KYBERNETIKA — VOLUME 7 (1971), NUMBER 2

Linear Adaptive System for Medical Diagnosis

ARNOŠT VESELÝ

In the present paper a mathematical model of medical diagnostic process, based on the theory of linear adaptive systems is described. According to this model a Diagnostic System as a program for computer MINSK 22 has been realized. The proper part of the Diagnostic System is a net of linear adaptive neurons. To each neuron there corresponds some disease. On the base of patient's symptomcomplex and after finishing the adaptation, the neuron can determine, if the patient is suffering from the corresponding disease. In the present paper the Diagnostic System is described including its application in rheumathology.

1. INTRODUCTION

In order to use computers in medical diagnosis, it is first necessary to consider mathematical models of the diagnostic process. These mathematical models form the basis of computer programs. For this purpouse various mathematical models have been used. The most usual of them make use of mathematical logic and methods of probability theory [1, 2, 3]. In this paper we use theory of linear adaptive systems, based on principal works of McCulloch, Pitts and Rosenblatt [4, 5, 6]. According to this theory we had worked out a mathematical model of the diagonistic process and we used it as a base of a program for computer MINSK 22. In the paper this program, called Diagnostic System, will be described together with its application in rheumathology.

2. MATHEMATICAL MODEL OF THE DIAGNOSTIC PROCESS

First let us consider a diagonistic process. Suppose that a patient was examined, i.e. that it was stated, which symptoms of a set of all symptoms under consideration, the patient had. The state of the patient is described by a vector $\mathbf{S} = (s_1, \ldots, s_n)$ with n components. The component s_i of a vector $\mathbf{S}$ has a value 1 if the patient had the corresponding symptom. If the patient had not the symptom corresponding to the component s_i, the value of s_i is 0. The vector $\mathbf{S}$ will be called the symptomcomplex $\mathbf{S}$. The set of all symptomcomplexes will be denoted by $\mathscr{S}$. Assume that patients can suffer from one or more diseases of a set of all possible diseases $D_1, D_2, \ldots, D_m$. To each disease D_j there corresponds a set of symptomcomplexes $\mathscr{S}_j = \{\mathbf{S}_{ij}, \ldots, \mathbf{S}_{kj}\}$, $j = 1, 2, \ldots, m$, of those patients with disease D_j. Denote by $\mathscr{S}^0$ the set of symptomcomplexes of the patients without any disease $D_1, D_2, \ldots, D_m$, i.e. the set of symptomcomplexes of healthy persons. Then $\mathscr{S}^0$ will be the complement of the set $\mathscr{S}_1 \cup \cup \mathscr{S}_2 \cup \ldots \cup \mathscr{S}_m$ in the set $\mathscr{S}$. To determine the right diagnosis of the patient (i.e. to determine the patient's disease), it means to determine either which sets $\mathscr{S}_j$ ($j = = 1, 2, \ldots, m$) patient's symptomcomplex $\mathbf{S}$ belongs to, or if it belongs to $\mathscr{S}^0$. Therefore our task is to carry out the decomposition of the set $\mathscr{S}$ of all possible symptomcomplexes $\mathbf{S}$ into its subsets $\mathscr{S}_j$ on the base of statistical data.

For this purpose we use a net of adaptive linear neurons. A linear adaptive neuron is a finite automaton with n binary inputs and one output. Its state is described by a vector with $n + 1$ components: $\mathbf{W} = (w_0, w_1, \ldots, w_n)$ which is said to be a weight vector. The behavior of a linear adaptive neuron can be described by a linear weight function

$$g(\mathbf{X}, \mathbf{W}) = g(x_1, x_2, \ldots, w_0, w_1, \ldots, w_e) = w_0 + \sum_{i=1}^{n} x_i w_i$$

where $\mathbf{X}$ is an input vector (the value of its ith component equals to the value of the ith input of the linear neuron) and $\mathbf{W}$ is its weight vector. If $\mathbf{X}_1$ is an input vector, $\mathbf{W}_1$ a weight vector and δ a positive real number, then the value of the output of the linear neuron is:

1. 1 if $g(\mathbf{X}_1, \mathbf{W}_1) > \delta$,
2. 0 if $g(\mathbf{X}_1, \mathbf{W}_1) < -\delta$,
3. 2 if $-\delta \leqq g(\mathbf{X}_1, \mathbf{W}_1) \leqq \delta$.

