

Towards Application of FML in Suspicion of Non-Common Diseases

Giovanni Acampora*, Tatiana Kiseliova[†], Karaman Pagava[‡] and Autilia Vitiello*

*Dept. of Computer Science

University of Salerno, Fisciano, Salerno 84084

Email: {gacampora, avitiello}@unisa.it

[†]Dept. of Computer Sciences

Ivane Javakhishvili Tbilisi State University, Tbilisi, Georgia 0143

Email: t.kiseliova@web.de

[‡] Child and Adolescent Medicine

Tbilisi Medical State University, Tbilisi, Georgia 0177

Abstract—In this paper we present the preliminary results of application of Fuzzy Markup Language (FML) to suspect a non-common disease. Under non-common diseases we understand rare diseases. From the broad point of view this problem belongs to the computer-assisted decision support in medical diagnostics and can be supported by fuzzy logic controllers. We can use conventional methods to diagnose a rare disease if it can be exhibited by outstanding symptoms. For example, there are several search machines and data banks that allow to find a rare disease clearly exhibited by a patient's symptoms/signs. But it is very difficult to diagnose a rare disease if it masks as a common disease. Diagnostic of rare diseases is connected with lack, uncertainty and imprecision of knowledge, medical mistake and even medical failure. Additionally, very often a common disease is also established with some degree of belief, thus, the expressions such as “it is possible that a patient has a particular disease” rather often present in the daily medical practice. It is clear that if we would know the common diseases, then deviations from them can be considered as a sign of non-common diseases. In this paper we investigate such deviations with the help of FML. We show how FML mechanism can be adjusted to suspect a rare disease, and discuss the appropriateness of the available operators.

Index Terms—Fuzzy Systems, Fuzzy Markup Language, Medical Diagnosis, Rare diseases.

I. INTRODUCTION

In this paper we present the preliminary results of application of Fuzzy Markup Language (FML) to suspect a non-common disease. Under non-common diseases we understand rare diseases. A rare disease is any disease that affects a small percentage of the population. There is no single, widely accepted definition for rare diseases (RD)¹. According to different definitions the prevalence of RD should be between 1/200000 and 1/1000, usually about 1/2000. There are several thousand rare diseases. Their total number is quite significant: the 5-6% of population have some rare diseases².

From the broad point of view this problem belongs to the computer-assisted decision support in medical diagnostics and can be supported by fuzzy logic controllers. It is possible to diagnose a rare disease if it can be exhibited by outstanding symptoms. For example, there are several search machines and data banks that allow to find a rare disease clearly exhibited by a patient's symptoms/signs³. But it is very difficult to diagnose a rare disease if it masks as a common disease. Diagnostic of rare diseases is connected with lack, uncertainty and imprecision of knowledge, medical mistake and even medical failure. From another side, a common disease also is established with some degree of belief, so the expressions such as “it is possible that a patient has a particular disease” rather often present in the daily medical practice.

It is clear that if we would know the common diseases, then deviations from them can be considered as a sign of non-common diseases. If these deviations reach a certain level, a suspicious of a rare disease exists. Thus, a physician obtains a sign of “non-normality” of a case in hand. Then a search of a rare disease can be continued in the specific databases, cited above.

In our approach we use a methodology based on fuzzy logic, because the elements of the diagnostic process (symptoms, diagnoses, connections between both) and a diagnostic process itself are mostly described imprecisely and approximately. First, we define what we understand under common (normal) diseases and then consider possible deviations from them. Our description of common (normal) diseases is based on the terminology used in CADIAG-like systems [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], whereas, we investigate the deviations from normal diseases with the help of FML [13], [14]. In the last years, FML (Fuzzy Markup Language) is emerging as one of the most efficient and useful language to define fuzzy control thanks to its capability of modeling Fuzzy Logic Controllers in a human-readable and hardware inde-

¹http://en.wikipedia.org/wiki/Rare_disease

²<http://rarediseases.info.nih.gov/>

<http://rarediseases.about.com/>

http://ec.europa.eu/health-eu/health_problems/rare_diseases/index_en.htm/

³<http://www.orpha.net/consor/cgi-bin/index.php?lng=EN>

<http://www.crdnetwork.org/>

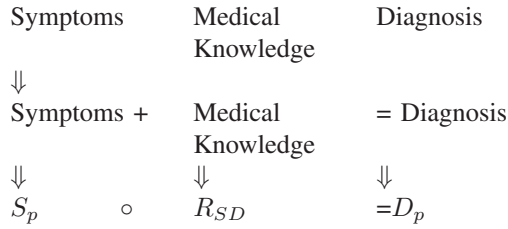
<http://www.rarediseases.org/search/rdblist.html>

pendent way, i.e., the so-called Transparent Fuzzy Controllers (TFCs) [15]. We show how FML mechanism can be adjusted to suspect a rare disease, and discuss the appropriateness of the available operators.

In next Section we recall some necessary definition and denotations. In the Section III we present a brief FML description. The results of FML applications to suspect a rare disease on examples are reported in the Section IV. Section V concludes our work.

