Application of Multi-dimensional Fuzzy Analysis to Decision Making

Alexandr A. Savinov

Laboratory of AI Systems
Institute of Mathematics, Moldavian Academy of Sciences
str. Academiei 5, MD-2028 Kishinev, Moldavia
Phone: 3732-73-81-30, Fax: 3732-73-80-27
E-mail: savinov@math.md
http://www.geocities.com/ResearchTriangle/7220/

Key words: Multi-dimensional fuzzy analysis, Fuzzy decision support systems, Fuzzy resolution, Rule induction, Fuzzy logical inference.

Abstract

The goal of multi-dimensional fuzzy analysis consists in discovering different properties in multi-dimensional fuzzy distributions represented either extensionally (database) or intensionally (knowledge base). In this paper we show how this approach can be applied to such problems as decision making and knowledge discovery in databases. For uniform and efficient representation of fuzzy knowledge and data we propose a technique of sectioned matrices. For carrying out logical inference we use a new operation of fuzzy resolution which is a generalization of the conventional resolution. With the help of this operation we find fuzzy prime disjunctions which later are used for making decisions in concrete situations. For discovering hidden dependencies in data a new fuzzy covering method is used.

1. Introduction

Let us consider the following general problem. Given n variables $x_1, x_2, ..., x_n$ each of which takes its values from the sets $X_1, X_2, ..., X_n$, respectively. The universe of discourse O is equal to the Cartesian product of all sets of the values

$$O = X_1 \times X_2 \times ... \times X_n$$

and consists of n-tuples

$$o = \langle x_1, x_2, ..., x_n \rangle$$

There is some distribution over the universe of discourse. The general goal of multi-dimensional analysis consists in finding different properties of distributions

and their transformations to other forms.

Depending on the character (modality) of the distribution and the form of its representation multi-dimensional analysis can be divided into subdomains. For example, traditional multi-dimensional statistical analysis aims at discovering properties of probabilistic distributions, e.g., mean values or correlation, represented extensionally in the form of a table of data. (Such an important case as the analysis of intensionally represented multi-dimensional probabilistic distributions, e.g., by means of rules, is not studied in this subdomain.)

Note that the key point in the whole approach is *multi-dimensionality*, i.e., it is precisely multi-dimensionality that is responsible for non-triviality of problems and difficulty of procedures. In particular, such a well known problem as finding satisfiability can be easily reformulated as a problem of multi-dimensional analysis of two-valued distributions over n-dimensional hyper-cube which are represented by means of CNF or DNF. Although multi-dimensional analysis can be used for extensionally represented distributions, e.g., to analyze tables of data, the most interesting and difficult problems arise in analyzing intensionally represented distributions. In the former case the distribution is represented pointwisely, i.e., each statement represents the value of the distribution in one point of the universe of discourse. In the latter case we use elementary distributions over the values of individual variables and some operations to represent multi-dimensional distributions. Thus *intensionality* of representation is the second key point of multi-dimensional analysis.

In this paper we consider multi-dimensional analysis of fuzzy finite distributions [13]. We suppose that the sets $X_1, X_2, ..., X_n$ called domains are finite and consist of $n_1, n_2, ..., n_n$ members, respectively. If the variable x_i has only two values (they are supposed to be 0 and 1, i.e., $X_i = \{0,1\}$ and $n_i = 2$), then it is said to be boolean variable. In finite case the power of the universe of discourse is equal to $n_1 \times n_2 \times ... \times n_n$. In the case of all boolean variables the universe of discourse is equal to n-dimensional hyper-cube with the power 2^n.

The second our supposition concerns the character of the distributions being studied. Namely, we suppose that the distributions take their values from the interval [0,1]. If the distribution is crisp, i.e., takes only values 0 and 1, it is said to be boolean. If only boolean variables and boolean distributions are used then we obtain just the case well studied in such fields as boolean algebra, propositional calculus, switching theory etc. Multi-dimensional analysis differs from these and other similar directions in that it has its own goal — finding properties of distributions, although many of these approaches often have very similar and even identical main notions and procedures.

The third main supposition is that the operations of minimum and maximum are used to combine fuzzy distributions. In particular, this property distinguishes fuzzy case from the probabilistic one.

According to these suppositions we assume that any problem domain being

considered is described in terms of attributes and their values. If such a suitable set of attributes and values exists then it is referred to as an *attribute model* of the problem domain [1]. All combinations of the values of attributes represent possible objects or states of the problem domain. By knowledge we mean some fuzzy constraints on possible states of the problem domain represented by the corresponding multi-dimensional fuzzy distribution. If there is not any knowledge about the problem domain then any combination of attribute values is possible (meaningful) and is assigned value 1. Otherwise some states are fuzzily disabled, i.e., they are assigned values which are less than 1 (0 means that the state or object is completely impossible, meaningless).