Let $\mathscr{X} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ be a set, elements of which are vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$. Let this set $\mathscr{X}$ be classified in such a way, that each vector $\mathbf{X}_i \in \mathscr{X}$ belongs just to the one of subsets $\mathscr{X}_1 \subset \mathscr{X}$ or $\mathscr{X}_2 \subset \mathscr{X}$ (i.e. $\mathscr{X}_1 \cap \mathscr{X}_2 = \emptyset$ and $\mathscr{X} = \mathscr{X}_1 \cup \mathscr{X}_2$). If there exists a weight function $g(\mathbf{X}, \mathbf{W}')$ such that $g(\mathbf{X}_i, \mathbf{W}') > \delta$ holds for all $\mathbf{X}_i \in \mathscr{X}_1$ and $g(\mathbf{X}_i, \mathbf{W}') < -\delta$ holds for all $\mathbf{X}_i \in \mathscr{X}_2$, the subsets $\mathscr{X}_1, \mathscr{X}_2$ are said to be linearly separable. If subsets $\mathscr{X}_1, \mathscr{X}_2$ are linearly separable, then there exist methods for obtaining a weight function $g(\mathbf{X}, \mathbf{W}')$ (with properties described above) from an arbitrary weight function $g(\mathbf{X}, \mathbf{W}_0)$, after a finite number of steps.

Theorem. *Let subsets $\mathscr{X}_1, \mathscr{X}_2$ of a set $\mathscr{X}$ be linearly separable. Let $\mathbf{S} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots\}$ be a sequence of vectors $\mathbf{X}_i \in \mathscr{X}$ in which each vector $\mathbf{X}_i$ repeats infinitely many times and c be a positive real number. Construct the sequence of vectors $\mathbf{W}_0, \mathbf{W}_1, \mathbf{W}_2, \ldots$ in the following way:*

1. *Let $\mathbf{W}_0$ be arbitrary (i.e. let its each component $w_{0,j}$ be an arbitrary real number).*

2. a) *If one of the following conditions*

(1) $$g(\mathbf{X}_k, \mathbf{W}_k) > \delta \quad and \quad \mathbf{X}_k \in \mathscr{X}_1\,,$$

(2) $$g(\mathbf{X}_k, \mathbf{W}_k) < -\delta \quad and \quad \mathbf{X}_k \in \mathscr{X}_2$$

is satisfied, then $\mathbf{W}_{k+1} = \mathbf{W}_k$.

2. b) *If the condition*

(3) $$g(\mathbf{X}_k, \mathbf{W}_k) \leqq \delta \quad and \quad \mathbf{X}_k \in \mathscr{X}_1$$

holds, the vector $\mathbf{W}_{k+1}$ *is constructed according to*

(4) $$w_{k+1.0} = w_{k.0} + c \quad and \quad w_{k+1.j} = w_{k.j} + cx_{k.j} \quad \text{for} \quad j = 1, 2, \ldots, n\,.$$

2. c) *If the condition*

(5) $$g(\mathbf{X}_k, \mathbf{W}_k) \geqq -\delta \quad and \quad \mathbf{X}_k \in \mathscr{X}_2$$

holds, the vector $\mathbf{W}_{k+1}$ *is constructed according to*

(6) $$w_{k+1,0} = w_{k,0} - c \quad and \quad w_{k+1,j} = w_{k,j} - cx_{k,j} \quad for \quad j = 1, 2, \ldots, n\,.$$

