SELECTION OF THE DIAGNOSTIC HYPOTHESES THROUGH FUZZY DECISION IN A MEDICAL DIAGNOSIS SYSTEM

Sabina MUNTEANU

Computer Science and Applied Informatics Department,
"Dunarea de Jos University" of Galati
Domneasca-47, Galati 800008, Romania
Sabina.Munteanu@ugal.ro

Abstract: The paper presents an evaluation of the diagnostic hypotheses' selection step within the medical diagnosis hybrid system DiaMed. Medical diagnoses are described in DiaMed through fuzzy decision functions that are used as discrimination functions for the selection. This type of selection is compared theoretically to a generic fuzzy inference model for medical diagnosis, and is evaluated on a differential diagnosis example, using the C discrimination index of Harrell.

Keywords**:** fuzzy decision, fuzzy inference systems, diagnosis, discrimination index.

1. INTRODUCTION

The diagnosis of complex systems needs a quick focus on relevant directions. This focus corresponds, in general to the hypotheses' selection step and uses intelligent efficient techniques that are less precise and transparent (neural systems, rule-based systems, fuzzy models). This lack of transparency makes it necessary to add an extra level for explanation, especially in the critical fields (as it is the case with medical diagnosis), finally leading to a combinative hybrid structure (for instance, the CHECK system – Torasso and Console, 1989).

The present paper presents an evaluation of the selection step in the original hybrid system DiaMed. The selection is modeled by fuzzy decision (Munteanu, 2005), and it is compared to its closest correspondent, which is the fuzzy inference system. The discriminative power of the fuzzy decision functions is evaluated with the statistic index C (Harrell, Lee and Mark, 1996), and the results are very good.

2. FUZZY INFERENCE SYSTEMS IN MEDICAL DIAGNOSIS

Any medical diagnosis problem can be naturally modeled as a classification with N classes (diagnoses) (Castellano, 2003)):

$$\Delta=\{d_1,\dots,d_N\} \quad (2.1)$$

The set of symptoms used for classification can be written as a K-dimensional vector $o=(o_1,\dots o_K)$, and the diagnosis is done with the classification function:

$$D:A \subseteq \Re^K \rightarrow \Delta, \quad (2.2)$$

where A can be defined by a hyper-interval:

$$A = \prod_{j=1}^{K} dom_j, \quad o_j \in dom_j=[i_j,I_j]$$

This paper was recommended for publication by Severin Bumbaru

The classification function D is based upon a fuzzy rules base that usually contains rules of the following form:

IF o is G_r THEN $\tilde{D}$ (x) is $d_1(v_{r1})$, …, $d_N(v_{rN})$, (2.3)

where G_r is a K-dimensional fuzzy relation.

Measurements vector o's degree of membership to G_r defines the firing degree of rule r (and can be viewed as a measure of the similarity between vector o and the prototype vector of G_r).

The rules above define the classification function:

$$\tilde{D}: A \rightarrow [0,1]^M \quad (2.4)$$

that is, $\tilde{D}$ (o)=(v_1,…,v_N), where each v_j represents the degree of presence of diagnosis d_j at a patient with symptoms vector o. This degree can have different semantics:

- How typical the case o is for diagnosis d_j;
- How serious disease d_j is in case o;
- The support for hypothesis that d_j is the true diagnosis in case o, deduced from the available evidence;
- The probability that d_j is the true diagnosis for o.

The degree of membership of vector o to diagnosis d_j, computed with R rules of type 2.3 can be written as following:

$$v_j = \frac{\sum_{r=1}^{R} G_r(o) v_{rj}}{\sum_{r=1}^{R} G_r(o)} \quad (2.5)$$

(where $G_r(o)$ is the degree of membership of o to the fuzzy relation G_r).
The final decision for diagnosis may choose the class with a maximum index, or it may keep in the list of possible hypotheses all the diagnoses whose degree is greater than a given threshold.

