UDC 004.93.1; 681.142
[1,2]Заяць В.М., д.т.н., професор; [1]Маєвський Яцек, PhD. доцент; [1]Марціняк Томаш, PhD. доцент; [3]Рибицька Ольга Мар'янівна, к.фіз.-мат.н., доцент; [3]Заяць Марія Михайлівна, ст. викладач.
[1]Університет Технологічно-Природничий (UTP), м. Бидгощ, Польща.
[2]Національний університету водного господарства та природокористування, м. Львів, Україна.
[3]НУ "Львівська політехніка", м. Львів, Україна.

MODELS AND ALGORITHMS FOR PROCESSING OF FUZZY INFORMATION

Zaiats V.M., Majewski J., Marciniak T., Rybytska O.M., Zaiats M.M. Models and algorithms for processing of fuzzy information. Methods and algorithms for the solution of applied problems of processing of statistical information on the basis of the theory of fuzzy integrals and blurry logic are proposed. The developed approaches provide a good choice of belonging function, the criterion of likelihood and fuzzy measure based on the knowledge and intuition of the expert. It is substantiated the expediency of using declarative programming languages for the processing of statistical information.
Keywords: fuzzy information; fuzzy set; blurry logic; belonging function; criterion of likelihood; declarative programming languages.

Заяць В.М., Маєвський Я., Марціняк Т., Рибицька О.М., Заяць М.М. Моделі та алгоритми для опрацювання нечіткої інформації. Запропоновано методи та алгоритми до розв'язання прикладних задач обробки нечіткої інформації на основі теорії нечітких інтегралів та розмитої логіки. Розроблені підходи забезпечують вдалий вибір функції приналежності, критерію ймовірності та нечіткої міри, заснованої на знанні та інтуїції експерта. Обгрунтовується доцільність використання декларативних мов програмування для обробки статистичної інформації.
Ключові слова: нечітка інформація; нечітка множина; розмита логіка; функція приналежності; критерій вірогідності; декларативні мови програмування.

Заяць В.М., Маевский Я., Марциняк Т., Рыбыцька О.М., Заяць М.М. Модели и алгоритмы для обработки нечеткой информации. Предложены методы и алгоритмы к решению прикладных задач обработки размытой информации на основе теории нечетких интегралов и размытой логики. Разработанные подходы обеспечивают удачный выбор функции принадлежности, критерия вероятности и нечеткой меры, основанной на знании и интуиции эксперта. Обосновывается целесообразность использования декларативных языков программирования для обработки статистической информации.
Ключевые слова: нечеткая информация; нечеткое множество; размыта логика; функция принадлежности; критерий достоверности; декларативные языки программирования.

I. Formulation of the problem. The problems of mathematical and computer modeling today lie, in particular, in the inappropriateness of clear logic and task model with clearly defined input parameters in cases where, for some reason, there are contradictions, uncertainty or lack of clarity of information about the object under study , system or phenomenon [1, 4, 6, 8, 11, 13-17].

Uncertainty is known to arise because of the lack of knowledge relating to a particular event [2]. She is present to the experiment. Mathematical models for information uncertainty objects are based on the probability theory, the theory of opportunities, measures of confidence, fuzzy logic and several others.

The phenomenon of fuzziness arises in the process of combining into one whole of objects, which have a common property:

$$X = \{x | owns\ \varphi\}$$

where *x* - all the elements of a certain universal set .

Given that there are always elements in the reality that it is unclear whether they possess the specified property or not, it is not a plural in the classical sense. Any attempt to interpret the general description leads to fuzzy concepts, since the exact description contains an excess of details. Increasing the accuracy of the description leads to an increase in the amount of information, the content of which decreases until the time when the accuracy and meaningfulness do not become mutually exclusive. For the first time, L.A. Zadeh [3] stressed the need for uncertainty for the transmission of content. It was the ideas of this American scientist give pushed for the development of "fuzzy mathematics" [4], which, along with the apparatus of fuzzy sets, contains other methods of work with uncertainty.