The first problem is how to represent such knowledge in attribute model by means of multi-dimensional fuzzy distributions. In section 2 we propose a method of sectioned matrices which allows us to represent knowledge and data uniformly and efficiently. The second problem is how to carry out logical inference and make conclusions proceeding from some knowledge and data. It is described in section 3. The third problem considered in the paper is how to generate knowledge from available data. In section 4 we propose a method to transform a sectioned matrix of data into a sectioned matrix of knowledge.

2. Knowledge and Data Representation

In this section we describe a sectioned matrix technique [2, 3] for representing multi-dimensional fuzzy distributions. An elementary distribution is a distribution over one domain and represents the corresponding proposition about the attribute. For example, if there is an attribute ALCOHOLISM STAGE with the values from the set {First, Second, Third}, then the elementary proposition

ALCOHOLISM STAGE = {First : 1, Second : 0.2, Third : 0}

means that the corresponding fuzzy distribution has value 1 in the point First, value 0.2 in the point Second and value 0 in the point Third.

Elementary propositions can be represented only extensionally since they are 1-dimensional, i.e., they can be represented only by enumerating their values in all points of the domain of definition. We will write elementary propositions in the form of one *section*, i.e., simply enumerating all its values, e.g., {1, 0.2, 0}. One number in the section is said to be a *component*. Thus the section {1, 0.2, 0} consists of three components 1, 0.2 and 0.

A sectioned *vector* is defined as a set of sections for all attributes. It can be interpreted as disjunction or conjunction. If the vector is interpreted as disjunction then all its sections are supposed to be combined with the operation of maximum. (From the logical point of view its propositions are combined with the connective OR.) If the vector is interpreted as conjunction then all its sections are supposed to be combined with the operation of minimum. For example, the following constructions are fuzzy sectioned vectors for an attribute model consisting of 3

attributes:

$$\{1, 0.2, 0\}, \{0, 1\}, \{0, 0, 0\}$$
$$\{0, 1, 0.5\}, \{0.5, 0.7\}, \{1, 1, 1\}$$

Depending on its interpretation, i.e., on the connective used for combining sections, each sectioned vector represents one or another multi-dimensional distribution over the universe of discourse O. If the sectioned vector $\mathbf{d}$ is interpreted as disjunction then it defines the following multi-dimensional fuzzy distribution over the space O:

$$\mathbf{d}(o) = \mathbf{d}(\langle x_1, x_2, ..., x_n \rangle) = \max[\mathbf{d}_1(x_1), \mathbf{d}_2(x_2), ..., \mathbf{d}_n(x_n)]$$

where $\mathbf{d}(o)$ is equal to the value of the distribution in the point o and $\mathbf{d}_i$ are sections of the vector. In other words, to find the value $\mathbf{d}(o)$ we have to choose the maximal of n components (one component from each section) corresponding to the values of variables. Note that the section consisting of all 0 is equivalent to its absence.

If the sectioned vector $\mathbf{k}$ is interpreted as conjunction then it defines a multi-dimensional fuzzy distribution over the space O in the dual way:

$$\mathbf{k}(o) = \mathbf{k}(\langle x_1, x_2, ..., x_n \rangle) = \min[\mathbf{k}_1(x_1), \mathbf{k}_2(x_2), ..., \mathbf{k}_n(x_n)]$$

To compute its value in some point we have to take minimal of n components corresponding to the values of variables. If some section consists of all 1 then it is equivalent to its absence.

With the help of disjunctions we can represent general dependencies between attributes. A number of such disjunctions combined with the operation of minimum represents a knowledge base.

Let us consider an example from the field of screening alcoholism [4] where the goal is to determine a patient alcoholism stage (the methods of treatment depend on this diagnosis). The problem domain can be described by the following three attributes:

DEPENDENCE = {Psychical, Psychico-physical, Physical}
LOSS OF SELF-CONTROL = {Quantitative, Situational}
ALCOHOLISM STAGE = {First, Second, Third}

If there are not known dependencies between these attributes then any state of the problem domain is possible, i.e., any combination of their values is considered meaningful, and the knowledge base represents the multi-dimensional distribution which is equal to 1 in all points of the universe of discourse.

One way to express our knowledge about the problem domain consists in indicating a state (or a fuzzy interval of states) which is disabled. Then the corresponding combinations of the attribute values are meaningless. For example, if nobody has ever seen anything like the object

$$\langle \text{Psychical, Situational, Third} \rangle$$

(i.e., a patient at the third stage of alcoholism with the psychical dependence and situational loss of self-control) then such a knowledge can be written as the disjunction

