

# A New Method for Solving Hard Diagnosis Problems

Ghazi Shakah

Faculty of Information Technology, Ajloun National University, PO box 43, Ajloun-26810, Jordan

E-mail: [g.shakah@anu.edu.jo](mailto:g.shakah@anu.edu.jo)

## Abstract

The paper highlights an approach to solving problems of medical diagnosis. The problems are formulated in terms of pattern recognition theory. An original algorithm is proposed. The combined information in the form of objects and rules is used for its training. Indirect justification of the algorithm with the use of the resolution method is presented.

**Keywords:** pattern recognition, algorithms, resolution method, Diagnosis problems, decision support systems.

**DOI:** 10.7176/CEIS/10-1-03

## 1. Introduction

Diagnosis problems often occur when describing different AI systems. Actually, they are the main body of systems oriented to management decision support. Peculiarities, functional possibilities and technical characteristics of such systems greatly depend on the way of problem solution, i.e when developing similar systems inference is one of the main problems. Existing inference algorithms have some drawbacks. For example, usual logical inference (resolution), in spite of the possibility of its justification in a number of cases, is too categorical [1, 2]. This is due to the fact that in the majority of cases one has to deal with problems that are not fully defined. Other methods (e.g. a statistical one) lack rigorous justification altogether and require rather specific and most often inaccessible information [5-7].

The study suggests an ingenious technique means of deduction. The essence of justification is rather simple and quite evident, i.e. the used mechanism of deduction should approximate the resolution method in verity domain of the latter one and dominate in the domain where the resolution does not allow making a decision. The paper illustrates mathematical aspects of problem [9-15].

## 2. Analysis and Problem Statement

The majority of diagnosis problems in there essences have the following statement:

In a certain set of objects  $X$  of arbitrary nature the following data is specified:

- finite number a of subsets (classes)  $X_1, \dots, X_l$  ( $l \in \mathbf{N}$ ).
- Initial information about objects  $x \in X$  possessing the property  $P_i(x) = "x \in X_i"$  for all  $i = 1, \dots, l$ .

It is necessary to build an algorithm  $A$  (defined on the whole set  $X$ ) that calculates the value of

Properties  $P_1(x), \dots, P_l(x)$  for each  $x \in X$ .

It is clear that each particular problem mainly depends on peculiarities of the subject area and requirements that are imposed upon algorithm  $A$ . We will consider below diagnosis problems which comply with the following requirements

1. The number of classes is finite. Information about classes can be encoded with the help of the system of binary characters  $\mathbf{B}_2^n$  ( $\mathbf{B}_2 = \{0,1\}$ ) and it is possible to indicate subspaces of independent and incoherent features  $\mathbf{B}_2^{n_1}, \mathbf{B}_2^{n_2}, \dots, \mathbf{B}_2^{n_k}$  ( $n_1 + \dots + n_k = n$ ), where  $k \in \mathbf{N}$  is determined by a

particular of classes. Each class  $X_i$  in the subspace  $\mathbf{B}_2^{n_j}$  is associated with a set of rules  $S_{ji} (j = 1, \dots, k; i = 1, \dots, l)$  where their verity domain determines the initial information about the class. Verity domains

$$X_i^0 = \{x \in \mathbf{B}_2^n : P_i(x) = S_{1i}(x) \wedge \dots \wedge S_{ki}(x) = \mathbf{truth}\}$$

have the property:

$$\bigcap_{i=1}^l X_i^0 \subset \mathbf{B}_2^n \quad (1)$$

When forming the rules, as it turned out, we may restrict ourselves to statement calculus, i.e. Boolean algebra without quantifiers [1,17,21].

Nothing is known about the subset  $X_i \cap X_j$  and classes, in general, may intersect.

The sought for algorithm  $A$  should be defined on the whole set  $\mathbf{B}_2^n$  so that results of its operation can be treated as realization of properties  $(P_1, \dots, P_l)$ . And the algorithm should have certain monotony, i.e. it should calculate values of properties on partial and redundant descriptions.

