From cognitive models to neurofuzzy systems - the mind space approach.

Włodzisław Duch
Department of Computer Methods, Nicholas Copernicus University,
ul. Grudziądzka 5, 87–100 Toruń, Poland.
e-mail: duch @ phys.uni.torun.pl
WWW: http://www.phys.uni.torun.pl/kmk/

Phenomenological theory based on simple concepts related to human cognition is introduced. These concepts are directly related to neurophysiological brain events but may also be extended to explain higher cognitive functions realized by the mind. This theory on the one hand solves fundamental problems in cognitive sciences, explaining puzzling behavior of human conscious experience, and on the other hand leads to useful models of mind in form of neurofuzzy systems. Such systems can compete in pattern recognition and classification tasks with neural networks and can compete in reasoning tasks with the expert systems.

I. INTRODUCTION

There are two distinct approaches to understanding of human intelligence and human mind. Artificial intelligence aims at building intelligent systems starting from the processing of symbols [1]. There are serious problems at the very foundation of such an approach, starting with the famous mind-body problem (how can the non-material mind interact with matter), the symbol grounding problem (how can the meaning be defined in a self-referential symbolic system) or the frame problem (catastrophic breakdowns of intelligent behavior for “obvious” tasks) [2].

Neural networks bring new possibilities to the field of artificial intelligence. So far neural networks seem to be suited best for the low-level cognitive tasks, such as vision or auditory processes, or for simple classification tasks, while they are somehow restricted in their abilities to realize predefined knowledge structures and in using such structures in sequential reasoning processes. There is no doubt that higher cognitive functions result from the activity of the brain and thus should be implementable by neural networks [3]. It is clear that the present shortcomings of neural networks are connected with the lack of modularity and low complexity of the models rather then with the inherent limitations of the neural modeling itself. Much is know about the details of neural processes responsible for brain functions and neurodynamics [4] as well as the computational cognitive neurosciences are thriving fields [5]. Recently even developmental psychology has been influenced by neurodynamics [6].

Can we understand higher mental activity directly in terms of neural processes in the brain ? It does not seem likely. Even in chemistry and physics phenomenological concepts that are not easily reducible to fundamental interactions are still used, in fact experimental chemists have hardly been affected by the development of quantum chemistry. Macroscopical theories are reducible only in principle to microscopical descriptions, in practice phenomenological approach to complex systems is most fruitful. Language of neuroscience and language of psychology is quite different. Since the brain is very complex intermediate theories, between neural and mental, physical and symbolic, are needed. An example of such a theory is sketched in this paper.

II. COGNITIVE MODELING

Cognitive modeling approach lies between the symbolic, rule-based methods of artificial intelligence, and distributed, associative processing of neural networks, combining the best of both worlds. The goal of this approach is to:

1) Create precise mathematical language directly describing cognitive states (mental events). Basic concepts of this language should be constrained by neurobiology and should be reducible, at least in principle, to neurodynamics of the brain.

2) Use this language to build a theory of cognitive systems.

3) Apply this theory to explain features of human cognitive processes, such as identification, association, generalization, reasoning, various states of mind, empirical facts related to consciousness.

4) Construct adaptive systems according to specifications, systems that will recognize the incoming signals, categorize, learn from examples and from general laws, self-organize, reason, use natural language and perform other cognitive functions.

Attractor neural networks [7] and other neurodynamical models offer good models of the brain's activity and should be used to understand basic mental events. Approximations and simplifications of such models are necessary to understand higher-order cognition. The low level cognitive processes, realized mostly by various topographical maps, define features of internal representations (some of which are hidden from the external world) [3]. These features represent many types of data: analog sensory signals, linguistic variables, numbers, visual images. Real mind objects are composed primarily of preprocessed sensory data, iconic representations, perception-action multidimensional objects. They seem to correspond to a stable attractors of global dynamic of the brain, realized by the transcortical neural cell assemblies (TNCAs).

The concept of neural cell assemblies was introduced already in 1949 by Donald Hebb in his seminal book [8]. The cerebral cortex has indeed a very modular structure [9,10]. Macrocolumns, distinguishable using neuroanatomical techniques, contain between $10^4 - 10^5$ neurons in a $1-2$ mm high column spanning six layers of the neocortex, within the cortical area of a fraction of mm^2. Axons of some NCA neurons spread horizontally on several millimeters enabling mutual excitation of different NCAs. Within a macrocolumn one may distinguish minicolumns, much smaller functional groups of neurons with inhibitory connections. They have a diameter of 30 μm and only 110 neurons, except in the primary visual cortex (V1 area), where they contain about twice as many neurons in orientation columns. These minicolumns behave as oscillators and recurrent excitations of such oscillators leads to entrainment and synchronization of neuronal pulses [11]. Vertical connections inside these minicolumns are largely excitatory and the density of these connections is of an order of magnitude higher than of the connections with neurons outside of the column.

