Distributed Prediction and Hierarchical Knowledge Discovery by ARTMAP Neural Networks

Gail A. Carpenter

Department of Cognitive and Neural Systems, Boston University
677 Beacon Street, Boston, MA 02215 USA
gail@cns.bu.edu

Abstract

Adaptive Resonance Theory (ART) neural networks model real-time prediction, search, learning, and recognition. ART networks function both as models of human cognitive information processing [1,2,3] and as neural systems for technology transfer [4]. A neural computation central to both the scientific and the technological analyses is the *ART matching rule* [5], which models the interaction between top-down expectation and bottom-up input, thereby creating a focus of attention which, in turn, determines the nature of coded memories.

Sites of early and ongoing transfer of ART-based technologies include industrial venues such as the Boeing Corporation [6] and government venues such as MIT Lincoln Laboratory [7]. A recent report on industrial uses of neural networks [8] states: "[The] Boeing … Neural Information Retrieval System is probably still the largest-scale manufacturing application of neural networks. It uses [ART] to cluster binary templates of aeroplane parts in a complex hierarchical network that covers over 100,000 items, grouped into thousands of self-organised clusters. Claimed savings in manufacturing costs are in millions of dollars per annum." At Lincoln Lab, a team led by Waxman developed an image mining system which incorporates several models of vision and recognition developed in the Boston University Department of Cognitive and Neural Systems (BU/CNS). Over the years a dozen CNS graduates (Aguilar, Baloch, Baxter, Bomberger, Cunningham, Fay, Gove, Ivey, Mehanian, Ross, Rubin, Streilein) have contributed to this effort, which is now located at Alphatech, Inc.

Customers for BU/CNS neural network technologies have attributed their selection of ART over alternative systems to the model's defining design principles. In listing the advantages of its THOT® technology, for example, American Heuristics Corporation (AHC) cites several characteristic computational capabilities of this family of neural models, including fast on-line (one-pass) learning, "vigilant" detection of novel patterns, retention of rare patterns, improvement with experience, "weights [which] are understandable in real world terms," and scalability (www.heuristics.com).

Design principles derived from scientific analyses and design constraints imposed by targeted applications have jointly guided the development of many variants of the basic networks, including fuzzy ARTMAP [9], ART-EMAP [10], ARTMAP-IC [11],

Gaussian ARTMAP [12], and distributed ARTMAP [3,13]. Comparative analysis of these systems has led to the identification of a *default ARTMAP* network, which features simplicity of design and robust performance in many application domains [4,14]. Selection of one particular ARTMAP algorithm is intended to facilitate ongoing technology transfer.

The default ARTMAP algorithm outlines a procedure for labeling an arbitrary number of output classes in a supervised learning problem. A critical aspect of this algorithm is the distributed nature of its internal code representation, which produces continuous-valued test set predictions distributed across output classes. The character of their code representations, distributed vs. winner-take-all, is, in fact, a primary factor differentiating various ARTMAP networks. The original models [9,15] employ winner-take-all coding during training and testing, as do many subsequent variations and the majority of ART systems that have been transferred to technology. ARTMAP variants with winner-take-all (WTA) coding and discrete target class predictions have, however, shown consistent deficits in labeling accuracy and post-processing adjustment capabilities. The talk will describe a recent application that relies on distributed code representations to exploit the ARTMAP capacity for one-to-many learning, which has enabled the development of self-organizing expert systems for multi-level object grouping, information fusion, and discovery of hierarchical knowledge structures. A pilot study has demonstrated the network's ability to infer multi-level fused relationships among groups of objects in an image, without any supervised labeling of these relationships, thereby pointing to new methodologies for self-organizing knowledge discovery.

References

[1]   S. Grossberg, "The link between brain, learning, attention, and consciousness," *Consciousness and Cognition,* vol. 8**,** pp. 1-44, 1999. ftp://cns-ftp.bu.edu/pub/diana/Gro.concog98.ps.gz

[2]   S. Grossberg, "How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex,"*Behavioral and Cognitive Neuroscience Reviews*, in press, 2003, http://www.cns.bu.edu/Profiles/Grossberg/Gro2003BCNR.pdf

[3]   G.A. Carpenter, "Distributed learning, recognition, and prediction by ART and ARTMAP neural networks," *Neural Networks*, vol. 10, pp. 1473-1494, 1997, http://cns.bu.edu/~gail/115_dART_NN_1997_.pdf

[4]   O. Parsons and G.A. Carpenter, "ARTMAP neural networks for information fusion and data mining: map production and target recognition methodologies," *Neural Networks*, vol. 16, 2003, http://cns.bu.edu/~gail/ARTMAP_map_2003_.pdf

[5]   G.A. Carpenter and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Computer Vision, Graphics, and Image Processing*, vol. 37, pp. 54-115, 1987.

