic modifiers which, seemingly, have no crisp core. It regards such quantitative verbal expressions like "many", "several", "few", "some amount" and others. In fact, even these verbal quantities express the distribution of uncertainty (usually relatively wide one) related to some not named but intuitively existing real value given by the actual application. The expression "much money" means something completely different in the case of family expenses and of state budget but in each of these cases its semantic content refers to the uncertainty connected with some numerical value. Also the verbal variable "something more than 10" mentioned in the previous paragraph means a composition of simple verbal variables "10" PLUS "something", where the former one is deterministic (its shape is degenerated to zero) meanwhile the latter one is anonymous. It can be, due to the actual application, connected with a crisp core (in our case, let us say, equal to 1) and a (probably rather asymmetric) shape function representing the considered verbal expression.

Anyhow, both, composed and anonymous, verbal variables can be expressed by means of well chosen combination of the simple ones.

3. REPRESENTATION OF VERBAL VARIABLES BY FUZZY QUANTITIES

It is to be stressed that evidently the most adequate mathematical representation of verbal quantity A is a fuzzy quantity a with membership function μ_a. The above decomposition of A into crisp core x_A and shape φ_A is not only to illustrate the inner structure of A but it can be used for the construction of that fuzzy quantity a which best corresponds to the quantitative and qualitative component of A. The formalization of the correspondence between verbal variable A and fuzzy quantity a is analyzed, e. g., in [6, 7, 8] and [9] and its character can be briefly summarized as follows.

If A is a simple verbal variable with the crisp core x_A, and simple shape φ_A (fulfilling (5), (6)) then we write $A = (x_A, \varphi_A)$ and the fuzzy quantity a with membership μ_a, corresponding to it, is defined by

$$\mu_a(x) = \varphi_A(x - x_A), \quad x \in R. \tag{7}$$

For example, if we consider verbal variable $A =$ "approximately 8" with the crisp core $x_A = 8$ and shape φ_A defined as

$$\varphi_A(x) = \max(0,\ 1 - |x|)$$

i. e.

$$\varphi_A(x) = x + 1 \quad \text{for } x \in [-1, 0], \qquad \varphi_A(x) = 1 - x \quad \text{for } x \in [0, 1]$$
$$\varphi_A(x) = 0 \quad \text{for } x \notin [-1, 1],$$

then the corresponding fuzzy quantity a with membership function μ_a is defined by

$$\mu_a(x) = \max(0,\ 1 - |x - 8|) = \varphi_A(x - 8).$$

The fuzzy quantity corresponding to composed verbal variables can be derived from the ones connected with their simple components. So, e. g., the verbal variable

$C =$ "approximately integer multiple of 10"

mentioned in the previous section, can be connected with composed shape

$$\varphi_C(x) = \max(\varphi_n(x),\ n = 0, \pm 1, \pm 2 \ldots)$$

where $\varphi_n(x) = \varphi_A(x - 10 \cdot n)$ and φ_A is the shape of "approximately", defined, e. g., as $\varphi_A(x) = \max(0,\ 1 - |x|)$. Then the fuzzy quantity c connected with C is determined by its membership function $\mu_c(x) = \varphi_C(x - x_C)$ under the natural assumption that we have put $x_C = 10$. Then also $\varphi_n(x) = \varphi_A(x - x_C \cdot n)$.

We have already mentioned that the shape φ_A of a verbal variable A usually has formal properties of membership function of some fuzzy quantity with modal value in 0. Let us call, for this moment, this fuzzy quantity α. Then it is evident that (7) implies $a = \alpha \oplus \langle x_A \rangle$, where the notation $\langle x_A \rangle$ was introduced in Section 1.

4. CALCULATION WITH DECOMPOSED VERBAL VARIABLES

The decomposition of a verbal variable into its quantitative and semantic components opens the way to the main methodological consequence of such treatment. Namely, we are able to part also the processing of the quantitative and semantic component with respect to their natural structure. It is evident that each of these two components demands a specific approach to its handling following from the specificity of its role in the description of vague verbal data.

If we consider a computational procedure with verbally described vague data at the input then the classical theory of fuzzy quantities offers to represent each verbal variable by a fuzzy quantity (e. g., by using (7)) and then to use arithmetic (and eventually other) operations based on the extension principle). This procedure is usually computationally complex and it must be repeated whenever new set of verbal data appears at its input. Moreover, the formal structure of the resulting fuzzy quantity need not be intuitively satisfactory. For example, the extent of its possible values can be too large, there can appear hardly interpretable gaps in them, etc. These intuitively percepted discrepancies are frequently connected with the specific logical structure of the extension principle.