II. PRELIMINARIES. NOTIONS AND DENOTATIONS

We present a medical decision support mechanism based on the fuzzy approach generally as follows. Three components take part in the process: symptoms/signs of the patient to whom a diagnosis should be established, medical knowledge (written as “IF S THEN D ”, but formalized as fuzzy relations) available before the process starts and diagnosis that we are trying to find. A schema of the process is presented below.



R_{SD} is a fuzzy relation between symptoms/signs and diseases is defined as

$$R_{SD} : \Sigma \times \Delta \rightarrow [0, 1] \quad (1)$$

$$S_p : \Sigma \rightarrow [0, 1] \quad (2)$$

$$D_p : \Delta \rightarrow [0, 1] \quad (3)$$

where $\Sigma = \{s_1, \dots, s_m\}$ is the crisp set of symptom-signs, $\Pi = \{p_1, \dots, p_r\}$ is the crisp set of patients, $\Delta = \{d_1, \dots, d_n\}$ is the crisp set of diseases, each element of the fuzzy set S_p shows to which degree it is true, that a patient $p \in \Pi$ has symptom/sign s_i , or a degree of possibility of the presence of the symptom, or its severity; symbol D_p is used for the final diagnosis for a patient $p \in \Pi$ and each element of the fuzzy set D_p shows to which degree it is true, that a patient $p \in \Pi$ has a given disease d_j , or a degree of possibility with which we can attach each relevant diagnostic label to the patient [16]. R_{SD} can be of different types, it depends what information is available from experts or manuals. For example, R_{SD} can be considered as *confirmation* relation (R_{SD}^c), *occurrence* relation (R_{SD}^o), *exclusion* relation (R_{SD}^e), *temporal* or *time of manifestation* relation (R_{SD}^t). R_{SD}^c shows the degree to which a symptom/sign confirms the presence of disease, R_{SD}^o shows the frequency of occurrence of symptom/sign with disease, R_{SD}^e indicates the degree in which the present symptom/sign (combination) excludes (or disconfirms) the disease, R_{SD}^t relation shows to which degree it is true that exhibition of a symptom leads to immediate manifestation of a disease (or a disease is manifested after some period of time).

	d_1	d_2	...	d_n
s_1	$f(s_1 d_1)$	$f(s_1 d_2)$	...	$f(s_1 d_n)$
s_2	$f(s_2 d_1)$	$f(s_2 d_2)$	...	$f(s_2 d_n)$
$\vdots$	$\vdots$	$\vdots$	$f(s_i d_j)$	$\vdots$
s_m	$f(s_m d_1)$	$f(s_m d_2)$	...	$f(s_m d_n)$

TABLE I: Numerical representation of symptom/sign - disease connections for R_{SD}^o

never,	very seldom,	seldom,
sometimes,	unspecified,	occasionally,
often,	very often,	always

TABLE II: Linguistic values

$R_{SD}^o(s_i, d_j)$ and $R_{SD}^c(s_i, d_j)$ can be derived from relative frequencies, i.e., numerically [5]:

$$R_{SD}^o(s_i, d_j) = f(s_i|d_j) \quad (4)$$

and

$$R_{SD}^c(s_i, d_j) = f(d_j|s_i) \quad (5)$$

where $f(s_i|d_j) = \frac{f(d_j \cap s_i)}{f(d_j)}$, $f(d_j|s_i) = \frac{f(d_j \cap s_i)}{f(s_i)}$; $f(s_i|d_j)$ is a conditional frequency of s_i given d_j , $f(d_j|s_i)$ is a conditional frequency of d_j given s_i , $f(d_j \cap s_i)$ is the absolute frequency of joint occurrence of d_j and s_i , $f(d_j)$ and $f(s_i)$ are absolute frequencies of d_j and s_i , correspondingly. We call this numerical interpretation.

An example of a numerical description is presented in the Table I.

A linguistic way opens a possibility to estimate relations R_{SD}^o , R_{SD}^c , R_{SD}^e , R_{SD}^t (linguistic variables) using fuzzy sets (values of linguistic variables), e.g., presented in the Table II.

The fuzzy sets (Table II) are defined as mappings from $[0, 1]$ to $[0, 1]$. An example of a linguistic description of symptom/sign - disease connection is presented in the Table III. Each linguistic value can be defined by an expert [17]. For example, an expert can define his own fuzzy set *often* and supports the interpretation with the following sentence: symptom/sign and disease meet each other often, if it happens in, approximately, 60%. Such way fuzzy relations of type II [16] are defined. An example of linguistic values from the scale (Table II) as triangular fuzzy sets is shown in the Table IV.

