

A Reply to Honavar's Book Review of *Neural Network Design and the Complexity of Learning*

J. STEPHEN JUDD

Siemens Corporate Research, 755 College Road East, Princeton, NJ 08540

I want to take this opportunity to reinforce a point that is perhaps not stated clearly enough in my book. It is a point that Honavar (and another reviewer) have rightfully pursued.

The NP-completeness is with respect to the size of the network and the amount of data in the training set to be loaded. Honavar points out that this measure of complexity of the input is somewhat suspect because "it does not reflect the . . . inherent regularities of the data items being loaded; nor can it capture possible effects of choosing a particular network topology for loading a given set of data items." Very true. If we let the devil specify the network architecture (and the data to be loaded) then he can make our lives exceedingly difficult, but we should not take this result to indicate that our lives have to be difficult.

In the domain of connectionist research, it is the researcher who gets to play the role of the devil, and he is certainly not trying to be malicious. If he knows what he is doing, he can presumably make the problem easy; but, no one yet has a theory of how best to accomplish this. The themes of my results are that it is easy to get into trouble, and that we need to elaborate the theory further. (Maybe the common all-to-all connectivity scheme is a solution for small nets, but that claim has to be proved and it may introduce concomitant generalization problems.)

In the neurological domain, the network design is not under our control; it is nature that has played the role of devil and we must see what hand she has dealt us. For my explorations I picked a particular design family (shallow nets, grid-like structures) to model what seems to be the first-order aspect of her handiwork. Am I off base here? No one has yet suggested so. Even though the architectures are extremely regular, the intractability still shines through, and we are left to conclude that the only types of data that can be loaded in there have to have some kind of fortuitous peculiarity. Honavar's concern about "the interactions between the network architecture and intrinsic structure of the data set to be loaded" are natural and important. My theorems on simple regular architectures are an attempt to address this concern, although such is not well elaborated in the book.

Blum and Rivest responded to my early theorems (which had large and complicated architectures as well as specially constructed data sets) by finding an apparently smaller, very simple, and very regular architecture that nevertheless also has intractable data sets. They too were worried that the intractability was somehow due to a devious mismatch between the data and the architecture, but their result also seems to mute this hypothesis. However, neither they nor myself have any crisp statements about this.

The reviewer makes one slight overstatement about generalization being difficult whenever memorization is, but has otherwise understood the book very well. I thank Honavar for reading the book so closely and for coming to many of the same questions as I think are important; his attention is flattering.