

Explicit Learning Relationships within Neural Systems

Talib S. Hussain
Intelligent Distributed Computing Dept.
BBN Technologies
Cambridge, MA, 02474

Abstract – The explicit representation and use of diverse learning relationships among diverse neural components is essential for the development of neural systems that can learn to solve complex real-world tasks.

I. PROBLEM STATEMENT

The limited learning capabilities of existing neural networks are a key barrier to the wider use of neural systems for complex, real-world domains. Most neural networks incorporate a simple learning mechanism that is used to adapt neural parameters or neural structures in simple ways over the course of large numbers of training trials. For instance, the most common approach is the use of a single “learning rule” that is applied homogeneously throughout the network and thus imposes a single fixed learning relationship among the neural elements. While such an approach is useful for some classes of problems, it is intrinsically limited. A complex real-world task may, for example, decompose into multiple sub-tasks that have qualitatively different properties and/or dependencies that vary with time. Such a problem may require a system to learn what the subtasks are, learn how to solve each subtask, and learn to adapt to changes in the task decomposition. No single “learning rule” will suffice.

An important step in advancing the state of the art is to develop techniques for engineering complex neural systems with functionally and/or structurally diverse neural components (e.g., modules) and with diverse functional and learning relationships among those components. At this juncture, we need not limit ourselves only to known artificial or biological neural mechanisms for adaptation. While we may have a good understanding of many aspects of the biological brain, we are far from a full understanding of how different neural components interact to produce the rich learning capabilities that natural systems, and in particular humans, demonstrate. Thus, a variety of social and cognitive learning relationships may prove valuable in the development of more robust neural systems.

Explicit representation and use of diverse learning relationships among neural components is a key need. Certain modular network models [1-2] and the

Attribute Grammar Encoding approach [3-5] have shown promise in addressing this issue. Further work along these lines may enable us to produce neural models capable of applying the right types of learning processes to the right types of learning objectives at the right time.

II. CURRENT RESEARCH

A variety of neural learning mechanisms have been developed over the years, and may be classed broadly into unsupervised weight modifying learning rules [6-7], supervised weight modifying learning rules [8-9], structural learning rules such as growing [10] and pruning [11] algorithms, modular learning mechanisms [1-2] and evolutionary neural network systems (ENNSs). Within ENNSs, evolutionary algorithms may be used to adapt a variety of neural properties, including weights [12], structure [13], neural learning rules [14], and neural network development rules [15-16].

The MINOS network [1] and the Hierarchical Mixtures of Experts (HME) [2] are modular models in which certain neural components explicitly modulate the learning activity of others. Within MINOS, multiple MINOS modules interact with an authority module. The MINOS modules report a result as well as a self-confidence measure, and the authority module allocates specific learning events to specific MINOS modules based on their error and confidence. Similarly, within HME, multiple competing modules interact with a gating network that modulates the amount of learning that occurs within the modules based on their error. Both approaches use an explicit learning relationship to enable task specialization.

Attribute Grammar Encoding (AGE) [3-5] provides a mechanism for the specification of a wide variety of neural network structure and behaviours. Attribute grammar [17] productions define rules that generate neurons, specify the internal functionality of those neurons, create connections among the neurons, and organize the network into modular structures with explicit functional relationships. To achieve this flexibility, a generalized neuron model is adopted that allows a neuron to receive, transmit and internally manipulate an arbitrary number of signal types, and to have arbitrary number of memory variables. Thus, in

addition to traditional activation and feedback signals, the grammar may define an arbitrary number of control signals that allow one neuron to explicitly modulate the behavior of another. Further, a Generic Neural Markup Language (GNML) was created, based on this neuron model and using XML [18], to allow the specification of arbitrary structural and functional components through nesting of different levels of detail. A neural interpreter was developed that accepts GNML specifications and produces functional neural networks capable of learning.

The grammar productions within AGE may be crafted to define neural properties to an arbitrary level of detail. At the highest level of detail, broad structural properties (e.g., modular organization) or functional relationships (e.g., patterns of feedback or control signal connectivity) may be identified. At the lowest level of detail, the functionality of a specific neuron may be composed by identifying mathematical operations, operation sequencing or nesting, and the input/output signal types and memory stores operated upon. The grammar produces a GNML specification of a complete neural network. Generally, terminals in the grammar encode GNML strings as attributes and productions manipulate and merge those strings via attribute evaluation rules to create larger and larger specifications. To simplify the grammar, terminals may encode arbitrarily complex GNML strings.

AGE has been applied to generate grammars for a variety of common neural networks models, as well as novel models that combine a variety of structural and functional properties. Modular networks are easily specified, and the modules may be arbitrarily different in their internal properties.