Another possibility to define linguistic values (simple one) is to assign to them an ordered set of numbers, e.g., from $[0, 1]$ as shown in the Table V.

	d_1	d_2	...	d_n
s_1	often	always	...	never
s_2	never	often	...	always
$\vdots$	$\vdots$	$\vdots$	f_{ij}	$\vdots$
s_m	often	seldom	...	always

TABLE III: Linguistic representation of symptom/sign - disease connections

<i>never</i>	<i>very seldom</i>	<i>seldom</i>
(0, 0.1, 0.2)	(0.1, 0.2, 0.3)	(0.2, 0.3, 0.4)
<i>sometimes</i>	<i>unspecified</i>	<i>occasionally</i>
(0.3, 0.4, 0.5)	(0.4, 0.5, 0.6)	(0.5, 0.6, 0.7)
<i>often</i>	<i>very often</i>	<i>always</i>
(0.6, 0.7, 0.8)	(0.7, 0.8, 0.9)	(0.8, 0.9, 1)

TABLE IV: Representation of linguistic terms as triangular fuzzy sets

<i>never</i>	<i>very seldom</i>	<i>seldom</i>
0	0.2	0.3
<i>sometimes</i>	<i>unspecified</i>	<i>occasionally</i>
0.4	0.5	0.6
<i>often</i>	<i>very often</i>	<i>always</i>
0.7	0.8	1.0

TABLE V: Representation of linguistic terms as numbers

In our work we use the scale (shown in Table II) to estimate the relationships of the symptoms/signs and diseases.

A max – min composition of fuzzy relations [18], [19] or its generalised version t -conorm- t -norm composition can be used for medical diagnosing. Thus,

$$D_p = S_p \circ R_{SD} \quad (6)$$

is a composition of a fuzzy set and a fuzzy relation.

Moreover $\forall d_j \in \Delta$

$$D_p(d_j) = \vee_{s_i \in \Sigma} \wedge \{S_p(s_i); R_{SD}(s_i, d_j)\} \quad (7)$$

where $D_p : \Delta \rightarrow [0, 1]$ are inferred possible diagnoses for the patient $p \in \Pi$, $\vee$ is a t -conorm and $\wedge$ is a t -norm.

In general, the aggregation operators can be used in this inference mechanism [16].

Working with a decision-support system, a physician expects from a computer program a tip, a help, what diagnosis it can be for a patient at hand. In this way, the system should alarm if some things are outside of its normal functioning, i.e., if the case is neither confirmed nor possible, or has “strange” values. Thus, such behaviour of a system could be considered as a sign of a possible rare disease.

It is obvious that conclusions about the patient’s diagnosis with a help of a computer program is made based on the information available, in particular, what initial information is contained in the knowledge base and what information about a patient at hand we have. It is possible to investigate how to suspect a rare disease under the presence of different initial information, in our denotation R_{SD}^o , R_{SD}^c , R_{SD}^e or R_{SD}^t . In the section IV we consider relation R_{SD}^o and simulate the inference process with FML. For this reason, a description of this novel language is provided in the following section.

III. FUZZY MARKUP LANGUAGE: A DESCRIPTION

Fuzzy Markup Language (FML) is a XML-based domain-specific language proposed by Acampora and Loia [13] whose main aim is to model fuzzy systems by directly dealing with

fuzzy concepts, fuzzy rules and fuzzy inference engines [8]. In particular, it is used to model two well-known kind of fuzzy controllers: Mamdani and Takagi-Sugeno-Kang (TSK). FML programs are coded through a collection of correlated semantic tags capable of modeling different components of a controller by exploiting abstraction benefits offered by XML tools. Differently from other similar approaches used to describe fuzzy controllers such as Fuzzy Control Language (FCL) or MATLAB Fuzzy Inference System (FIS) developed by the Matworks, FML allows fuzzy designers to simply code their ideas on heterogeneous hardware without understanding details related to the different platforms. In short, thanks to FML it is possible to model a fuzzy controller in human-readable and hardware independent way. The exploitation of FML is resulted in a new fuzzy control design philosophy named *transparent fuzzy controller* [20] based on the integration among fuzzy systems, labeled tree theory [21], and XML methodology.

In the following, a brief description of the used FML tags is given. The root tag of FML program is `<FuzzyController>` which has two attributes: name and ip. The *name* attribute permits to specify the name of fuzzy controller and *ip* is used to define the location of controller in a computer network. The fuzzy knowledge base is defined by means of the tag `<KnowledgeBase>` which maintains the set of fuzzy concepts used to model the fuzzy system. In order to define a fuzzy concept, the following set of nested tags is used:

- `<FuzzyVariable>` defines the fuzzy concept;
- `<FuzzyTerm>` defines a linguistic term describing the fuzzy concept;
- a set of tags defining a shape of fuzzy sets related to fuzzy terms.

The attributes of `<FuzzyVariable>` tag are: name, scale, domainLeft, domainRight, type and, for only output concept, accumulation, defuzzifier and defaultValue. The *name* attribute defines the name of fuzzy concept; *scale* is employed to define the scale used to measure the fuzzy concept, *domainLeft* and *domainRight* are used to model the universe of discourse of fuzzy concept, i.e, the set of real values related to fuzzy concept; the position of fuzzy concept into rule (consequent part or antecedent part) is defined by *type* attribute (input/output); *accumulation* attribute defines the method of accumulation that is a method that permits the combination of results of a variable of each rule in a final result; *defuzzifier* attribute defines the method used to execute the conversion from a fuzzy set, obtained after aggregation process, into a numerical value to give it in output to system; *defaultValue* attribute defines a real value used only when no rule has fired for the variable at issue. `<FuzzyTerm>` uses two attributes: *name* used to identify the linguistic value associate with fuzzy concept and *complement*, a boolean attribute that defines, if it is true, it is necessary to consider the complement of membership function defined by given parameters. Fuzzy shape tags, used to complete the definition of fuzzy concept, are several, in our case of study, the triangular membership functions has been used through `<TriangularShape>` tag. Each shaping tag uses a set of attributes which defines the real outline of corresponding

