DOI: 10.5604/20830157.1121328

ACTUALIZATION OF THE DISTRIBUTED KNOWLEDGE BASE OF ERGATIC SYSTEM USING THE METHOD OF FUZZY CLASSIFICATION

Viktor Perederii, Eugene Borchik
Nikolaev Sukhomlinskii National University, Department of Computer Systems and Networks

Abstract: *In the article a method of actualization the distributed knowledge base of ergatic system using the method of fuzzy classification is proposed. As an example we consider the request choice formation of an alternative of decision-making from the knowledge base, according to the values of the input parameters. Genetic algorithm is used for finding optimal solutions. For automation of calculations MATLAB software package was used.*

Keywords: knowledge base, fuzzy classification, membership functions, objects

AKTUALIZACJA ROZPROSZONEJ BAZY WIEDZY SYSTEMU ERGATYCZNEGO ZA POMOCĄ METODY KLASYFIKACJI ROZMYTEJ

Streszczenie*. W pracy zaproponowano metodę aktualizacji rozproszonej bazy wiedzy systemu ergatycznego (system maszyna-człowiek) używając rozmytej klasyfikacji. Rozważono przykłady formułowania zapytań, wybór alternatywnych decyzji z bazy wiedzy, zgodnie z wartościami parametrów wejściowych. Celem znalezienia optymalnych rozwiązań zastosowano algorytmy genetyczne. Do automatyzacji obliczeń zastosowano pakiet MATLAB.*

Słowa kluczowe: baza wiedzy, klasyfikacja rozmyta, funkcje przynależności, obiekty

Introduction

One of the most important components of ergatic decision support systems, at the management of complex technical objects, is a Knowledge Base (KB), which is realized as a special kind of database, developed for operating knowledge. KBs should contain structured information covering some area of knowledge for using with a specific purpose. Modern KBs contain not only factual information, but also the rules of inference, reasoning about newly inputted facts, meaningful information processing but also work together with the information retrieval systems and have classification structure and format of knowledge representation.

The most important and time-consuming affair in creating knowledge base is to support its relevance. Considered in [2-5, 8] the ways of supporting the relevance of KB in our opinion are quite complex, involving technical knowledge, moreover, presupposes the existence of large arrays of address databases received from various sources. In [1, 7] there are considered the methods of examination and diagnosis in order to maintain the relevance of KB.

All considered expert methods are widely used and is described in detail in the modern literature.

The disadvantages of expert methods are subjectivism, the limited application, the high costs of their conduct, so these methods are appropriate to use at the initial stage of filling KB.

This paper proposes a method for updating knowledge base to build alternatives to decision-making in real time, taking into account the cognitive state of users based on fuzzy classification method.

1. Methods

In a distributed KB there is stored information in the form of alternatives, which maintains records of direct and indirect data influencing the process of making relevant decisions. Full information structure of formation of the alternatives presented in figure 1, where:

- X1 - information on the characteristics of the external environment, which directly affects DM (noise, temperature, low-frequency vibration, light level, etc).
- X2 - information about psychological characteristics of DM (test results, physiological state of DM, pulse, pressure, etc.);
- X3 - information about technological process, as well the range of deviations.

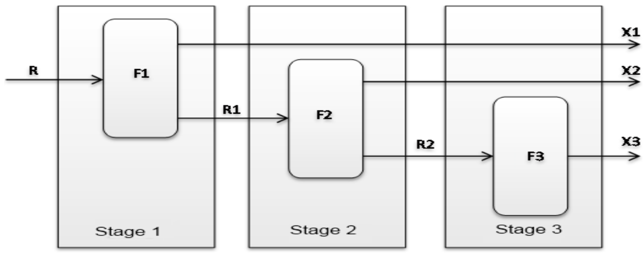

Fig. 1. The information structure of formation the alternatives. Request for the alternative of accepting solution «R» passes through three stages

The first stage is a function (F1) of separation from the main flow of the set of "X1", which carries information about the characteristics of the external environment impact on DM. This function handles the flow of R into variable flow R1 and an array of values X1. Function (F1):

$$X1 = R / R1 \quad (1)$$

The second stage is a function (F2) of separation from variable flow R1 of the set of "X2", in which there is stored information about the psychological characteristics of DM. This function handles the thread R1 into variable flow R2 and the array value of X2. Function (F2):

$$X2 = R1 / R2 \quad (2)$$

The third stage is represented by the function (F3) of converting the variable flow R2 in into the set of "X3", in which there is stored information about the technological process. This function handles the flow of R2 into the array value of X3. Function (F3):

$$X3 = R2 \quad (3)$$

Consider the example of choice on request of Ri from KB alternative decision-making, in accordance with the established dependencies, from the values of the parameters x_1, x_2, x_3, and which may vary in limits, presented in Table 1.

