

AN EXPERT SYSTEM IN MEDICAL DIAGNOSIS

R. Raboanary*, R. Andriambololona*, J. Soffer*, J. Raboanary*

* Madagascar –Institut des Sciences et Techniques Nucléaires (Madagascar-ISTN)
P.O Box 4279 101 Antananarivo MADAGASCAR

• CNRS Luminy Case 907 Centre de Physique Théorique F-13288 Marseille France.

° Institut Supérieur Polytechnique de Madagascar (ISPM) Antsobolo Antananarivo
MADAGASCAR

Abstract

AN EXPERT SYSTEM IN MEDICAL DIAGNOSIS.

Health problem is still a crucial one in some countries. It is so important that it becomes a major handicap in economic and social development. In order to solve this problem, we have conceived an expert system that we called MITSABO, which means TO HEAL, to help the physicians to diagnose tropical diseases. It is clear that by extending the data base and the knowledge base, we can extend the application of the software to more general areas.

In our expert system, we used the concept of "self organization" of neural networks based on the determination of the eigenvalues and the eigenvectors associated to the correlation matrix XX^t . The projection of the data on the two first eigenvectors gives a classification of the diseases which is used to get a first approach in the diagnosis of the patient. This diagnosis is improved by using an expert system which is built from the knowledge base.

Key words: Expert system, neural networks, medical diagnosis, tropical diseases, data base, knowledge base, self organization.

1 Introduction

In order to help physicians to make a better and faster diagnosis of diseases in particular tropical diseases, we have conceived an expert system that we called MITSABO, that means TO HEAL. An expert system, by definition, works like a human expert. Our work is divided in three parts. In the first one, we show the different steps of a physician's diagnosis of a patient. The second part is related to the architecture of the expert system, including the application of the method of data analysis. The results thus obtained are made more precise by the application of the inference engine which itself uses the knowledge base system. In the third part we give the clinical conclusion.

2 The diagnosis process of a patient

When a human expert does the diagnosis of a patient, three steps are performed:

1. Information about the patient:
 - Name, age, address...
 - The reason of his visit to the physician
 - The family health history
 - The evolution of the disease.

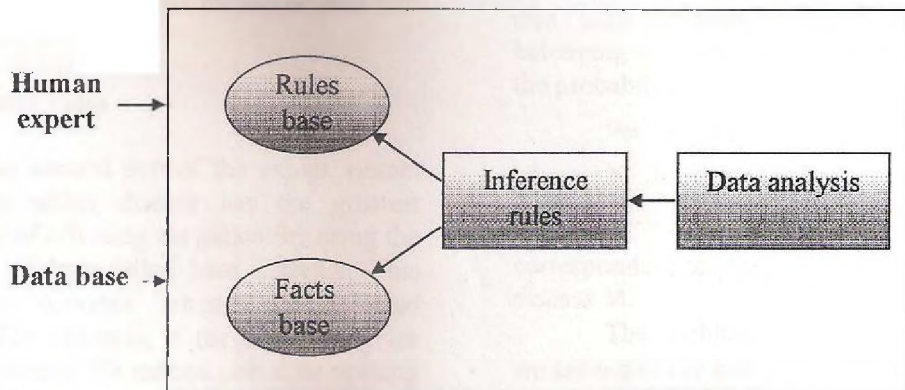
After this first step, the physician does a clinical exam I:

- general signs (temperature, pulse,...)
 - functional signs.
2. - clinical exam II: physical signs
 - biological exam.
 3. - clinical conclusion
 - discussion of the diagnosis.

3 Architecture of the expert system MITSABO

The figure 1 gives the general scheme of the expert system MITSABO, essentially composed of the kernel in relation with the external environment.

Figure 1 Synoptical scheme of the expert system MITSABO



In fact, our expert system follows the logic given in the section 2. The kernel of the the expert system is composed first of the data analysis which corresponds to the first part of the diagnosis, while the second one corresponds to the application of the inference rules. The third part is concerned with the clinical conclusion.

