Comment

Elie Bienenstock and Stuart Geman

According to the authors, this paper has three principal goals: "informs a statistical readership about Artificial Neural Networks (ANNs), points out some of the links with statistical methodology and encourages cross-disciplinary research...." It seems to us that the authors have been spectacularly successful with regards to the first two of these goals, and it is likely that this paper will do much to further stimulate the already active scientific exchange between the statistics and neural modeling communities.

As Cheng and Titterington made clear, neural networks, at least the very popular examples reviewed in their paper, are not really new inasmuch as they represent variations on common statistical themes, especially nonparametric and semiparametric estimation and classification. Furthermore, Cheng and Titterington suggest that the tie to real neurons may be somewhat tenuous (we will amplify on this shortly). Nevertheless, despite this dubious biological connection and strong ties to already well-studied statistical methods, this field has attracted wide attention from within the government (principally the Department of Defense but also other branches including the Department of Commerce) as well as many sectors of industry. It has drawn many top science students at our top schools. In the meantime, many statistics departments complain that it is hard to find first-rate graduate students.

We would like to use this discussion to speculate about the reasons behind the fantastic growth of the neural modeling field, especially in light of the close ties to well-studied areas of statistics which have themselves been received with substantially less enthusiasm. There are many reasons for the remarkable popularity and visibility of neural networks. We will propose a few and suggest that some of them may be based partly on misconceptions.

THE APPEAL OF BRAIN MODELING

The endeavor is nearly irresistible: building models and machines possessing a measure of human intelligence, working through the puzzles of perception and cognition and "explaining" the brain. Indeed, many researchers in the neural modeling community believe that the kinds of networks discussed by Cheng and Titterington are meaningfully connected with biology, providing a starting point from which we can begin to organize and understand the overwhelmingly complex anatomical and physiological data, and from which new kinds of theoretically-directed biological experiments will emerge. Still, most neural modelers would agree that these attempts are nothing more than the crudest of approximations not to be taken seriously as models of real neurons or real neuronal interactions at the level of any important detail. Cheng and Titterington have already remarked that "it is clear that the brain does not learn by the generalized delta rule." It is also clear that there is very little in the way of feedforward networks in the brain (virtually all substantial pathways are reciprocated making it clear that the dynamics is not that of a feedforward network) and that the real equations of synaptic modification are a good deal more complicated than a Hebbian or gradient-descent rule. In short, ANNs are hardly neural.

THE APPEAL OF "GENERALIZATION"

Model-free generalization has served as a kind of Holy Grail in neural modeling: begin with a more-or-less *tabula rasa* (blank slate, or, in statistical parlance, "nonparametric") architecture and a realistically-sized training set for some challenging classification or estimation task and devise a learning rule powerful enough to discover the regularities and invariants that would extrapolate good performance beyond the training data. Such a device might be used to "beat the stock market" or solve the automatic target recognition (ATR) problem which has resisted many years of expensive R&D effort. But statisticians know that generalization (good performance on samples not in the training set) depends almost entirely on the extent to which the training set is representative, and/or the structure of the problem happens to accommodate the models used. It is too much to expect statistical methods to "discover," by themselves, complex and nontrivial structure such as the structure

Elie Bienenstock is Visiting Associate Professor (on leave from CNRS, Paris, France) and Stuart Geman is Professor, Division of Applied Mathematics, Brown University, Box F, Providence, Rhode Island 02912.