

Fuzzy Systems of Logical Inference and Their Applications

Oleksandr Provotar

University of Rzeszow, Rzeszow, 35-310, Poland
aprowata1@bigmir.net,

Abstract. The approaches to the solution of various problems of artificial intelligence methods are proposed. In particular, the problem of knowledge representation by means of fuzzy specifications in expert systems, the problem of recognizing the structures of the proteins of different organization levels and the problem of building linguistic models in fuzzy Boolean variables logic are considered. All methods are based on the ideas of inductive mathematics. To investigate a reliability of these methods is possible only with the help of the theory of probability or possibility theory.

Key words: fuzzy sets, problem of recognizing, fuzzy Boolean variable

1 Introduction

It is known that the fuzzy sets [2], [5], [6], [7], [8] are the convenient tool to present knowledge in information systems. Using the fuzzy sets it is possible to outline, for instance, the picture of symptoms of the patient in expert diagnostics systems. Determination of the diagnosis in such systems requires using the mechanisms of logical inference. In particular, in case of the fuzzy specifications as symptoms as well diagnostics such mechanisms can be so called the fuzzy systems of logical inference which are built on the basis of ideas and methods of inductive mathematics [1].

The fuzzy specification of problem means ordered set of fuzzy instructions (fuzzy rules, linguistic rules). The fuzzy specification of the problem with the algorithm during fulfilling which the approximate (fuzzy) solution of the problem is received will be called as fuzzy system of logical inference.

Let x_1, \dots, x_n are input linguistic variables and y – output linguistic variable [6]. The ordered set of fuzzy instructions looks like as following

if x_1 *is* $A_{11} \wedge \dots \wedge x_n \in A_{1n}$ **then** y *is* B_1 ,
if x_1 *is* $A_{21} \wedge \dots \wedge x_n \in A_{2n}$ **then** y *is* B_2 ,
.....
if x_1 *is* $A_{m1} \wedge \dots \wedge x_n \in A_{mn}$ **then** y *is* B_m ,

where A_{ij} and B_i – fuzzy sets, symbol “ \wedge ” is interpreted as t -norm of fuzzy sets.

The algorithm of calculating the output of such specification under the inputs A'_1, \dots, A'_n consists in performing such steps:

1. Calculate the truth level of the rules:

$$\alpha_i = \min \left[\max \left(A'_1(x_1) \wedge A_{i1}(x_1) \right), \dots, \max \left(A'_n(x_n) \wedge A_{in}(x_n) \right) \right];$$

2. Calculate outputs of each rule:

$$B'_i(y) = \min(\alpha_i, B_i(y));$$

3. Calculate aggregated output:

$$B(y) = \max \left(B'_1(y), \dots, B'_m(y) \right).$$

2 Expert Diagnostics Systems

Let $X_1 = \{5, 10, 15, 20\}$, $X_2 = \{5, 10, 15, 20\}$, $X_3 = \{35, 36, 37, 38, 39, 40\}$ – spaces for determining the values of linguistic variables:

$$\begin{aligned} x_1 &= \text{“Coughing”} = \{\text{“weak”}, \text{“moderate”}, \text{“strong”}\}, \\ x_2 &= \text{“Runningnose”} = \{\text{“weak”}, \text{“moderate”}, \text{“strong”}\}, \\ x_3 &= \text{“Temperature”} = \{\text{“normal”}, \text{“raised”}, \text{“high”}, \text{“veryhigh”}\}, \end{aligned}$$

accordingly.

Determine the elements of these sets:

$$\begin{aligned} \text{“Coughing”} : \text{“weak”} &= 1/5 + 0.5/10; \\ &\text{“moderate”} = 0.5/5 + 0.7/10 + 1/15; \\ &\text{“strong”} = 0.5/10 + 0.7/15 + 1/20. \\ \text{“Runningnose”} : \text{“weak”} &= 1/5 + 0.5/10; \\ &\text{“moderate”} = 0.5/10 + 1/15; \\ &\text{“strong”} = 0.7/15 + 1/20. \\ \text{“Temperature”} : \text{“normal”} &= 0.5/35 + 0.8/36 + 0.9/37 + 0.5/38; \\ &\text{“raised”} = 0.5/37 + 1/38; \\ &\text{“high”} = 0.5/38 + 1/39; \\ &\text{“veryhigh”} = 0.8/39 + 1/40. \end{aligned}$$