There is a fairly good evidence that it is the temporal synchronization of a number of neural cell assemblies, corresponding to an attractor in the global dynamics of the brain, that corresponds to a perception or a motoric action. Although such NCAs play important role in the brain models description of their functions requires rather complex dynamical models.

What we are really interested in is the classification of the existing attractor states, transition probabilities between them and formation of the new ones during learning. A drastic, although quite natural simplification of the neurodynamics, leads to a discreet finite automata (DFA), such as the hidden Markov models [12]. Here I will present a different approach, linking directly neurodynamics with the psychological theory of feature (conceptual) spaces. Feature spaces [13] are used in cognitive psychology to discuss mental representation. Mind space approach to cognition [14] is based on the assumption that the objects in the feature spaces correspond directly to the attractors of the global dynamics of the brain.

Imagine [14] a coordinate system based on the features of mental representations obtained from lower-level modules, such as topographical maps or computational maps based on population coding. This coordinate system defines a multidimensional space, called here "the mind space", serving as an arena on which mind events take place. In this space a "mind function" is defined, describing the "mind objects" as a fuzzy areas where the mind function has nonzero values. Each of these mind objects corresponds to an attractor of the brain's neurodynamics. There are two time scales in this model. Changes in the features of internal representations themselves are slow and depend on the plasticity of the brain, i.e. real physical changes during the learning processes. At the mind space level these changes should be reflected in the number of dimensions, the character of the axis and the topology of the space. In the first approximation I will neglect all such changes.

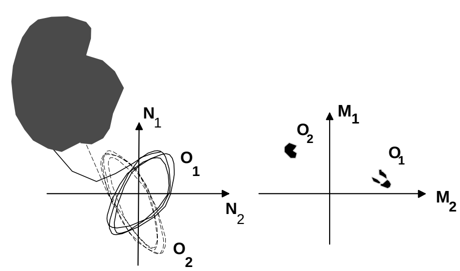

FIG. 1. Relation between attractors representing correlation of the spiking activity of a group of neurons (here just two) and objects in the feature space.

Faster time scale is connected with the actual dynamics of activation of mental representations, trains of thoughts or perceptions, going from one mind object to another. At a given moment combination of features obtained from lower-level processing modules activates only one (or in large systems only a few) attractors in the neurodynamics. This corresponds to certain “mind state” given usually by a fuzzy point (since the corresponding neurodynamics is always to some degree chaotic) in the mind space. Mind objects in the region pointed at by the mind state are “activated” or “recognized”. Evolution of the mind state is equivalent to a series of activations of objects in the mind space (a searchlight beam lighting successive objects is a good metaphore here). To model a real mind corresponding to the real brain one would have to know all possible attractors of the brain’s dynamics, clearly an impossible tasks. The number of required dimensions would also be very large, one could even consider the possibility of continuous number of dimensions [15]. Experimental techniques of cognitive psychology, such as probing the immediate associations and measuring the response time give enough information to place basic mind objects corresponding to some concepts or perceptions in the mind space. In the simplified version of the model mind objects are created and positioned using unsupervised as well as supervised methods of learning, similar to the learning vector quantization [16] or other local learning techniques [17].

The idea of a “mind space” or “conceptual space” is not more metaphorical than such concept of physics as space-time or a state of the system described by a wavefunction. A proper mathematical description of the mind space is very difficult because of high dimensionality of this space and of a complicated metric that has a non-Euclidean character. Human conceptual space seems to be better approximated by a number of linked mental spaces rather than one universal space. Nevertheless, simple approximations are quite useful in many situations. Associations among mind objects, coresponding to the transition probabilities between different attractors, should be based on the distance between mind objects and should take into account not only the features of representations but also the spatio/temporal correlations. “Intuition” is based on the topography of the mind space. Instead of a logical reasoning dynamical evolution of the mind state (activation of a series of mind objects) is considered in this approach. Logical and rule-based reasoning is only an approximation to the dynamics of the state of mind.

Mind space is used as a container of the mind objects, including memories reflecting the states of a total system (i.e. states of an organism in biological terms). A natural practical realization of this idea is obtained by modular neural networks, with nodes specializing in description of groups of objects in the mind space (coded in different attractor states of the same transcortical neural cell assembly). The function of each node of the network is an approximation to the activity of an attractor neural network, or a fragment of the neurocortex that responds to stimulations by stable reverberations of persistent spiking activity. Nodes of such network do not represent single neurons but rather averaged activity of neural cell assemblies. Such network may be considered from two points of view: as a neural network based on localized processing functions or as a fuzzy expert system based on representation of knowledge by fuzzy sets.

