Computer-Aided Sequential Diagnosis Using Fuzzy Relations – Comparative Analysis of Methods

Marek Kurzynski[1,2] and Andrzej Zolnierek[1,2]

[1] Wroclaw University of Technology, Faculty of Electronics, Chair of Systems and Computer Networks, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland
`{marek.kurzynski,andrzej.zolnierek}@pwr.wroc.pl`
[2] The Witelon University of Applied Sciences,
Ul. Sejmowa 5A, 59-220 Legnica, Poland

Abstract. A specific feature of the explored diagnosis task is the dependence between patient's states at particular instants, which should be taken into account in sequential diagnosis algorithms. In this paper methods for performing sequential diagnosis using fuzzy relation in product of diagnoses set and fuzzified feature space are developed and evaluated. In the proposed method first on the base of learning set fuzzy relation is determined as a solution of appropriate optimization problem and next this relation in the form of matrix of membership grade values is used at successive instants of sequential diagnosis process. Different algorithms of sequential diagnosis which differ with as well the sets of input data as procedure are described. Proposed algorithms were practically applied to the computer-aided recognition of patient's acid-base equilibrium states where as an optimization procedure genetic algorithm was used. Results of comparative experimental analysis of investigated algorithms in respect of classification accuracy are also presented and discussed.

1 Introduction

Multiple diagnoses of the patient's state based on results of successive examinations is one of the most frequent and typical medical diagnosing tasks. Such a task, henceforth called the sequential diagnosis, involves dealing with a complex decision problem. This is caused by the dependence of the patient's state at a given time on the preceding states and on the already applied treatment. Although there remains no doubt as to the very existence of this dependence, it may be of a diversified nature and range; its simplest instance can be a one-instant-backwards dependence to so complex arrangements as those in which the current state depends on the whole former course of the disease.

We have to take into account these sequential diagnosing dependencies when we intend to support diagnosing tasks using a computer. In other words, when constructing an appropriate decision algorithm we must not limit our approach

to only the narrow information channel that concerns just the current symptoms but we have to consider all the available measurement data instead, as they may contain important information about the patient's state at a given instant. The measurement data comprise all the examination results obtained so far, the applied treatment procedures, as well as the diagnoses formulated at previous diagnosis instants. Thus the bulk of data is very rich and grows over time from one instant to another; this can be viewed at as both an advantage and disadvantage, depending on the viewpoint. Every medical practitioner will recognize the great usefulness of such data. Performing the sequential diagnosis he/she will inevitably ask the patient not only about the former symptoms and applied treatment but about the previous diagnostic statements. On the other hand, however, the lavishness and incremental nature of the available data make it impossible to comprehend them completely, as a result of which fact various simplifications and compromises must be made.

A specific feature of the explored diagnosis task is the dependence between patient's states at particular instants, which dependence is, specifically again, taken into account in sequential diagnosis algorithms. The dependence can be included at an as early stage as that of formulating a mathematical model for the diagnosis task, or as late as at the stage of selecting the appropriate input data set in the decision algorithm which otherwise does not differ from the classical diagnosis task. An example for the former case can be the probabilistic approach which offers the effective, as it has turned out, description of the dependences and actual treatment, in the form of a controlled Markov chain ([4], [5]). We call it the effective description because it leads to a constructive algorithm. However, its form is quite complex and depends on the context range that is taken into account in the actual model.

The other case occurred when the approximated inference engine was applied based on a fuzzy rule system, and when artificial neural networks were applied ([1]). Both methods deal with a well-know procedure used either for fuzzy rule construction based on empirical data, or for neural network training. The specificity of the investigated diagnostic task reveals itself here exclusively in the form of input data which are not associated only with the direct symptoms that manifest the current state, but comprise up to an extent the "historic" information that regards the preceding course of the disease. For this case we do not know how far backwards the examined input data should spread into the past; the "the more the better" rule need not necessarily be true here. As far now, there are no analytical evidence to be used in this issue, whilst any attempts to answer the question are under way of experimental research.