Let $Y = \{\text{influenza}, \text{sharp respiratory disease}, \text{angina}, \text{pneumonia}\}$ – space for determining the value of linguistic variable y . Then the dependence of the patient’s disease on his symptoms can be described by the following systems of speci-

fications:

if x_1 is “weak” \wedge x_2 is “weak” \wedge x_3 is “raised” **then**
 y is “0.5/*influenza* + 0.5/*sharp respiratory disease* + 0.4/*angina* +
 +0.8/*pneumonia*“;

if x_1 is “weak” \wedge x_2 is “moderate” \wedge x_3 is “high” **then**
 y is “0.8/*influenza* + 0.7/*sharp respiratory disease* + 0.8/*angina* +
 +0.3/*pneumonia*“;

if x_1 is “weak” \wedge x_2 is “moderate” \wedge x_3 is “very high” **then**
 y is “0.9/*influenza* + 0.7/*sharp respiratory disease* + 0.8/*angina* +
 +0.2/*pneumonia*“;

If to the input x_1 of this algorithm to supply value $A'_1 = 1/5 + 0.7/10$, to the input x_2 – value $A'_2 = 1/5 + 0.5/10$, to the input x_3 – value $A'_3 = 1/36 + 0.9/37$, then in accordance with procedure of fulfilling the algorithm of the fuzzy system of logical inference the fuzzy solution of the problem will be received

$$B = 0.5/*influenza* + 0.5/*sharp respiratory disease* +
 +0.4/*angina* + 0.5/*pneumonia*.$$

The problem of searching the symptoms using fuzzy diagnosis can be inverse to this problem. Specifically, let output of the fuzzy system of logical inference with inputs $A'_1 = x_1/5 + x_2/10 + x_3/15 + x_4/20$, $A'_2 = 1/5 + 0.5/10$, $A'_3 = 1/36 + 0.9/37$ is fuzzy set $B = 0.5/*influenza* + 0.5/*sharp respiratory disease* + 0.4/*angina* + 0.5/*pneumonia*$. Aggregation of the individual outputs leads to the next system of fuzzy relation equations:

$$\begin{aligned} \min [\max(x_1 \wedge 1, x_2 \wedge 0.5), 0.4] &= 0.4, \\ \min [\max(x_1 \wedge 1, x_2 \wedge 0.5), 0.5] &= 0.5. \end{aligned}$$

Solving it the value of the symptom “*Coughing*” will be received, which is described by the fuzzy set “*Coughing*” = 0.5/5 + 1/10.

To determine the probability of event A in the space of elementary events X , the notion of probable measure is introduces. It is numerical function P , which puts number $P(A)$ to the elementary event A , besides that

$$0 \leq P(A) \leq 1, P(X) = 1, P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

for any A_1, A_2, \dots such, as $A_i \cap A_j = \emptyset$, with $i \neq j$.

The fuzzy event A in space X will be called the fuzzy set

$$A = \{(x, \mu_A(x)), x \in X\},$$

where $\mu_A : X \rightarrow [0, 1]$ – membership function of fuzzy set A . Then probability of event A can be calculated using the formula

$$P(A) = \sum_{x \in A} \mu_A(x)P(x).$$

Considering this, it is possible to calculate the probabilities of any fuzzy event with provided probability measure. In particular, the probability of symptom “*Coughing*“ at probabilities distribution

$$P(5) = 0.4, P(10) = 0.4, P(15) = 0.1, P(20) = 0.1$$

can be calculated in the following way:

$$P(\text{“Coughing“}) = 0.5 \cdot 0.4 + 1 \cdot 0.4.$$

3 Bioinformatics

It is known [1], [4] that the problem of recognizing the structures of the proteins of different organization levels is rather complicated. To solve it the different methods and approaches, including experimental (based on physics of chemical relations creation), machine teaching (used the data bases of experimentally found secondary structures as learning samples), probabilistic (on the basis of the Bayes procedures and Markov chains) are used.