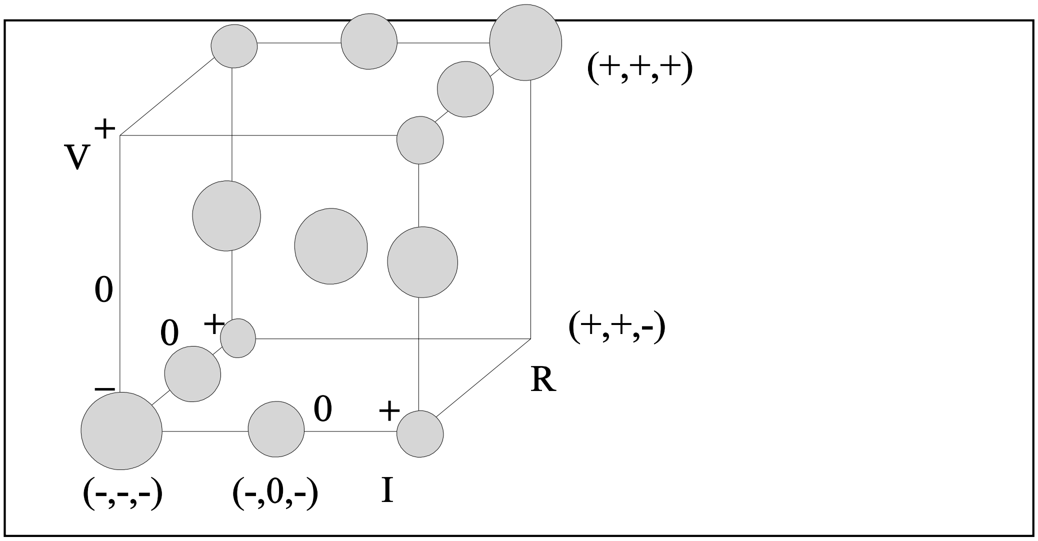

FIG. 2. Representation of the Ohm's law $V = I * R$ in the mind space model. The axes illustrate only one feature of variables, their change: $-$ for decreasing, 0 for constant and + for increasing. This representation of a small subspace of the mind space is created in an unsupervised way from examples or directly from the corresponding rules. Such knowledge representation is very effective in the reasoning process, for example in qualitative analysis of electrical circuits.

It is useful to discriminate between the static and the dynamic cognitive functions. Static functions are related to the knowledge that is readily available, intuitive, used in recognition and immediate evaluation. Dynamic functions of mind are used in reasoning and problem solving. The mind space approach is sufficient to describe the static aspects of human cognition. How well can the dynamical aspects of human thinking and problem solving be modeled using such systems? There are some reasons to be optimistic even in this case. Systems based on the concept of mind space try to avoid full description of the underlying dynamical brain processes that can be properly modeled only in a huge phase space. Transition probabilities between attractors in dynamical systems are approximated by the overlaps of the mind objects representing these attractors in the mind space. Adding hidden dimensions (corresponding to internal features that influence the dynamics but are not accessible through inputs or outputs of the system) allows to model arbitrary transition probabilities (associations of mind objects). Such problem solving tasks as playing chess seem to be based on a large memory (large number of mind objects) and on a memory-based reasoning [18] with a rather limited exploration of the search space. Memory-based reasoning is related to probabilistic neural networks and outperforms in many cases other learning methods, including neural networks [18]. It is not clear how much human thinking is dominated by learned skills; transfer of general thinking skills seems to be an illusion and some experts even ask if humans are rational at all [19].

Symbolic approach to dynamics, although a drastic simplification, gives very interesting results even for chaotical systems [20], therefore mind space approximation to neurodynamics should also be fruitful. In this paper only the static functions of the mind are modeled. The following sections present a particular realization of the ideas described above, in a form of a modular system that may be applied to a variety of cognitive tasks.