In the paper we propose a novel approach to the problem of algorithmization of sequential diagnosis task based on a concept of fuzzy relation given by matrix of membership grade values. The paper is a sequel to the author's earlier publications ([1], [2], [3], [4]) and it yields new results dealing with the application of fuzzy inference systems to the decision making at the successive instants of sequential diagnosis procedure. The contents of the work are as follows. In section 2 necessary background is introduced and the sequential diagnosis method

is described and formalized. In section 3 algorithms for sequential diagnosis are discussed. In the proposed approach fuzzy relation in the product of set of decisions and feature space is determined as a solution of appropriate optimization problem. In the presented example the genetic algorithm was applied to find optimal solution. In section 4 we discuss the results of application of proposed fuzzy decision systems to the computer-aided recognition of acid-base equilibrium states.

2   Preliminaries and the Problem Statement

We will treat the sequential diagnosis task as a discrete dynamical process. The patient (object) is at the n-th instant in the state $j_n \in \mathcal{M}$, where $\mathcal{M}$ is an M-element set of possible states numbered with the successive natural numbers. Thus

$$j_n \in \mathcal{M} = \{1, 2, ..., M\}. \tag{1}$$

Obviously, the notion of instant has no specific temporal meaning here, as its interpretation depends on the character of the case under consideration. The actual measure used may be minutes, hours, days, or even weeks.

The state j_n is unknown and does not undergo our direct observation. What we can only observe is the indirect symptoms (also called features or tokens) by which a state manifests itself. We will denote a d-dimensional symptom value vector by $x_n \in \mathcal{X}$, for symptoms measured at the n-th instant (thus $\mathcal{X}$ is the symptom space); let us also denote by u_n the therapy chosen from the therapy set $\mathcal{U}$ to be applied at the n-th instant.

As already mentioned, the patient's current state depends on the history and thus in the general case the decision algorithm must take into account the whole sequence of the preceding symptom values, $\bar{x}_n = \{x_1, x_2, \ldots, x_n\}$, and the sequence of applied therapies, $\bar{u}_{n-1} = \{u_1, u_2, \ldots, u_{n-1}\}$. It must be underlined here that sometimes it may be difficult to include all the available data, especially for bigger n. In such cases we have to allow various simplifications (e.g. make allowance for only several recent values in the $\bar{x}_n$ and $\bar{u}_{n-1}$ vectors), or compromises (e.g. substituting the whole disease history segment that spreads as far back as the k-th instant, i.e. the $\bar{x}_k$ and $\bar{u}_{k-1}$ values, with data processed in the form of a diagnosis established at that instant, say i_k).

Apart from the data measured for a specific diagnosed patient we need some more general information to take a valid diagnostic decision, namely the *a priori* knowledge concerning the general associations that hold between diagnoses on the one hand, and symptoms and the applied treatment schemata, on the other. This knowledge may have multifarious forms and various origins. From now on we assume that it has the form of a so called training set, which in the investigated decision task consists of m training sequences:

$$\mathcal{S} = \{S_1, S_2, ..., S_m\}. \tag{2}$$

A single sequence:

$$S_k = ((x_{1,k}, u_{1,k}, j_{1,k}), (x_{2,k}, u_{2,k}, j_{2,k}), ..., (x_{N,k}, u_{N,k}, j_{N,k})) \tag{3}$$

denotes a single-patient disease course that comprises N symptom observation instants, the applied treatment, and the patient's state.

Analysis of the sequential diagnosis task implies that, when considered in its most general form, the explored decision (diagnostic) algorithm can in the n-th step make use of the whole available measurement data (perhaps partly substituted with former diagnoses), as well as the knowledge included in the training set. In consequence, the algorithm is of the following form:

$$\Psi_n(\bar{x}_n, \bar{u}_{n-1}, \mathcal{S}) = i_n. \tag{4}$$

The next chapter describes in greater detail the construction of the diagnostic algorithm (4) using concept of fuzzy relation in product of feature space for various input data and decision set.

3   Algorithms of Sequential Diagnosis

In the presented method of sequential diagnosis, first on the base of learning sequences (2) we find fuzzy relation between fuzzified feature space and class number set as a solution of appropriate optimization problem. Next this relation expressed as matrix which elements represent intensity of fuzzy symptoms for each diagnosis (patient's state) can be used to make decision in particular steps of the whole sequential diagnosis process.

