# Modular Artificial Neural Networks Bibliography

Compilation Date: May, 1995

Compiled by: Bernardo Morcego
Automatic Control and Computer Engineering Dept.
Universitat Politècnica de Catalunya
c/ Pau Gargallo, 5
08021 Barcelona, Spain

email: bernardo@esaii.upc.es

[1] Albus, J.S., A new approach to manipulator control: the cerebellar model articulation controller (CMAC), ASME J. Dynamic Systems Measurement Control, 97(3):220-227, 1975.

[2] Albus, J.S., A robot conditioned reflex system modeled after the cerebellum, Proceedings fall joint computer conference, XLI : 1095-1104, 1972.

[3] Albus, J.S., A theory of cerebellar functions, Maths. Biosci, 1(1):25-61, 1971.

[4] Albus, J.S., Brains, Behaviour, Robotics, Peterborough, N.H, Byte Books, 1981.

[5] Albus, J.S., Data storage in the cerebellar model articulation controller (CMAC), ASME J. Dynamic Systems Measurement Control, 97(3):228-223, 1975.

[6] Albus, J.S., Mechanisms of Planning and Problem solvingx in the Brain, Mathematical Biosciences 45 pp247-293, 1979.

[7] Albus, J.S., Theoretical and experimental aspects of a cerebellar model, Ph.D Thesis, University of Maryland, USA, 97(3):220-227, 1972

[8] Bengio, Y. & P. Frasconi, Input/Output Hidden Markov Models for Sequence Processing, Advances in Neural Information Processing Systems 4, (J. E. Moody and S. J. Hanson and R. P. Lippman eds.), 1992.

`ftp ftp.iro.umontreal.ca/pub/neuro/nips94-iohmms.ps.Z`

`URL http://www.iro.umontreal.ca/labs/neuro/`

[9] Bennani, Y. and P. Gallinari, Task Decomposition Through a Modular Connectionist Architecture: a Talker Identification System, Artificial Neural Networks 2, I Aleksander and J. Taylor (Eds), Elsevier, 1992.

[10] Boers, E.J:W. and H. Kuiper, Biological Metaphors and the Design of Modular Artificial Neural Networks, Master's Thesis, Leiden University, 1992.

[11] Eldracher, M. and B. Baginski, Neural Subgoal Generation Using Backpropagation, World Congress on Neural Networks, Lawrence Erlbaum, 1993.

[12] Fogelman, F., B. Lamy and E. Viennet., Multi-Modular Neural Network Architectures for Pattern Recognition: Applications in Optical Character Recognition and Human Face Recognition, Advances in Pattern Recognition Systems using Neural Networks, Eds. I. Guyon and P.S.P. Wang, IJPRAI, vol. 7, number 4, 1993.

[13] Fogelman, F., E. Viennet, B. Lamy, Multi-Modular Neural Network Architectures: Applications in Optical Character and Human Face Recognition, Int. Journal of Pattern Recognition and Artificial Intelligence vol 7, No 4: 721-755, 1993.

[14] Fox, D., V. Heinze, K. Möller, S. Thrun, G. Veenker, Learning by error-driven decomposition, Proceedings of International Conference on Artificial Neural Networks, 1991

[15] Fuller, T.H., Jr. and Takayuki D. Kimura, Supervised Competitive Learning, Journal of Intelligent Material Systems and Structures, 1994.

Abstract:

Supervised Competitive Learning (SCL) assembles a set of learning modules into a supervised learning system to address the stability-plasticity dilemma. Each learning module

acts as a similarity detector for a prototype, and includes prototype resetting (akin to that of ART) to respond to new prototypes. SCL has usually employed backpropagation networks as the learning modules. It has been tested with two feature abstractors: about 30 energy-based features, and a combination of energy-based and graphical features (about 60). About 75 subjects have been involved. In recent testing (15 college students), SCL recognized 99% (energy features only) of test digits, 91% (energy) and 96.6% (energy/graphical) of test letters, and 85% of test gestures (energy/graphical). SCL has also been tested with fuzzy sets as learning modules for recognizing handwritten digits and handwritten gestures, recognizing 97% of test digits, and 91% of test gestures.

`ftp archive.cis.ohio-state.edu/pub/neuroprose/fuller.scl.ps.Z`

[16] Gasser, M., "Modularity in a Connectionist Model of Morphology Acquisition" in COLING 94 (Proceedings of the 15th annual conference on Computational Linguistics) Vol.1 pp 214-220 (1994)

[17] Goudreau, M.W. and C.L. Giles, "Using Recurrent Neural Networks to Learn the Structure of Interconnection Networks," Neural Networks, 1996.