The model of vague verbal data described above offers another, often more effective, approach. The input verbal variables need not be transformed into fuzzy quantities but their components – crisp cores and shapes – can be processed by means of methods corresponding with their characters. Then a resulting crisp core and resulting shape are derived and only they are transformed, using (7), into the resulting fuzzy quantity as a representation of the vague result of the realized procedure. This method displays several advantages.

First, it respects the principal differences between the quantitative and qualitative (semantic) component of verbal variable. Their separate processing allows to use for each of them the operations which correspond with the demands of the actual application.

Further, especially the choice of the operations applied to the shapes enables us to keep, e. g., the extent of possible values of the result in acceptable and intuitively justifiable limits.

And, last but not least, in many practical applications the relevant computational algorithm is repeated many times, always for a new set of vague verbal input data which differ only in their quantitative components but preserve the semantic structure of uncertainty (naturally, they are produced by identical or similar sources). Then the separate processing of both components means that for each new set of the input data only the part of the algorithm which processes the crisp cores is to be inovated. The result of the processing of the shapes (which is usually computationally more complicated) can be taken from the previous applications of the same algorithm and put into (7) together with the new resulting crisp core.

These advantages could be taken into consideration when we evaluate the applicability of the submitted model of verbal variable to practical situations.

The actual methods of processing crisp cores and shapes can follow from the situation – meanwhile the processing of the crisp cores evidently turns into the classical deterministic procedures over real numbers (or real-valued vectors), the processing of shapes opens much more possibilities.

If φ_A, φ_B are (not necessarily simple) shapes of verbal variables A and B, respectively, then it is possible to use, due to the realistic interpretation of the relation between A and B in the structure of the considered algorithm, some operations based on the fuzzy logic, e. g.,

$$\min(\varphi_A(x), \varphi_B(x)), \quad \max(\varphi_A(x), \varphi_B(x)),$$

or operations of more algebraic type, like

$$\varphi_A(x) \cdot \varphi_B(x), \quad \varphi_A(x) + \varphi_B(x) - \varphi_A(x) \cdot \varphi_B(x), \quad (\varphi_A(x) + \varphi_B(x))\,/2,$$

or more structural operations inspired, e. g., by the extension principle

$$\begin{aligned} \varphi(x) &= \sup_{y \in R} \left(\min(\varphi_A(y), \varphi_B(x-y))\right), \\ \varphi(x) &= \sup_{\substack{y \in R \\ y \neq 0}} \left(\min(\varphi_A(y), \varphi_B(x/y))\right). \end{aligned} \tag{8}$$

It is worth mentioning that, e. g., in the case of the "extension-principle-like" multiplication the final result is much more realistic (i. e., much more concentrated to realistically expected values) than in the case where the same extension principle was applied to the usual fuzzy quantities. For example, if we are to multiply two verbal variables A, B represented by fuzzy quantities a, b, such that

$$x_A = x_B = 100$$
$$\{x \in R : \varphi_A(x) > 0\} = \{x \in R : \varphi_B(x) > 0\} = (-10, 10)$$

i. e., the measures of their support are 20, then we may proceed either by means of (3) and to calculate fuzzy quantity $c = a \odot b$ such that

$$\{x \in R : \mu_c(x) > 0\} = (8100,\ 12100),$$

with measure of support equal to 4000, or we may calculate the shape φ_C of a verbal variable $C = A \cdot B$, where $x_C = 10000$ and the shape φ_C is calculated by (8) with

$$\{x \in R : \varphi_C(x) > 0\} = (-100,\ 100).$$

Using (7) we construct a fuzzy quantity c^* with μ_c^* such that

$$\{x \in R : \mu_c^*(x) > 0\} = (9900,\ 1100),$$

i. e., with measure of support equal to 200. It is significant to note that in many practical applications the result represented by c^* appears much more realistic and corresponding with the practical experience.

