KYBERNETIKA — VOLUME 26 (1990), NUMBER 3

ON NEURAL NETWORKS

JIŘÍ BENEŠ

Typical neural network models representing a broad survey of actual trends are considered in an unifying way as Complexes subject to specific control systems, called Formators which may have a two-level arrangement and incorporate human operators. The importance of injected random noise in the process of organization is stressed. Potential applications of neural networks include their use as parts of formators for pattern classification in situational control. Prospective research areas are stated. One possible trend is to bring closer together the theory of neural networks and that of generalized cellular automata.

1. INTRODUCTION

The renewal of interest in this interdisciplinary field is certainly due to the progress in microelectronics which enables to consider the realization of complex models of neural networks. Such endeavour to build artefacts possibly embodying some of the properties of biological neural tissues is stimulated also by progress in neurology. Still there remains the fundamental obstacle in attempting to build such models, i.e. the absence of the understanding the very essence of thought processes in living beings. Sound but very cautious optimism may lead us to consider both theoretical modelling and empirical simulation on a wide but selected population of neural networks as means of carrying on research to approach an elucidation of thinking and its constituents, such as situation assessment, different forms of memories, decision making etc., but also of its important form: of the creative thinking manifesting itsel in giving form (shape) and function (behaviour) to objects. Meanwhile, and without yet considering such a naturally related essential goal, the attainment of which would bring mankind undoubtedly very far ahead, there is a wealth of novel practical applications of models of neural networks as artefacts, e.g. in information processing as: pattern recognition, specific computation tasks, or vector quantization techniques for communication purposes, etc. which explain the recent wide and intense interest in neural networks.

2. NEURAL NETWORKS AS PARTS OF CYBERNETIC SYSTEMS

A neural network model is essentially a Complex constituted by a large number of Elements, linked by connections which have their weights. The Elements are nodes and may be represented as carrying out a summation of their input signals and then a subsequent typical nonlinear transformation. The dynamics of the neural network concerns the changes of its form (structure), which is connected generally with changes of the connection weights, and of its function.

The generally accepted relative simplicity of its Elements (nodes, processing elements), which is a major asset of the neural network model and which distinguishes it from complex networks of computers (processors) or from complexes of many-state automata, does not permit to class such a model among Cybernetic systems with autonomous automatic organization (which occur e.g. in chemistry) and where the process of self-control is based upon the differentiated properties of the elements and their mutual interactions.

Not only for theoretical but also for practical reasons (in order to be able to reach intentionally a variety of complicated and sophisticated goals) it is wise to consider the neural network model as a Complex which is a part of a Cybernetic system with automatic control of organization, where the process of organization of a Complex of Elements is based upon the interaction of a special control system, called Formator, and the Complex. The Cybernetic System is formed by a feed-back system arrangement, where the Formator influences the organization of the Complex through its acting variables (Fig. 1).

It is advantageous that the control of neural network models can thus be considered as a special case of the formator control of a distributed net of centres. This enables us to use a more general approach, known from large and complex system control. In conformity with the principles of formator control, even man can be a constituent part of the formator. In the case of neural network models he can fulfil the function of the designer or of the teacher or of the supervisor.

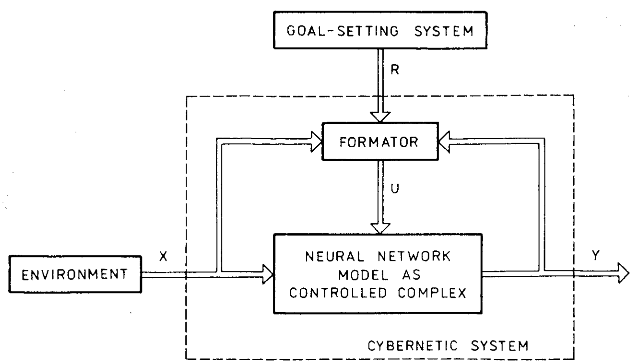

Fig. 1. Neural network model as part of a Cybernetic System.

Even if the taxonomy of neural networks knows certain types of networks, which are functioning without a teacher or a supervisor, we still consider for practical reasons as useful to recommend in the general scheme of the Cybernetic system, constituted by the complex of the neural network and by the formator, a hierarchically layered arrangement of the formator into two levels (Fig. 2).