The method of recognition of the secondary structure of DNA using fuzzy systems of logical inference is proposed. The problem is the following: it is necessary to build the fuzzy system of logical inference which using random amino acid sequence would define (as an fuzzy set) the secondary structure of central remainder (of the amino acid) of the input sequence.

To solve this problem at first it is necessary to design the fuzzy specification of the problem according to learning samples. One of the methods to build the system of fuzzy instructions according to numerical data consists of the following. Let’s the rules base with n inputs and one output is created. There are learning dates (samples) as the sets of pairs for that

$$(x_1(i), x_2(i), \dots, x_n(i); d(i)), i = 1, 2, \dots, m,$$

where $x_j(i)$ – inputs and $d(i)$ – output, at that $x_j(i) \in \{a_1, a_2, \dots, a_k\}$, $d(i) \in \{b_1, b_2, \dots, b_l\}$. It is necessary to build the fuzzy system of logical inference which would generate the correct output data according to random input values. The algorithm of solving of the provided problem consists in the following sequence of steps:

1. Dividing the space of inputs and outputs for areas (dividing learning data for groups on m_1, \dots, m_k lines, which means, each input and output is divided for $2N + 1$ cuts where N for each input is selected individually. Separate areas (segments) will be called in the following way:

$$M_N(\text{left } N), \dots, M_1(\text{left } 1), S(\text{medium}), D_1(\text{right } 1), \dots, D_N(\text{right } N).$$

Determination membership function for each areas.

To build the fuzzy rules of logical inference the learning samples from 15 remains of the protein MutS [2] are used which look like the following:

```

K V S E G G L I R E G Y D P D
e - - - h h h h h h h h h h h h
V S E G G L I R E G Y D P D L
- - - h h h h h h h h h h h h
S E G G L I R E G Y D P D L D
- - h h h h h h h h h h h h
E G G L I R E G Y D P D L D A
- h h h h h h h h h h h h h
G G L I R E G Y D P D L D A L
h h h h h h h h h h h h h

```

The prediction belongs to the central remain, besides the following denotations are used: h – for spiral, e – for cylinder, “-” – other.

According to the algorithm, the teaching data is divided, for example, for 3 groups:

```

K V S E G G L I R E G Y D P D
V S E G G L I R E G Y D P D L;

S E G G L I R E G Y D P D L D
E G G L I R E G Y D P D L D A;

G G L I R E G Y D P D L D A L;

```

and compared to each group with according fuzzy sets

$$A_i^{(m_1)}, A_i^{(m_2)}, A_i^{(m_3)}, B^{(m_1)}, B^{(m_2)}, B^{(m_3)}.$$

Then fuzzy specification of the recognition problem will look like:

$$\begin{aligned}
R^{(1)} &: \text{if } x_1 \text{ is } A_1^{(m_1)} \text{ and } x_2 \text{ is } A_2^{(m_1)} \text{ and } \dots \text{ and } x_{14} \text{ is } A_{14}^{(m_1)} \text{ then } y \text{ is } B^{(m_1)}, \\
R^{(2)} &: \text{if } x_1 \text{ is } A_1^{(m_2)} \text{ and } x_2 \text{ is } A_2^{(m_2)} \text{ and } \dots \text{ and } x_{14} \text{ is } A_{14}^{(m_2)} \text{ then } y \text{ is } B^{(m_2)}, \\
R^{(3)} &: \text{if } x_1 \text{ is } A_1^{(m_3)} \text{ and } x_2 \text{ is } A_2^{(m_3)} \text{ and } \dots \text{ and } x_{14} \text{ is } A_{14}^{(m_3)} \text{ then } y \text{ is } B^{(m_3)},
\end{aligned}$$

Using the algorithm of solving the specification, we will find the output received system of fuzzy instructions, if to the input the following amino acid sequence is supplied:

L K V S E G G L I R E G Y D P.

In accordance with the procedure of executing the algorithm we will get that the secondary structure of the remainder L is h .

4 Fuzzy Boolean Variables

The algebra of statement is one of the chapters of classic mathematical logic. The statement means the variable which can be of two possible values – 0 or 1. Such variable is

called Boolean. In some cases generalizing the notion of Boolean variable to the notion of fuzzy Boolean variable is useful [3]. The fuzzy Boolean variable is called a variable p , which takes a value from the interval $[0, 1]$.