[6] T.P. Caudell, S.D.G. Smith, R. Escobedo, and M. Anderson, "NIRS: Large scale ART 1 neural architectures for engineering design retrieval," *Neural Networks*, vol. 7, pp. 1339-1350, 1994, `http://cns.bu.edu/~gail/ NIRS_Caudell_1994_.pdf`

[7] W. Streilein, A. Waxman, W. Ross, F. Liu, F., M. Braun, D. Fay, P. Harmon, and C.H. Read, "Fused multi-sensor image mining for feature foundation data," In *Proceedings of 3rd International Conference on Information Fusion,* Paris, vol. I, 2000.

[8] P. Lisboa, "Industrial use of saftey-related artificial neural netoworks," *Contract Research Report 327/2001,* Liverpool John Moores University, 2001. `http://www.hse.gov.uk/research/crr_pdf/2001/crr01327. pdf`

[9] G.A. Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds, and D.B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks*, vol. 3, pp. 698-713, 1992, `http://cns.bu.edu/~gail/ 070_Fuzzy_ARTMAP_1992_.pdf`

[10] G.A. Carpenter and W.D. Ross, "ART-EMAP: A neural network architecture for object recognition by evidence accumulation," *IEEE Transactions on Neural Networks*, vol. 6, pp. 805-818, 1995, `http://cns.bu. edu/~gail/097_ART-EMAP_1995_.pdf`

[11] G.A. Carpenter and N. Markuzon, "ARTMAP-IC and medical diagnosis: Instance counting and inconsistent cases," *Neural Networks*, vol. 11, pp. 323-336, 1998. `http://cns.bu.edu/~gail/ 117_ARTMAP-IC_1998_.pdf`

[12] J.R. Williamson, "Gaussian ARTMAP: A neural network for fast incremental learning of noisy multidimensional maps," *Neural Networks*, vol. 9, pp. 881-897, 1998, `http://cns.bu.edu/~gail/ G-ART_Williamson_1998_.pdf`

[13] G.A. Carpenter, B.L. Milenova, and B.W. Noeske, "Distributed ARTMAP: a neural network for fast distributed supervised learning," *Neural Networks,* vol. 11, pp. 793-813, 1998, `http://cns.bu.edu/~gail/ 120_dARTMAP_1998_.pdf`

[14] G.A. Carpenter "Default ARTMAP," *Proceedings of the International Joint Conference on Neural Networks (IJCNN'03),* 2003. `http://cns.bu.edu/~gail/Default_ARTMAP_ 2003_.pdf`

[15] G.A. Carpenter, S. Grossberg, and J.H. Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *Neural Networks*, vol. 4, pp. 565-588, 1991, `http://cns.bu.edu/~gail/054_ ARTMAP_1991_.pdf`.

Biography

Gail Carpenter (`http://cns.bu.edu/~gail/`) obtained her graduate training in mathematics at the University of Wisconsin (PhD, 1974) and taught at MIT and Northeastern University before moving to Boston University, where she is a professor

in the departments of cognitive and neural systems (CNS) and mathematics. She is director of the CNS Technology Lab and CNS director of graduate studies; serves on the editorial boards of *Brain Research*, *IEEE Transactions on Neural Networks*, *Neural Computation*, *Neural Networks*, and *Neural Processing Letters*; has been elected to successive three-year terms on the Board of Governors of the International Neural Network Society (INNS) since its founding, in 1987; and was elected member-at-large of the Council of the American Mathematical Society (1996-1999). She has received the INNS Gabor Award and the Slovak Artificial Intelligence Society Award. She regularly serves as an organizer and program committee member for international conferences and workshops, and has delivered many plenary and invited addresses. Together with Stephen Grossberg and their students and colleagues, Professor Carpenter has developed the Adaptive Resonance Theory (ART) family of neural networks for fast learning, pattern recognition, and prediction, including both unsupervised (ART 1, ART 2, ART 2-A, ART 3, fuzzy ART, distributed ART) and supervised (ARTMAP, fuzzy ARTMAP, ART-EMAP, ARTMAP-IC, ARTMAP-FTR, distributed ARTMAP, default ARTMAP) systems. These ART models have been used for a wide range of applications, such as remote sensing, medical diagnosis, automatic target recognition, mobile robots, and database management. Other research topics include the development, computational analysis, and application of neural models of vision, nerve impulse generation, synaptic transmission, and circadian rhythms.