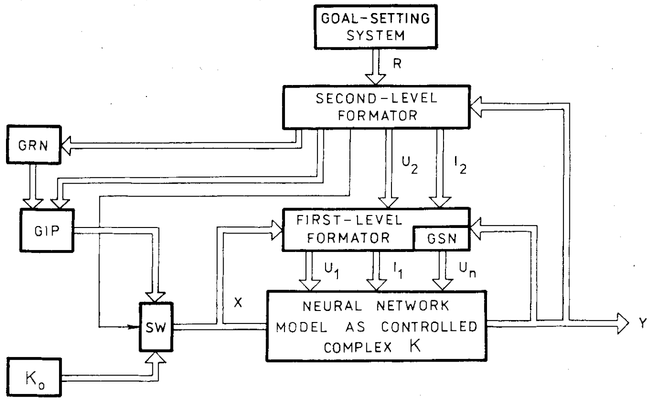

Fig. 2. A two-level arrangement of formators for the control of the neural network model.

This enables e.g. to incorporate a human operator into the second-level formator and to charge him with such functions (of a designer, of a teacher, of a supervisor) which he may fulfil together with the aid of the first-level formator.

Here again, in the control element of the first-level formator, a human operator may be present, exerting decisions about the acting variables U_1 and further informational command variables I_1 to be applied by the first-level formator upon the neural network complex. An important class of these acting variables may be injected noise U_n from a random process generator GSN simulating the so-called synaptic noise.

The second-level formator gets its command variables R from a hierarchically higher placed system for goal-setting. This formator controls the first-level formator by its acting variables U_2 and by its further informational command variables I_2. It can operate a switch SW conveying to the input of the complex of the neural network (and simultaneously upon the first-level formator) either the output variables K_0 of the external environment, or the output of a generator of input patterns (situations) GIP, which may be distorted intentionally by a generator of random noise GRN.

Generally, the acting variables of the first-level formator can influence these parameters of the model:

a) the threshold of the elements. Its value can be differentiated in space and time according to the zones of the complex;

b) the connection weights. They are changed by the formator according to learning rules which are specific for the type of the neural network used. Generally an initialization of the network is required, i.e., the weights are to be set at the beginning as a requisite for their subsequent adaptive changes;

c) the injected noise.

3. THE FARLEY AND CLARK MODEL

This is one of the first models, simulating the evolution of a neural network on a digital computer. It stems from the Lincoln Laboratory at M.I.T. as early as 1954 [1]. The formator of the self-organizing system is called there the modifier. This model embodies basic principles and relations between the parameters in neural networks which are valid and still of major interest. Particularly it uses the injection of random Gaussian noise into the Complex in order to modify its behaviour, to make it less stiff (to randomize its structure) and to permit hereby the establishment of new connections between the elements. Noise is thus used as one of the control agents to ease the organization of the Complex.

J. Faber [2] has stressed the importance of this principle in neurology and has elaborated a comparison of the brain to self-organizing Cybernetic models, where the cortex is the Complex, the sensory-thalamic system is a signal generator, the limbic system is a discriminatory or analytical unit and where the ascending activation reticular system, the raphe nuclei, and the locus caeruleus and nucleus gigantocellularis are formators.

4. A SHORT SURVEY OF THEORETICAL APPROACHES

At the present state of research, let us point out these selected approaches:

a) the Hopfield model and its extensions [3], [4], [8] (see Section 5);
b) the additive continuous neural network model [13] (see Section 6);
c) the back-propagation training algorithm [5] (see Section 7);
d) the adaptive resonance theory [6], [8] (see Section 8);
e) the theory of chaotic switches in the modeling of neural networks [7] (see Section 9).

5. THE HOPFIELD MODEL AND ITS EXTENSIONS

The original model by J. J. Hopfield (1982) [3] is a neural network with binary input variables. Each of the N elements may acquire one of the two states: $s_i = +1$ or $s_i = -1$. For such network it is possible to introduce for the special symmetric case $b_{ij} = b_{ji}$ a so-called global energy function

$$E = -\tfrac{1}{2} \sum_{i,j=1}^{N} b_{ij} s_i s_j \,, \tag{1}$$

where b_{ij} is the weight of the connection from the jth element (neuron) to the ith neuron.

There is an essential analogy between this neural network and the Ising spin system in physics [9], where E has then the meaning of energy and an expression analogous to (1) is valid for the deterministic system at absolute zero temperature. At a finite temperature the spin system in this theory has a stochastic behaviour.

J. J. Hopfield has suggested to use the local minima of the global energy function (1) as attractors for the storage of the input patterns presented to the neural network, acting thus as associative memory. With the use of so-called Hebbian learning rule he has deduced an expression for the pertinent connection weights:

$$b_{ij} = \frac{B}{N} \sum_{m=1}^{M} x_i^m x_j^m\,, \quad i \neq j\,, \quad B > 0\,, \tag{2}$$

where $\{x_i^m\}$ for $i = 1, \ldots, N$ is the mth pattern from the ensemble of M memorized input patterns.

