WHAT'S WRONG WITH NEURAL NETWORKS ?

G. DREYFUS, L. PERSONNAZ, E. BIENENSTOCK

Ecole Supérieure de Physique et de Chimie Industrielles
de la Ville de Paris
Laboratoire d'Electronique
10, rue Vauquelin
75005 PARIS
FRANCE

Let us face it: the neural network craze has crossed the Atlantic. It is right here now, and it might be a good idea to capitalize on past examples and try not to make the same mistakes as before. This position paper is a brief critical assessment of the present situation in neural network research, based on an extensive experience in the field and on many discussions with foreign and French researchers alike, informal or within organized reflexion groups.

NEURAL NETWORKS: AN OLD STORY

The concept of a formal neuron originated in the 1940's; the popular McCulloch-Pitts neuron was designed in order to try to model the operation of the nervous system, but the idea of attempting to mimick the brain in order to build powerful machines spread very quickly. This generated a lot of excitement in the 1960's (not unlike what we observe right now, incidentally) but the famous - some would say infamous - book by Minsky and Papert entitled "Perceptrons" put an end to this period. Some research went on during twenty years, with a low level of activity. The whole field was renewed by (i) the efficient use of concepts originating from statistical mechanics in order to predict and analyze the behaviour of complex, nonlinear systems, and (ii) the introduction of a method for training multilayer feedforward networks, known as the backpropagation algorithm. These two contributions turned out to be very useful, although they are widely different in nature: the first one gives sound theoretical bases for the analysis of learning and retrieval, both in feedback and in feedforward networks, whereas the backpropagation technique provides an algorithm which is useful when handled properly, with distressingly little theoretical background.
The activity started very slowly in 1982, but increased exponentially during the last

three years. The number of research papers followed the trend, generating new specialized journals, and of course, a host of conferences and workshops, of widely varying scopes and scientific level. This excitement raises three basic problems:
(i) the relevance of formal neurons to real biology,
(ii) the relevance of formal neurons to real applications,
(iii) the relevance of formal neurons to real scientific research.
We shall try to tackle these questions by first stating the hopes that formal neural networks raise in the field of neurobiology and the criticisms that they face; in a second part, we shall comment on the links between concepts in neurobiology and potential applications. Finally, we shall issue a few words of caution about future research, especially applied research, in the field.

WHAT NEUROBIOLOGISTS SAY TO FORMAL NEURAL NETWORK RESEARCHERS

After an initial phase of outright skepticism, an increasing number of neurobiologists display an earnest interest in the potential benefits of modelling the operation of the nervous system - or small subsystems thereof - with very simple units in the spirit of the McCulloch-Pitts formal neuron. However, the basic problem is the same in this field as in any other scientific field: the value of a theory lies in its ability, not only to **analyze** scientific data, but mainly to **predict** phenomena and inspire new experiments whose results should be in agreement with theory. Suggesting a model which accounts for previous experimental results is one, usually non-trivial, thing; suggesting original experiments which will unravel new behaviours predicted by the model is another, often more difficult, task. It must be stated clearly that the number of examples of such a complete, bidirectional interplay is very small. However, there are indeed examples of such successful interactions, and there is no theoretical reason why such things should not happen more frequently in the future, this evolution being spurred both by the emergence of new experimental techniques for observing the nervous system, and by the recent theoretical advances in the analysis of complex systems of formal neurons. Clearly, the task ahead is very difficult, and no one seriously thinks that the mechanisms underlying the workings of the brain will be fully understood within a predictable time span. Nevertheless, the right choice of the subsystems to be studied, of the experimental and theoretical tools, and of the partner research groups, may lead to substantial advances.
The main deficiency of the present models of neural networks is their crudeness. Taking into account all the intricacies of real neurons is a hopeless task, and would be

at odds with the very idea of modelling; however, it is clear that the McCulloch-Pitts model leaves aside too many important features of real neurons, and, in addition, fails to account for the variety of neurons which are present in living systems. The fact that these crude models can mimick, sometimes in an impressive way, the behaviour of some biological systems, is very exciting and definitely encouraging. But there is a long way to a complete theory of the brain.

To summarize, the main impact of formal neural networks on real neurobiology has been to develop contacts between neurobiologists and researchers from other fields, and to show that some concepts, arising mainly from statistical mechanics, can be helpful in neurobiological modelling.

WHAT FORMAL NEURAL NETWORK RESEARCHERS SAY TO NEUROBIOLOGISTS

In the previous section, we mentioned formal neural network researchers who aim at producing models of biological systems. There is, of course, another brand of neural network researchers, who aim at building powerful artificial machines, and want to draw inspiration from biological systems in order to solve hard problems in vision, speech recognition, and the like, for which conventional computers seem to be inadequate. The basic characteristics which look appealing in this context are the adaptation properties, the massive parallelism, and the simplicity of the elementary unit. The real biological relevance of the models are not a main concern here, and might even hamper, in a sense, the development of this applied research: the end product will definitely be silicon-based, yet it is not at all clear that the hardware solutions found by nature during evolution are appropriate for silicon implementations. Thus, the interest is really in the principles of operation of, say, the visual, the auditory, or the olfactory system, not in the details of the implementation in "wetware". Once the basic operation mechanisms have been clarified, engineers feel free to get away from biology since the technological constraints that they must face have nothing to do with the constraints that living systems have. Indeed, if the biological inspiration becomes very remote, one might even question the use of the term "neural network", since there are many adaptive and/or parallel systems which are not "neural" at all. This ambiguity has been the origin of various misunderstandings between biologists and researchers in the field of formal neurons, but these misunderstandings are now clearing off.

