The Third International Conference "Problems of Cybernetics and Informatics"
September 6-8, 2010, Baku, Azerbaijan. Section #3 "Modeling and Identification"
www.pci2010.science.az/3/31.pdf

ON THE BASIS OF FUZZY LOGIC IN A BROAD SENSE THE CHOICE OF SUPPORTING SUBSETS IN FUZZY CLUSTER ANALYSIS

Adalat Karimov[1], Ulviyya Rzayeva[2]

Azerbaijan State Economic University, Baku, Azerbaijan
[1]*adalat_kerim@mail.ru*, [2]*ulyarzayeva@rambler.ru*

The problem of modeling by syntagmas of supporting subsets of objects with fuzzy signs, accepting linguistic values is considered. For the decision the fuzzy logic in a broad sense (*FLb*) *is used. The special class of linguistic syntagmas in FLb is defined and the set of linguistic syntagmas is under construction.*

The majority of various mathematical decision methods of pattern recognition problems breaks up on two groups, one of which can be treated from positions of the theory of decisions (the discriminate approach), and another - within the limits of syntactic (structural or logic) the approach. Pattern recognition (reference of each object to some class) is usually carried out by means of partition of space of signs into areas. Development of researches on pattern recognition for last decade has been mostly connected with the discriminate approach and its applications. The structural approach is applied to problems of pattern recognition in which the information describing structure of each object is important. And from recognition procedure it is required giving the chance not only to carry object to a certain class (to classify it), but also to describe those parties of object which exclude its reference to other class.

Logic methods of recognition have found application in the decision of practical problems from the most different areas of human activity: biology, medicine, military science, the nuclear physics etc. It is possible to consider system of logic pattern recognition consisting of three supporting parts, namely: from the block of preprocessing, the block of the description or representation of object and the block of syntactic analysis [4].

Specificity of the problems concerning various areas, is shown mainly at a stage of formalization and construction of mathematical model of a problem. To this day there is no criterion for an establishment of completeness of space of signs, therefore for a true establishment as much as possible signs undertake, then synthesis of signs their optimum quantity gets out. It is known, that the information contains not only in separate signs, generally it contains in their combinations (informative combinations) [1]. Purpose of this paper consists in modeling of information combinations by means of syntagmas within the limits of FLb.

Any law classifies only some part of objects. Having united certain quantity of laws in a composition, it is possible to receive the algorithm, capable to classify any objects. We will give the formal description of algorithm of estimations calculation (AEC) as algorithm of classification on the basis of laws. The principle of classification applied in AEC, is called also as a principle partial precedent - the object x concerns that class in which is available more objects close to x by informative parts of feature descriptions (supporting subsets).

1. Distance functions ρs are set: $X \times X \to R_+$, $s = 1..., n,$ generally as

$$\rho s\ (x, x') = |fs\ (x) - fs\ (x')\ |,\ s = 1..., n, \qquad (1)$$

where $f1..., fn$ are descriptions of objects by some signs.

2. The system of supporting subsets is set

$$\Omega = \{\omega \mid \omega \subseteq \{1,...,n\}\}$$

3. The binary threshold function of affinity estimating similarity steams of objects is entered x, $x' \in X$ on supporting subset $\omega \subset \Omega$,

$$B_\omega(x,x') = \bigwedge_{s\in\omega}\left[\rho_s(x,x')\right] \le \varepsilon_s$$

где ε_s - non-negative numbers named thresholds, *s = 1..., n*. In (1) threshold ε_s names accuracy of measurement of a sign *fs*.

4. The estimation of affinity of object $x \in X$ to a class $c \in Y$ as result of the weighed voting nearness of object *x* to all objects of a class *c* on all supporting subsets is entered:

$$\Gamma_c(x) = \sum_{i:y_i=c\omega\subset\Omega}\sum a_{\omega_i} B_\omega(x_i,x)$$

where weight a_{ω_i} are assumed normalized on unity:

$$\sum_{i:y_i=c\omega\subset\Omega}\sum a_{\omega_i} = 1$$

5. The algorithm of classification *A(x)* carries object *x* to that class which has typed the greatest sum of voices:

$$a(x) = \arg\max_{c\in Y}\Gamma_c(x)$$

So, the algorithm of calculation of estimations is set by system supporting subsets Ω, thresholds ε_s, *s = 1, ..., n* and scales $\alpha_{\omega i}$, $\omega \subset \Omega$, *i = 1..., ℓ*. [5]

Definition. The syntagma is a combination of two sentence parts, the connected themes or other relation with an unequal orientation of members where one member is defined, and another - defining.