`URL http://www.neci.nj.nec.com/homepages/giles.html`

[18] Gruau, F. and D. Whitley, The Cellular Developmental of Neural Networks: the Interaction of Learning and Evolution, Ecole Normale Superieure de Lyon, Research Report 93-04, 1993.

[19] Hampshire, J.B. and A. Waibel, The Meta-Pi Network: Building Distributed Knowledge Representations for Robust Multi-Source Pattern Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, no. 7, pp. 751-769, July, 1992.

[20] Happel, B.L.M, & J.M.J. Murre (submitted). Evolving complex dynamics in modular interactive neural networks. Submitted to Neural Networks.

`ftp ftp.mrc-apu.cam.ac.uk/pub/nn/murre/chaos1.ps`

[21] Happel, B.L.M. and J.M.J Murre, Designing Modular Network Architectures Using a Genetic Algorithm, Artificial Neural Networks 2, I Aleksander and J. Taylor (Eds), Elsevier, 1992.

[22] Happel, B.L.M., & J.M.J. Murre (1994). The design and evolution of modular neural network architectures. Neural Networks, 7, 985-1004.

`ftp ftp.mrc-apu.cam.ac.uk/pub/nn/murre/nnga1.ps`

[23] Heemskerk, J.N.H., J. Hoekstra, J.M.J. Murre, L.H.J.G. Kemna, & P.T.W. Hudson (1994). The BSP400: a modular neurocomputer. Microprocessors and Microsystems, 18, 67-78.

[24] Jacobs, R.A. and Jordan, M.I. (1992) Computational Consequences of Bias toward Short Connections, Journal of Cognitive Neuroscience, 4-4, 323-336.

[25] Jacobs, R.A. and M.I. Jordan, Learning Piecewise Control Strategies in a Modular Neural Network Architecture, IEEE Trans. on Systems, Man, and Cybernetics, Vol 23, No 2, 1993.

[26] Jacobs, R.A., Jordan, M.I. and Barto, A.G. (1991) Task Decomposition Through Competition in a Modular Connectionist Architecture: The What and Where Vision Tasks, Cognitive Science, 15, 219-250.

[27] Jacobs, R.A., Task Decomposition Through Competition in a Modular Connectionist Architecture, University of Massachusetts, COINS TR 90-44, 1990.

[28] Jordan, Michael I. and Jacobs, Robert A., Hierarchies of adaptive experts, Advances in Neural Information Processing Systems 4, (J. E. Moody and S. J. Hanson and R. P. Lippman eds.), 1992.

[29] Miikkulainen, R. and M.G. Dyer, Natural Language Processing with Modular PDP Networks and Distributed Lexicon, Cognitive Science 15(3): 343 - 399, 1991.

[30] Mumford, D. (1992) On the computational architecture of the neocortex: The role of cortico-cortical loops, Biological Cybernetics, 66, 241-251.

[31] Murray-Smith, R., A Local Model Network Approach to Nonlinear Modelling, Department of Computer Science, University of Strathclyde, Ph.D. Thesis, 1994.

ftp psyche.mit.edu/pub/rod/phd/*.ps.Z

[32] Murre, J.M.J. (1992). Categorization and learning in modular neural networks. Hemel Hempstead: Harvester Wheatsheaf. Co-published by Lawrence Erlbaum in the USA and Canada (Hillsdale, NJ).

[33] Murre, J.M.J. (1993). Transputers and neural networks: an analysis of implementation constraints and performance. IEEE Transactions on Neural Networks, 4, 284-292.

[34] Murre, J.M.J., & S.E. Kleynenberg (1990). The MetaNet Network Environment for the development of modular neural networks. In: B. Widrow & B. Angeniol (Eds.) Proceedings of the International Neural Network Conference, INNC-90-Paris, 717-720. Dordrecht: Kluwer.

[35] Murre, J.M.J., R.H. Phaf, & G. Wolters (1989). CALM networks: a modular approach to supervised and unsupervised learning. Proceedings of the International Joint Conference on Neural Networks, Washington DC, 1, 649-656. (New York: IEEE Press, IEEE Catalog Number 89CH2765-6.)

[36] Murre, J.M.J., R.H. Phaf, & G. Wolters (1992). CALM: Categorizing And Learning Module. Neural Networks, 5, 55-82.

[37] Perrone, M.P. and L. N Cooper, Learning from what's been learned: Supervised learning in multi-neural network systems, Proceedings of the World Congress on Neural Networks, 1993.