3.1 Data analysis

The data analysis does the selection of the disease among the known possible tropical diseases. At the end of this first part, one, two or three diseases will be retained.

Eighteen tropical diseases are collected (paludism, leprosy, plague,..), and we have considered 23 clinical signs (fever, headache, perspiration,..), but the number of diseases as well as clinical signs may be augmented. [1], [2], [3], [4].

Following the mathematical techniques used in data analysis, we construct the Table 1, where in the first line we put 35 elements, composed of the clinical signs, followed by continued observations such as temperature, pulse... In the first column, we have the diseases M_i and the names of the patients. The inputs of the table are 0 and 1, where 0 means the absence of clinical sign, whereas 1 means its existence.

Let X be the matrix obtained from this table. We calculate $R=XX^t$. Since R is a definite positive matrix, we obtain, by using Jacobi algorithm [5], two eigenvectors \vec{u}_1, \vec{u}_2

and their corresponding positive eigenvalues λ_1 and λ_2 , which are the two largest ones.

Each disease and each patient are vectors with 35 components.

We project the diseases and the patients in the plane (\vec{u}_1, \vec{u}_2) . (Figure 2).

We obtain N classes of points surrounding the diseases M_i which are approximately the centres of each class. In our case, $N=35$. The question is now to decide in which class one patient may belong.

For this purpose, we introduce the fuzzy logic.

In our analysis, each class is a fuzzy set.

A fuzzy set A is defined in the following way:

$A=\{(x, \mu(x)), x \in X\}$, where X is the universe and $\mu(x)$, called the grade of membership of x to the fuzzy set A , is a real number between 0 and 1.

Now let:

$$x_i \in X, i=1,..,n, \text{ with } |X|=n, \\ \Gamma_i(\theta)=\{x \in X, d(x,x_i) < \theta\} \text{ with } \theta > 0,$$

we define:

$$\mu_A(x_i) = \frac{|\Gamma_i(\theta)|}{n}$$

Let P be the point representing a patient whose disease is to be analysed in the plane (\vec{u}_1, \vec{u}_2) , and suppose that M_i, M_j, M_k , are the centres of classes which are in the immediate neighbourhood of P . P may then contract one of the diseases M_i , or M_j , or M_k with the respective grades of membership $\mu_{M_i}(P), \mu_{M_j}(P)$ and $\mu_{M_k}(P)$. Here, $\mu_{M_i}(P)$,

$\mu_{M_j}(P)$, $\mu_{M_k}(P)$ are respectively the grade of membership of P to the class which has respectively M_i , M_j and M_k as centres. These three diseases are retained for the second part of the diagnostic.

3.2 Inference rules

The second part of the expert system determines which disease has the greatest probability of affecting the patient by using the technique of knowledge base systems. This disease is chosen among the selected diseases. For example, if the first part gives only one disease, the second part does nothing but confirming this choice. The restriction to two or three diseases limits the field of the diagnostic and allows more precise results.

The method is as follows:

The interaction between the physician and the patient is represented by the tree of the figure 3. The root is the starting point of this interactive conversation. Each one of the nodes represents a given rule describing the partial sign of the disease (PSD). A rule is composed of an hypothesis and a conclusion (if...then...). The hypothesis may be the combination of many conditions connected by the logical connectives AND or OR.

At the first question, the physician, from the answer of the patient, must choose a rule s_i^1 , $i=1, \dots, n_1$, where n_1 is the number of nodes at the first level of the tree. Let α_i^1 be the probability of conclusion of the rule s_i^1 .

More generally, if we are at level $k-1$, then we have to choose the rule PSD s_k^i at level k , let α_k^i be the corresponding probability of the conclusion of the rule. This latter is computed by using the function MAX whenever the logical connective OR appears, and the function MIN is used for the logical connective AND.

For example, let be given the rule:

if cond₁ AND (cond₂ OR cond₃) then conclusion.

If c_1 , c_2 , and c_3 are the probability associated to cond₁, cond₂, and cond₃ then $c = \text{MIN}(c_1, \text{MAX}(c_2, c_3))$.