Besides the symmetry condition $b_{ij} = b_{ji}$, there are further conditions for the validity of (2): a) the presented input patterns are random and mutually uncorrelated; b) the connection weights are the result of learning only (the so-called premiss of tabula rasa at the beginning); c) each Element (neuron) is connected with all others; d) the number N is very large (for a thermodynamic analogy $\lim N \to \infty$ would even be required).

It is reported that it has been experimentally confirmed that Hopfield's neural network functions well as a part of a classifier, even if condition (c) is by far not fulfilled. Besides this type of network is steady enough against a further interruption of connections between the elements.

Let us consider the formator control of Hopfield's neural model, used as a part of a classifier of input patterns (situations). The classifier is here a Cybernetic System with feedback as in Fig. 1, the neural network is the controlled Complex, the Control System is a formator with two levels as in Fig. 2.

The classes into which we wish to distribute the patterns are chosen in the second-level formator: it is a set of M templates $\{x_i^m\}$, $i = 1, \ldots, N$. The information about them is passed to the first-level formator, where the connection weights b_{ij} are computed according to (2) and used as acting variables for the setting of the weights in the neural network model. A further required step is the setting of the starting non-equilibrium state of the neural network (the so-called initialization of the network) which is done by the acting variables of the first-level formator by setting $y_i(0) = x_i$, $i = 1, \ldots, N$, where x_i is the value of the input variable of the ith element when a pattern $\{x_i\}$ is presented from the generator of input patterns GIP which is controlled by the second-level formator (Fig. 2). Then, the neural network is left to its free development, during which following iterations take place [8]:

$$y_j(t + 1) = f\left[\sum_{i=1}^{N} b_{ij}\, y_i(t)\right]. \tag{3}$$

The iterations follow until one of the local equilibrium states is reached, when the values y_j do not change in time any more. The set $\{y_j\}$ is then the mth template (from the total number of M templates) which is the nearest to the presented pattern $\{x_i\}$; hereby the classification is performed.

An extension of Hopfield's network results from the injection of noise. H. Sompolinsky [4] has investigated a network, where instead of (2), the expression for the connection weights was:

$$b_{ij} = \frac{B}{N} \sum_{m=1}^{M} x_i^m x_j^m + \delta_{ij}\,, \quad i \neq j\,, \quad B > 0\,, \tag{4}$$

where δ_{ij} is the synaptic noise which is here an analogy to noise in spin glasses (with symmetric random (positive and negative) interactions between the spins) in physics [10], and which has a Gaussian distribution with mean values in time: $\overline{\delta_{ij}} = 0$ and $\overline{\delta_{ij}^2} = \delta^2/N$. These values are not ensemble means when sets of input patterns are presented and δ_{ij} is not correlated with these patterns. In Fig. 2 it is assumed that the noise is introduced into the model of the neural network from GSN, a generator of synaptic noise in the first-level formator as one of its acting variables. The investigation of the frequently very positive influence of the injected noise upon the automatic organization of the neural network is known already from the Farley and Clark model, quoted in Section 2 [1], [2].

Another extension of Hopfield's model is the randomly diluted version which results from a randomly effected interruption of a finite number of its connections [4]. The weights are then:

$$b_{ij} = \frac{c_{ij}}{Nc} \sum_{m=1}^{M} x_i^m x_j^m\,, \quad i \neq j\,, \tag{5}$$

where the average number of connections of one element after the completion of the random cuts is Nc, and where each c_{ij} (with the symmetry condition $c_{ij} = c_{ji}$) is a random variable which takes the value 1 or 0 with the probability c or $1 - c$.

The choice of the probability c when studying the reliability (steadiness of the model against the cutting of connections) would be done by the second-level formator which would pass this information to the first-level formator which would calculate the pertinent weights after (5) as acting variables upon the neural network.

6. THE ADDITIVE CONTINUOUS NEURAL NETWORK MODEL

It has been recently further investigated by A. Guez et al., cf. [13]. The dynamics of this network of N elements as a nonlinear system can be expressed by these equations:

$$\frac{\mathrm{d}\mathbf{s}(t)}{\mathrm{d}t} = -\mathbf{T}(t)\,\mathbf{s}(t) + \mathbf{B}f[\mathbf{s}(t)] + \mathbf{x}\,, \tag{6}$$

where **T** is a diagonal matrix, the diagonal elements of which are τ_i, where τ_i is the time constant of the change of the state s_i of the ith element; **B** is an $N \times N$ matrix, the elements of which are b_{ij} – the connection weights for the transmission from the jth element to the ith element. Let us stress that this matrix need not be symmetrical. This is different from a major premiss of Hopfield's theory. Further, in eq. (6): $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_N(t)]^T$ is the state vector of the ensemble of elements in an N-dimensional Euclidean space, where $s_i(t)$ is the state of the ith element (neuron),

$$\mathbf{f}[\mathbf{s}(t)] = \{f_1[s_1(t)], f_2[s_2(t)], \ldots, f_N[s_N(t)]\}^T,$$

where $f_i(s_i)$ is a function expressing the dependence of the output variable of the ith element upon its state s_i. This function is to be differentiable in s_i, but may not be a sigmoid, as encountered in many neural networks. Further: $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_N(t)]^T$ – the vector of the input variables of the neural network.