Two procedures has been proposed which differ exclusively with the relevant selection of input data. The first algorithm includes k-instant-backwards-dependence ($k < N$) with full measurement data. It means, that decision at the nth instant is made on the base of vector of features

$$\bar{x}_n^{(k)} = (x_{n-k}, x_{n-k+1}, ..., x_{n-1}, x_n). \tag{5}$$

In the second approach however, we also include k-instant- backward- dependence, but using the previous diagnoses in lieu of the previous symptom values. In the both approaches information about the sequence of applied therapies

$$\bar{u}_n^{(k)} = (u_{n-k}, u_{n-k+1}, ..., u_{n-1}). \tag{6}$$

plays the role of peculiar "switch" which allows to select appropriate fuzzy relation (matrix).

Before we will describe both algorithms let first introduce sets $\mathcal{S}^{\bar{u}^{(k)}}$ and $\mathcal{S}^{\bar{u}^{(k)}}_{\bar{j}^{(k)}}$ denoting sequences of $(k+1)$ learning patterns from $\mathcal{S}$ in which at the first k position the sequence of therapies $\bar{u}^{(k)} \in \mathcal{U}^k$ (Cartesian product of $\mathcal{U}$) appears and additionally sequence of classes $\bar{j}^{(k)} \in \mathcal{M}^k$ appears, respectively.

3.1   Algorithm with kth Order Depencence (AkD)

The algorithm with the full measurement features can be expressed in the following points:

1. Cover the space of features $\mathcal{X}$ with fuzzy regions. In the further example we use triangular fuzzy numbers with 3 regular partitions [2]. Obtained fuzzy sets correspond to the linguistic values of features which state fuzzified feature space $\mathcal{X}_F$.
2. For each sequence of therapies $\bar{u}^{(k)} \in \mathcal{U}^k$:
   - Determine observation matrix $O^{\bar{u}^{(k)}}$, i.e. fuzzy relation in the product of Cartesian product of fuzzified feature space $\mathcal{X}_F^k$ and learning subset $\mathcal{S}^{\bar{u}^{(k)}}$. The ith row of observation matrix contains grades of memberships of features $\bar{x}^{(k)}$ of ith learning sequence from $\mathcal{S}^{\bar{u}^{(k)}}$ to the fuzzy sets of space $\mathcal{X}_F^k$.
   - Determine decision matrix $D^{\bar{u}^{(k)}}$, i.e. relation in product of learning sequences $\mathcal{S}^{\bar{u}^{(k)}}$ and the set of decisions (diagnoses) $\mathcal{M}$. Elements of ith row of decision matrix are equal to zero with except the position corresponding to the last class number of ith sequence in the set $\mathcal{S}^{\bar{u}^{(k)}}$ which is equal to one.
   - Find matrix $E^{\bar{u}^{(k)}}$, so as to minimize criterion

$$\rho(O^{\bar{u}^{(k)}} \circ E^{\bar{u}^{(k)}}, D^{\bar{u}^{(k)}}), \tag{7}$$

   where operator $\circ$ denotes max-min-norm composition of relations, i.e. multiplication of matrices O and E with $\times$ and $+$ operators replaced by min and max operators (more general by t-norm and s-norm operators) ([7]). Criterion (7) evaluates difference between matrices A and B, i.e. $\rho(A, B) \geq 0$ and $\rho(A, B) = 0$ iff $A = B$. In the further experimental investigations we adopt

$$\rho(A, B) = \sum_{i,j}(a_{ij} - b_{ij})^2 \tag{8}$$

   and as method of minimization (7) the genetic algorithm will be applied.

Matrix $E^{\bar{u}^{(k)}}$ is a fuzzy relation in product of decision set $\mathcal{M}$ and feature space $\mathcal{X}_F^k$ for objects (patients) subject to therapies $\bar{u}^{(k)}$ in which reflects knowledge contained in the learning sequences.

