System, Modelling, Control. Zakopane, Poland, May 1-5, 1995

FROM COGNITIVE MODELS TO NEUROFUZZY SYSTEMS.

Wlodzislaw Duch
Department of Computer Methods, Nicholas Copernicus University
Grudziadzka 5, 87-100 Toruñ, Poland
e-mail: duch@phys.uni.torun.pl

Abstract

Phenomenological theory of mind based on simple concepts related to human cognition is introduced. Basic concepts of this theory are directly related to neurophysiological events in the brain and 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 in reasoning tasks with expert systems.

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. There are serious problems at the very foundation of such an approach, starting with the famous mind-body problem (how can the 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). On the other hand there is no doubt that higher cognitive functions are a function of the brain activities and much is know about the details of neural processes responsible for these functions. Can we understand higher mental activity directly in terms of brain processes? It does not seem likely; even in chemistry and physics phenomenological concepts that are not easily reducible to fundamental interactions are still used. Macroscopical theories are reducible only in principle to microscopical descriptions, but in practice phenomenological approach to complex systems is most fruitful. Since the brain is very complex intermediate theories, between neural and mental, physical and symbolic, are needed. Such a theory is sketched in this paper.

COGNITIVE MODELING

Our approach [1-3] lies between the symbolic, rule-based methods of artificial intelligence, and distributed, associative processing of neural networks, combining best of both worlds. Our goal is to:

1) Create precise mathematical language describing cognitive states (mental events).

2) Use this language to derive general theory of cognitive systems.

3) Apply this theory to: a) explanation of human cognitive processes: identification, association, generalization, reasoning, various states of mind, empirical facts related to consciousness; b) construction of adaptive systems according to specifications, systems that will: recognize, categorize, learn from examples, self- organize, reason, use natural language ...

Attractor neural networks [4] offer good models of 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). These features may represent many types of data: analog sensory signals, numbers, linguistic variables. We can imagine [2] a coordinate system based on these features defining a multidimensional space, called here "the mind space". In this space a "mind function" is defined, describing the "mind objects" as a fuzzy areas where the mind function has nonzero values. Real mind objects are primarily composed of preprocessed sensory data, iconic representations, perception-action multidimensional objects. They correspond to stable attractors of brain's dynamics realized by the transcortical neural cell assemblies (TNCAs).

Features of internal representation of data may change slowly with time but active features change rapidly. Their values at a given moment represent "the mind state" corresponding to a point in the mind space. If there is a mind object in this region the object is "activated" or "recognized". Evolution of the mind state is equivalent to a series of activations of objects in the mind space. These objects are created and positioned using unsupervised as well as supervised methods of learning, similar to the learning vector quantization [5] or other local learning techniques [6-9]. The idea of a "mind space" or "conceptual space" is not more metaphorical than the concept of space-time or other concepts in physics. A proper mathematical description of the mind space is very difficult because of high dimensionality of this space and complicated metric that has a non-Euclidean character. Simple approximations may work quite well in many situations.

Associations among mind objects are based on the distance between them and 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. 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, memories reflecting states of the total system (i.e. 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. 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. 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.

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. We are confident that 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? 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 the phase space. There are some reasons to be optimistic even in this case. 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). 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 [10]. Symbolic approach to dynamics, a drastic simplification, gives very interesting results even for chaotical systems [11].

Since dynamic functions are more difficult to model we will restrict our attention to the static functions now.

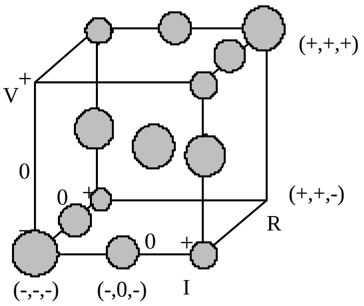

Fig. 1 Representation of the Ohm's law $V = I * R$ in the mind space model. The axes illustrate only one feature of variables, their change: – for decreasing, 0 for constant and + for increasing. This representation of a small subspace of the mind space is created in an unsupervised way from examples or directly from the corresponding rules. Such knowledge representation is very effective in the reasoning process, for example in qualitative analysis of electrical circuits [3].