Using results of the carried out analysis, the above-mentioned problem statement was revised and conditions concerning algorithm  $A$  were defined. As for problem statement, actually, a space of the formation of objects  $X \subseteq \mathbf{B}_2^n$  and a way of specifying initial information by rules were revised. Other features of the problem statement remained the same. Conditions for choosing an algorithm can be described in the following manner. Each algorithm  $A$  is a representation

$$A: \mathbf{B}_2^n \rightarrow \{0,1\}^l \quad (2)$$

and if we denote  $A(x) = (\mu_A^1(x), \dots, \mu_A^l(x))$ , where  $\mu_A^i(x)$  is a value of property  $P_i$  for an object  $x \in X$  (that can be called a degree of property confirmation, because  $\mu_A^i(x) \in [0,1]$ ), then the condition of monotony of algorithm  $A$  can be written as

$$\forall x \in \bigcap_{j=1}^l X_j^0 \quad \forall i = 1, \dots, l, \quad \left\{ \begin{array}{l} \mu_A^i(\bar{x}) \leq \mu_A^i(x), \end{array} \right. \quad (3)$$

Where  $\bar{x}$  - are objects in  $X$  that are obtained from  $x$  by eliminating and adding meaningful features (i.e. that are equal to 1).

Thus, we obtain a problem of decision-making (diagnostics and treatment) that is solved by algorithms of type (2) with the restriction (3) and condition that objects can be described in space  $\mathbf{B}_2^n$  and the initial information is specified by rules in the language of statement calculus in such a way that (1) is satisfied.

### 3. Problem Solution

Notice that any algorithm solving the problem stated above can be realized and used only when it withstands a certain testing, approbation. Methodology of such testing is well known and greatly depends on aims of development [2- 4,20]. As a practical matter it means that solution should be found in a set of algorithms

possessing parameters (in the model), because in this case the choice of algorithms becomes considerably easy. Besides, when using parameters we may expect to obtain some additional "good" properties of algorithms, e.g. convergence, etc. It is precisely this reasoning that determined the approach to problem solution described below.

### 3.1. Canonical Algorithm

One can discuss for a long time that algorithms solving the stated problem are inductive both by the nature of information and principles of their construction [3, 6]. But, it appears that the problem has at least two very simple (deductive by techniques) solutions: with the help of resolution methods and exhaustion. In the latter case it is necessary first to transform the initial information from representation in the form of rules to "object"

representation in the space  $\mathbf{B}_2^n$  or  $\mathbf{B}_2^{n_1} \times \dots \times \mathbf{B}_2^{n_k}$ . To carry out this transformation we may use an algorithm for building a disjunctive normal form (DNF). Equivalence between the initial and obtained representations is quite evident [5,7,19]. We will not extensively discuss details of such algorithms because they can be found practically in any manual on artificial intelligence or logic's. Results of their operation are equivalent and can be presented in the following manner using designations mentioned above

$$\mu_A^i(x) = \begin{cases} 1, & x \in X_i^0 \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

It is easy to notice that these algorithms  $A$  solve the problem stated above. At the same time  $\mu_A^i(x) \in \{0,1\}$ . The algorithms also satisfy the condition (3).

For us such algorithms will be source (canonical) ones and that is why we denote them by a special symbol  $A_0$ . The solution will be sought in a set of algorithms that are as good as  $A_0$  from the viewpoint of the following definition

We call algorithm  $A_0$  dominating  $A$  (let us denote  $A \phi A_0$ ) if

$$\forall x \in X \quad \forall i \in \{1, \dots, l\} \quad (\mu_A^i(x) \geq \mu_{A_0}^i(x)).$$

It is easy to see that any algorithm dominating  $A_0$  also solves the problem under consideration, but is less categorical outside the sample  $\cup X_i^0$ . Besides, since legitimacy of using  $A_0$  for the problem solution is beyond question, then proof of domination also lets us make a similar conclusion about algorithm  $A$ . Although we realize that such conditions are necessary but insufficient for categorical justification. But, apparently, any categorical justification is possible not earlier then the inductive conclusion as a whole is justified [6,17,22].

### 3.2 Dominating Algorithm

Below we will describe one set of algorithms  $A$  and show that any of them dominates  $A_0$ . In the following we will consider some properties of the constructed algorithms. Where it is possible, without sacrifice to understanding, we will leave evident details and restrict our consideration to references.