$$\mathbf{1} = \{0, 1, 1\}, \{1, 0\}, \{1, 1, 0\}$$

which pricks a whole in the unit level distribution (Fig. 1).
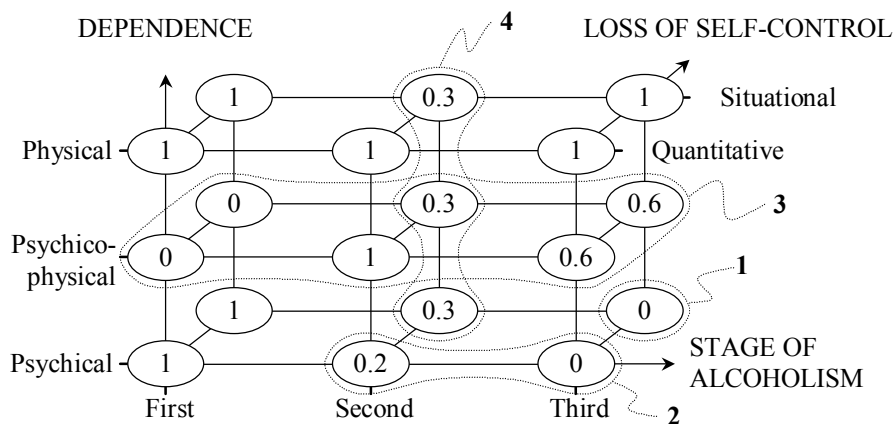

Figure. 1. Fuzzy distribution over the universe of discourse

Another way to represent knowledge is by means of conventional implications. Each implication can be easily transformed into disjunction by negating its left part. For example, the following two implications

DEPENDENCE = {Psychical : 1, Psychico-physical : 0, Physical : 0} AND
LOSS OF SELF-CONTROL = {Quantitative : 1, Situational : 0} →
ALCOHOLISM STAGE = {First : 1, Second : 0.2, Third : 0}

DEPENDENCE = {Psychical : 0, Psychico-physical : 1, Physical : 0} →
ALCOHOLISM STAGE = {First : 0, Second : 1, Third : 0.6}

can be written as the following two disjunctions

$$\mathbf{2} = \{0, 1, 1\} \{0, 1\} \{1, 0.2, 0\}$$
$$\mathbf{3} = \{1, 0, 1\} \{0, 0\} \{0, 1, 0.6\}$$

Rules represented in the form of implication usually describe dependence of the goal from the rest of the attributes. Sometimes it is convenient to express our knowledge in the form of a backward dependency where we describe properties (symptoms, situations) which are typical of the concrete conclusion (decision). In

other words, we suppose that a concrete decision has been made and then ask the question what other conditions it is characterized by. For example, let us suppose that we know one more dependency where situational loss of self-control is usually (degree 0.3) not typical of the second stage of alcoholism:

ALCOHOLISM STAGE = {First : 0, Second : 1, Third : 0} →
LOSS OF SELF-CONTROL = {Quantitative : 1, Situational : 0.3}

Then it can be written in the form of disjunction

$$\mathbf{4} = \{0, 0, 0\}, \{1, 0.3\}, \{1, 0, 1\}$$

The whole matrix representing our knowledge consists of 4 disjunctions **1**, **2**, **3**, **4** and looks as follows

$$
\mathbf{D} =
\begin{array}{ccc|c}
x_1 & x_2 & x_3 & \\
\{0, 1, 1\} & \{1, 0\} & \{1, 1, 0\} & \mathbf{1} \\
\{0, 1, 1\} & \{0, 1\} & \{1, 0.2, 0\} & \mathbf{2} \\
\{1, 0, 1\} & \{0, 0\} & \{0, 1, 0.6\} & \mathbf{3} \\
\{0, 0, 0\} & \{1, 0.3\} & \{1, 0, 1\} & \mathbf{4}
\end{array}
$$

and the corresponding multi-dimensional fuzzy distribution is shown in Fig. 1.

Data is represented by means of sectioned vectors interpreted as conjunctions. Each section of one conjunction represents fuzzy constraints on the values of the corresponding attribute independent of the other attributes. Thus observable data used as an input for logical inference and making decision can be represented as one conjunction each section of which contains information about one attribute. If there is not any observable data then such vector of data contains all 1 (all values are possible). If we have known that some attribute cannot take certain value then the corresponding component is decreased in proportion to our degree of confidence. For example, if it is known that the patient does not have physical dependence than it can be written as the sectioned vector

$$\{1, 1, 0\}, \{1, 1\}, \{1, 1, 1\}$$

In addition, if we suppose with the degree 0.5 that he does not have quantitative loss of self-control then we obtain the vector

$$\{1, 1, 0\}, \{0.5, 1\}, \{1, 1, 1\}$$

Thus as we obtain and enter new information, components of the vector of data (and the whole distribution values) are decreased. Obviously, each section must contain at least one 1, otherwise our data is inconsistent.

If all components of conjunction are equal to 0 except for the one component in each section, for example,

$$\{1, 0, 0\}, \{1, 0\}, \{1, 0, 0\}$$
$$\{0.7, 0, 0\}, \{0.7, 0\}, \{0.7, 0, 0\}$$

then such a vector represents the distribution which is equal to 0 in all except for the one points of the universe of discourse where it is equal to the minimal of n non-0 components in each section. Thus by means of such vector we can represent one element of data, e.g., one row of a database, along with the corresponding degree of membership, and with the help of a set of such conjunctions we can represent any multi-dimensional distribution.