FEATURE SPACE MAPPING SYSTEM

FSM network [3] has some unique properties, rather different from those of most artificial neural network models. It uses separable processing functions for localized description of fuzzy data in the mind space. In the special case when gaussian processing functions are used by the network nodes (gaussians are the only radial basis functions that are separable [6]) this model belongs to the family of the growing Hyper Basis Function (gHBF) networks. Localized processing functions representing the mind objects are initially centered on the data vectors $\mathbf{D}=(D_1, D_2, \ldots D_N)$ with dispersions of each gaussian 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:

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

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

$$\begin{aligned} M(\mathbf{X},\mathbf{D},\mathbf{D}) &= \Sigma_p\, W_p s(\mathbf{X};\mathbf{D}^p,\mathbf{D}^p) \\ &= \Sigma_p\, W_p\, \Pi_i\, g\left(X_i; D_i^p, \Delta_i^p\right) \end{aligned}$$

and does not vanish only around the data vectors $\mathbf{D}$ stored in the M function. The weights $\mathbf{W}$ and the dispersions $\mathbf{D}$ are the adaptive parameters defining the mind function for a given set of $\mathbf{D}$ input values. If the input data values are noisy the centers $\mathbf{D}$ are treated as adaptive parameters as it is done in clustering algorithms, such as the LVQ model [5]. 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 [6], RAN networks [7] or other vector quantization methods [8]. Initial value of the adaptive parameters is obtained from the k-nearest neighbor heuristics or from 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 $\mathbf{Y}$ should form one mind object, therefore in FSM inputs and outputs are treated on an equal footing. HBF approximation of one dimensional function is given by Y=HBF($\mathbf{X}$); in FSM this relation is always fuzzy and the most probable function is obtained from maximization:

$$\max_Y M(\mathbf{X}, Y) \Leftrightarrow Y = F(\mathbf{X})$$

Thus for a given X value a whole range of probable Y values is obtained. FSM network (Fig. 2) 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 gradient. The network reaches a stable state when local maximum is found, therefore FSM is 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.

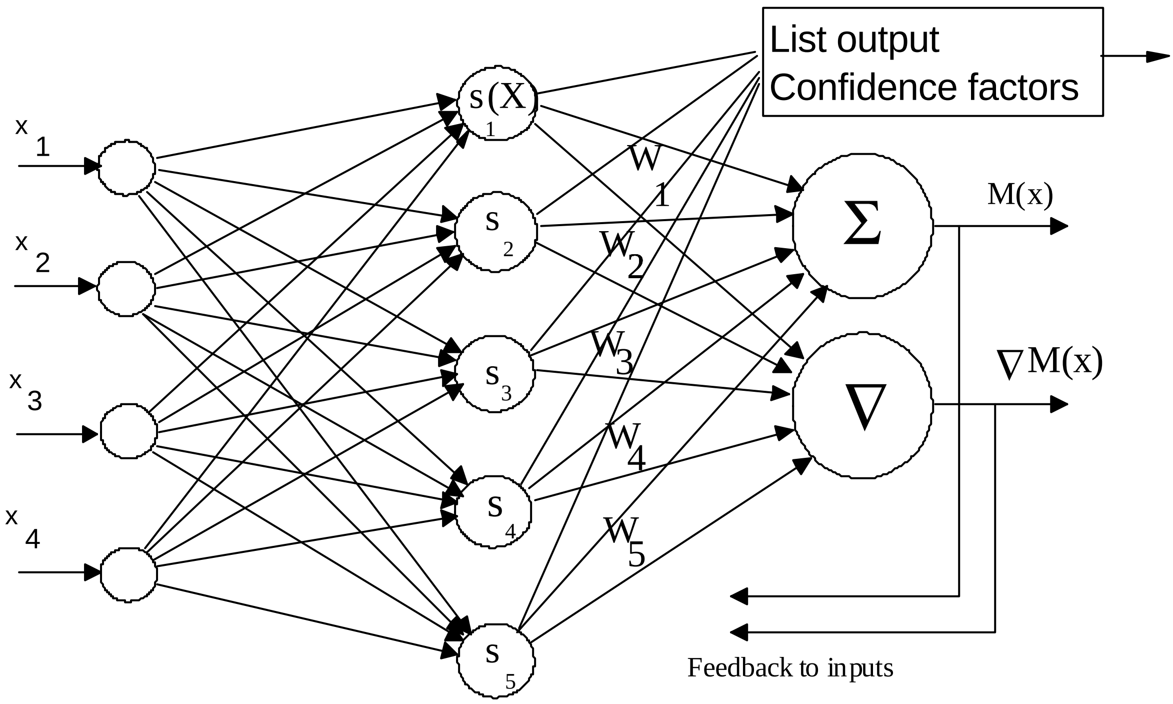

Fig. 2 Example of a network realizing the Feature Space Mapping (FSM) model.