The application of the theory of fuzzy sets and measures is a step towards the convergence of the precision of classical mathematics with a false inaccuracy of the environment, an attempt to overcome the linguistic barrier between a person whose judgment and evaluation are approximate and fuzzy with technical means (by software), which can only carry out precise instructions [5].

II. Analyses of last researches and aim of work. A device that allows you to work with fuzzy logic, blurred" parameters of models, is a Fuzzy-technology device [1-3, 8]. The Fuzzy-Technology division has some kind of expert systems [5, 6].

Fuzzy output systems are designed to implement the fuzzy output process and serve as a conceptual basis for all modern fuzzy logic. Fuzzy output systems allow you to solve the tasks of automated control, data classification, image recognition, decision-making, and many others.

Linguistic variables extend the ability to represent knowledge. They are determined by fuzzy sets, whose values are determined by functions of fidelity. Affiliation functions can be obtained through expert assessments [6], or by analysing fuzzy clusters. According to [6], fuzzy expert systems can be implemented when the cost of obtaining accurate information, that is, information in absolute terms, exceeds the maximum revenue from the restructuring of a model or is practically impossible.

It is known that the initial stage of constructing artificial intelligence on the basis of the use of natural language is based on ambiguous logic and the mechanism of output with rigid rules. First-generation expert systems work on the following set of rules:

If $\boldsymbol{X}$ *it belongs* $\boldsymbol{A_k}$*, then* $\mathbf{Y_k}$ *it belongs* $\boldsymbol{C_k}$
were $k = 1, \dots, N$.

Here, there is a set of exact numerical or symbolic values that can be linguistic meanings of linguistic variables.

At present, in expert systems, productive models are the most used formalism for representing knowledge

*"If **A**, then **B**".*

The second generation of expert systems has at least two peculiarities: fuzzy representation of knowledge and fuzzy calculations. One of the most common tasks of logical output in the conditions of fuzziness can be formulated as follows

Given (fuzzy) production rule "If A then B".
Observed A '(A to some extent). What should be B?

After obtaining a fuzzy set of conclusions, we find a specific numerical means (conducting dephasing).

The purpose of this work is to develop a fuzzy data analysis algorithm, which consists of creating a fuzzy set of data model, creating a fuzzy knowledge data base model and selecting the membership functions for the purpose of effective classification of data and their statistical processing.

III. Building models with blurry logic. Consider an object with one output and type inputs

$$y = f(x_1, x_2, \dots, x_n), \tag{1}$$

where $x_1 \dots . x_n$ - *the* set of values of input variables; y - output variable.

For construct a mathematical model on the basis of establishing the relationship between input and output variables in accordance with experimental data we shell make the operation of pacification and quantitative and qualitative variables are translated into linguistic terms.

$$U_i = \left[\underline{u_i}, \overline{u_i}\right], i = \overline{1, n} \tag{2}$$

$$Y = [\underline{y}, \overline{y}] \tag{3}$$

where $\underline{u_i}, \overline{u_i}$ - the smallest and the highest possible value y of the variable x_i; $\underline{y}, \overline{y}$ - the smallest and the highest possible value of the output variable y.

To solve the problem (1) it is necessary to apply a method of making a decision by means of which a fixed vector of input variables $x^* = (x_1^*, x_2^*, ..., x_n^*)$, $x_i^* \in U_i$ would unambiguously be placed in accordance with the solution $y^* \in Y$. For the formal establishment of this dependence we shall consider the input x_i, $i = \overline{1,n}$ and output parameters y as linguistic variables given on universal sets (2), (3). To evaluate linguistic variables x_i, $i = \overline{1,n}$, and y we will use qualitative terms from the following term sets: $A_i = \{a_i^1, a_i^2, ..., a_i^{p_i}\}$ – term - set of input variable x_i, $i = \overline{1,n}$; $D = \{d_1, d_2, ..., d_m\}$ – term -set of output variable y.

For the construction of term -sets, one can apply, for example, the method proposed in [16].

