

## A CURSORY INTRODUCTION TO THE PHYSICISTS' NEURAL NETWORKS

L. PERSONNAZ, E. BIENENSTOCK and G. DREYFUS

*Ecole Supérieure de Physique et de Chimie Industrielles de la ville de Paris, Laboratoire d'Electronique, 10, rue Vauquelin, F-75005 Paris, France*

Résumé - Nous présentons une très courte introduction aux réseaux de neurones formels, mettant en relief les principaux problèmes et les voies de recherche actuelles, et indiquant quelques références bibliographiques.

Abstract - The paper is a very short introduction to networks of formal neurons, intended to outline the main problems and trends and provide some basic references.

In the past half-century, many attempts have been made to model various parts of the brain, particularly those responsible for such functions as perception, movement, memory and learning. One of the first major contributions to the mathematical modeling of brain function was a paper published in 1943, where Warren S. McCulloch and Walter H. Pitts introduced their "formal neuron", a highly stylized version of the real thing /1/. They showed that networks of interconnected formal neurons could perform specific logical tasks. Less mathematical but extremely influential was the work of the psychophysicist Donald O. Hebb, who in 1949 introduced the notion of a "cell assembly", a distributed representation of an external event inside a neural network /2/. Hebb proposed that cell assemblies are created, and continually modified during learning, as a consequence of the modification of the synaptic efficacies, i.e., the conductances of the contacts between neurons. According to Hebb's rule, repeated activation of one neuron by another leads to an increase of the strength of the synapses which take part in the activation process. McCulloch and Pitts's formal neuron, as well as Hebb's learning rule, are still widely used in physicists' models.

The first attempts at designing machines inspired from the nervous system were made in the 1960's, but these attempts were only partly successful: limitations became soon apparent, that could not be overcome with the technical and conceptual tools available at that time /3/. The level of activity in the field remained fairly low for some twenty years. In 1982, the physicist John J. Hopfield published his seminal paper in which he recognized the analogy between Hebb's cell assemblies and disordered magnetic systems /4/. The proof that the proposed model was analytically solvable /5/ triggered a large research effort in recent years. The number of papers on neural networks, particularly of the Hopfield type, originating from the physics community, is probably well over one hundred at the present day /6/. An even larger number of papers, many of which are authored by computer scientists or electrical engineers, describe potential applications of various types of neural networks /5/, in such fields as artificial vision, speech understanding and synthesis, combinatorial optimization, interpolation and prediction...

The elementary component in most of these models is the McCulloch-Pitts formal neuron, or a somewhat elaborate version thereof. Such processing units are infinitely less complex than real neurons. It is, however, the very simplicity of these models which makes them attractive, for the following reasons:

- as mentioned above, the tools of statistical physics can be used to analyze and predict the behaviour of large assemblies of formal neurons; one may in particular investigate the information storage properties and the dynamics of learning and retrieval of such networks acting as associative memories;
- despite their crudeness, these models are able to account for some properties of biological neural assemblies; such approaches are gradually gaining acceptance among neurobiologists, although debates are far from being settled /7/;
- electronic or opto-electronic networks of formal neurons can be built, thus opening the

way to potential industrial applications.

The neural networks studied today belong essentially to two families:

- Feedforward neural networks, in which information flows from an input layer to an output layer through a small number (one to three) of "hidden layers". These networks use various rules for modifying their synapses, including the popular "error back-propagation algorithm" /6/. They exhibit interesting properties for classification and related problems. Although their behaviour is far from being completely understood, they are good candidates for applications; they are probably less relevant as models of biological nervous systems.

- Feedback neural networks, whose archetype is the Hopfield network. These are dynamical systems, which evolve in time until they reach a stationary state. They have been the favorite subject of investigation for physicists. The initial Hopfield model has been greatly refined in order to gain biological plausibility, one of the major steps being the derivation of analytic results for systems with asymmetric couplings /8/. Moreover, results of great importance concerning the limitations of the storage capacity of such networks were derived /9/.

Although the level of activity in the field has undergone a tremendous increase in recent years, many issues are still to be addressed. In particular, much needs to be done to bring these models closer to recent experimental data from neurobiology and cognitive sciences: as was mentioned above, some of the key ingredients in recent works date back to the 1940's. More elaborate models are still considered untractable, but this will presumably change in the future. Another type of work that is called for concerns the limitations in the capabilities of formal neural networks. As was mentioned above in connection with ref. 8, results concerning the limitations of associative memory networks are now available. Similar results on the learning and generalization capabilities of feedforward networks are still needed; this lack of solid conceptual bounds might lead to unreasonable hopes, especially in the field of short-term applications.

## REFERENCES

- /1/ McCulloch, W.S., Pitts, W.H., Bull. Math. Biophys. **5**, 115 (1943).
- /2/ Hebb, D.O., "The Organization of Behavior" (J. Wiley, New York, 1949).
- /3/ Minsky, M. and Papert, S., "Perceptrons" (MIT Press, 1969).
- /4/ Hopfield J.J., Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).
- /5/ Amit, D., Gutfreund, H., Sompolinsky, H., Phys. Rev. A **32**, 1007 (1985).
- /6/ It is impossible to give a comprehensive bibliographical review of this blossoming field. The following book, which ignores all the developments arising from the physics community, provides basic information on feedforward networks:  
Rumelhart, D.E., McClelland J.L., "Parallel Distributed Processing" (MIT Press, 1986).  
A number of recent conference proceedings contain a wealth of scientific and bibliographical data :  
"Parallel models of associative memory", Hinton, G.E., Anderson, J.A., eds. (Lawrence Erlbaum Associates, 1981).  
"Disordered Systems and Biological Organization", Bienenstock, E., Fogelman Soulié, F., Weisbuch, G., eds. (Springer Verlag, 1986).  
"Neural Networks for Computing", Denker, J., ed. (American Institute of Physics, New York, 1986).  
"Neural Computers", Eckmiller, R., von der Malsburg, C., eds. (Springer Verlag, 1987).  
"Neural Information Processing Systems", Anderson, D.Z., ed. (American Institute of Physics, New York, 1988).  
Proceedings of the IEEE International Conference on Neural Networks (San Diego, 1987).  
Proceedings of the IEEE International Conference on Neural Networks (San Diego, 1988).  
"Neural Networks : from Models to Applications", Personnaz, L., Dreyfus, G., eds. (IDSET, Paris, 1989).
- /7/ Toulouse G., Proceedings of the Bat-Sheva Seminar (Jerusalem, 1988).
- /8/ Derrida, B., Gardner, E., Zippelius, A., Europhys. Lett. **4**, 167 (1987).
- /9/ Gardner, E., Europhys. Lett. **4**, 481 (1987).