The manner of utilize the matrices E for decision making is obvious. At the nth step of sequential diagnostic process, first for known sequence (6) we select matrix $E^{\bar{u}_n^{(k)}}$ and for sequence of feature observations (5) the row-matrix of fuzzy observation $O(\bar{x}_n^{(k)})$ is determined. Next we calculate the row-matrix of soft decisions:

$$O(\bar{x}_n^{(k)}) \circ E^{\bar{u}_n^{(k)}} = D(\bar{x}_n^{(k)}, \bar{u}_n^{(k)}) \tag{9}$$

and final diagnosis is made according to the maximum rule.

It must be emphasized that proposed procedure leads to the very flexible sequential recognition algorithm due to optional value of k. In particular the value of k need not be constant but it may dynamically change from step to step. It means next, that choice $k = n - 1$ for nth instant of sequential diagnosis

denotes the utilization of the whole available information according to the general form of diagnostic rule (4). On the other side however, such concept - especially for bigger n - is rather difficult for practical realization.

3.2 Reduced Algorithm with kth Order Depencence and Crisp History (RkDC)

In this approach for diagnosis at the nth instant, we substitute the whole disease history segment which - as previously - covers the k last instances, i.e. $(x_{n-k}, x_{n-k+1}, ..., x_{n-1})$ values with data processed in the form of diagnoses established at these instances, say

$$\bar{i}_n^{(k)} = (i_{n-k}, i_{n-k+1}, ..., i_{n-1}). \tag{10}$$

Such a concept significantly simplifies the computational procedure since - the sequence of previous diagnoses can play exactly the same role as a sequence of previous therapies in the algorithm AkD. Thus, we get identical procedure for determining matrices E in which set $\mathcal{S}^{\bar{u}^{(k)}}$ is replaced with set $\mathcal{S}_{\bar{j}^{(k)}}^{\bar{u}^{(k)}}$ and fuzzified product feature space $\mathcal{X}_F^k$ with simply $\mathcal{X}_F$.

As a consequence for each sequence of possible therapies (6) and diagnoses $\bar{j}^{(k)}$, on the base of learning sequences (2) via optimization procedure (7) the matrix $E_{\bar{j}^{(k)}}^{\bar{u}^{(k)}}$ is determined, which applied in the formula

$$O(x_n) \circ E_{\bar{i}_n^{(k)}}^{\bar{u}_n^{(k)}} = D(x_n, \bar{u}_n^{(k)}, \bar{i}_n^{(k)}). \tag{11}$$

leads to the vector of soft decision at the nth instant and next after maximum defuzzification procedure, to the crisp diagnosis.

3.3 Reduced Algorithm with kth Order Depencence and Soft History (RkDS)

In the RkDC algorithm with crisp history, matrix E for given sequence of therapies $\bar{u}_n^{(k)}$ was univocally determined by observed sequence of previous diagnoses (10). In the concept of algorithm with soft history however, we take into account soft decisions at previous instances, i.e. sequence of diagnoses for previous instances before defuzzification procedure (row-matrices D containing grades of membership for particular diagnoses) instead of sequence (10) of crisp decisions.

Let

$$D_{n-i} = (d_{n-i}^{(1)}, d_{n-i}^{(2)}, ..., d_{n-i}^{(M)}), \tag{12}$$

be the vector of membership grades for all classes produced by diagnostic algorithm at the $(n-i)$th instant $(i = 1, 2, ..., k)$.

In the RkDS algorithm at the nth instant we replace in (11) matrix $E_{\bar{i}_n^{(k)}}^{\bar{u}_n^{(k)}}$ for observed sequence of previous diagnoses (10) with the weighted sum of matrices

for all possible sequences $\bar{i}_n^{(k)} \in \mathcal{M}^k$, viz.

$$O(x_n) \circ \sum_{\bar{i}_n^{(k)} \in \mathcal{M}^k} w_{\bar{i}_n^{(k)}} \times E_{\bar{i}_n^{(k)}}^{\bar{u}_n^{(k)}} = D(x_n, \bar{u}_n^{(k)},), \tag{13}$$

where weight coefficients are equal to product of elements of vectors (12) corresponding to the elements of vector (10), namely

$$w_{\bar{i}_n^{(k)}} = d_{n-k}^{(i_{n-k})} \cdot d_{n-k+1}^{(i_{n-k+1})} \cdot \ldots \cdot d_{n-1}^{(i_{n-1})}. \tag{14}$$