III. FEATURE SPACE MAPPING SYSTEM.

Feature Space Mapping (FSM) network [21,22] has some unique properties, rather different from those of most artificial neural network models. The network receives pre-processed multidimensional input data. The goal is to create a feature space representation of the incoming data, therefore flexible functions with small number of adaptive parameters should be used as node transfer functions. FSM uses separable processing functions for localized description of fuzzy data in the mind space (in this section the mind space and the feature space are synonymous). In the special case when gaussian processing functions are used by the network nodes (gaussians are the only radial basis functions that are separable [23]) this model belongs to the family of the growing and shrinking Hyper Basis Function (HBF) networks. Localized processing functions representing the mind objects are initially centered on the data vectors $\mathbf{D}=(D_1, D_2, ...D_N)$ with dispersion (fuzziness) of each component proportional to the error or uncertainty of the variables D_i. Many types of separable functions may be used by the nodes of FSM system, including localized products of pairs of sigmoidal functions that for N-dimensions have the form:

$$s(\mathbf{X};\mathbf{D},\Delta) = \prod_{i=1}^{N} \sigma(X_i - D_i + \Delta_i/2)(1 - \sigma(X_i - D_i - \Delta_i/2)) \tag{1}$$

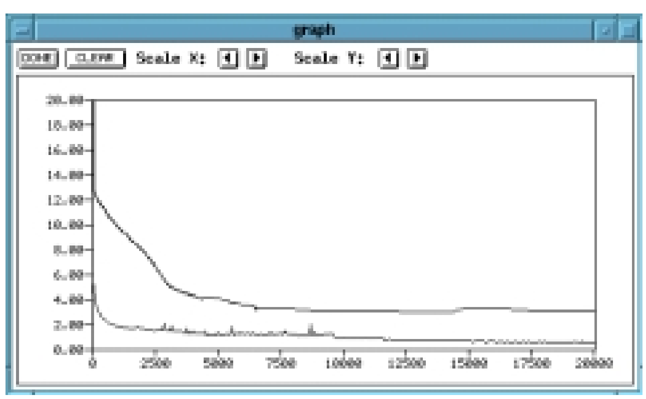

FIG. 3. Convergence rate of the RBF (upper curve) and FSM (lower curve) on the two-spiral benchmark with 100 nodes.

These functions are more flexible than gaussian functions in description of multidimensional densities of arbitrary shapes. Each variable X_i defines a new dimension, the data vector

$\mathbf{X}$ is a point in the feature space and the input data vector together with the associated uncertainties of the inputs defines a fuzzy region in the feature space, described by the values of the $s(\mathbf{X};\mathbf{D},\Delta)$ function. The mind function M for a collection of mind objects $\mathbf{D} = \{D^p\}$ has the following general form:

$$M(\mathbf{X},\mathbf{D},\Delta) = \sum_p W_p s(\mathbf{X};\mathbf{D^p},\Delta^p) = \sum_p W_p \prod_i g\left(X_i; D_i^p, \Delta_i^p\right) \tag{2}$$

and does not vanish only around the data vectors $\mathbf{D}$ stored in the M function. The weights W and the dispersions Δ are the adaptive parameters defining the mind function for a given set of input values. If the input data samples are noisy or they are too numerous to store them the centers $\mathbf{D}$ are also treated as adaptive parameters, as it is done in many clustering algorithms such as the LVQ model [16]. In the learning process the shapes of the mind objects and their mutual positions are adjusted by local learning procedures reflecting the structure of the incoming data. This stage is quite similar to the learning in the Hyper Basis Functions networks [23], RAN networks [24] or vector quantization methods [16]. Initial value of the adaptive parameters is obtained from the k-nearest neighbor heuristics or from the information about an intrinsic scale and uncertainty of the input data. However, the structure of FSM network differs in several respects from the structure of HBF or RAN networks.

Functions processed by different nodes of FSM network may be different while in RBF, HBF or RAN networks they are of the same type. In FSM inputs $\mathbf{X}$ and outputs $\mathbf{Y}$ form one mind object, therefore they are treated on an equal footing. HBF approximation of one dimensional function is given by $\mathbf{Y} = HBF(\mathbf{X})$; in FSM this relation is always fuzzy and the most probable function is obtained from maximization:

$$\max_Y M(\mathbf{X},\mathbf{Y}) \Leftrightarrow \mathbf{Y} = FSM(\mathbf{X}) \tag{3}$$

Thus for a given value of $\mathbf{X}$ a whole range of probable $\mathbf{Y}$ values is obtained, close to the maximum of the M function, giving an estimation of the uncertainty of the unswer. FSM network (Fig. 4) has two outputs, one giving the value of the M function, and another giving the value of the gradient of M. These values are used to find the local maximum in the mind space by changing the inputs along the direction of the gradient. The network reaches a stable state when a local maximum is found, there is no requirement for a global extremum points as in the backpropagation networks. FSM may be regarded as a special kind of a recurrent network in which output is connected to input and all positions and sizes of basins of attractors are explicitly defined.