[38] Perrone, M.P. and N. Intrator, Unsupervised Splitting Rules for Neural Tree Classifiers, Proceedings of the International Joint Conference on Neural Networks, 1992.

[39] Perrone, M.P., A Novel Recursive Partitioning Criterion, Proceedings of the International Joint Conference on Neural Networks, 1991.

[40] Perrone, M.P., A Soft-Competitive Splitting Rule for Adaptive Tree-Structured Neural Networks, Proceedings of the International Joint Conference on Neural Networks, 1992.

[41] Perrone, M.P., and L. N Cooper, When Networks Disagree: Ensemble Method for Neural Networks, Artificial Neural Networks for Speech and Vision, ed. R. J. Mammone, 1993.

[42] Perrone, M.P., Averaging Techniques for Neural Networks, The Handbook of Brain Theory and Neural Networks, MIT Press, 1993.

[43] Perrone, M.P., General Averaging Results for Convex Optimization, Proceedings of the 1993 Connectionist Models Summer School, Eds. M. C. Mozer and P. Smolensky and D. S. Touretzky and J. L. Elman and A. S. Weigend, 1994.

[44] Perrone, M.P., Pulling It All Together: Methods for Combining Neural Networks, nips6:1188-1189, 1994.

[45] Phaf, R.H. (1994). Learning in natural and connectionist systems: experiments and a model. Dordrecht: Kluwer Academic Publishers.

[46] Plaut, D.C, Double Dissociation without Modularity: Evidence from Connectionist Neuropsychology, to appear in Journal of Clinical and Experimental Neuropsycology, 1994.

[47] Pozarlik, R. (1994) Integrating Symbolic and Parallel Distributed Processes in a Jigsaw Neural Network, Proc. of the WCNN'94 (San Diego), vol 4, 550-555.

[48] Schmidhuber, J.H. and D. Prelinger, Discovering Predictable Classifications, Neural Computation vol. 5:625-635, 1993.

[49] Schmidhuber, J.H. and R. Huber, Learning to Generate Artificial Fovea Trajectories for Target Detection, International Journal of Neural Systems vol. 2:135-141, 1991.

[50] Schmidhuber, J.H., Curious Model-Building Control Systems, Proc. International Joint Conference on Neural Networks, Singapore, 1458-1463, 1991.

[51] Schmidhuber, J.H., Learning Complex, Extended Sequences Using the Principle of History Compression, Neural Computation vol. 4:234-242, 1992.

[52] Schmidhuber, J.H., Learning Factorial Codes By Predictability Minimization, Neural Computation vol. 4(6) :863-879, 1992.

[53] Schmidhuber, J.H., Learning to Generate Sub-Goals for Action Sequences, Artificial Neural Networks, (eds.) T. Kohonen, K. Mäkisara, O. Simula and J. Kangas, Elsevier, 1991.

[54] Schmidhuber, J.H., Learning Unambiguous Reduced Sequence Descriptions, Advances in Neural Information Processing Systems 4, (J. E. Moody and S. J. Hanson and R. P. Lippman eds.), 291-298, 1992.

[55] Schmidhuber, J.H., M. C. Mozer and D. Prelinger, Continuous History Compression, Proc. of Intl. Workshop on Neural Networks, RWTH Aachen, Augustinus Pub., 1993.

[56] Schmidhuber, J.H., Planning Simple Trajectories Using Neural Subgoal Generators, Proc. of the 2nd International Conference on Simulation of Adaptive Behavior, MIT Press, 1992.

[57] Schmidhuber, J.H., Reinforcement Learning in Markovian and Non-Markovian Environments, Advances in Neural Information Processing Systems 3, (D. S. Lippman and J. E. Moody and D. S. Touretzky eds.), 500-506, 1991.

[58] Schmidhuber, J.H.,Learning to Control Fast-Weight Memories: An Alternative to Recurrent Nets, Neural Computation vol. 4:131-139, 1992.

[59] Sharkey, N.E. & Sharkey, A.J.C. (1992) A Modular design for Connectionist Parsing. In Proceedings of Workshop on Language Technology, Twente. (Eds) M.F.J. Drosaers and A. Nijholt.

[60] Smieja, F., Multiple network systems (MINOS) modules: Task division and module discrimination, Proceedings of the 8th AISB conference on Artificial Intelligence, Leeds, 1991.