For this sequence of vectors $\mathbf{W}_0, \mathbf{W}_1, \ldots$ *the following statement is true: There exists such positive integer* n_0*, that from this* n_0 *all the following equations are satisfied:* $\mathbf{W}_{n_0} = \mathbf{W}_{n_0+1} = \ldots$

The proof of this theorem can be found for example in [4, 5]. Evidently for the weight function $g(\mathbf{X}, \mathbf{W}_r)$ the following holds: For all $r > n_0$ and for all $\mathbf{X}_i \in \mathscr{X}_1$ we have $g(\mathbf{X}_i, \mathbf{W}_r) > \delta$ and for all $r > n_0$ and for all $\mathbf{X}_i \in \mathscr{X}_2$ we have $g(\mathbf{X}_i, \mathbf{W}_r) < -\delta$. Therefore if a weight vector of an adaptive linear neuron, arbitary at the beginning, is changed in the way mentioned in the preceding theorem, then after n_0 steps the output of the neuron will be 1 for all $\mathbf{X}_i \in \mathscr{X}_1$ and 0 for all $\mathbf{X}_i \in \mathscr{X}_2$.

The procedure of changing the weight vector is called *adaptation with constant correction*.

Another way of adjustement of the weight vector is sometimes used. This adjustement of the weight vector differs from the preceding adjustement in the following way.

In 2. b) (see preceding theorem) the sequence of vectors $\mathbf{W}_k^0, \mathbf{W}_k^i, \ldots, \mathbf{W}_k^r$, is constructed ($c$ is a real positive number):

α) $\mathbf{W}_k^0 = \mathbf{W}_k$,

β) if $g(\mathbf{X}_k, \mathbf{W}_k^i) \leqq \delta$ then $\mathbf{W}_k^{i+1}$ is constructed according to

(7) $$w_{k,0}^{i+1} = w_{k,0}^i + c \quad \text{and} \quad w_{k,j}^{i+1} = w_{k,j}^i + cx_{k,j} \quad \text{for} \quad j = 1, 2, \ldots, n\,.$$

Obviously there exists such i for which

(8) $$g(\mathbf{X}_k, \mathbf{W}_k^i) > \delta$$

holds (observe that c is positive).

Let $i = r$ be the least i, for which (8) holds. Then put $\mathbf{W}_{k+1} = \mathbf{W}_k^r$.

In 2. c) the sequence $\mathbf{W}_k^0, \mathbf{W}_k^1, \ldots, \mathbf{W}_k^s$ is constructed where:

α) $\mathbf{W}_k^0 = \mathbf{W}_k$,

β) if $g(\mathbf{X}_k, \mathbf{W}_k^i) \geqq -\delta$ then $\mathbf{W}_k^{i+1}$ is constructed according to

(9) $$w_{k,0}^{i+1} = w_{k,0}^i - c \quad \text{and} \quad w_{k,j}^{i+1} = w_{k,j}^i - cx_{k,j} \quad \text{for } j = 1, 2, \ldots, n\,.$$

Evidently there exists the least $i = s$ for which

(10) $$g(\mathbf{X}_k, \mathbf{W}_k^i) < -\delta$$

holds. Then put $\mathbf{W}_{k+1} = \mathbf{W}_k^s$.

This adjustement of weight vector is called the *adaptation with absolute correction*. For this adaptation procedure it is easy to prove (on account of the preceding theorem) a similar convergence theorem as for the adaptation procedure with constant correction.