For each term of each linguistic variable, based on expert knowledge, the functions of membership (trapezoidal, triangular, rectangular, sinusoidal, parabolic, etc.) [5] are constructed based on expert knowledge, where $\mu^{a_i^p}(x)$ - the degree of belonging of the element $x \in U_i$ to the term $a_i^p \in A_i$, $i = \overline{1,n}$; $p = \overline{1, p_i}$; $\mu^{d_j}(y)$ – of the measure of the membership of the element y to the term $d_j \in D$, $j = \overline{1,m}$.

The definition of linguistic estimates of variables and the functions of membership necessary for their formalization is the first stage in the construction of a fuzzy model of the object being studied. In the literature on fuzzy logic, he received the name of the phasing of variables (English fuzzification) [8].

IV. Approach to building Models of Fuzzy Knowledge Bases. The next step is to create a base of fuzzy knowledge. Let the object (8) know the rules that connect its inputs and output using vectors such as:

$$N = k_1 + ... + k_j + ... + k_m, \qquad (4)$$

where k_j – the number of experimental data corresponding to the same value d_j of the term-set of the output variable y; m - the total number of terms of the output variable, and in the general case $k_1 \neq ... \neq k_m$.

We will assume that the number of available experimental data $N < p_1 \, p_2 ... p_n$, is less than the total overview of various combinations of possible terms of input variables p_i, $i = \overline{1,n}$.. Then the knowledge base is a table formed according to the following rules:

1. The table's size $N \times (n+2)$, is equal to the number of columns $n+2$ – and N – the number of rows;
2. Each row of the matrix - a combination of input variables assigned by the expert to one of the possible values of the term-set of the output variable y. In this case, the first k_1 lines correspond to the value of the output variable $y = d_1$, the following k_2 lines - the value $y = d_2$, etc., and the last k_m lines - the value $y = d_m$;
3. The first n columns of the matrix correspond to the input variables $x_i, i = \overline{1,n}$; $(n+1)$ -st - the weight of the rule W_{jp}, $j = \overline{1,m}$, $p = \overline{1,k_j}$, $(n+2)$ -nd - the value d_j of the output term-set of the variable y, $j = \overline{1,m}$, corresponding to the combination of values in the first $(n+1)$ columns.
4. The element located at the intersection of the i-th column and the jp-line corresponds to the linguistic evaluation of the parameter x_i in the row of knowledge matrix with the number jp. In this case, the linguistic assessment a_i^{jp} is chosen with the term-set corresponding to the variable x_i, i.e. $a_i^{jp} \in A_i$, $i = \overline{1,n}$; $j = \overline{1,m}$; $p = \overline{1,k_j}$.

Table 1. General view of the fuzzy knowledge base

Number Incoming combination	**Input variables**				**Weight**	**Output variable**
	x_1	x_2	$\dots x_i \dots$	x_n	w	y
11	a_1^{11}	a_2^{11}	a_i^{11}	a_n^{11}	w_{11}	
12	a_1^{12}	a_2^{12}	a_i^{12}	a_n^{12}	w_{12}	d_1
…	…	…	…	…	…	
$1k_1$	$a_1^{1k_1}$	$a_2^{1k_1}$	$a_i^{1k_1}$	$a_n^{1k_1}$	w_{1k_1}	
…	…	…	…	…	…	…
j_1	a_1^{j1}	a_2^{j1}	a_i^{j1}	a_n^{j1}	w_{j1}	
j_2	a_1^{j2}	a_2^{j2}	a_i^{j2}	a_n^{j2}	w_{j2}	d_j
…	…	…	…	…	…	
jk_j	$a_1^{jk_j}$	$a_2^{jk_j}$	$a_i^{jk_j}$	$a_n^{jk_j}$	w_{jk_j}	
…	…	…	…	…	…	…
m_1	a_1^{m1}	a_2^{m1}	a_i^{m1}	a_n^{m1}	w_{m1}	
m_2	a_1^{m2}	a_2^{m2}	a_i^{m2}	a_n^{m2}	w_{m2}	d_m
…	…	…	…	…	…	
mk_m	$a_1^{mk_m}$	$a_2^{mk_m}$	$a_i^{mk_m}$	$a_n^{mk_m}$	w_{mk_m}	

When an expert creates linguistic rules such as "IF THING" that form the basis of fuzzy knowledge about a particular object, the expert's confidence in each rule may be different. If one rule in the opinion of an expert can serve as an undeniable truth, then according to another rule in the same expert there may be some doubts.

