Summary of NIPS-90 Workshop:
Constructive and Destructive Learning Algorithms

Workshop Leader: Scott E. Fahlman
School of Computer Science
Carnegie-Mellon University
Pittsburgh, PA 15213
Phone: (412) 268-2575
Internet: scott.fahlman@cs.cmu.edu

This was a two-day workshop on Nov. 30 and Dec. 1, 1990, devoted to presentation and discussion of "dynamic" learning algorithms: algorithms which change the topology of a network as learning progresses. Some of these algorithms, called "destructive" or "subtractive" algorithms, start with excess connections and remove any that are not needed; others, the "constructive" or "additive" algorithms, start with a sparse network and add new hidden units and connections as needed, sometimes in multiple layers. In either case, the user is relieved of the burden of guessing in advance precisely what network topology will best fit a given problem. In addition, many of these algorithms claim improved learning speed and generalization. Growing interest in this topic was reflected in the number of people attending: as many as 60 at times, and a hard core of 25-30 people who made it to all four sessions, despite scheduling conflicts with other workshops.

We began with an introduction of the workshop participants and their particular interests. This was followed by a brief discussion of network topologies and how they relate to the network's expressive power and ability to generalize. The question of generalization kept recurring in our discussions. As a rule, a network with too many free, independent parameters and too little training data will eventually begin to overfit small random details in the training data, leading to poor performance on new cases that were not part of the training set. In a dynamic architecture, it is possible to counteract this tendency by keeping only those free parameters that play a large role in reducing the training set error. If the network is constructed gradually, it is possible to employ a separate test-set and to stop the training as soon as performance on the test set begins to deteriorate.

The relation of network topology to generalization can be complex, however: under some conditions, excess capacity in the network will give rise to redundancy (and hence to some degree of fault-tolerance) rather than to over-fitting. A truly minimal network, by definition, cannot be fault-tolerant. A few ideas about how to promote fault-tolerance in dynamic networks were discussed, but this is an area that requires more research.

Another important theme was the possibility of employing "greedy" learning in some of the constructive algorithms. In a greedy scheme (Cascade-Correlation is one example) each new hidden unit tries to deal with as much as possible of the existing problem. It is then frozen, and subsequent units only deal with the unsolved residue. This is in contrast to schemes such as standard backprop, in which all of the units remain fluid and must interact continuously in order to work out which parts of the overall task will be handled by which units or coalitions. Such negotiations can lead to somewhat better solutions than the

greedy scheme, but they are generally quite slow and expensive. Hybrid schemes are possible in which units are plastic when first added, but gradually harden; these allow some negotiation among multiple units, but without the kind of total free-for-all that we see in backprop.

An interesting dichotomy was seen between networks that used units with "global" receptive fields (e.g. sigmoid) and those whose units respond only to some local patch of the existing input space. The latter category includes radial basis and gaussian ellipsoid units and also the whole family of nearest-neighbor schemes. The local schemes tend to have more free parameters, so more care is required to avoid over-fitting. On the other hand, the local schemes often can learn much faster because a new unit does not have to fit *all* of the training cases.

The following people then gave presentations, either on dynamic learning schemes that they have developed or on their experience in using schemes developed by others:

- Steve Hanson, Siemens
- Fernando Tenorio, Purdue Univ.
- Vasant Honavar, Iowa State Univ.
- Guo-Zheng Sun, Univ. of Maryland
- John Platt, Synaptics Inc.
- Scott Fahlman, Carnegie Mellon Univ.
- Chris Lebiere, Carnegie-Mellon Univ.
- Ira Smotroff, MITRE
- Steve Gustafson, Univ. of Dayton
- Masumi Ishikawa, Kyushu Inst. of Technology
- Hans Thodberg, Danish Meat Research Inst.
- Steve Gallant, Northeastern Univ.
- Peter Dayan, Edinburgh (presenting work of M. Frean)
- Steve Lane, Princeton

This workshop was largely tutorial and exploratory in nature: most of us wanted to learn more about the algorithms and experiences of the others. Consequently, no dramatic conclusions were reached. However, the following concluding statements seemed to appeal to most of the group:

- The constructive algorithms seem to have some advantages over the purely destructive ones in terms of the time and effort required to train them. Most promising of all are the hybrid schemes that add units but only after dissolving away less important connections to those units.
- The choice between local and global unit-types depends on the nature of the feature space, and on the desired tradeoff between learning speed and smoothness of the decision surface. It might be worth exploring schemes that combine local and global unit-types in one network.
- Dynamic architectures are exciting because they choose their own topology and because they offer some significant advantages in learning speed and generalization. There is a big space of possible designs, only partially explored.