The probability of the conclusion of the rule is $\alpha_k^i = c \times \delta$, where δ is the probability of the rule. The conclusion of a rule will be one of the condition (hypothesis) of the rules

that belong to the nodes below and which are their sons.

In order to determine the disease of the patient, we start from the root and go down the tree until we reach one particular disease M belonging to the leaves. We may then compute the probability

$$\alpha_M = \alpha_{i_1}^1 \alpha_{i_2}^2 \dots \alpha_{i_N}^N,$$

where N is the depth of the tree and $\alpha_{i_1}^1, \alpha_{i_2}^2, \dots, \alpha_{i_N}^N$ are the labels corresponding to the path from the root to the disease M.

The problem of this approach is that we are not really sure that the probability α_M is the highest possible value. Indeed, another path from the root to a disease M' may give a better value $\alpha_{M'} > \alpha_M$. Then, we may need to do a backtracking. Backtracking is itself nondeterministic, and may lead to untractable problems i.e NP complete problems.

In our approach, we know already that three diseases M_i , M_j and M_k have been identified from the data analysis. In order to choose a disease among them, we proceed as follows:

Consider a disease, say M_i , which belongs to the leaves of the tree. It's well known that there is one and only one path from M_i back to the root. Once this path is identified, we go down from the root until we reach the disease M_i and compute α_{M_i} . We do the same

for α_{M_j} and α_{M_k} .

3.3 Clinical conclusion

Let M_i , M_j , and M_k be the selected diseases from the data analysis. Then we retain the disease $M \in \{ M_i, M_j, M_k \}$ such that:

$$P(M,P) = \text{Max}(\mu_{M_i}(P)\alpha_{M_i}, \mu_{M_j}(P)\alpha_{M_j}, \mu_{M_k}(P)\alpha_{M_k}).$$

4 Conclusion

Our algorithm is faster and gives more precise diagnostic than usual expert system. We hope that it will render the physicians and all the population a great service.

Table 1 Data for the construction of the matrix R

	Fever	Headhache	Diarrhoea
Paludism 1 M_1	1	1			0
Paludism 2 M_2	1	1			0
.....					
.....					
Leprosy M_{18}	0	0			0
Rakoto	1	0			1
Rabe	0	1			1
.....					