It is required to select an alternative for any set of parameter values x_1, x_2, x_3, wherein the values of some or all of the parameters do not belong the intervals, specified in Table 1.

To solve this problem we use the method of fuzzy classification [8]. Classification problem consists in assigning an object, specified by the vector of informative signs $V = (x_1, x_2, ..., x_n)$, to one of advance certain classes $\{A_1, A_2, ..., A_m\}$, that is, consists in performing of mapping the form: $V=(x_1, x_2,\dots, x_n)\rightarrow U \in \{A_1,A_2,\dots,A_m\}$.

Table 1. Limits of variation of the input parameters

x_1	x_2	x_3	A_i
[10,15]	[1,2]	[80,120]	A_1
[10,15]	[7,8]	[80,140]	A_2
[10,15]	[11,13]	[80,140]	A_3
[10,15]	[1,2]	[250,320]	A_4
[10,15]	[7,8]	[250,320]	A_5
[10,15]	[11,13]	[250,320]	A_6
[10,15]	[1,2]	[420,500]	A_7
[10,15]	[7,8]	[420,500]	A_8
[10,15]	[11,13]	[420,500]	A_9
[30,40]	[1,2]	[80,120]	A_{10}
[30,40]	[7,8]	[80,120]	A_{11}
[30,40]	[11,13]	[80,120]	A_{12}
[30,40]	[1,2]	[250,320]	A_{13}
[30,40]	[7,8]	[250,320]	A_{14}
[30,40]	[11,13]	[250,320]	A_{15}
[30,40]	[1,2]	[420,500]	A_{16}
[30,40]	[7,8]	[420,500]	A_{17}
[30,40]	[11,13]	[420,500]	A_{18}
[60,70]	[1,2]	[80,120]	A_{19}
[60,70]	[7,8]	[80,120]	A_{20}
[60,70]	[11,13]	[80,120]	A_{21}
[60,70]	[1,2]	[250,320]	A_{22}
[60,70]	[7,8]	[250,320]	A_{23}
[60,70]	[11,13]	[250,320]	A_{24}
[60,70]	[1,2]	[420,500]	A_{25}
[60,70]	[7,8]	[420,500]	A_{26}
[60,70]	[11,13]	[420,500]	A_{27}

Classification based on fuzzy inference is made on the knowledge base in the form:

$$\bigcap_{i=1}^{n} x_i = \tilde{a}_{ij} \ \text{ with weight } \ w_j \to u_j = A_j, \quad j = \overline{1,m}, \qquad (4)$$

where $u_j \in \{A_1, A_2, \ldots, A_m\}$ – consequent value of j-th rule; $\tilde{a}_{ij}$ - a fuzzy term, valuating the criteria $x_j (i = \overline{1,n})$ in the j-th rule.

The degree of membership of classification object, informative characteristics is given by a vector $V^* = (x_1^*, x_2^*, \ldots, x_n^*)$, classes A_j from knowledge base, is calculated as:

$$\mu_{A_j}(V^*) = w_j \wedge_{i=\overline{1,n}} [\mu_j(x_i^*)], \ j = \overline{1,m}, \qquad (5)$$

where $\mu_j(x_i^*)$ - is a degree of membership values x_i^* of fuzzy therms $\tilde{a}_{ij}$; $\wedge$ - operation of finding the minimum.

As a solution, there is selected class with a maximum degree of membership:

$$U^* = \underset{\{A_1, A_2, \ldots, A_m\}}{\arg} \max(\mu_{A_1}(V^*), \mu_{A_2}(V^*), \ldots, \mu_{A_m}(V^*)) \qquad (6)$$

In the problem under consideration, vector of informative features of object classification is V=(x_1, x_2, x_3), and alternatives $A_1, A_2, \ldots, A_{27}$ - classes of solutions. The fuzzy knowledge base of mapping $V \to U \in \{A_1, A_2, \ldots, A_{27}\}$ we build, taking into account dependency, presented in Table 2, where by L, A, H are denoted the terms "low", "average", "high".