Let $\mathscr{X}$ be a set and $\mathscr{X}_1, \mathscr{X}_2, \ldots, \mathscr{X}_m$ its subsets. Let $\mathscr{X}^0$ be the complement of the set $\mathscr{X}_1 \cup \mathscr{X}_2 \cup \ldots \cup \mathscr{X}_m$ in the set $\mathscr{X}$. Construct a net of m linear adaptive neurons, in which the ith inputs of all neurons are connected together. Such net of m adaptive neurons has n inputs and m outputs. The ith neuron of this net is capable, after the procedure of adaptation, to discriminate between the subset $\mathscr{X}_i$ and its complement $\mathscr{X}_i^0$ in the set $\mathscr{X}$, if the condition of linear separability is satisfied. Let $\mathscr{X}_1, \mathscr{X}_2, \ldots, \mathscr{X}_m$ be subsets and $\mathscr{X}_1^0, \mathscr{X}_2^0, \ldots, \mathscr{X}_m^0$ their complements in the set $\mathscr{X}$. Let m classifications of the set $\mathscr{X} : \{\mathscr{X}_1, \mathscr{X}_1^0\}, \ldots, \{\mathscr{X}_i, \mathscr{X}_i^0\}, \ldots, \{\mathscr{X}_m, \mathscr{X}_m^0\}$ be given. If the condition of linear separability is satisfied for each classification $\{\mathscr{X}_i, \mathscr{X}_i^0\}$, then the net of m linear adaptive neurons is capable, after the procedure of adjustement of weight vectors, to determine correctly for all vectors $\mathbf{X} \in \mathscr{X}$ and for all classifications $\{\mathscr{X}_i, \mathscr{X}_i^0\}$ if $\mathbf{X} \in \mathscr{X}_i$ or $\mathbf{X} \in \mathscr{X}_i^0$. The ith output of the net of neurons is 1 or 0 depending on wheter $\mathbf{X} \in \mathscr{X}_i$ or $\mathbf{X} \in \mathscr{X}_i^0$ respectively.

Return to the problem of the stating of medical diagnosis. Let $\mathscr{X}$ be a set $\mathscr{S}$ of all possible symptomcomplexes and let $\mathscr{X}_1, \ldots, \mathscr{X}_m$ be subsets $\mathscr{S}_1, \ldots, \mathscr{S}_m$ corresponding to diseases $\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_m$. If all couples of subsets $\{\mathscr{S}_1, \mathscr{S}_1^0\}, \ldots, \{\mathscr{S}_m, \mathscr{S}_m^0\}$ are linearly separable, it is possible to construct a net of m linear adaptive neurons, which will correctly determine (after a finite step adaptation) if $\mathbf{S}$ belongs to $\mathscr{S}_i$ or $\mathscr{S}_i^0$, for each $i = 1, 2, \ldots, m$ and all $\mathbf{S}$. It means to determine, according to the symptomcomplex $\mathbf{S}$ of the patient, from which of diseases $\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots \boldsymbol{D}_m$ the patient is suffering or if he is a healthy person. Thus the problem of the diagnosis stating is solved.

3. DIAGNOSTIC SYSTEM AND ITS APPLICATION

In the preceding chapter the adaption algorithm for the net of linear adaptive neurons has been described. This net of neurons is a part of Diagnostic System realized on the computer MINSK 22 in the symbolic language SADRS. In the following this Diagnostic System (DS) will be briefly described.

DS suppose that symptomcomplexes of patients and weight vectors of the net of neurons are stored on magnetic tapes. At the beginning of the adaptation procedure

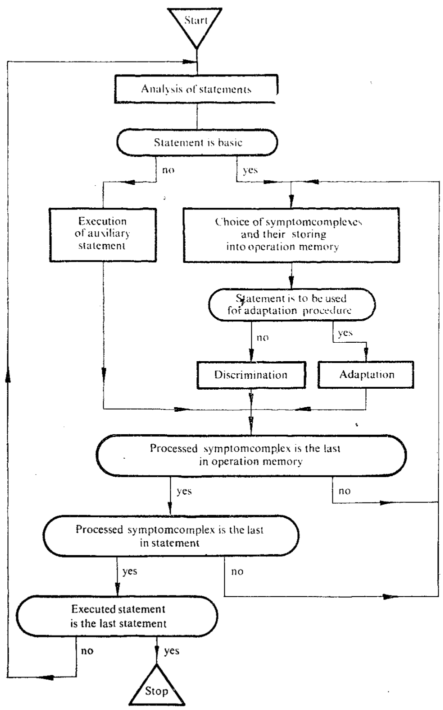

Fig. 1.