fuzzy set. In particular, the $\langle \text{TriangularShape} \rangle$ tag uses three parameters. The rule base is modeled by $\langle \text{RuleBase} \rangle$ tag which defines a fuzzy rule set. The $\langle \text{RuleBase} \rangle$ tag uses five attributes: name, type, activationMethod, andMethod and orMethod. Obviously, the *name* attribute uniquely identifies the rule base. The *type* attribute permits to specify the kind of fuzzy controller (Mamdani or TSK) respect to the rule base at issue. The *activationMethod* attribute defines the method used to implication process; the *andMethod* and *orMethod* attribute define, respectively, the and and or algorithm to use by default. In order to define the single rule the $\langle \text{Rule} \rangle$ tag is used. The attributes used by the $\langle \text{Rule} \rangle$ tag are: name, connector, operator and weight. The *name* attribute permits to identify the rule; *connector* is used to define the logical operator used to connect the different clauses in antecedent part (and/or); *operator* defines the algorithm to use for chosen connector; *weight* defines the importance of rule during inference engine time. The definition of antecedent and consequent rule part is obtained by using $\langle \text{Antecedent} \rangle$ and $\langle \text{Consequent} \rangle$ tags. $\langle \text{Clause} \rangle$ tag is used to model the fuzzy clauses in antecedent and consequent part. This tag use the attribute *modifier* to describe a modification to term used in the clause. To complete the definition of fuzzy clause the nested $\langle \text{Variable} \rangle$ and $\langle \text{Term} \rangle$ tag have to be used.

The FML tags represent the set of lexemes used to build a fuzzy controller. However, in order to realize a well-formed XML-based language, an FML context-free grammar is necessary. The FML context-free grammar is modeled by XML file in the form of a XML Schema Document (XSD). Moreover, since an FML program realizes only a static view of a fuzzy system, the so-called eXtensible Stylesheet Language Translator (XSLT) is provided to change this static view to a computable version. In particular, the XSLT technology is used to convert a fuzzy controller description into a general-purpose computer language to be computed on several hardware platforms.

Finally, the FML is essentially composed by three layers:

- XML in order to create a new markup language for fuzzy logic control;
- a XML Schema Document in order to define the legal building blocks;
- extensible stylesheet language transformations (XSLT) in order to convert a fuzzy controller description into a specific programming language.

IV. FML IN SIMULATION

In this section, the simulation with FML of the inference process used to suspect a rare disease under the presence of the initial information R_{SD}^o is shown and discussed.

In particular, we illustrate the use of FML to suspect a rare disease by means of an example. Assume, that a knowledge base contains R_{SD}^o as, for example, in the Table VI and corresponding numerical form (due to the scale Table V) as singletons and/or triangular fuzzy sets. In this case $\Delta = \{d_1, d_2\}$, $\Sigma = \{s_1, s_2, \dots, s_6\}$. We assume that a rare

	<i>Bronchitis</i>	<i>Pneumonia</i>
<i>wheezing</i>	sometimes	occasionally
<i>breathlessness</i>	very seldom	often
<i>anosmia</i>	seldom	seldom
<i>fever</i>	sometimes	often
<i>sinusitis</i>	sometimes	sometimes
<i>chronicotitis</i>	seldom	seldom

TABLE VI: Linguistic descriptions for “normal” diseases *Bronchitis/Pneumonia*

	d_1	d_2
s_1	0.4	0.6
s_2	0.2	0.8
s_3	0.3	0.3
s_4	0.4	0.8
s_5	0.4	0.4
s_6	0.3	0.3

TABLE VII: Corresponding numerical descriptions for “normal” diseases (see Table VI)

disease is not in the knowledge base, it means that it is not in Δ .

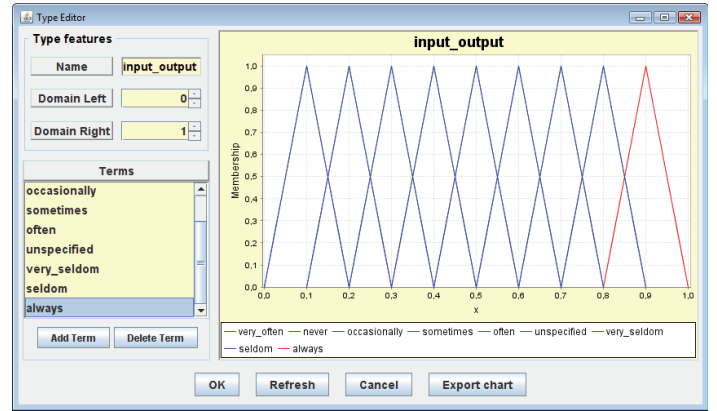


Fig. 1: Presentation of linguistic values from Table V in FML

We define nine linguistic values from Table IV as shown in Figure 1.

Thus, we use the same type of input and output variables in FML, in particular, triangular shape.

