

Introduction

Structured Connectionist Systems

ALEX WAIBEL

waibel@cs.cmu.edu

Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213

In recent years, connectionism has reemerged as an active and lively area of research. Originally motivated in part by a desire to understand and mimic the information processing of the nervous system, it has by now grown and spread into subdisciplines for which cognitive or biological plausibility no longer represent the *only* concern. Research on application driven connectionist learning systems are amongst those. They have led to marriages between classical non-connectionist and connectionist approaches and have generated novel modeling strategies by which successful operation in practical domains can be attained. Previous knowledge of the task or the domain is then used to impose structure on the learning algorithm or the best combination of connectionist and classical processing strategies.

Early work has already branched out in this direction. Amongst these are mechanisms to deal with time (Elman, 1990), sequential information (Bourlard & Wellekens, 1988; Jordan, 1986), modeling invariances (Waibel, et al., 1989; LeCun, et al., 1989; Simard, et al., 1992) and the learning of architectural parameters of a connectionist network in constructive or destructive learning algorithms (Hanson, 1990; Fahlman & Lebiere, 1990; Bodenhausen, 1991; LeCun, et al., 1990; Waibel, 1989). Combinations of symbolic and connectionist learning methods have also been undertaken (Shavlik & Towell, 1989; Masuoka, et al., 1990). Many more interesting avenues have been tried and reported. Sadly, they cannot all be cited here, and the small set listed here should by no means be viewed as representative or complete.

In this special issue, we present four papers that are examples of structured approaches to connectionist learning. The common thread among them is that they are connectionist learning systems that impose, assume, or derive structure in connectionist learning. They limit and constrain learning, rather than hoping that the learner will come up with all relevant representational, structural and algorithmic constraints by itself or by accident.

Two of the given papers apply connectionist learning in practical systems that have classically not always been viewed as pattern recognition problems: robot arm control and software engineering. Here suitable network architectures, training databases, optimization criteria, and objective functions are chosen as well as an appropriate embedding of the connectionist learner in the overall system. The other two papers differ in that they impose constraints on the learning algorithm itself. These constraints help the algorithm identify relevant knowledge important to the task and ignore irrelevant variability in the data.

The first paper by Pomerleau, applies backpropagation to learn control functions from visual input in robot control problems where precise analytical derivation of control signals is impossible due to torque and inertia of the manipulator. The choice of data

representation and presentation and the choices of objective functions are examples of design and modeling choices. By analyzing visual feedback, a neural net learns to provide fine control adjustments to complement classical control in a more reliable space station manipulator.

The paper by Schwanke and Hanson applies neural network components to emulate subjective judgments of similarity to produce effective criteria for automatic modularization of software. The system learns to assign useful categorization of software into modules and could act as a software engineer's advisor or "intelligent assistant." Here too, successful integration of connectionist pattern processing leads to benefits in a classically symbolic activity.

The paper by Spirkovska and Reid explores an approach based on higher order neural networks that specifically eliminates irrelevant variability in the input data. In particular, the networks presented impose invariance toward rotation, translation, and scaling for use in 2D and 3D visual object recognition. This eliminates the need to discover such invariances by extra learning and extra data.

The paper by Cohn, Atlas, and Ladner presents a self-directed learning algorithm, by which a network takes the initiative and queries (or explores) a feature space to seek out useful properties and boundaries of a task. Rather than stumbling on the distinctive aspects of its world, the learning algorithm queries or searches for the "interesting" regions of the space. In doing so, the learner can assemble knowledge of the task more rapidly and generalize more effectively. Query learning may not only deliver competitive performance results on practical tasks, but it leads us closer to an understanding of more plausible models of learning; learning algorithms that explore the world actively rather than watching it go by.

Building connectionist learning systems for competitive real world applications is a challenge that combines what we learn about connectionist learning with design and careful system integration. Keeping an eye on both is a challenging perspective that puts new ideas to the test but rewards us with performance and insight into large scale learning processes.

References

- Bodenhausen, U. (1991). Learning the architecture of neural networks for speech recognition. *Proceedings ICASSP* (pp. 117-120) IEEE.
- Bourlard, H. & Wellekens, C.J. (1989). Links between markov models and multilayer perceptrons. D. S. Touretzky (Ed.) *Advances in Neural Network Information Processing Systems, 1*. San Francisco, CA: Morgan Kaufmann.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.
- Fahlman, S. & Lebiere, C. (1990). The cascade-correlation learning architecture. In D. S. Touretzky (Ed.) *Advances in Neural Information Processing Systems, 2*. San Francisco, CA: Morgan Kaufmann.
- Hanson, S. (1990). Meiosis networks. In D. S. Touretzky (Ed.) *Advances in Neural Information Processing Systems, 2*. San Francisco, CA: Morgan Kaufmann.
- Jordan, M. L. (1986). *Serial order: A parallel distributed processing approach* (Technical Report.) San Diego, CA: University of California at San Diego, Institute for Cognitive Science.
- LeCun, Y., Boder, B., Denker, J., Henderson, D., Howard, R.E., Hubbard, W., and Jackel, L.D. (1989). Handwritten digit recognition with a back-propagation network. In D. S. Touretzky (Ed.) *Advances in Neural Information Processing Systems, 1*. San Francisco, CA: Morgan Kaufmann.

- LeCun, Y., Denker, J., Solla, S., Howard, R.E., and Jackel, L.D. (1990). Optimal brain damage. In D. S. Touretzky (Ed.) *Advances in Neural Information Processing Systems*, 2. San Francisco, CA: Morgan Kaufmann.
- Masuoka, R., Watanabe, N., Kawamura, A., Owada, Y., & Asakawa, K. (1990). Neurofuzzy system—Fuzzy inference using structured neural network. *Proceedings of the international conference on fuzzy logic & neural networks* (pp. 173–177).
- Shavlik, J. W. & Towell, G. G. (1989). An approach to combining explanation-based and neural learning algorithms. *Connection Science*, 1. Reprinted in J. Shavlik & T. G. Dietterich (Eds.), *Readings in Machine Learning*, San Francisco, CA: Morgan Kaufmann, 828–839, 1990.
- Simard, P., Le Cun, Y., Victorri, B. & Denker, J. (1992). Tangent Prop—A formalism for specifying selected invariances in an adaptive network. In J. E. Moody, S. J. Hanson, & R. P. Lippmann (Eds.) *Advances in neural information processing systems*, 4. San Francisco, CA: Morgan Kaufmann.
- Waibel, A. (1989). Modular construction of time-delay neural networks for speech recognition. *Neural Computation*, 1 39–46.
- Waibel, A., Hanazawa, T., Hinton, G., Shikano, K. & Lang K. (1989). Phoneme recognition using time-delay neural networks, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37, 328–339.