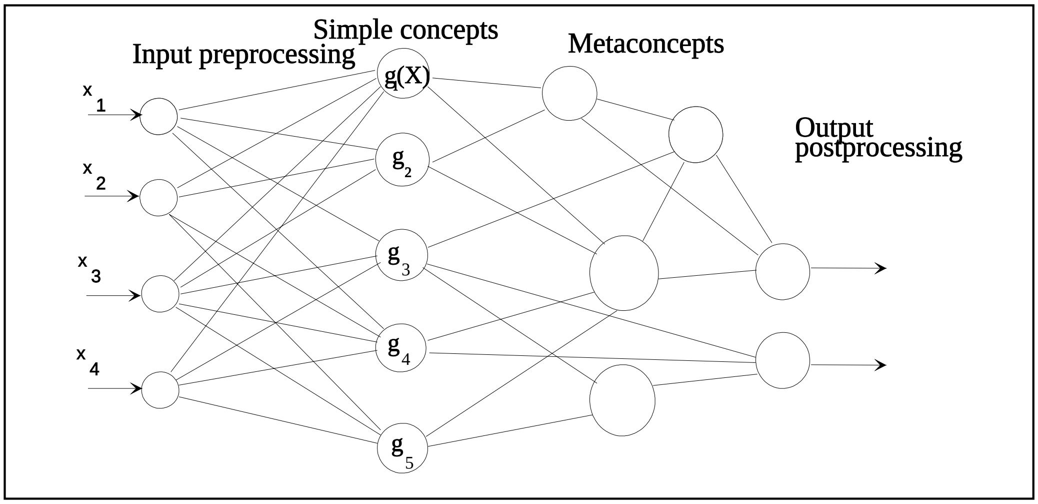

FIG. 4. Structure of the FSM network.

After the initial nodes of the network are established (using the "best prototypes" or preselected input data) on-line learning is performed, with the new data patterns constantly presented to the system. The problem may be stated in the following way: given the approximating function $F^{(n-1)}$ realized by the adaptive system and the new data $(\mathbf{X}^n, \mathbf{Y}^{(n)})$, find the best new estimate $F^{(n)}$. Parameters of the existing nodes are changed to take account of the new data and new nodes are added only if:

$$\min_k \left\| \mathbf{X}^{(n)} - \mathbf{D}^{(k)} \right\| > d_{\min};\ Y_i^{(n)} - F_W^{(n-1)}(\mathbf{X}_n; \mathbf{D}, \Delta)_i > \varepsilon_i \tag{4}$$

Here d_{min} is the resolution of the data in the input space while ε_i is the maximum acceptable error of the i-th output component. Initial values for the dispersions Δ are frequently based on the nearest neighbor heuristic. If both criteria given above are not satisfied for the new input data gradient adaptation of the weights, centers and dispersions of the node functions is performed. Only the local gradient estimation is used here for the $(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})$ data (as is also done in RAN and in the function estimation approach [24]). The weights are changed according to:

$$\mathbf{W} \leftarrow \mathbf{W} + \eta \left(Y^{(n)} - F_W^{(n-1)}(\mathbf{X}_n; \mathbf{D}, \Delta) \right) \times \nabla_{W,D} F_W^{(n-1)}(\mathbf{X_n}; \mathbf{D}, \Delta) \tag{5}$$

where η is the adaptation step size and only one-dimensional output is assumed for simplicity. The dispersions of the node functions should be rather large to obtain a smooth approximating function and to avoid overfitting of a noisy data. If the new node is not needed positions of the maxima in the mind space are changed according to:

$$\mathbf{D} \leftarrow \mathbf{D} + \eta_d \left(\mathbf{X} - \mathbf{D} \right) \tag{6}$$

This solution leads to self-organization of data clusters in the feature space reflecting the probability distribution of the incoming data. A small change in the dispersions is also performed. From the formal point of view equations for learning procedure may be derived from regularization theory [23] using tensor product stabilizers. The FSM adaptive system with N nodes tries to minimize a local error function

$$E[M_W] = \sum_{i=1}^{N} \sum_{j \in O\left(D^{(i)}\right)} K_i \left(X^{(j)} - D^{(i)}\right) (Y^{(j)} - M_W(X^{(j)}))^2 \tag{7}$$

where the kernel functions K_i and the neighborhood definitions $O(D^{(i)})$ depend on the problem while W symbolize all adaptive parameters. This error function may also include a proper stabilizer although in practice it is smipler to add noise to the input data to get smoother approximations.