WHAT SOME FORMAL NEURAL NETWORK RESEARCHERS SAY TO THE REST OF THE WORLD

The excitement which was generated by the first achievements in the field of neural network research prompted the creation of a host of small companies in the United States, mainly vendors of software simulators and of "neurocomputers", which are basically personal computers with add-on cards having some number-crunching abilities (featuring a digital signal processing chip or even simply a microprocessor). Venture capital was injected into the game, attracted by market studies which are sometimes akin to pure fantasy, predicting an exponential increase of the "neural network market", whatever that might be! Related commercial activities such as extravagantly priced "newsletters" and organized trips to some US neural network laboratories, which were willing to encourage such activities, appeared very soon. The latest advance in the field is the emergence of so-called "neural network analysts".

Is all this a real problem for the future of the domain, or is it just an epiphenomenon which will quickly fade away? We think that there is a real danger in presenting as a "proven technology" a set of concepts which are neither proven nor technological. Let us explain this.

The concepts are not proven in that the theoretical background of formal neural networks, although infinitely more secure and richer than in the 1960's, is still not sufficient. The methods of statistical mechanics have been very fruitful because they showed a number of exciting generic properties of dynamic neural networks; yet most of this conceptually important work is ignored by that segment of the neural network community which is devoted to applications, partly because statistical mechanics is difficult, partly because of the different publication habits of the two communities. Also, these methods have not yet been applied to static networks trained by backpropagation, so that the latter still lack a safe conceptual basis; things are improving, and some very nice pieces of work have been published, dealing with what can or cannot be learnt by neural networks, and with their generalization properties. Obviously, these topics are still in their infancy. Thus, it must be stated clearly that there is no safe theoretical basis, as yet, for the operation and training of the most popular neural networks, but that much difficult work, which generally does not make the headlines, is being performed in various places.

It might be argued that one does not need a full-fledged theory in order to develop applications, and that neural networks are proven through the success of their applications. This is definitely wrong. There are many **potential** applications of neural networks, and many prototype applications working satisfactorily in various research

centers, in universities and in industry; there are a number of reports on successful attempts - failures will never be reported, of course. However, commercial, civil applications are nowhere to be found as yet. This definitely does not mean that applications are doomed to fail; it means that neural networks are still in the realm of research and development, and that it is too early to assess, in any reasonable way, the commercial future of neural networks.
Let us turn now to the second term in "proven technology". Why are neural networks not a technology? Because, at the present time, the vast majority of investigations have been performed with simulations on standard computers, possibly with an add-on card. This can hardly be termed a new technology. Again, there are prototypes of really innovative products, such as simulators using parallel architectures (transputer-based, for instance), or such as silicon chips with original architectures, and a lot more is to come in the near future; but these are definitely research prototypes, not commercial products. If it can be shown in the next years that neural networks can indeed be valuable for large-scale applications, then a neural network technology may evolve, with, for instance, new advances in analog techniques, or with new materials for manufacturing synapses; right now, there is nothing like a neural network "technology".

To summarize: what is the message that neural network researchers who are not concerned mainly with short-term profit should try to communicate to outsiders? Basically, that the field of neural networks is not mature enough for straightforward applications without a prior extensive reflexion phase on the idiosyncrasies of the problems that one wants to solve, on the way the problems are solved classically, or on the real reasons why they are not solved classically. Much can be learnt from standard non-parametric statistical methods and their relationships to neural networks. It should be stated clearly that there is nothing like a single magic network that will solve just any problem provided the computer is left to run for days. Indeed, there are many different types of networks: a particular application may require the design of a specific neural network, and also, sometimes more importantly, the choice of an appropriate data representation. In other words, success in any application will stem from collaborations between (i) real neural network professionals and (ii) engineers or researchers who are real professionals in their own fields.

CONCLUSION

There is nothing wrong with neural networks. What is definitely wrong is hype, careless claims, outright lies, amateurish work, lax publication policies. What this field

needs is a lot of high-quality scientific and technical work, a lot of interactions between really professional people in neurobiology, psychology, mathematics, computer science, electronics, both in Industry and in Universities. **But there is no magic in weighted sums**. No one should expect that any really difficult problem will be solved by blindly using weighted sums and non-linearities, even if they are provided by the fanciest and most picturesque simulator; first and foremost, solving a problem with networks of formal neurons requires a lot of thinking and exercising our own little brain cells, in the framework of the difficult, old-fashioned, perhaps not-too-glamorous scientific method.