All the decision algorithms that are depicted in this chapter have been experimentally tested as far as the decision quality is concerned. Measure for the decision quality is the frequency of correct diagnoses for real data that are concerned with recognition of human acid-base equilibrium states. The purpose of our research and associated tests was not only the comparative analysis of the presented algorithms but also answering the question whether including the inter-state dependence (whatever its form would be) would yield a better decision quality as compared to algorithms that did not take into account such a dependence. The next chapter describes the performed tests and their outcome.

4   Practical Example: Sequential Diagnosis of Acid-Base Equilibrium States

In the course of many pathological states, there occur anomalies in patient organism as far as both hydrogen ion and carbon dioxide production and elimination are concerned, which leads to disorders in the acid-base equilibrium (ABE). Thus we can distinguish acidosis and alkalosis disorders here. Either of them can be of metabolic or respiratory origin, which leads to the following ABE classification: metabolic acidosis, respiratory acidosis, metabolic alkalosis, respiratory alkalosis, correct state.

In the process of treatment, correct recognition of these anomalies is indispensable, because the maintenance of the acid-base equilibrium, e.g. the pH stability of the fluids is the essential condition for correct organism functioning. Moreover, the correction of acid-base anomalies is indispensable for obtaining the desired treatment effects.

In medical practice, only the gasometric examination results are made use of to establish fast diagnosis, although the symptom set needed for correct ABE estimation is quite large. The utilized results are: the pH of blood, the pressure of carbon dioxide, the current dioxide concentration.

An anomalous acid-base equilibrium has a dynamic character and its changes depend on the previous state and the therapy applied, and in consequence they require frequent examinations in order to estimate the current ABE state. It is clear now that the sequential decision methodology presented above well suits the needs of computer aided ABE diagnosing.

The current formalization of the medical problem leads to the task of the ABE series recognition, in which the classification basis in the n-th moment

constitutes the quality feature consisting of three gasometric examinations and the set of diagnostic results $\mathcal{M}$ is represented by 5 mentioned acid-base equilibrium states. This model can be completed also with therapeutic possibilities (controlling) which patient might undergo. Assuming the certain simplification, these therapies could be divided into three following categories: respiratory treatment, pharmaceutical treatment, no treatment.

In order to study the performance of the proposed concept of sequential diagnosis, the algorithms presented in previous section were applied for the ABE state sequential recognition. Experiments have been worked out on the basis of evidence material that was collected in Neurosurgery Clinic of Medical Academy of Wroclaw and constitutes the set of training sequences (2). The material comprises 78 patients (78 sequences) with ABE disorders caused by intracranial pathological states for whom the following data were regularly put down on the 12-hour basis: 1. gasometric examination results, 2. the correct ABE state diacrisis, and 3. the decision concerning the therapy to be applied. There were around 20 examination cycles for each patient, yielding the total of 1486 single examination instances.

In order to find matrices E for different therapies the genetic algorithm was applied, which is a popular method in optimization and can improve the search procedure ([8]). The genetic algorithm proceeded as follows.

- *Coding method* – the values of elements of matrix E were directly coded to the chromosome.
- *The fitness function* – was defined as follows:

$$Fit = Q - \rho(A, B), \tag{15}$$

  where ρ is as in (8) and Q is suitably selected positive constant.
- *Initialization* – the initial population of chromosomes with which the search begins was generated randomly. The size of population – after trials – was set to 40.
- *Reproduction* – roulette wheel with elitism.
- *Crossover and mutation* – a two-point crossover was used and probability of mutation was 0.05.
- *Stop procedure* – evolution process was terminated after 1000 generations.

The outcome is shown in Table 1. It includes the frequency of correct diagnoses for the investigated algorithms (their names being explained in previous section). Additionally, the algorithm A0D includes neither inter-state-dependences nor the influence the applied therapy has exerted on a patient's

Table 1. Frequency of correct diagnosis for various diagnostic algorithms

Algorithm	A0D	A1D	A2D	R1DC	R2DC	R1DS	R2DS
Result	80.3%	88.4%	91.3%	85.9%	87.7%	85.6%	88.2%

state but utilises only the current symptom values instead. Thus it will be obtained by putting $k = 0$ in the algorithm AkD.