From the point of view of formator control we see that we can essentially change the properties of this network by influencing the elements of matrices **T** and **B** by the acting variables of the formator. The question arises how to choose these matrix elements if we wish that the endpoint of an a priori selected state vector $\mathbf{s}^*$ of the ensemble of the network elements be a stable equilibrium point of the neural network, which we intend to exploit in order that the network recognizes a class of input patterns (situations) a priori selected by us.

In 1988 Guez, Protopopsecu and Barhen [13] showed that in order that $\mathbf{s}^* \in \mathbb{R}^N$ meets this requirement, it is necessary that τ_i, b_{ij} for $i, j = 1, \ldots, N$ comply with certain conditions. Similarly conditions have been formulated for the case when we wish a certain number M of a priori selected points to be stable equilibrium points of the neural network. This permits the determination of suitable values τ_i and b_{ij} which we can consider directly as values of the acting variables of the formator control.

In Fig. 2 these acting variables can be identified with the acting variables of the first-level formator which would also compute τ_i and b_{ij}, respecting the constraints, imposed upon them, whereas the choice of the ensemble of equilibrium points would be done by the second-level formator, particularly on the basis of information about the aims of the function of the whole Cybernetic System, coming from the Goal-setting system.

7. THE BACK-PROPAGATION TRAINING ALGORITHM

One of the general principles which can be applied is the partition of the complex of the neural network into zones. These zones can be acted upon in a selective way by the acting variables of the formator. In some cases these zones can form layers, e.g. in the case of the layered network which has a layer of inputs at the bottom, a number of intermediate layers and a layer of output elements at the top. Such layered neural networks are used when applying the method of training by back-propagation algorithm (E. Rumelhart et al., [5]). It is assumed that there are no connections between the Elements within a layer. A typical arrangement is a three-layer perceptron with a layer of continuous-valued inputs, two intermediate (hidden) and one output layer of Elements, where three sets of connection weights (to the elements of each layer) are to be adjusted. Moreover, the thresholds of the elements can be adjusted by considering them as weights of connections from specific constant inputs. The nonlinearities of the Elements are continuously differentiable, generally sigmoids.

When the back-propagation algorithm is used to train such a network, this can be considered as a form of supervised learning. In such a case the total error in the response of the neural network, observed by the second-level formator is reduced by adjusting the connection weights through the acting variables of the first-level formator. This total error $\mathscr{E}$ is

$$\mathscr{E} = \tfrac{1}{2} \sum_{c} \sum_{j} (y_{j,c} - d_{j,c})^2 \tag{7}$$

where y_j is the actual value of the output variable of the jth output element; d_j is the desired value of this variable; c is an index labelling the input-output pairs. The reduction of $\mathscr{E}$ is done by the recursive back-propagation algorithm, described in [8] and used in the first-level formator for the computation of the weight adjustments. This algorithm computes weights starting at the output layer Elements back to the Elements in the intermediate layers and to the inputs. Theoretically, such kind of three-layer perceptrons has versatile properties, even if the number of presentations of training data may be large [8]. Still the appropriate number of layers of such neural networks and its arrangement are a matter of research (M. A. Jones, 1988).

8. THE ADAPTIVE RESONANCE THEORY [6] [8]

This is a further developing theory of adaptation of networks, basically trained without supervision, and called self-organizing networks. The theory aims at securing so-called autonomous learning of a type when learning remains adaptive to significant events and at the same time is stable in the presence of irrelevant events. This may be secured by properly focusing attention from the higher level upon information coming from beneath. G. A. Carpenter and S. Grossberg [6] have elaborated a number of schemes using the adaptive resonance theory architecture, aiming at quick and stable learning in a nonstationary environment. R. P. Lippmann [8] has described a Carpenter-Grossberg classifier, showing that feed-back connections are provided from the output elements (nodes) to the input elements. An interesting item is the setting of a matching threshold ranging between 0 and 1 and called vigilance. The resistance to noise of this classifier is to be improved.