In order to reflect these different degrees of confidence in the base of fuzzy knowledge, the weighting of the rules is introduced - these are numbers from the interval [0, 1], which characterize the expert's confidence in each specific rule chosen by him to make a decision. The general view of the knowledge base is given in Table. 1.

After building the knowledge base, you need to carefully check in Table 1 the presence of opposite lines of content, that is, rules that the same input variables have different output values. The introduced matrix of knowledge defines a system of logical utterances such as "YES - THEN, OTHER", which associate the values of input variables with one of the possible output values $d_j,\ j=\overline{1,m}$:

IF $(x_1 = a_1^{11})$ **AND** $(x_2 = a_2^{11})$ **AND…AND** $(x_n = a_n^{11})$ (with weight w_{11}),
OR $(x_1 = a_1^{12})$ **AND** $(x_2 = a_2^{12})$ **AND** … **AND** $(x_n = a_n^{12})$ (with weight w_{12}),
OR …
AND $(x_1 = a_1^{1k_1})$ **AND** $(x_2 = a_2^{1k_1})$ **AND** … **AND** $(x_n = a_n^{1k_1})$ (with weight w_{1k_1}),
THEN $y = d_1$, OTHER
IF $(x_1 = a_1^{21})$ **AND** $(x_2 = a_2^{21})$ **AND** … **AND** $(x_n = a_n^{21})$ (with weight w_{21}),
OR ………………………..……………….
OR $(x_1 = a_1^{2k_2})$ **AND** $(x_2 = a_2^{2k_2})$ **AND** … **AND** $(x_n = a_n^{2k_2})$ (with weight w_{2k_2}),
THEN $y = d_2$, OTHER …
IF $(x_1 = a_1^{m1})$ **AND** $(x_2 = a_2^{m1})$ **AND** … **AND** $(x_n = a_n^{m1})$ (with weight w_{m1}),
OR …………………………………………..
OR $(x_1 = a_1^{mk_m})$ **AND** $(x_2 = a_2^{mk_m})$ **AND** … **AND** $(x_n = a_n^{mk_m})$ (with weight w_{mk_m}), THEN $y = d_m$.

A similar system of logical expressions is called a fuzzy knowledge base. Using the operations $\cup$ (OR) and $\cap$ (AND) described system of logical statements can be rewritten in a more compact form:

$$\bigcup_{p=1}^{k_j}\left[\bigcap_{i=1}^{n}\left(x_i = a_i^{jp}\right)\right] \rightarrow y = d_j,\ \ j=\overline{1,m}. \qquad (5)$$

Thus, the input relation (1), which establishes the connection between the input parameters and the output variable, is formalized in the form of a system of fuzzy logical statements (5) based on the created matrix of knowledge. The rules for fuzzy considered of the described system, in which the degree of truth is different from zero, named active.

In [10], a method is proposed to use fuzzy logic equations that are based on a knowledge matrix or isomorphic system of logical expressions (5) and allow us to calculate the values of the membership function of the output variable for the fixed values of the inputs of the object.

Linguistic estimates a_i^{jp} of the variables $x_1,..x_n,$ contained in the logical statements (5) will be considered as fuzzy sets defined on universal sets (2). We introduce the following notation:

$\mu^{a_i^{jp}}(x_i)$ – function-membership of parameter x_i to fuzzy term $a_i^{jp},\ i=\overline{1,n},\ j=\overline{1,m},\ p=\overline{1,k_j}$; $\mu^{d_j}(x_1,x_2,...x_n)$ – function-membership vector of input variables $x=(x_1,x_2,..x_n)$ to term of the output variable $y=d_j,\ j=\overline{1,m}.$

Thus, we have two types of functions, the relationship between which is determined by the base of fuzzy knowledge (5), on the basis of which you can output a system of logical equations, which can be submitted in a compact form:

$$\mu^{d_j}(x_1,x_2,...,x_n) = \bigvee_{p=1}^{k_j}\left(w_{jp}\left[\bigwedge_{i=1}^{n}\mu^{a_i^{jp}}(x_i)\right]\right),\ \ j=\overline{1,.m}; \qquad (6)$$

where $\vee -$ logical «OR»; $\wedge -$ logical «**I**».