Let's start from the general scheme of algorithm  $A$ . For this purpose we will introduce additional designations and assumptions. Let's consider that:

1. when dividing  $\mathbf{B}_2^n$  into subsets  $\mathbf{B}_2^{n_1}, \dots, \mathbf{B}_2^{n_k}$  each of them has features  $I_j \subset \{1, \dots, n\}$  and  $|I_j| = n_j$ . Without loss of generality we assume that the set  $\{1, \dots, n\}$  is numbered in the following manner:  $n_1$  features from  $\mathbf{B}_2^{n_1}$  (i.e.  $I_1 = \{1, \dots, n_1\}$ ) come first, followed by  $n_2$  features from  $\mathbf{B}_2^{n_2}$  (i.e.  $I_2 = \{n_1 + 1, \dots, n_1 + n_2\}$ ), etc.;
2. sets  $\{1, \dots, l\} \times I_j$  are associated with:

– vectors  $(a_{im+1}, \dots, a_{im+n_j}) \in \mathbf{R}^{n_j}$  (where  $m = \sum_{u=1}^{j-1} n_u$ ) characterized by  $a_{iu} \geq 0$  for all

suitable  $i, u$ ;

– sets of objects  $X_{ij}^0$  that exhaust verity domain of rules from  $S_{ij}$  and that are obtained by transformation of the rules to DNF. It is evident that between conjuncts of such DNF and objects from  $X_{ij}^0$  there is one-to-one onto function [6].

Now  $A$  can be described as a sequence of the following steps.

**Algorithm:**

**Step 1.** For the specified object  $x \in \mathbf{B}_2^n$  we perform

Step 1.1 We fix the number  $u \in \{1, \dots, k\}$  and pass to the next step.

Step 1.2 For each  $i \in \{1, \dots, l\}$  and for all  $x' \in X_{ui}^0$  we calculate

$$\mu_A^{i,u}(x, x') = \max \left\{ 0, \left( \sum_{v \in I_u} (-1)^t \cdot a_{iv} \right) \cdot \left( \sum_{v \in I_u} a_{iv} \right)^{-1} \right\}$$

where

Thus,

$$\left( \sum_{v \in I_u} (-1)^t \cdot a_{iv} \right) \cdot \left( \sum_{v \in I_u} a_{iv} \right)^{-1} = \begin{cases} 1, & \text{if } t = 1 \\ -1, & \text{if } t = 2 \end{cases} \quad t = \begin{cases} 1, & \text{if } x_v \neq x'_v \\ 2, & \text{otherwise.} \end{cases}$$

Step 1.3 If all  $x' \in X_{ui}^0$  are exhausted, we calculate

$$\mu_A^{i,u}(x) = \max_{x' \in X_{ui}^0} \{ \mu_A^{i,u}(x, x') \}$$

Step 1.4 If all numbers  $i$  are not exhausted, we return to step 1.2. Otherwise, we pass to the next step.

Step 1.5 If all  $u$  are not exhausted, we return to step 1.1. Otherwise, we pass to step 2.

**Step 2.** For each  $i \in \{1, \dots, l\}$  we calculate

$$\mu_A^i = \left( \sum_{u=1}^k \gamma_u \cdot \mu_A^{i,u}(x) \right) \cdot \left( \sum_{u=1}^k \left( \gamma_u \cdot \sum_{v \in I_u} a_{iv} \right) \right)^{-1}$$

where  $\gamma_u \in \mathbf{R}$  ( $\gamma_u \geq 0$ ) is a measure of "importance" of the subset of features  $I_u$ .

**Step 3.** If a set of objects  $x$  is not exhausted, we return to step 1. Otherwise, the algorithm finishes its work.

It is easy to see that a particular algorithm  $A$ , in accordance with this scheme, is associated with the choice of

parameters  $a_{iv}$  and  $\gamma_u$ . In reference to the choice we may say that it is mainly determined by reasons over and above the algorithm: the choice of description space of objects from  $X$ , the desire to assign some meaningful interpretation to numbers  $a_{iv}$  and  $\gamma_u$ , means of formalization of the notion of object likeness (step 1.2), etc. Irrespective of the choice (within permissible limits) it is possible to prove that the following main assertion takes place.

Assertion [8,16]. Each algorithm  $A$  solves the problem and  $A \phi A_0$ .

### 3.3 Determining Parameters and some Properties Of $A$

As it was stated above, algorithm and values calculated by it greatly depend on the choice of parameters. It is clear, that this choice should be such that the results of the operation of the algorithm we may interpret in terms of the problem, analyze the initial information and better understand the problem. Moreover, from the scheme of constructing it is evident that in the set of algorithms under consideration the question of complexity of realization is of importance because, in accordance with the methodology, is based on the complete exhaustion of objects from . We will show below that in the chosen way of determining parameters instead of we may restrict ourselves to objects from not only in the scheme of constructing .