As input variables we consider 6 symptoms: wheezing, breathlessness, anosmia, fever, sinusitis, chronicotitis. As output variables we consider two diseases: Bronchitis and Pneumonia. Due to the Table VI we have constructed the following rule base (see Figure 2).

In the listing 1, a portion of the FML-based fuzzy controller to suspect a rare disease is presented.

As we speak about a rare disease, we consider how “far” is this case from the common representation.

In the Table VII we present corresponding singletons for linguistic values from Table VI for the explanation reason.

We can predefined an excluded diagnoses in the case if a symptom is fully present/absent with a corresponding full

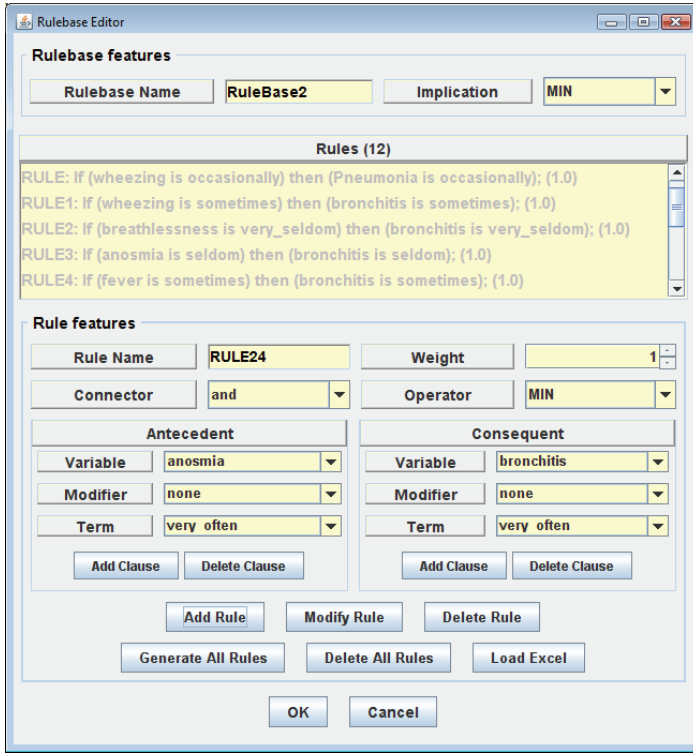


Fig. 2: 12 rules describing relation of six symptoms and two diseases in FML

occurrence/absent of occurrence. Simply speaking, at least one element of the Table VII should be 1 or 0 and at least one membership degree of fuzzy set S_p should be 1 or 0 as well. It means, that, to define such case, we need a crisp information. It is a clear situation, but is not always a case.

As we have already mentioned, the difficulty to suspect a rare disease is because of its symptomatic similarity to the “normal cases” and, of course, absence of the necessary information. If a physician has enough experience with a computer program, s/he usually has been working, and s/he mentions that the values for possible diagnoses are much below or much above the usual representation, it can be also the sign of alarm “a rare disease” for a physician.

We need to define appropriate operations and thresholds (that allow to detect a deviation from the “normal” case). If we take max-min composition of fuzzy relations (7), it restricts the maximum membership degree of possible diseases for a patient by values in Table VII. For example, if a patient at hand has

$$S_p^1 = \{(s_1, 0.4), (s_2, 0.2), (s_3, 0.3), (s_4, 0.4), (s_5, 0.4), (s_6, 0.2)\}$$

the possible diagnosis d_1 has the membership degree 0.4 (see (8)). If a patient at hand has

$$S_p^2 = \{(s_1, 1.0), (s_2, 1.0), (s_3, 0.2), (s_4, 0.3), (s_5, 0.2), (s_6, 0.2)\}$$