9. THE THEORY OF CHAOTIC SWITCHES IN THE MODELING OF NEURAL NETWORKS

This theory, which has till now been only sketched by M. Peschel et al. in 1988 [7] is based on a synthesis of a number of principles. One of them concerns the dynamics of mutually interacting elements, each described by $dy/dt = y.f(x_1, x_2, \ldots, x_k)$, where x_i are inputs and y is the element's output. These elements are coupled together into Lotka-Volterra networks. Another principle is the assumption of specific chaos in the behaviour of the individual elements (neurons), which are modelled by so-called Evolon modules. From the point of view of formator control, following acting variables upon such neural network model are available: the change of the chaos intensity parameters of the individual elements and the change of the connection parameters between the elements. A two-level arrangement of formators could again be used, the second-level formator embodying the decision functions.

10. THE TECHNOLOGY OF ADAPTING THE WEIGHTS

This is one of the key problems in the realization of neural network models.

One example of a solution is the VLSI chip of AT & T Bell Laboratories with 54 neurons and about 3000 programmable synapses. It uses complementary metal oxide semiconductor (CMOS) technology with 2·5 micrometer design rules. The chip contains 75 000 transistors and has an area of $(6 \times 6)\,mm^2$. The computations are analog, but the input, output and control signals are digital [11].

There is another development at Bell Communications Research of an experimental chip for the so-called Boltzmann machine by J. Alspector et al. with 6 neurons and 15 connections. The circuit is analog and a digital control permits small discrete changes of the synaptic weights.

At the MIT Lincoln Laboratory chips with analog synaptic weights are developed (Sage et al.), using MNOS/CCD technologies.

At the Jet Propulsion Laboratory a programmable 32×32 binary synaptic CMOS chip has been developed.

A whole application area of adjustable weights is represented by the coprocessors of neurocomputers, embodying the hybrid concept.

Still, as [11] reports, circuit chips for neuronal network models now have at most a few hundred electronic neurons.

A trend has appeared to use optical techniques for the connections between the elements, as they might be done in 3 dimensions. This constitutes a major challenge for optoelectronics.

Another problem is the control of the internal threshold of the Elements by which their states may be influenced.

11. APPLICATIONS OF NEURAL NETWORKS

Already three decades ago there have been expectations of the use of Perceptrons for quick target recognition. The book of M. Minsky and S. Papert about Perceptrons (1969) has had a critical impact upon these hopes.

Nowadays, after years of Research and Development, some of the potential applications of neural networks are:

A) A broad class of pattern classification tasks using vector quantization technique. One of them is e.g. massive image compression for video transmission over a single-integrated-services digital network (ISDN) (E. C. Posner, 1986). Another application if for quick retrieval in an associative memory or for a quick complex situation assessment. Another field is the compression of speech.

At this point let us express more in detail the use of the principle of vector quantization for the pattern (situation) recognition function of the Cybernetic System composed by the neural network and the formator. Let the N-dimensional vector of the input variables of the neural network $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$ be formed by the values of continuous random variables which enter the individual Elements of the network. As a vector of the output variables we get an N-dimensional vector $\mathbf{y}$, the coordinates of which have discrete amplitudes. This output vector is $\mathbf{y} = q(\mathbf{x})$, where q is the so-called quantization operator. If provision is taken, $\mathbf{y}$ belogs to a finite set $\mathbf{Y} = \{y_m, m = 1, \ldots, M\}$, where the mth code vector or template is $\mathbf{y}_m = [y_{m1}, y_{m2}, \ldots, y_{mN}]^T$. In conformity with the terminology of speech coding, the set $\mathbf{Y}$, i.e. the set of classes in which the classification is performed, is a code book, and M is the size of the code book. In order to construct such a code book it is necessary to divide the considered part of the N-dimensional euclidean space, where the random vector $\mathbf{x}$ is situated, into M zones and to adjoint a typical template vector $\mathbf{y}_m$ to each zone.

The ability of the model of the neural network together with its formator to classify the input vectors $\mathbf{x}$ into a priori chosen M typical classes, the ensemble of which is $\mathbf{Y}$, is given essentially by the fact that a priori (and this in the second-level formator in Fig. 2) local minima of the global energy function of this neural network (generally the stable equilibrium points of this network) have been selected as the individual $\mathbf{y}_m$, i.e. as the individual templates, and that the neural network has been set by the acting variables of the first-level formator in conformity with this selection and this e.g. by the adjustment of the weights between the nonlinear elements of the network (possibly also by control of their internal thresholds);

B) their use in Neurocomputers (R. Hecht-Nielsen) and in neural computations of some problems, e.g. in linear programming, in combinatorial optimization etc.;

C) their use in robotics, e.g., in sensory-motor control. Here the perceptron of J. C. Albus called Cerebellar Model Function Controller, a system which was able to control a robot arm with seven degrees of freedom, is to be quoted particularly;

D) their use in specific information treatment systems, as the "Wisard" (I. Aleksander) for face recognition, the "Neural Phonetic Typewriter" (T. Kohonen), the "Neocognitron" (K. Fukushima), the "Net Talk" (C. R. Rosenberg and T. Sejnowski) and o.m.;

E) their use in formators and for the formator control of large scale and complex systems, especially of networks of interconnected centres (see Section 12).