Representation of data by fuzzy regions of high density in the mind space make the FSM system equivalent to a fuzzy expert system. The rules of the fuzzy expert systems are of the following type:

$$\begin{aligned} &\mathrm{IF}(x_1 \in X_1 \wedge x_2 \in X_2 \wedge \ldots x_N \in X_N) \\ &\qquad \mathrm{THEN}\,(y_1 \in Y_1 \wedge y_2 \in Y_2 \wedge \ldots y_M \in Y_N) \end{aligned} \tag{8}$$

The rules in fuzzy expert systems are unique, i.e. the same IF part should not have a few different THEN parts. These rules may be directly programmed in the FSM network if many outputs from network nodes are allowed. More general rules of the type

$$\begin{aligned} &\mathrm{IF}\left(x_1 \in X_1^{(1)} \wedge \ldots x_N \in X_N^{(1)}\right) \vee \left(x_1 \in X_1^{(2)} \wedge \ldots x_N \in X_N^{(2)}\right) \vee (\ldots) \\ &\qquad\qquad\qquad \mathrm{THEN}\,(y_1 \in Y_1 \ldots \wedge y_M \in Y_M) \end{aligned} \tag{9}$$

may also be used in the FSM system. Therefore queries addressed to the system may contain logical operators that are used to restrict the search in the mind space. FSM system may contain a number of recognition, or identification, modules and some modules implementing logical operations (Fig.5). This design corresponds to a number of separate mind spaces, each with its own coordinate system based on the unique input features. The results of identifications in these mind spaces are integrated in spaces that are higher in the hierarchy.

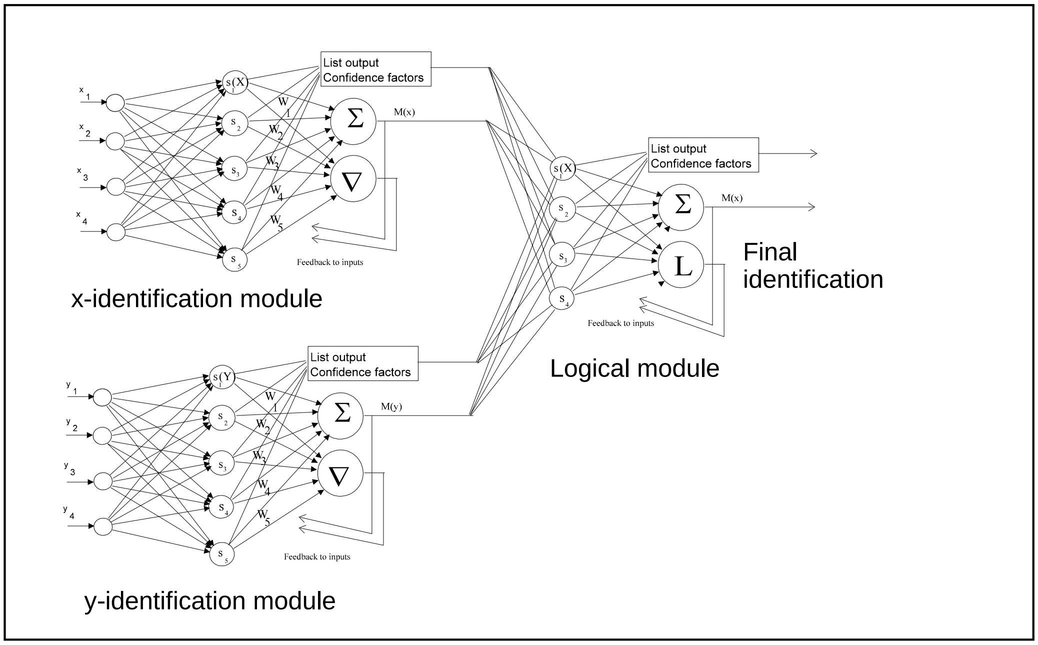

FIG. 5. Two recognition and one logical module of FSM.

To reduce the complexity of search in highly dimensional mind spaces a technique based on dynamical scaling is used. If gradients of the M-function at point $\mathbf{X}$ are small, making the nearest mind object hard to find, fuzziness of all mind objects is temporarily increased at the beginning of the search, leaving only a rough representation of mind objects. This corresponds to the initial orientation step in human information processing, determining what the problem is about. After the local maximum of the M-function is found the FSM system focuses on the problem by changing the fuzziness of all objects to standard values and performing more detailed search. Several answers may be found by switching off temporarily the mind objects corresponding to the solutions found so far and repeating the search procedure. In addition local two-dimensional maps of the mind space objects around the solution found help to visualize the multidimensional relations among mind objects. These maps are obtained by minimization of the measure of topography preservation, reuivalent to the multidimensional scaling procedure used in mathematical psychology [25].