the possible diagnosis d_1 has the membership degree 0.4

```
<?xml version="1.0" encoding="UTF-8">
<FuzzyController name="Mamdani_controller" ip="127.0.0.1">
  <KnowledgeBase>
    <FuzzyVariable name="wheezing" domainleft="0.0" domainright="1.0" scale=""
      type="input">
      <FuzzyTerm name="always" complement="false">
        <TriangularShape Param1="0.8" Param2="0.9" Param3="1.0"/>
      </FuzzyTerm>
      <FuzzyTerm name="never" complement="false">
        <TriangularShape Param1="0.0" Param2="0.1" Param3="0.2"/>
      </FuzzyTerm>
      <FuzzyTerm name="occasionally" complement="false">
        <TriangularShape Param1="0.5" Param2="0.6" Param3="0.7"/>
      </FuzzyTerm>
      <FuzzyTerm name="often" complement="false">
        <TriangularShape Param1="0.6" Param2="0.7" Param3="0.8"/>
      </FuzzyTerm>
      <FuzzyTerm name="seldom" complement="false">
        <TriangularShape Param1="0.2" Param2="0.3" Param3="0.4"/>
      </FuzzyTerm>
      <FuzzyTerm name="sometimes" complement="false">
        <TriangularShape Param1="0.3" Param2="0.4" Param3="0.5"/>
      </FuzzyTerm>
      <FuzzyTerm name="unspecified" complement="false">
        <TriangularShape Param1="0.4" Param2="0.5" Param3="0.6"/>
      </FuzzyTerm>
      <FuzzyTerm name="very_often" complement="false">
        <TriangularShape Param1="0.7" Param2="0.8" Param3="0.9"/>
      </FuzzyTerm>
      <FuzzyTerm name="very_seldom" complement="false">
        <TriangularShape Param1="0.1" Param2="0.2" Param3="0.3"/>
      </FuzzyTerm>
    </FuzzyVariable>
    .....
    <FuzzyVariable name="bronchitis" domainleft="0.0" domainright="1.0"
      scale="" defaultvalue="0.0" accumulation="MEAN"
      defuzzifier="MM" type="output">
      <FuzzyTerm name="always" complement="false">
        <TriangularShape Param1="0.8" Param2="0.9" Param3="1.0"/>
      </FuzzyTerm>
      <FuzzyTerm name="never" complement="false">
        <TriangularShape Param1="0.0" Param2="0.1" Param3="0.2"/>
      </FuzzyTerm>
      <FuzzyTerm name="occasionally" complement="false">
        <TriangularShape Param1="0.5" Param2="0.6" Param3="0.7"/>
      </FuzzyTerm>
      <FuzzyTerm name="often" complement="false">
        <TriangularShape Param1="0.6" Param2="0.7" Param3="0.8"/>
      </FuzzyTerm>
      <FuzzyTerm name="seldom" complement="false">
        <TriangularShape Param1="0.2" Param2="0.3" Param3="0.4"/>
      </FuzzyTerm>
      <FuzzyTerm name="sometimes" complement="false">
        <TriangularShape Param1="0.3" Param2="0.4" Param3="0.5"/>
      </FuzzyTerm>
      <FuzzyTerm name="unspecified" complement="false">
        <TriangularShape Param1="0.4" Param2="0.5" Param3="0.6"/>
      </FuzzyTerm>
      <FuzzyTerm name="very_often" complement="false">
        <TriangularShape Param1="0.7" Param2="0.8" Param3="0.9"/>
      </FuzzyTerm>
      <FuzzyTerm name="very_seldom" complement="false">
        <TriangularShape Param1="0.1" Param2="0.2" Param3="0.3"/>
      </FuzzyTerm>
    </FuzzyVariable>
    .....
  </KnowledgeBase>
  <RuleBase name="RuleBase1" andMethod="MIN" orMethod="MAX"
    activationMethod="MIN" type="mamdani">
    <Rule name="RULE1" connector="and" operator="MIN" weight="1.0">
      <Antecedent>
        <Clause>
          <Variable>wheezing</Variable>
          <Term>sometimes</Term>
        </Clause>
      </Antecedent>
      <Consequent>
        <Clause>
          <Variable>bronchitis</Variable>
          <Term>sometimes</Term>
        </Clause>
      </Consequent>
    </Rule>
    .....
  </RuleBase>
</FuzzyController>
```

Listing 1: FML code to model a portion of the FML-based fuzzy controller to suspect a rare disease

as well. It means, that the values from the initial tabel (in our case R_{SD}^0) bound from above the possible exhibitions of symptoms/signs of the patient. With operator min is not possible to see deviations.

	d_1	d_2
S_p^1	0.4	0.4
S_p^2	0.4	0.8
S_p^3	0.4	0.7

TABLE VIII: Possible degrees for “normal” diseases under max-min composition

$$D_p(d_1) = \max \left\{ \begin{array}{l} \min\{0.4, 0.4\}; \\ \min\{0.2, 0.2\}; \\ \min\{0.3, 0.3\}; \\ \min\{0.4, 0.4\}; \\ \min\{0.4, 0.4\}; \\ \min\{0.3, 0.3\}; \end{array} \right\} = 0.4 \quad (8)$$

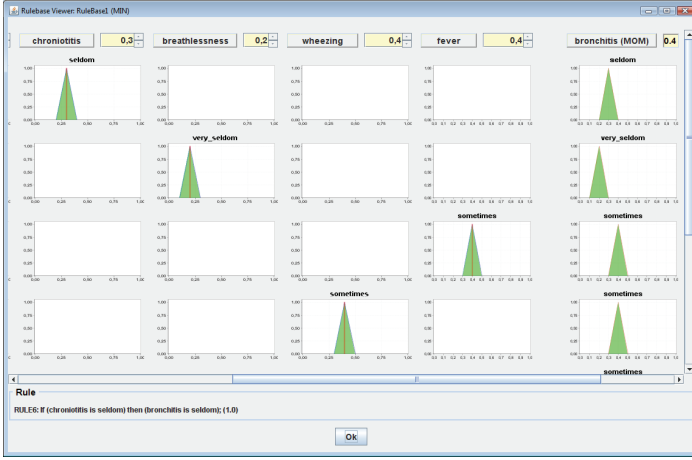


Fig. 3: Results of execution of FML under S_p^1 for d_1

This standard max – min composition in our case guarantees that the degree of occurrence can be less than 0.4, but not more (see Figure 3, Table VIII). Using FML, the antecedent and consequent of a rule in the rule base should be defined as fuzzy sets. We, however, consider R_{SD} relation is a given one, we do not construct it from fuzzy sets described symptoms/signs and diseases. The omission of this step can be solved by assigning the same linguistic values to the antecedent variable and the consequent variable. In general, due to our approach to define deviations from “common” diseases as a sign of non-normality, and consequently, a suspicious of a rare disease, we need operators, that allow to estimate such deviation. For example, in controllers, realized in the FML, an input value and antecedent value are connected by ‘min’ operator, but if we need to know, how far is an input value from the antecedent value, we need a corresponding operator that estimates the distance between two fuzzy sets.

