N93-295207

161801

p. 8

THE AXIOMATIC DEFINITION OF A LINGUISTIC SCALE FUZZINESS DEGREE, ITS MAJOR PROPERTIES AND APPLICATIONS

Alexander P.Ryjov

Abstract.Model of human estimate of real objects as measuring procedure in fuzzy linguistic scales (FLS) is being considered in the report. The definition of FLS fuzziness degree and its major properties is given in the report. Definitions of information loses and noise while user works with data base (or knowledge base), containing linguistic description of objects are being introduced and described, and proven, that this value gives linear connection with degree of fuzziness.

Key words: estimate of real object, fuzzy linguistic scales, degree of fuzziness, quality of information search.

INTRODUCTIONS

Model of human estimate of real objects as measuring procedure in fuzzy linguistic scales (FLS) /1/ is being considered in the report. While describing objects some human being can't use any measuring devices, he makes it in terms of some sensible properties, and he has some doubts while giving some value to a property.

If there are a lot of property's values the trouble of choice is that there are some of them, which are "just equally" suitable for the object description. And if there are little of values the trouble is that all of them are "just equally" unsuitable to describe some object.

General study object of this works is a set of scale's value of a linguistic scale /1/. Example of scale's value for linguistic scale "Height" is given an Fig.1.

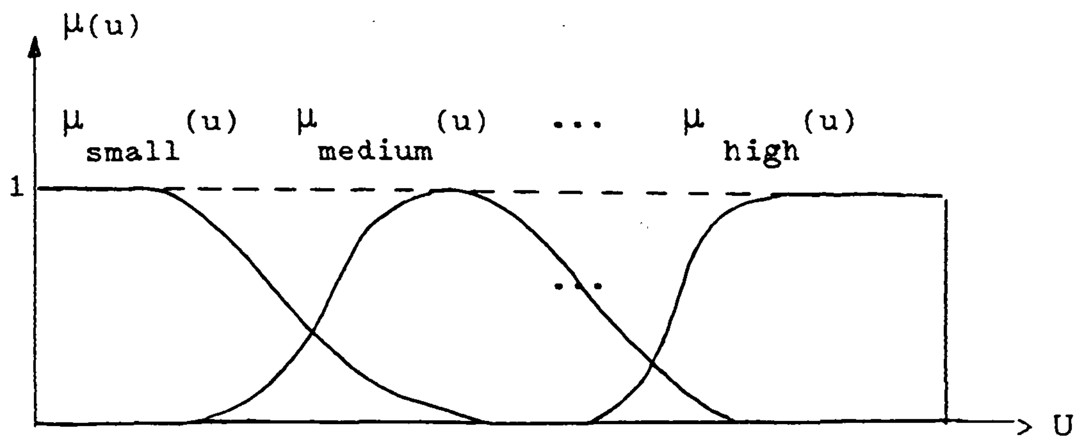

Fig.1

Such structures can be also interpreted as a set of different alternatives in problem solving and decision-making /2,3,4/ or a descriptions of classes in fuzzy classification and clustering /5,6/ or a representation of term-sets of linguistic values /7/ and etc. However, first interpretation (in the same way /8/) is the most preferable for application in information systems.

1. FLS FUZZINESS DEGREE: DEFINITION, EXAMPLE AND PROPERTY

The definition of FLS fuzziness degree is given in the paper under some matter-of-fact restrictions on membership function form, and the set of such functions, which create the FLS.

Let's assume, that membership functions for FLS l_t (where t - number of scale values) are defined on some segment $U \in \mathbb{R}^1$ and meets following requirements:

1) normal /9/: $\forall j \ (1 \le j \le t) \ \exists \ U_j^1 \ne \emptyset$,

where $U_j^1 = \{u \in U: \mu_j(u) = 1\}$, U_j^1 are segment;

2) increasing from the left U_j^1 and decreasing from the right U_j^1.

The requirements are quite natural for membership functions of notions gathered in some FLS's scale values set. Actually, the first means that there's at least one object for each scale value, which is typical or ideal for the notion; and the second may be interpreted as requirement of gradual changing of the notion limits.

Characteristic functions we'll be mentioned in the article. Let's assume, that :

3) those functions can have not more than two break points of second sort.