There are many other important application areas, some of which are not yet well defined (creative thinking, extension of human faculties, etc.).

12. NEURAL NETWORKS AS PARTS OF FORMATORS

Even with cautious evaluation of the functions which may fulfil neural networks, it may be stated that some of them have proved to be quick and effective parts of pattern classifiers. Such neural networks could form an essential part of the formator, mainly of its three types of analyzers ANX, ANR, ANY (of the input, control and output variables) and also of its control element CE. See the scheme

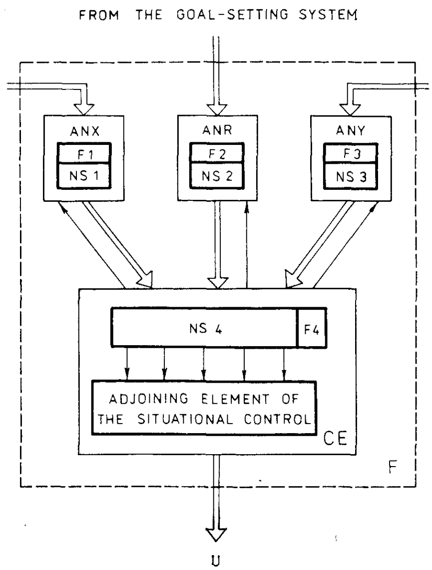

Fig. 3. Neural networks and their formators used as constituent parts of a formator.

of the formator in Fig. 3. The neural network NS 1 of its analyzer of input variables ANX together with its formator F1 is set as a classifier into M_1 classes, similarly NS2 together with its formator F2 is set to classify into M_2 classes, similarly NS3 together with its formator F3 is set to classify into M_3 classes. The Control Element CE is provided with a neural network NS4 with its own formator. The number of input variables of this network is $M_1 + M_2 + M_3$. This network with its formator F4 is a classifier into M_4 classes. To each of these output classes, each of which expresses a specific class of situations, will then be connected, by an adjoining (associative) element, according to the method of situational control, a specific decision in the form of control, i.e. of the vector $\boldsymbol{U}$ of the acting variables of the formator. This decision is a beforehead prepared alternative of control, which has been worked out e.g. on the basis of previous experience, possibly with the use of an expert system and which suits best the given class of situations (the so-called macrosituation).

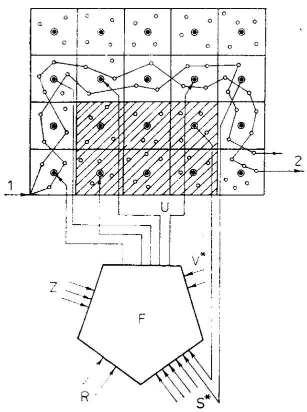

Fig. 4. Routing in a complex of zones with randomly distributed mobile stations.

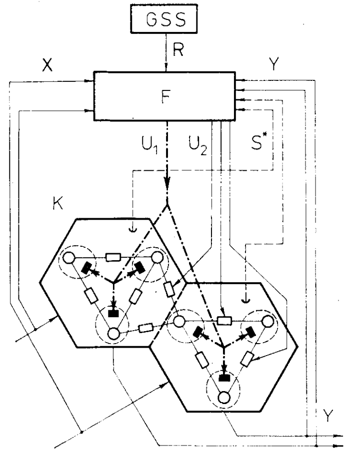

Fig. 5. Formator control of the state and interconnection of zonal elements in a geographical Complex.

The use of controlled neural nets in the formator (Fig. 3) shows a trend towards the realization of very effective formators, e.g. for the control of large and complex systems. A potential application of such a formator is in Fig. 4 for the routing of flows of signals in a complex by means of zonal radio stations (double circles) which are in contact with the randomly spread mobile elements. Another application is in Fig. 5 where the formator F acts upon the complex K by applying two sets of acting variables: U_1 acting upon the states of the zonal subelements (the full rectangles are again formators) and U_2 acting upon the interconnections between the elements. This advanced scheme aims at a synthesis of cooperative, evolutionary and situational control of a large-scale system.

13. RELATION TO CELLULAR AUTOMATA

A possible trend in future research could aim at bringing nearer together the theory of neural networks and that of cellular automata.

Technological requirements lead naturally to the elimination of one of the distinctions: neural network models will have elements located on a regular lattice.

Another distinction concerning the states of the neighbors which determine the state of a cell can be lessened by a generalization of the cellular automaton, introducing instead of the cellular space a demonstration space [14]. Here, a newly introduced concept is also the neighbourhood space.