`ftp borneo.gmd.de/pub/as/janus/pre_4.ps`

`URL http://borneo.gmd.de/AS/janus/publi/pre_4.html`

[61] Smieja, F., The Pandemonium system of reflective agents, IEEE Transactions on Neural Networks, 1995

`URL http://borneo.gmd.de/AS/pages/janus.html`

[62] Sun, R., Schemas, Logics, and Neural Assemblies, To appear in: Applied Intelligence, Special issue on connectionist models Vol.5, No.2. Feb. 1995 (edited by Michael Dyer)

Abstract:

To implement schemas and logics in connectionist models, some form of basic-level organization is needed. This paper proposes such an organization, which is termed a discrete neural assembly. Each discrete neural assembly is in turn made up of discrete neurons (nodes), that is, a node that process inputs based on a discrete mapping instead of a continuous function. A group of discrete neurons (nodes) closely interconnected form an assembly and carry out a basic functionality. Some substructures and superstructures of such assemblies are developed, to enable complex symbolic schemas to be represented and processed in connectionist networks. The paper shows that logical inference can be performed precisely, when necessary, in these networks and with certain generalization, more flexible inference (fuzzy inference) can also be performed. The development of various connectionist constructs demonstrates the possibility of implementing symbolic schemas, in their full complexity, in connectionist networks.

`ftp archive.cis.ohio-state.edu/pub/neuroprose/sun.schema.ps.Z`

[63] Sung-Bae, C. and J.H. Kim, Combining multiple neural networks by fuzzy integral for robust classification, IEEE Trans. Systems, Man and Cybernetics, vol. 25, no. 2, 380-384, 1995.

[64] Sung-Bae, C. and K. Shimohara, Toward evolvable model of modularized neural networks, Proc. the Fifth Annual Conference of Japanese Neural Network Society, 117-118, Tsukuba, 1994.

[65] Sung-Bae, C., A neurocomputing framework: from methodologies to application, Neurocomputing, 1995. (To appear)

[66] Sung-Bae, C., Cooperation of modularized neural networks by fuzzy integral with OWA operators, Proc. 3rd Int. Conf. Fuzzy Logic, Neural Nets, and Soft Computing, 95-96, Iizuka, 1994.

[67] Sung-Bae, C., Exploiting modularity in neural networks, PSYCOLOQUY, psycoloquy. 94.5.61.categorization.12.cho, ISSN 1055-0143, 1994.

[68] Tresp, Volker and Michiaki Taniguchi, Combining Estimators Using Non-Constant Weighting Functions, G. Tesauro, D. S. Touretzky and T. K. Leen, eds., "Advances in Neural Information Processing Systems 7", MIT Press, Cambridge MA, 1995.

Abstract:

This paper discusses the linearly weighted combination of estimators in which the weighting functions are dependent on the input. We show that the weighting functions can be derived either by evaluating the input dependent variance of each estimator or by estimating how likely it is that a given estimator has seen data in the region of the input space close to the input pattern. The latter solution is closely related to the mixture of experts approach and we show how learning rules for the mixture of experts can be derived from the theory about learning with missing features. The presented approaches are modular since the weighting functions can easily be modified (no retraining) if more estimators are added. Furthermore, it is easy to incorporate estimators which were not derived from data such as expert systems or algorithms.

`ftp archive.cis.ohio-state.edu/pub/neuroprose/tresp.combining.ps.Z`

[69] Waibel, A., Modular Construction of Time-Delay Neural Networks for Speech Recognition, Neural Computation 1, 39-46, 1989.

[70] Wildberger, A.M., Induction of knowledge from data by clustering, Proceedings of the 1993 Summer Computer Simulation Conference, 1993.

Abstract:

Knowledge acquisition for the development of expert systems is usually performed by interviewing one or more human experts with the expectation of eliciting from them an intuitive grasp of the domain area that has been distilled from their experience. On the other hand, cybernetic methods of knowledge acquisition attempt to create an expert system by automating the induction of general rules from data derived from the domain area. Cluster analysis is a time-honored method of classifying, categorizing, grouping or dividing raw data into "chunks" that are in some sense "natural" and which serve to produce a generalization for which the data are examples. Clustering competes with other cybernetic methods of knowledge acquisition including those derived from the connectionist school of artificial intelligence, such as artificial neural networks and genetic algorithms, as well as the formal methods of mathematical statistics. Hierarchical clustering methods can automatically construct the "class and object" (or frame and semantic network) representations of knowledge, complete with their conventional top-down inheritance schema. It is also possible to use clustering to induce qualitative rules directly from quantitative data, resulting in the more common rule-based expert system. By introducing probabilistic elements into the measure of similarity, clustering can also be used to generate membership functions for fuzzy set models (36 Refs.)