It has to be stressed that the approach to the computation with fuzzy (verbal) quantities described in this paper represents an extension of the method based on the classical extension principle. If we use the extension principle for shapes then it can be easily verified that the result of (8) and (7) for addition is equal to the result of (2). On the other hand, there may exist a significant difference between the results of (3) and (7), (8) for multiplication as shown in the previous paragraphs. Note, please, that the application of (7), (8) for shapes is computationally simpler, and also the obtained results may be intuitively more acceptable. Of course, the possibility to combine the shapes of the input data also by means of other rules that the potential extension principle is, represents the most significant extension of the computational methods (with different shapes of the results). The choice of the actually applied computational procedure can reflect the character of the specific applications.

5. CONCLUSIVE REMARKS

The main purpose of the above sections is to remember and briefly motivate an alternative approach to vague data, usually consequently represented by fuzzy quantities. It summarizes the most significant of the ideas dispersed in the referred papers.

It is possible to conclude that the suggested method of processing vague verbal quantities which is based on the separate processing of their quantitative and qualitative components displays several advantages in comparison with direct processing of fuzzy quantities.

In addition to its relative simplicity it offers a wide scale of possible modifications adapted to the given applied problems which are to be solved. This wide choice of particular procedures means also the flexibility of the general model, respecting the variability of the modelled situations. The submitted approach to the verbal variables, by means of the variability of its particular procedures, also openly admits the role of the subjectivity and personal reflection of the modelled situation in the process of computing real applied algorithms with the vague input data.

The model of verbal variables presented above is in certain sense a simplified version of the model treated in some of the referred papers. Namely, it is possible to include in it, without any change of the interpretation and methods formulated above, also a component reflecting the quality or reliability of the sources of the processed verbal data. It is to reflect the fact that, e. g., the verbal expression "approximately x" means something quite else for small or large values of x, as well as for different procedures (subjective estimation, measurement by different tools, etc.) of the generation of the "approximate" values.

This variability of the data sources can be included into the model (see [6, 8, 9]) by means of so called "*scale function*". It is an increasing and continuous real-valued function $f : R \to (-\infty, \infty)$ such that $f(0) = 0$. The influence of the quality of the data source on the form of the fuzzy quantity representing the processed verbal variables can be expressed by a modification of (7) into the form

$$\mu_a(x) = \varphi_a\left(f_A(x) - f_A(x_A)\right) \tag{9}$$

where f_A is the scale characterizing the source of the verbal variable A.

Obviously, the higher the gradient (first derivative) of the scale is the more exact or more reliable appears to be the source of "approximate x" or other verbal description of the considered variable. The processing (mutual combination or composition) of particular scales during the process of calculation with verbal variables represents a specific problem of modelling verbal variables (see [6, 8, 9]).

(Received March 21, 2002.)

REFERENCES

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

[2] E. E. Kerre, M. Mareš, and R. Mesiar: Orderings of generated fuzzy quantities. In: Proc. IPMU'98, La Sorbone, Paris 1998, Vol. I, pp. 250–254.

[3] M. Mareš: Dual meaning of verbal quantities (abstract). In: IDA'2000 – Abstracts. Vienna University of Technology, Vienna 2000, p. 47.

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

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

[6] M. Mareš: Verbal quantitative information in economic reality. Acta Oeconomica Pragensia *5* (1997), 1, 121–134.

[7] M. Mareš and R. Mesiar: Composition of shape generators. Acta Mathematica et Informatica Universitatis Ostraviensis *4* (1996), 1, 37–45.

[8] M. Mareš and R. Mesiar: Calculation over verbal variables. In: Computing With Words in Information/Intelligent Systems (J. Kacprzyk and L. A. Zadeh, eds.), Physica–Verlag, Heidelberg 1999, pp. 409–427.

[9] M. Mareš and R. Mesiar: Vagueness of verbal variable. In: Soft Computing in Financial Engineering (R. Ribeiro, J. Kacprzyk, R. R. Yager, and L. A. Zadeh, eds.), Physica–Verlag, Heidelberg 1999, pp. 3–20.

[10] L. A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning. Information Sci., Part I, 8, 199–249, Part II, 8, 301–357, Part III, 9, 43–80.

[11] H. J. Zimmermann: Soft Computing in Financial Engineering. Physica–Verlag, Heidelberg 1999, 3–20.

Prof. RNDr. Milan Mareš, DrSc., 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: mares@utia.cas.cz

Prof. RNDr. Radko Mesiar, DrSc., Faculty of Civil Engineering. Slovak Technical University, Radlinského 11, 813 68 Bratislava. Slovak Republic.
e-mail: mesiar@cvt.stuba.sk