3. Multi-attribute Decision Making

The problem of fuzzy decision making in concrete situation for an attribute model of the problem domain can be formulated as follows [5, 6]. There is some data on the observable situation represented in the form of fuzzy constraints on the possible attribute values and written as a vector-conjunction **f**. In addition, there is a knowledge about fuzzy dependencies between attributes which is represented by means of a sectioned matrix **D**. The problem is to find fuzzy constraints on the values of a goal attribute(s) which are equal to the projection of the joint multi-dimensional distribution including data and knowledge onto the goal variable(s). Thus logical inference is reduced to the problem of finding projection of some multi-dimensional distribution. In other words, if we know that some *combinations* of attribute values are disabled (knowledge about the problem domain) and, in addition, certain values of *individual* attributes are disabled (data on the observable situation) then we have to find disabled values of the goal attribute(s).

Once we have built our knowledge base we are interested in carrying out consultations, i.e., obtaining concrete conclusions on necessary attributes in concrete situations. For example, we might be interested in obtaining possible stages of alcoholism for a patient for whom we know that he does not have physical dependence, i.e., the situation is described by the conjunction

$$\{1, 1, 0\},\{1, 1\},\{1, 1, 1\}$$

In principle, we could solve this problem extensionally using combination and projection principle [7, 8]. For this it is necessary to calculate the values of 4 distributions corresponding to 4 disjunctions from the knowledge base and the values of 1 distribution corresponding to 1 conjunction representing data in all points from the universe of discourse. Then we have to combine all 5 distributions with the help of minimum operation and to find the ultimate distribution representing both knowledge and data. And finally we have to find projection of this distribution onto the third variable by finding maximal values in 3 subspaces of the universe of discourse each of which corresponds to one value of the goal attribute. In our example we can easily infer (Fig. 1) that the projection on the goal attribute is equal to $\{1, 1, 0.6\}$, i.e., in this situation the third stage of alcoholism is disabled with the degree 0.6. Obviously, this method is inapplicable to real examples since the universe of discourse is very large space and it is not possible to compute the distribution values in all its points. Some usefulness of this procedure is that we deeper understand the essence of multi-dimensional logical inference at the level of one point ("pointwise" logical inference). Thus we have to

use intensional approaches to logical inference which would explicitly use disjunctions to make conclusions.

One such method also uses combination and projection principle but in other form. In this method we try to make conclusions from individual disjunctions and then combine all results by finding conjunction of all projections. After the first pass when we found all projections of individual disjunctions we should take the conclusions along with the data and use it as new data for the next pass. Then, at the next pass we make new conclusions from individual disjunctions and again add them to the previous data. The process is stopped when no new conclusions can be obtained.

For this method we need a procedure of finding projection of one disjunction and one conjunction. It can be proved that this projection is computed in two steps. First, we have to find componentwise conjunction (minimum) of two sectioned vectors. Then we have to take the sectioned vector interpreted as disjunction resulted from this step and find its projection onto the goal attribute. Projection of one disjunction is equal to its goal section to which we add with the help of maximum operation the maximal component in all the rest of the sections. (For the sake of simplicity in this paper we do not take into account the issue of handling disjunction constant which is connected with the notion of reduced disjunctions [3, 5, 6, 9, 13].)

In our example, first, we find componentwise conjunction of the vector of data with each disjunction and obtain the matrix

$$\mathbf{D} = \begin{array}{ccc} x_1 & x_2 & x_3 \\ \begin{vmatrix} \{0, 1, 0\} & \{1, 0\} & \{1, 1, 0\} \\ \{0, 1, 0\} & \{0, 1\} & \{1, 0.2, 0\} \\ \{1, 0, 0\} & \{0, 0\} & \{0, 1, 0.6\} \\ \{0, 0, 0\} & \{1, 0.3\} & \{1, 0, 1\} \end{vmatrix} & \begin{matrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \end{matrix} \end{array}$$

(In this matrix 3rd components in 1st section are equal to 0.) Then we have to find 4 projections of the individual disjunctions onto 3rd attribute. Since all 4 disjunctions involve 1 either in 1st or in 2nd section (non-goal sections), we obtain that all their projections are equal to {1, 1, 1}, i.e., they do not have non-trivial projections onto the goal attribute (addition to any section of the value 1 with the help of maximum operation results in the section consisting of all 1). We have not inferred any conclusion and therefore the procedure is stopped. Comparing this result with that of the previous procedure (pointwise logical inference) we see that this procedure (it could be called "disjunctionwise" logical inference) does not guarantee obtaining exact result. Note that the method of inference on individual disjunctions (rules) followed by combination of conclusions dominates on the application to inference engines. It is caused by its simplicity and that exactness of inference is not the most important requirement in many systems — frequently it is sufficient if rules work independently.