The decision $d^* \in D\{d_1, d_2, ..., d_m\},$ that corresponds to a fixed vector of the values of input variables $x^* = (x_1^*, x_2^*, ..., x_n^*)$ will be carried out in accordance with the following algorithm constructed using the apparatus of the fuzzy (blurry) logic [10]:

1. The possible range of change of controlled parameters will determined, a knowledge base is created with the use of expert data and a system of fuzzy logic equations is derived (6).
2. The vector of the values of the input variables $x^* = (x_1^*, x_2^*, ..., x_n^*)$ is fixed.
3. We specifies the function - membership of fuzzy term for different controlled parameters.
4. Using logical equations (6), the values of many parametric functions membership $\mu^{d_j}(x_1^*, x_2^*, ..., x_n^*)$ of vector X for all values $d_j,\ j = \overline{1,m}$ of the output variable y are calculated. In this case, the logical operations $\cup$ (OR) and $\cap$ (AND) over the membership functions are replaced by the operations max and min:

$$\mu(a) \vee \mu(b) = \max[\mu(a), \mu(b)], \qquad (7)$$

$$\mu(a) \wedge \mu(b) = \min[\mu(a), \mu(b)]. \qquad (8)$$

That is, first find the minimum values of belonging functions in each rule, and then among them they choose the highest value of the membership function among all rules for each value, which corresponds to the original variable *y*. Thus, the conclusion is made that the origin variable belongs *y* to a term d_j^*, whose membership function is maximal.

The proposed algorithm uses the idea of identifying the linguistic term by the maximum of membership function and generalizes this approach to the entire knowledge matrix. The computational part of this algorithm is easily realized by simply applying operations max and min.

To obtain a clear number $[\underline{y}, \overline{y}]$ from an interval that corresponds to a fuzzy value of the output variable, it is necessary to apply a dephasing operation. This clear number y^* can be determined, for example, by the center of gravity method:

$$y^* = \frac{\int_{Min}^{Max} y\mu(y)dy}{\int_{Min}^{Max} \mu(y)dy} \qquad (9)$$

Obviously, to implement the described approach to the processing of fuzzy information it is advisable to use declarative programming languages. These include Lisp, Prolog or their modifications depending on a specific objective problem) [13-15], since they are most successfully adapted for the realization of functions of the form (2), (3), (5) - (9) which can be both analytic, so and descriptive (functional, logic rules, fuzzy sets). Such software allows solving tasks related to qualitative recognition and analysis of complex structures (handwriting recognition, handwriting, psychophysiological state of the person, construction and analysis of storage systems, processing, information security, automated proofing of the theorems, environmental monitoring and decision-making, social protection and economic development).

V. Results of numerical Analysis. Note that results computer modeling and numerical experiments a proud efficient applied described methods and algorithms to analysis fuzzy information.

In particular, the task of analyzing the efficiency of the pharmaceutical turnover of a pharmaceutical company under conditions of limited financing and short terms on the supply of goods was considered. This approach proved to be effective, since it allows not only to minimize company expenses, but also to track the dynamics of turnover throughout the operating period.

One can propose a different approach to processing large volumes of fuzzy data under conditions of incomplete certainty of the vector of input variables (primary characteristics). The essence of the approach is based on conducting a simulation of the behavior of the investigated system and an expert assessment of the complementation of the existing knowledge base with new informative data and the establishment of the vector of input characteristics. Obviously, such an approach is iterative and it is necessary to take care of the convergence of the calculation process to achieve the goal with minimal cost and limited error.