3. FUZZY DECISION IN MEDICAL DIAGNOSIS

An original modeling of diseases is described herein, by fuzzy decision functions, which was used in DiaMed. Let $\Delta = \{d_1, ..., d_N\}$ be a restricted context of disorders taken into consideration, and $M = \{m_1, ..., m_K\}$ a complete set of manifestations characterizing them. It is also considered that Δ is a complete set of causes for *M*. The matrix $WEIGHTS = (w_{ij})_{j=1..K}^{i=1..N}$ contains the relevance indices of symptoms within the disorders' definitions, either statistically computed or defined by an expert.

The main originality of the present approach resides in defining diseases by transforming a fuzzy decision model, as described following:

- Each symptom m_j is defined by a fuzzy function $M_j : dom_j \rightarrow [0,1]$, ($dom_j$ represents the domain of the M_j function, and could be a continuous or discrete range, depending on the symptom), and $A = \prod_{j=1}^{K} dom_j$ is the set of all possible observations (Π is the Cartesian product; observations are considered over the range of all possible symptoms, although in practice only a small part is effectively used, as the rest of them are absent or not observed yet on a given patient);
- Observations are represented by a K-dimensional point $o = (M_1(o), ..., M_K(o))$; this vector can subsequently undergo modifications, as new information is gained by testing;
- Diagnostics are defined by means of the fuzzy weighted decision functions: $D_1^{w_1}, ..., D_N^{w_N}$, as following:

$$D_i^{w_i} = h_i(M_{i_1}(o), ..., M_{in_i}(o)) \quad (3.1)$$, where:

- $Symptoms(d_i) = \{m_{i_1}, ..., m_{in_i}\}$, and $i_1, ...i_{n_i} \in \{1, ..., K\}$ represent the indices of the symptoms relevant for disease d_i;
- w_i is the weight vector of symptoms within definition of disorder d_i;
- h_i is an aggregation function defined by means of fuzzy operators which model a human expert's way of reasoning.

The key to the representation above is to regard symptoms as fuzzy criteria that accomplish a diagnostic, should they be satisfied. The indices $D_i^{w_i}(o)$ of the degree of match between the observations' vector and the complex criterion that defines diagnostic d_i are computed in the end, inducing a hierarchy for the diagnostic hypotheses, based on the particular evidences of a given case. The final diagnoses' set is determined by applying a significance threshold, which should be disorder-dependent (in order to overcome the inconvenience of the variability of the total sum of weights from a disorder to another).

The detailed aspects of the computation are described following. Let "∧" , "∨" represent a T-norm and a S-norm, respectively (Sousa and Kaymak, 2002). Studies within the fields of fuzzy logic and decision making have proved that conjunctive and disjunctive aggregation are just not enough to model decision behavior properly. Many practical situations suggest a compensatory way of thinking: good properties of some criteria compensate for poor properties of others (medical diagnosis represents a very suitable example, as one doesn't need to observe but few symptoms out of a great variety in order to diagnose a certain disease).

An usual approach to the compensatory aggregation is given by fuzzy integrals. These are more appropriate to be used in computing the score for those specific disorders where clear diagnostic criteria do not exist (like those cases where symptoms are divided into more classes, and the final result is given by the conjunctive/ disjunctive aggregation of some terms of the form "at least x from class C_x") .

Definition 3.1 (Sousa and Kaymak, 2002) Let $C=\{c_1,\ldots,c_m\}$ be the set of decision criteria. Let P(C) be the set of the parts of C. A *fuzzy measure* over C is a function $g:P(C)\rightarrow[0,1]$ that satisfies:

1. $g(0)=0, g(C)=1$;
2. $A\subset B\subset C \Rightarrow g(A)\leq g(B)$. (3.2) ■

Definition 3.2. **(Sousa and Kaymak, 2002)** Given a fuzzy measure g on C, the *Sugeno* fuzzy *integral* of a function $p:C\rightarrow[0,1]$ with respect to g is defined as:

$$S_g(p(c_1),\ldots,p(c_m)) = \bigvee_{i=1}^{m}(p(c_{(i)}) \wedge g(A_{(i)})). \quad (3.3)$$