Let us consider a method of logical inference which guarantees that the conclusion, i.e., projection obtained, will be exact. This method requires that the matrix of

fuzzy disjunctions be equivalently transformed into the form where it contains all *prime* disjunctions [3, 5, 6, 9, 13]. Informally prime disjunctions are the most strong propositions about the problem domain provided that all of them are consequences of the initial matrix, i.e., if any component of prime disjunction is decreased then it is not already a consequence of the matrix. Thus before we can carry out logical inference we have to generate all prime disjunctions for our knowledge base.

One approach to building prime disjunctions is based on a new original operation of fuzzy resolution [3, 5, 6, 9-11, 13, 14] which generalizes traditional resolution in that it can be used (i) in the case of multi-valued variables (proposed in [2]), and (ii) for fuzzy distributions (proposed in [3]). If the resolution is applied to k-th section of two disjunctions $\mathbf{u}$ and $\mathbf{v}$ then the resolvent $\mathbf{w}=\mathbf{u}\langle x_k\rangle\mathbf{v}$ resulted from this operation is computed as follows:

1. k-th section $\mathbf{w}_k$ is equal to conjunction of two corresponding sections $\mathbf{u}_k$ and $\mathbf{v}_k$, i.e., each component in $\mathbf{w}_k$ is equal to the minimum of two corresponding components in $\mathbf{u}_k$ and $\mathbf{v}_k$.

2. non-k-th sections $\mathbf{w}_i$ are equal to disjunction of two corresponding sections $\mathbf{u}_i$ and $\mathbf{v}_i$, i.e., each component in $\mathbf{w}_i$ is equal to the maximum of two corresponding components in $\mathbf{u}_i$ and $\mathbf{v}_i$.

The whole procedure for generating prime disjunctions consists in applying the resolution operation to all disjunction pairs on all variables and adding new disjunctions to the matrix. Of course, it is natural to check if new disjunction is a consequence of one of those which are already in the matrix. When all disjunctions have been generated the matrix has to be simplified by removing (absorbing) disjunctions which follow from others. Then new pass has to be carried out. The procedure is stopped when any new resolvent is absorbed by disjunctions in the matrix, i.e., we are not able to generate new disjunctions.

After the first pass of this procedure we obtain the matrix

$$
\mathbf{D}_1 =
\begin{array}{|ccc|l}
x_1 & x_2 & x_3 & \\
\{0, 1, 1\} & \{0, 1\} & \{1, 0.2, 0\} & \mathbf{2} \\
\{1, 0, 1\} & \{0, 0\} & \{0, 1, 0.6\} & \mathbf{3} \\
\{0, 0, 0\} & \{1, 0.3\} & \{1, 0, 1\} & \mathbf{4} \\
\{0, 1, 1\} & \{0, 0\} & \{1, 1, 0\} & \mathbf{5}=\mathbf{1}\langle x_2\rangle\mathbf{2} \models \mathbf{1} \\
\{0, 0, 1\} & \{1, 0\} & \{1, 1, 0.6\} & \mathbf{6}=\mathbf{1}\langle x_1\rangle\mathbf{3} \\
\{0, 1, 1\} & \{1, 0\} & \{1, 0.3, 0\} & \mathbf{7}=\mathbf{1}\langle x_3\rangle\mathbf{4} \\
\{0, 0, 1\} & \{0, 1\} & \{1, 1, 0.6\} & \mathbf{8}=\mathbf{2}\langle x_1\rangle\mathbf{3} \\
\{0, 1, 1\} & \{0, 0.3\} & \{1, 0.2, 1\} & \mathbf{9}=\mathbf{2}\langle x_2\rangle\mathbf{4} \\
\{1, 0, 1\} & \{1, 0.3\} & \{0, 0, 0.6\} & \mathbf{10}=\mathbf{3}\langle x_3\rangle\mathbf{4}
\end{array}
$$

Note that the disjunction **1** has been absorbed by the resolvent $\mathbf{5}=\mathbf{1}\langle x_2\rangle\mathbf{2}$, i.e., resolvent absorbed one of its parents. The matrix $\mathbf{D}_1$ consists of 9 disjunctions and no one of them is a consequence of another. At the second pass we have to find all

resolvents of the disjunctions in matrix $\mathbf{D}_1$. After absorbing disjunctions **5**, **7**, **6**, **8**, **9** we obtain the matrix

$$
\mathbf{D}_2 =
\begin{array}{ccc|l}
x_1 & x_2 & x_3 & \\
\{0, 1, 1\} & \{0, 1\} & \{1, 0.2, 0\} & 2 \\
\{1, 0, 1\} & \{0, 0\} & \{0, 1, 0.6\} & 3 \\
\{0, 0, 0\} & \{1, 0.3\} & \{1, 0, 1\} & 4 \\
\{1, 0, 1\} & \{1, 0.3\} & \{0, 0, 0.6\} & 10{=}3\langle x_3\rangle 4 \\
\{0, 1, 1\} & \{0, 0\} & \{1, 0.3, 0\} & 11{=}2\langle x_2\rangle 7 \models 5, 7 \\
\{0, 0, 1\} & \{0, 0\} & \{1, 1, 0.6\} & 12{=}3\langle x_1\rangle 5 \models 6, 8 \\
\{0, 0, 1\} & \{1, 0\} & \{1, 0.3, 0.6\} & 13{=}4\langle x_3\rangle 6 \models 6 \\
\{0, 1, 1\} & \{0, 0.3\} & \{1, 0.2, 0\} & 14{=}5\langle x_3\rangle 9 \models 9 \\
\end{array}
$$