Table 2. The fuzzy knowledge base of mapping

If X_1=L and X_2=L and X_3=L then $u_1=A_1$;
If X_1=L and X_2=A and X_3=L then $u_2=A_2$;
If X_1=L and X_2=H and X_3=L then $u_3=A_3$;
If X_1=L and X_2=L and X_3=A then $u_4=A_4$;
If X_1=L and X_2=A and X_3=A then $u_5=A_5$;
If X_1=L and X_2=H and X_3=A then $u_6=A_6$;
If X_1=A and X_2=L and X_3=H then $u_7=A_7$;
If X_1=A and X_2=A and X_3=H then $u_8=A_8$;
If X_1=A and X_2=H and X_3=H then $u_9=A_9$;
If X_1=A and X_2=L and X_3=L then $u_{10}=A_{10}$;
If X_1=A and X_2=A and X_3=L then $u_{11}=A_{11}$;
If X_1=A and X_2=H and X_3=L then $u_{12}=A_{12}$;
If X_1=A and X_2=L and X_3=A then $u_{13}=A_{13}$;
If X_1=A and X_2=A and X_3=A then $u_{14}=A_{14}$;
If X_1=A and X_2=H and X_3=A then $u_{15}=A_{15}$;
If X_1=A and X_2=L and X_3=H then $u_{16}=A_{16}$;
If X_1=A and X_2=A and X_3=H then $u_{17}=A_{17}$;
If X_1=A and X_2=H and X_3=H then $u_{18}=A_{18}$;
If X_1=H and X_2=L and X_3=L then $u_{19}=A_{19}$;
If X_1=H and X_2=A and X_3=L then $u_{20}=A_{20}$;
If X_1=H and X_2=H and X_3=L then $u_{21}=A_{21}$;
If X_1=H and X_2=L and X_3=A then $u_{22}=A_{22}$;
If X_1=H and X_2=A and X_3=A then $u_{23}=A_{23}$;
If X_1=H and X_2=H and X_3=A then $u_{24}=A_{24}$;
If X_1=H and X_2=L and X_3=H then $u_{25}=A_{25}$;
If X_1=H and X_2=A and X_3=H then $u_{26}=A_{26}$;
If X_1=H and X_2=H and X_3=H then $u_{27}=A_{27}$;

Membership functions of terms of the input variable x_1 are shown in Figure 2.

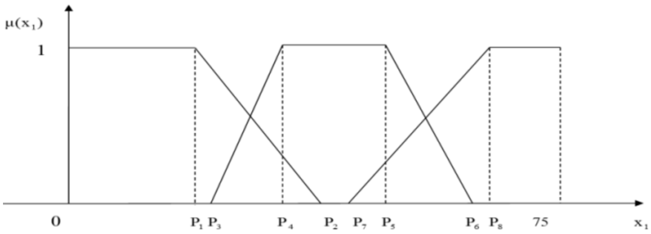

Fig. 2. Membership functions of therms of the input variable x_1

The membership functions of the terms of input variables x_2, x_3, we will choose similar to the corresponding functions of the variable x_1. This will reduce the total number of parameters of membership functions of the terms of all the variables from 24 to 8, which is significantly while minimizing function (7).

As a criterion of training fuzzy classifier [8] let us choose the simplest criterion:

$$\frac{100\%}{N}\sum_{k=1}^{N}\Delta_k(P) \rightarrow min, \quad (7)$$

where:

$\Delta_k(P) = \begin{cases} 1, & \text{if } u_k \neq F(P,V_k) \\ 0, & \text{if } u_k = F(P,V_k) \end{cases}$ an error of classification of the object V_k;

N - the number of pairs of input-output (V_k, u_k), $k=\overline{1,N}$ training sample;

P is the vector of the parameters of the membership function of the fuzzy terms of knowledge base (1);

F(P, V_k) is the result of the classification on the fuzzy basis with parameters P if the input value is V_k.

Training fuzzy classifier, therefore, is to find the vector P that minimizes the distance between the results of the logical inference and experimental data from the sample (V_k , u_k).

When training fuzzy classifier in considered task, do the following. First, select the parameter vector membership function of a fuzzy terms $\mu(x_1)$ as follows: P_0=(15,25,15,30,45,60,50,60).

Choose from the k-th row of table 1 random values x_1, x_2, x_3, belonging to the respective intervals of changing the values of variables in the considered string. Get input vector V_k. According to Table 1, it belongs to the class u_k=A_k. Considering all rows we will receive 27 pairs of "input-output" (V_k, u_k), $k=\overline{1,27}$ training sample.