In equation 3.3, the inferior index "$_{(i)}$" indicates that the indices have been permuted such that $0 \leq p(c_{(1)}) \leq \ldots \leq p(c_{(m)}) \leq 1$, and $A_{(i)} = \{c_{(i)},\ldots,c_{(m)}\}$. ■

For those disorders that do not have clear criteria as the one above, let C be the set of their relevant symptoms, $A\subseteq C$ and *g(A)* the normalized weight of the subset A (i.e. the sum of weights of symptoms in A divided by the sum of weights of symptoms in C). $p(m_j)$ is the value of criterion (symptom) m_j at a given patient (system) *p*. This interpretation allows us to give the following definition.

Definition 3.3 (Munteanu, 2005). Let $d_i\in\Delta$, *p* be a patient under observation. The *score of the disease* d_i at a given patient p with associated observation vector *o* is defined as:

$$Score(p, d_i)=S_g(M_{i_1}(o),\ldots,M_{in_i}(o)) \quad (3.4)$$

(where g is defined using the weights w_i of symptoms from M inside disease's d_i definition– weight is 0 if symptom doesn't occur in the definition). The fuzzy measure of a set of symptoms is the normalized sum of their weights, while S_g is the Sugeno fuzzy integral given by 3.2). ■

4. AN EVALUATION OF THE DISCRIMINATIVE POWER OF A DIAGNOSIS TEST

Harrell's C discrimination index is used to evaluate the selection step in DiaMed (Harrell, Lee and Mark, 1996) (this index is the equivalent of the area under the ROC curve -Receiver Operator Characteristic – Bamber, 1075). The index measures the discriminative power of the system, that is, the power to distinguish between cases from different classes, and it will be used in practice for differential diagnosis. If the value of the C index is close to 0.5, then the test is useless, its quality being the better the more the index approaches 1 (1 means perfect discrimination).

Assume that one has access to a "gold standard" (the true result of the classification) – that may be 0 (healthy patient) or 1 (patient suffers of disease D). The system's output is a real number between 0 and 1. A threshold is chosen (for instance, 0.5) and classification is made by comparing test results with this threshold (if score is higher than 0.5, then the patient is considered ill).

The following notations are made for the population used to evaluate the test:

- RP= the number of true positives (correctly classified ill persons, suffering of disease D);
- FN= the number of false negatives (they are ill but the test finds them healthy);
- FP= the number of false positives (they are healthy but the test classifies them as ill);
- RN= the number of true negatives (correctly classified healthy persons).

Definition (Fletcher, 1988). The sensitivity, and respectively the specificity of a diagnosis test are defined by:

- Sensitivity: SN = RP/(RP+FN) (the ability to detect ill persons)
- Specificity: SP = RN/(RN+FP) (the ability to reject healthy persons)

The interpretation of the index (area) is represented by means of a simple example.

	Healthy (gold standard 0)	**Ill (gold standard 1)**
Test results	0.3	0.8
	0.2	0.2
	0.5	0.5
	0.1	0.7
	0.7	0.9

All possible 0-1 pairs are formed afterwards, to compare the system's estimates for each pair:

Healthy	**Ill**	
0.3	0.8	concordant
<		
0.2	0.2	discordant
0.5	0.5	concordant
0.1	0.7	concordant
0.7	0.9	concordant

The pair (0.3, 0.8) is concordant because the test has assigned a higher score to a really ill person than to a really healthy one. One finally gets 18 concordant pairs, 4 discordant and 3 ties, so it can be written:

$$Index\_C = \frac{Concordant + \frac{1}{2}Tie}{Total} = 0.78 \quad (4.1),$$

which means the test was good enough.

5. CASE STUDY

For our example, a pair of disorders is chosen that rises recognition problems (*cirrhosis* and *congestive heart failure*), as there are many cases when congestive heart failure is confused with cirrhosis, manifesting similar symptoms. The C index, obtained by the fuzzy decision classification scheme in Section 3, is compared for different pairs of dual norms (which is used to compute the fuzzy integrals). The calibration index in the last column measures the sensitivity of the computation (it consists of a simple measure of the "sum of differences" type that supports the great majority of the calibration methods).