At the next pass we generate only one disjunction which absorbs disjunction **10**

$$
\mathbf{D}_3 =
\begin{array}{ccc|l}
x_1 & x_2 & x_3 & \\
\{0, 1, 1\} & \{0, 1\} & \{1, 0.2, 0\} & 2 \\
\{1, 0, 1\} & \{0, 0\} & \{0, 1, 0.6\} & 3 \\
\{0, 0, 0\} & \{1, 0.3\} & \{1, 0, 1\} & 4 \\
\{0, 1, 1\} & \{0, 0\} & \{1, 0.3, 0\} & 11{=}2\langle x_2\rangle 7 \models 5, 7 \\
\{0, 0, 1\} & \{0, 0\} & \{1, 1, 0.6\} & 12{=}3\langle x_1\rangle 5 \models 6, 8 \\
\{0, 0, 1\} & \{1, 0\} & \{1, 0.3, 0.6\} & 13{=}4\langle x_3\rangle 6 \models 6 \\
\{0, 1, 1\} & \{0, 0.3\} & \{1, 0.2, 0\} & 14{=}5\langle x_3\rangle 9 \models 9 \\
\{1, 0, 1\} & \{1, 0\} & \{0, 0.3, 0.6\} & 15{=}3\langle x_3\rangle 13 \models 10 \\
\end{array}
$$

This is the final matrix of all prime disjunctions since we are not able to generate stronger disjunctions — all new resolvents are absorbed.

The fact that the matrix contains all prime disjunctions guarantees that projection on any variable obtained as a conjunction of projections of individual disjunctions is exactly equal to the real projection of the corresponding multi-dimensional distribution. Thus to carry out logical inference we impose our data represented as the conjunction **f** onto each prime disjunction and then combine all their projections onto the goal attribute. For example, in our case we have to change all 3-rd components in 1-st section (i.e., 3-rd column) onto 0 and then it becomes obvious that only disjunction **12** has non-trivial projection onto the attribute x_3 which is equal to the exact value $\{1, 1, 0.6\}$. If, in addition, we suppose that the patient does not have quantitative loss of self-control, i.e., the whole situation is described by the conjunction

$$\{1, 1, 0\}, \{0, 1\}, \{1, 1, 1\}$$

then two more disjunctions fire. Disjunction **4** has projection $\{1, 0.3, 1\}$, and disjunction **13** — $\{1, 0.3, 0.6\}$. The final conclusion is characterized by the projection $\{1, 0.3, 0.6\}$. We also can find additional information by finding projections on other attributes. In the above situation we see that disjunction **15** gives projections $\{1, 0.6, 1\}$ onto the attribute x_1, i.e., we can conclude that the patient does not have psychico-physical dependence (2-nd value of the 2-nd

attribute is disabled with the degree 0.6). By means of the matrix of all prime disjunctions we can also analyze knowledge base. In particular, we can find hidden dependencies, i.e., dependencies which follow from explicitly formulated disjunctions, or find symptoms which can influence one or another attribute.

It is obvious that the method of inference based on finding all prime disjunctions is inapplicable to real problem domains since their number is extremely high. However, it is not necessary to build all prime disjunctions — for decision support systems it is enough to carry out only several passes especially on early stages of debugging (more disjunctions can be generated when the knowledge base has been created and is ready for use). To compensate the loss of exactness multiple passes of inference can be conducted, i.e., conclusions obtained at earlier passes are used as data along with the old data at the next pass.

This approach to fuzzy knowledge representation and logical inference has been implemented in the fuzzy expert system shell EDIP for MS-Windows 3.x [12, 15].

4. Knowledge Discovery in Databases

Let us suppose that a database is represented by means of the matrix of conjunctions **K** which represents some distribution over the universe of discourse O. The problem is to express this distribution in the form of dependencies between attributes represented by disjunctions. Thus it is necessary to transform the matrix of conjunctions **K** representing data into the matrix of disjunctions **D** representing knowledge. Such a possibility is provided by a generalized covering method. This method just like the above described fuzzy resolution operation generalizes the conventional covering method in two directions: (i) it allows using multi-valued variables instead of only boolean ones, and (ii) it can be used with fuzzy matrices representing fuzzy distributions.