Let us continue to discuss our considerations on examples.

If we take the *absolute difference* (for two numbers a and b their *absolute difference* is defined as $|a - b|$) instead of min-operation (max-operation remains unchanged) we obtain more flexibility to detect deviations. For example, for S_p^1 the

	d_1	d_2
S_p^1	0.0	0.13
S_p^2	0.23	0.24
S_p^3	0.40	0.26

TABLE IX: Possible degrees for “normal” diseases under *mean – absolute difference* composition

membership degree of d_1 in this case is 0 and for S_p^2 we obtain 0.8. If instead of the *max* operator we take another aggregation operator, such as the *arithmetic mean* of deviations (the *arithmetic mean* of $a_1, a_2, \dots, a_n$ is $\frac{a_1 + a_2 + \dots + a_n}{n}$), for d_1 in the case of S_p^1 we obtain 0 as well and 0.23 in the case of S_p^2 . We interpret 0 as a sign of most believable diagnosis, and every deviation from 0 points less believable diagnosis. The question is, as always in applications, how to detect a threshold, under (or above) which a suspicious of a rare disease arises. If we take *max – absolute difference* composition of S_p and R_{SD} and a result is different from 0, the final value shows that at least one symptom/sign differs from the normal case, but does not say anything about the rest of symptoms. When a rare disease can be pointed by a representative symptom (i.e., such symptom which presence/absence is a sure sign of the non-normality of a diagnosis) max- *absolute difference* is suitable.

The *mean – absolute difference* composition of S_p and R_{SD} shows the mean value of all deviations. As it is often with rare diseases, they are described with the same symptoms/signs as “normal” diseases and only a certain deviation can point a rare disease, the application of *mean – absolute difference* composition of S_p and R_{SD} seems to be reasonable. Notice, that in general, a p – *mean*, and/or other distance metrics can be considered.

$$S_p^3 = \{(s_1, 0.7), (s_2, 0.7), (s_3, 0.7), (s_4, 0.7), (s_5, 0.8), (s_6, 0.7)\}$$

The possible degrees for d_1, d_2 under S_p^1, S_p^2, S_p^3 using max – min composition are presented in the Table VIII.

The possible degrees for d_1, d_2 under S_p^1, S_p^2, S_p^3 using *mean – absolute difference* composition are presented in the Table IX.

For this particular example 0.25 can be chosen as a threshold to point a suspicion of a rare disease, i.e., if all numbers in one string in the Table IX are above 0.25 a physician should think about “non-normality” of a possible diagnosis for a patient at hand. It means, that we found an approach to suspect a rare disease if the diagnoses from the initially defined list seems to be possible. The choice of a threshold remains on experts. There are many ways to formalise this procedure[22], [23], [24], which approach has to be used is a subject of a separate investigation.

Notice, that the difference between symptoms of a given patient and the predefined symptoms/signs picture of a disease may be also calculated as ratio $\frac{a}{b}$ and then also a mean value of these ratios may be calculated. In this case a value for a normal case will be 1.

So, we need an operation that will calculate a distance

between two fuzzy sets in the inference mechanism. It is a subject of our future applications.

It can be also argued that one number, as a threshold, does not reflect a flexibility in diagnostics. In forthcoming time we investigate interval- and fuzzy sets estimations of linguistic expressions for a rare disease diagnostic problem. In this paper our intend was to discuss the principal approach to diagnostics of rare diseases and show the first steps for its realisation with the help of FML.

V. CONCLUSIONS

The problem to suspect a rare disease (in particular, in pediatrics) with a help of computer - assisted system is an important problem and till now is not well investigated. This problem is connected with absence, insufficiency, uncertainty and imprecision of information and, therefore, belongs to the category of problems, that can be solved by fuzzy logic approaches.

Our main idea to suspect a rare disease is based on the assumption, that deviation from a "normal" disease is a sign to warn a rare disease. Thus, well acquaintance with normal diseases involves a recognition of a rare disease. We suppose, that a physician has already established a preliminary diagnosis for a patient, but s/he hesitates about its correctness. Then a physician, based on the patient's data, checks the deviations from the "normal" cases, and if these deviations reach certain level, a suspicion of a rare disease arises.

We test our approach on FML and preliminary results have shown that we are able successfully model the process of suspicious of a rare disease, but for flexibility we need additional operators, that allow to estimate a deviation of one fuzzy set from another. It is a subject of our recent collaborative work.

ACKNOWLEDGMENT

This work was supported by Georgia National Science Foundation, Ministry of Education and Science of Georgia, Grant GNSF/ST08/6-460, Grant STCU-GNSF #5015 and the Grant #GNSF(ST08/1-361), Georgian National Science Foundation. Any idea in this publication is possessed by the authors and may not represent the opinion of the Georgian National Science Foundation itself.