III. KEY AVENUES

Consider the wide variety of learning relationships that exist among intelligent creatures in nature, such as: learning from observation, learning from mistakes, learning by doing, parenting, mimicry, mentoring, indirect feedback, specific feedback, approval, penalty, collaboration, teaming, brainstorming, active learning via exploration or questions, one-to-many lectures, one-on-one tutoring, and so on. Such types of learning relationships may form the basis for neural learning mechanisms that direct the interactions between different neural components.

Efforts to develop more robust neural learning systems should not only explore neural mechanisms based on these and other learning relationships, but more importantly should explore the use of multiple learning approaches within a single system. Human abilities are the sum of all their learning capabilities, not just one or two. While it is good to model

simplifying underlying mechanisms, the extreme simplification, such as down to one or two learning rules within a network, limits our understanding of the rich interplay between the different types of learning approaches. Research on the synergies between neural modules with diverse processing capabilities has led to an improvement in the capability, ease of development, and applicability of neural solutions. It has also furthered a deeper understanding of functional properties of neural systems. Similarly, studies on neural systems with diverse learning properties should lead to significant advances in the state of the art.

These studies will be facilitated by the creation of representations that encompass multiple learning mechanisms. AGE has demonstrated one approach with a high degree of generality. However, the generality was achieved in part by relaxing the traditional definition of a neuron and of the signals that are transmitted within a network. AGE also requires a complex neural interpreter to execute the networks. Alternative approaches to the representation and incorporation of diverse learning mechanisms should be explored.

REFERENCES

- [1] Smieja, F.J. and Mühlenbein, H. Reflective Modular Neural Network Systems. Technical Report: GMD #633, German National Research Centre for Computer Science, 1992.
- [2] Jordan, M.I. and Jacobs, R.A. Hierarchical Mixtures of Experts and the EM Algorithm. Technical Report: A.I. Memo #1440, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1993.
- [3] Hussain, T.S. and Browse, R.A. "Evolving neural networks using attribute grammars," 2000 IEEE Symposium on Combinations of Evolutionary Computation and Neural Networks, p. 37-42, 2000.
- [4] Hussain, T.S. Attribute Grammar Encoding of the Structure and Behaviour of Artificial Neural Networks. Ph.D. Thesis, Queen's University at Kingston, 2003.
- [5] Hussain, T.S. "Generic Neural Markup Language: Facilitating the design of theoretical neural network models," 2004 IEEE International Joint Conference on Neural Networks (IJCNN) Piscataway, NJ: IEEE, p. 235-242, 2004.
- [6] Kohonen, T. Self-Organization and Associative Memory. Berlin: Springer-Verlag, 1984.
- [7] Hebb, D.O. The Organization of Behavior: A Neuropsychological Theory. New York: Wiley, 1949.
- [8] Rumelhart, D.E., Hinton, G.E., and Williams, R.J. "Learning internal representations by error propagation," Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations. D.E. Rumelhart and J.L. McClelland (Eds.), Cambridge, MA: MIT Press, p. 318-362, 1986.
- [9] Grossberg, S. Studies of Mind and Brain. Boston, MA: D. Reidel Publishing, 1982.

- [10] Fahlman, S.E. and Lebiere, C. "The Cascade-Correlation learning architecture," *Advances in Neural Information Processing Systems 2*. D.S. Touretsky (Ed.), San Mateo, CA: Morgan Kaufmann, p. 524-532, 1990.
- [11] LeCun, Y., Denker, J.S. and Solla, S.A. "Optimal brain damage," *Advances in Neural Information Processing Systems 2*. D.S. Touretsky (Ed.), San Mateo, CA: Morgan Kaufmann, p. 598-605, 1990.
- [12] Miller, G.F., Todd, P.M. and Hegde, S.U. "Designing neural networks using genetic algorithms," *Proceedings of the Third International Conference on Genetic Algorithms and Their Applications*. J.D. Schaffer (Ed.), San Mateo, CA: Morgan Kaufmann, p. 379-384, 1989.
- [13] Kitano, H. "Designing neural networks using genetic algorithms with graph generation system," *Complex Systems*, 4, p. 461-476, 1990.
- [14] Bengio, S., Bengio, Y., and Cloutier, J. "Use of genetic programming for the search of a new learning rule for neural networks," *First Conference on Evolutionary Computation*, p. 324-327. 1994.
- [15] Jacob, C. "Genetic L-system programming," *Parallel Problem Solving from Nature III, Lecture Notes in Comp. Sci.* 866, Berlin: Springer, p. 334-343, 1994.
- [16] Gruau, F. "Automatic definition of modular neural networks," *Adaptive Behavior*, 3(2), p. 151-183, 1995.
- [17] Knuth, D. E. "The semantics of context-free languages," *Mathematical Systems Theory*, 2(2), p.127-145, 1968.
- [18] Bray, T., Paoli, J., Sperberg-McQueen, C.M. and Maler, E. *Extensible Markup Language (XML) 1.0 (Second Edition)*, World Wide Web Consortium (W3C) Technical Report: Recommendation Oct. 6, 2000.