Using (2), (3), produce a classification based on fuzzy input data V_k. The calculation is made in MATLAB environment; assume the weights w_j equal to zero. While classification result F (P, V_k) on the fuzzy basis with the parameters P_0 when the input value V_k differs from a clear exit u_k 15% of the trials.

To increase the accuracy of the fuzzy classification we will use the criterion of (4).

As the minimized function is an integer, the most appropriate is the genetic algorithm for finding extreme. For its realization we use "gatool" MatLAB function.

Using components of the vector P_0 for setting in "gatool" the limits of changing the function arguments (4), we obtain the solution

$$P=(18{,}53;23{,}665;14{,}15;34{,}876;45{,}63;59{,}72;46{,}44;58{,}88) \quad (8)$$

Now differences between the results of fuzzy classification and a clear exit are not observed.

Below are the results of the fuzzy selection of alternatives for the vector of parameters of membership functions (5) and some input values x_1, x_2 and x_3:

Table 3. The results of calculations of forming alternatives

x_1	19	22	21	22,5	22,3	22,1	22	2
x_2	6	6	5	7	13	2	4,2	13
x_3	420	400	410	410	385	380	430	380
A_i	A_8	A_{14}	A_5	A_{14} or A_{17}	A_{15}	A_{13}	A_{13} or A_{16}	A_6

In the case, for example, when V = (22, 5; 7, 410) selection of alternatives is ambiguous: either A_{14} or A_{17}. As a solution we should accept the new formed alternative that after assessing the relevance will be recorded into a knowledge base:

$$A^* = A_{14} \cup A_{17}.$$

2. Results

The proposed method allows increasing the efficiency of the process of actualization the distributed knowledge base through automation fuzzy classification components of alternatives in multilevel ergatic decision support systems in the management of complex technical objects in the real time.

References

[1] Asuncion A., Newman D.J.: UCI Machine Learning Repository. Irvine (USA), University of California, School of Information and Computer Science, 2007.

[2] Kovalenko I.I.: Koncepciya postroeniya sistemnyx informacionnyx texnologij podderzhki ekspertnogo analiza scenariev [The concept of building system information technologies of support expert analysis of scenarios]. Kovalenko I.I., Perederij V.I., Shved A.V.: Sistemni texnologiї [System technologies]. Regional interuniversity collection of scientific works, 6 (71), 2010, p. 74–88 (in Russian).

[3] Perederij V. I.: Zastosuvannya merezhi Bajesa z ocinki stupenya znachimosti vplivayuchix faktoriv na operatora v avtomatizovanix sistemax prijnyattya relevantnix rishen [Application of Bayesian network assess the significance of the influencing factors on the operator in automated systems relevant decision-making]. Perederij V.I., Litvinenko V.I.: Komp'yuterni nauki ta informacijni texnologiї [Computer Science and Information Technology]. Bulletin of the National University "Lviv Polytechnic". 733/2012, pp. 120-128 (in Ukrainian).

[4] Rotshtejn A.P.: Intellektualnye texnologii identifikacii: nechetkaya logika, geneticheskie algoritmy, nejronnye seti [Intellectual identification technologies: fuzzy logic, genetic algorithms, neural networks]. Universum, Vinnitsa 1999 (in Russian).

[5] Sapronov Yu.G.: Ekspertiza i diagnostika obektov i sistem servisa [Examination and diagnostics of objects and systems of service]. Academia, Moscow 2008 (in Russian).

[6] Shtovba S.D.: Vvedenie v teoriyu nechetkix mnozhestv i nechetkuyu logiku [Introduction to the theory of fuzzy sets and fuzzy logic]. http://matlab.exponenta.ru/fuzzylogic/book1, (in Russian).

[7] Uzga-Rebrovs O.: Nenoteiktibu parvaldisana 3 [Uncertainty in management 3] Rezekne, 2010 (in Latvian).

[8] matlab.exponenta.ru/fuzzylogic/book1/1_7_5_5.php

Ph.D. Viktor Perederiy
e-mail: viperkms@yandex.ua

Ph.D, assistant professor, head of Department of Computer Systems and Networking, Nikolaev National V.A. Sukhomlinsky University, Ukraine, Nikolaev, Nikolskaya str., 24, tel.: 050-9516143.

Ph.D. Eugene Borchik
e-mail: borchikeu@gmail.com

Ph.D., assistant professor, Nikolaev National V.A. Sukhomlinsky University, Ukraine, Nikolaev, Nikolskaya str., 24, tel.0972305264.

otrzymano/received: 01.06.2014 *przyjęto do druku/accepted: 23.06.2014*