Let's assume that L is the set of functions satisfying requirements 1)- 3). The set L is a subset of a set of functions integral able on some measurable set of functions L_2, and therefore, a measure can be introduced on L. For example :

$$d(f,g) = \int_U |f(u) - g(u)| du, \ f \in L, \ g \in L.$$

Let's introduce some restrictions on a set of functions from L, which are creating a set value of FLS l_t. And let's assume that a set of such functions suit following requirements:

4) completeness: $\forall \ u \in U \quad \exists j \quad (1 \le j \le t): \ \mu_j(u) \ne 0$

5) orthogonally: $\forall\ u \in U \quad \sum_{j=1}^{t} \mu_j(u) = 1$

These restrictions are quite natural too. Assuming that 4) isn't true then a set $U' = \{u \in U: \forall j\ (1 \le j \le t)\ \mu_j(u) = 0\}$ may be harmlessly deleted from , therefore a set $U \setminus U'$ may be considered instead of universum. That means that there's no scale values associated with any point from U' set, and scale has improper definition.

Restriction 5) was described in /2/. Scales built under 5) are not only useful for theoretical analysis, but they must be the most spread in use, because the restrictions mean that:

- used notions (scale values) are quite differ from each other;
- they do not describe the same objects.

Let's call a set of FLS with scale values under 4) and 5) $G(L)$ - scales.

We can introduce a measure on $G(L)$ too.

Lemma 1. Let's assume that
$\{\hat{\mu}_1(u), \hat{\mu}_2(u), \ldots, \hat{\mu}_t(u)\}$ - a set of scale values $\hat{l}_t$;
$\{\mu_1(u), \mu_2(u), \ldots, \mu_t(u)\}$ - a set of scale values l_t;
$d(f,g)$- a measure in L .

Then $\rho(l_t, \hat{l}_t) = \sum_{i=1}^{t} d(\mu_i, \hat{\mu}_i)$ - is a measure in $G(L)$.

To formulate axioms we should define a scale, which is based on some FLS and is "unfuzzy", meaning that the scale's value is a set of characteristic functions, produced with membership functions of FLS.

Thus, assuming that $l_t \in G(L)$, is a FLS defined on U and consisting of membership functions $\hat{\mu}_1(u), \ldots, \hat{\mu}_t(u)$. Let's construct some "unfuzzy" set value l_t. l_t - is a set of characteristic functions $h_1(u), \ldots, h_t(u)$, where

$$h_i(u) = \begin{cases} 1, & \text{if } \max\limits_{1 \le j \le t} \mu_j(u) = \mu_i(u) \\ 0, & \text{otherwise} \end{cases}$$

Call $\hat{l}_t$ - the nearest "unfuzzy" scale, based on FLS $l_t \in G(L)$.

Let's assume that fuzziness degree of FLS, whose scale values are defined upon universum U, is the value of functional $\xi(l_t)$, defined on the membership function scale values set and satisfying following axioms:

A1. $0 \le \xi(l_t) \le 1 \ \forall \ l_t \in G(L)$.

A2. $\xi(l_t)=0 \Leftrightarrow \forall \ u \in U \ \exists i \ (1 \le i \le t): \mu_i(u)=1, \ \mu_j(u)=0 \ \forall \ j \ne i$.

A3. $\xi(l_t)=1 \Leftrightarrow \forall \ u \in U \ \exists i_1, i_2 \ (1 \le i_1, i_2 \le t)$:

$$\mu_{i_1}(u) = \mu_{i_2}(u) = \max_{1 \le j \le t} \mu_j(u).$$

A4. Let's assume that FLS l_t and $l'_{t'}$ are defined on universumes U and U' correspondingly; t and t' can be equal and not equal and not equal to each other.

$$\xi(l_t) \le \xi(l'_{t'}), \text{ if } \rho(l_t, \hat{l}_t) \le \rho(l'_{t'}, \hat{l}'_{t'}),$$

where $\rho(\cdot,\cdot)$ - some metric in $G(L)$.

Axiom A1 defines domain of values for functional $\xi(l_t)$, or fuzziness measuring borders.

Axioms A2 and A3 describes the scales where $\xi(l_t)$ assumes minimal and maximal values, or maximal "unfuzzy" and maximal "fuzzy" scales correspondent.

Axiom A4 defines the fuzziness degree comparison rule for each pare scales. It may be expressed in such a way: the nearer given FLS to its nearest unfuzzy scale, the less it's fuzziness degree.