Let p and q – fuzzy Boolean variables. Logical operations with such variables are determined like that:

$$\begin{aligned}\bar{p} &= 1 - p, \\ p \wedge q &= \min(p, q), \\ p \vee q &= \max(p, q).\end{aligned}$$

From determination of the operations we get, that the following laws

$$p \vee \bar{p} = 1, p \wedge \bar{p} = 0$$

are broken.

Let p_1, p_2, \dots, p_n – fuzzy Boolean variables. Function $f(p_1, \dots, p_n)$ is called a function of fuzzy Boolean variables if it takes value on the interval $[0, 1]$.

Function f of fuzzy Boolean variables is called analytical, if it can be presented by the formula, which includes operations \neg, \wedge, \vee .

Since in the fuzzy logic above-mentioned laws are violated, function

$$p \rightarrow q = \bar{p} \cdot \bar{q} \vee \bar{p} \cdot q \vee p \cdot q$$

cannot be simplified.

One of the tasks Boolean variables function analysis consists in the following. It is necessary to find out under the which conditions the values of analytical function, for example $f(p, q) = p \wedge q$, includes in a given interval $[\alpha, \beta]$ of the segment $[0, 1]$ under the condition, that $p \in [a_1, a_2] \subseteq [0, 1]$ and $q \in [b_1, b_2] \subseteq [0, 1]$. Solution to this problem gives the possibility to calculate probabilities of varying weather in the following way.

It is know, that weather forecasters evaluate their forecasts using probability theory point out their forecasts with expressions:

$$\begin{aligned}&\text{“sunny“ with probability } p \in [0.7, 0.8]; \\ &\text{“windy“ with probability } q \in [0.3, 0.5]; \\ &\text{“cloudy“ with probability } h \in [0.8, 0.9].\end{aligned}$$

Let us consider Boolean analytical function

$$f(p, q) = p \rightarrow q = \bar{p} \cdot \bar{q} \vee \bar{p} \cdot q \vee p \cdot q.$$

Let us suppose, that $p \in [0.7, 0.8)$, $q \in [0.3, 0.5)$. It is necessary to find out in which intervals of the segment $[0, 1]$ values of the function includes. In other words, it is necessary to find which interval contains the probability that the weather will be “sunny“ and “windy“, or “not sunny“ and “windy“, or “not sunny“ and “not windy“.

Considering, that $p \rightarrow q = \bar{p} \cdot \bar{q} \vee \bar{p} \cdot q \vee p \cdot q$ and that

$$p \in [0.7, 0.8], q \in [0.3, 0.5]$$

we find that

$$\bar{p} \in [0.2, 0.3], \bar{q} \in [0.5, 0.7].$$

So

$$\begin{aligned}\bar{p} \cdot \bar{q} &= \min(\bar{p}, \bar{q}) \in [0.2, 0.3], \\ \bar{p} \cdot q &= \min(\bar{p}, q) \in [0.2, 0.3], \\ p \cdot q &= \min(p, q) \in [0.3, 0.5].\end{aligned}$$

Then

$$\begin{aligned}\bar{p} \cdot \bar{q} \vee \bar{p} \cdot q &\in [0.2, 0.3], \\ \bar{p} \cdot q \vee \bar{p} \cdot \bar{q} \vee p \cdot q &\in [0.3, 0.5].\end{aligned}$$

That is why value of the function $f(p, q)$ will includes to the interval $[0.3, 0.5]$.

Let us consider another option of the problem. Let us suppose, that as earlier, $p \in [0.7, 0.8]$, $q \in [0.3, 0.5]$ analytical Boolean function is unknown, but is known the interval of the segment $[0, 1]$, in which the values of this function are includes. In this case – this is the interval $[0.3, 0.5]$. It is necessary to find out in which interval of the segment $[0, 1]$ the values of this unknown function included, if for example $p \in [0.5, 0.6]$, $q \in [0.3, 0.5]$.