Figure2 Projection of the diseases and of the patients

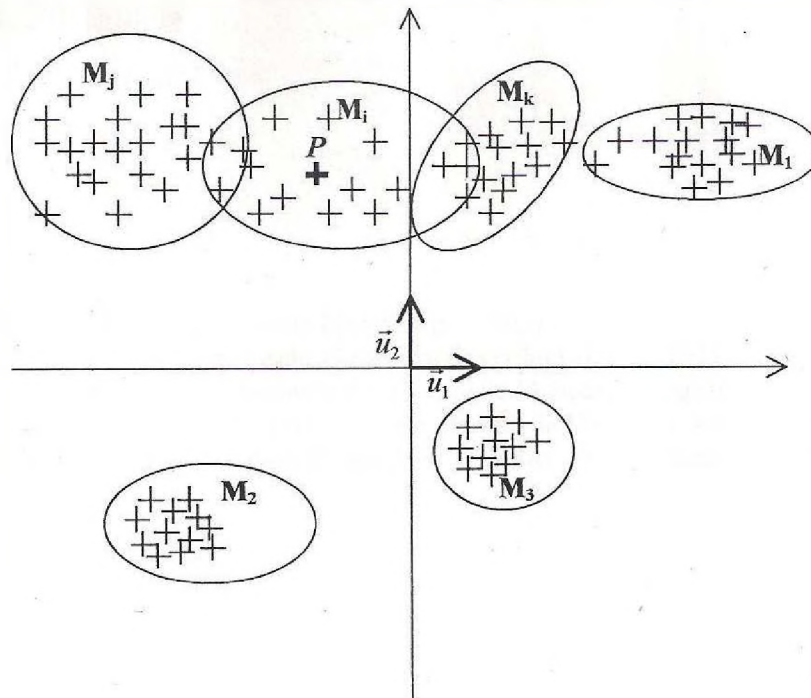
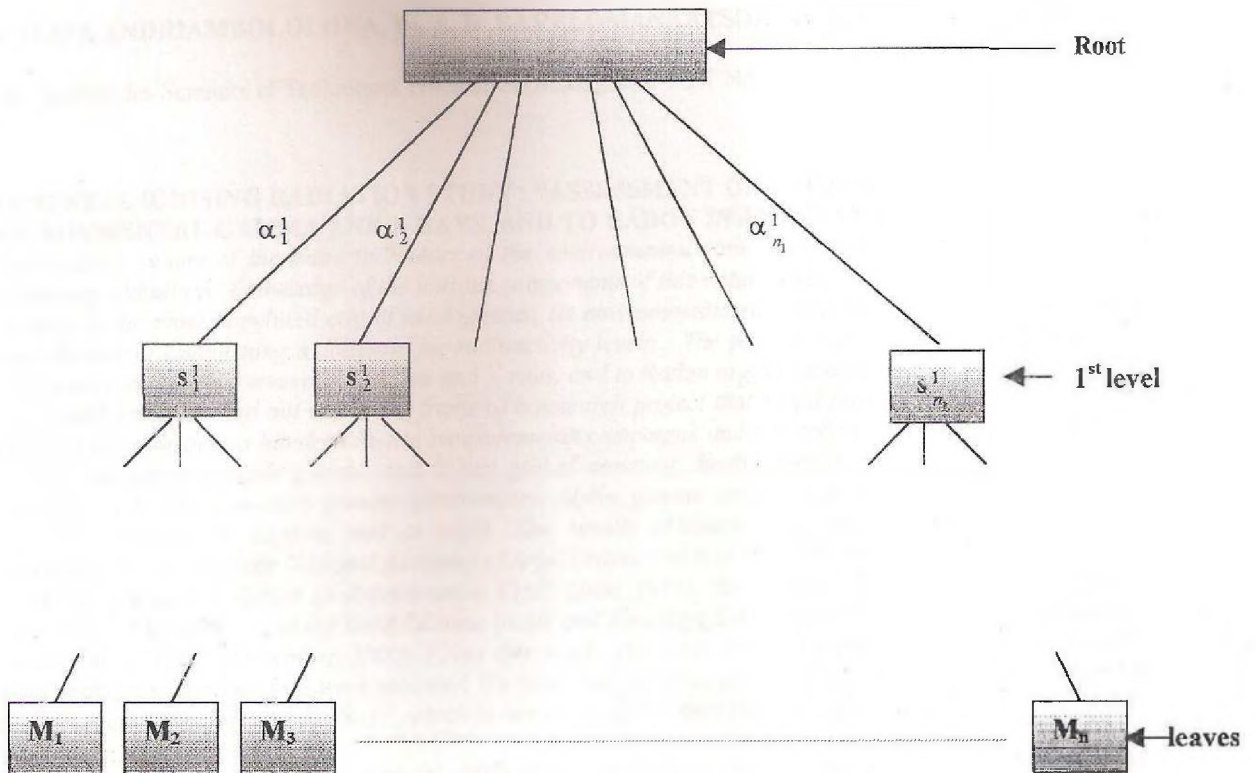


Figure 3 Architecture in tree of the system expert MITSABO



References

- [1] Marc GENTILINI Médecine tropicale Flammarion (1995)
- [2] F. AUBERT, P. GUITTARD L'essentiel médical de poche Ellipses (1995)
- [3] H. RAJAONA Gastro-entérologie médicale, Faculté de Médecine Antananarivo (1993)
- [4] M. RAZANAMPARANY Sémiologie médicale, Faculté de Médecine Antananarivo (1989)
- [5] R. L. Burden, J.D. FAIRES, Numerical Analysis, 3rd edition PWS Publishers (1985)