Let's give an answer for question of existence a functional satisfying those axioms.

Theorem 1. Assume that $l_t \in G(L)$. Then functional

$$\xi(l_t) = \frac{1}{|U|} \int_U f(\mu_{i_1^*}(u) - \mu_{i_2^*}(u))du,$$

here $\mu_{i_1^*}(u) = \max_{1 \le j \le t} \mu_j(u)$, $\mu_{i_2^*}(u) = \max_{\substack{1 \le j \le t \\ j \ne i_1^*}} \mu_j(u)$,

f satisfies following requirement:

F1: $f(0) = 1$, $f(1) = 0$;

F2: f decreases,

is fuzziness degree l_t, i.e. satisfies A1 - A4.

It's easy to prove, that the only linear function satisfying F1, F2 is a function $f(x) = 1 - x$.

A subset of polynomials of degree 2, satisfying F1, F2, can be described. Those are expressions of the following type:

$$f_a(x) = ax^2 - (1 + a)x + 1.$$

Subset of functions of other types (logarithmic, trigonometric etc.) satisfying conditions F1, F2 may be defined in a similar way. Let's use those functions in formula for $\xi(l_t)$, and get some functionals, satisfying A1 - A4, i.e. it is a fuzziness degree.

FLS fuzziness degree properties for linear f are being described in the report. In this case

$$\xi(l_t) = \frac{1}{|U|}\int_U (1 - (\mu_{i_1^*}(u) - \mu_{i_2^*}(u)))du, \qquad (*)$$

here $\mu_{i_1^*}(u) = \max_{1 \le j \le t} \mu_j(u)$, $\mu_{i_2^*}(u) = \max_{\substack{1 \le j \le t \\ j \ne i_1^*}} \mu_j(u)$.

This fuzziness degree measurement functional was introduced at the first time in /10/ for the task of optimal quality properties values set choice in human-machine systems.

Let's define the following subset of function set L:

$\bar{L}$- a set of functions from L, which are part-linear and linear on

$$\tilde{U} = \{u \in U: \forall j \ (1 \le j \le t) \ \ 0 < \mu_j(u) < 1\};$$

$\hat{L}$ - a set of functions from L, which are part-linear on U (including $\tilde{U}$).

Theorem 2. Let $l_t \in G(\bar{L})$. Then $\xi(l_t) = \frac{d}{2|U|}$, where $d = |U^*| = |\{u \in U: \forall j \ (1 \le j \le t) \ \} \ \mu_j(u) \ne 1\}|$

Theorem 3. Let $l_t \in G(\hat{L})$. Then $\xi(l_t) = C\frac{d}{|U|}$, where $C < 1$, $C = \text{Const}$.

The fuzziness degree of a fuzzy set induced by $\xi(l_t)$ is defined as fuzziness degree of a trivial FLS, determined with a fuzzy set $\mu(u)$:

$$\xi(\mu) = \frac{1}{|U|}\int_U (1 - |2\mu(u) - 1|)du$$

It's easy proved, that $\xi(\mu)$ satisfies all the axioms for the set's fuzziness degree /11/. It may show that the introduced in the report more general notion $\xi(l_t)$ had been correctly defined.

It's easy shown, that the functional may be considered as an average human doubts degree while describing some real object (situations) /4,12/.

2. APPLICATION FLS FUZZINESS DEGREE TO INFORMATION SEARCH

The results were published at the first time in /11/.

Definitions of information loses and noise while user works with data base, containing linguistic description of objects are being introduced and described in the report. While interacting with the system user formulates his query and gets an answer according to the search request. And if he knew real (not linguistic) values of object characteristics, he, possibly, would defeat some of displayed objects (noise) and he would add some others from data base (loses). Information noise and losses appear because of fuzziness of scale elements.

Because of volume restrictions and taking into account the illustrative character of the chapter we stop at the main results. In the next work we are going to describe the problems of formalization of fuzzy database information retrieval quality rations in complete.