One of the approaches to solve this problem consists in building and researching it is so called linguistic model [2], [8]. In this model the variables p, q, f will be considered as linguistic variables, and appropriate intervals will be described by the fuzzy sets of the space $[0, 1]$. So, our fuzzy model will look like the following:

$$R : \text{if } p \in A_1 \wedge q \in A_2 \text{ then } f \in B$$

where $A_1 = 1/[0.7, 0.8]$, $A_2 = 1/[0.3, 0.5]$, $B = 1/[0.3, 0.5]$. It is necessary to find the output B' of this fuzzy rule with inputs $A'_1 = 1/[0.6, 0.8]$, $A'_2 = 1/[0.3, 0.5]$. After calculating according to the procedure we will have:

$$B'(y) = 1/[0.3, 0.5].$$

One more position of the task consists in the following. Let $p \in [0.5, 0.6]$.

Let us calculate the probability of windy under the condition that $p \rightarrow q \in [0.3, 0.5]$.

Let us suppose, that $q \in [b_1, b_2] \subseteq [0, 1]$. Then

$$\begin{aligned}p \wedge q &\in [\min(0.5, b_1), \min(0.6, b_2)], \\ \bar{p} \wedge q &\in [\min(0.4, b_1), \min(0.5, b_2)], \\ \bar{p} \wedge \bar{q} &\in [\min(0.4, 1 - b_2), \min(0.5, 1 - b_1)], \\ p \cdot q \vee \bar{p} \cdot q &\in [\max(\min(0.5, b_1), \min(0.4, b_1)), \max(\min(0.6, b_2), \min(0.5, b_2))], \\ p \wedge q \vee \bar{p} \cdot \bar{q} \vee \bar{p} \cdot q &\in [\max(\min(0.5, b_1), \min(0.5, b_1), \min(0.5, 1 - b_1)), \\ &\max(\min(0.6, b_2), \min(0.4, b_2), \min(0.4, 1 - b_2))].\end{aligned}$$

Here we get:

$$\begin{aligned} \max(\min(0.5, b_1), \min(0.5, 1 - b_1)) &> 0.3, \\ \max(\min(0.6, b_2), \min(0.4, b_2), \min(0.4, 1 - b_2)) &\leq 0.5. \end{aligned}$$

From first correlation we find b_1 :

$$\begin{aligned} 0.5 \leq b_1 &\Rightarrow \\ & b_1 - \text{arbitrary} \\ b_1 \leq 0.5 &\Rightarrow \end{aligned}$$

From second correlation we find b_2 :

$$\begin{aligned} 0.4 \leq 0.6 \leq b_2 &\Rightarrow \\ 0.4 \leq b_2 \leq 0.6 &\Rightarrow 0 \leq b_2 \leq 0.5 \\ b_2 \leq 0.4 \leq 0.6 & \end{aligned}$$

So, the probability of windy weather $q \in [0, 0.5]$.

5 Conclusions

The proposed approach to solving problems (based on fuzzy models) allows to simplify the methods of solving problems. Other approaches may be based on possibility theory. But there is a necessity for additional studies of the reliability issues in the first and in the second case.

References

1. Gupal, A.M., Sergienko, I.V.: Optimal Recognition Procedures (in Russian). Naukova Dumka, Kyiv (2008)
2. Katerynych, L., Provotar, A.: Neural Networks Diagnostics in Homeopath System. International Journal "Information Theories & Applications". 15, 89–93 (2008)
3. Konyshcheva, L.K., Nazarov, D.M.: Foundations of Fuzzy Sets Theory (in Russian). SPB Piter, Moscow (2011)
4. Lesk, A.: Introduction to Bioinformatics (in Russian). Laboratoria Znaniy, Moscow (2009)
5. Leski, J.: Systemy Neuronowo-Rozmyte (in Polish). Naukowo-Techniczne, Warszawa (2008)
6. Provotar, A.I., Lapko, A.V., Provotar, A.A.: Fuzzy Inference Systems and Their Applications. International Scientific Journal "Cybernetics and Systems Analysis". 49, 517–525 (2013)
7. Rutkowska, D., Piliński, M., Rutkowski, L.: Sieci Neuronowe, Algorytmy Genetyczne i Systemy Rozmyte (in Polish). Wydawnictwo Naukowe PWN, Warszawa (1999)
8. Zadeh, L.A.: Fuzzy Sets as a Basis for a Theory of Possibility. Fuzzy Sets and Systems. 1, 3–28 (1978)