These results imply the following conclusions:

1. Algorithm A0D that does not include the inter-state dependences and treats the sequence of states as independent objects is worse than those that were purposefully designed for the sequential medical diagnosis task, even for the least effective selection of input data. This confirms the effectiveness and usefulness of the conceptions and algorithm construction principles presented above for the needs of sequential diagnosis.
2. There occurs a common effect within each algorithm group: the model of the second order dependency (A2D, R2DC, R2DS) turns out to be more effective than the first order dependence approach (A1D, R1DC, R1DS).
3. Algorithms A1D and A2D that utilize the original data (i.e. symptoms along with therapy) yield always better results than those which substitute the data with diagnoses.
4. There is no essential difference between the algorithms with crisp and soft history.

5   Conclusions

This work presents fuzzy approach to the algorithmization of the sequential diagnosis task and the implied decision algorithms which take into account a specific feature of the sequential diagnosis, i.e. dependence between patient's states at particular instants. The fuzzy methods which have been already applied to the computer-aided sequential diagnosis, use fuzzy rules generated from the learning set and Mamdani inference engine ([1]). Recognition method presented in this work however, deals with fuzzy relation concept developed to the sequential diagnosis process.

The comparative analysis presented above for the sequential diagnosis algorithms is of the experimental nature. We have carried out a series of experiments on the basis of a specific exemplar that concerns acid-base unequilibrium state diagnosing using rich enough set of real-life data. The objective of our experiments was to measure quality of the tested algorithms that was defined by the frequency of correct decisions. The algorithm-ranking outcome cannot be treated as one having the ultimate character as a that of a law in force, but it has been achieved for specific data within a specific diagnostic task. However, although the outcome may be different for other tasks, the presented research may nevertheless suggest some perspectives for practical applications. All the experiments show that algorithms which are appropriate for sequential diagnosis, i.e. ones that include - in whatever form - the inter-state dependences are much more effective as far as the correct decision frequency is concerned than algorithms which do not include the actual associations. This testifies that the proposed conceptions are correct, and the constructed algorithms effective, for computer aided sequential medical diagnosis.

References

1. Kurzynski, M.W.: Benchmark of Approaches to Sequential Diagnosis. In: Lisboa, P., Ifeachor, J., Szczepaniak P.S. (eds.): Perspectives in Neural Computing. Springer-Verlag, Berlin Heidelberg New York (1998) 129-140
2. Kurzynski, M.W.: Multistage Diagnosis of Myocardial Infraction Using a Fuzzy Relation. In: Rutkowski, L., Tadeusiewicz, R. (eds): Artificial Intelligence in Soft Computing, Lecture Notes in Artificial Intelligence, Vol. 3070, Springer-Verlag, Berlin Heidelberg New York (2004) 1014-1019
3. Kurzynski, M.W., Zolnierek, A.: A Recursive Classifying Decision Rule for Second-Order Markov Chains, Control and Cybernetics **18** (1990) 141-147
4. Zolnierek, A.: The Empirical Study of the Naive Bayes Classifier in the Case of Markov Chain Recognition Task. In: Kurzynski, M.W., Wozniak M. (eds.): Computer Recognition Systems CORES 05, Springer-Verlag, Berlin Heidelberg New York (2005) 329-336
5. Devroye, L., Gyorfi, P., Lugossi, G.: A Probabilistic Theory of Pattern Recognition. Springer-Verlag, Berlin Heidelberg New York (1996)
6. Duda, R., Hart, P., Stork, D.: Pattern Classification. John Wiley and Sons, New York (2001)
7. Czogala, E., Leski, J.: Fuzzy and Neuro-Fuzzy Intelligent Systems, Springer-Verlag, Berlin Heidelberg New York (2000)
8. Goldberg, D.: Genetic Algorithms in Search, Optimization and Machine Learning. Adison-Wesley, New York (1989)