In the classical theory of cellular automata the neighborhood index of one cell N remains independent upon the location of the cell-automaton and remains constant during its evolution. On the contrary, in the case of the mentioned generalization, the neighbourhood of each cell in the demonstration space is a function of its instantaneous configuration and changes with each step. The neighbourhood space is here the set of all possible links between the elements of the demonstration space.

Aiming at respecting random phenomena, a statistical mechanics of cellular automata as models of self-organization has been developed (S. Wolfram, 1983).

14. CONCLUSION

Here we have considered neural network models as parts of Cybernetic Systems. Connected with them are prospective research areas and problems, e.g.:

A) Further elaboration and application of the theory of attractors in nonlinear dynamic models, specifically to neural networks.

B) The use of information-theoretical approach to design considerations of the pertinent Cybernetic systems (Fig. 1), involving the choice of different thresholds of resolution, marked by ε in Fig. 6 [15], [16]. Particularly, applying the so-called Kolmogorov-Tichomirov inequality, following conditions must be fulfilled. (G. M.

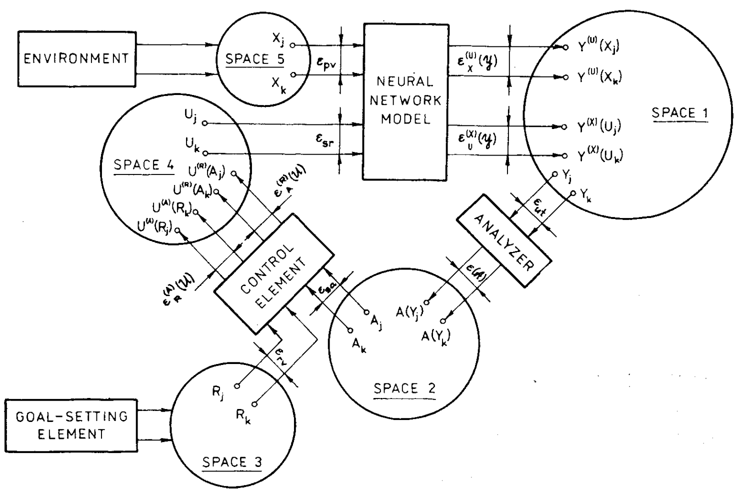

Fig. 6. The thresholds of resolution in a simplified Cybernetic system including the neural network model.

Ulanov, 1970): $H_{\varepsilon_S}(\mathscr{R}) < \log_2 J < L_{\varepsilon_{rv}}(\mathscr{R})$, where $\mathscr{R}$ is the set of command variables, J is the number of states of the Complex, ε_S is the maximum admissible deviation between **R** and **Y**, ε_{rv} is the threshold of resolution of the states of the Command Element by the Control Element through the command variables. The dimensions of the Euclidean spaces 1 and 3 are assumed equal. H_{ε_S} stands for minimum ε_S-entropy and $L_{\varepsilon_{rv}}$ is the ε_{rv}-capacity of the set $\mathscr{R}$. An important design indication is $\varepsilon_{rv} = \varepsilon_{ut}$ which is the utilitary threshold of resolution of the states of the Complex by the Analyzer through the measured variables of the Complex.

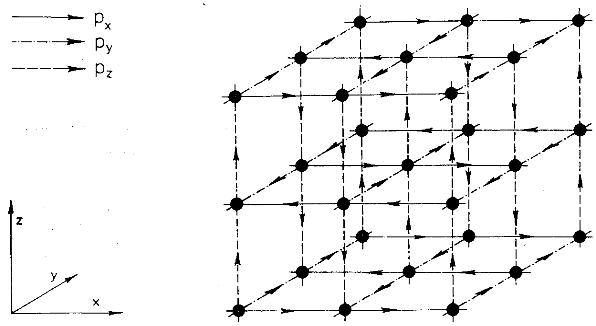

Fig. 7. A specific three-dimensional lattice of interconnected (non-neuronal) elements.

C) The routing problem in neural network models. Even if the processes in a neural network have a distributed character it is profitable to have some knowledge about the oriented spread of impulses in neural networks. This certainly needs simulation studies on such networks. In Fig. 7 a three-dimensional symmetrical lattice of simple (non-neuronal) elements, deduced from the crystallographic group $I\bar{4}3m$, belonging to the Cubic System has proved under the conditions given in [13] particularly apt to the oriented propagation of impulses governed by 3 probabilities p_x, p_y, p_z indicated. It would be of interest to investigate such stochastic spread of impulses through different types of networks but with nonlinear elements (neurons).