Theorem 4. Assume that $l_t \in G(\bar{L})$, $\xi(l_t)$ - degree of fuzziness of l_t; $\Pi_x(U)$, $H_x(U)$ - average information loses and noise, appearing during information search with search attribute value set X, equal to l_t -scale values set; U - universum l_t; N(u) - number of objects, whose definitions are in a database and which having a real characteristic value equal to u, - is a constant. Furthermore, assume that all of property values are equally preferable for user, meaning that request probabilities for all the property values are equal. Then

$$\Pi_x(U) = H_x(U) = \frac{2N}{3t}\,\xi(l_t),\ N = \text{Const}.$$

Theorem 5. Assume that $l_t \in G(L)$, $N(u) = N = \text{Const}$ and request probabilities for all the property values are equal. Then

$$\Pi_x(U) = H_x(U) = \frac{c}{t}\xi(\mu),$$

where c - a constant, which depends on N only.

Thus, α% - fuzziness degree decrease leads to the same decrease of average information loses and noise if the number of property values is constant. Simultaneous fuzziness degree decrease of properties values number lead to even more substantial decrease of information loses and noise.

The following method of property values set choosing for fuzzy databases, can be evaluated from the given results:

1. To generate all possible sets of property values.
2. To represent each of with FLS scale values set.
3. To evaluate the degree of fuzziness for each of the property values sets according to (*).

4. Chose the set of property value set, which has the minimal ratio of fuzziness degree and number of elements. Your choice will provide the minimal information loses and noise of information retrieval using the property.

CONCLUSIONS

Some method to calk the fuzziness degree of the combination of fuzzy sets (defined upon the same universum) has been given in the article. The axioms for such measure of uncertainty have been formulated, its interpretation has been given. The theorem of existence has been proven and some properties of fuzziness degree have been described.

The problems of using of the results in information applications (fuzzy retrieval systems) have been discussed. It is described that the fuzziness degree has linear dependence with the indicator of retrieval quality. Taking into account the result the methodic of choosing the optimal values has been suggested. Using the method some user may describe objects to achieve better results of finding information in fuzzy data bases. Under these circumstances a person - a source of information - would suffer minimal difficulties (uncertainties) to describe real objects.

The results may be used also in some tasks to construct knowledge bases, decision-making tasks under fuzzy conditions and pattern recognition.

References.

1. Decision-making models on the base of linguistic variable / Ed. by Borisov A.N. and etc. - Riga: Zinatne, 1982 (in Russian).

2. Fuzzy sets in models of control and artificial intelligence / Ed. by Pospelov D.A. - Moscow: Nauka, 1986 (in Russian).

3. Capocelli R., De Luca A. Fuzzy sets and decision theory // Information and Control. - 1973. - v.23. - p. 43 - 50.

4. Ryjov A. About a method of fuzzy object and situation optimal description in decision-making systems // International symposium on fuzzy approach to reasoning and decision-making. Abstracts. - Technical university Liptovsky Nikolach, 1990. - p. 81.

5. Bezdek J.S. Numerical taxonomy with fuzzy sets // Journal of Mathematical Biology. - 1974. - v.1. - p. 57 - 71.

6. De Luca A., Termini S. Measures of ambiguity in the analysis of complex systems // Lectures Notes on Computer Sciences. - 1977. - v. 53. - p. 58 - 73.

7. Zadeh L.A.The concept of a linguistic variable and its application to approximate reasoning, Part 1,2,3, //Inform.-Sci.8,199-249;8,301-357;9,43-80(1975).

8. Ryjov A.P. The degree of fuzziness of linguistic scale and its properties // Fuzzy sets in informatics: Abstracts of Moscow International Conference. - Moscow, September 20 - 23. - Moscow, 1988. - p. 57.

9. Kaufmann A. Introduction to the the theory of fuzzy subset, v.1 - N.Y.: Academic Press, 1975. - 643 p.

10. Ryjov A. About degree of fuzziness of fuzzy characteristics // VII All-Union conference on problem of theoretical cybernetics. September 18, 1985 - September 20, 1985. Abstracts.Part I - Irkutsk,1985. - p. 53-55 (in Russian).

11. Ryjov A. The degree of fuzziness of fuzzy descriptions // Mathematical cybernetics and its application to biology / Ed. by Krushinsky L.V., Yablonsky S.V., Lupanov O.B. - Moscow, Moscow University Publiching, 1987. - p.60-77 (in Russian).

12. Ryjov A., Averkin A. The Axiomatic definition of a linguistic scale fuzziness degree and its general properties // II All-Union conference "Artificial intelligence - 90". Abstracts. Part I. - Minsk,1990. - p.162-165 (in Russian).