IV. APPLICATIONS

FSM system, described above as an example of application of the cognitive modeling approach, is a universal neurofuzzy system based on the concept of the mind space. It may be used in all neural networks and expert systems types of applications. Among applications of this system currently pursued by our group [22] one should mention:

Classification of stellar spectra: modern telescopes, including Hubble Space Telescope, produce large amounts of stellar spectra. Classification of these spectra is still usually done by experts, although first applications of patter recognition and neural networks classifiers have been reported [26]. In this case the main problem is with the quality of data for training since databases contain spectra that need special treatment to be useful. The data may be presented in the form of histograms, with error bars for each value of the histogram, and transformed via Fourier or Hadamard procedure to a set of a few hundred numbers (this is also the dimension of the feature space used). The main purpose of this classification is to find unusual spectra for further processing.

Classification of chemical spectra: a large database of chemical spectra contains 25.000 infrared spectra and many other types of spectra. Similar normalization procedure as for the stellar spectra is used. The system should find the name of the molecule if its spectrum was contained in the training set. It also should analyze more complex spectra, finding those that correspond to molecular fragments contained in the target molecule, performing deconvolution of the given spectrum into the component spectra and finally simulating the given spectrum using these components.

More sophisticated applications include:

Testing theories about human intuition by measuring the length of time for the correct response and analyzing the errors that students make in problems involving qualitative physics. It is possible to recreate proximity data from the cognitive psychology questionnaires training the FSM system using the same type of contextual descriptors as are given to humans and mapping the multidimensional feature space objects on two-dimensional maps.

Classification of personality types using raw as well as pre-processed data from personality inventories such as MMPI (more than 500 questions with five possible answers each).

We are also considering application of FSM in natural language analysis. The idea of mental spaces [27] has proved to be very fruitful in the study of reference problems in linguistics, although mental spaces were so far constructed as ordered sets and relations.

V. SUMMARY

Cognitive modeling approach is quite fruitful not only for understanding of the human mind but also as an approach to design practical systems for technical applications. Attractive features of the FSM system include:

- Direct modeling of knowledge represented in the mind space by the fuzzy multidimensional objects. Mind objects in the feature space, together with their mutual relations, are a new form of knowledge representation.
- Symbolic interpretation of the individual processing nodes of FSM network is usually possible. Neural realization in form of adaptable network allows for unsupervised creation of nodes (learning) from examples. These nodes may represent attractor dynamics of transcortical neural cell assemblies.
- Adjusting overlaps and fuzziness of mind objects allows for full control over associations between these objects (clusters or knowledge items) and control over generalization by metanodes.
- Supervised and unsupervised learning methods for self-organization of mind space objects are used. Learning may proceed from examples, as in neural networks, or from general laws, as in expert systems.

- Implementation of a typical expert system production rules is straightforward; rules have the form:

  $IF\,(FACT_1.and.FACT_2.or.FACT_3...)$ than $(FACT_N)$

- Reasoning may take form of one-dimensional searches (if separable functions are used), focusing successively on single variables, with the depth of search equal to the number of unknown features. The system performs memory-based reasoning.

- Gradient or nearest neighbor techniques for finding associations between mind objects are used. Temporarily increasing the fuzziness of all mind objects helps to concentrate on relevant parts of the mind space (defocusing and focusing on the problem) by following the gradient to nearest interesting areas.

- FSM uses a constructive algorithm. Adding and removing mind objects (expanding and pruning the network) reduces the complexity of the model.

- Fine tuning of object representations allows to use FSM in pattern recognition and adaptive control tasks.

- Hierarchies of objects leading to categories and metaconcepts are formed spontaneously.

- The scaling of the complexity of the system is linear with the number of mind objects, making FSM ideal for parallel processing.

Acknowledgments
Support by the Polish Committee for Scientific Research, grant 8T11F 00308, is gratefully acknowledged.

[1] Newell A, Simon H. A. (1976) *Computer Science as empirical inquiry: symbols and search.* Communic. of the ACM 19: 113-126; A. Newell, *Unified theories of cognition.* (Harvard Univ. Press, Cambridge, MA 1990)

[2] Harnad, S. (1990) *The symbol grounding problem.* Physica D 42: 335-346; Harnad, S. (1993) *Problems, problems: the frame problem as a symptom of the symbol grounding problem.* PSYCOLOQUY 4 (34) frame-problem.11; Rakover, S.S. (1993). *Precise of Metapsychology: Missing Links in Behavior, Mind, and Science.* PSYCOLOQUY 4(55) metapsychology.1.rakover.