D) The problem of the complexity of the Elements of the neural network. The Elements used in the models are generally very simplified in comparison with the wealth of physiological knowledge e.g. about the adaptivity of neurons (A. A. Frolov, I. P. Muravyov, 1987). One possible way would therefore be to start with models of neurons reflecting even their subcellular constitution (some of the subcellular mechanisms of control and homeostasis could again be represented by Cybernetic Systems) and to study then their ensembles using simulation on networks of processors or appropriate neurocomputers.

E) Further research directed to the extension of generalized cellular automata theory (see Section 13) to the modelling of neural networks, respecting their complexity (see point D above) seems to be rationally indicated.

Neural networks have been objects of research for decades. There is now a renaissance of this field, bringing new incentives for interdisciplinary and world wide research, free from prejudice. These network models are constituents of Cybernetic systems. This leads us to remember here the 25th anniversary of the death of Norbert Wiener (26. 11. 1984 – 18. 3. 1964), the founder of Cybernetics, who has himself much contributed to this field.

(Received September 14, 1989.)

REFERENCES

[1] B. G. Farley and W. A. Clark: Simulation of self-organizing systems by digital computer. Trans. IRE *PGIT-4* (1954), 76—84.

[2] J. Faber: Modelling of control systems: Vigilance, sleep and epilepsy. In: Fourth Formator Symp. on Mathematical Methods for the Analysis of Large Scale Systems (J. Beneš, L. Bakule, eds.), Academia, Prague 1983, pp. 181—205.

[3] J. J. Hopfield: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA *79* (1982), 2554—2558.

[4] H. Sompolinsky: The theory of neural networks: The Hebb rule and beyond. In: Heidelberg Colloquium on Glassy Dynamics, 1986 (J. L. v. Hemmen, I. Morgenstern, eds., Lecture Notes in Physics 275), Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo 1987, pp. 485—527.

[5] D. E. Rumelhart, G. E. Hinton and R. J. Williams: Learning representations by back-propagating errors. Nature *323* (1986), 533—536.

[6] G. A. Carpenter and S. Grossberg: The Art of adaptive pattern recognition by a self-organizing neural network. Computer *21* (1988), 3, 77—88.

[7] M. Peschel, H. M. Voigt, W. Mende, J. Wolf, H. Herzel and F. Breitenecker: Chaotische Schalter — Ein neuer Zugang zur Modellierung von Neuronennetzen. Messen-steuern-regeln *31* (1988), 1, 36—38.

[8] R. P. Lippmann: An introduction to computing with neural nets. IEEE Acoust. Sign. Speech Processing Magazine (1987), April, 4—22.

[9] P. Peretto: Collective properties of neural networks: A statistical physics approach. Biol. Cybernet. *50* (1984), 51—62.

[10] G. Toulouse, S. Dehaene and J.-P. Changeux: Spin glass models of learning by selection. Proc. Natl. Acad. Sci. USA *83* (1986), 1695—1698.

[11] L. D. Jackel, H. P. Graf and R. E. Howard: Electronic neural network chips. Appl. Optics *26* (1987), 23, 5077—5080.

[12] A. A. Vedenov, A. A. Ježov, L. A. Knižnikova, E. B. Levčenko and Ju. G. Černov: Nonlinear systems with memory and the modeling of the functions of neuronal assemblies (in Russian). In: Intellektualnyie processy i ikh modelirovanie (E. P. Velichov, A. V. Černavskij, eds.), Nauka, Moscow 1987, pp. 229—248.

[13] A. Guez, V. Protopopsecu and J. Barhen: On the stability, storage capacity, and design of nonlinear continuous neural networks. IEEE Trans. Systems Man Cybernet. *18* (1988), 1, 80—87.

[14] J. Beneš and P. Kolář: Automatic configuration control for goal-oriented roving complex. In: Preprints VIIth IFAC Symp. on Automatic Control in Space. VDI/VDE-Ges. Mess- und Regelungstechnik, Düsseldorf 1976, pp. 739—752.

[15] J. Beneš: The concept of adaptive thresholds of resolution. In: Second Formator Symposium on Mathematical Methods for the Analysis of Large Scale Systems (J. Beneš, L. Bakule, K. Sladký, eds.), Academia, Prague 1975, pp. 61—89.

[16] E. C. Posner and E. R. Rodemich: Differential entropy and tiling. J. Statist. Phys. *1* (1969), 1, 57—69.

[17] J. Beneš and J. Grim: Automatic stochastic control of impulses on a three-dimensional crystallographic lattice. Kybernetika *12* (1976), 5, 355—362.
[18] J. Beneš: On systems with automatic control of configuration. In: Automatic and Remote Control. Proc. Second Congress IFAC, Basle, Switzerland, 1963, Butterworths, London—Oldenbourg, Munich—Vienna 1964, Vol. Theory, pp. 656—663.

Doc. Ing. Dr. techn. Jiří Beneš, DrSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.