General covering method is intended for finding prime disjunctions given a matrix of conjunctions (DNF). Prime disjunctions are built gradually by means of searching the space of all possible disjunctions starting from the initial disjunction consisting of all 0 and then increasing its different components. As components increase they influence corresponding columns of the matrix and some its lines are *covered*. When all lines of the matrix are covered we say that the disjunction obtained is prime one (formally it has to be guarantied by the procedure). An order in which the procedure searches the space of disjunctions is determined by the order in which we increase the components. A line of the matrix is said to be covered if it is a consequence of the current disjunction being built.

An important property of this procedure is that it allows us to search only a part of the whole space of disjunctions which contains required dependencies between attributes. For example, if we are searching for dependencies between the first and the third attributes then we increase components in the corresponding sections only.

Let us suppose that we do not have an expert knowledge about stages of alcoholism but there is a database of case histories which can be transformed to the sectioned matrix (fuzzy DNF) consisting of 14 conjunctions

$$\mathbf{K} = \begin{array}{ccc} x_1 & x_2 & x_3 \\ \left| \begin{array}{ccc} \{1, 0, 0\} & \{1, 0\} & \{1, 0, 0\} \\ \{0.2, 0, 0\} & \{0.2, 0\} & \{0, 0.2, 0\} \\ \{1, 0, 0\} & \{0, 1\} & \{1, 0, 0\} \\ \{0.3, 0, 0\} & \{0, 0.3\} & \{0, 0.3, 0\} \\ \{0, 1, 0\} & \{1, 0\} & \{0, 1, 0\} \\ \{0, 0.6, 0\} & \{0.6, 0\} & \{0, 0, 0.6\} \\ \{0, 0.3, 0\} & \{0, 0.3\} & \{0, 0.3, 0\} \\ \{0, 1, 0\} & \{0, 0.6\} & \{0, 0, 0.6\} \\ \{0, 0, 1\} & \{1, 0\} & \{1, 0, 0\} \\ \{0, 0, 1\} & \{1, 0\} & \{0, 1, 0\} \\ \{0, 0, 1\} & \{1, 0\} & \{0, 0, 1\} \\ \{0, 0, 1\} & \{0, 1\} & \{1, 0, 0\} \\ \{0, 0, 0.3\} & \{0, 0.3\} & \{0, 0.3, 0\} \\ \{0, 0, 1\} & \{0, 1\} & \{0, 0, 1\} \end{array} \right| \begin{array}{c} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \\ \mathbf{7} \\ \mathbf{8} \\ \mathbf{9} \\ \mathbf{10} \\ \mathbf{11} \\ \mathbf{12} \\ \mathbf{13} \\ \mathbf{14} \end{array} \end{array}$$

(This matrix is semantically equivalent to the above matrix of disjunctions $\mathbf{D}$, i.e., they represent the same 3-dimensional fuzzy distribution.) It is required to discover a dependence between two attributes x_1 and x_3.

Disjunctions representing dependencies of two attributes contain components in the corresponding sections only (more exactly, they can contain components which are less than 1 also in other sections but for the sake of simplicity we will not touch this issue). Thus we have to gradually increase components in sections 1 and 3. It can be noticed that the third column (3-rd value of the 1-st attribute) contains the maximal number of non-0 components therefore it is natural to cover the maximum number of lines by increasing the corresponding component. Let us set 3-rd component of the disjunction to 1 and cover 6 last conjunctions **9**, **10**, **11**, **12**, **13**, **14** of the matrix which after removing them is equal to

$$\mathbf{K}_1 = \begin{array}{ccc} x_1 & x_2 & x_3 \\ \left| \begin{array}{ccc} \{1, 0, 0\} & \{1, 0\} & \{1, 0, 0\} \\ \{0.2, 0, 0\} & \{0.2, 0\} & \{0, 0.2, 0\} \\ \{1, 0, 0\} & \{0, 1\} & \{1, 0, 0\} \\ \{0.3, 0, 0\} & \{0, 0.3\} & \{0, 0.3, 0\} \\ \{0, 1, 0\} & \{1, 0\} & \{0, 1, 0\} \\ \{0, 0.6, 0\} & \{0.6, 0\} & \{0, 0, 0.6\} \\ \{0, 0.3, 0\} & \{0, 0.3\} & \{0, 0.3, 0\} \\ \{0, 1, 0\} & \{0, 0.6\} & \{0, 0, 0.6\} \end{array} \right| \begin{array}{c} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \\ \mathbf{7} \\ \mathbf{8} \end{array} \end{array}$$