the weight vectors are the zero vectors. DS is controlled by statements. The basic statements determine either the adaptation procedure or the discrimination procedure, in which DS joins corresponding diseases to a given symptomcomplex. For example a statement u (10, 20−30);, is a statement for adaptation procedure with constant correction, regarding to the symptomcomplexes with number of order on magnetic tape 10, 20−30. There are some auxiliary statements, for example the statement for output of weight vectors *out v*;, or the statement for giving initial weight vectors at the beginning of the adaptation procedure *pv*;. DS consists of the direction part and of the proper adaptive system (net of linear adaptive neurons). At first the direction part analyses a given statement. If it is some of auxiliary statements, it is carried out

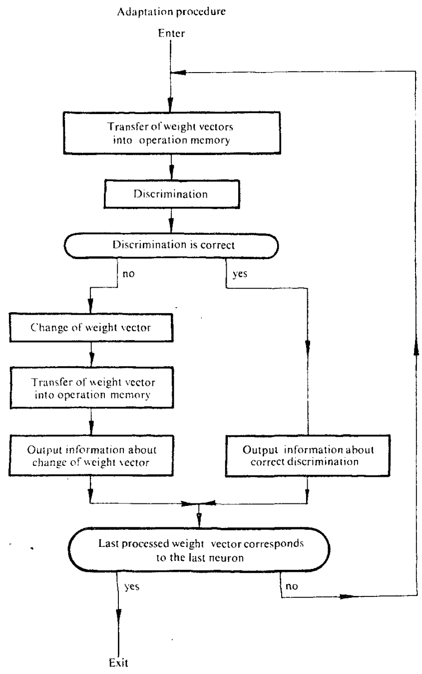

Fig. 2.

immediately. If some basic statement is concerned, direction part analyses symptomcomplex part of this statement, it picks up corresponding symptomcomplexes from magnetic tape and it stores them into operation memory. Then the direction is handed over the proper adaptive system, which processes symptomcomplexes stored in the operation memory. When all these symptomcomplexes are processed, the direction part picks up following symptomcomplexes from magnetic tape. Then these new symptomcomplexes are processed by the proper adaptive system. When the last processed symptomcomplex is the last symptomcomplex recorded in the symptomcomplex part of the statement, the following statement is analysed. The operation of DS is obvious from the flow chart of Fig. 1, 2, 3. DS can process symptomcomplexes with less than 2000 binary symptoms and it is able to realize a linear adaptive net consisting maximally of 50 nerons (it means that it can discriminate maximally among

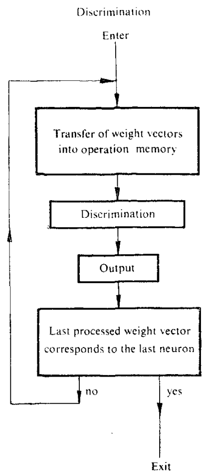

Fig. 3.

50 diseases, what seems to be, from a point of view of supposed applications, quite satisfactory).

DS was applied in rheumathology, in collaboration with Institute of Rheumathology in Prague. In the following some results of this application will be mentioned. We used symptomcomplexes of two different groups of patients. In the first group there were patients with rheumathoid arthritis, in the second group there were healthy persons. The symptomcomplexes were composed of non-specific symptoms, it means that on the base of these symptoms a physician was not able to predict with certainty, if the patient would fall ill with rheumathoid arthritis or would not. DS was adapted to discriminate between these two groups. We had a group of 360 symptomcomplexes (180 symptomcomplexes of patients with rheumathoid darthritis and 180 symptomcomplexes of healthy persons). DS was adapted on 260 symptomcomplexes (130 symptomcomplexes of patients with rheumathoid arthritis and 130 symptomcomplexes of healthy persons). A quality of discriminating was tested on a group of 100 symptomcomplexes. There were applied both algorithms for adaptation. The results are given in Table 1 showing per cents of symptomcomplexes of the tested group determined correctly, uncorrectly and per cents of undecided symptomcomplexes.