[3] P.S. Churchland, T.J. Sejnowski, *The computational brain* (MIT, Bradford Book 1992)

[4] N. Rashevsky, *Mathematical Biophysics* (Dover, NY 1960)

[5] M. S. Gazzaniga, ed. *The Cognitive Neurosciences* (MIT, Bradford Book 1995)

[6] E. Thelen, L.B. Smith, *A Dynamic Systems Approach to the Development of Cognition and Action* (MIT Press 1994)

[7] D.J. Amit, *Modeling brain function. The world of attractor neural networks* (Cambridge Univ. Press 1989); T.L.H. Watkin, A. Rau, M. Biehl, Rev. Modern Phys. 65 (1993) 499

[8] D. Hebb, *The Organization of Behavior* (J. Wiley, NY 1949)
[9] Y. Burnod, *An adaptive neural network: the cerebral cortex* (Prentice Hall 1990)
[10] W.H. Calvin, *Cortical columns, modules and Hebbian cell assemblies*, in: M. A. Arbib, Editor, *The Handbook of Brain Theory and Neural Networks* (MIT Press 1995), pp. 269-272
[11] W. Singer, *Synchronization of neuronal responses as a putative binding mechanism*, in: M. A. Arbib, Editor, *The Handbook of Brain Theory and Neural Networks* (MIT Press 1995), pp. 960-964
[12] M.P. Casey, *Computation in Discrete-Time Dynamical Systems* (PhD thesis, UCSD 1995, available in neuroprose).
[13] I. Roth, V. Bruce *Perception and Representation*, (Open University Press, 2n ed, 1995)
[14] W. Duch (1994), *Towards artificial minds*, Proc. of I National Conference on neural networks and applications, Kule, April 1994, pp. 17-28; W. Duch, *Transparent theory of consciousness - is there a problem?* Behavioral and Brain Sciences, 1995 (submitted); W. Duch (1994) UMK-KMK-TR 1/94 report, *A solution to the fundamental problems of cognitive sciences*, available from ftp.phys.uni.torun.pl/pub/papers/kmk
[15] B. MacLennan, *Field computation in the brain*, CS-92-174 (Univ. of Tennessee, Knoxville, TN 37996)
[16] T. Kohonen, *An Introduction to Neural Computing.* Neural Networks 1 (1988) 3-16; T. Kohonen, *Self-organization and Associative Memory* (Springer-Verlag 1984, 3rd edition: 1989); T. Kohonen, *Self-organizing Maps* (Springer-Verlag 1995).
[17] L. Bottou, V. Vapnik, *Local learning algorithms*, Neural Comput. 4 (1992) 888-901; V. Vapnik, L. Bottou, *Local Algorithms for Pattern Recognition and Dependencies Estimation*, Neural Comput, 1993, v.5, pp. 893-909
[18] D.L. Waltz, *Memory-based reasoning*, in: M. A. Arbib, Editor, *The Handbook of Brain Theory and Neural Networks* (MIT Press 1995), pp. 568-570
[19] A. Garnham and J. Oakhill, *Thinking and reasoning* (Oxford, Blackwell 1994)
[20] T. Bedford, M. Keane and C. Series, *Ergodic theory, symbolic dynamics and hyperbolic spaces* (Oxford University Press 1991)
[21] W. Duch, G.H.F. Diercksen, *Feature Space Mapping as a universal adaptive system*. Computer Physics Communications **87** (1995) 341-371; W. Duch, *Floating Gaussian Mapping: a new model of adaptive systems*, Neural Network World 4 (1994) 645-654
[22] W. Duch, R. Adamczak, N. Jankowski and A. Naud, *Feature Space Mapping: a neurofuzzy network for system identification*, Engineering Applications of Neural Networks, Helsinki 1995, pp. 221–224
[23] T. Poggio and F. Girosi, *Networks for approximation and learning*. Proc. of the IEEE 78 (1990) 1481; T. Poggio and F. Girosi, *A theory of networks for approximation and learning*. Center for Biological Information Processing (CBIP), Paper No.31, 1994; F.Girosi, M. Jones and T. Poggio, *Priors, stabilizers and basis functions: from regularization to radial, tensor and additive splines.* CBIP Paper No.75, 1994
[24] J. Platt, *A resource-allocating network for function interpolation.* Neural Comput, 1991, v.3, p. 213; V. Kadirkamanathan, M. Niranjan, *A function estimation approach to sequential learning with neural networks.* Neural Comput, 1993, v.5, p. 954
[25] W. Duch, *Quantitative measures for the Self-Organizing Topographic Maps*. Open Systems and Information Dynamics 3 (1994) 295-302
[26] O. Lahav, A. Naim, L. Sodre, M.C. Storrie-Lombardi, *Neural computations as a tool for galaxy classification: methods and examples* (Cambridge University, preprint CAP 9508002, August 1995)
[27] G. Fauconniere, *Mental Spaces* (Cambridge Univ. Press 1994)