When dealing with non-physical data in artificial intelligence problems, the construction of recognition systems, expert systems, medical and parametric diagnostics, the creation of logical-linguistic models, the most successfully adapted declarative programming languages. The logical statements and functional-logical dependencies provide the opportunity to describe the problem with fuzzy formulated data and obtaining solutions in the form of logical sequences, new functional dependences or probabilistic characteristics with definite mathematical hope and dispersion of the input signal. Ultimately, the initial vector of primary attributes should be refined, which will ensure reliable processing of fuzzy data.

VI. Conclusion and perspective forward researches. In the work, the authors proposed approaches to the processing of fuzzy information in the conditions of incomplete definition of the vector of input characteristics, which are based on the theory of fuzzy sets and measures, the construction of membership functions and the application of languages declarative programming.

The further development of the proposed approaches can be achieved by conducting statistical studies of specific application problems related to the need to assess both the quantity and value of information based on a figurative approach and theories of fuzzy sets.

The proposed approach is advisable to use in applied problems whose mathematical description is difficult or completely impossible. This approach will contribute to the development, in fact, of the methods of recognition theory and identification, processing of blurry data and the theory of information and coding.

1. Ueno X. *Introduction to knowledge engineering*. Tokyo: Omsya.- 1985. – 360 p.
2. Novak V., Perfiljeva I., Mochkorzh I., *Mathematical Principles of Fuzzy Logic*. M: Fizmalit, 2006. - 347 p.
3. Zadeh L.A."*Outline of a New Approach to the Analysis of Complex Systems and Decision Processes"*. IEEE. Trans. Syst. Van and Cybern, 1, 1973. - P. 28-44.
4. Konysheva L.K., Nazarov D.M. *Fundamentals of the theory of fuzzy sets*. M: Peter, 2011. - 190 p.
5. Syvalko M. *Mathematics of hidden possibilities*. Ostrog: Publishing House of the National University "Ostroh Academy", 2011. - 394 p.
6. Turksen I.B. Fuzzy expert systems / Ed. M. Zheleny - St. Petersburg: Peter, 2002. - 1120 p.
7. Shapiro D.I., Blishchun A.F. "*Choosing solutions with fuzzy state of the system"*. Algorithms and Programs, 1978, No. 1. - p.75.
8. Zimmermann H.J. *Fuzzy Set Theory and its Applications*. Kluner: Academic Publishers, Dordrecht, Boston, MA, 1991. - 315 p.
9. Dubois D. *Theory of Opportunity. Appendices to the representation of knowledge in computer science*. M.: Radio and Communication, 1990. - 288 p.
10. Rothstein O.P., Zhupanova MO, Sheverda V.M. "*Differential diagnosis of ischemic heart disease based on fuzzy logic*". Bulletin of the Ministry of Education and Science, No. 3, 1994. - P. 32-40.
11. Kryukov AP, Rodionov A.Ya. *The language of analytical computing*. Moscow: From Moscow State University, 1984. - 176 p.
12. Dyakonov V. *Matcad 2001*. SPB: Peter, 2001. - 624 pp.
13. Zaiats V.M. *Function programming*. Lviv: Publisher "Beskyd Beat", 2003. - 160 p.
14. C.In., D. Solomon. *Programming in Turbo Prolog*. M.: Svit, 1996. - 640 p.
15. Zaiats V.M., Zaiats M.M. Logical and functional programming. Tutorial. The Ministry of Education and Science of Ukraine (the list №1 / 11-16135 for 13 October 2014) provide Griff. Kamyanets-Podilsky: Publisher "Ruta". 2016. ISBN 978-617-738-138-8. - 400 p. (in Ukraine).
16. Zaiats V.M., Rybitska O. M The hidden possibilities of mathematics for the statistical processing of information // Materials International scientific and technical conf. "System Analysis and Information Technologies" (SAIT-2013), May 27-31, 2013, NTUU "KPI". Kiev, 2013. - P. 317-318.