Table 1.

Symptomcomplexes	Adaptation with	δ	Decision		Without Decision
			Correct	Uncorrect	
Group of 50 healthy persons	Absolute correction	5	72	2	26
		0	84	8	8
	Constant correction	5	68	2	30
		0	84	16	0
Group of 50 patients with rheumathoid arthritis	Absolute correction	5	80	2	18
		0	82	16	2
	Constant correction	5	64	8	28
		0	84	8	8
Both groups together	Absolute correction	5	76	2	22
		0	83	12	5
	Constant correction	5	66	5	29
		0	84	12	4

4. CONCLUSION

Convergence theorem (chapter 2) for algorithm of a net of linear adaptive neurons, which is DS based on, was proved under the condition of linear separability of couples of sets $\{\mathscr{S}_1, \mathscr{S}_1^0\}, \ldots, \{\mathscr{S}_m, \mathscr{S}_m^0\}$. If this condition is not satisfied, properties of this algorithm are unknown. There was constructed a simple example of two linearly nonseparable sets, for which this algorithm (for arbitrary sequence of input vectors $\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots,$ presented in adaption process) did not converge. It oscilated between several weight vectors and none of them was the most optimal (in the sense of minimalization of the number of uncorrect decisions). But also for application on linearly nonseparable sets, this algorithm can be usefull when the number of uncorrect decisions is relatively small. The quality of decision making has to be tested.

Diagnostic System was made for continuens processing of medical diagnostic data. For application of DS medical data are systematically collected in rheumathology. At the present time, data are collected also for other rheumathic diseases. A comparison of decision making quality of DS and a group of physicians is in the preparation, too.

(Received November 21, 1969.)

REFERENCES

[1] М. Л. Быховский: Вероятностная логика построения самообучаевося диагностического процеса на математических машинах. Эксп. хирургия *7* (1962), 1, 3—11.

[2] R. S. Ledley, L. B. Lusted: Reasoning Foundations of Medical Diagnosis. Science *130* (1959), 3366, 9—21.

[3] R. S. Ledley: Use of computers in Biology and Medicine. McGraw-Hill, New York 1965.

[4] N. J. Nilsson: Learning Machines. McGraw-Hill, New York 1965.

[5] A. Novikoff: On Convergence Proofs for Perceptrons. Proceedings of the Symposium of Mathematical Theory of Automata, New York 1962.

[6] F. Rosenblatt: Principles of Neurodynamics. Washington D.C. 1962.

VÝTAH

Lineární adaptivní systém pro lékařskou diagnostiku

Arnošt Veselý

Je popsán matematický model stanovení lékařské diagnózy, využívající teorie adaptivních systémů. Na základě tohoto matematického modelu byl na počítači MINSK 22 realizován diagnostický systém. Vlastní částí diagnostického systému je síť lineárních adaptivních neuronů, při čemž každý neuron této sítě přísluší jedné z chorob, mezi kterými rozlišujeme. Váhové vektory jednotlivých neuronů jsou v procesu učení upraveny tak, aby každý neuron byl schopen stanovit, na základě symptomkomplexu pacienta a po ukončení procesu učení, má-li pacient jemu příslušnou chorobu. Diagnostický systém je v článku popsán zároveň s jednou jeho dílčí aplikací v reumatologii.

Ing. Arnošt Veselý, Ústav hematologie a krevní transfuse (Institute of Hematology and Blood Transfusion), U nemocnice 1, Praha 2.