Let's first describe an algorithm of parameters determination. In so doing we consider that samples for all are constructed by reduction of the corresponding set of rules to DNF.

#### Algorithm.

**Step 1.** For the specified  $I_u$  and sets of objects  $X_{ui}^0$  we perform the following sequence of steps.

Step 1.1 We fix the number of the feature  $v \in I_u$  and for each  $i \in \{1, \dots, l\}$  we calculate

$$b_{iv} = \left( \sum_{x_i \in X_{ui}^0} x_{iv} \right) \cdot \left( |X_{ui}^0| \right)^{-1}$$

Step 1.2 If all features are not exhausted, we return to step 1.1. Otherwise, we pass to step2.

**Step 2.** For all  $v \in I_u$ ,  $i \in \{1, \dots, l\}$  we calculate

$$b_v = \left( \sum_{i=1}^l b_{iv} \right) \cdot l^{-1}, \quad a_{iv} = |b_{iv} - b_v|,$$

$$\gamma_u = \min_{i \in \{1, \dots, l\}} \max_{v \in I_u} \{a_{iv}\}.$$

**Step 3.** End of work algorithm.

Now let's consider some properties of an algorithm for parameters construction and relation of the parameters with properties of dominating algorithms . We restrict our attention only to those algorithms that relate to the question of applicability. Let's first check that the formed parameters can be used. But it is evident because the following property can be observed.

#### Property 1:

The formed parameters are permissible, i.e.  $a_{iv} \geq 0$  and  $\gamma_u \geq 0$  for

all  $i \in \{1, \dots, l\}$ ,  $u \in \{1, \dots, k\}$ ,  $v \in I_u$ .

Now let's discuss the following aspect. Obviously, when constructing  $X_i^0$  by reducing formula

$S_{li}(x) \wedge \dots \wedge S_{ki}(x)$  in (4) to DNF we will obtain the same set of objects of space  $\mathbf{B}_2^n$  as compared with direct summation  $X_{li}^0 \oplus \dots \oplus X_{ki}^0$  of samples  $X_{ui}^0$  of the corresponding DNF of rule  $S_{ui}$ , i.e. sample  $X_i^0$  differs in the set of features  $I_u$  in the number of similar members of the sample  $X_{ui}^0$ . The number of such members can be easily calculated and it is rather large. But there arises a question: is it possible when constructing parameters to restrict ourselves to the sample  $X_{ui}^0$  or it is necessary to take into consideration the suitable part of the sample  $X_i^0$ ? In the latter case this would lead to exponential growth of the number of operations, memory capacity and impossibility to use algorithm  $A$ . But, it appears that we can answer in the affirmative concerning the first part of the formulated question because the following property can be observed.

**Property 2:**

Parameters  $a_{iv}$  and  $\gamma_u$  for samples  $\bigcup_{t=1}^{t_k} X_{uk}^0$  ( $k = 1, \dots, l$ ) do not depend on  $t_k \in \mathbf{N}$ .

Now let's consider some properties of parameters  $a_{iv}$  the value of which depends on different nature of rules  $S_{ui}$  and, consequently, of samples  $X_{ui}^0$ . Intuitively it is clear that the greater the difference among the rules the greater is the difference in parameters  $a_{iv}$  and the greater is the relative weight of the set of features  $I_u$ . We will assign a formal meaning to this property by proving the validity of its negation.

**Property 3:**

If  $X_{ul}^0 = \dots = X_{ul}^0$  for some  $u \in \{1, \dots, k\}$ , then  $\gamma_u = 0$  and  $a_{iv} = 0$  for all  $i \in \{1, \dots, l\}$ ,  $v \in I_u$ .

It is easy to see that the corresponding features  $I_u$  do not in any way influence the formation of the assessment  $\mu_A^i(x)$ . And this, on the whole, agrees with intuitive knowledge of the operation of  $A$ . But it should also alert because when calculating  $\mu_A^i(x)$  the sum of such values is in the denominator. But it turns out that equality to zero of parameters  $a_{iv}$  by the condition  $X_{ul}^0 = \dots = X_{ul}^0$  is the only one because the following property can be observed.

**Property 4:**

From  $a_{iv} = 0$  for all  $v \in I_u$ ,  $X_{ul}^0 = \dots = X_{ul}^0$  follows.