Between members of syntagmas there can be different relations from which the *predicative* is the main and expresses dependence of two members with obligatory communication of time and an inclination.

Let statements (syntagmas) $Z_1,...,Z_l$ mean any properties of distinguished objects, for example various illnesses, latent defects and malfunctions in the difficult mechanism etc. Let statements $X_1,...,X_n$ - registered in experiment a sign, by their certain sets (informative combinations) it is possible to establish presence of one and absence of other properties of distinguished objects. Aprioristic data on them expressing, on the one hand, communication between statements $Z_1,...,Z_l$ и $X_1,...,X_n$, with another - dependence only between elements $Z_1,...,Z_l$ or only between elements $X_1,...,X_n$ in general case can be presented in the form of one parity of equivalence:

$$E(X_1,...,X_n;Z_1,...,Z_l) = 1 \qquad (2)$$

where *E* is certain Bool function.

The recognition observable object is to establishment, leaning against aprioristic dependences (2) and experimental data $X_1,...,X_n$, which properties among $Z_1,...,Z_l$ belongs or does not belong to this object. According to that properties of objects are characterized by *l* elements Z_j, *j=1, ..., l*, there can be the greatest 2^l various on the properties of objects' types. If each type of object to consider as a class and to designate

$$\Omega_1 = Z_1 \cdot Z_2 \cdot ... \cdot Z_l,$$
$$\Omega_2 = \overline{Z}_1 \cdot Z_2 \cdot ... \cdot Z_l,$$
$$\text{… … … … … … … ….}$$
$$\Omega_2{}^l = \overline{Z}_1 \cdot \overline{Z}_2 \cdot ... \cdot \overline{Z}_l,$$

then we will come to standard statement of a problem of recognition: on the basis of dependence of a kind (2) and informative combinations $X_1,...,X_n$ which also are supporting

subsets, to define, to what class $\Omega_1, \Omega_2..., \Omega_2^l$ the given object concerns. We will assume, that as a result of experiment (supervision of object, measurements of parameters, etc.) some data concerning values of the validity of combinations $X_1,...,X_n$ by which the recognized object is characterized have been obtained, and these data are presented as Bool function

$$G(X_1,...,X_n) = 1 \tag{3}$$

Methodically the decision of a problem of recognition in the resulted statement is reduced to the consecutive decision of some typical problems which is based on the determined solving rule (3) [2].

Number of supporting subsets of capacity q of signs' sets equally C_N^q. In this paper considered number of combinations of the supporting sets modeled by syntagmas, no more, than C_N^3, that is $q \leq 3$.

Let S_1, S_2, S_3 are system of supporting sets. Sizes

$$\begin{aligned}
&\Gamma(\omega, \Omega_1) = \Gamma_{S_1}(\omega, \Omega_1) + \Gamma_{S_2}(\omega, \Omega_1) + ... + \Gamma_{S_l}(\omega, \Omega_1) = \sum_{S_A} \Gamma(\omega, \Omega_1) \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&\Gamma(\omega, \Omega_m) = \Gamma_{S_1}(\omega, \Omega_m) + \Gamma_{S_2}(\omega, \Omega_m) + ... + \Gamma_{S_l}(\omega, \Omega_m) = \sum_{S_A} \Gamma(\omega, \Omega_m)
\end{aligned} \tag{4}$$

represent functions of an accessory of object ω to corresponding class $\Omega_1, ..., \Omega_m$. Then in our case the number summands in the formula defining quantity $\Gamma_{S_A}(\omega, \Omega_i)$ (4), is equal $(r_i - r_{i-1})$ C_N^3 [2].

The problem of objects' clusterization in the basis of fuzzy logic in a broad sense and construction of clusterization algorithm in terms of linguistic syntagmas is completely opened in [3].

References

1. Dorodnitsin A.A. Problems of mathematical modeling in descriptive sciences - Cybernetics, 1983, №4, pp. 6-10.
2. Gorelik A.L., Gurevich I.B., Skripkin V. A. A modern state of recognition problems. Moscow, «Radio and communication», 1985, pp. 77-86.
3. Kerimov A.K., Rzayeva U.Sh. About one problem of objects fuzzy clusterization on the basis of fuzzy logic in a broad sense. The Materials of the Third Congress of the World Mathematical Society of Turkic Countries, Almaty, 2009, pp. 27-32.
4. Novikova N.M. Structural recognition pattern. 2008.
5. Vorontsov K.V. Lectures on logic algorithms of classification, http://www.ccas.ru/voron/download/LogicAlgs.pdf