S-norm	**T-norm**	**C - Index**	**Calibration index**
max	min	0.9856	546.6659
a+b-ab	ab	0.9856	1186.5055
min(a+b, 1)	max(a+b-1, 0)	0.9856	1632.4998

Table 5.1. C-Index and calibration for "**cirrhosis**"

These experiments determined us to choose the pair (max, min) for our system, because it gives a smaller calibration index. Its superiority probably resides in being less sensitive to symptoms' weights, which might not have been quite accurate.

S-norm	**T-norm**	**C - Index**	**Calibration index**
max	min	0.956	392.1201
a+b-ab	ab	0.9088	908.2604
min(a+b, 1)	max(a+b-1, 0)	0.9064	1243.3226

Table 5.2. C-Index and calibration for "**congestive heart failure**"

6. CONCLUSIONS AND FUTURE WORK

The fuzzy decision model used in DiaMed to select diagnostic hypotheses represents a viable, more efficient alternative to the approach above. By its means, each fuzzy rules base corresponding to an output variable *y* can be re-written as a fuzzy decision function. So instead of aggregating the score v_j of a disorder from R rules in Section 2, one can directly compute it by the fuzzy decision function of the disease.

This means that if one has *m* outputs, the selection of hypotheses will only need *m* fuzzy decision functions. Moreover, these functions use more readable, more natural and more complex fuzzy aggregation operators (like fuzzy integrals, compensatory operators etc.), than a basis formed exclusively with the AND operator.

Another great advantage of the fuzzy decision functions described above is given by the fact they are easier to build than fuzzy inference systems, as long as one gets clear and precise enough expert knowledge. In order to achieve that, an appropriate model is needed: the value "small" or "big" for a given element is no longer considered a flaw in itself, but the value of an element is represented as a function of other elements in the system, on which it depends (respecting the structure of a causal net).

In the end, instead of aggregating different degrees of realization for the fuzzy values (*big*, *small* etc.) of one and the same output variable, one straightforward gets a number considered *small* or *big* with respect to a given context (related to the other outputs' values), rather than on an absolute scale.

Therefore, the advantages of the approach are its simplicity, efficiency of the computation (which takes into account the specificities of symptoms for a given disorder), and the quick focalization of the search towards relevant directions. Besides, it can accurately express a great diversity of vague criteria (for instance, *most*, *at least x out of n*, *a significant number out of*... etc.).

The main difficulty encountered in evaluating the selection step is still an open problem: the lack of a rigorously defined index to measure the discriminative power for multiple response-diagnosis. If one must deal with the whole problem (not only a single disorder), the C index should be adjusted accordingly, and the notion of discrimination itself should be restated within the new context.

REFERENCES

Bamber D. "The area above the ordinal dominance graph and the area below the receiver operating graph", *Journal of Mathematical Psychology*, 12, :387-415, 1975.

Castellano G., Fanelli A., Mencar C. "A Fuzzy Clustering Approach for Mining Diagnostic Rules", *Proceedings of 2003 IEEE International Conference on Systems, Man and Cybernetics*, 1, :2007-2012, Washington D.C., USA, 2003.

Fletcher R.H., Fletcher SW., Wagner E.H.*"Clinical Epidemiology: The Essentials"*, 2nd edn. Baltimore: Williams &Wilkins, 1988.

Harrell F.E., Lee K.L., Mark D.B. "Tutorial in Biostatistics, Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors", *Statistics in Medicine*, 15, : 361-387, 1996.

Munteanu, S. "A Hybrid Model for Diagnosing Multiple Disorders", *International Journal of Hybrid Intelligent Systems*, 2, No.1 IOSPress, :35-55, 2005.

Sousa, J., Kaymak, U. *"Fuzzy Decision Making in Modeling and Control",* World Scientific Pub Co, 2002.

P.Torasso, L.Console. "A multilevel architecture for diagnostic problem-solving", *Computational Intelligence*,1 :101-112, 1989.