Now we see that the first two columns both have 4 non-0 components. We can choose any of them (and we must do it in the case of searching for all dependencies) but let us choose the second column since it has greater sum of the components. Then the disjunction $\{0, 1, 1\}, \{0, 0\}, \{0, 0, 0\}$ covers 4 last lines **5**, **6**, **7**, **8** in the matrix $\mathbf{K}_1$ and we obtain

$$\mathbf{K}_2 = \begin{array}{ccc} x_1 & x_2 & x_3 \\ \left| \begin{array}{ccc} \{1, 0, 0\} & \{1, 0\} & \{1, 0, 0\} \\ \{0.2, 0, 0\} & \{0.2, 0\} & \{0, 0.2, 0\} \\ \{1, 0, 0\} & \{0, 1\} & \{1, 0, 0\} \\ \{0.3, 0, 0\} & \{0, 0.3\} & \{0, 0.3, 0\} \end{array} \right| & \begin{array}{c} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \end{array} \end{array}$$

Now we see that this matrix can be covered by the third section $\{1, 0.3, 0\}$ (the component 1 covers disjunctions **1**, **3**, and the component 0.3 — **2**, **4**). Thus we obtain that the disjunction

$$\{0, 1, 1\}, \{0, 0\}, \{1, 0.3, 0\}$$

covers the whole matrix of conjunctions and as such it represents a fuzzy dependence between attributes x_1 and x_3. Notice that it is equal to the prime disjunction **11** from the matrix $\mathbf{D}_3$.

5. Conclusion

In the paper it has been described how new original results from the field of multi-dimensional fuzzy analysis can be applied to the problem of decision making and rule induction. In particular, a uniform and efficient method of sectioned matrices has been proposed for extensional and intensional representation of multi-dimensional fuzzy distributions. If knowledge is represented by means of such a matrix interpreted as fuzzy CNF and data is represented by means of a sectioned vector interpreted as conjunction, then logical inference consists in finding projection of the whole fuzzy distribution on the goal attribute. There has been described a procedure for solving this problem based on the operation of fuzzy resolution. If data is represented by a sectioned matrix interpreted as fuzzy DNF, then the rule induction consists in finding the most general disjunctions which cover the fuzzy distribution it represents. We described a generalized fuzzy covering method to find such rules. The main problem of multi-dimensional fuzzy analysis consists in computational difficulty. For two algorithms described in the paper we have proposed an approximate variant which can be applied to real tasks.

6. References

[1] Levchenko, V.I., 1990, Diagnostic system based on finite predicates, Preprint, Institute of Mathematics and CC AS Moldavia, Kishinev, Shtiintsa (in Russian).

[2] Zakrevsky, A.D., 1989, Logical inference in finite predicates, Preprint No.6, Institute of Technical Cybernetics, AS Belorussia, Minsk (in Russian).

[3] Savinov, A.A., 1991, Matrix representation of fuzzy knowledge in attribute models, Preprint, Institute of Mathematics and CC, AS Moldavia, Kishinev, Shtiintsa (in Russian).

[4] Levchenko, V.I., and Savinov, A.A., 1994, Diagnosis by fuzzy constraints in attribute model, 2nd Eur. Congr. on Intelligent Techniques and Soft Computing—

EUFIT'94, Aachen, Germany, September 20-23, 382-385.

[5]   Levchenko, V.I., and Savinov, A.A., 1993, The matrix representation of fuzzy knowledge and its application to the expert systems design, Computer Sci. J. of Moldova **1**(1), 62-84.

[6]   Levchenko, V.I., and Savinov, A.A., 1993, Matrix representation of fuzzy predicates and its application in expert systems, Izvestia RAN, Tehnicheskaia kibernetika No.5, 1993, 126-140 (in Russian).

[7]   Zadeh, L.A., 1975, The concept of a linguistic variable and its application to approximate reasoning — Part I. — Information Sciences **8**, pp. 301-357.

[8]   Zadeh, L.A., 1979, A theory of approximate reasoning, In: Machine Intelligence **9** (Hayes J.E., Michie D. and Mikulich L.I., Eds.). — New York: Elsevier, pp. 149-194.

[9]   Savinov, A.A., 1993, Fuzzy propositional logic, Fuzzy Sets and Systems **60**(1), 9-17.

[10]   Savinov, A.A., 1993, Fuzzy propositional logic for the knowledge representation, First European Congress on Fuzzy and Intelligent Technologies—EUFIT'93, Aachen, Germany, September 7-10.

[11]   Savinov, A.A., 1996, Some properties of new resolution rule in the logic of possibility distributions, 4th European Congress on Intelligent Techniques and Soft Computing—EUFIT'96, Aachen, Germany, September 2-5, 178-182.

[12]   Levchenko, V.I., and Savinov, A.A., 1992, The representation of fuzzy knowledge in the diagnostic expert system shell EDIP, Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Networks—IIZUKA'92, Iizuka, Japan, July 17-22.

Online references

[13]   Savinov, A.A., 1997, Fuzzy Multi-dimensional Analysis, online paper, http://www.geocities.com/ResearchTriangle/7220/fmda_0.html.

[14]   Savinov, A.A., 1997, New Fuzzy Resolution Operation, online paper, http://www.geocities.com/ResearchTriangle/7220/fresol.html.

[15]   Fuzzy expert system shell EDIP, http://www.geocities.com/ResearchTriangle/7220/edip.zip