**4. Conclusion**

The study did not aim at fully describing and studying the introduced set of dominating algorithms. We focused on the following reasoning. If the behavior of the algorithm is substantiated by the proof of some understandable properties, such algorithms have the right to exist. Eventually, practical application is determined by results of experiments, to have only proofs are obviously insufficient.

**References**

Davis R, McGraw-Hill: (advanced computer science series 2016): Knowledge Based Systems in Artificial Intelligence. - New York. p-132-147.

Thayse A., Gribomont P: (2001): Approche logique de l'intelligence artificielle. 1 De la logique classique a la programmation logique. Bordas, Paris.

- Lauriere J.-L. Intelligence Artificielle. Resolution de Problemes parl’Homme et la Machine. Eyrolles, Paris.
- Wright G.H. (2005):A Treatise on Induction and Probability. Brace and World, New York.p.21-27
- Hintikka J. Towards a Theory of Inductive Generalization. Reading in the philosophy of science, New York.
- Popper K. (2002) :The Logic of Scientific Discovery. London.p.29-33.
- Obraztsov V. Krasnoproshin V. Fuzzy Algorithms in Decision-Making and Management Problems. Proc. Of the 2nd Congress SIGEF, S. De Compostela, , v.2, pp 68-79.
- Krasnoproshin, V, Obraztsov V.( 1996):The Problem of Algorithms Choosing in Pattern Recognition, Pattern Recognition and Image Analysis. Vol 6. no 2. -p.p.188-199.
- Krasnoproshin V., Obraztsov V., Vissia H.(, 2000):Knowledge as an Object of Mathematical Formalization, Special Issue of International Association AMSE, Leon, France–p.p.27-43.
- Obraztsov V., Vissia H Krasnoproshin V.,( 2010). Solution of applied problems: formalization, methodology and justification, “Computational Intelligence in Business and Economics”, - London: World Scientific, , Vol. 3 – pp 57-64.
- Vissia H. Krasnoproshin V., Obraztsov V., (2011).Overall Approach to the Solution of Applied Problems, Special Issue of International Association AMSE, Leon, France, – p.p. 166-176
- Obraztsov V.( 2012).Problems of Choice and Pattern Recognition within Binary Spaces, Journal of Computational Optimization in Economics and Finance.. vol. 4 № 2/3, – p.p.34-45
- Rjazanov.V, Krasnoproshin V., Obraztsov V.,, Vissia H. (2016).Artificial Intelligence in Knowledge–Based Technologies and Systems, Computer Science and Information Technologies. Volume 4, #1, – p.p. 27 – 32, DOI: 10.13189/csin.. 040105
- Krasnoproshin V., Obraztsov V., Rjazanov V., Vissia H.( 2017):Decision-Making in Sports Traumatology, Sports Management as an Emerging Economic Activity Trends and Best Practices, Springer. – ISBN 978-3-319-63907-9. – P. 207-219.
- Munehiro Nakamura, Atsushi Otsuka.(2014):Haruhiko Kimura, "Automatic selection of classification algorithms for non-experts using meta-features", China-USA Business Review, vol. 13, no. 3, pp. 199-205.
- Shakah G. Valvachev A.N. (2008):Active System Management Under Uncertainty, International Journal of Computing. Vol. 7, Issue 3, 95-98,
- Jain A., Duin R., Man J. (2000):Statistical Pattern Recognition: A Review. IEEE Transactions on Pattern Analysis and Machine Learning, Vol. 22 4-37.

Kotthoff Lars, Ian P. Gent, Ian Miguel.(2012):"An evaluation of machine learning in algorithm selection for search problems", *AI Communications*, vol. 25, no. 3, pp. 257-270,

Adomavicius, G, Tuzhilin, A. (2005): Personalization technologies: a process-oriented perspective. *Commun. ACM* 48(10), 83–90.

Chris A. Lawson, Charles W. Kalish.(2009):Volume 37, Issue 5, pp 596–607| Cite as Sample selection and inductive generalization

Jiwen Lu ; Venice Erin Liong ; Xiuzhuang Zhou ; Jie Zhou.(2015):Volume: 37 Issue: 10 Learning Compact Binary Face Descriptor for Face Recognition.p.2041-2056.

Nitin Pise, Parag Kulkarni. (2016):Algorithm selection for classification problems: SAI Computing Conference (SAI), IEEE.p.203-211.