After the initial nodes of the network are established 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 $(\mathbf{X}_n, Y_n)$, 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:

$$\min_k \|\mathbf{X}_n - \mathbf{D}_k\| > d_{\min};\; Y_n - F_W^{(n-1)}(\mathbf{X}_n; \mathbf{D}, \mathbf{s}) > \varepsilon$$

Here d_{min} is the resolution of the data in the input space. The value for the dispersion σ_k is frequently based on the nearest neighbor heuristic. When the new data does not satisfy both criteria given above, gradient adaptation of the weights, centers and fuzziness of the node functions is performed. Only the local gradient estimation is used here for the $(\mathbf{X}_n, Y_n)$ data (as is also done in RAN and in the function estimation approach [7]). The weights are changed according to:

$$\mathbf{W} \leftarrow \mathbf{W} + \eta\left(Y_n - F_W^{(n-1)}(\mathbf{X}_n; \mathbf{D}, \mathbf{s})\right) \times \nabla_{W,D} F_W^{(n-1)}(\mathbf{X}_n; \mathbf{D}, \mathbf{s})$$

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

$$\mathbf{D} \leftarrow \mathbf{D} + \eta_d\, (\mathbf{X} - \mathbf{D})$$

This solution leads to self-organization of data clusters in the mind 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 [6] using tensor product stabilizers. The FSM adaptive system tries to minimize a local error function

$$E[M_W] = \sum_{i=1}^{N} \sum_{j \in O(Ci)} K_i(X_j - C_i)(Y_j - M_W(X_j))^2$$

where the kernel functions K_i and the neighborhood definitions $O(C_i)$ depend on the problem while W symbolize all adaptive parameters. This error function may also include a proper stabilizer although in practice we add noise to the input data to get smooth 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:

$$\mathrm{IF}(x_1 \in X_1 \wedge x_2 \in X_2 \wedge \ldots x_N \in X_N)$$
$$\mathrm{THEN}(y_1 \in Y_1 \wedge y_2 \in Y_2 \wedge \ldots y_M \in Y_N)$$

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 a given node are allowed. More general rules of the type

$$\mathrm{IF}\left(x_1 \in X_1^{(1)} \wedge \ldots x_N \in X_N^{(1)}\right)$$
$$\vee\left(x_1 \in X_1^{(2)} \wedge \ldots x_N \in X_N^{(2)}\right) \vee (\ldots)$$
$$\mathrm{THEN}\,(y_1 \in Y_1 \ldots \wedge y_M \in Y_M)$$

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.

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 the basic features of mind objects. This corresponds to a general orientation step in human information processing. After the local maximum 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 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 [12].

APPLICATIONS

FSM system, described above as an example of application of the general 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 pursued by our group [13] we should mention:

Classification of stellar spectra: modern telescopes, including Hubble Space Telescope, produce large amounts of stellar spectra. Classification of these spectra is still done manually or by correlating the position of the star with the entry in the catalog of known stars. 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. They are 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.

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).

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;

symbolic interpretation, neural realization;

full control over associations and generalizations by adjusting overlaps and fuzziness of mind objects;

supervised and unsupervised learning methods for self-organization of mind space objects;

learning from examples, as in neural networks, and learning from general laws, as in expert systems;

straightforward implementation of a typical expert system production rules in the form:

IF (FACT 1.and.FACT2.or.FACT3...)

than (FACT_N)

reasoning may take form of one-dimensional searches (if separable functions are used), focusing on single variable, with the depth of search equal to the number of unknown features;

fast retrieval gradient techniques for finding associations with the multi-scale approach (focusing and defocusing) to concentration on relevant parts of the mind space;

adding and removing mind objects (network nodes) to reduce complexity of the model;

fine tuning of object representations for pattern recognition and adaptive control;

spontaneous formation of hierarchies of objects leading to categories and metaconcepts;

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

REFERENCES

[1] W. Duch (1994), Proc. of I National Conference on neural networks and applications, Kule, April 1994, pp. 17-28; W. Duch, Floating Gaussian Mapping: a new model of adaptive systems, Neural Network World **4** (1994) 645-654

[2] W. Duch, Transparent theory of consciousness - is there a problem? Behavioral and Brain Sciences, 1995 (in print); A solution to fundamental problems of cognitive science, PSYCOLOQUY, 1994 (submitted)

[3] W. Duch and G.H.F Diercksen, Feature Space Mapping as a Universal Adaptive System, Comp.Phys.Comm. 1995 (in print).

[4] D. Amit, Modeling Brain Function, Cambridge University Press 1989

[5] T. Kohonen, Self-organization and Associative Memory. (Springer-Verlag, New York, 1984, 3rd edition: 1989).

[6] 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

[7] 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

[8] B. Fritzke, Vector quantization with growing and splitting elastic net, in: ICANN '93: Proc. of the International Conference on artificial neural networks, Amsterdam, 1993

[9] 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

[10] A. Garnham and J. Oakhill, Thinking and reasoning, Oxford, Blackwell 1994

[11] T. Bedford, M. Keane and C. Series, Ergodic theory, symbolic dynamics and hyperbolic spaces, Oxford, Oxford University Press 1991

[12] W. Duch, Quantitative measures for the Self-Organizing Topographic Maps. Open Systems and Information Dynamics **3** (1994) 295-302

[13] 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 (in print).