REFERENCES

- [1] K.-P. Adlassnig and G. Kolarz. CADIAG-II: computer-assisted medical diagnosis using fuzzy subsets. In M.M. Gupta and E. Sanchez, editors, *Approximate Reasoning in Decision Analysis*, pages 495–505. North Holland Publishing Company, 1982.
- [2] K. Boegl. *Design and Implementation of a Web-Based Knowledge Acquisition Toolkit for Medical Expert Consultation Systems*. PhD thesis, Technische Universität Wien, 1997.
- [3] K. Boegl, K.-P. Adlassnig, Y. Hayashi, T. Rothenfluh, and H. Leitich. Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system. *Artificial Intelligence in Medicine*, 30:1–26, 2004.
- [4] M. Daniel. Theoretical comparison of inference in CADIAG and MYCIN-like systems. *Tatra Mountains Mathematical Publications*, 16:255–272, 1999.
- [5] M. Daniel, P. Hájek, and H. Nguyen. CADIAG-2 and MYCIN-like systems. *Artificial Intelligence in Medicine*, 9:241–259, 1997.
- [6] P. Hájek, T. Havránek, and R. Jiroušek. *Uncertain Information Processing in Expert Systems*. CRC Press, 1992.
- [7] P. Hájek and J. Valdés. An analysis of MYCIN-like expert systems. *Mathware and Soft Computing*, 1:45–68, 1994.
- [8] T. Kiseliöva. A theoretical comparison of Disco and CADIAG-II-like systems for medical diagnoses. *Kybernetika*, 42(6):723–748, 2006.
- [9] T. Kiseliöva and K. Pagava. New approach to estimate different drugs and/or other medical interventions effectiveness based on fuzzy logic principles. *Georgia Medical News*, 3(156):65–68, March 2008.
- [10] T. Kiseliöva and K. Pagava. Modelling uncertainty in the connection of socio-economic factors and health-related behavior of adolescents in Georgia. In *Proceedings of the 28th North American Fuzzy Information Processing Society Annual Conference (NAFIPS2009)*, Cincinnati, OH, USA, June 14–17, 2009.
- [11] H. Leitich, K.-P. Adlassnig, and G. Kolarz. Evaluation of two different models of semi-automatic knowledge acquisition for the medical consultant system CADIAG-II/RHEUMA. *Artificial Intelligence in Medicine*, 25:215–225, 2002.
- [12] R. Seising, C. Schuh, and K.-P. Adlassnig. Medical knowledge, fuzzy sets and expert systems. In *Workshop on Intelligent and Adaptive Systems in Medicine*, Prague, <http://cyber.felk.cvut.cz/EUNITE03-BIO>, March 31 – April 1, 2003.
- [13] G. Acampora and V. Loia. Fuzzy control interoperability and scalability for adaptive domotic framework. In *IEEE Transactions on Industrial Informatics*, volume 1.
- [14] Mei-Hui Wang, Chang-Shing Lee, Kuang-Liang Hsieh, Chin-Yuan Hsu, Giovanni Acampora, and Chong-Ching Chang. Ontology-based multi-agents for intelligent healthcare applications. *Journal of Ambient Intelligence and Humanized Computing*, 1(2), 2010. DOI: 10.1007/s12652-010-0011-5.
- [15] D. Acampora, V. Loia, and A. Vitiello. Enhancing Transparent Fuzzy Controllers through Temporal Concepts: An Application to Computer Games. In *2010 International Conference on Technologies and Applications of Artificial Intelligence*, pages 405–412, 2010.
- [16] G. J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic. Theory and Applications*. Prentice Hall PTR, 1995.
- [17] H. Phagava. Numerical estimates of probability expressions to be used in medical expert programs. In *Georgian Medical News*, volume 2, pages 130–134. February 2002.
- [18] E. Sanchez. Medical diagnosis and composite fuzzy relations. In M. M. Gupta, R. K. Ragade, and R. R. Yager, editors, *Advances in Fuzzy Set Theory and Applications*, pages 437–444. North-Hollands, New York, 1979.
- [19] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning, Part I, Part II, Part III. *Information Sciences*, 8,9:199–249, 301–357, 43–80, 1975.
- [20] G. Acampora and V. Loia, "Using fml and fuzzy technology in ambient intelligent environments," *International Journal of Computational Intelligence Research*, vol. 1, no. Issue, pp. 171-182, 2005.
- [21] G. Acampora, V. Loia, "An Open Integrated Environment for Transparent Fuzzy Agents Design," *Open Source Development, Communities and Quality, IFIP International Federation for Information Processing*, Springer Boston, vol. 275/2008, pp. 1571-5736, 2008.
- [22] C. Naylor. How to build an inferencing engine. In R. Forsyth, editor, *Expert Systems, Principles and Case Studies*, chapter 5, pages 85–105. Chapman and Hall, 1989.
- [23] J. W. Grzymala-Busse. *Managing Uncertainty in Expert Systems*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers Boston/Dordrecht/London, University of Kansas, 1991.
- [24] A. Kandel. *Fuzzy expert systems*. CRC Press, Inc, 2000 Corporate Blvd., N.W., Boca Raton, Florida 33431, 1991.